

Methods in Pharmacology
and Toxicology

Springer Protocols

Kunal Roy *Editor*

Multi-Target Drug Design Using Chem- Bioinformatic Approaches

EXTRAS ONLINE

 Humana Press

METHODS IN PHARMACOLOGY AND TOXICOLOGY

Series Editor

Y. James Kang

**Department of Pharmacology &
Toxicology, University of Louisville
Louisville, Kentucky, USA**

For further volumes:

<http://www.springer.com/series/7653>

Methods of Pharmacology and Toxicology publishes cutting-edge methods and protocols in all areas of pharmacological and toxicological research. Each book in the series offers time tested laboratory protocols and step by step methods for reproducible lab experiments to aid toxicologists and pharmaceutical scientists in laboratory testing. With an emphasis on the molecular biology of toxicological methods, Methods of Pharmacology and Toxicology focuses on topics with wide ranging implications to human health, such as Immunotoxicology, Drug Metabolism, and Metabolomics to provide investigators with highly useful compendiums of key strategies and approaches to successful research in drug development.

More information about this series at <http://www.springer.com/series/7653>

Multi-Target Drug Design Using Chem-Bioinformatic Approaches

Edited by

Kunal Roy

Department of Pharmaceutical Technology, Jadavpur University, Kolkata, West Bengal, India

Dedication

For Aatreyi, Arpit and Chaitali

Preface

Despite significant development of novel rational drug design strategies and high-throughput screening methods, the cost of drug development has sharply increased, and at the same time, the rate of failures in clinical trials has escalated [1]. The “one drug, one target, one disease” approach has failed to appreciate the complexities of disease pathways and the system-wide effects of drugs [2]. Diseases are often multifactorial involving a combination of constitutive and/or environmental factors, and they result from the breakdown of robust physiological systems due to multiple genetic and/or environmental factors, leading to the establishment of robust disease conditions. Thus, complex disorders are more likely to be healed or alleviated through simultaneous modulation of multiple targets. Until now, there are still not fully effective drugs for treating complex, multifactorial diseases, such as cancer, metabolic diseases, and neurological diseases [1]. Polypharmacology that addresses small-molecule interactions with multiple targets has generated a great interest in drug discovery [3]. This approach allows for studies of off-target activities and the facilitation of drug repositioning. Multi-target drugs expand the number of pharmacologically relevant target molecules by introducing a set of indirect, network-dependent effects [4]. Moreover, low-affinity binding of multi-target drugs eases the constraints of druggability and significantly increases the size of the druggable proteome. Multi-target agents are a promising strategy to face complex, multifactorial disorders and drug resistance issues. Compared to combination therapies, they present several advantages, including more predictable pharmacokinetics, lower probabilities of drug interactions, and higher patient compliance [5]. Several already existing efficient drugs, such as nonsteroidal anti-inflammatory drugs, antidepressants, anti-neurodegenerative agents, and multi-target kinase inhibitors, affect many targets simultaneously [4]. Hybridization of drugs is also a powerful tool to develop better treatments for several human diseases, as this can provide combination therapies in a single multifunctional agent in a more specific and powerful way than conventional treatments [6].

In polypharmacology, one of the most important goals is to rationally design compounds that act on multiple key targets driving the pathogenesis of a given disease. Therefore, targeting multiple proteins simultaneously stands a good chance to increase drug efficacy and decrease the possibility of drug resistance. In order to achieve these goals, it would be necessary to develop state-of-the-art computational techniques for data curation, model development, and quantitative predictions [2]. Computational approaches are capable of predicting the activity profile of ligands to a set of targets, anticipating potential selectivity issues, and discovering desired multi-target activities early in the iterative design and optimization steps typical of a preclinical drug discovery project. These approaches are based on 2D or 3D shape and chemical similarity, pharmacophore mapping, target and binding site similarity assessment, docking experiments, bioinformatics, graph theory and modeling, machine learning algorithms, and chemogenomics [3]. These approaches can be classified into statistical data analysis and bioinformatics, ligand-based, and structure-based approaches, all of which are well-documented in the literature. The structure-based methods include inverse docking, binding site similarity analysis, inverse pharmacophore modeling, molecular dynamics simulation, etc., while the ligand-based methods include similarity ensemble approach, extended-connectivity fingerprint, fragment-based shape similarity

search, etc. which can be used in combination with a variety of machine learning methods including deep learning [2]. Systems biology approaches and cellular networks help to understand complex diseases and their mechanisms and offer a lot of possibilities to point out the key elements as potential drug targets and thus suggest new therapeutic treatment strategies. Proteochemometric modeling (PCM) simultaneously considers the bioactivity of multiple ligands against multiple targets and permits exploration of the selectivity and promiscuity of ligands on biomolecular systems of different complexity [7].

Computational modeling including quantitative structure-activity relationship (QSAR), pharmacophore mapping, docking, virtual screening, and other cheminformatics and proteochemometric approaches play a vital role in finding and optimization of leads in any drug discovery program. Computational modeling helps to understand the important molecular features contributing to the binding interactions with the target proteins, thus facilitating design of new potential compounds and prediction of activity of designed compounds which have not yet been tested. These approaches can save time, money, and more importantly animal sacrifice in the complex, long, and costly drug discovery process.

This volume (*Multi-target Drug Design Using Chem-Bioinformatic Approaches*) intends to showcase the recent advances in computational design of multi-target drug candidates involving various ligand- and structure-based strategies. Different chem-bioinformatic modeling strategies that can be applied for design of multi-target drugs have been discussed in this book. Apart from a few literature reviews on the application of chemometric and cheminformatic modeling tools for multi-target drug design, several case studies are also presented. Important databases and web servers in connection with multi-target drug design are also discussed. There are a total of 21 chapters in this book.

The first chapter “Cheminformatics Approaches to Study Drug Polypharmacology” provides a tutorial overview on selected cheminformatics methods useful for assembling, curating/preparing a chemical database, and assessing its diversity and chemical space. This chapter also discusses the methods for evaluating the structure-activity relationships and polypharmacology.

The second chapter “Computational Predictions for Multi-target Drug Design” highlights the current state-of-the-art methodologies used in multi-target identification for therapeutic effects of known drugs or new drug candidates. This chapter emphasizes experimental validation of model-derived predictions.

The third chapter “Computational Multi-target Drug Design” discusses multi-target or polypharmacological drug discovery and several in silico methodologies like quantitative structure-activity relationship (QSAR), pharmacophore modeling, and molecular docking used in the process of discovery of multi-targeted drugs.

The fourth chapter “Multi-target Drug Design for Neurodegenerative Diseases” presents an overview of multi-target computational methods as well as of their successful applications to neurodegenerative diseases. This chapter recommends application of virtual screening encompassing both structure-based and ligand-based techniques for effective multi-target drug design.

The fifth chapter “Molecular Docking Studies in Multi-target Antitubercular Drug Discovery” gives an overview of various targets for antitubercular drug development followed by a literature survey of application of docking studies for the development of multi-target compounds for developing new promising drug candidates against tuberculosis.

The sixth chapter “Advanced Chemometric Modeling Approaches for the Design of Multi-target Drugs Against Neurodegenerative Diseases” discusses the recent advances in chemometric techniques in multi-target anti-neurodegenerative drug design. This chapter

recommends the use of proteochemometric modeling for multi-target-directed ligand design.

The seventh chapter “Computational Studies on Natural Products for the Development of Multi-target Drugs” provides an overview of the currently used computational methods in natural product research, with special reference to multi-target drug design. This chapter discusses that pan-assay interference compounds (PAINS) are for the most part not extraordinarily promiscuous and should not be disregarded prematurely.

The eighth chapter “Computational Design of Multi-target Drugs Against Alzheimer’s Disease” provides the basic background about the molecular targets implicated in the pathogenesis of Alzheimer’s disease. Furthermore, the chapter reviews structure-activity relationships (SAR), 2D and 3D quantitative structure-activity relationships (QSAR), as well as other computational modeling studies performed on multi-target agents for Alzheimer’s disease.

The ninth chapter “Design of Multi-target-Directed Ligands as a Modern Approach for the Development of Innovative Drug Candidates for Alzheimer’s Disease” reviews some examples of the exploitation of the multi-target-directed ligand approach in the rational design of novel drug candidate prototypes for the treatment of Alzheimer’s disease.

The tenth chapter “Virtual Screening for Dual Hsp90/B-Raf Inhibitors” describes a computational strategy leading to the identification of the first dual inhibitors of heat shock protein 90 (Hsp90) and protein kinase B-Raf, both being validated targets for anticancer drug discovery.

The eleventh chapter “Strategies for Multi-target-Directed Ligands: Application in Alzheimer’s Disease (AD) Therapeutics” presents several *in silico* strategies adopted for the development of multi-target anti-Alzheimer compounds followed by a case study leading to their *in vitro* validation.

The twelfth chapter “Computational Design of Multi-target Kinase Inhibitors” summarizes two effective computational strategies to identify multi-target kinase inhibitors. The first approach involved a combination of merged pharmacophore matching, database screening, and molecular docking to reliably identify potential multi-target kinase inhibitors. The second strategy employed ensemble pharmacophore-based screening (EPS) of a compound database, post-EPS filtration (PEPSF) of the ligand hits, and multiple dockings.

The thirteenth chapter “Proteochemometrics for the Prediction of Peptide Binding to Multiple HLA Class II Proteins” discusses “proteochemometrics” (PCM) as a method for deriving QSAR. This chapter presents a protocol applied to a set of peptides binding to seven polymorphic HLA class II proteins from locus DP.

The fourteenth chapter “Linked Open Data: Ligand-Transporter Interaction Profiling and Beyond” presents a workflow for retrieving and curating information for multiple drug targets from the open domain, provides insights into how the retrieved data can be employed in ligand- and structure-based approaches, and discusses the hurdles to consider with respect to data analysis.

The fifteenth chapter “Design of Novel Dual-Target Hits Against Malaria and Tuberculosis Using Computational Docking” reviews different approaches (knowledge-based and screening-based) for designing multi-target inhibitors. Additionally, a step-by-step guide (protocol) and different computational resources are also discussed in detail to design multi-target hits for malaria and tuberculosis.

The sixteenth chapter “Computational Design of Multi-target Drugs Against Breast Cancer” presents protocols and computational practices for screening of multi-target drug molecules for breast cancer receptors. However, the authors emphasize that validation of the screened molecules is essential in the *in vitro* and *in vivo* conditions.

The seventeenth chapter “Computational Methods for Multi-target Drug Designing Against *Mycobacterium tuberculosis*” presents available strategies for computational multi-target drug designing with their advantages and disadvantages. This chapter also discusses an easy, fast, and accurate protocol for multi-target drug designing against the *Mycobacterium tuberculosis*.

The eighteenth chapter “Development of a Web Server for Identification of Common Lead Molecules for Multiple Protein Targets” presents a computational protocol that involves screening, docking, and scaffold-based optimization of hit molecules from a variety of compound libraries against any two specified protein targets. The protocol is made available via a web server named “Multi-target Ligand Design.”

The nineteenth chapter “Computational Method for Prediction of Targets for Breast Cancer Using siRNAs Approach” discusses the development and application of a web-based database, BOSS, for selection of potential RNAi based on the sequences that have been used and validated for RNAi-mediated suppression of breast oncogenes. This database includes the latest information regarding used RNAi molecules that can be cost-effective and less time-consuming.

The twentieth chapter “Historeceptomics: Integrating a Drug’s Multiple Targets (Polypharmacology) with Their Expression Pattern in Human Tissues” presents “historeceptomics” as a new, integrative informatics approach to describing the mechanism of action of drugs in a holistic, in vivo context. The chapter discusses that this approach may give new insights into drug mechanism of action, drug repurposing, and prediction of adverse effects, including the design and development of multi-target drugs or drug combinations.

The twenty-first chapter “Networking of Smart Drugs: A Chem-Bioinformatic Approach to Cancer Treatment” reviews the existing network of “smart drugs” by using a chem-bioinformatic approach toward cancer treatment. According to the authors, an application of computational tools in smart drug designing for cancer treatment will be path-breaking in the future.

I am sure that this collection of 21 chapters will be useful to the researchers working in the field of drug discovery and development.

Kolkata, India

Kunal Roy

References

1. Lu J-J, Pan W, Hu Y-J, Wang Y-T (2012) Multi-target drugs: the trend of drug research and development. PLoS One 7(6):e40262. doi:10.1371/journal.pone.0040262
2. Chaudhari R, Tan Z, Huang B, Zhang S (2017) Computational polypharmacology: a new paradigm for drug discovery. Expert Opin Drug Discov 12(3):279–291, doi: 10.1080/17460441.2017.1280024
3. Rastelli G, Pinzi L (2015) Computational polypharmacology comes of age. Front Pharmacol 6:157. doi: 10.3389/fphar.2015.00157
4. Korcsmáros T, Szalay MS, Böde C, Kovács IA, Csermely P (2007) How to design multi-target drugs. Expert Opin Drug Discov 2:799–808. doi: 10.1517/17460441.2.6.799
5. Talevi A (2015) Multi-target pharmacology: possibilities and limitations of the “skeleton key approach” from a medicinal chemist perspective. Front Pharmacol 6:205. doi: 10.3389/fphar.2015.00205
6. Bérubé G (2016) An overview of molecular hybrids in drug discovery. Expert Opin Drug Discov 11:281–305. doi: 10.1517/17460441.2016.1135125
7. Cortes-Ciriano I, van Westen GJP, Murrell DS, Lenselink EB, Bender A, Malliavin TE (2015) Applications of proteochemometrics—from species extrapolation to cell line sensitivity modeling. BMC Bioinform 16(Suppl 3):A4. doi: 10.1186/1471-2105-16-S3-A4

Contents

<i>Preface</i>	<i>vii</i>
<i>Contributors</i>	<i>xiii</i>

PART I CHEM-BIOINFORMATIC TOOLS

Cheminformatics Approaches to Study Drug Polypharmacology	3
<i>J. Jesús Naveja, Fernanda I. Saldívar-González, Norberto Sánchez-Cruz, and José L. Medina-Franco</i>	
Computational Predictions for Multi-Target Drug Design	27
<i>Neelima Gupta, Prateek Pandya, and Seema Verma</i>	
Computational Multi-Target Drug Design	51
<i>Azizeh Abdolmaleki, Fereshteh Shiri, and Jahan B. Ghasemi</i>	

PART II COMPUTATIONAL MULTI-TARGET DRUG DESIGN: LITERATURE REVIEWS

Multitarget Drug Design for Neurodegenerative Diseases	93
<i>Marco Catto, Daniela Trisciuzzi, Domenico Alberga, Giuseppe Felice Mangiatordi, and Orazio Nicolotti</i>	
Molecular Docking Studies in Multitarget Antitubercular Drug Discovery	107
<i>Jéssika de Oliveira Viana, Marcus T. Scotti, and Luciana Scotti</i>	
Advanced Chemometric Modeling Approaches for the Design of Multitarget Drugs Against Neurodegenerative Diseases	155
<i>Amit Kumar Halder, Ana S. Moura, and M. Natália D. S. Cordeiro</i>	
Computational Studies on Natural Products for the Development of Multi-target Drugs	187
<i>Veronika Temml and Daniela Schuster</i>	
Computational Design of Multitarget Drugs Against Alzheimer's Disease	203
<i>Sotirios Katsamakas and Dimitra Hadjipavlou-Litina</i>	
Design of Multi-target Directed Ligands as a Modern Approach for the Development of Innovative Drug Candidates for Alzheimer's Disease	255
<i>Cindy Juliet Cristancho Ortiz, Matheus de Freitas Silva, Vanessa Silva Gontijo, Flávia Pereira Dias Viegas, Kris Simone Tranches Dias, and Claudio Viegas Jr.</i>	

PART III CASE STUDIES

Virtual Screening for Dual Hsp90/B-Raf Inhibitors	355
<i>Andrew Anighoro, Luca Pinzi, Giulio Rastelli, and Jürgen Bajorath</i>	

Strategies for Multi-Target Directed Ligands: Application in Alzheimer's Disease (AD) Therapeutics	367
<i>Sucharita Das and Soumalee Basu</i>	
Computational Design of Multi-target Kinase Inhibitors	385
<i>Sinoy Sugunan and Rajanikant G. K.</i>	
Proteochemometrics for the Prediction of Peptide Binding to Multiple HLA Class II Proteins.	395
<i>Ivan Dimitrov, Ventsislav Jordanov, Darren R. Flower and Irini Doytchinova</i>	
Linked Open Data: Ligand-Transporter Interaction Profiling and Beyond.	405
<i>Stefanie Kickinger, Eva Hellsberg, Sankalp Jain, and Gerhard F. Ecker</i>	
Design of Novel Dual-Target Hits Against Malaria and Tuberculosis Using Computational Docking	419
<i>Manoj Kumar and Anuj Sharma</i>	
Computational Design of Multi-Target Drugs Against Breast Cancer.	443
<i>Shubhendra Tripathi, Gaurava Srivastava, and Ashok Sharma</i>	
Computational Methods for Multi-Target Drug Designing Against <i>Mycobacterium tuberculosis</i>	459
<i>Gaurava Srivastava, Ashish Tiwari, and Ashok Sharma</i>	
PART IV DATABASES AND WEB SERVERS	
Development of a Web-Server for Identification of Common Lead Molecules for Multiple Protein Targets	487
<i>Abhilash Jayaraj, Ruchika Bhat, Amita Pathak, Manpreet Singh and B. Jayaram</i>	
Computational Method for Prediction of Targets for Breast Cancer Using siRNA Approach	505
<i>Atul Tyagi, Mukti N. Mishra, and Ashok Sharma</i>	
PART V SPECIAL TOPICS	
Historeceptomics: Integrating a Drug's Multiple Targets (Polypharmacology) with Their Expression Pattern in Human Tissues	517
<i>Timothy Cardozo</i>	
Networking of Smart Drugs: A Chem-Bioinformatic Approach to Cancer Treatment.	529
<i>Kavindra Kumar Kesari, Qazi Mohammad Sajid Jamal, Mohd. Haris Siddiqui, and Jamal Mohammad Arif</i>	
<i>Index</i>	557

Contributors

- AZIZEH ABDOLMALEKI • *Department of Chemistry, Tuyserkhan Branch, Islamic Azad University, Tuyserkhan, Iran*
- DOMENICO ALBERGA • *Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari “Aldo Moro”, Bari, Italy*
- ANDREW ANIGHORO • *Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms- Universität, Bonn, Germany*
- JAMAL MOHAMMAD ARIF • *Department of Bioscience, Integral University, Lucknow, Uttar Pradesh, India*
- JÜRGEN BAJORATH • *Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany*
- SOUMALEE BASU • *Department of Microbiology, University of Calcutta, Kolkata, West Bengal, India*
- RUCHIKA BHAT • *Department of Chemistry, Indian Institute of Technology Delhi, New Delhi, India; Supercomputing Facility for Bioinformatics & Computational Biology, Indian Institute of Technology Delhi, New Delhi, India*
- TIMOTHY CARDOZO • *New York University School of Medicine, NYU Langone Health, New York, NY, USA*
- MARCO CATTO • *Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari “Aldo Moro”, Bari, Italy*
- M. NATÁLIA D. S. CORDEIRO • *LAQV@REQUIMTE/Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, Porto, Portugal*
- SUCHARITA DAS • *Department of Microbiology, University of Calcutta, Kolkata, West Bengal, India*
- KRIS SIMONE TRANCHES DIAS • *PeQuiM, Laboratory of Research in Medicinal Chemistry, Institute of Chemistry, Federal University of Alfenas, Alfenas, Brazil*
- IVAN DIMITROV • *Faculty of Pharmacy, Medical University of Sofia, Sofia, Bulgaria*
- IRINI DOYTCHINOVA • *Faculty of Pharmacy, Medical University of Sofia, Sofia, Bulgaria*
- GERHARD F. ECKER • *Department of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria*
- DARREN R. FLOWER • *School of Life and Health Sciences, Aston University, Birmingham, UK*
- MATHEUS DE FREITAS SILVA • *PeQuiM, Laboratory of Research in Medicinal Chemistry, Institute of Chemistry, Federal University of Alfenas, Alfenas, Brazil; Programa de Pós-Graduação em Química, Federal University of Alfenas, Alfenas, Brazil*
- RAJANIKANT G. K. • *School of Biotechnology, National Institute of Technology Calicut, Calicut, Kerala, India*
- JAHAN B. GHASEMI • *Drug Design in Silico Lab, Chemistry Faculty, University of Tebran, Tebran, Iran*
- VANESSA SILVA GONTIJO • *PeQuiM, Laboratory of Research in Medicinal Chemistry, Institute of Chemistry, Federal University of Alfenas, Alfenas, Brazil; Programa de Pós-Graduação em Química, Federal University of Alfenas, Alfenas, Brazil*

- NEELIMA GUPTA • *Centre of Advanced Study, Department of Chemistry, University of Rajasthan, Jaipur, India*
- DIMITRA HADJIPAVLOU-LITINA • *Department of Pharmaceutical Chemistry, School of Pharmacy, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece*
- AMIT KUMAR HALDER • *LAQV@REQUIMTE/Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, Porto, Portugal*
- EVA HELLSBERG • *Department of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria*
- SANKALP JAIN • *Department of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria*
- QAZI MOHAMMAD SAJID JAMAL • *Department of Health Information Management, College of Applied Medical Sciences, Buraydah Colleges, Buraydah, Al Qassim, Saudi Arabia; Novel Global Community Educational Foundation, Sydney, NSW, Australia*
- ABHILASH JAYARAJ • *Department of Chemistry, Indian Institute of Technology Delhi, New Delhi, India; Supercomputing Facility for Bioinformatics & Computational Biology, Indian Institute of Technology Delhi, New Delhi, India*
- B. JAYARAM • *Department of Chemistry, Indian Institute of Technology Delhi, New Delhi, India; Supercomputing Facility for Bioinformatics & Computational Biology, Indian Institute of Technology Delhi, New Delhi, India; Kusuma School of Biological Sciences, Indian Institute of Technology Delhi, New Delhi, India*
- J. JESÚS NAVEJA • *Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, Mexico City, Mexico; PECEM, School of Medicine, Universidad Nacional Autónoma de México, Mexico City, Mexico*
- SOTIRIOS KATSAMAKAS • *Department of Pharmaceutical Chemistry, School of Pharmacy, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece*
- KAVINDRA KUMAR KESARI • *Department of Applied Physics, Aalto University, Helsinki, Finland; Department of Bioproduct and Biosystem, Aalto University, Helsinki, Finland*
- STEFANIE KICKINGER • *Department of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria*
- MANOJ KUMAR • *Department of Chemistry, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, India; Department of Chemistry and Chemical Biology, McMaster University, Hamilton, ON, Canada*
- GIUSEPPE FELICE MANGIATORDI • *Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari "Aldo Moro", Bari, Italy*
- JOSÉ L. MEDINA-FRANCO • *Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, Mexico City, Mexico*
- MUKTI N. MISHRA • *Biotechnology Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, Uttar Pradesh, India*
- ANA S. MOURA • *LAQV@REQUIMTE/Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, Porto, Portugal*
- ORAZIO NICOLOTTI • *Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari "Aldo Moro", Bari, Italy*
- JÉSSICA DE OLIVEIRA VIANA • *Federal University of Paraíba, Health Center, João Pessoa, PB, Brazil*
- CINDY JULIET CRISTANCHO ORTIZ • *PeQuiM, Laboratory of Research in Medicinal Chemistry, Institute of Chemistry, Federal University of Alfenas, Alfenas, Brazil; Programa de Pós-Graduação em Química, Federal University of Alfenas, Alfenas, Brazil*

- PRATEEK PANDYA • *Amity Institute of Forensic Sciences, Amity University, Noida, India*
- AMITA PATHAK • *Department of Chemistry, Indian Institute of Technology Delhi, New Delhi, India; Supercomputing Facility for Bioinformatics & Computational Biology, Indian Institute of Technology Delhi, New Delhi, India*
- LUCA PINZI • *Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy*
- GIULIO RASTELLI • *Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy*
- FERNANDA I. SALDÍVAR-GONZÁLEZ • *Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, Mexico City, Mexico*
- DANIELA SCHUSTER • *Institute of Pharmacy/Pharmacognosy and Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria; Department of Pharmaceutical and Medicinal Chemistry, Institute of Pharmacy, Paracelsus Medical University Salzburg, Salzburg, Austria*
- LUCIANA SCOTTI • *Federal University of Paraíba, Health Center, Teaching and Research Management—University Hospital, João Pessoa, PB, Brazil*
- MARCUS T. SCOTTI • *Federal University of Paraíba, Health Center, João Pessoa, PB, Brazil*
- ANUJ SHARMA • *Department of Chemistry, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, India*
- ASHOK SHARMA • *Biotechnology Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, Uttar Pradesh, India*
- FERESHTEH SHIRI • *Department of Chemistry, University of Zabol, Zabol, Iran*
- MOHD. HARIS SIDDIQUI • *Department of Bioengineering, Integral University, Lucknow, Uttar Pradesh, India*
- MANPREET SINGH • *Supercomputing Facility for Bioinformatics & Computational Biology, Indian Institute of Technology Delhi, New Delhi, India*
- GAURAVA SRIVASTAVA • *Biotechnology Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, Uttar Pradesh, India*
- SINOY SUGUNAN • *School of Biotechnology, National Institute of Technology Calicut, Calicut, Kerala, India*
- NORBERTO SÁNCHEZ-CRUZ • *Department of Pharmacy, School of Chemistry, Universidad Nacional Autónoma de México, Mexico City, Mexico*
- VERONIKA TEMML • *Institute of Pharmacy/Pharmacognosy and Center for Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria*
- ASHISH TIWARI • *Biotechnology Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, Uttar Pradesh, India*
- SHUBHANDRA TRIPATHI • *Biotechnology Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, Uttar Pradesh, India*
- DANIELA TRISCIUZZI • *Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari “Aldo Moro”, Bari, Italy*
- ATUL TYAGI • *Biotechnology Division, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, Uttar Pradesh, India*
- SEEMA VERMA • *Centre of Advanced Study, Department of Chemistry, University of Rajasthan, Jaipur, India*
- FLÁVIA PEREIRA DIAS VIEGAS • *PeQuiM, Laboratory of Research in Medicinal Chemistry, Institute of Chemistry, Federal University of Alfenas, Alfenas, Brazil*

- CLAUDIO VIEGAS JR. • *PeQuiM, Laboratory of Research in Medicinal Chemistry, Institute of Chemistry, Federal University of Alfenas, Alfenas, Brazil; Programa de Pós-Graduação em Química, Federal University of Alfenas, Alfenas, Brazil*
- VENTSISLAV YORDANOV • *Faculty of Pharmacy, Medical University of Sofia, Sofia, Bulgaria*

Part I

Chem-Bioinformatic Tools



Cheminformatics Approaches to Study Drug Polypharmacology

J. Jesús Naveja, Fernanda I. Saldívar-González, Norberto Sánchez-Cruz, and José L. Medina-Franco

This work is dedicated to the loving memory of Nicolás Medina Sandoval.

Abstract

Herein is presented a tutorial overview on selected chemoinformatics methods useful for assembling, curating/preparing a chemical database, and assessing its diversity and chemical space. Methods for evaluating the structure–activity relationships (SAR) and polypharmacology are also included. Usage of open source tools is emphasized. Step-by-step KNIME workflows are used for illustrating the methods. The methods described in this chapter are applied onto a chemical database especially relevant for epipolypharmacology that is an emerging area in drug discovery. However, the methods described herein could be extended to other therapeutic areas and potentially to other areas of chemistry.

Keywords Chemoinformatics, ChemMaps, Chemical space, Data mining, Epigenetics, Epinformatics, KNIME, Molecular diversity, Open-access, Polypharmacology, Structure–activity relationships, SmARt

1 Introduction

The rapid growth of chemical information demands efficient and reliable computational algorithms to analyze the accumulated data. Similarly, current trends in drug discovery such as polypharmacology [1, 2] demand the organization and efficient mining of multiple drug–target interactions and study of structure–multiple activity relationships (SMARt) efficiently [3]. Indeed, a plethora of methods and resources for exploiting SMARt and other data relevant to polypharmacology have been published, and many of them are open access [4]. This review includes methodological details for implementing scalable KNIME cheminformatics workflows for:

- a. Curating a chemical database;
- b. Computing chemical descriptors;

Electronic supplementary material: The online version of this article (https://doi.org/10.1007/7653_2018_6) contains supplementary material, which is available to authorized users.

- c. Analyzing and comparing database diversity, and
- d. Visualizing their chemical space.

Of note, KNIME is an open-access initiative intended for generating data mining pipelines or workflows, which are capable of integrating multiple tools [5].

Although sufficiently detailed, this review aims at being a quick practical guide. More comprehensive tutorials in chemoinformatics can be found elsewhere [6, 7]. Additionally, web applications for cheminformatics methods that have been developed by our research group are mentioned in the respective subsections. These applications are part of the D-Tools initiative for generating open cheminformatics resources (available at <https://www.difacquim.com/d-tools/>). The D-Tools usage is further described elsewhere [4, 8–11], and these are not the focus of this review.

2 Methods

2.1 Construction and Curation of a Compound Database

Due to the increase in the amount of chemical information, where it is common to the concept of big data [12], the efficient management of information represents a challenge today. This is of particular importance in polypharmacology where large compound datasets contain information of screening across several biological endpoints. In response to this need, the construction of compound and other databases can be a convenient way to sort information according to the data available and the specific objectives of the study.

In chemoinformatics, construction of databases is a fundamental practice to perform various computational studies like the design of chemical libraries, characterization and comparison of the chemical space, the study of the structure–activity relationships (SAR), and virtual screening studies, among others.

Currently, web pages of large public databases such as DrugBank [13], ChEMBL [14], ZINC [15], and BindingDB [16] allow the user to download their own databases (complete or partial downloads) with information on approved drugs, drugs in the experimental phase, commercially available compounds, molecular targets, etc. However, these databases are not always updated, so they can be enriched with new information published in books or in scientific articles.

Also, in research groups devoted to the synthesis, isolation from natural sources and/or evaluation of new chemical entities can be carried out for the construction of completely new compounds' databases. Such collections are usually referred to as in-house databases.

The process of building and annotating chemical databases is not trivial. Each organization has its own rules, conventions, and

procedures. However, the steps that are considered essential are listed below:

1. Identify compounds and resources that contain information required, e.g., journals and databases with chemical information [4, 17].
2. In a spreadsheet, it is recommended that the user has the following information for each compound:
 - a. Name of each compound. This can be searched in public databases.
 - b. A number that identifies this compound in the database that has been consulted, for example, ChemSpider ID, Substance or Compound ID (SID, CID in PubChem, the CAS registry number, or an internal and consistent code if building an in-house collection).
 - c. Structure input. An example of this is the use of Canonical SMILES notation used for encoding molecular structures that can be imported to other molecular editing systems. It is worth noting the relevance of creating a single computational representation. This can be achieved by using various algorithms in a process known as canonicalization.
3. Once this information is collected in the spreadsheet, save the database preferably in *.csv* format (comma delimited). Other database formats with chemical information and compatible with most computer programs as KNIME are *sdf* (structure data file), *mol* (molecular data file), and *mol2* (tripos mol2 file).

For the management and analysis of databases, the KNIME Example Server provides access to many explanatory workflows. The example server is accessible via the KNIME Explorer panel within the KNIME workbench and represents a great help when starting a new workflow.

Some of the nodes to start working with files with chemical information are: *Molecule Type Cast*, a node useful for reading chemical data from a *.csv* file or database, and this node casts any string as a chemical type (i.e., It tells KNIME “This is a smiles string”) and *Marvin MolConverter*, a node provided by Chemaxon/Infocom that translates seamlessly between types (smiles \leftrightarrow sdf \leftrightarrow mrv).

An important aspect to consider when analyzing molecular databases generated by other scientists is that these may contain wrong information or unnecessary information for the intended application or project. Therefore, cleaning or curating the information is highly relevant to enhance the quality of the data and to avoid erroneous results [18].

As in the construction of databases, there is no widely accepted standard protocol for the preparation of small molecules. However,

hereunder are described the essential points in the preparation and curation of databases:

1. Normalize the chemical structures. In this step, each chemical structure is checked for valid atom types, valence checks, and functional groups such as nitro groups are converted to a consistent representation. This is followed by a standardization step in which chemical structures are converted to a canonical tautomeric form, aromatic structures are kekulized, placement of stereo bonds is standardized, and all implicit hydrogens are converted to explicit hydrogens [19].
2. Remove duplicates. After the molecules have been properly standardized, it is appropriate to detect duplicates. InChiKeys is a useful method to identify several states of protonation and tautomers of a molecule.
3. Discard inorganic and organometallic atoms or molecules if these are not the object of study. It is worth mentioning that the majority of the chemoinformatics programs currently available are developed to process small organic molecules.
4. Wash the compound database by applying to each molecule a set of rules of “cleaning” such as the elimination of salts and the adjustment of the protonation states. The purpose of this step is to ensure that each chemical structure is in a form suitable for the subsequent modeling.
5. Enumerate tautomers and stereoisomers. This step is important in virtual screening studies, particularly when using search methods such as docking or pharmacophore.
6. Optimize the geometry and minimize the energy if the database will be used to evaluate the potential of each compound to bind to a receptor or enzyme, or to calculate descriptors which depend on the three-dimensional conformation of the molecule. The specific method to optimize the geometry will largely depend on the type, quantity of molecules to optimize, and, most importantly, on the specific application.

In addition, if the quantity of compounds is too large to be examined or tested with the resources available, different strategies can be employed to reduce the number of compounds in a rational and consistent manner. Such strategies include: filtering—essentially imposing secondary search criteria to eliminate compounds, clustering—taking a representative subset of a larger set, and human inspection of the compound structures (with or without extra data) [20].

In several articles, the impact of the use of duplicates and inconsistencies in the molecular structures in prediction models had already been discussed [21]. For this reason, the project CERAPP (Collaborative Estrogen Receptor Activity Prediction Project) has

developed a workflow to curate databases [22]. A similar workflow can be found at the link <https://github.com/zhu-lab/curation-workflow/blob/master/Structure%20Standardizer2.zip>.

Gally et al. also report a workflow designed to prepare molecular databases but focused on studies of virtual screening [23]. In addition to carrying out of the standardization of chemical structures, the workflow of Gally et al. has implemented filters (based on molecular property distribution) to characterize specific subsets of chemical libraries such as drug-like, lead-like, or fragment-like subsets of compounds.

See Workflow 1 in the Supplementary Information for an example in KNIME.

The following analyses use an epigenomics chemical database that has already been curated and published [24].

2.2 Diversity Analysis

In drug discovery projects focused on one single target or multiple targets, it is of high relevance quantifying the structural diversity of compound datasets. For instance, if the goal of a high-throughput screening campaign is to identify hit compounds with a desirable polypharmacological profile, it is desirable to screen a compound collection with high diversity. This will increase the possibilities to find active molecules with a desirable profile. If the goal of the screening campaign is to further develop a focused library (e.g., increase the structure–activity information of a focused region in chemical space [25]), it is desirable to screen a compound dataset with high internal similarity (low diversity).

The diversity in a chemical library can be assessed in multiple ways, mainly depending on the data under scrutiny. In addition to the diversity metric, a key aspect of diversity analysis is molecular representation [26, 27]. The most common ways to represent molecules in chemoinformatic applications are molecular descriptors (including physicochemical properties and molecular fingerprints), and chemical scaffolds [28]. Depending on the type of descriptor and the level of accuracy desired (considering the time of computation and the number of compounds to analyze), the input structures can be in two or three dimensions (the latter requires conformational analysis). The choice of molecular representation depends on the goals of the study.

A more detailed description on how to use molecular descriptors and scaffolds as an input for diversity analysis follows in the next paragraphs. See Workflow 2 in the Supplementary Information for an exemplary diversity analysis in KNIME.

2.2.1 Molecular Descriptors

Molecular descriptors capture information of the whole molecule and are usually straightforward to interpret. Also, whole molecular properties such as physicochemical properties of pharmaceutical interest are usually part of empirical rules for drug likeness that aids to guide drug discovery programs. KNIME includes RDKit,

CDK, and Indigo nodes, with which complexity descriptors (e.g., chiral carbons, and fraction of sp^3 carbon atoms), and physicochemical properties of pharmaceutical interest (including molecular weight, number of hydrogen bond donors and acceptors, number of rotatable bonds, logarithm of octanol–water partition coefficient, and topological polar surface area) [28].

Starting with curated databases (discussed in Sect. 2.1), the steps for quantifying diversity with molecular descriptors are:

1. Select the features to be evaluated (usually the six commonest physicochemical properties of pharmaceutical relevance, *vide supra*).
2. Scale the data using a Z -transformation. This transforms the data to dimensional units. The purpose is to improve the comparability of the variables and give a similar weight to all of them independently of the units with which they were originally measured.
3. Compute pairwise euclidean distance. For a database with n compounds, $n \times (n - 1)/2$ pairwise comparisons are to be computed. Euclidean distance is calculated with the formula:

$$D(A, B) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2},$$

where $D(A, B)$ is the euclidean distance between compound A and B, a_i and b_i are the i -th descriptor, and n the total number of descriptors [29]. $D(A, B)$ can take any positive real number as value.

4. Compute a central tendency statistic (e.g., mean or median) for all the pairwise comparisons. The larger the mean or median, the more diverse the dataset is [30].
5. Finally, for comparison, the statistic can be computed for other reference databases or looked up at the literature if already reported.

2.2.2 Molecular Fingerprints

Many structural features escape the very general information obtained with physicochemical and complexity descriptors. Molecular fingerprints are vectors that aim towards a more comprehensive set of features (usually more than a hundred) to compare molecules. Every feature is encoded as a Boolean variable, where “0” represents absence and “1” represents presence of the feature. Therefore, repeated motifs are not generally acknowledged. For every molecule, a Boolean vector of features is obtained, and these are susceptible of standard set operations [31–33]. However, molecular fingerprints do have limitations, for example, they could be more difficult to interpret intuitively, and therefore pose a greater difficulty for extracting insights relevant for medicinal chemistry.

The steps for computing diversity based on fingerprints are:

1. Select a molecular fingerprint. Although the selection of the “best” fingerprint could be different from case to case, it has been consistently found that MACCS keys 166-bits [34] are useful for quantifying database diversity. In turn, extended connectivity fingerprints of diameter 4 (ECFP4) [32] as well as other circular fingerprints are, overall, better suited for virtual screening, activity landscape modeling, and SAR studies in general.
2. Compute pairwise Tanimoto similarity [27, 35]. For a database with n compounds, $n \times (n - 1)/2$ pairwise comparisons are to be computed. Tanimoto similarity is calculated with the expression:

$$T(A, B) = \frac{c}{a + b - c}$$

where $T(A, B)$ is Tanimoto similarity with possible values being any real number between 0 and 1, c is the number of features for which both molecules A and B have a “1” value, a is the number of features for which molecule A has a “1” value, and b is the number of features for which molecule B has a “1” value. Dissimilarity matrices implemented in KNIME are quite efficient at these calculations. However, by default they compute values as dissimilarities, the complement of similarities, or distance matrices. Conversion from Tanimoto dissimilarity to similarity is accomplished by just subtracting the value from 1 ($T_s = 1 - T_d$, where T_s is Tanimoto similarity and T_d is Tanimoto dissimilarity).

3. Compute a central tendency statistic (e.g., mean or median) for all the pairwise comparisons. Conversely to Euclidean distance (and any distance metric in general), the smaller the mean or median, the more diverse the dataset is [30].
4. Finally, for comparison, the statistic can be computed for other reference databases or looked up at the literature if already reported.

2.2.3 Molecular Scaffolds

KNIME has nodes for finding Murcko scaffolds [36, 37]. By definition, Murcko scaffolds contain all the cyclic systems in a molecule as well as the linkers between them. All other decorations and ramifications are omitted. The greatest benefit of working with scaffolds data is that, unlike molecular fingerprints, they are readily interpreted by medicinal chemists. Nonetheless, the representation is rougher and loses information from the side chains. Also, more advanced methods must be applied to account for the structural relations among the scaffolds.

It is logical and generally accepted that a dataset is more diverse when it has a large number of different scaffolds, and the proportions of compounds with each scaffold are evenly distributed. The procedure for measuring scaffold diversity is as follows:

1. Find Murcko scaffolds for every molecule in the dataset.
2. Compute a frequency table of the scaffolds.
3. From here, there are a number of different methods for assessing the diversity [38]:
 - a. Order the scaffolds by their frequency of occurrence and compute the median (i.e., the minimum number of scaffolds in the database that contain at least 50 % of the total entries). Lower values in this statistic mean higher diversity.
 - b. Order the scaffolds by their frequency of occurrence. This order would be an index from 1 to n , where n is the total number of different scaffolds in the dataset. Divide all indexes by n , such that the highest index value is 1. Using scaffold indexes in the x -axis and their respective cumulative proportions in the y -axis, compute the area under the curve as a diversity statistic. This statistic admits as value any real number in the domain [0.5, 1.0]. Lower values in this statistic mean higher diversity.
 - c. Compute scaled Shannon entropy (SSE) with the formula:

$$\text{SSE} = \frac{\text{SE}}{\log_2 n'}$$

where $\text{SE} = \sum_{i=1}^n -p_i \log_2 p_i$,

where p_i is the proportion in the dataset of the i -th scaffold (calculated by dividing the occurrence of this i -th scaffold by the total number of entries/molecules), SE is the Shannon entropy, and n is the total number of scaffolds in the dataset. SSE takes as value a real number in the range [0,1]. For this statistic, higher values mean higher scaffold diversity.

4. Finally, the statistic can be computed for other reference databases for comparison.

2.2.4 Consensus Diversity Plots

In the light of numerous variables that can be used to quantify diversity, visual representations have been built in order to summarize multiple of them simultaneously. These are the consensus diversity plots (CDPs). A CDP, as defined by González-Medina et al. [10], renders 2D diversity measured by scaffolds, fingerprints, physicochemical properties, and the number of compounds in the databases. It is also possible to integrate 3D data [24]; however, we will not emphasize on 3D data usage here. The steps required for plotting a CDP from data are:

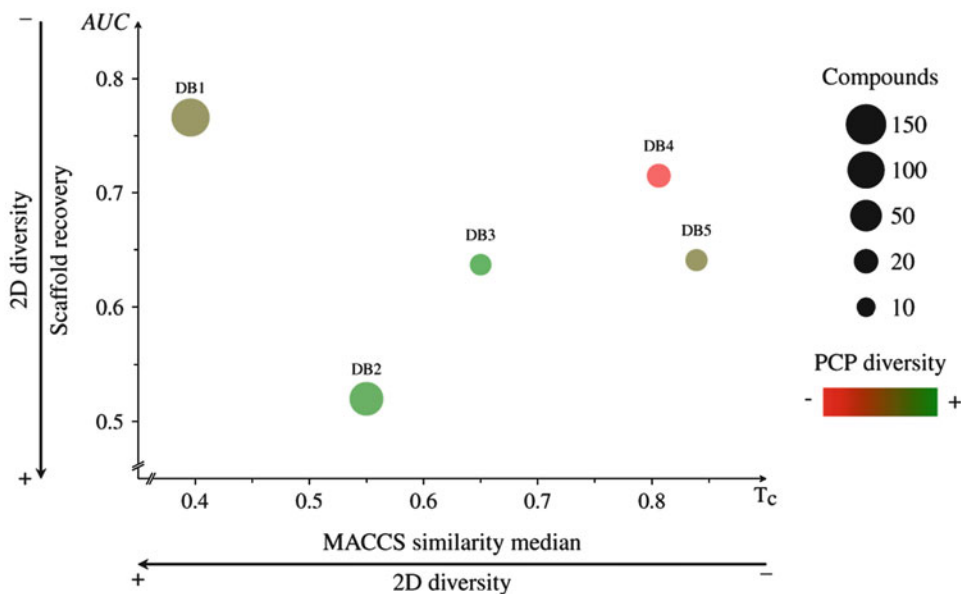


Fig. 1 An exemplary consensus diversity plot (CDP). Each data point represents a compound database. Molecular fingerprints diversity is plotted in the x -axis, the scaffold diversity in the y -axis, the physicochemical properties diversity in a color continuous scale, and the relative number of compounds in the database as the data point size. *AUC* area under the curve, *PCP* physicochemical properties

1. Curate databases; calculate diversity with physicochemical properties, molecular fingerprints, and scaffolds (see above for details).
2. Plot the molecular fingerprints diversity in the x -axis, the scaffold diversity in the y -axis, the physicochemical properties in a color continuous scale, and the number of compounds in the database as the data point size. Every data point represents a database. (See Fig. 1 and Supplementary KNIME Workflow 3 for a few examples.)

As an alternative, an online server was developed for generating CDPs and is also available in D-Tools (see Sect. 1). A video tutorial is available at <https://youtu.be/lruo1ypKGBE>, and detailed written instructions about how to use it can be found at <http://132.248.103.152:3838/CDPlots/>.

2.3 Structure–Activity Relationship Analysis

A common assumption in virtual screening is that similar molecules are expected to have similar properties, e.g., comparable biological activity. This assumption is called the similarity principle. Although virtual screening is often useful for detecting active compounds, it is reassuring to verify whether the similarity principle is valid for the molecules under scrutiny. Such insights can be obtained through a subtype of SAR analysis, activity landscape modeling. SAR analysis of chemical libraries, for which activity against a biological target is

known, can also reveal substructures that are relevant for inhibiting the target in question. The next paragraphs give details onto some useful methods for assessing SAR of single and multiple libraries simultaneously. Workflow 4 in the Supplementary Information illustrates a KNIME implementation of the methods described below.

2.3.1 Structure–Activity Similarity Maps

Structure–activity similarity (SAS) maps are bidimensional activity landscape representations that contrast structural similarity (e.g., measured with Tanimoto coefficient of molecular fingerprints) and activity similarity (for example, as pIC_{50} or pKi). Systematic pairwise compound comparisons are included in the plot [39]. Each point in a SAS map represents a pair of compounds and is colored according to the most active compound of the pair. The sequence of steps for generating and ultimately interpreting a SAS map is as follows:

1. Given n compounds in a library, compute the $n \times (n - 1)/2$ paired chemical similarity as described in Sect. 2.2.2.
2. Similarly, for the same paired comparisons calculate the absolute difference in potency. All compounds should have potency in pIC_{50} units. It is calculated from IC_{50} measurements in nanomolar concentration units with the formula (ideally, all compounds should have IC_{50} values measured under the same protocol and assay conditions):

$$\text{pIC}_{50} = -\log_{10}(\text{IC}_{50} [\text{nM}]).$$

3. Plot the structural similarity in the x -axis and the potency difference in the y -axis. The color of the data points can also be set to render more information, for example, the maximum potency value in the pair.
4. The resultant plot, illustrated in Fig. 2, can be divided into four quadrants with thresholds defined a priori: (a) smooth (high structural similarity and low activity difference), (b) activity cliffs (high structural similarity but high activity difference), (c) scaffold hops (low structural similarity but low activity difference), and (d) uncertainty (low structural similarity and high activity difference) [40–42]. Typical potency thresholds are 2 for deep activity cliffs and 1 for shallow activity cliffs. In the case of structural similarity, 1 or 2 standard deviations above the mean could be used.

Alternatively, a web application for plotting SAS maps can be found at D-Tool under <https://unam-shiny-difacquim.shinyapps.io/ActLSmaps/>. A video tutorial is available at <https://youtu.be/52jHCcg5mXU>.

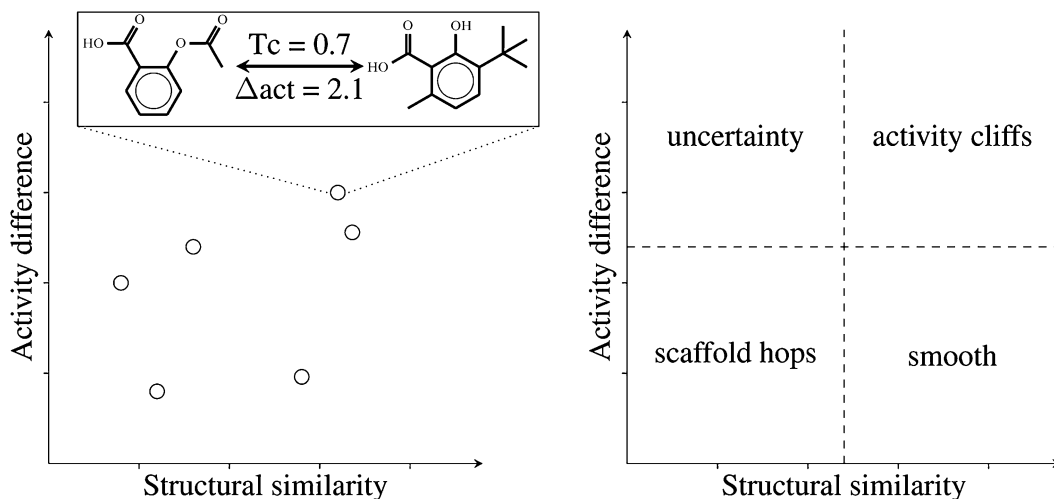


Fig. 2 Structure–activity similarity (SAS) maps. Each data point represents a pair of compounds. The *x*-axis plots the structural similarity, while the *y*-axis plots the activity difference. Four quadrants are formed as described in Sect. 2.3.1. A color scale might be added to represent density of points or the maximum activity value in the pair. *Tc* Tanimoto coefficient

2.3.2 Scaffold Enrichment Factor

SAR can also be explored based on chemical scaffolds. For every dataset with activity annotations against a particular biological target, every scaffold could be considered as a cluster of molecules. At this point, it is interesting to find which clusters have a higher or lesser proportion of active molecules, pointing towards clusters of highly related molecules that tend to be more or less active than the average. This is the basis of the calculation of enrichment factors (EF) for scaffolds, which are obtained as follows:

1. If activity is represented quantitatively in the dataset, a threshold of activity should be set a priori. Often, a pIC_{50} of 5–6 or more is useful for defining a compound as active.
2. Essentially, the EF is an odds ratio, i.e., a ratio of proportions. Specifically, the proportion of active compounds with a given scaffold is divided by the proportion of active compounds in the general dataset. A more formal definition would be that, for every scaffold λ , an EF is calculated using the equation [43]:

$$\text{EF}(C_\lambda) = \frac{\text{Act}(C_\lambda)}{\text{Act}(C)}$$

$$\text{where } \text{Act}(C_\lambda) = \frac{|C_\lambda^+|}{|C_\lambda|}$$

$$\text{and } \text{Act}(C) = \frac{|C^+|}{|C|},$$

where, in turn, C is the total number of compounds tested, C^+ the number of compounds active, C_λ the number of total compounds with a scaffold λ tested, and C_λ^+ the number of

compounds with a scaffold λ active against the target. Values above 1 imply a positively enriched scaffold (i.e., a scaffold that has a higher proportion of active compounds than the general dataset), while values below 1 have the opposite meaning.

3. EFs are susceptible of hypothesis testing. For finding statistically significant enriched scaffolds, chi-squared tests can be run using a 2×2 contingency table for the compounds considering as variables whether they have a given scaffold and whether they are active. Since sometimes values in the cells might be lesser than 5, and this interferes with the analytic calculation of the chi-squared statistic, simulated values can be obtained.
4. After running all p -values for every scaffold, the false discovery rate correction (or other method for correcting for multiple hypothesis testing) should be applied.

2.3.3 Degree of Polypharmacology

The methods for SAR analysis mentioned above are useful for single target studies. However, sometimes inhibition data of multiple targets are available for single compounds. These data could lead to polypharmacology studies. Maggiora and Gokhale recently formalized the notion of polypharmacology and polyspecificity [44]. In practical terms, the degree of polypharmacology of a molecule equals the number of different targets against which the molecule is active, while the analogous degree of polyspecificity of a target equals the number of different molecules that are active against the target.

2.3.4 Multiple Structure–Activity Relationship Analysis

A review addressing SmARt analysis in epigenetics was recently published [3]. Two of the most useful SmARt tools are methodologically explained in the following paragraphs: dual-activity difference (DAD) maps and structure–promiscuity index difference (SPID). Similarly as for other SAR analyses, Workflow 4 in the Supplementary Information contains practical tools for computing them.

Dual-Activity Difference Maps

DAD maps are designed to compare at once the activity of compounds against two biological endpoints, in contrast to SAS maps [45]. However, DAD maps lose structural information, which is accounted for with SAS maps. The procedure for generating a DAD map is straightforward:

1. Select a library of compounds with the activity of each independently measured against two different endpoints.
2. Plot in the x -axis one of the measurements and on the y -axis the other. A general form of a DAD map is presented in Fig. 3.

Structure–Promiscuity Index Difference

Aiming towards a statistic for quantifying the relationship between structural similarity and polypharmacology (or promiscuity), the SPID was created [46]. It is computed with the formula:

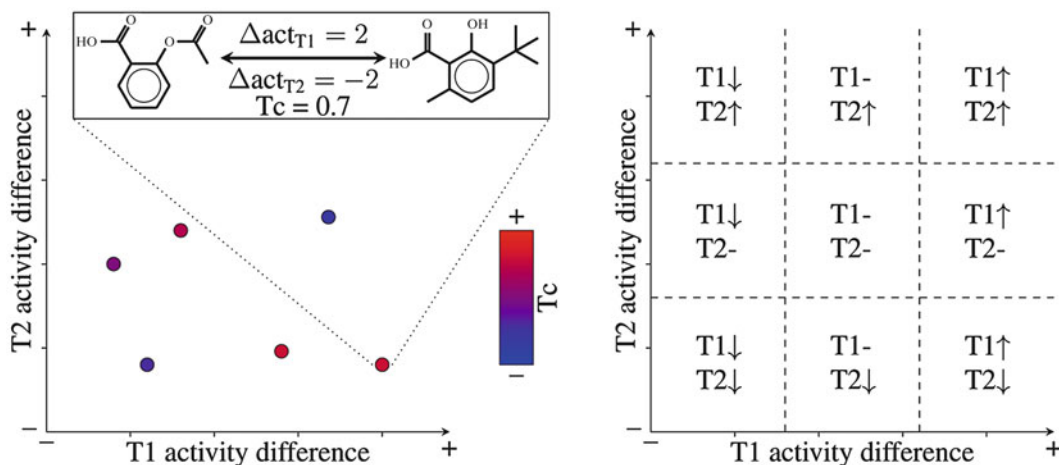


Fig. 3 Dual-activity difference (DAD) maps. Each data point represents a pair of compounds. The *x*-axis plots the activity difference of target 1, while the *y*-axis the activity difference of target 2. A color continuous scale might be added to the plot to represent chemical similarity of each pair of compounds. Up to nine regions can be distinguished depending on whether activity is conserved, increased, or decreased for any of the two targets. *T_c* Tanimoto coefficient, *T1* target 1, *T2* target 2

$$SPID(A, B) = \frac{|P_A - P_B|}{1 - T(A, B)}$$

where *A* and *B* are chemical compounds, P_A and P_B are the potencies of compounds *A* and *B*, respectively, and $T(A, B)$ is the Tanimoto similarity of compounds *A* and *B* computed as in Sect. 2.2.2 using molecular fingerprints.

3 Chemical Space

Visual representations of the relationships of the compounds in a database are often useful for assessing libraries' diversity and SAR. Furthermore, the recent development of database fingerprints (DFPs) (described below) has made easier to chart multiple target-focused libraries in the chemical space, thereby providing polypharmacology insights [24]. Workflow 5 in the Supplementary Information illustrates a KNIME implementation of the methods described in this section.

3.1 Principal Components Analysis for Charting Compounds

There are no universal methods for chemical space representations [47, 48]. A commonly used approach involves calculating similarity matrices, which capture all the pairwise comparisons. These matrices are squared and have *n* columns and rows, with *n* equal to the number of compounds in the dataset. Finally, principal components analysis (PCA) as well as other dimensionality reduction methods is useful to compress most of the relevant information in a few

variables. This makes possible to obtain visualizations of the chemical space. The concrete steps for creating visualizations of the chemical space using the approach presented above are as follows:

1. Select the set of descriptors with which the similarity or distance will be calculated. Common sets are: physicochemical properties (see Sect. 2.2.1) and molecular fingerprints (see Sect. 2.2.2). Compute the similarity matrix accordingly.
2. Apply PCA to the matrix. Select two or three principal components for plotting. It is useful to consider the percentage of variance captured with each principal component.

This method may become impractical for large datasets (>1000 compounds). See Sect. 3.3 for a chemical space visualization method that is less computationally expensive.

3.2 Comparing Multiple Libraries in the Chemical Space

DFPs are a recently introduced approach to simplify the representation of all compounds in a dataset using a single bit-vector for each database, thereby summarizing every individual fingerprints it contains. DFPs retain the predominant information captured in the molecular fingerprints of the molecules in a given chemical dataset. Briefly, if a given bit had a “1” value in at least 50 % of the compounds in the dataset, it is set to “1” in the DFP, or as “0” otherwise. Further details of the DFPs standardization are described elsewhere [49]. This adds only one step prior to chemical space visualization as commented in Sect. 3.1. If it is intended to include SAR in these plots, libraries could be filtered to include only active compounds. Figure 4 shows schematically the concept of DFPs.

3.3 ChemMaps

Several chemical space visualizations are based upon pairwise similarity measurements. Remarkably, computation of similarity matrices has exponential complexity. Thus, sometimes calculation times make impractical to chart the chemical space of more than 1000 compounds. ChemMaps aim at simplifying the computational task, by adaptively selecting some molecules in the database as comparison references or “satellites.” This method reduces up to 30 % of the time needed for generating a visualization of the chemical space, depending on the size and diversity of the database [50]. The method is as follows:

1. Select at random 25 % of the compounds in a library to use as satellites.
2. Compute the pairwise similarity matrix of all the compounds against the satellites.
3. Perform PCA on the matrix and select the first two or three principal components.
4. Using the principal components as descriptors, compute the distance matrix for all the charted compounds or a subset.

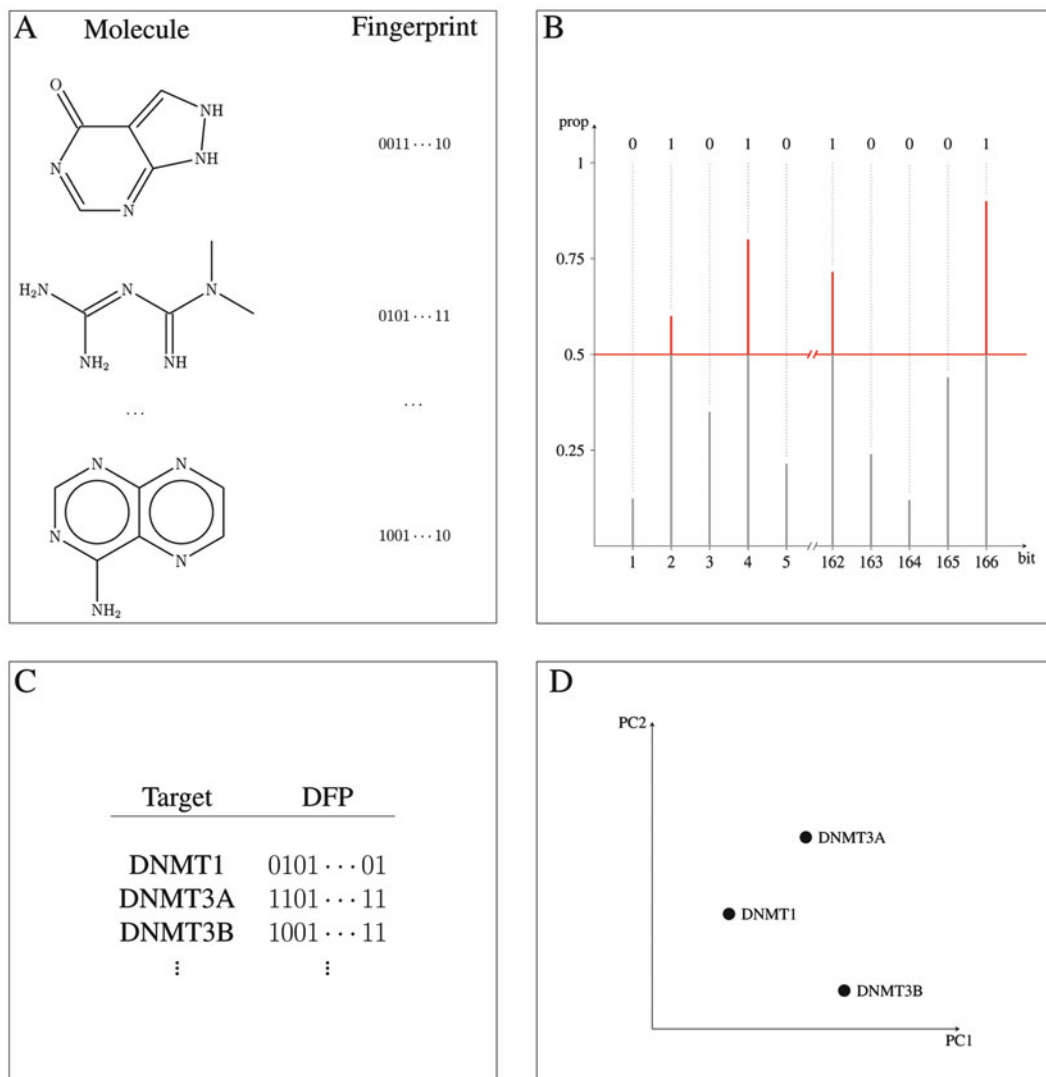


Fig. 4 Database fingerprint (DFP). (a) For every compound in a chemical database, different kind of fingerprints might be obtained. (b) Usually, fingerprints store data in bits. If 50 % or more of the compounds in the database have a value of “1” for a given bit, then it is set as “1” in the DFP, otherwise it is set as “0.” (c) This procedure could be applied to many target-focused libraries. (d) DFPs of multiple libraries can be visualized to represent the chemical space of such libraries. DFPs can also be used for other applications, such as virtual screening. *DFP* database fingerprint

5. Add another 5 % of the database compounds to be used as satellites and repeat steps 2–4.
6. Calculate the correlation between the distances obtained with the PCA as descriptors and repeat step 5 until a correlation of 0.9 or higher is achieved.
7. Plot the chemical space. See Fig. 5.

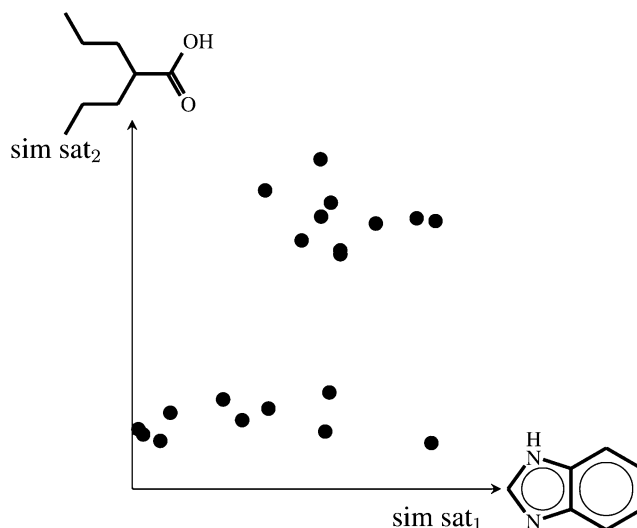


Fig. 5 ChemMaps concept. Chemical space is charted relative to adaptive chemical satellites. Two satellites are used in the example

3.4 Activity Landscape Sweeping

It is common that some structural clusters tend to form when analyzing the chemical space of libraries. Moreover, these clusters may also have different SAR morphologies, with a smoother or rougher application of the similarity principle [11, 51]. The SAR studies and their use for selecting clusters of molecules from a given library are named “activity landscape sweeping.” Such approach is useful to characterize discrete regions in the chemical space where predictive methods that heavily rely upon the similarity principle could be applied. The method is quite straightforward:

1. As a baseline, compute the general SAS map for the whole library as described in Sect. 2.3.1.
2. Plot the chemical space as described in either Sect. 3.1 or 3.3.
3. For defining clusters in the chemical space, apply some method for unsupervised clustering, such as k -means. K -means method could use many principal components for defining the clusters. For selecting a number of principal components to use, a rule of thumb is to plot the contribution of variances of the principal components and select the “elbow” of the curve (i.e., the inflexion point whereupon adding more principal components do not significantly add information). Given that k -means also requires to a priori define the number of clusters, a similar procedure as that for selecting the number of principal components could be applied. However, instead of plotting the variances contribution, the within groups sum of squares is used. However, the number of clusters can also be defined visually by manually adjusting it.

4. Once that clusters of compounds are defined, individual SAS maps per cluster are plotted as described in Sect. 2.3.1.
5. The SAS maps and the proportions of activity cliffs are compared, in order to identify regions with smoother SAR.

4 Target Fishing

In polypharmacology, the identification of all the likely targets for a given chemical compound is of utmost importance and has been an active area of research in recent years [52]. This problem is known as reverse virtual screening or “target fishing” [53]. There is a plethora of computational approaches applied in this field. Cheminformatics methods are mostly based on the principle of SAR [54] which suggests that similar compounds are likely to overlap between the sets of targets that they show activity against [55].

This identification of targets for a given compound can be carried out based on the similarity it presents with other compounds that are known to be active or inactive against some targets. If quantitative and comparable activity values are available, it is possible to build quantitative structure–activity relationships (QSAR) models [21, 56] for every target of interest. If the activity values are not completely reliable, a better alternative is the use of the categorical form of them to build machine-learning models for clustering and classification [57]. Although the general objective of most of these methodologies is the identification of targets for a given compound, the amount and type of biological information available can lead to various applications. This section describes the methodologies implicated in them.

4.1 Target Identification

The most general application of target fishing strategies consists of predicting all the possible targets for a given compound, or at least all of them for which bioactivity data is known. Most of these strategies treat the target fishing problem as a multi-label classification problem, in which every target is a label that a given compound belongs to and for which a predictive model is constructed [52, 58]. The main differences between different approaches are the molecular representation employed and the predictive models used. This work is not intended to provide a detailed description on the construction of these models, which can be found in several other works [59, 60], but of the general strategy for their application.

4.1.1 Multi-label Classifiers

One of the most used alternatives to face the target fishing problem is by building a multi-label classifier. The general steps to build such model are described below:

1. Given a set of targets of interest, a set of compounds, and a defined bipartite activity relation between them, construct and

curate compound databases for each target according to the methods discussed in Sect. 2.1.

2. Build and validate a binary classifier for each database, which allows to distinguish between active and inactive compounds. At this point lies the main difference between distinct models, because the pertinence of a compound to one class or another can be defined according to a priori defined thresholds for a given score. For instance, a similarity coefficient when dealing with similarity searches (discussed in Sect. 2.2), an activity value in the case of QSAR models, or the probability coming from a machine-learning model.
3. Finally, evaluate a compound of interest with all binary models. The targets associated to that compound will be those for which the binary classifiers assign a score higher than the established threshold.

The general scheme of a multi-label classifier is presented in Fig. 6a. The application of these types of strategies in drug design projects is discussed in other works [21, 61, 62] and currently there are several web implementations of these methods [58, 63].

4.1.2 Cluster Analysis

Another methodology to address the multi-label classification problem of target fishing is clustering, which is the task of grouping objects (compounds) such a way that objects belonging to the same group are more similar to each other in comparison to those belonging to other groups. This kind of methodologies only take into account the structure and properties of compounds known to

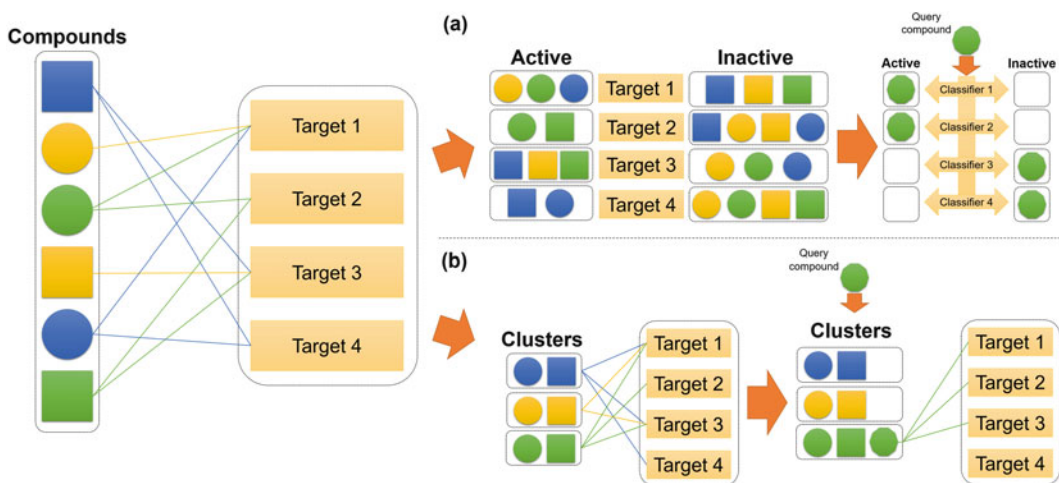


Fig. 6 (a) Representation of a multi-label classifier. The targets associated to the query compound are those for which the corresponding classifiers identify them in the active class. (b) Representation of a clustering analysis. The targets associated to the query compound are those associated to the cluster in which such compound is grouped

be active against each target of interest. The general strategy is as follows:

1. Given a set of targets of interest, a set of compounds, and a defined bipartite activity relation between them, construct and curate a database considering only the compounds known to be active against at least one target.
2. Split the compound database into multiple groups by using a clustering algorithm. This grouping task can be performed according to different criteria, for example, by scaffolds or by molecular similarity, discussed in Sect. 2.2, or employing a machine-learning algorithm like *k*-means, discussed in Sect. 3.4.
3. For each cluster, identify all the targets against which at least one compound in the cluster is active, those will be the targets associated to that cluster.
4. Finally, assign a compound of interest to one cluster by using the same criteria involved in step 2, the targets associated to that cluster will be the predicted targets for the query compound.

Figure 6b presents the general scheme of a cluster analysis. Recent applications of this type of approaches in different research areas and web implementations are discussed in other publications [64, 65].

4.2 Target Deconvolution

Although the knowledge of compounds with activity against one or several targets is fundamental for the development of the strategies presented in Sect. 4.1, these are not the only bioactivity data available. In addition to data from target-based methodologies, the amount of data from cell-based phenotypic screenings has increased considerably in recent years [66]. One of the major advantages of this kind of information is that it provides a more direct view of the responses taking place in the context of a complex biological system, such as a cell [67].

Identifying the molecular targets of active hits from phenotypic screens is a required process to understand the mechanisms of action involved and thus direct the optimization of such compounds. This task is referred as target deconvolution and the cheminformatic approaches to address the problem are essentially the same as those presented in the previous section, with the major difference being that the set of targets to analyze is reduced to those relevant for the phenotype under study [64, 68].

5 Future Prospects

The increasing awareness of polypharmacology in drug discovery and developments will continue demanding the application of cheminformatics approaches to accelerate the process. Computational

methods initially developed for drug discovery focused on a single target are being adapted to develop compounds for multiple targets. Typical examples are SMART and inverse virtual screening or target fishing. In this regard, it is expected that such approaches are further refined to improve accuracy. It is also expected that new computational approaches will emerge to boost the development of polypharmacological drugs.

Acknowledgements

This work was supported by the Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) grant IA203718 and National Council of Science and Technology (CONACyT), Mexico grant number 282785. JJN, FIS-G, and NS-C are thankful to CONACyT for the granted scholarships number 622969, 629458, and 335997, respectively.

References

1. Rosini M (2014) Polypharmacology: the rise of multitarget drugs over combination therapies. *Future Med Chem* 6:485–487. <https://doi.org/10.4155/fmc.14.25>
2. Méndez-Lucio O, Naveja JJ, Vite-Caritano H et al (2016) Review. One drug for multiple targets: a computational perspective. *J Mex Chem Soc* 60:168–181
3. Saldívar-González FI, Naveja JJ, Palomino-Hernández O, Medina-Franco JL (2017) Getting SMART in drug discovery: chemoinformatics approaches for mining structure–multiple activity relationships. *RSC Adv* 7:632–641. <https://doi.org/10.1039/C6RA26230A>
4. González-Medina M, Naveja JJ, Sánchez-Cruz N, Medina-Franco JL (2017) Open chemoinformatic resources to explore the structure, properties and chemical space of molecules. *RSC Adv* 7:54153–54163. <https://doi.org/10.1039/C7RA11831G>
5. Berthold MR, Cebron N, Dill F et al (2009) KNIME – the Konstanz information miner. *SIGKDD Explor Newsl* 11:26. <https://doi.org/10.1145/1656274.1656280>
6. Varnek A (2017) Tutorials in chemoinformatics. <https://doi.org/10.1002/9781119161110>
7. Saldívar-González FI, Hernández-Luis F, Lira-Rocha A, Medina-Franco JL (2017) Manual de Quimioinformática, 1st edn. Universidad Nacional Autónoma de México, Mexico City
8. González-Medina M, Medina-Franco JL (2017) Platform for unified molecular analysis: PUMA. *J Chem Inf Model* 57:1735–1740. <https://doi.org/10.1021/acs.jcim.7b00253>
9. González-Medina M, Méndez-Lucio O, Medina-Franco JL (2017) Activity landscape plotter: a web-based application for the analysis of structure-activity relationships. *J Chem Inf Model* 57:397–402. <https://doi.org/10.1021/acs.jcim.6b00776>
10. González-Medina M, Prieto-Martínez FD, Owen JR, Medina-Franco JL (2016) Consensus diversity plots: a global diversity analysis of chemical libraries. *J Cheminform* 8:63. <https://doi.org/10.1186/s13321-016-0176-9>
11. Naveja JJ, Oviedo-Osornio CI, Trujillo-Minero NN, Medina-Franco JL (2017) Chemoinformatics: a perspective from an academic setting in Latin America. *Mol Divers*. <https://doi.org/10.1007/s11030-017-9802-3>
12. Richter L, Ecker GF (2015) Medicinal chemistry in the era of big data. *Drug Discov Today Technol* 14:37–41. <https://doi.org/10.1016/j.ddtec.2015.06.001>
13. Law V, Knox C, Djoumbou Y et al (2014) Drug-Bank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* 42:D1091–D1097. <https://doi.org/10.1093/nar/gkt1068>
14. Bento AP, Gaulton A, Hersey A et al (2014) The ChEMBL bioactivity database: an update. *Nucleic Acids Res* 42:D1083–D1090. <https://doi.org/10.1093/nar/gkt1031>
15. Irwin JJ, Shoichet BK (2005) ZINC – a free database of commercially available

- compounds for virtual screening. *J Chem Inf Model* 45:177–182. <https://doi.org/10.1021/ci049714+>
- Liu T, Lin Y, Wen X et al (2007) BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res* 35:D198–D201. <https://doi.org/10.1093/nar/gkl999>
 - Lavecchia A, Cerchia C (2016) In silico methods to address polypharmacology: current status, applications and future perspectives. *Drug Discov Today* 21:288–298. <https://doi.org/10.1016/j.drudis.2015.12.007>
 - Fourches D, Muratov E, Tropsha A (2016) Trust, but verify II: a practical guide to chemogenomics data curation. *J Chem Inf Model* 56:1243–1252. <https://doi.org/10.1021/acs.jcim.6b00129>
 - Hersey A, Chambers J, Bellis L et al (2015) Chemical databases: curation or integration by user-defined equivalence? *Drug Discov Today Technol* 14:17–24. <https://doi.org/10.1016/j.ddtec.2015.01.005>
 - Miller MA (2002) Chemical database techniques in drug discovery. *Nat Rev Drug Discov* 1:220–227. <https://doi.org/10.1038/nrd745>
 - Cherkasov A, Muratov EN, Fourches D et al (2014) QSAR modeling: where have you been? Where are you going to? *J Med Chem* 57:4977–5010. <https://doi.org/10.1021/jm4004285>
 - Mansouri K, Abdelaziz A, Rybacka A et al (2016) CERAPP: collaborative estrogen receptor activity prediction project. *Environ Health Perspect* 124:1023–1033. <https://doi.org/10.1289/ehp.1510267>
 - Gally J-M, Bourg S, Do Q-T et al (2017) Vsprep: a general KNIME workflow for the preparation of molecules for virtual screening. *Mol Inform*. <https://doi.org/10.1002/minf.201700023>
 - Naveja JJ, Medina-Franco JL (2017) Insights from pharmacological similarity of epigenetic targets in epipolypharmacology. *Drug Discov Today*. <https://doi.org/10.1016/j.drudis.2017.10.006>
 - Medina-Franco JL, Martinez-Mayorga K, Meurice N (2014) Balancing novelty with confined chemical space in modern drug discovery. *Expert Opin Drug Discov* 9:151–165. <https://doi.org/10.1517/17460441.2014.872624>
 - Sheridan RP, Kearsley SK (2002) Why do we need so many chemical similarity search methods? *Drug Discov Today* 7:903–911. [https://doi.org/10.1016/S1359-6446\(02\)02411-X](https://doi.org/10.1016/S1359-6446(02)02411-X)
 - Medina-Franco JL, Maggiora GM (2013) Molecular similarity analysis. In: Bajorath J (ed) *Cheminformatics for drug discovery*. Wiley, Hoboken, NJ, pp 343–399. <https://doi.org/10.1002/9781118742785.ch15>
 - Singh N, Guha R, Giulianotti MA et al (2009) Chemoinformatic analysis of combinatorial libraries, drugs, natural products, and molecular libraries small molecule repository. *J Chem Inf Model* 49:1010–1024. <https://doi.org/10.1021/ci800426u>
 - Xu J, Hagler A (2002) Chemoinformatics and drug discovery. *Molecules* 7:566–600. <https://doi.org/10.3390/70800566>
 - Gortari EF, Medina-Franco JL (2015) Epigenetic relevant chemical space: a chemoinformatic characterization of inhibitors of DNA methyltransferases. *RSC Adv* 5:87465–87476. <https://doi.org/10.1039/C5RA19611F>
 - Eckert H, Bajorath J (2007) Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches. *Drug Discov Today* 12:225–233. <https://doi.org/10.1016/j.drudis.2007.01.011>
 - Rogers D, Hahn M (2010) Extended-connectivity fingerprints. *J Chem Inf Model* 50:742–754. <https://doi.org/10.1021/ci100050t>
 - Ewing T, Baber JC, Feher M (2006) Novel 2D fingerprints for ligand-based virtual screening. *J Chem Inf Model* 46:2423–2431. <https://doi.org/10.1021/ci060155b>
 - Durant JL, Leland BA, Henry DR, Nourse JG (2002) Reoptimization of MDL keys for use in drug discovery. *J Chem Inf Comput Sci* 42:1273–1280. <https://doi.org/10.1021/ci010132r>
 - Jaccard P (1901) Etude Comparative de la Distribution Florale dans une Portion des Alpes et des Jura. *Bull Soc Vaudoise Sci Nat* 37:547–579
 - Bemis GW, Murcko MA (1996) The properties of known drugs. 1. Molecular frameworks. *J Med Chem* 39:2887–2893. <https://doi.org/10.1021/jm9602928>
 - Xu Y-J, Johnson M (2002) Using molecular equivalence numbers to visually explore structural features that distinguish chemical libraries. *J Chem Inf Comput Sci* 42:912–926
 - Medina-Franco J, MartÁ-nez-Mayorga K, Bender A, Scior T (2009) Scaffold diversity analysis of compound data sets using an entropy-based measure. *QSAR Comb Sci* 28:1551–1560. <https://doi.org/10.1002/qsar.200960069>
 - Shanmugasundaram V, Maggiora GM (2001) Characterizing property and activity landscapes using an information-theoretic approach. CINF-032. In 222nd ACS National Meeting,

- Chicago, IL, USA; August 26–30, 2001; American Chemical Society, Washington, DC
40. Guha R (2012) Exploring structure-activity data using the landscape paradigm. Wiley Interdiscip Rev Comput Mol Sci. <https://doi.org/10.1002/wcms.1087>
 41. Bajorath J, Peltason L, Wawer M et al (2009) Navigating structure-activity landscapes. *Drug Discov Today* 14:698–705. <https://doi.org/10.1016/j.drudis.2009.04.003>
 42. Medina-Franco JL (2012) Scanning structure-activity relationships with structure-activity similarity and related maps: from consensus activity cliffs to selectivity switches. *J Chem Inf Model* 52:2485–2493. <https://doi.org/10.1021/ci300362x>
 43. Medina-Franco JL, Petit J, Maggiora GM (2006) Hierarchical strategy for identifying active chemotype classes in compound databases. *Chem Biol Drug Des* 67:395–408. <https://doi.org/10.1111/j.1747-0285.2006.00397.x>
 44. Maggiora G, Gokhale V (2017) A simple mathematical approach to the analysis of polypharmacology and polyspecificity data. [version 1; referees: 3 approved, 1 approved with reservations]. F1000Res. <https://doi.org/10.12688/f1000research.11517.1>
 45. Pérez-Villanueva J, Santos R, Hernández-Campos A et al (2011) Structure–activity relationships of benzimidazole derivatives as anti-parasitic agents: dual activity-difference (DAD) maps. *Med Chem Commun* 2:44–49. <https://doi.org/10.1039/C0MD00159G>
 46. Yongye AB, Medina-Franco JL (2012) Data mining of protein-binding profiling data identifies structural modifications that distinguish selective and promiscuous compounds. *J Chem Inf Model* 52:2454–2461. <https://doi.org/10.1021/ci3002606>
 47. Osolodkin DI, Radchenko EV, Orlov AA et al (2015) Progress in visual representations of chemical space. *Expert Opin Drug Discov* 10:959–973. <https://doi.org/10.1517/17460441.2015.1060216>
 48. Medina-Franco J, Martinez-Mayorga K, Giulianotti M et al (2008) Visualization of the chemical space in drug discovery. *Curr Comput Aided Drug Des* 4:322–333. <https://doi.org/10.2174/157340908786786010>
 49. Fernández-de Gortari E, García-Jacas CR, Martínez-Mayorga K, Medina-Franco JL (2017) Database fingerprint (DFP): an approach to represent molecular databases. *J Cheminform* 9:9. <https://doi.org/10.1186/s13321-017-0195-1>
 50. Naveja JJ, Medina-Franco JL (2017) Chem-Maps: towards an approach for visualizing the chemical space based on adaptive satellite compounds [version 1; referees: 1 approved, 2 approved with reservations]. F1000Res. <https://doi.org/10.12688/f1000research.12095.1>
 51. Naveja JJ, Medina-Franco JL (2015) Activity landscape sweeping: insights into the mechanism of inhibition and optimization of DNMT1 inhibitors. *RSC Adv* 5:63882–63895. <https://doi.org/10.1039/C5RA12339A>
 52. Wale N, Karypis G (2009) Target fishing for chemical compounds using target-ligand activity data and ranking based methods. *J Chem Inf Model* 49:2190–2201. <https://doi.org/10.1021/ci9000376>
 53. Jenkins JL, Bender A, Davies JW (2006) In silico target fishing: predicting biological targets from chemical structure. *Drug Discov Today Technol* 3:413–421. <https://doi.org/10.1016/j.ddtec.2006.12.008>
 54. Hansch C, Maloney PP, Fujita T, Muir RM (1962) Correlation of biological activity of phenoxyacetic acids with hammett substituent constants and partition coefficients. *Nature* 194:178–180. <https://doi.org/10.1038/194178b0>
 55. Nettles JH, Jenkins JL, Bender A et al (2006) Bridging chemical and biological space: “target fishing” using 2D and 3D molecular descriptors. *J Med Chem* 49:6802–6810. <https://doi.org/10.1021/jm060902w>
 56. Cramer RD (2012) The inevitable QSAR renaissance. *J Comput Aided Mol Des* 26:35–38. <https://doi.org/10.1007/s10822-011-9495-0>
 57. Lavecchia A (2015) Machine-learning approaches in drug discovery: methods and applications. *Drug Discov Today* 20:318–331. <https://doi.org/10.1016/j.drudis.2014.10.012>
 58. Yao Z-J, Dong J, Che Y-J et al (2016) Target-Net: a web service for predicting potential drug-target interaction profiling via multi-target SAR models. *J Comput Aided Mol Des* 30:413–424. <https://doi.org/10.1007/s10822-016-9915-2>
 59. Nidhi, Glick M, Davies JW, Jenkins JL (2006) Prediction of biological targets for compounds using multiple-category Bayesian models trained on chemogenomics databases. *J Chem Inf Model* 46:1124–1133. <https://doi.org/10.1021/ci060003g>

60. Kawai K, Fujishima S, Takahashi Y (2008) Predictive activity profiling of drugs by topological-fragment-spectra-based support vector machines. *J Chem Inf Model* 48:1152–1160. <https://doi.org/10.1021/ci7004753>
61. Nikolic K, Mavridis L, Djikic T et al (2016) Drug design for CNS diseases: polypharmacological profiling of compounds using cheminformatic, 3D-QSAR and virtual screening methodologies. *Front Neurosci* 10:265. <https://doi.org/10.3389/fnins.2016.00265>
62. Rognan D (2010) Structure-based approaches to target fishing and ligand profiling. *Mol Inform* 29:176–187. <https://doi.org/10.1002/minf.200900081>
63. Awale M, Reymond J-L (2017) The polypharmacology browser: a web-based multi-fingerprint target prediction tool using ChEMBL bioactivity data. *J Cheminform* 9:11. <https://doi.org/10.1186/s13321-017-0199-x>
64. Kunimoto R, Dimova D, Bajorath J (2017) Application of a new scaffold concept for computational target deconvolution of chemical cancer cell line screens. *ACS Omega* 2:1463–1468. <https://doi.org/10.1021/acsomega.7b00215>
65. Reker D, Rodrigues T, Schneider P, Schneider G (2014) Identifying the macromolecular targets of de novo-designed chemical entities through self-organizing map consensus. *Proc Natl Acad Sci U S A* 111:4067–4072. <https://doi.org/10.1073/pnas.1320001111>
66. Zheng W, Thorne N, McKew JC (2013) Phenotypic screens as a renewed approach for drug discovery. *Drug Discov Today* 18:1067–1073. <https://doi.org/10.1016/j.drudis.2013.07.001>
67. Lee J, Bogyo M (2013) Target deconvolution techniques in modern phenotypic profiling. *Curr Opin Chem Biol* 17:118–126. <https://doi.org/10.1016/j.cbpa.2012.12.022>
68. Mugumbate G, Mendes V, Blaszczyk M et al (2017) Target identification of mycobacterium tuberculosis phenotypic hits using a concerted chemogenomic, biophysical, and structural approach. *Front Pharmacol* 8:681. <https://doi.org/10.3389/fphar.2017.00681>



Computational Predictions for Multi-Target Drug Design

Neelima Gupta, Prateek Pandya, and Seema Verma

Abstract

Computational techniques have proven to be an essential tool in modern drug discovery research. These tools offer powerful methods for prediction of ligand–receptor interaction events at atomic details, without attempting exhaustive experimental setup. Single ligand–single target strategies for the discovery of new drug molecules have become outdated due to the factors like drug resistance, increased side effects, reduced efficacy, etc., in addition to the involvement of long time period for validation of a new molecule by toxicology and pharmacokinetic studies. Multi-target drug designing approach can offer a paradigm shift for alternative usage of known drugs for complex diseases. These approaches combine knowledge of complex disease networks, chemical and physical characteristics of drugs, and biological receptors. With the availability of advanced computational resources, a number of tools have been developed that help in the identification of new and multiple targets for the already known or new drugs. In the present chapter, an attempt has been made to highlight the current state-of-the-art methodologies used in multi-target identification for therapeutic effects of known drugs or new drug candidates.

Keywords Binding interactions, Machine learning, Molecular docking, Molecular dynamics, Multi-target drug design (MTDD), Polypharmacology, QM–MM approach, QSAR, Systems approach

1 Introduction

Drug discovery for complex diseases continues to present several challenges to researchers as these diseases are known to possess complex network of pathways involving diverse proteins as functional units. Examples of such diseases are cancer, certain neurological disorders, inflammatory diseases, and metabolic and cardiovascular diseases; even addiction-related problems encompass networks of multiple pathways [1]. These networks are often so robust that any effort to tamper them through modulating a single biomolecular target results in the activation of alternative pathways for effective management of network functioning [2, 3]. Even with the most state-of-the-art technology, the drug discovery process is still an extremely expensive and time consuming prospect [4–7]. It has been recognized that an effective strategy to counter such robustness of biological networks is to attack them simultaneously at multiple levels [8]. Such multi-level targeted attacks require drug molecules that can bind with multiple biomolecules—mostly proteins, but also nucleic acids in some cases. The search for such

multi-target agents could be achieved by identifying promiscuous compounds that have the tendency to bind with several proteins. During recent times, one effective way to fill the gap in the development of new therapeutic agents has been the repurposing of the existing approved drugs for different diseases. The strategy has been extensively supported by computational modeling approaches for faster prediction of efficacy of known drugs with new targets [9–15] by elimination of steps involving toxicological and pharmacokinetic assessments. For example, multi-target interaction studies have proven to be useful in the prediction of potential ligands that target multiple proteins in malaria [11]. Another option to recognize a multi-target drug may be the identification of alternate targets for a known drug with some degree of side effects, consequently leading to the optimization of particular drug for a complex disease. Thereby, drug repurposing or polypharmacology is another way to find out alternative targets of existing drugs [16, 17].

Multi-target drugs have become popular during recent years owing to the following advantages:

1. Improved efficacy
2. Fewer side effects
3. Better bioavailability
4. Reduced risk of drug resistance
5. Less burden on discovery

Efficacy of a drug depends on its interaction with biological receptors. Binding of drug-like small molecules with biological receptors involves several bond types, most notably among them are weak interactions originating in weak hydrogen bonds, van-der-Waals interactions, hydrophobic associations, etc. As a result, the molecule naturally possesses affinity towards more than one biomolecules with varying degree of strength. Such multi-target binding raises an obvious point that in case the drug molecule has affinity to bind with several other biomolecules, shouldn't they be optimized appropriately for modulation with desired targets in biological systems?

Two main approaches in computational multi-target drug discovery being practiced are:

1. Ligand-based virtual screening
2. Structure-based virtual screening

Different methodologies under the above two types of approaches have been presented in Fig. 1.

Machine Learning (ML) methods depend on understanding the differences in known active and inactive leads. These methods comprise several approaches such as linear discriminant analysis

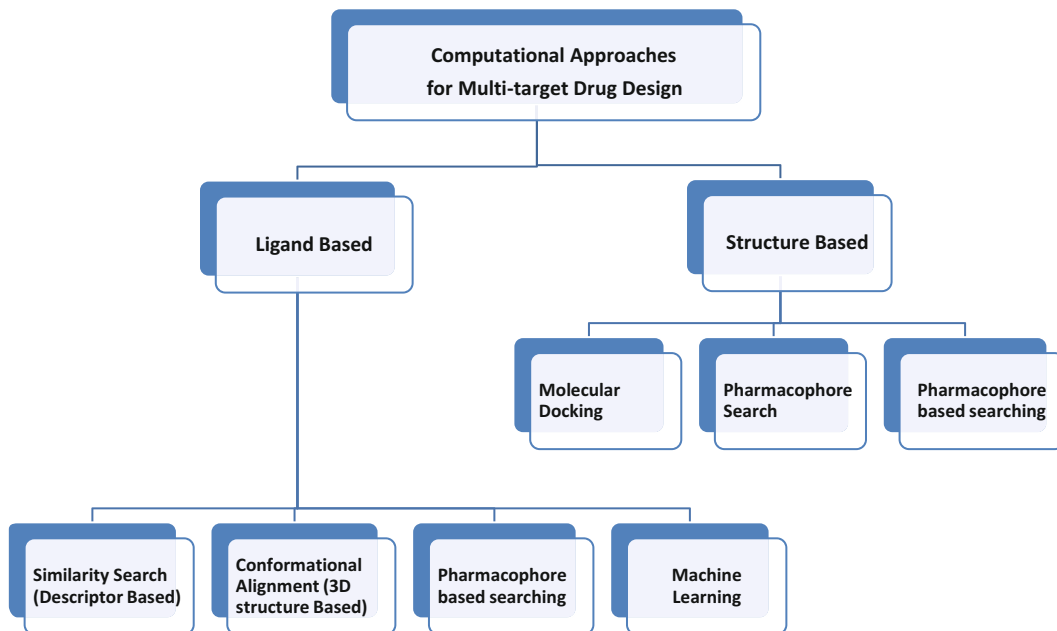


Fig. 1 An overview of computational approaches for multi-target drug discovery

(LDA), k -nearest neighbors, binary kernel discrimination, Bayesian methods, decision trees, random forests, artificial neural networks (ANN), and support vector machines. The reliability of these methods depends on the quality of datasets.

Considerable progress has taken place in the development of structure-based drug discovery approaches due to advancements in computational techniques and resource utilization. Depending on the ligand- and structure-based approaches, various available computational tools are utilized for multi-target drug design. Here, in the following sections, the details of some important computational methods and methodologies that are helpful in speeding up the multi-target drug discovery process are explained.

2 Multi-Target Drug Design with Systems Approach

Human biology is full of complex networks consisting of proteins, nucleic acids, etc. Each node of a network signifies a biological macromolecule such as protein, DNA, or RNA while each link represents their mutual interactions within pathways. Most multi-target drugs are weak binders as they interact with multiple receptors with weaker interactions. It is known that limited inhibition of a small number of targets can provide useful therapeutic effects and it has been suggested that often the inclusive effect of partial drug action at multiple target sites is better than the exclusive drug action at a single target. Several multi-target drugs such as

sunitinib, lapatinib, and dasatinib have been discovered and used effectively for cancer treatment [18, 19]. These drugs are known to effectively block the alternative escape mechanism of disease networks [20–22].

Systems approach makes use of the available knowledge of disease processes and drug actions to integrate biological system network and data mining to identify the drug for treatment of diseases by modulation of more than one target, either one drug binding to multiple targets or multiple drugs binding to different targets within a network. Network theory-based methods such as network-based inference (NBI), target-based similarity inference (TBSI), and drug-based similarity inference (DBSI) have been effectively used to identify new targets for existing drugs [23]. Recently, a few ligand-based target similarity methods have been developed that predict similar protein targets for known drug molecules. For instance, investigations with Similarity Ensemble Approach [24] reveal that functional genomics and ligand similarity are complimentary, and can be assessed by employing network-based similarity searching approaches. Some freely available tools and servers employing Systems Approach are listed in Table 1.

Table 1
List of some freely available tools and servers for systems approach

Program	Details	Link
PharmMapper Server	Identifies potential target candidates for small molecules using pharmacophore mapping approach	Honglinli's group, School of Pharmacy, East China University of Science and Technology http://lilab.ecust.edu.cn/pharmmapper/index.php
Stitch Server (search tool for interactions of chemicals)	Provides information about interactions from metabolic pathways, crystal structures, binding experiments, and drug–target relationships	EMBL http://stitch.embl.de/
Pocketome	Ensembles of conformations of druggable binding sites that can be identified experimentally from co-crystal structures in the Protein Data Bank	University of California at San Diego http://www.pocketome.org/
SystemsDock	Network pharmacology-based prediction and analysis, uses high-precision docking simulation and molecular pathway map	http://systemsdock.unit.oist.jp/iddp/home/index
BINDNET	Prediction of binding network between proteomes and small molecules	Computational Biology Research Group (CompBio), State University of New York http://protinfo.compbio.buffalo.edu/cando/servers/bindnet/

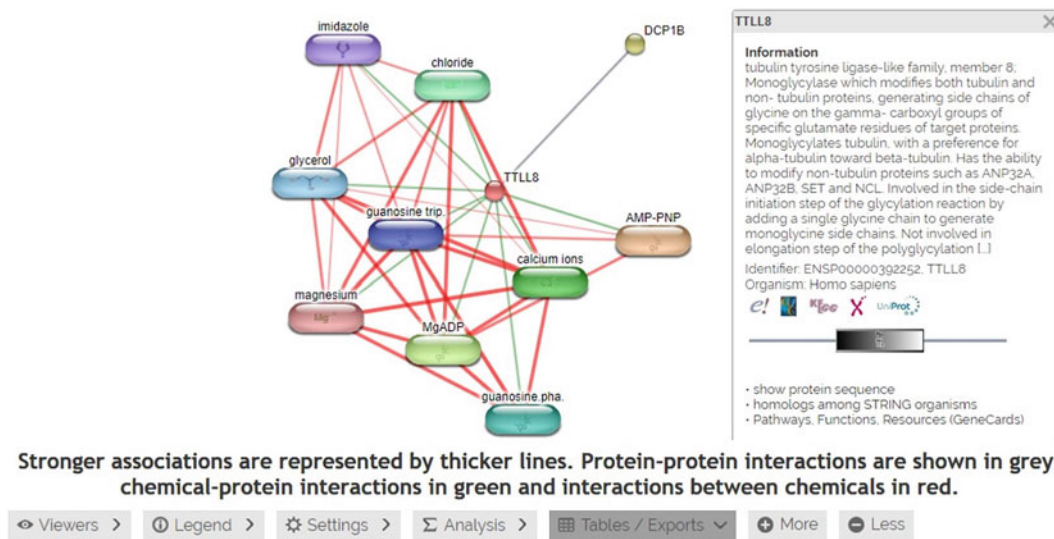


Fig. 2 Interaction network of tubulin tyrosine ligase-like family protein using Stich

Figure 2 presents network linkages of tyrosine ligase-like family protein (TTL8) with other macromolecules in a metabolic pathway as obtained from Stich program. Such information is valuable in finding multiple targets of known drugs as it provides multiple options for drug targeting in a concerted framework.

3 Molecular Docking

Ligand–target docking is an essential tool in modern drug discovery and development efforts. Docking is generally employed to identify binding affinity and appropriate conformations of drug molecules in the bound form within a receptor binding site [25–33]. Molecular docking essentially relies on the three-dimensional structural information of both the receptor and ligand. There are two main steps in the docking protocol that determine its output in terms of binding free energy estimates and ligand conformations, viz., conformational search and scoring of ligand based on free energy estimates. In some cases, docking can also be utilized to find out the binding mode of a drug when the binding site information about a target is unknown [28, 30, 31]. Several studies have shown that docking helps in the determination of the most effective locations for suitable binding of the target. Some of the challenges that are faced by docking calculations include receptor flexibility, solvation effects, conformational search, scoring criteria, covalent binding, etc. Recent attempts have been made to address the problem of target and ligand flexibility using novel algorithm design.

Most of the available docking programs consider drug flexibility, while ignoring target flexibility to provide a rigid-body system as the target site for drug molecules [27, 28, 31, 34–37]. Although, these methods have produced reasonable estimates in most cases, they often fail to account for precise simultaneous conformational flexibility of ligand and binding site atoms as both of them may change during the interaction process. Biological systems are essentially dynamic systems, where solution structures of molecules remain flexible for environmental and other molecular effects. A normal DNA molecule remains structurally intact but conformationally flexible in solution. Similarly, proteins also possess flexible regions to accommodate ligands for metabolic and other pathways. Therefore, it becomes essential to consider the flexibility of target sites in computational calculations. Molecular dynamics enhanced docking can be useful in generating structural ensemble to map flexibility of protein sites [38, 39].

3.1 Methodology

A conventional protocol for rigid-body docking includes, preparation of ligand, receptor, docking box, setting up calculation parameters, search algorithm, scoring function, number of iterations, etc. Several different types of strategies are applied for calculation pertaining to the binding of small molecules with biological receptors in various docking tools. Some of the important steps followed in a docking protocol [37, 40] are as follows:

1. Generation of 3D structures of ligand and receptor (Fig. 3).
2. Preparation of ligand and receptor: addition of polar hydrogens, atom identification, charge assignment, etc.

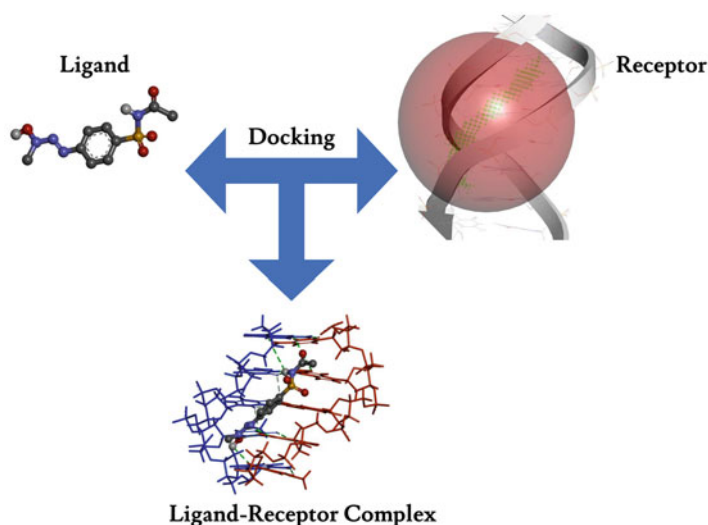


Fig. 3 Overview of obtaining docked complex from interaction of receptor and ligand

3. Assignment of force field.
4. Generation of grid in binding site of the receptor.
5. Searching of stable conformations of ligand in binding site based on various degrees of freedom by systematic or stochastic algorithms (Fig. 3).
6. Scoring and ranking: evaluating binding free energies and subsequently ranking the binding poses.

The steps mentioned above can be suitably modified when multiple receptors are screened *viz-a-viz* single ligand on massive parallel processing machines to reduce large-scale scanning [41]. The platforms, such as VinaMPI (massively parallel Message Passing Interface) [41] and INVDOCK [42], make use of inverse docking to identify multiple targets for selected ligands [43, 44].

Another approach for dealing with multi-target identification by docking is to use it in combination with molecular dynamics simulations to sample ligand and receptor flexibility [38, 39]. Binding affinity prediction is the key aspect in drug target interaction forecast which is ultimately used to correlate the therapeutic utility. Programs such as CANDO [45] provide another unique approach in which multiple targets are screened using automated fragment-based docking and molecular dynamics approach to predict binding energies of drug–target systems with more accuracy. CANDO identifies multiple targets using knowledge-driven fragment-based docking followed by dynamics simulations. CANDO is a comprehensive package of automated tools to screen large number of protein targets and includes protein structure prediction tools such as I-TASSER, MODELLER, etc., for proteins without known crystal structures. Some popular open-source available docking programs are listed in Table 2.

3.1.1 General Methodology of Docking Experiment Using AutoDock-Vina

The programs MGLTools, AutoDock-Vina 1.1.2 (Table 2), and any of the following visualization tools: *viz.*, UCSF-Chimera, Discovery Studio Visualizer, PyMol, etc., are required for docking experiment.

The following steps summarize the working of a docking experiment and obtaining output results using AutoDock-Vina.

Step I—A folder to be created, where the input files, e.g., Target.pdbqt, Ligand.pdbqt, and instruction files, are to be saved and the calculations carried out.

Step II—Creation of Target (Protein/DNA) input file

1. Open MGLTools
2. File | Read Molecule | location of Target.pdb
3. Edit | Hydrogens | Add | Polar only | remove water molecules (if required)

Table 2
List of common freely available docking programs

Software	Description	Link
AutoDock	For prediction of binding of small molecules, or drug candidates, to a receptor of known 3D structure	The Scripps Research Institute http://autodock.scripps.edu/
ParDock	Automated Server for Protein Ligand Docking by an all-atom energy-based Monte Carlo	SCFBio at IIT, Delhi http://www.scfbio-iitd.res.in/dock/pardock.jsp
DNADock	Online DNA-Ligand Docking	SCFBio at IIT, Delhi http://www.scfbio-iitd.res.in/dock/dnadock.jsp
SwissDock	EADock DSS based protein–ligand interaction prediction server	Swiss Bioinformatics Institute http://www.swissdock.ch/
DOCK	Rigid-body docking to find lowest energy binding mode using geometric matching algorithm to superimpose the ligand onto a negative image of the binding pocket	University of California, San Francisco http://dock.compbio.ucsf.edu/
DOT	Docking macromolecules, including proteins, DNA, and RNA	University of California at San Diego http://www.sdsc.edu/CCMS/DOT/
GRAMM (<i>Global Range Molecular Matching</i>)	Open-source program for protein docking using only the atomic coordinates of the two molecules (two proteins/a protein and a smaller compound/two transmembrane helices, etc.)	Vakser Lab, University of Kansas http://vakser.compbio.ku.edu/main/resources_gramm1.03.php
Hex Protein Docking	Open-source interactive protein docking and molecular superposition program	Dave Ritchie http://hex.loria.fr/
ZDOCK	Automated Fast Fourier Transform-based protein docking program	University of Massachusetts Medical School http://zdock.umassmed.edu/
Cluspro	Protein–protein docking	Structural Bioinformatics Lab, Boston University and Stony Brook University
MolFit	Protein–protein docking using Fast Fourier Transform	Weizmann Institute of Science http://www.weizmann.ac.il/Chemical_Research_Support/molfit/downloads.html
MOLS 2.0	Open-source program for peptide modeling and protein–ligand docking	University of Madras https://sourceforge.net
DOCK BLASTER	Online server for structure-based ligand discovery	University of California, San Francisco https://blaster.docking.org/

(continued)

Table 2
(continued)

Software	Description	Link
GEMDOCK	For computing a ligand conformation and orientation relative to the active site of target protein	Institute of Bioinformatics, National Chiao Tung University, Taiwan http://gemdock.life.nctu.edu.tw/dock/
FORECASTER Suite	Suite of programs for drug design and discovery that includes FITTED docking program and the sites of metabolism prediction program IMPACTS	Molecular Forecaster Inc. and McGill University www.fitted.ca/ and http://molecularforecaster.com/
CANDO	Server for fragment-based docking with dynamics, multi-targeting, and drug repurposing	Computational Biology Research Group, State University of New York, Buffalo http://ram.org/compbio/protinfo/cando/

4. Grid | Macromolecule | Choose | Select the Target.pdb (this opens a save dialogue) Save in the same folder in pdbqt format.
5. Setting the grid box size—Go to MGLTools again | Grid | Grid Box | Adjust Spacing if required | Center grid Box (by changing x,y,z center) | reduce or increase the box size in $x, y,$ and z directions around the active site | Note down all box parameters | close Grid Dialog Box

Step III—Creation of input file for Ligand

1. Go to MGLTools again | Input | Open | Location of Ligand.pdb
2. Add polar Hydrogens, if needed
3. Hide | Target | select Ligand
4. Ligand | Torsion Tree | Choose Torsion (this will show all the torsion angles and if required, change torsion angle by clicking on torsion angle) | Done
5. Ligand | output | save as pdbqt (in the same folder having Target.pdbqt file)

Step IV—Creation of Instruction file

Create an instruction file named conf.txt for AutoDock-Vina in notepad (Fig. 4).

Step V—Docking calculation

1. At this point, the calculation folder (created in Step I, e.g., VinaTest) should contain three files: Target.pdbqt; Ligand.pdbqt; and conf.txt

```

conf.txt - Notepad
receptor = Receptor_1.pdbqt
ligand = Ligand_1.pdbqt
exhaustiveness = 32
center_x = 3.078
center_y = -6.842
center_z = 24.489
size_x = 10
size_y = 8
size_z = 10
out = out.pdbqt
cpu = 4
num_modes = 20

log.txt - Notepad
#####
# If you used AutoDock Vina in your work, please cite:
#
# O. Trott, A. J. Olson,
# AutoDock Vina: improving the speed and accuracy of docking
# with a new scoring function, efficient optimization and
# multithreading, Journal of Computational Chemistry 31 (2010)
# 455-461
# DOI 10.1002/jcc.21334
#
# Please see http://vina.scripps.edu for more information.
#####
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: 1871602512
Performing search ... done.
Refining results ... done.

mode | affinity | dist from best mode
      | (kcal/mol) | rmsd l.b. | rmsd u.b.
-----|-----|-----|-----
1 | -5.2 | 0.000 | 0.000
2 | -5.1 | 2.792 | 4.144
3 | -5.1 | 11.093 | 11.995
4 | -5.1 | 2.650 | 5.433
5 | -5.0 | 1.128 | 2.800
6 | -4.9 | 2.941 | 5.548
7 | -4.8 | 3.340 | 5.899
8 | -4.8 | 2.517 | 4.471
9 | -4.6 | 2.517 | 4.471
  
```

Fig. 4 Examples of an instruction file (conf.txt) and output file (log.txt) from an AutoDock-Vina calculation

Table 3
Some common freely available tools useful in drug designing

S. No.	Software	Description
1	Discovery Studio Visualizer	Visualization and analysis
2	UCSF-Chimera	Visualization and analysis
3	PyMol	Visualization and analysis
4	OpenBabel	File format convertor
5	ACD/ChemSketch	Sketching and generating ligand files in various formats

- Open command line and reach the calculation folder, Example—<C:\Users\Science\Documents\VinaTest>
- Run the calculation by invoking the Vina program into the calculation folder. Example, <C:\Users\Science\Documents\VinaTest> “\Program Files(x86)\The Scripps Research Institute\Vina\vina.exe” --conf conf.txt --log log.txt
Note that “--conf” is an instruction to read input file named “conf.txt”. “--log” is the instruction to generate results in a file named “log.txt”. Results of the docking calculation are provided in the log.txt file generated by the program (Fig. 4).

The input and output geometries of ligand, target, and docked complex can be visualized using any of the visualization tools listed in Table 3. The graphical visualization of target HSA and

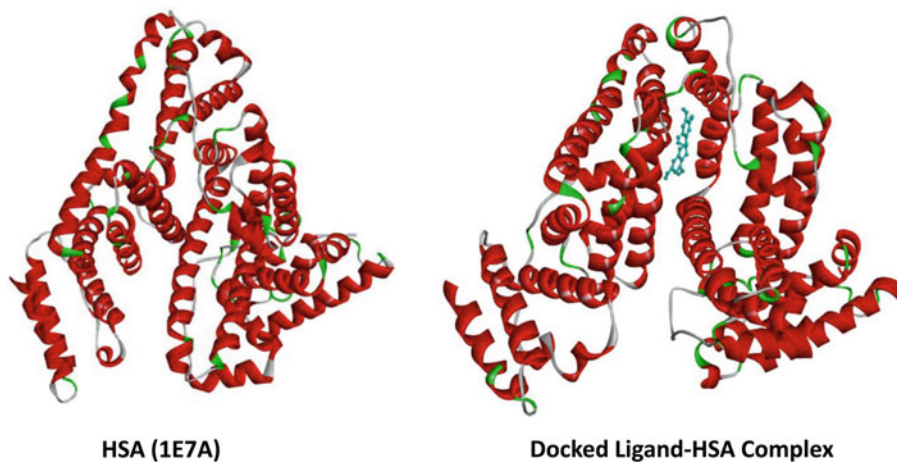


Fig. 5 Visualization of protein (crystal structure extracted from PDB) and docked complex using Discovery Studio program

ligand–HSA complex using Discovery Studio tool has been depicted in Fig. 5.

4 Hyphenated Methods

4.1 QM–MM Approach

Majority of the currently used docking methods describe the ligand–receptor interactions by making use of the classical force field potential of Molecular Mechanics (MM), having the limitation of fixed charge treatment of electrostatic interactions. On the other hand, the Quantum Mechanics (QM) calculations, generally used for studying chemical reaction mechanism, are intensive and time consuming; hence, their use for faster prediction of potential drug candidates has not been explored rigorously. QM method is particularly suitable to study the transition states (TS) and intermediates formed during a reaction. It might be interesting to note that such transition states could be well suited as inhibitors of enzymes. Since QM calculations are computationally costly, it has been found appropriate to study intermolecular interactions between two large molecules such as ligand and receptor using a hybrid QM–MM method involving the reaction site at more accurate QM and rest of the system at MM level. Such QM–MM method has been used for studying chemical reactions of enzyme systems [46, 47]. In conventional non-covalent binding event, it is often desirable to identify the true minimum energy conformation that a small drug-like molecule can attain, in order to form a stable ligand–receptor complex. The true level of low energy conformation can be obtained best by employing QM method to the ligand. Recently, we have employed the hybrid QM–MM approach using a two-layer ONIOM method in an attempt to identify the

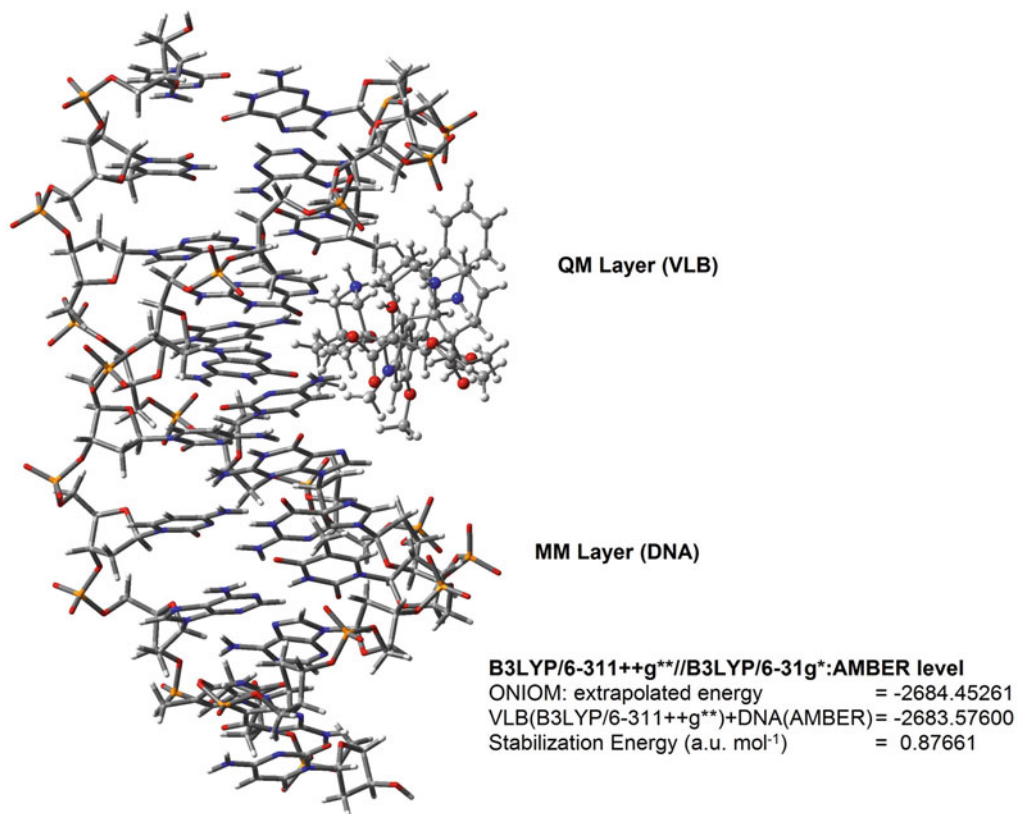


Fig. 6 Visualization of the drug–DNA complex resulting from QM–MM approach using ONIOM calculation

conformational strain of ligand molecule in docked state, treating the latter with QM in bound form [27]. The geometry of drug (DRG)–receptor (R) complex in bound state, obtained from docking experiment, was subjected to QM–MM treatment in an ONIOM calculation [27] (Fig. 6). The ONIOM(B3LYP/6-31G*:**AMBER**) calculation was done on DRG–R complex using a two-layer method considering the receptor R as rigid molecule and subjecting the DRG in the bound state to energy minimization using density functional theory (DFT)-level calculations with B3LYP functional and 6-31G* basis set. Interaction energy E_5 of the DRG–R complex, as a measure of stabilization, was obtained from the following equation sequences:

$$E_1(\text{DRG-R})_{\text{ONIOM}} = E_2(\text{DRG-R})_{\text{MM}} + E_3(\text{DRG})_{\text{QM}} - E_4(\text{DRG})_{\text{MM}}$$

where $E_2(\text{DRG-R})_{\text{MM}}$ is the energy of DRG–R complex obtained from AMBER force field treatment and $E_3(\text{DRG})_{\text{QM}}$ and $E_4(\text{DRG})_{\text{MM}}$ are the energy of the DRG in bound form obtained from DFT (B3LYP/6-31G*) and AMBER treatment, respectively.

$$E_5 = E_1(\text{DRG-R})_{\text{ONIOM}} - \left[E_6(\text{DRG})_{\text{QM}} + E_7(\text{R})_{\text{MM}} \right]$$

$E_6(\text{DRG})_{\text{QM}}$ is the QM energy of isolated DRG and $E_7(\text{R})_{\text{MM}}$ the MM energy of R alone.

Major shortcomings of MM docking algorithms arise due to its incapability of treating the role of charge polarization in determining the ligand's bound conformation, henceforth, limiting the prediction of the correct binding mode. Schrödinger suite of programs [48] has introduced QM-Polarized Ligand Docking (QPLD) algorithm to its QM-MM software QSite. QPLD uses ab initio charge calculations to consider charge polarization induced by the protein environment making the docking results even more accurate. The QM-MM approach provides the conformational landscape of drugs in the bound form that are not easily obtained through conventional docking methods. If applied suitably, such approach can indirectly provide an insight about the actual binding mode of small molecules when experimental details are not available.

5 Quantitative Structure–Activity Relationship Approach

Drug discovery itself remains a complex and time consuming process and over the period of time during recent decades, the search of new drugs for complex diseases has shifted towards multi-target drugs as with the invent of cheminformatics it has become possible to identify multiple targets for a potential drug candidate [49, 50]. Quantitative structure–activity relationship (QSAR) utilizes the correlation between molecular structure of chemical species and their corresponding activity viz-*a-viz* biological receptors [51, 52]. This is possible by establishing a mathematical relationship with the structural features (known as molecular descriptors) of small molecules and biological targets. These descriptors are used in calculations based on dimensionality. Features such as molecular weight, types of atoms, functional groups, bonding patterns, and substitution are considered in zero- and one-dimensional QSAR calculations. Similarly, more complex descriptors such as molecular connectivity indices, molecular surface, molecular volume, electronic, steric, shape, volume, rotatable bonds, interatomic distances, electronegativity, atom distribution, aromaticity, solvation properties, etc. are considered in 2D and 3D QSAR calculations. Based on the dimensionality of the applied descriptors, two most applied 3D-QSAR models are CoMFA (Comparative Molecular Field Analysis) and CoMSIA (Comparative Molecular Similarity Index Analysis). Further, more complex descriptors also include the conformational variation of small molecules and other experimental insight into the structure of the molecules. QSAR

calculations essentially require large datasets of small molecules and in some cases biological macromolecules as well. These datasets are divided into training and test sets in order to validate the calculations. Multi-dimensional data is often analyzed using multivariate analysis [53]. To accomplish this, advanced data mining approaches are utilized in evaluation of the descriptors and their corresponding correlation with the activities [53, 54].

Conventional QSAR methods explore the correlation with respect to a single receptor to identify potential small molecules to be used as therapeutic agents. However, in order to identify multiple targets for a set of chemical compounds, significant modifications have been done in the form of multi-target QSAR method (mt-QSAR) [55–57]. Normal course of computer-aided drug design approach uses QSAR dataset for standard support vector regression to solve the constrained optimization problem in search for a drug-like molecule having high binding affinity with a desired biological target. Rosenbaum et al. have elaborated a multi-target QSAR approach using multi-task algorithms to exploit the similarity between several targets to translate information between the target-specific QSAR models [58]. mt-QSAR models are used to link the structures of potential drugs with biological activity against a number of diverse targets [59–62]. Much recently, advancements in chemoinformatics has led to two novel *in silico* approaches—first having ability to integrate multiple chemical and biological data into a single multi-tasking model for quantitative structure–biological effect relationships (mtk-QSBER) and the second resulting from combination of the perturbation theory with quantitative structure–property relationships modeling tools (pt-QSPR) [63]. The mt-QSBER model is based on a particular modeling technique such as linear discriminant analysis (LDA), multiple linear regression (MLR), or machine learning tools like artificial neural networks (ANN). Pt-QSPR model combines the moving average approach with perturbation theory to address the deviations by considering all possible combinations of the dataset cases [63]. Another novel approach for descriptors, MARCH-INSIDE has been developed to incorporate specific information about biphasic partition system, biological species, and chemical structure [62, 64]. Using advanced QSAR methods, many biological activities such as pharmacological and toxicity effects have been successfully predicted [65]. Therefore, there exists enormous potential for making use of mt-QSAR studies in the development of multi-target drug designing.

General methodology followed in a QSAR experiment is depicted in Fig. 7. The QSAR modeling generally involves following steps: (1) Collection of experimental data and designing of training and test sets of chemicals, (2) Generation of descriptors to relate appropriately the chemical structure to biological activities, (3) Creation of a QSAR model based on variable analysis or statistical methods using regression analysis and its evaluation for predictive

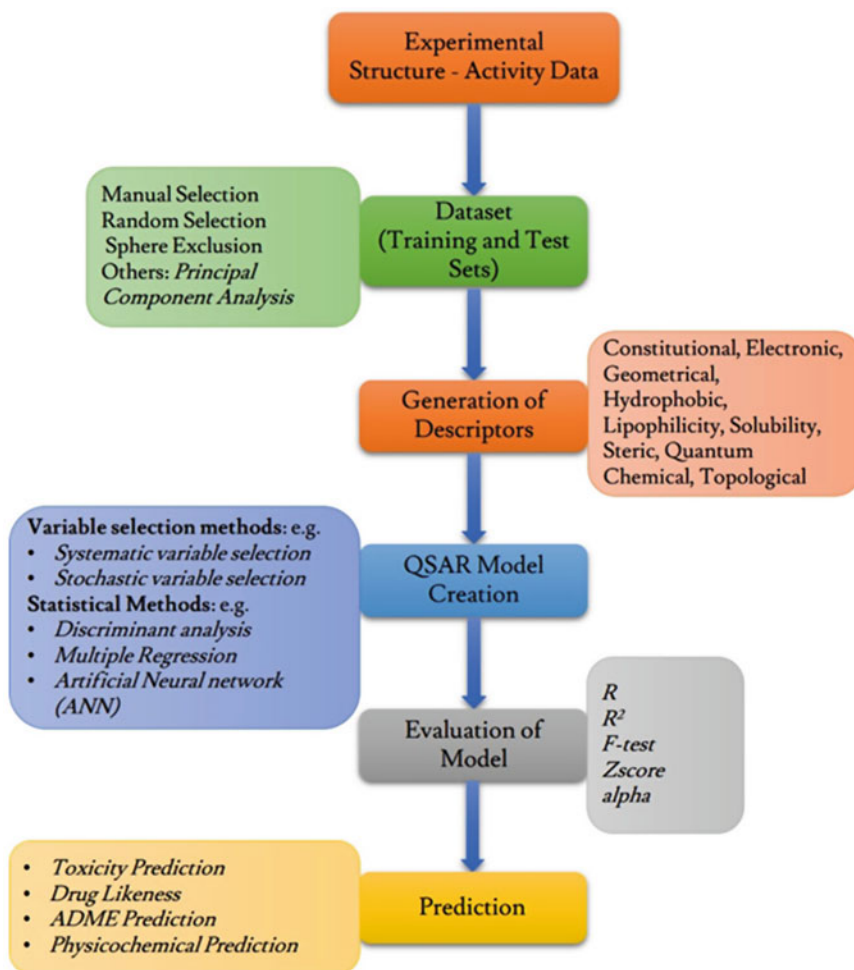


Fig. 7 Overview of quantitative structure–activity relationship (QSAR) methodologies

capacity in terms of goodness of fit, stability as well as robustness of model, and (4) Application of validated model for evaluation using appropriate software for prediction of biological activity.

6 Molecular Dynamics Approach

One major drawback of rigid-body docking is that the estimates obtained do not always conform to the exact binding energy of the complex and unable to reflect the true structural profile of ligand *viz-a-viz* receptor. This is largely due to the fact that during actual interaction, not only ligand but also the conformation of the receptor atoms may change to accommodate the ligand. This is especially true to the binding site region of the receptor where structural disturbances are often accompanied by dislodging of solvent

molecules and conformational rearrangement of a few side chains. In many cases, water molecules occupy an important place in the structural integrity of a receptor. For example, DNA minor groove possesses a spine of hydration which needs to be disturbed in order for the drug molecule to bind in the minor groove. Similarly, studies have revealed that sampling of such solvent molecules is an essential prerequisite to estimate properly the structural feature of ligand–receptor complexes. Conformational changes due to bonded as well as non-bonded long-range van der Waals and electrostatic molecular interactions are also needed to be addressed by considering dynamic properties of biological ensembles. Conventional docking protocols often fail to achieve this and may be treated with molecular dynamics simulations [66] of both types—classical and quantum. In the classical mechanics treatment, the atoms are treated as balls and chemical bonds as springs and the laws of classical mechanics are applied to define the system dynamics. In the quantum mechanics-based MD simulations, quantum nature of the chemical bond is taken into account and the electron density functional for the valence bonding electrons is treated using quantum equations, whereas the dynamics of ions is treated with classical mechanics.

Using state-of-the-art software tools and computational power, one can easily predict the forces acting on each atom by every other atom in the system and consequently obtain the information about the structure of the molecule at any given point in time. Thus, molecular dynamic simulations provide information about conformational states of molecules as a function of time and are usually achieved by applying classical Newtonian movement. The trajectories of atoms and molecules are usually determined by numerically solving Newton's equations of motion for a system of interacting particles, where forces between the non-bonded particles are often calculated using interacting atomic pair empirical potentials, mainly comprising of the mathematical models based on Lennard–Jones (LJ) and Coulomb models [67, 68]. LJ potential is a simple mathematical model that describes the approximation of van der Waals interactions between two uncharged particles with an advantage of easy computability. The LJ equation, considering the distance between the two particles to calculate the potential energy of the interaction, is given below:

$$E_{LJ} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (1)$$

In above equation, E_{LJ} is the intermolecular potential between the two atoms separated by the distance r . ϵ is the potential energy well depth and is a measure of attraction between two particles. σ is the closest distance at which the intermolecular potential between

the two particles is zero and may be referred to as the van der Waals radius.

The electrostatic interactions between non-bonded charged particles are considered using Coulomb's electrostatic potential [69]. The equation for calculation of the Coulomb interaction potential by taking into account the charges q_1 and q_2 of two atoms and the distance r between them is given below:

$$E_p = \frac{q_1 q_2}{4\pi\epsilon_0 r} \quad (2)$$

where E_p is Coulombic potential energy, and ϵ_0 is the permittivity of vacuum. When $E_p < 0$, then the interaction is attractive, while if $E_p > 0$, then the interaction is repulsive.

To run MD simulations, the force on each particle is required to be known so that the acceleration is calculated as a function of gradient of potential energy, since each single atom is affected by potential energy function of every other bonded or non-bonded atom in the system. Force field may be presented by a mathematical expression describing the dependence of the energy of a system on the coordinates of its particles and is expressed in terms of interatomic potential energy due to bonded and non-bonded interactions.

$$V(R) = E_{\text{bonded}} + E_{\text{non-bonded}} \quad (3)$$

As mentioned above, the non-bonded potential may be considered as a combination of LJ and Coulomb's electrostatic potentials (Eqs. 1 and 2), the bonded potential is the result of three intramolecular degrees of freedom due to stretching along bond, bending motion of bonds, and rotation around the bonds. A typical MD force field equation given below, therefore, consists of the following potential terms:

$$V(R) = \sum_{\text{bonds}} \frac{a_i}{2} (l_i - l_{i0})^2 + \sum_{\text{angles}} \frac{b_i}{2} (\theta_i - \theta_{i0})^2 + \sum_{\text{torsions}} \frac{c_i}{2} (1 + \cos(n\omega_i - \gamma_i)) \\ + \frac{1}{2} \sum_{\substack{i=1 \\ \text{non bonded}}}^N \sum_{\substack{j \neq 1 \\ \text{non bonded}}}^N 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{1}{2} \sum_{\substack{i=1 \\ \text{non bonded}}}^N \sum_{\substack{j \neq 1 \\ \text{non bonded}}}^N k \frac{q_i q_j}{r_{ij}} \quad (4)$$

The first three terms in Eq. 4 represent internal degrees of freedom of a molecule through the bonded atoms. The first term calculates the energy contribution due to change in bond lengths while the second term signifies the potential due to valence angles of the molecule. Similarly, the third term denotes the torsional potential due to fluctuation in energy originating from change in torsion angles. The fourth term in the force field gives estimates for van der Waal's interactions while the last term estimates the

Coulombic interaction potential between charged pairs. In principle, the interaction potential between every non-bonded pair of atoms should be computed. Nevertheless, the interactions between atom pairs separated by a distance larger than the predefined cut-off distance are ignored to speed up the computation without affecting much the extent of accuracy. Further, water as the solvent in biological systems influences the structure and dynamics of interacting molecules mainly by screening their electrostatic interaction. Solvent effect is usually implied in MD simulation by including the dielectric screening constant in the electrostatic term of the potential energy function (Eq. 5) and not including the explicit water molecules.

$$V_{\text{elect.}} = \frac{q_i q_j}{\epsilon_{\text{eff}} r_{ij}} \quad (5)$$

where q_i , q_j are the partial atomic charges, ϵ_{eff} is the effective dielectric constant, and r_{ij} is the relative distance between the two particles. Often the value of effective dielectric constant (ϵ_{eff}) itself is taken to be distance (r_{ij}) dependent. MD simulations are also better suited for such structural contingencies in which the role of solvent could be crucial [70]. In recent times, several continuum electrostatic theory-based solvent models have also been developed.

In MD simulations, the time-dependent behavior of a set of interacting atoms is obtained by integration of Newton's equations of motion:

$$\mathbf{F}_i = m_i \frac{d^2 \mathbf{r}_i(t)}{dt^2} \quad (6)$$

where $\mathbf{r}_i(t)$ is the position vector of i th atom, \mathbf{F}_i is the force acting upon i th atom at time t , and m_i is the mass of the atom. Integration of Newton's equation of motion is used to define position $\mathbf{r}_i(t + \Delta t)$ of the i th atom at time $t + \Delta t$. Several algorithms are available for its solution, but most commonly used one is the Verlet algorithm due to its simplicity. Thereby, MD simulations result in a time-dependent series of conformations known as MD trajectories which are defined by both positions and velocities. The MD trajectories of moving particles may be analyzed for getting three types of properties, i.e., thermodynamic, structural, and dynamical. Therefore, the functional properties of the biomolecular system that are influenced by the dynamic events can be estimated at the atomic level by MD simulations. Molecular dynamics essentially provides several conformations of a receptor molecule to sample thermodynamically viable receptor flexibility. Another important benefit of molecular dynamics simulation is that it can help in the identification of transition states in a complexation event between a ligand and receptor [71].

A Molecular Dynamics simulation involves the following steps. An initial configuration of the system at $t = 0$ is selected. For biomolecule, mostly an X-ray crystal structure from the Protein Databank is taken as initial structure. If required, an energy minimization of the structure is done prior to a molecular dynamics simulation. A valid force field is selected, a cut-off radius is selected, and in case of the presence of partial charges, the method to treat the electrostatic interactions and an integration algorithm are chosen. Coordinates and velocities of the system from an MD simulation are saved for further analysis. Average structures are calculated and visualized to reflect the conformational changes. Time-dependent properties are displayed graphically with one of the axis corresponding to time. Having access to the positions of atoms, velocities, and forces as function of time, the properties such as mean energy, rms difference, rms fluctuation, etc. can be computed. Commonly available programs for performing MD simulations are given in Table 4, the reliability of results depends on the force field

Table 4
List of commonly used programs for MD simulations

Software	Description	Link
NAMD	Distributed freely, parallel molecular dynamics code designed for high-performance simulation of large biomolecular system uses VMD visualize for simulation setup and trajectory analysis [72]	http://www.ks.uiuc.edu/Research/namd/
GROMACS (GROningenMACHine for Chemical Simulations)	Public domain high-performance program capable to simulate the Newtonian equations of motion for systems with hundreds to millions of particles [73]	http://www.gromacs.org/Downloads
CHARMM (Chemistry at HARvard Macromolecular Mechanics)	For application to many-particle systems with a comprehensive set of energy functions. <i>Multi-scale techniques</i> including QM/MM, MM/CG, and several implicit solvent models available [74]	http://charmm.chemistry.harvard.edu/
AMBER	Freeware package of programs for molecular dynamics simulations of proteins and nucleic acids [75]	http://ambermd.org/GetAmber.php
GROMOS (GROningen MOlecular Simulation)	Freeware program package for the dynamic modeling of biomolecules using methods of molecular dynamics, stochastic dynamics, energy minimization, and path integral formalism [76]	http://www.gromos.net
TINKER	Public domain software for molecular mechanics and dynamics calculations, with some special features for biopolymers [77, 78]	https://dasher.wustl.edu/tinker/

employed to model the intra- and intermolecular interactions and the accuracy is usually validated by comparison with experimental results. VMD, gOpenMol, and Rasmol are some freely available tools for visualization and analysis of input and output results.

7 Convergence of Theoretical and Experimental Results

Computational methods are best at predicting the interaction patterns of small molecules with biological receptors and latest advancements in computational techniques have made them particularly useful in identifying multiple targets for already approved drug molecules. The techniques and methods discussed above have clearly established the usefulness of computer-driven methods in enhancing the pace of multi-target drug discovery research in particular. However, the ultimate test of all such computational methods rests on the ability of these methods to provide accurate estimates that can be verified in the living system. Experimental validation, therefore, becomes essential for every prediction. Several studies have taken into consideration both the experimental and computational aspects of drug–receptor interactions [25–32] thereby providing a useful insight as to the extent the prediction methods apply in real-world scenarios.

References

1. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. *Cell* 100:57–70. [https://doi.org/10.1016/S0092-8674\(00\)81683-9](https://doi.org/10.1016/S0092-8674(00)81683-9)
2. Oechsner M, Buhmann C, Strauss J, Stuerenburg HJ (2002) COMT-inhibition increases serum levels of dihydroxyphenylacetic acid (DOPAC) in patients with advanced Parkinson's disease. *J Neural Transm* 109(1):69–75. <https://doi.org/10.1007/s702-002-8237-z>
3. Hartman IVJL, Garvik B, Hartwel L (2001) Principles for the buffering of genetic variation. *Science* 291:1001–1004. <https://doi.org/10.1126/science.1056072>
4. Bonander N, Bill RM (2009) Relieving the first bottleneck in the drug discovery pipeline: using array technologies to rationalize membrane protein production. *Expert Rev Proteomics* 6:501–505. <https://doi.org/10.1586/epr.09.65>
5. Gillespie SH, Singh K (2011) XDR-TB, what is it; how is it treated; and why is therapeutic failure so high? *Recent Pat Antiinfect Drug Discov* 6:77–83
6. Horst JA, Laurenzi A, Bernard B, Samudrala R (2012) Computational multitarget drug discovery. In: *Polypharmacology in drug discovery*. Wiley, New York, pp 263–301
7. Sacks LV, Behrman RE (2009) Challenges, successes and hopes in the development of novel TB therapeutics. *Future Med Chem* 1:749–756. <https://doi.org/10.4155/fmc.09.53>
8. Keith CT, Borisy AA, Stockwell BR (2005) Innovation: multicomponent therapeutics for networked systems. *Nat Rev Drug Discov* 4:71–78. <https://doi.org/10.1038/nrd1609>
9. Ekins S, Williams AJ, Krasowski MD, Freundlich JS (2011) In silico repositioning of approved drugs for rare and neglected diseases. *Drug Discov Today* 16. <https://doi.org/10.1016/j.drudis.2011.02.016>
10. Jenwitheesuk E, Horst JA, Rivas KL, Van Voorhis WC, Samudrala R (2008) Novel paradigms for drug discovery: computational multitarget screening. *Trends Pharmacol Sci* 29:62–71. <https://doi.org/10.1016/J.TIPS.2007.11.007>
11. Jenwitheesuk E, Samudrala R (2005) Identification of potential multitarget antimalarial drugs. *JAMA* 294:1487. <https://doi.org/10.1001/jama.294.12.1490>

12. Minie M, Chopra G, Sethi G, Horst J, White G, Roy A, Hatti K, Samudrala R (2014) CANDO and the infinite drug discovery frontier. *Drug Discov Today* 19:1353–1363. <https://doi.org/10.1016/j.drudis.2014.06.018>
13. Ren J, Xie L, Li WW, Bourne PE (2010) SMAP-WS: a parallel web service for structural proteome-wide ligand-binding site comparison. *Nucleic Acids Res* 38:W441–W444. <https://doi.org/10.1093/nar/gkq400>
14. Swamidass SJ (2011) Mining small-molecule screens to repurpose drugs. *Brief Bioinform* 12:327–335. <https://doi.org/10.1093/bib/bbr028>
15. Xu K, Cote TR (2011) Database identifies FDA-approved drugs with potential to be repurposed for treatment of orphan diseases. *Brief Bioinform* 12:341–345. <https://doi.org/10.1093/bib/bbr006>
16. Anighoro A, Bajorath J, Rastelli G (2014) Polypharmacology: challenges and opportunities in drug discovery. *J Med Chem* 57(19):7874–7887. <https://doi.org/10.1021/jm5006463>
17. Hopkins AL, Mason JS, Overington JP (2006) Can we rationally design promiscuous drugs? *Curr Opin Struct Biol* 16(1):127–136. <https://doi.org/10.1016/j.sbi.2006.01.013>
18. Krug M, Hilgeroth A (2008) Recent advances in the development of multi-kinase inhibitors. *Mini Rev Med Chem* 8(13):1312–1327. <https://doi.org/10.2174/138955708786369591>
19. Adrian G, Marcel V, Robert B, Richard T (2007) A comparison of physicochemical property profiles of marketed oral drugs and orally bioavailable anti-cancer protein kinase inhibitors in clinical development. *Curr Top Med Chem* 7(14):1408–1422. <https://doi.org/10.2174/156802607781696819>
20. Nahta R, Yu D, Hung MC, Hortobagyi GN, Esteva FJ (2006) Mechanisms of disease: understanding resistance to HER2-targeted therapy in human breast cancer. *Nat Clin Pract Oncol* 3(5):269–280. <https://doi.org/10.1038/ncponc0509>
21. Sergina NV, Rausch M, Wang D, Blair J, Hann B, Shokat KM, Moasser MM (2007) Escape from HER-family tyrosine kinase inhibitor therapy by the kinase-inactive HER3. *Nature*. <https://doi.org/10.1038/nature05474>
22. Tabernero J (2007) The role of VEGF and EGFR inhibition: implications for combining anti-VEGF and anti-EGFR agents. *Mol Cancer Res* 5:203–220
23. Cheng F, Liu C, Jiang J, Lu W, Li W, Liu G, Zhou W, Huang J, Tang Y (2012) Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput Biol* 8. <https://doi.org/10.1371/journal.pcbi.1002503>
24. O’Meara MJ, Ballouz S, Shoichet BK, Gillis J (2016) Ligand similarity complements sequence, physical interaction, and co-expression for gene function prediction. *PLoS One* 11(7):e0160098. <https://doi.org/10.1371/journal.pone.0160098>
25. Bhattacharjee P, Sarkar S, Pandya P, Bhadra K (2016) Targeting different RNA motifs by beta carboline alkaloid, harmalol: a comparative photophysical, calorimetric, and molecular docking approach. *J Biomol Struct Dyn* 34(12):2722–2740. <https://doi.org/10.1080/07391102.2015.1126694>
26. Sarkar S, Pandya P, Bhadra K (2014) Sequence specific binding of beta carboline alkaloid harmalol with deoxyribonucleotides: binding heterogeneity, conformational, thermodynamic and cytotoxic aspects. *PLoS One* 9(9):e108022. <https://doi.org/10.1371/journal.pone.0108022>
27. Pandya P, Agarwal LK, Gupta N, Pal S (2014) Molecular recognition pattern of cytotoxic alkaloid vinblastine with multiple targets. *J Mol Graph Model* 54:1–9. <https://doi.org/10.1016/j.jmgm.2014.09.001>
28. Masum AA, Chakraborty M, Pandya P, Halder UC, Islam MM, Mukhopadhyay S (2014) Thermodynamic study of rhodamine 123-calf thymus DNA interaction: determination of calorimetric enthalpy by optical melting study. *J Phys Chem B* 118(46):13151–13161. <https://doi.org/10.1021/jp509326r>
29. Islam MM, Chakraborty M, Pandya P, Al Masum A, Gupta N, Mukhopadhyay S (2013) Binding of DNA with Rhodamine B: spectroscopic and molecular modeling studies. *Dyes Pigments* 99(2):412–422. <https://doi.org/10.1016/j.dyepig.2013.05.028>
30. Pandya P, Gupta SP, Pandav K, Barthwal R, Jayaram B, Kumar S (2012) DNA binding studies of Vinca alkaloids: experimental and computational evidence. *Nat Prod Commun* 7(3):305–309
31. Pandya P, Islam MM, Kumar GS, Jayaram B, Kumar S (2010) DNA minor groove binding of small molecules: experimental and computational evidence. *J Chem Sci* 122(2):247–257
32. Islam MM, Pandya P, Kumar S, Kumar GS (2009) RNA targeting through binding of small molecules: studies on t-RNA binding by the cytotoxic protoberberine alkaloid coralayne.

- Mol Biosyst 5(3):244–254. <https://doi.org/10.1039/b816480k>
33. Islam MM, Pandya P, Chowdhury SR, Kumar S, Kumar GS (2008) Binding of DNA-binding alkaloids berberine and palmatine to tRNA and comparison to ethidium: spectroscopic and molecular modeling studies. *J Mol Struct* 891(1–3):498–507. <https://doi.org/10.1016/j.molstruc.2008.04.043>
 34. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31(2):455–461. <https://doi.org/10.1002/jcc.21334>
 35. Chang MW, Ayeni C, Breuer S, Torbett BE (2010) Virtual screening for HIV protease inhibitors: a comparison of AutoDock 4 and Vina. *PLoS One* 5(8):e11955. <https://doi.org/10.1371/journal.pone.0011955>
 36. Shaikh SA, Jayaram B (2007) A swift all-atom energy-based computational protocol to predict DNA-ligand binding affinity and Delta Tm. *J Med Chem* 50(9):2240–2244. <https://doi.org/10.1021/jm060542c>
 37. Gupta A, Sharma P, Jayaram B (2007) ParDOCK: an all atom energy based Monte Carlo docking protocol for protein-ligand complexes. *Protein Pept Lett* 14(7):632–646. <https://doi.org/10.2174/092986607781483831>
 38. Lin J-H, Perryman AL, Schames JR, McCammon JA (2002) Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *J Am Chem Soc* 124(20):5632–5633. <https://doi.org/10.1021/ja0260162>
 39. Amaro RE, Baron R, McCammon JA (2008) An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J Comput Aided Mol Des* 22(9):693–705. <https://doi.org/10.1007/s10822-007-9159-2>
 40. Yuriev E, Agostino M, Ramsland PA (2011) Challenges and advances in computational docking: 2009 in review. *J Mol Recognit* 24(2):149–164. <https://doi.org/10.1002/jmr.1077>
 41. Ellingson SR, Smith JC, Baudry J (2013) VinaMPI: facilitating multiple receptor high-throughput virtual docking on high-performance computers. *J Comput Chem* 34(25):2212–2221. <https://doi.org/10.1002/jcc.23367>
 42. Pan JB, Ji N, Pan W, Hong R, Wang H, Ji ZL (2014) High-throughput identification of off-targets for the mechanistic study of severe adverse drug reactions induced by analgesics. *Toxicol Appl Pharmacol* 274(1):24–34. <https://doi.org/10.1016/j.taap.2013.10.017>
 43. Amaro RE, Schnauffer A, Interthal H, Hol W, Stuart KD, McCammon JA (2008) Discovery of drug-like inhibitors of an essential RNA-editing ligase in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* 105(45):17278–17283. <https://doi.org/10.1073/pnas.0805820105>
 44. Hui-Fang L, Qing S, Jian Z, Wei F (2010) Evaluation of various inverse docking schemes in multiple targets identification. *J Mol Graph Model* 29(3):326–330. <https://doi.org/10.1016/j.jmgm.2010.09.004>
 45. Chopra G, Samudrala R (2016) Exploring polypharmacology in drug discovery and repurposing using the CANDO platform. *Curr Pharm Des* 22:3109–3123. <https://doi.org/10.2174/1381612822666160325121943>
 46. Lodola A, De Vivo M (2012) The increasing role of QM/MM in drug discovery. *Adv Protein Chem Struct Biol* 87:337–362. <https://doi.org/10.1016/B978-0-12-398312-1.00011-1>
 47. Gleeson MP, Gleeson D (2009) QM/MM calculations in drug discovery: a useful method for studying binding phenomena? *J Chem Inf Model* 49(3):670–677. <https://doi.org/10.1021/ci800419j>
 48. Schrödinger L (2017) Schrödinger Suite 2017–4 QM-Polarized Ligand Docking protocol; Glide, Schrödinger, LLC, New York, NY, 2017; Jaguar, Schrödinger, LLC, New York, NY, 2017; QSite, Schrödinger, LLC, New York, NY, 2017
 49. Speck-Planche A, Cordeiro MN (2015) Multi-tasking models for quantitative structure-biological effect relationships: current status and future perspectives to speed up drug discovery. *Expert Opin Drug Discov* 10(3):245–256. <https://doi.org/10.1517/17460441.2015.1006195>
 50. Liu X, Zhu F, Ma X, Shi Z, Yang S, Wei Y, Chen Y (2013) Predicting targeted polypharmacology for drug repositioning and multi-target drug discovery. *Curr Med Chem* 20(13):1646–1661
 51. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, Dearden J, Gramatica P, Martin YC, Todeschini R, Consonni V, Kuz'min VE, Cramer R, Benigni R, Yang C, Rathman J, Terfloth L, Gasteiger J, Richard A, Tropsha A (2014) QSAR modeling: where have you been? Where are you going to? *J Med Chem* 57(12):4977–5010. <https://doi.org/10.1021/jm4004285>

52. Roy K, Kar S, Das RN (2015) Selected statistical methods in QSAR. In: Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment. Academic Press, Amsterdam, pp 191–229. <https://doi.org/10.1016/b978-0-12-801505-6.00006-5>
53. Nantasenamat C, Worachartcheewan A, Jamsak S, Preeyanon L, Shoombuatong W, Simeon S, Mandi P, Isarankura-Na-Ayudhya C, Prachayasittikul V (2015) AutoWeka: toward an automated data mining software for QSAR and QSPR studies. *Methods Mol Biol* 1260:119–147. https://doi.org/10.1007/978-1-4939-2239-0_8
54. Yang Y, Lin T, Weng XL, Darr JA, Wang XZ (2011) Data flow modeling, data mining and QSAR in high-throughput discovery of functional nanomaterials. *Comput Chem Eng* 35(4):671–678. <https://doi.org/10.1016/j.compchemeng.2010.04.018>
55. Vina D, Uriarte E, Orallo F, Gonzalez-Diaz H (2009) Alignment-free prediction of a drug-target complex network based on parameters of drug connectivity and protein sequence of receptors. *Mol Pharm* 6(3):825–835. <https://doi.org/10.1021/mp800102c>
56. Geronikaki A, Druzhilovsky D, Zakharov A, Poroikov V (2008) Computer-aided prediction for medicinal chemistry via the Internet. *SAR QSAR Environ Res* 19(1–2):27–38. <https://doi.org/10.1080/10629360701843649>
57. Marzaro G, Chilin A, Guiotto A, Uriarte E, Brun P, Castagliuolo I, Tonus F, Gonzalez-Diaz H (2011) Using the TOPS-MODE approach to fit multi-target QSAR models for tyrosine kinases inhibitors. *Eur J Med Chem* 46(6):2185–2192. <https://doi.org/10.1016/j.ejmech.2011.02.072>
58. Rosenbaum L, Dorr A, Bauer MR, Boeckler FM, Zell A (2013) Inferring multi-target QSAR models with taxonomy-based multi-task learning. *J Cheminform* 5(1):33. <https://doi.org/10.1186/1758-2946-5-33>
59. Prado-Prado FJ, Gonzalez-Diaz H, de la Vega OM, Ubeira FM, Chou KC (2008) Unified QSAR approach to antimicrobials. Part 3: First multi-tasking QSAR model for input-coded prediction, structural back-projection, and complex networks clustering of antiprotozoal compounds. *Bioorg Med Chem* 16(11):5871–5880. <https://doi.org/10.1016/j.bmc.2008.04.068>
60. Prado-Prado FJ, Martinez de la Vega O, Uriarte E, Ubeira FM, Chou KC, Gonzalez-Diaz H (2009) Unified QSAR approach to antimicrobials. 4. Multi-target QSAR modeling and comparative multi-distance study of the giant components of antiviral drug-drug complex networks. *Bioorg Med Chem* 17(2):569–575. <https://doi.org/10.1016/j.bmc.2008.11.075>
61. Cruz-Monteagudo M, Borges F, Cordeiro MN, CagideFajin JL, Morell C, Ruiz RM, Canizares-Carmenate Y, Dominguez ER (2008) Desirability-based methods of multiobjective optimization and ranking for global QSAR studies. Filtering safe and potent drug candidates from combinatorial libraries. *J Comb Chem* 10(6):897–913. <https://doi.org/10.1021/cc800115y>
62. Prado-Prado FJ, Garcia-Mera X, Gonzalez-Diaz H (2010) Multi-target spectral moment QSAR versus ANN for antiparasitic drugs against different parasite species. *Bioorg Med Chem* 18(6):2225–2231. <https://doi.org/10.1016/j.bmc.2010.01.068>
63. Speck-Planche A, Cordeiro M (2017) Advanced in silico approaches for drug discovery: mining information from multiple biological and chemical data through mtk-QSBER and pt-QSPR strategies. *Curr Med Chem* 24(16):1687–1704. <https://doi.org/10.2174/0929867324666170124152746>
64. Gonzalez-Diaz H, Aguero G, Cabrera MA, Molina R, Santana L, Uriarte E, Delogu G, Castanedo N (2005) Unified Markov thermodynamics based on stochastic forms to classify drugs considering molecular structure, partition system, and biological species: distribution of the antimicrobial G1 on rat tissues. *Bioorg Med Chem Lett* 15(3):551–557. <https://doi.org/10.1016/j.bmcl.2004.11.059>
65. Lagunin A, Zakharov A, Filimonov D, Poroikov V (2011) QSAR modelling of rat acute toxicity on the basis of PASS prediction. *Mol Inform* 30(2–3):241–250. <https://doi.org/10.1002/minf.201000151>
66. Allen MP (2004) Introduction of molecular dynamics simulation. In: Attig N, Binder K, Grubmuller H, Kremer K (eds) *Computational soft matter: from synthetic polymers to proteins*, Lecture notes, NIC series, vol 23. John von Neumann Institute for Computing, Julich, pp 1–28
67. Sagui C, Darden TA (1999) Molecular dynamics simulations of biomolecules: long-range electrostatic effects. *Annu Rev Biophys Biomol Struct* 28:155–179. <https://doi.org/10.1146/annurev.biophys.28.1.155>
68. Verlet L (1967) Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys Rev* 159(1):98

69. Atkins P, Paula J (2006) *Physical chemistry for the life sciences*. W H Freeman & Co, New York
70. Zhao H, Cafisch A (2015) Molecular dynamics in drug design. *Eur J Med Chem* 91:4–14. <https://doi.org/10.1016/j.ejmech.2014.08.004>
71. Hospital A, Goñi JR, Orozco M, Gelpí JL (2015) Molecular dynamics simulations: advances and applications. *Adv Appl Bioinform Chem* 8:37–47. <https://doi.org/10.2147/AABC.S70333>
72. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26(16):1781–1802. <https://doi.org/10.1002/jcc.20289>
73. Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun* 91(1–3):43–56. [https://doi.org/10.1016/0010-4655\(95\)00042-e](https://doi.org/10.1016/0010-4655(95)00042-e)
74. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4(2):187–217. <https://doi.org/10.1002/jcc.540040211>
75. Case DA, Cerutti DS, Cheatham TE III, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Greene D, Homeyer N, Izadi S, Kovalenko A, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Mermelstein D, Merz KM, Monard G, Nguyen H, Omelyan I, Onufriev A, Pan F, Qi R, Roe DR, Roitberg A, Sagui C, Simmerling CL, Bello-Smith WM, Swails J, Walker RC, Wang J, Wolf RM, Wu X, Xiao L, York DM, Kollman PA (2017) AMBER 2017, 17th edn. University of California, San Francisco
76. Scott WRP, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Krüger P, van Gunsteren WF (1999) The GROMOS biomolecular simulation program package. *J Phys Chem A* 103(19):3596–3607. <https://doi.org/10.1021/jp984217f>
77. Lagardère L, Jolly L-H, Lipparini F, Aviat F, Stamm B, Jing ZF, Harger M, Torabifard H, Cisneros GA, Schnieders MJ, Gresh N, Maday Y, Ren PY, Ponder JW, Piquemal J-P (2018) Tinker-HP: a massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields. *Chem Sci* 9(4):956–972. <https://doi.org/10.1039/c7sc04531j>
78. Harger M, Li D, Wang Z, Dalby K, Lagardere L, Piquemal JP, Ponder J, Ren P (2017) Tinker-OpenMM: absolute and relative alchemical free energies using AMOEBA on GPUs. *J Comput Chem* 38(23):2047–2055. <https://doi.org/10.1002/jcc.24853>



Computational Multi-Target Drug Design

Azizeh Abdolmaleki, Fereshteh Shiri, and Jahan B. Ghasemi

Abstract

Multi-target (mt) therapy is an attractive approach as well as a challenging task in drug discovery research and pharmaceutical industry. The multi-target drug design strategy is an opportunity to find new drugs for the treatment of two or more targets simultaneously. Advanced characterization of bioactive molecules, computational science, and molecular biology have contributed to planning of new bioactive compounds and evaluating different features of multi-targeted drugs. Computational methods have different roles in drug candidate searching, analysis, and prediction in this field. Here, we discuss several *in silico* methodologies and computer-aided drug design (CADD) as structure-activity relationship (SAR), quantitative SAR (QSAR), pharmacophore modeling, and molecular docking in the process of drug discovery in the field of multi-targeted drugs (MTDs). Computational efficiency of each method has been stated in relation to their key strength and weakness. These capacities for binding affinity prediction are rationally effective with promising potential in easing drug discovery directed at selective multiple targets.

Keywords CADD, Drug discovery, Molecular docking, MTD/MTDD, Multi-target, Pharmacophore, QSAR, SAR

1 Multi-Targeted or Polypharmacological Drug Discovery

The key idea of “one drug-one target-one disease” is single protein target identification whose inhibition can lead to success in the treatment of the identified disease. There is no side effect based on the principle assumption in highly selective ligands via binding to secondary non-therapeutic targets. Recently, the results of network biology and post-genomic studies have revealed that proteins are not isolated and act as a portion of a network with a highly connected robust system [1]. Thus, a growing interest in emerging therapeutics that influences on multiple targets at the same time has developed.

Multi-targeted drugs (MTDs) are generally accepted due to the ability of the polypharmacology of these drugs to modulate the bioactivity of multiple targets [2, 3]. Polypharmacology is the use or design of a single drug that works on the “multiple targets” or “disease pathways.” It offers therapeutic agents that can modulate “multiple targets.” Accidental drug-target interactions could produce adverse effects or toxicity [4]. Some of these effects have been

connected to unintended contacts with specific off-targets, i.e., the cardiac hERG channel for astemizole, terfenadine, and grepafloxacin [5]; the serotonin 5HT_{2B} receptor for fenfluramine [6]; and the M₂ receptor for rapacuronium [7]. Polypharmacology opens a window to logically design subsequent production of therapeutic agents with more efficiency and less toxicity [4, 8]. Therefore, it would be either responsible for side effects or the best to reach a high level of control over drug specificity [9, 10]. In this regard, Morphy described MTD in several publications using the term “designed multiple ligands.” This term exhibits at least something responsible for MTD efficacy [11, 12]. Other terms such as “dirty drug” or “promiscuous” have been used for their multiple bioactivities that can be helpful for improved efficacy or may not be positive for unwanted pharmacology or toxicity. Analysis of the physicochemical data of candidate drugs indicates that toxicity risk and clinical failure of “promiscuous or dirty compounds” are due to high lipophilicity [12]. It is essential to point out that the phrase “promiscuous” has also been applied in the area of high-throughput screening (HTS) to represent HTS analytes’ bioactivity associated with typically undesirable physicochemical or chemical features. This fact may be linked to covalent bond formation of ligand-target complex [11] or undesirable biophysical effects in vitro such as aggregation of colloidal particles [13]. In these two cases, the observed effects are unwanted biological activities. But, in case of in vivo situation, this colloidal aggregation may be favorable in the gastrointestinal region for improving oral absorption [14]. In our opinions, the covalent bonding potential of a ligand with a target may be favorable in dedicated drug research and development (R&D).

Today, it is well-known that single drug targets fail in complicated diseases such as cancer; central nervous system (CNS) diseases, i.e., Parkinson’s and Alzheimer’s diseases; and rheumatoid arthritis (RA) illnesses, and this affects many different cell types or tissues. Drugs with a better balance in safety and effectiveness have preference for regulating multiple targets in contrast to a one-target drug. The success of MTDs in some therapeutic scopes such as infectious diseases, diabetes, and cancer requires efficient hard work (Fig. 1) to identify novel combination of medicines with different indications [12, 16].

Comparing the therapeutics that resulted from MTD and “combinations of single-target drugs,” it appears that MTDs have advantages such as the following:

- Higher efficacy
- Improved safety profile
- Easier administration
- Further expected pharmacokinetics

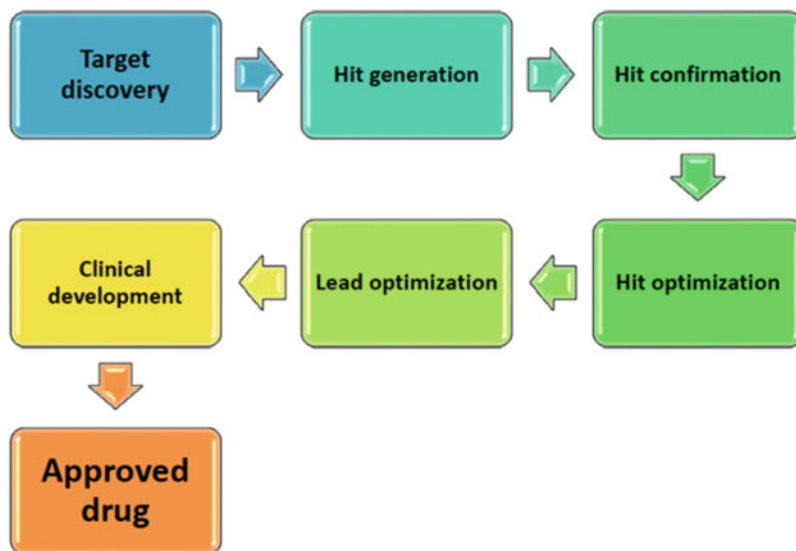


Fig. 1 Drug processing pathway from target discovery to clinical application [15]

- “Lower probabilities of drug interactions” [17]
- “Higher patient compliance” [17]

As mentioned before, various studies in different branches such as medicinal chemistry, computational studies, and systems biology are done in order to find the best drug candidates for pharmaceutical molecule engineering. It is the main task to produce effective treatment regimens. Complicated diseases, like rheumatoid arthritis and cancer, are certainly multifactorial in nature. Hornberg et al. [18] described cancer complexity and difficulty of predicting many molecular processes of cancer therapy. Systems biology is required for complete understanding of relationships among systems behavior with the systems components and all interactions among them [18, 19]. Generally, three principles have led to success of clinical treatment [18]:

- Drug transport, pharmacokinetics, and reaching of a drug to its target site
- Selectivity (related to side effects and drug dose toxicity)
- Efficacy (e.g., the amount of sufficient effect on cancer cells by drug-target attack)

To propose new medicines, various *in silico* methods have been planned in order to ease the analysis of detailed data of target molecules and to offer a selective drug. To our knowledge, ligand- and structure-based CADD methodologies (Fig. 2) are used in a wide range in academia and industry investigations. Virtual drug screening strategies have been applied to recognize sources of

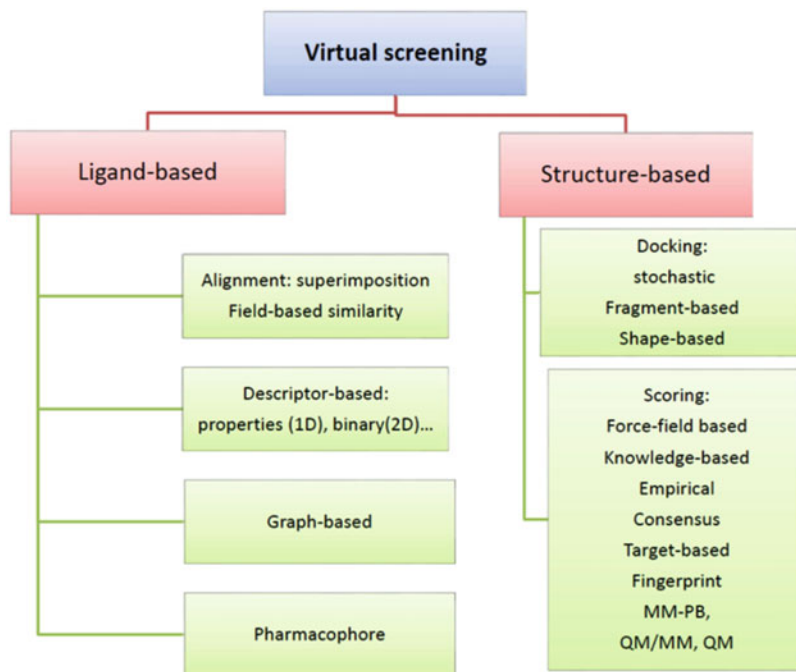


Fig. 2 Different in silico methods for discovery of therapeutics agents [15]

off-target activity profile of drugs and examine their capacity causing desired or adverse side effects [20, 21].

CADD methods are commonly used for making decisions of exploring the bioactive compounds and their fundamental physicochemical properties [15]. They try to know how the bimolecular associates work to justify their functions in a unit, so pharmacokinetic understanding and prediction of four key issues including “absorption, distribution, metabolism, excretion, and toxicity (ADME/Tox)” assist designing of a drug. As a significant result, CADD techniques, namely, SAR, QSAR, docking, and others, have become important tools to remove experimental troubles to understand biological activity networks of drugs and to gain better insight into it [12, 15].

2 Multi-Target Drug Discovery/Design (MTDD) Strategy

As mentioned in the previous section, the progress of studies and efforts to aid the discovery of multi-target ligands for treatment of complex diseases has been observed recently [22]. In silico drug design as a suitable tool has been used to identify hits in the initial step to lead optimization at the final step. Many CADD methods (Fig. 3) have been extended to address different parts of drug design [17, 23]. A number of published reviews have examined in

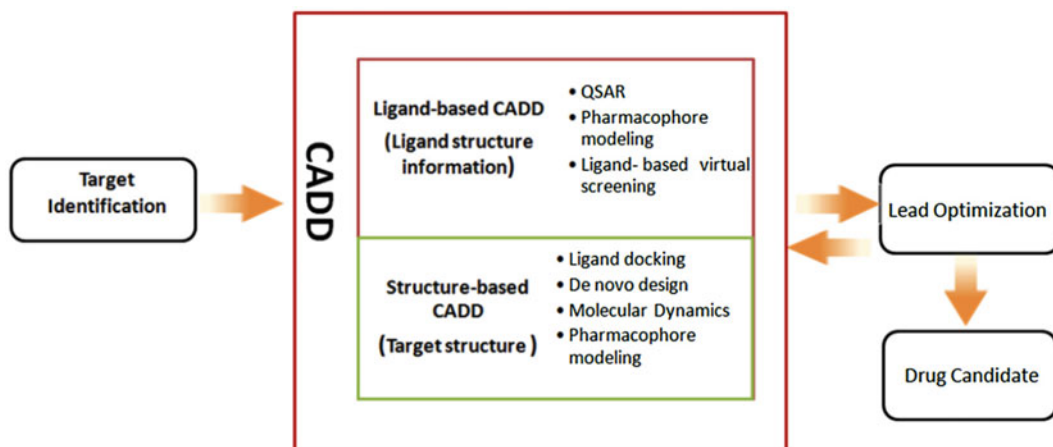


Fig. 3 Different strategies of CADD for drug discovery [12]

silico methods for polypharmacology encompassing structure-based, ligand-based, and data-driven MTDD. Frequently, most of these studies focus on utilization of existing drugs for repurposing (i.e., usage of well-known drugs to discover new targets) [4, 8, 12, 15, 23, 24]. Zhang et al. have classified CADD methods applied to MTD discovery into “target centric” or “ligand centric” [23].

The methods mentioned above could be used to detect interactions among molecular paths that may possibly be leveraged for medical treatment. It is needed to emphasize that combinations of computational data analysis methods with usage of different machine learning approaches give access to maximum information of bimolecular systems.

This chapter presents a summary of important theoretical features behind in silico methods for multi-target drug design (MTDD). It familiarizes the reader with basic concepts of computational processes as a routine tool and a guideline for multi-target drug discovery. The chapter also focuses on the popular procedures including structure-activity relationships (SARs), quantitative structure-activity relationships (QSARs), pharmacophore searches, and molecular docking using 3D receptor structures. Description of each method is followed by some uses in drug design. The text has main sections consisting of mt-SARs, mt-QSAR method, molecular docking, and pharmacophore modeling.

2.1 Multi-Target Structure-Activity Relationships (Mt-SARs)

Lead compound generation is an important phase in the process of drug discovery [25]. Structure-activity relationships (SARs) are well-known in current drug research projects and have been mostly applied for the new lead finding, receptor optimization with respect to physicochemical properties, pharmacokinetics, and scaffold generation [25]. Thus, understanding of this concept is important for practical medicinal chemistry as well as for successful in silico

optimization of a bioactive substance. Once a SAR plan is accessible, it is conceivable to do drug design computationally. In particular, any fruitful use of the systems needs an artificial optimization of a basic feature of the original search space, to be specific the “principle of strong causality” [26]. Maggiora and Johnson renamed this idea in drug design as the similarity rule of chemicals. Similarity in chemistry is a key concept to consider so many aspects of molecular properties, i.e., “biological activities,” “reactivity characteristics,” “structural features,” and so on.

It should be noted that drug design scientists try to find the relationships between the chemical structure and activity (SAR) of the chemicals by data mining practices applied on “molecular fragments.” SAR identifies the molecular structure changes that can increase beneficial drug effectiveness. Generally, a few changes in the lead compound structure for modification continue with practical tests to measure the biological activity variations [27].

By examining many derivatives of one pharmacophore, SARs can be formed to direct other medicinal chemists to create reasonable drug candidates. Regardless of their usefulness, SARs do not connect drug candidates to their active site within a cell [12]. In drug discovery research, it is important to link a drug candidate to its assumed *in vivo* target. Auxiliary information signifying the specificity of this drug-target interaction will aid the SAR principally at branch points where chemical changes may have significantly altered the original pharmacophore. Solving the SAR method restrictions will make a notable progress in the initial step of drug finding [28].

Bio-data collection and organizing based on conventional SARs are a common tactic for lead optimization in the medicinal chemistry toolbox. For molecules with low chemical similarity and chemical scaffold clustering, computations can also be imperative for screening of libraries and analysis of rough data [29]. In lead optimization or hit-to-lead plan, molecular series are commonly explored case-by-case to deduce SAR data and generate actual analogues. In fact, ligands are often active against a number of closely connected targets. This type of cross-reactivity is sometimes favored but is often considered as lacking molecular specificity [30]. Accordingly, if one intends to focus on a single target, one usually attempts to extract compounds with multi-target activity made target-selective through chemical optimization. Consequently, mt-SARs must be considered and evaluated. This task is slightly difficult for analogue design processes performed conventionally. We reviewed previous examples of SAR/QSAR of multi-target drugs using the following descriptors [12]:

- Quantum-chemical descriptors for determining electronic effects
- Molecular quantum similarity (MQS) to determine local reactivity indexes

Computational methods have seldom been used to study mt-SARs due to difficulty of mt-SARs analyses for bioactivity vs. a number of targets in the ligand series. mt-SAR understanding is needed for optimization and designing of a ligand [31]. Activity cliffs considered through characterization of the most important features of activity landscapes play a key task in the inclusive SAR study [32].

The molecular descriptors play an important role in drug design as follows:

- (a) For predicting chemical properties
- (b) For classifying chemical structures
- (c) For searching for similarities between chemical structures

It is thought that with the increase and continuous growth of different descriptors, significant problems in SAR studies would be answered [33]. Selective ligands and associated SAR information are required for framework combination [34].

The term “framework combination” as a knowledge-based lead generation method combines two pharmacophores or frameworks of ligands in one molecule leading to a dual-acting bioactive molecule. Each one of these pharmacophores is selective for a target. SAR information from one-target drug finding projects can be utilized to direct this combination process. MTDs emerging from above concept can be categorized into “linked,” “fused,” or “merged” frameworks. The size of overlapping between two frameworks determines which dual molecules can be generated. Several methods of designing MTDs (Fig. 4) have been reported in our previous review [15].

Generally, a maximum overlapping tries to produce a simpler and smaller analogue with a better chance of bioactivity. Therefore, a high level of similarity of two pharmacophores will require a higher amount of overlap. SAR flexibility is another important term for a medicinal chemist in the field of MTDD, and generally it appears to be valid in some scope to the monoamine-binding targets. It is possibly connected to the presence of basic nitrogen that binds to a key aspartate residue in the target. An anchor set with the contribution of a big quantity of the binding energy may allow larger flexibility in a different place in the molecule [35–37].

Reported SAR samples of the framework combination procedure in the literature indicate wide-ranging information of selective drug SARs allowing the rational designing of MTDs. In this regard, proteomic families like nuclear receptors, GPCRs, transporters, oxidases, and proteases are examples [38].

2.2 Multi-Target Quantitative Structure- Activity Relationships (Mt-QSARs)

A well-known CADD method is QSAR mostly built up by correlating chemical structure information of compounds with response data. Therefore, a standard method of QSAR presents regression/classification equations between a set of molecular descriptors with

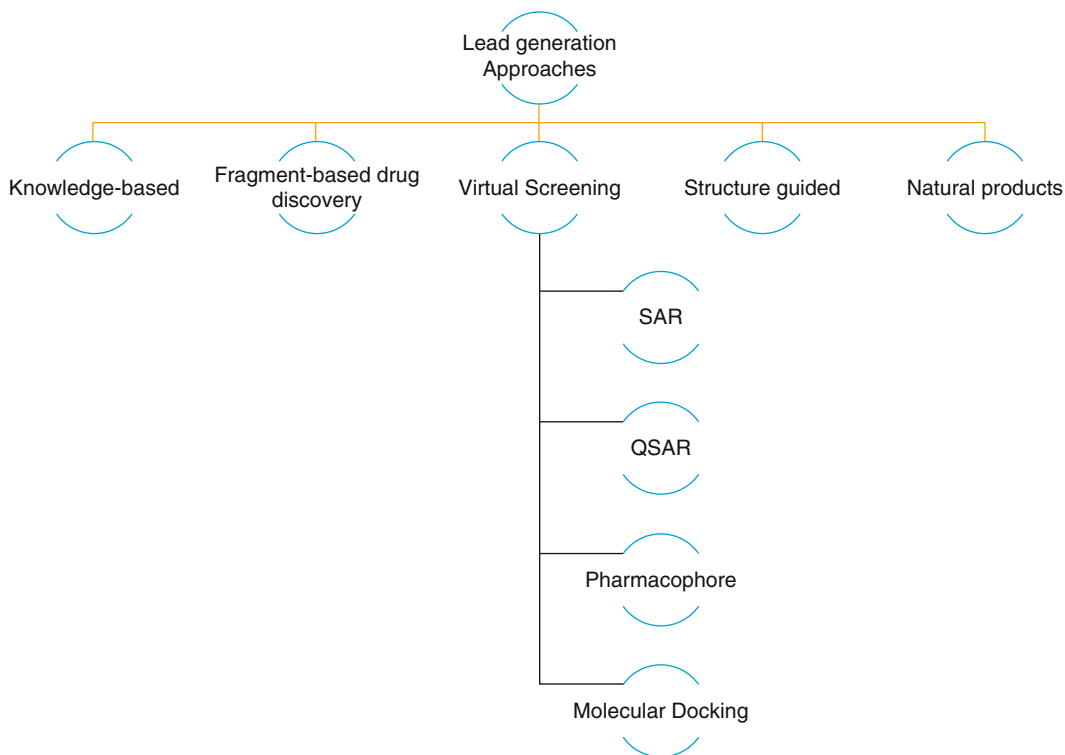


Fig. 4 Lead generation procedures in drug discovery

the experimental values of the properties under study. Several steps of this procedure include:

- (a) Data collection
- (b) Molecular descriptor assembly
- (c) Descriptor selection
- (d) Model building
- (e) Model interpretation
- (f) Model evaluation

Dataset gathering, descriptor generation, and descriptor thinning continue with dividing the data into test and training sets. The model is constructed in a step involving a linear or nonlinear data analysis method. In the end, the model quality, predictability, and its robustness are assessed. Therefore, the QSAR models detect drug leads by guessing the activity on the basis of statistically significant correlation between molecular activities and their structures [12].

Thus, common QSAR models use chemicals acting against one specific target. Consequently, one uses many models against several targets to predict a given activity for a set of chemicals. The method of multi-target or multitasking QSAR (mt-QSAR) is capable of

tackling this limitation by means of making a single equation that relates multiple properties. In summary, a single mt-QSAR model may predict multiple outputs. This is the best characteristic aspect of mt-QSAR. Thus, it is an advantage of the new method for exploring and searching of new drugs, including new leads [39]. We can compare the two methods by comparing the mentioned steps of classical QSAR with an mt-QSAR searching process. The second method involves the steps [39] as follows:

- Data collection (group of chemicals for study and/or analysis, biological property against various targets, parasites, bacteria, etc.)
- Descriptor generation (topological, physicochemical, quantum chemical, etc.)
- Searching for the statistical models and their validation (either through multilinear techniques, such as LDA, MLR, PLS, etc., or nonlinear, such as ANN)
- Model building in search of new bioactives

Generally, an mt-QSAR or multitasking QSAR can lead to predictions of multiple outputs by a single model. Thus, we can predict bioactivity against different microbial species to any drug using a single model [40]. Further description of mt-QSAR details as a flowchart is given in the next section. The mt-QSAR is a relationship between the drugs structure to their activities against different targets.

2.2.1 Molecular Structure Descriptors

Descriptor generation is the initial phase next to the collection of compound structures for QSAR model building. Molecular descriptors quantify physicochemical and structural properties of compounds, their substituents, or electrostatic fields with the correlation to their binding activities. 1D, 2D, and 3D descriptor classes translate chemical composition, topology, 3D shape, and functionality, respectively. Structural indices are used to differentiate structures in order to explain non-covalent molecular interactions, thus relating to properties and activities. There are various chemical descriptors that can be generalized and used to solve the mt-QSAR problems.

At present, there are many marketed and free available software for descriptor calculation [41]. MOE, CODESSA, DRAGON, TOMOCOMD, TOPS-MODE, or MARCH-INSIDE(MI) [42] are the software tools that can be used for descriptor calculation and mt-QSAR or multiplexing QSAR (mx-QSAR) model building. “For the pickup of the best mt-QSAR model, we consider the prior ideas about variable selection method.” “We also consider the rule of parsimony, in which the best model was that with high statistical significance but having few molecular descriptors as possible” [42]. Speck-Planche et al. applied TOPS-MODE and DRAGON

to search mt-QSAR/mx-QSAR models [42]. Alonso and co-workers defined a multiplexing QSAR (mx-QSAR) model for multiplexing tests of anti-Alzheimer drugs based on the MARCH-INSIDE (MI) method [12].

2.2.2 QSAR and Statistical Technique(s) for the Model Building

In 3D-QSAR studies, 3D alignment of chemical structures causing overlapping of the chemical features is required. Other versions (e.g., 4D, 5D, and 6D) integrate new dimensions or new degrees of freedom. Thus, with the added refined analysis on the active site modification of an enzyme to the ligand topology, and vice versa, these methods can be signified well again [33]. There are several reports of 3D-QSAR for single target in combination with molecular docking studies, GRid-INdependent Descriptors (GRIND) [43], CoMFA and CoMSIA descriptors [44, 45], and CoMSIA descriptors for 4D-QSAR [46].

mt-QSAR models relate the drug structure to the bioactivity against different targets [47, 48]. The mt-QSAR permits to compute the probability of activity of a given chemical against different pharmacological or biological targets. It means “a single equation for multiple outputs” [39]. Frequently, chemical compounds have multiple bioactivities “(cf. polypharmacology) that may be interrelated (not to mention multiple physical properties that are frequently the subject of prediction by QSAR approaches).” However, QSAR models are studied for each target property separately, “without utilizing knowledge that can be extracted from QSAR models for other activities of the same compounds. Individual QSAR models of this sort should not be viewed as separate entities but rather as nodes in a network of interrelated models.” This concept is accounted for in an inductive knowledge transfer approach realized in a multitask learning (MTL) and feature net (FN) methods. MTL treats several tasks in parallel and uses a shared representation of data. This can be performed using machine learning methods yielding models with several outputs, such as neural networks, PLS, or SVM with special kernels. FN treats different tasks sequentially when predictions made by previously developed models are used as descriptors for the main task [49]. Figure 5 shows a typical flowchart of processing of multi-target QSAR modeling.

Nowadays data mining methods are used for bioactivity prediction of substances. These methods find drug leads with the statistical analysis of basic correlations between activity status and molecular structures; regression methods can be integrated for estimating levels of activity [50]. Compared to chemical similarity methods, these methods do pattern recognition in vast cluster compounds and chemogenomic space that deliver similarities in multidimensional space. Some of the most popular machine learning methods [50, 51] include:

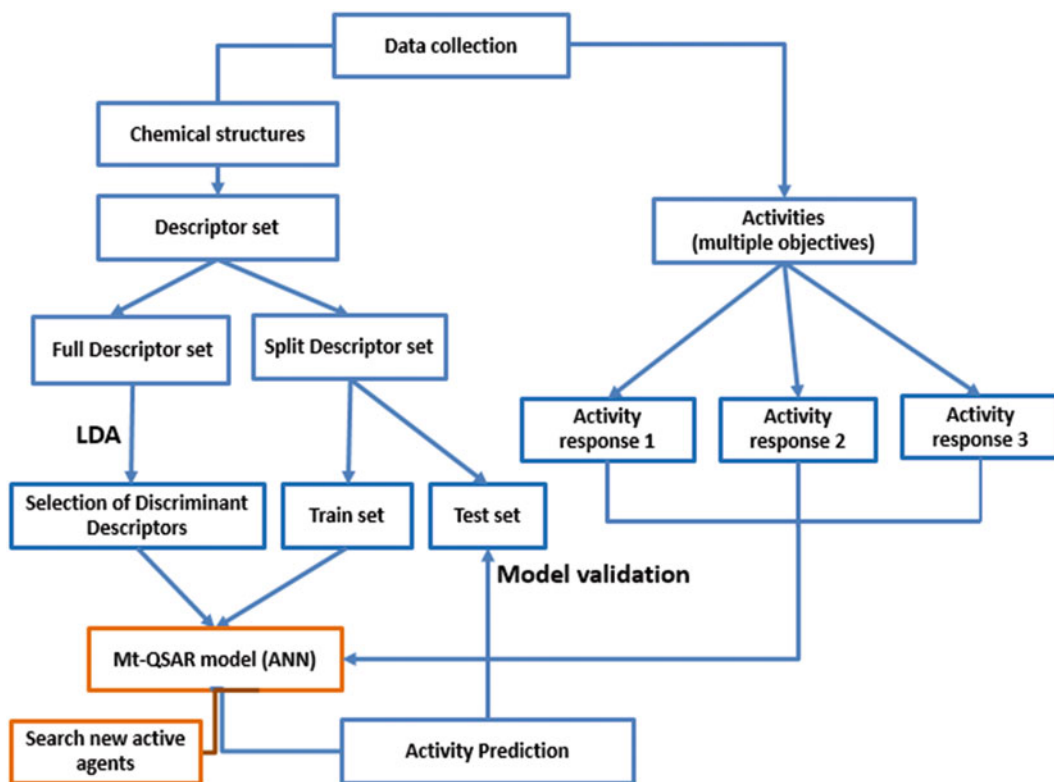


Fig. 5 Typical flowchart of multi-target QSAR methods

- Linear discriminant analysis (LDA)
- Logistic regression (LR)
- Binary kernel discrimination
- k-nearest neighbor (kNN)
- Decision trees (DT)
- Bayesian methods
- Random forests (RF)
- Support vector machines (SVM)
- Artificial neural networks (ANN)

It is easy to apply linear and nonlinear statistical tools such as LDA or ANN to create mathematical models enabling a good prediction of different pharmacological and biological properties of compounds, such as antifungal or antiparasitic activity and bacterial susceptibility, just to mention a few. It is to be expected that there will be an important spread of this sort of methods in the near future given their great performance and low computational cost [39]. In this regard, we can cite examples of mt-QSAR study reported by Cruz-Monteaquedo in a multi-objective optimization

(MOOP) of desired activity or properties of drugs against different targets [52, 53]. MOOP technique can find the global optimal solution by simultaneous optimization of numerous dependent properties [54]. Several instances of mt-QSAR model for bacterial, fungal, and viral species using ANN and linear discriminate analysis (LDA) methods exist in recent published papers [12]. Zanni et al. have reviewed many mt-QSAR published research considering the above methods [39]. The performance of these practices depends on numerous factors [50] including the following:

- Training set variation, their capacity to deal with imbalanced datasets
- Applicability domain
- Parameter ranges in covering active and inactive chemical space

It should be noted that after mt-QSAR model building, one can use it for the search of new active agents (screening, computational design, etc.).

Some researchers have applied the QSAR theories to construct models for MTDD. Ajmani and Kulkarni used a new process of group-based QSAR (G-QSAR) for lead optimization of multi-kinase (PDGFR-beta, FGFR-1, and SRC) and scaffold hopping of multi-serotonin target (serotonin receptor 1A and serotonin transporter) inhibitors. Also, the established G-QSAR models propose key “fragment based features that can produce the building blocks to conduct combinatorial library design for exploring of the optimally potent multi-target inhibitors” [55].

Namasivayam et al. studied emerging chemical patterns (ECP) for classification “with overlapping or distinct activities of multi-target and different mechanisms for molecular interactions” [56]. Also, in another work, the ECP process was used by above researchers to discover molecular descriptors and value ranges. This processing distinguishes between “modeled conformations and experimental bioactive conformations of ligands experimentally” [57]. “For small training sets, the ECP process does better than decision trees and Bayesian QSAR methods” [58].

Prado-Prado et al. reported mt-QSARs using Markov model for encoding molecular backbone information. The term “MARCH-INSIDE, MARKovian CHEmicals IN Silico DEsign” refers to the mentioned process [59]. They developed mt-QSAR models for active agents against multiple fungal [60, 61], bacterial [40], and multi-target agent identification [62]. The model development was performed by joining multi-target/species changes of binding site specifications into the species-dependent descriptors or mt-dependent descriptors and stochastic Markov drug-binding process models. They computed new multi-target spectral moments to meet a QSAR model using Markov chain theory for active drugs against 40 viral species. In this way, they can present

“matrix invariants such as spectral moments, stochastic entropies, and potentials for molecular properties study” [63, 64]. The same author group in another mt-QSAR study applied Markov chain theory for determining a new mt-spectral moments to find a QSAR model which estimates activities by an mt-QSAR model for 500 drugs against 16 parasite species and other 207 drug-like compounds. “They used LDA for model development and classification of drugs as actives or non-actives against various tested parasite species. They derived four types of nonlinear artificial neural networks (ANN) which were evaluated as the mt-QSAR models. The total performance for the training set was 87% for the improved ANN model” [65].

Cruz-Monteagudo et al. have reported mt-QSAR studies using a multi-objective optimization (MOOP) of activity or properties of drugs against diverse targets. An effective methodology of the MOOP problem uses the “Derringer’s desirability function and several QSAR models for different objectives.” The MOOP process as a chemometric method is a useful tool that obtains the global optimal solution by “simultaneous optimization of numerous dependent properties” [54]. Various applications of MOOP techniques can be found “ranging from substructure mining” to molecular docking mainly on the basis of “weighted-sum-of-objective-functions (WSOF)” and Pareto optimality methods.” The above strategy considers the full spectra of objectives in order to avoid local optima. This aspect is a key advantage that leads to a more efficiency [12, 40]. Different desirability functions may be used to identify the response as maximized, minimized, or given a target value. Derringer and Suich presented a desirability function that can be useful in the field of drug discovery [62, 66]. Derringer’s desirability function is a tactic of the different techniques of multi-criteria decision-making [54]. The MOOP-DESIRE method on the basis of Derringer’s desirability function enables researchers of developing a global QSAR model. Thus, such QSAR models consider the pharmacokinetic, pharmacological, and toxicological profiles of a series of candidate molecules simultaneously [52, 67].

2.3 De Novo Method

De novo method is an MTDD plan with different target profiles. Virtual screening techniques need pre-existing ligand libraries, but de novo design procedures can produce a pool of drug candidates from simple building blocks (i.e., atoms, fragments) to obtain a focused library [23].

CADD-based de novo design offers a complementary methodology through screening virtual ligands. It greatly increases the size of the pool of molecules that are freely available for HTS (usually 1–3 million molecules). Here, the practice of molecule synthesis and testing must be performed by smart algorithms. Fragment-based virtual synthesis methods have been verified for their

particular applicability to this task [68, 69]. Two problems for de novo designing are:

- The assembly problem
- The scoring problem

Fragment-based methods provide a way to tackle the assembly problem and two plans for scoring problem through in situ design or disregarding receptor structure [70].

Practical or in silico fragment-based search could be applied to discover the primary starting point [71–73]. Moreover, it provides evidences for employing “ensemble linking” plans and supporting the efficiency of the “fragment linking” algorithms that link fragments occupying different subpockets. With the progress of automated robotic chemical synthesis and its pairing to de novo design software, designed ligands can be quickly produced and examined [74, 75]. Artificial intelligence methods such as deep learning have been used to predict ADMET properties, and more properties related to late-phase growth can be unified in the de novo design practice [76, 77].

In silico screening has been extended further along the biological dimension leading to new integrative methods capable of estimating the pharmacological profile of molecules on multiple targets [78]. Screening of compound libraries and virtual screening were established to more rational methods, such as de novo design with insight from liganded or unliganded protein structures, database mining employing pharmacophore models derived from known compounds, or based on mechanistic insight.

2.4 Molecular Docking

Molecular docking has become an interesting keyword in the computational multi-target drug design development. Molecular docking is in silico technique that can be used to model the interaction between a ligand and a target protein at the atomic level. Docking is widely applied to the study of biomolecular interactions and mechanisms, and it is used in structure-based drug design and virtual ligand screening methodologies. The aim of docking is to predict the best binding pose of a ligand to fit the binding site of a protein and evaluate its binding affinity using a scoring function.

Many available docking software tools have revealed their impressive growth, and the accuracy of each tool has been regularly improved by introducing new techniques. A docking software can generate correct binding poses in many cases if a reasonable template structure of a protein is chosen. The challenge is how to efficiently select a correct pose among the putative candidate poses generated by the software. There are several cavity detection programs or online servers for *apo* structure that can detect putative active sites within proteins in blind docking. Blind docking was introduced for the detection of possible binding sites and modes of

peptide ligands by scanning the entire surface of protein targets. For example, in the AutoDock program, when we do not know where the ligand binds, we can build a grid volume that is big enough to cover the entire surface of the protein using a larger grid spacing than the default value and more grid points in each dimension. Then, we can perform preliminary docking experiments to see if there are particular regions of the protein preferred by the ligand. This is sometimes referred to as blind docking, cavity detection, or pocket detection [79].

Docking programs have been developed based on different aspects as follows: open source or commercial [80], scoring [81, 82], posing [83–85], sampling [86, 87], optimization [88–90], and flexibility [91, 92]. Docking and cavity detection software tools developed in the last decade are listed in Table 1.

2.4.1 Receptor Representation in Docking

The traditional docking algorithm represents the receptor as a rigid body, leading to predict an incorrect binding pose resulting in meaningless ligand-binding scores. Different techniques for receptor flexibility have been developed. In some methods, receptor flexibility occurs by enumerating its conformations or modeling its changes (e.g., conformational searching and/or optimization approaches) during docking. The ensemble-based techniques depend on using multiple input receptor conformations into docking programs. It is noteworthy that using an ensemble approach is superior to a single-receptor conformation input.

GalaxyDock [90, 119] (Table 1) explains protein flexibility of residues selected previously in the receptor binding site by using global optimization, and we can make a favorable comparison by applying a wide range of programs (FLIPDock, AutoDock, Rosetta Ligand, and SCARE) in binding pose prediction accuracy with success rates of 80–87%. By doing so, it is possible to measure success rate as the fraction of docked poses with predicted rmsd of ≤ 2 Å compared with the experimental structures.

Consensus Induced Fit Docking (cIFD) [131] method as a way for the adaption of receptor binding site in order to accommodate multiple diverse ligands is a vital characteristic for virtual screening (VS) introduced by Kalid et al. The cIFD workflow undergoes two stages: (1) IFD of multiple ligands for preliminary binding mode determination and (2) receptor optimization in the presence of a “hybrid” ligand which combines selected poses of the IFD-docked ligands. cIFD was validated on three targets demonstrated before to be challenging for docking. A practical way is recommended to explain receptor flexibility covertly to avoid the additional computational costs of ensemble docking or accounting for full receptor flexibility during docking.

Ensemble docking which includes docking compounds into many conformations of a target receptor is a robust way to show receptor flexibility in a separate format. This approach copies the

Table 1
Software and performance details for molecular docking

Program name	Descriptions	Ref.
AutoDock Vina	It is an entirely separate code base and approach from AutoDock that was developed with a focus on runtime performance and ease of system setup. It uses a fully empirical scoring function and an iterated local search global optimizer to produce docked poses. It includes support for multi-threading and flexible side chains	[93]
APBS	It performs solvation free-energy calculations using the Poisson-Boltzmann implicit solvent method	[94]
AutoDock	It is an automated docking program that uses a physics-based semiempirical scoring function mapped to atom-type grids to evaluate poses and a genetic algorithm to explore the conformational space. It includes the ability to incorporate side-chain flexibility and covalent docking	[95]
Clusterizer-DockAccessor	They are tools for accessing the quality of docking protocols. It interfaces with a number of open source and free tools	[96]
DockoMatic	It provides a graphical user interface for setting up and analyzing AutoDock and AutoDock Vina docking jobs, including when to run on a cluster. It also includes the ability to run inverse virtual screens (find proteins that bind a given ligand) and support for homology model construction	[97]
DOVIS	It is an extension of AutoDock 4.0 that provides more efficient parallelization of large virtual screening jobs	[98]
Idock	It is a multi-threaded docking program that includes support for the AutoDock Vina scoring function and a random forest scoring function. It can output per-atom free-energy information for hotspot detection	[99]
MOLA	It is a prepackaged distribution of AutoDock and AutoDock Vina for deployment on multi-platform computing clusters	[100]
NNScore	It uses a neural network model to score protein-ligand poses	[101]
Paradocks	It is a parallelized docking program that includes a number of population-based metaheuristics, such as particle swarm optimization, for exploring the space of potential poses	[102]
PyRx	It is a visual interface for AutoDock and AutoDock Vina that simplifies setting up and analyzing docking workflows. Its future as an open-source solution is in doubt due to a recent shift to commercialization	[103]
rDock	It is designed for docking against proteins or nucleic acids and can incorporate user-specified constraints. It uses an empirical scoring function that includes solvent-accessible surface area terms. A combination of genetic algorithms, Monte Carlo, and simplex minimization is used to explore the conformational space. Distinct scoring functions are provided for docking to proteins and nucleic acids	[104]
RF-Score	It uses a random forest classifier to score protein-ligand poses	[105]

(continued)

Table 1
(continued)

Program name	Descriptions	Ref.
smina	It is a fork of AutoDock Vina designed to better support energy minimization and custom scoring function development (scoring function terms and atom-type properties can be specified using a runtime configuration file). It also simplifies the process of setting up a docking run with flexible side chains	[106]
VHELIBS	It assists the non-crystallographer in validating ligand geometries with respect to electron density maps	[107]
VinaLC	It is a fork of AutoDock Vina designed to run on a cluster of multiprocessor machines	[108]
VinaMPI	It is a wrapper for AutoDock Vina that uses Open MPI to run large-scale virtual screens on a computing cluster	[109]
Zodiac	It is a visual interface for structure-based drug design that includes support for haptic feedback	[110]
BetaDock	Prioritizing shape complementarity, based on the theory of b-complex and the Voronoi diagram, rigid bodies (both receptor and ligand)	[84]
bhDock	Identification of low-resolution binding sites, ligand posing via simulated annealing MD-based global optimization	[111]
EADock DSS	Efficient tree-based dihedral space sampling (DSS), based on hybrid sampling engine and multiobjective scoring	[87]
GENIUS	Binding constraints via essential interaction pairs (EIP)	[83]
H-DOCK	Docking by hydrogen bond matching and shape complementarity	[112]
LigDockCSA	Powerful global optimizer, conformational space annealing (CSA)	[89]
MedusaDock	Modeling both ligand and receptor flexibility simultaneously with sets of discrete rotamers	[92]
NeuroDock	Generation of docked poses by self-organization of atom coordinates without the need for input/seed conformation	[85]
ParaDockS	Open source and operating system-independent code	[102]
PythDock	A simple scoring function, a population-based search engine (particle-swarm optimizer algorithm)	[81]
SKATE	Decoupling sampling from scoring	[86]
VoteDock	Discriminant analysis, adaptive and radial sampling and clustering, consensus scoring	[81]
RigiDock + PoseMatch	A fast approximation scheme for the docking of rigid fragments	[113]
PD DOCK	Parameter optimization for docking scores	[114]
BSP-SLIM	Blind docking using low-resolution receptor structures, such as homology models	[115]
CovalentDock	Capacity to handle the molecular geometry constraints of the covalent bonding using special atom types and directional grid maps	[116]

(continued)

Table 1
(continued)

Program name	Descriptions	Ref.
CRDOCK	Increased speed and efficient treatment of ligand flexibility, by pre-generating the conformational libraries of ligands	[117]
FIPSDock	Implementation of the fully informed particle swarm optimization	[118]
GalaxyDock	Flexibility of preselected receptor side chains by global optimization of an AutoDock-based energy function trained for flexible side-chain docking	[90]
GalaxyDock2	Geometry-based generation of initial ligand conformations in conformational space annealing global optimization	[119]
HYBRID	Utilization of the knowledge of bound ligands	[120]
LiGenDock	Capacity to conduct de novo design	[121]
PRL-Dock	Increased speed, probabilistic relaxation labelling algorithm to search for potential intermolecular hydrogen bonding	[122]
QuickVina	Increased speed, improvements in the local search algorithm via heuristic prevention of some intermediate points undergoing local search	[123]
SAMPLE	It is based on the hybrid genetic algorithm, ability to handle intramolecular and intermolecular degrees of freedom and an arbitrary number of independent species	[124]
eFindSite	It is using homology modeling and machine learning predicts ligand-binding sites in a protein structure	[125]
fpocket	It detects and delineates protein cavities using Voronoi tessellation and is able to process molecular dynamics simulations	[126]
KVFinder	It is a PyMOL plug-in for identifying and characterizing pockets	[127]
McVol	It calculates protein volumes and identifies cavities using a Monte Carlo algorithm	[128]
PocketPicker	It is a PyMOL plug-in that automatically identifies potential ligand-binding sites using a grid-based shape descriptor	[129]
POVME	POcket volume MEAsurer is a tool for measuring and characterizing pocket volumes that includes a graphical user interface	[130]

dynamic behavior of the protein. As a consequence, the structural degree of freedom is attained where each ligand may find a matching receptor. Practically speaking, the latter depends on the conformational space covered by the ensemble. It is generally assumed that ensemble docking is superior to docking into a single-receptor conformation.

This hypothesis was examined by Craig et al. through probing VS enrichment using an ensemble of crystallographically derived

structures [132]. They tested their approach using Glide in SP mode with GlideScore against the aspartic protease β -secretase and the cAbl kinase domain. They employed an area under the receiver operating characteristics (ROC) curve (AUC) as an enrichment metric. The results indicate that crystallographically derived ensembles in comparison with all individual members produced better enrichment only in some cases. More frequently, the ensemble gave better results compared with a mean of the enrichments of the individual members. Thus, following a range of caveats and suggestions for further investigations (more test systems, more docking programs/scoring functions/ensemble construction strategies, and inclusion of receptor conformation energy into a scoring function), the authors concluded that it may be incorrect to assume that ensemble docking is always able to account reliably for protein flexibility. On one hand, this conclusion, though very cautious, seems to be quite uncommon; on the other hand, Craig and co-workers were not alone in questioning and finding faults in the ensemble-based treatment of receptor flexibility.

Several key features influencing ensemble docking with regard to pose prediction and VS performance were suggested by Korb et al. [133]. The features are as follows: sampling accuracy, choice of the scoring function, and the similarity of docked ligands to the ligands bound to the protein structures in an ensemble. In addition, they thoroughly assessed the ensemble performance compared with the performance of the individual ensemble members, and the following observations were made: (1) in almost all cases, ensembles had a better performance than the worst single structure; (2) in many cases ensembles performed better than the average single protein structure; and (3) in some cases ensembles performed better than the best single protein structure. Considering these findings, they determined that the rational prospective selection of optimum ensembles is a demanding task requiring further critical research. The most significant point to mention is that protocols are required to generate ensembles, in terms of both size and membership, leading to increased docking efficiency and reduced false-positive rate in VS.

The issue of rational ensemble construction was dealt with by Xu and Lill through considering three potential selection strategies, namely, clustering based on pairwise rmsd, pose prediction performance, and VS performance in terms of actives/decoys differentiation [134]. They used VS performance which appeared to be the most successful method for the selection criterion. Therefore, using the earlier developed Limoc concept, they were able to obtain a balance between the extent of protein flexibility accounted for and the risk of false positives resulting from the excessive ensemble size. Specifically, successful docking was shown to be performed using ensembles of relatively small size (3–5 receptor structures).

In case extensive and diverse crystallographic data are not accessible for a given target, simulation-based ensembles can be used. A widely agreed-upon viewpoint is that a greater sampling of biomolecular space is desirable for the construction of the optimal ensemble. The main obstacles to such coverage are the small configurational changes usually observed in relaxed complex schemes and induced fit methods, and, conversely, short time scales are generally accessible to molecular dynamics (MD) simulations [135]. Approaches employed to surpass the obstacles include the brute force hardware advances and algorithmic improvements. Conceptually, the simplest brute force improvements to the problem of configurational space coverage come from increasing speed and efficiency of computer hardware. The latest advances in this field come from the implementation of graphics processing unit (GPU)-accelerated computing [136], building special purpose supercomputers such as Anton, or moving calculations into the cloud [137]. More conceptually demanding are the advances in algorithms, allowing more efficient space coverage. Enhanced sampling methods such as temperature-accelerated replica exchange, Hamiltonian-based accelerated MD, umbrella sampling, metadynamics, and Markov state models can accelerate calculations by developing artificial biases into simulations.

Combining docking with molecular dynamics (MD) is exclusively useful in cases where binding causes significant conformational changes. Dynamic docking's predictive power in estimating the binding free energy can be assessed against experimental results, when available. The key step is thus to develop dynamic docking methods and techniques for correctly computing the binding free energy of a ligand to its target. This task is at the very core of challenge in developing new and better computational tools for drug discovery [138]. Whalen et al. developed the flexible enzyme receptor method by steered molecular dynamics [139]. They present a novel hybrid method for accurate and precise affinity rank ordering of ligands against a challenging enzyme drug target. This hybrid method combines ensemble docking via AutoDock 4, computationally efficient steered MD by YASARA 9.11.9, and binding affinity calculation by MM/PBSA. They demonstrated an improved accuracy in affinity predictions, considered poor when delivered by traditional scoring functions even when receptor flexibility is taken into account. Multi-targeted molecular dynamics (MTMD) is a unique technique in which important intermediate catalytic structures could be sampled out. MTMD studies of stable binding poses of substrates and inhibitors were further carried out to analyze the conformational changes in the residues forming stable binding interactions. A report was published using the MTMD techniques by Prajapati and Sangamwar [140]. They presented the mechanistic picture of translocation of P-glycoprotein as

a well-known multidrug resistance in drug therapy; MTMD studies gave sufficient details into conformational changes in drug-binding region, role of intracellular coupling helices and topological arrangements of TM helices in P-glycoprotein translocation, and quantitative shifts in TM helices providing valuable information for future experiments.

2.4.2 Ligand Representation in Docking

Ligand representation in docking, especially, conformation selection and specifications, has significant effects on docking results. Feher and Williams in their studies have shown that input ligand conformations alter the docking outcomes [135]. They tested this variability using GOLD, Glide, FlexX, and Surflex programs. It is concluded that there are two major effects leading to such variability: the adequacy of conformational search during docking (major) and random “chaotic” effects arising from sensitivity to small input perturbations (minor, but significant). Ligand flexibility is one of the main obstacles in failure of docking protocols to correctly predict the pose. Bohari and Sastry tested five programs (Glide, GOLD, FlexX, CDOCKER, and LigandFit) on a dataset of 199 FDA-approved drug-target complexes [141]. Even given the limited range of ligand conformational complexity in such dataset, they noted the dependence of pose prediction accuracy on ligand “size” and observed better performance for low or medium flexibility. Similarly, using docking with AutoDock 4 and AutoDock Vina, Houston and Walkinshaw suggested that there appears to be a certain ligand size that maximizes pose prediction accuracy because of optimum flexibility [142]. Using these programs [143, 144] and some others can result in increased failure rates while docking small and fragment-like molecules. This is generally accepted to be a scoring failure. Conversely, failing to correctly dock large, often highly flexible molecules is ascribed to be a shortcoming of sampling.

2.4.3 Scoring Functions

Scoring functions accounting for the potential of a compound to act as a ligand are mainly derived from three main approaches based on their characteristics.

Scoring functions in molecular docking are fast approximate mathematical methods to predict the strength of the non-covalent interactions (also referred to as binding affinity) between two molecules after having been docked. Ranking of compounds by their probability of being active (virtual screening) presents a different problem to predict the bound pose of a compound (pose prediction). Therefore, different scoring functions perform better for one or the other of these problems. Furthermore, scoring functions cannot predict binding affinity or binding free energy for two reasons: They calculate mostly enthalpic terms disregarding entropy particularly of the protein. Entropy is explicitly required to

estimate binding free energy. Scoring functions only know about the bound state of the protein-ligand system, not the unbound states of the protein and the ligand. Binding free energy can only be estimated using knowledge of the bound state and the unbound states of the binding partners.

Popular scoring functions are roughly grouped into four main categories: force field, empirical, knowledge-based, and machine learning scoring functions.

Force field scoring functions: Binding affinities are estimated by summing the strength of intermolecular van der Waals and electrostatic interactions between all atoms of the two molecules in the complex using a force field.

Empirical scoring functions: They are based on counting the number of various types of interactions between the two binding partners.

Knowledge-based scoring functions: They are based on a statistical analysis of protein-ligand complex structures.

Machine learning scoring functions: The functional form is inferred directly from the data. Machine learning scoring functions have consistently been found to outperform classical scoring functions at binding affinity prediction of diverse protein-ligand complexes.

Consensus scoring: Given the performance inadequacies of established and newly developed scoring functions, a common approach to attempt improving their accuracy is using scoring functions in combination (i.e., consensus scoring). The consensus scoring approach based on multiple scoring functions seems to perform better than a single scoring function.

The ranking mechanism which can correctly identify such modes or effectively distinguish between binders and nonbinders or active and inactive compounds still remains a challenge. To tackle this problem, many advances to scoring functions have been attempted by addressing the main obstacles to robust scoring: entropy and desolvation effects.

In an investigation by Huang and Zou, simple empirical terms were added, accounting for ligand configurational entropy to their knowledge-based scoring function ITScore to produce ITScore/SE [145]. ITScore/SE achieved a significant improvement in its performance compared with ITScore as well as 14 other scoring functions.

On the interface of biomolecular complexes, one factor that facilitates the increase in specificity and/or affinity and may enable promiscuous binding is water molecules. Thus, properly accounting for specific water molecules on the interface as well as for the general effect of solvation is an important aspect of docking [144].

The choice of the best docking program for a specific protein receptor could be a complicated issue. However, in order to make a

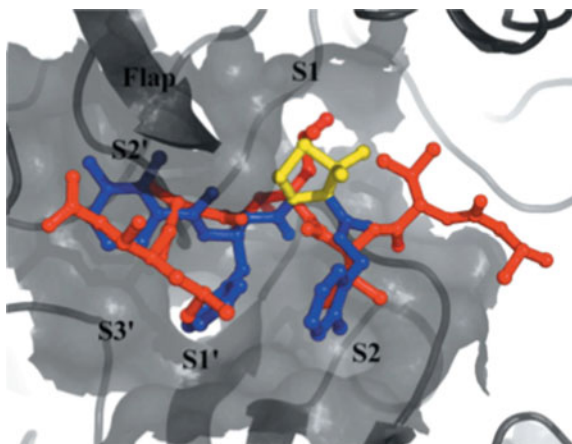


Fig. 6 Molecular docking strategy of multi-target inhibitor discovery [147]

possible regression between scoring function and bioactivity, the evaluation criterion is to be on the basis of the best regression of scoring function. The docking program has the potential to calculate only the binding affinity; hence, it cannot be argued that a good dock score guarantees a good inhibitor or an agonist since the real mechanism is both complicated and unknown. There is one thing we can say for sure as the concluding point, which is “the ligand binds well into the binding site of the protein.” After bioassay validations, it can be judged whether it is an inhibitor or an agonist [146].

Molecular docking has been applied in virtual screening against the individual targets in HIV and its associated opportunistic pathogens to find multi-target agents such as KNI-764 that inhibit both HIV-1 protease and malarial plasmeprin II enzyme (Fig. 6) [147].

In multi-targeted docking approach (mt-docking), the data obtained from a series of ligands tested on multiple targets of a specific disease are used to create a network; the extracted data provide information about the number and quality of the compound's interactions, helping to understand mechanism. Through degree centrality parameters, it becomes possible to characterize the importance of the nodes in a network (by molecule or target protein). A score-weighted docking prediction model reveals information concerning all of the bindings between a ligand and a receptor [148]. An investigation by Azam and collaborators [149] suggested 13 compounds from ginger (*Zingiber officinale*) against Alzheimer drug targets: acetylcholinesterase (AChE), butyrylcholinesterase (BuChE), β -site amyloid precursor protein-cleaving enzyme (BACE), glycogen-synthase-kinase-3 β (GSK), TNF- α converting enzyme (TACE), c-Jun N-terminal kinase (JNK), nitric oxide synthase (NOS), human carboxylesterase, *N*-methyl-D-

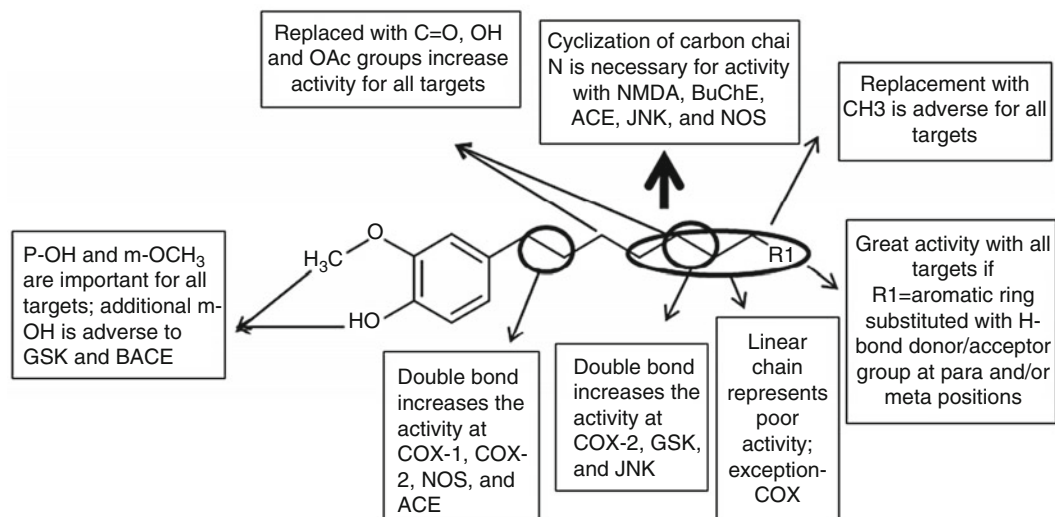


Fig. 7 Structural requirements for ginger compounds to interact with anti-Alzheimer targets [84]

aspartate (NMDA), cyclooxygenase-I (COX1), cyclooxygenase-II (COX2), phosphodiesterase-5, and the angiotensin-converting enzyme. They employed a rigid protein and a flexible ligand whose torsion angles were identified (for ten independent runs per ligand). The calculations, reliability, and reproducibility of the molecular docking methodology were validated; docking parameters showed a correlation coefficient of $r^2 = 0.931$. From the docking data, the authors built a “ginger” model from structural requirements and interactions with the receptors (Fig. 7). Multifunctional drug investigations report on the different binder classes and their various targets in many diseases, from respiratory infections to cancer, simultaneously displaying the great breadth of such research, as well as the immense applicability of multi-target drugs.

2.5 Pharmacophore Modeling

Ehrlich in 1909 defined pharmacophore as a molecular framework that carries (phoros) the essential features responsible for a drug’s (pharmacon) biological activity [150]. After a century, in 1998 Camille Wermuth submitted a refined definition to the IUPAC as follows: “A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supra-molecular interactions with a specific biological target structure and to trigger (or to block) its biological response” [151].

A pharmacophore model can be derived by either ligand-based or structure-based methods. Ligand-based methods generate pharmacophoric features by superposing a set of active molecules and subsequently extracting their common chemical features essential for binding activity. Structure-based methods construct pharmacophoric features by probing possible interaction points between target and ligands.

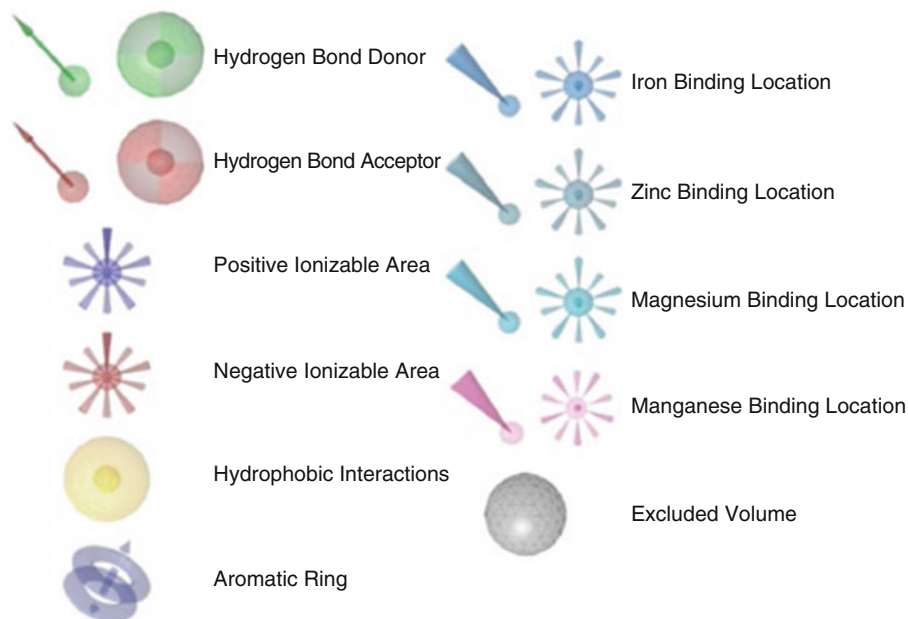


Fig. 8 Pharmacophore depictions in LigandScout. *Hydrogen-bonding interactions*: Hydrogen bonding is an attractive interaction of electropositive hydrogen atoms with an electronegative atom (H-bond acceptor) like oxygen, fluorine, or nitrogen. *Hydrophobic interactions*: Hydrophobic (lipophilic) interactions occur when nonpolar amino acid side chains in the protein come into close contact with lipophilic groups of the ligand. *Aromatic and cation- π interactions*: Electron-rich π -systems like aromatic rings are capable of forming strong attractive interactions with other π -systems (π -stacking) and adjacent cationic groups. *Ionic interactions*: Ionic interactions are strong attractive interactions (energies >400 kJ/mol) that occur between oppositely charged groups of the ligand and the protein environment [152]

Pharmacophore models must in some way uniformly represent the physicochemical properties and location of functional groups involved in ligand-target interactions. The most common representation of pharmacophores is a spatial arrangement of so-called chemical (or pharmacophoric) features that describe essential structural elements and/or observed ligand-receptor interactions by means of geometric entities.

Most common properties used to define pharmacophores are hydrogen bond acceptors, hydrogen bond donors, basic groups, acidic groups, partial charge, aliphatic hydrophobic moieties, and aromatic hydrophobic moieties [152].

For instance, LigandScout supports the derivation of 14 feature types whose graphical representation is shown in Fig. 8.

Pharmacophore features have been used extensively in drug discovery for virtual screening, de novo design, and lead optimization [153, 154].

Since pharmacophore-based queries are based on pharmacophore feature alignments rather than compound scaffold alignments, they are able to find hit molecules with diverse scaffolds

when compared to the original ligands used for the generating the pharmacophore models [155]. This is of special interest for researchers who need to find novel molecules that are patentable or lead candidates with better ADMET properties, and/or higher activity, and/or selectivity toward the target [156].

Several computational tools and programs which have established the pharmacophoric geometry for automated pharmacophore generators such as HipHop, HypoGen (Accelrys Inc.), DISCO, GASP, GALAHAD, PHASE, LigandScout, MOE, GBPM, Schrödinger, Pocket v.2, HS-Pharm, and Snooker have been developed. The main difference between these programs is in the algorithms applied for managing the flexibility of ligands and for the alignment of compounds [154].

Ligand- and structure-based pharmacophore models and molecular docking are commonly used in single-target drug discovery. Combined pharmacophore and molecular docking strategy named as combinatorial approaches were applied to discover multi-target inhibitors. Parallel searches against each individual target are directly conducted in these methods to find virtual hits that simultaneously interact with multiple targets. Based on the chemical structures of known ligands or the 3D structure of the binding site, a pharmacophore model with several key features can be built for each target. Multiple conformations of virtual ligands generated and mapped onto pharmacophore model and fitness are evaluated [23].

For docking-based methods, the ligands are put into the binding site and then evaluated by scoring functions. Figures 9 and 10 show the strategy and framework of these approaches. This approach is mentioned due to selection of common hits involving various computational screening steps that are computationally expensive and could present multiple challenges such as comparing binding scores or fitness scores for different targets.

Recently, an investigation by Zhang et al. suggested a common pharmacophore-guided multi-target drug design technique to overcome these limitations based on the findings that the number of bound conformations of a ligand even for different targets is limited [158, 159]. For systems where a ligand binds all targets having common features, this approach can be applied. As can be seen in Fig. 11, the pharmacophore model for each target can be first constructed from 3D structures of the targets or available potential ligands. Then, a common feature pharmacophore model is generated by aligning all combinations of possible pharmacophore. This technique can be applied to rapidly screen matching compounds. A molecular docking algorithm is performed for the top-ranking compounds to identify compounds that bind well to all targets in the initial screen. The common pharmacophore can also be used as a post-filter technique after multiple docking screening to discover compounds that may bind

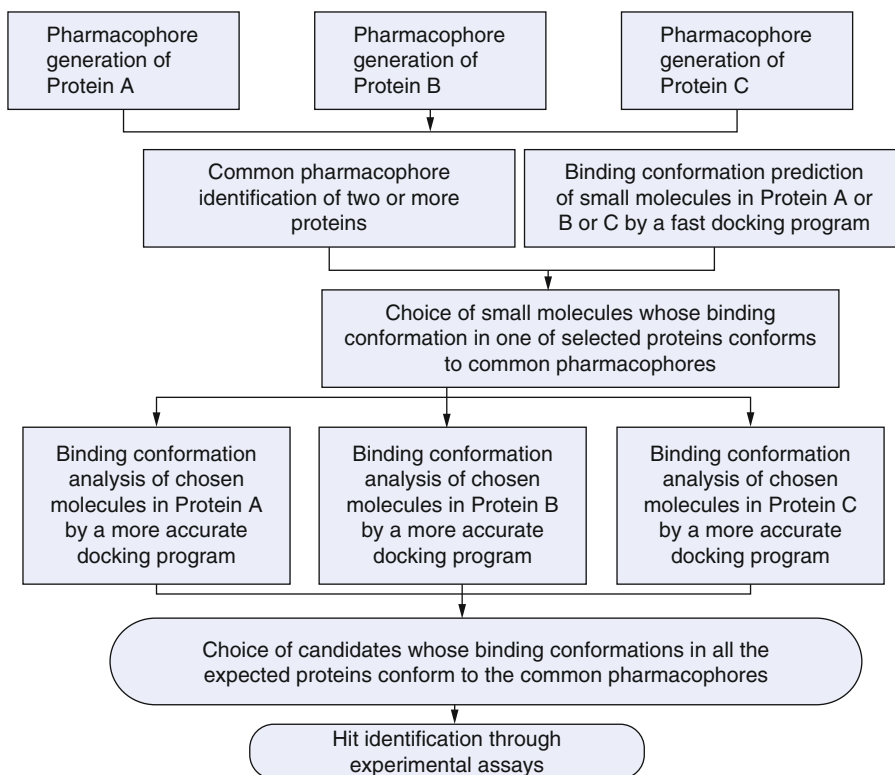


Fig. 9 Combined pharmacophore and molecular docking strategy of multi-target inhibitor discovery [157]

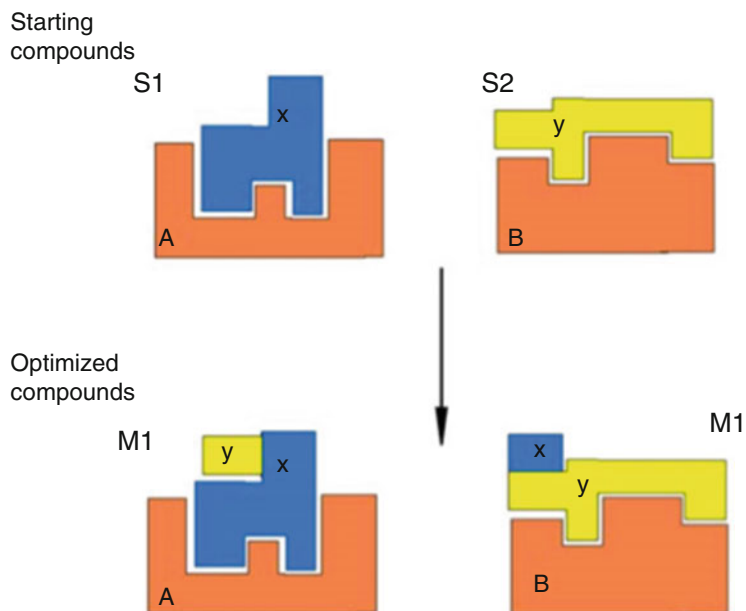


Fig. 10 Illustration of framework combination approach to multi-target drug discovery [157]

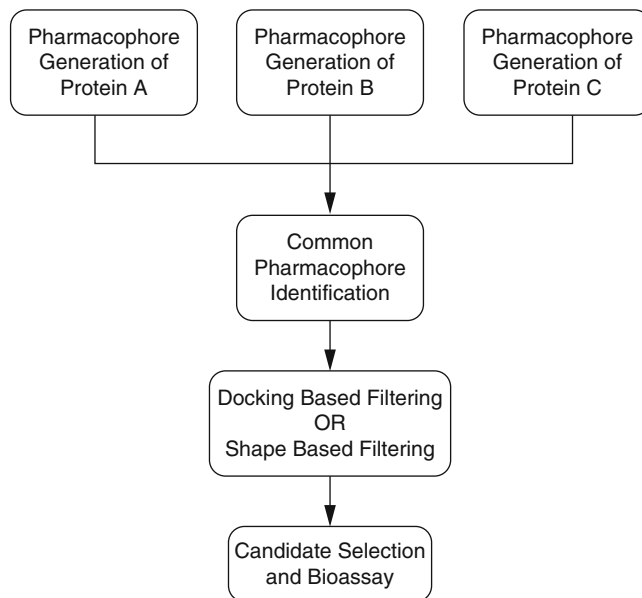


Fig. 11 Common pharmacophore-based multi-target drug design [23]

to all targets. Furthermore, a shape-based comparison can also be used as an alternative approach for docking [160].

This approach was used to design compounds that can simultaneously inhibit two human inflammation-related proteins, leukotriene A4 hydrolase (LTA4H-h) and human non-pancreatic secretory phospholipase A2 (hnpS-PLA2). The common pharmacophore was constructed using structure-based pharmacophores generated by Pocket V2 using co-crystal structures of LTA4H-h and sPLA2. The common pharmacophore model includes two hydrophobic sites and a metal-coordination site (Fig. 12).

The key step for this type of design strategy is generation of a common feature model. Rather than directly building a common pharmacophore from known pharmacophores, Hsu et al. used the information from a large-scale dataset of docked compounds to build a core site-moiety map followed by searching for multi-target inhibitors. They docked molecules to the orthologous proteins, shikimate kinase from *Mycobacterium tuberculosis* and *Helicobacter pylori* (MtSK and HpSK, respectively) [161].

In the end, which virtual libraries are to be searched should be carefully decided. Pockets with low similarities will make it difficult to identify an adequate number of multi-target candidates for further studies. A well-designed library with a broad range of chemical space and a sufficient number of molecular candidates is highly recommended.

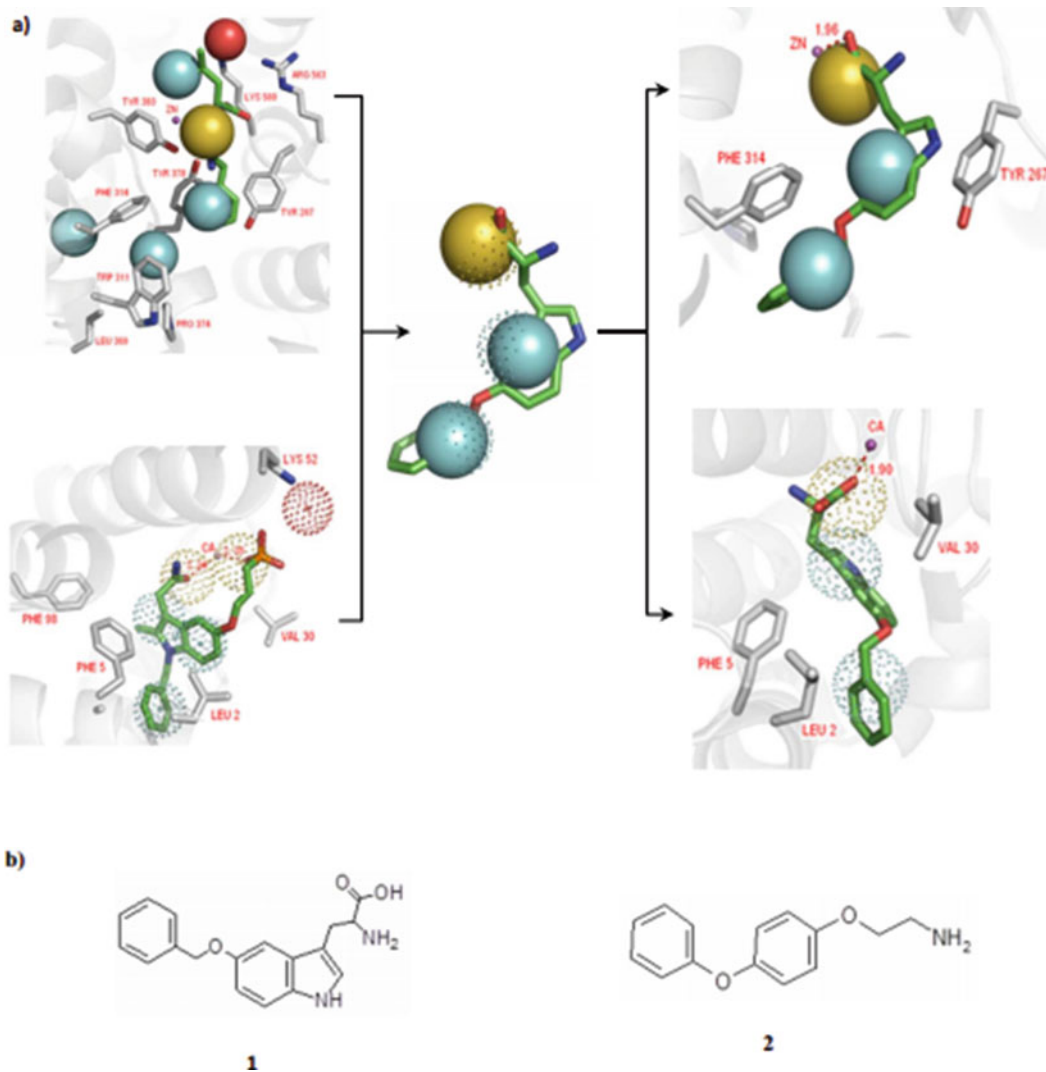


Fig. 12 The common pharmacophore-based design of a multi-target inhibitor of hnpS-PLA2 and LTA4H-h. **(a)** Compound 1 fit a common pharmacophore generated by two targets. The binding conformation was confirmed using a docking method. **(b)** Compound 1 and compound 2 showed the desired multi-target activities [23]

3 Challenges and Limitations of Computational MTDD

Although a rapid progress has been made in MTDD during the recent decade and many successful examples of its techniques and uses have been reported, several challenges still remain. For instance, target combinations were not selected since sparse quantitative data for network dynamic studies were found, and we could not construct a molecular network for the disease. As omics data

and quantitative measurements are gathered rapidly, soon more disease-connected models will be accessible to analyze target combinations. It is possible to use more than one target combination in most cases so that a disease network can be controlled. Numerous multiple target combinations will make it possible for the researchers to choose targets easily modulated by small molecules and at the same time obtaining the same level of network control [162]. The validation of MTDs in early experimental phase is significant for post-optimization. When we lay emphasis on polypharmacology as an important discovery strategy, we do not mean to choose any possible promiscuous leads for optimization. Computational methods are less expensive than HTS; however, it is necessary to carefully examine the selection of ligands from *in silico* methods to avoid non-specific activities [163]. Recently, the pan-assay interference compounds (PAINS) have been broadly employed to prevent false-positive ligands which have chemical reactions and may interfere with fluorescent assays or aggregate [164–166]. The PAINS criteria analysis is only responsible for possible false activity. The next type of non-specific false-positive is derived from colloidal aggregates where the aggregated compounds are either aqueous solutions or micelle formation.

Membrane and soluble proteins may adsorb aggregates resulting to inhibitory effects or infrequently activation [167–170]. A CADD method has been suggested for the detection of the similarity of compounds in well-known aggregators.

Also, some practical techniques such as hyphenated NMR with other well-known assays, flow cytometry light scattering, can detect aggregation [163]. Development of the MTD-based methods (e.g., pharmacophore and docking) as well as available drug candidates and exact scoring algorithms can reduce false-positive rates. Besides these, growing computational power will drop runtime, and more extensive searches are required to be carried out in the right time frame [171].

The development of *de novo* design both through ligand- and structure- based design will also benefit from the combined data. As computational and experimental tools are rapidly developed, devising MTDs for feasible target combinations is expected to be increasingly applicable in the near future.

The favorable efficacy of existing combination therapeutics illustrates that searches specifically designed to identify multi-target mechanisms can provide a new path forward in drug discovery. There are two approaches to create multi-target therapeutics, mixture of monotherapies and individual multi-targeted single agent [172]. The advantages and disadvantages of these two approaches are discussed in Table 2 [16].

Table 2
Advantages and disadvantages of different methods of MTDD

Advantages	Disadvantages
<i>Mixture of monotherapies</i>	
<ul style="list-style-type: none"> • Straightforward to tailor ratio of agents in the mixture—to account for differential potency at target and/or target stoichiometry • Opportunity for sequenced action or independently varying target exposure (e.g., using immediate versus extended release formulations of the components) • Speed to proof-of-concept clinical trials and cost advantage over multi-targeted single agent 	<ul style="list-style-type: none"> • Might need to align pharmacokinetics and/or pharmacodynamics of the component agents in co-formulated product • Approval might require “combination versus parts” factorial trial design • Must look for drug-drug interaction (DDI) relative to single agent
<i>Individual multi-targeted single agent</i>	
<ul style="list-style-type: none"> • Standard new chemical entity intellectual property position • Standard development program and regulatory approval process • Easier manufacture and formulation of an individual active pharmaceutical ingredient (API) compared with a mixture 	<ul style="list-style-type: none"> • Challenge to achieve multi-selective action without becoming nonselective • Challenge to optimize potency at two targets simultaneously in a single chemical structure—might only be able to achieve low potency at the targets • Difficult to achieve sequenced action at the targets

4 Future of MTDD

Nowadays, MTDs are a favorable plan to face complicated cases of diseases. MTD design is a promising alternative for single-target drugs in a more efficient manner for adjusting biological networks mainly in polygenic diseases or drug resistance. The search for more effective therapeutics with progressive advances of computer-based and theoretical tactics in bioinformatics science and linked disciplines would become possible. Using existing advanced computational ways may offer a more rapid and accurate mode to achieve this task.

Combination of the mentioned drug exploration/design with various data mining methods is useful for achieving maximum information of chemicals as well as lead discovery. The success rates and limitations of the MTDD technologies have made it an interesting multidisciplinary field of investigation for developers, students, and practitioners.

In this way, there is a great enthusiasm for the utilizing of “hybrid multi-target drugs” [173]. In silico screening has been extended further along the biological dimension leading to new integrative methods capable of estimating the pharmacological profile of molecules on multiple targets [72]. The beginning of

mt molecules is being attainable due to advanced techniques in various fields of medical and chemical sciences, for example, organic synthesis, evaluation of biological activities, natural products, and/or computational methodology, and different cheminformatics techniques.

In silico methods are efficient tools for pharmaceutical hypothesis testing/discovery that are used to obtain an insight to in vitro/in vivo models for building relationships (as qualitative or semi-quantitative) between activity and molecular structure. These methods include data mining, pharmacophores, databases, network analysis, de novo analysis, SAR/QSAR and other molecular modeling approaches, machine learning, and data analysis tools that use a computer [78]. These efforts display that multi-target ligands pave an important way in the progress of novel drug (lead) compounds.

All of these can provide a significant opportunity for future progress allowing rapid access to diverse series of multi-target molecules with higher activity, better selectivity, and lower toxicity than the common combination therapies. But the critical requirement for growth still remains with the efficacy of synthesis. Hence, the development of efficient synthetic methods allowing rapid access to diverse series of mt molecules must be further investigated by researchers [174].

Glossary

ADME/Tox	Absorption, distribution, metabolism, excretion, and toxicity
ANN	Artificial neural network
API	Active pharmaceutical ingredient
CADD	Computer-aided drug design
cIFD	Consensus Induced Fit Docking
CNS	Central nervous system
DDI	Drug-drug interaction
DT	Decision trees
ECP	Emerging chemical patterns
FN	Feature net
GPU	Graphics processing unit
GQSAR	Group-based QSAR
HTS	High-throughput screening
kNN	k-nearest neighbor
LDA	Linear discriminant analysis
LR	Logistic regression
MD	Molecular dynamics
MI	MARCH-INSIDE
MTD	Multi-targeted drugs
MTDD	Multi-target drug discovery/design

mt-docking	Multi-target docking
MTL	Multitask learning
MTL	Multitask learning
MTMD	Multi-targeted molecular dynamics
mt-QSARs	Multi-target quantitative structure-activity relationships
mt-SARs	Multi-target structure-activity relationships
ROC	Receiver operating characteristics
SARs	Structure-activity relationships
SVM	Support vector machines
VS	Virtual screening
WSOF	Weighted-sum-of-objective-functions

References

- Achenbach J, Proschak E (2011) Rational, computer-aided design of multi-target ligands. *J Chem* 3(S1):P10
- Metz JT, Hajduk PJ (2010) Rational approaches to targeted polypharmacology: creating and navigating protein–ligand interaction networks. *Curr Opin Chem Biol* 14(4):498–504
- Morphy R, Rankovic Z (2009) Designing multiple ligands-medicinal chemistry strategies and challenges. *Curr Pharm Des* 15(6):587–600
- Anighoro A, Bajorath J, Rastelli G (2014) Polypharmacology: challenges and opportunities in drug discovery: miniperspective. *J Med Chem* 57(19):7874–7887
- Yeung KS, Meanwell NA (2008) Inhibition of hERG channel trafficking: an under-explored mechanism for drug-induced QT prolongation. *ChemMedChem* 3(10):1501–1502
- Setola V, Roth BL (2005) Screening the receptorome reveals molecular targets responsible for drug-induced side effects: focus on ‘fen-phen’. *Expert Opin Drug Metab Toxicol* 1(3):377–387
- Jooste E, Zhang Y, Emala CW (2005) Rapacuronium preferentially antagonizes the function of M2 versus M3 muscarinic receptors in Guinea pig airway smooth muscle. *Anesthesiology* 102(1):117–124
- Reddy AS, Zhang S (2013) Polypharmacology: drug discovery for the future. *Expert Rev Clin Pharmacol* 6(1):41–47
- Morphy R, Rankovic Z (2005) Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem* 48(21):6523–6543
- Rastelli G, Pinzi L (2015) Computational polypharmacology comes of age. *Front Pharmacol* 6:157
- Baell JB (2010) Observations on screening-based research and some concerning trends in the literature. *Future Med Chem* 2(10):1529–1546
- Abdolmaleki A, B Ghasemi J, Ghasemi F (2017) Computer aided drug design for multi-target drug design: SAR/QSAR, molecular docking and pharmacophore methods. *Curr Drug Targets* 18(5):556–575
- Jadhav A, Ferreira RS, Klumpp C, Mott BT, Austin CP, Inglese J et al (2009) Quantitative analyses of aggregation, autofluorescence, and reactivity artifacts in a screen for inhibitors of a thiol protease. *J Med Chem* 53(1):37–51
- Doak AK, Wille H, Prusiner SB, Shoichet BK (2010) Colloid formation by drugs in simulated intestinal fluid. *J Med Chem* 53(10):4259–4265
- Abdolmaleki A, Ghasemi JB (2017) Dual-acting of hybrid compounds—a new dawn in the discovery of multi-target drugs: lead generation approaches. *Curr Top Med Chem* 17(9):1096–1114
- Zimmermann GR, Lehar J, Keith CT (2007) Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov Today* 12(1–2):34–42
- Talevi A (2015) Multi-target pharmacology: possibilities and limitations of the “skeleton key approach” from a medicinal chemist perspective. *Front Pharmacol* 6:205
- Hornberg JJ, Bruggeman FJ, Westerhoff HV, Lankelma J (2006) Cancer: a systems biology disease. *Biosystems* 83(2–3):81–90
- Khalil I, Hill C (2005) Systems biology for cancer. *Curr Opin Oncol* 17(1):44–48
- Keiser M, Setola V, Irwin J, Laggner C, Abbas A, Hufeisen S et al (2009) Predicting

- new molecular targets for known drugs. *Nature* 462:175–181
21. Jenwithesuk E, Samudrala R (2007) Identification of potential HIV-1 targets of minocycline. *Bioinformatics* 23(20):2797–2799
 22. Shang E, Yuan Y, Chen X, Liu Y, Pei J, Lai L (2014) De novo design of multitarget ligands with an iterative fragment-growing strategy. *J Chem Inf Model* 54(4):1235–1241
 23. Zhang W, Pei J, Lai L (2017) Computational multitarget drug design. *J Chem Inf Model* 57(3):403–412
 24. Lavecchia A, Cerchia C (2016) In silico methods to address polypharmacology: current status, applications and future perspectives. *Drug Discov Today* 21(2):288–298
 25. Andricopulo AD, Montanari CA (2005) Structure-activity relationships for the design of small-molecule inhibitors. *Mini Rev Med Chem* 5(6):585–593
 26. Guha R (2010) The ups and downs of structure-activity landscapes. In: *Chemoinformatics and computational chemical biology*. Springer, Heidelberg, pp 101–117
 27. Martins GR, Napolitano HB, Camargo LTFM, Camargo AJ (2012) Structure-activity relationship study of rutaearpine analogous active against central nervous system cancer. *J Braz Chem Soc* 23(12):2183–2190
 28. www.biolog.com
 29. Duffy BC, Zhu L, Decornez H, Kitchen DB (2012) Early phase drug discovery: cheminformatics and computational techniques in identifying lead series. *Bioorg Med Chem* 20:5324–5342
 30. Wassermann AM, Peltason L, Bojarath J (2010) Computational analysis of multitarget structure activity relationships to derive preference orders for chemical modifications toward target selectivity. *ChemMedChem* 5:847–858
 31. Chen YC (2015) Beware of docking! *Trends Pharmacol Sci* 36(2):78–95
 32. Dimova D, Bajorath J (2012) Design of multi-target activity landscapes that capture hierarchical activity cliff distributions. *J Cheminform* 4(Suppl 1):P4
 33. Santos CBR, Lobato CC, Alexandre M, Sousa C et al (2014) Molecular modeling: origin, fundamental concepts and applications using structure-activity relationship and quantitative structure-activity relationship. *Rev Theor Sci* 2:1–25
 34. Medina-Franco JL, Giulianotti MA, Welmaker GS, Houghten RA (2013) Shifting from the single to the multi-target paradigm in drug discovery. *Drug Discov Today* 18(9,10):495–501
 35. Angus D, Bingham M, Buchanan D, Dunbar N, Gibson L, Goodwin R et al (2011) The identification, and optimisation of hERG selectivity, of a mixed NET/SERT re-uptake inhibitor for the treatment of pain. *Bioorg Med Chem Lett* 21(1):271–275
 36. Bénardeau A, Benz J, Binggeli A, Blum D, Boehringer M, Grether U et al (2009) Aleglitzar, a new, potent, and balanced dual PPAR α/γ agonist for the treatment of type II diabetes. *Bioorg Med Chem Lett* 19(9):2468–2473
 37. Zhang W, Nan G, Wu H-H, Jiang M, Li T-X, Wang M et al (2017) A simple and rapid UPLC-PDA method for quality control of *Nardostachys jatamansi*. *Planta Med.* <https://doi.org/10.1055/s-0043-123655>
 38. Norman P (2008) Pfizer's dual-acting β_2 agonists/muscarinic M3 antagonists: Pfizer: WO2008041095. *Expert Opin Ther Pat* 18(9):1091–1096
 39. Zanni R, Galvez-Llompарт M, Galvez J, Garcia-Domenech R (2014) QSAR multitarget in drug discovery: a review. *Curr Comput Aided Drug Des* 10(2):129–136
 40. Prado-Prado FJ, Uriarte E, Borges F, González-Díaz H (2009) Multi-target spectral moments for QSAR and complex networks study of antibacterial drugs. *Eur J Med Chem* 44(11):4516–4521
 41. Yap CW (2010) Software news and update PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J Comput Chem* 32(7):1467–1474
 42. Alonso N, Caamaño O, Romero-Duran FJ, Luan F, D S Cardeiro MN, Yañez M, González-Díaz H, García-Mera X (2013) Model for high-throughput screening of multitarget drugs in chemical neurosciences: synthesis, assay, and theoretic study of Rasagiline carbamates. *ACS Chem Neurosci* 4(10):1393–1403
 43. Ahmadi P, Ghasemi JB (2014) 3D-QSAR and docking studies of the stability constants of different guest molecules with beta-cyclodextrin. *J Incl Phenom Macrocycl Chem* 79(3–4):423–435
 44. Ghasemi JB, Pirhadi S, Ayati M (2011) 3D-QSAR studies of 2-arylbenzoxazoles as novel cholesteryl ester transfer protein inhibitors. *Bull Kor Chem Soc* 32(2):645–650
 45. Abedi H, Ghasemi JB, Ebrahimzadeh H (2013) 3D-QSAR, CoMFA, and CoMSIA of new phenyloxazolidinones derivatives as

- potent HIV-1 protease inhibitors. *Struct Chem* 24:433–444
46. Ghasemi JB, Safavi-Sohi R, Barbosa EG (2012) 4D-LQTA-QSAR and docking study on potent gram-negative specific LpxC inhibitors: a comparison to CoMFA modeling. *Mol Divers* 16:203–213
 47. Prado-Prado FJ, Gonzalez-Diaz H, Vega OM, Ubeira FM, Chou KC (2008) Unified QSAR approach to antimicrobials. Part 3: First multi-tasking QSAR model for input-coded prediction, structural back-projection, and complex networks clustering of antiprotozoal compound. *Bioorg Med Chem* 16:5871
 48. Prado-Prado FJ, Vega OM, Uriarte E, Ubeira FM, Chou KC, Gonzalez-Diaz H (2009) Unified QSAR approach to antimicrobials. 4. Multi-target QSAR modeling and comparative multi-distance study of the giant components of antiviral drug-drug complex networks. *Bioorg Med Chem* 17:569
 49. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M et al (2014) QSAR modeling: where have you been? Where are you going to? *J Med Chem* 57 (12):4977–5010
 50. Yap CW, Li H, Ji ZL, Chen YZ (2007) Regression methods for developing QSAR and QSPR models to predict compounds of specific pharmacodynamic, pharmacokinetic and toxicological properties. *Mini Rev Med Chem* 7:1097–1107
 51. Koike A (2006) Comparison of methods for chemical-compound affinity prediction. *SAR QSAR Environ Res* 17:497–514
 52. Cruz-Monteagudo M, Borges F, Cordeiro MN, Cagide Fajin JL, Morell C, Ruiz RM et al (2008) Desirability-based methods of multiobjective optimization and ranking for global QSAR studies. Filtering safe and potent drug candidates from combinatorial libraries. *J Comb Chem* 10:897–913
 53. Cruz-Monteagudo M, Borges F, Cordeiro MN (2008) Desirability-based multiobjective optimization for global QSAR studies: application to the design of novel NSAIDs with improved analgesic, anti-inflammatory, and ulcerogenic profiles. *J Comput Chem* 29:2445–2459
 54. Sivakumar T, Manavalan R, Muralidharan C, Valliappan K (2007) Multi-criteria decision making approach and experimental design as chemometric tools to optimize HPLC separation of domperidone and pantoprazole. *J Pharm Biomed Anal* 43:1842–1848
 55. Ajmani S, Kulkarni SA (2012) Application of GQSAR for scaffold hopping and lead optimization in multi-target inhibitors. *Mol Inform* 31:473–490
 56. Namasivayam V, Hu Y, Balfer J, Bajorath J (2013) Classification of compounds with distinct or overlapping multi-target activities and diverse molecular mechanisms using emerging chemical patterns. *J Chem Inf Comput Sci* 53(6):1272–1281
 57. Auer J, Bajorath J (2008) Distinguishing between bioactive and modeled compound conformations through mining of emerging chemical patterns. *J Chem Inf Model* 48:1747–1753
 58. Auer J, Bajorath J (2006) Emerging chemical patterns: a new methodology for molecular classification and compound selection. *J Chem Inf Model* 46:2502–2514
 59. Prado-Prado FJ, Borges F, Uriarte E, Pérez-Montoto LG, González-Díaz H (2009) Multi-target spectral moment: QSAR for antiviral drugs vs. different viral species. *Anal Chim Acta* 651:159–164
 60. Gonzalez-Diaz H, Prado-Prado FJ (2008) Unified QSAR and network-based computational chemistry approach to antimicrobials. Part I: Multispecies activity models for antifungals. *J Comput Chem* 29:656–667
 61. Gonzalez-Diaz H, Prado-Prado FJ, Santana L, Uriarte E (2006) Unify QSAR approach to antimicrobials. Part I: Predicting antifungal activity against different species. *Bioorg Med Chem Lett* 14:5973–5980
 62. Vina D, Uriarte E, Orallo F, Gonzalez-Diaz H (2009) Alignment-free prediction of a drug-target complex network based on parameters of drug connectivity and protein sequence of receptors. *Mol Pharm* 6:825–835
 63. González-Díaz H, Agüero G, Cabrera MA, Molina R, Santana L, Uriarte E et al (2005) Unified Markov thermodynamics based on stochastic forms to classify drug considering molecular structure, partition system and biological species: distribution of the antimicrobial G1 on rat tissues. *Bioorg Med Chem Lett* 15:551–557
 64. Ramos de Armas R, González-Díaz H, Molina R, Pérez-González M, Uriarte E (2004) Stochastic-based descriptors studying peptides biological properties: modeling the bitter tasting threshold of dipeptides. *Bioorg Med Chem* 12(18):4815–4822
 65. Prado-Prado FJ, García-Mera X, González-Díaz H (2010) Multi-target spectral moment QSAR versus ANN for antiparasitic drugs

- against different parasite species. *Bioorg Med Chem* 18:2225–2231
66. Derringer G, Suich R (1980) Simultaneous optimization of several response variables. *J Qual Technol* 12:214
 67. Cruz-Monteagudo M, Borges F, Cordeiro M (2008) Desirability-based multiobjective optimization for global QSAR studies: application to the design of novel NSAIDs with improved analgesic, antiinflammatory, and ulcerogenic profiles. *J Comput Chem* 29 (14):2445–2459
 68. Jhoti H (2007) Fragment-based drug discovery using rational design. *Ernst Schering Found Symp Proc* (3):169–185
 69. Mauser H, Guba W (2008) Recent developments in de novo design and scaffold hopping. *Curr Opin Drug Discov Devel* 11 (3):365–374
 70. Schneider G, Böhm H-J (2002) Virtual screening and fast automated docking methods. *Drug Discov Today* 7(1):64–70
 71. Bottegoni G, Favia AD, Recanatini M, Cavalli A (2012) The role of fragment-based and computational methods in polypharmacology. *Drug Discov Today* 17(1–2):23–34
 72. Ferruz N, Harvey MJ, Mestres J, De Fabritiis G (2015) Insights from fragment hit binding assays by molecular simulations. *J Chem Inf Model* 55(10):2200–2205
 73. Joseph-McCarthy D, Campbell AJ, Kern G, Moustakas D (2014) Fragment-based lead discovery and design. *J Chem Inf Model* 54 (3):693–704
 74. Li J, Ballmer SG, Gillis EP, Fujii S, Schmidt MJ, Palazzolo AM et al (2015) Synthesis of many different types of organic small molecules using one automated process. *Science* 347(6227):1221–1226
 75. Reutlinger M, Rodrigues T, Schneider P, Schneider G (2014) Combining on-chip synthesis of a focused combinatorial library with computational target prediction reveals imidazopyridine GPCR ligands. *Angew Chem Int Ed* 53(2):582–585
 76. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436
 77. Xu Y, Dai Z, Chen F, Gao S, Pei J, Lai L (2015) Deep learning for drug-induced liver injury. *J Chem Inf Model* 55(10):2085–2093
 78. Ekins S, Mestres J, Testa B (2007) In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *Br J Pharmacol* 152(1):9–20
 79. Davis IW, Raha K, Head MS, Baker D (2009) Blind docking of pharmaceutically relevant compounds using RosettaLigand. *Protein Sci* 18(9):1998–2002
 80. Pirhadi S, Sunseri J, Koes DR (2016) Open source molecular modeling. *J Mol Graph Model* 69:127–143
 81. Chung JY, Cho SJ, Hah J-M (2011) A python-based docking program utilizing a receptor bound ligand shape: PythDock. *Arch Pharm Res* 34(9):1451
 82. Plewczynski D, Łazniewski M, Grotthuss MV, Rychlewski L, Ginalski K (2011) VoteDock: consensus docking method for prediction of protein–ligand interactions. *J Comput Chem* 32(4):568–581
 83. Takaya D, Yamashita A, Kamijo K, Gomi J, Ito M, Maekawa S et al (2011) A new method for induced fit docking (GENIUS) and its application to virtual screening of novel HCV NS3-4A protease inhibitors. *Bioorg Med Chem* 19(22):6892–6905
 84. Kim D-S, Kim C-M, Won C-I, Kim J-K, Ryu J, Cho Y et al (2011) BetaDock: shape-priority docking method based on beta-complex. *J Biomol Struct Dyn* 29 (1):219–242
 85. Klenner A, Weisel M, Reisen F, Proschak E, Schneider G (2010) Automated docking of flexible molecules into receptor binding sites by ligand self-organization in situ. *Mol Inform* 29(3):189–193
 86. Feng JA, Marshall GR (2010) SKATE: a docking program that decouples systematic sampling from scoring. *J Comput Chem* 31 (14):2540–2554
 87. Grosdidier A, Zoete V, Michielin O (2011) Fast docking using the CHARMM force field with EADock DSS. *J Comput Chem* 32 (10):2149–2159
 88. Brylinski M, Skolnick J (2010) Q-DockLHM: low-resolution refinement for ligand comparative modeling. *J Comput Chem* 31 (5):1093–1105
 89. Shin WH, Heo L, Lee J, Ko J, Seok C, Lee J (2011) LigDockCSA: protein–ligand docking using conformational space annealing. *J Comput Chem* 32(15):3226–3232
 90. Shin W-H, Seok C (2012) GalaxyDock: protein–ligand docking with flexible protein side-chains. *J Chem Inf Model* 52(12):3225–3232
 91. Ding F, Dokholyan NV (2012) Incorporating backbone flexibility in MedusaDock improves ligand-binding pose prediction in the CSAR2011 docking benchmark. *J Chem Inf Model* 53(8):1871–1879
 92. Ding F, Yin S, Dokholyan NV (2010) Rapid flexible docking using a stochastic rotamer

- library of ligands. *J Chem Inf Model* 50 (9):1623–1632
93. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31(2):455–461
 94. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98(18):10037–10041
 95. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS et al (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 30(16):2785–2791
 96. Ballante F, Marshall GR (2016) An automated strategy for binding-pose selection and docking assessment in structure-based drug design. *J Chem Inf Model* 56(1):54–72
 97. Bullock C, Cornia N, Jacob R, Remm A, Peavey T, Weekes K et al (2013) DockoMatic 2.0: high throughput inverse virtual screening and homology modeling. *J Chem Inf Model* 53(8):2161–2170
 98. Jiang X, Kumar K, Hu X, Wallqvist A, Reifman J (2008) DOVIS 2.0: an efficient and easy to use parallel virtual screening tool based on AutoDock 4.0. *Chem Cent J* 2 (1):18
 99. Li H, Leung K-S, Wong M-H (eds) (2012) idock: a multithreaded virtual screening tool for flexible ligand docking. In: *IEEE symposium on computational intelligence in bioinformatics and computational biology (CIBCB)*, 2012. IEEE
 100. Abreu RM, Froufe HJ, Queiroz MJR, Ferreira IC (2010) MOLA: a bootable, self-configuring system for virtual screening using AutoDock4/Vina on computer clusters. *J Chem* 2(1):10
 101. Durrant JD, McCammon JA (2011) NNScore 2.0: a neural-network receptor–ligand scoring function. *J Chem Inf Model* 51 (11):2897–2903
 102. Meier R, Pippel M, Brandt F, Sippl W, Baldauf C (2010) ParaDockS: a framework for molecular docking with population-based metaheuristics. *J Chem Inf Model* 50(5):879–889
 103. Dallakyan S, Olson AJ (2015) Small-molecule library screening by docking with PyRx. *Methods Mol Biol* 1263:243–250
 104. Ruiz-Carmona S, Alvarez-Garcia D, Foloppe N, Garmendia-Doval AB, Juhos S, Schmidtke P et al (2014) rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput Biol* 10(4):e1003571
 105. Li H, Leung KS, Wong MH, Ballester PJ (2015) Improving AutoDock Vina using random forest: the growing accuracy of binding affinity prediction by the effective exploitation of larger data sets. *Mol Inform* 34 (2–3):115–126
 106. Koes DR, Baumgartner MP, Camacho CJ (2013) Lessons learned in empirical scoring with smina from the CSAR 2011 benchmarking exercise. *J Chem Inf Model* 53(8):1893–1904
 107. Cereto-Massagué A, Ojeda MJ, Joosten RP, Valls C, Mulero M, Salvado MJ et al (2013) The good, the bad and the dubious: VHELIBS, a validation helper for ligands and binding sites. *J Chem* 5(1):36
 108. Zhang X, Wong SE, Lightstone FC (2013) Message passing interface and multithreading hybrid for parallel molecular docking of large databases on petascale high performance computing machines. *J Comput Chem* 34 (11):915–927
 109. Ellingson SR, Smith JC, Baudry J (2013) VinaMPI: facilitating multiple receptor high-throughput virtual docking on high-performance computers. *J Comput Chem* 34 (25):2212–2221
 110. Zonta N, Grimstead IJ, Avis NJ, Brancale A (2009) Accessible haptic technology for drug design applications. *J Mol Model* 15 (2):193–196
 111. Vorobjev YN (2010) Blind docking method combining search of low-resolution binding sites with ligand pose refinement by molecular dynamics-based global optimization. *J Comput Chem* 31(5):1080–1092
 112. Luo W, Pei J, Zhu Y (2010) A fast protein–ligand docking algorithm based on hydrogen bond matching and surface shape complementarity. *J Mol Model* 16(5):903–913
 113. Sadjad B, Zsoldos Z (2011) Toward a robust search method for the protein–drug docking problem. *IEEE/ACM Trans Comput Biol Bioinform* 8(4):1120–1133
 114. Takahashi O, Masuda Y, Muroya A, Furuya T (2010) Theory of docking scores and its application to a customizable scoring function. *SAR QSAR Environ Res* 21 (5–6):547–558
 115. Lee HS, Zhang Y (2012) BSP-SLIM: a blind low-resolution ligand–protein docking approach using predicted protein structures. *Proteins* 80(1):93–110
 116. Ouyang X, Zhou S, Su CTT, Ge Z, Li R, Kwok CK (2013) CovalentDock: automated covalent docking with parameterized covalent

- linkage energy estimation and molecular geometry constraints. *J Comput Chem* 34 (4):326–336
117. Cortés AC, Klett J, Dos HS, Perona A, Gil-Redondo R, Francis SM et al (2012) CRDOCK: an ultrafast multipurpose protein-ligand docking tool. *J Chem Inf Model* 52(8):2300–2309
 118. Liu Y, Zhao L, Li W, Zhao D, Song M, Yang Y (2013) FIPSDock: a new molecular docking technique driven by fully informed swarm optimization algorithm. *J Comput Chem* 34 (1):67–75
 119. Shin WH, Kim JK, Kim DS, Seok C (2013) GalaxyDock2: protein–ligand docking using beta-complex and global optimization. *J Comput Chem* 34(30):2647–2656
 120. McGann M (2012) FRED and HYBRID docking performance on standardized datasets. *J Comput Aided Mol Des* 26 (8):897–906
 121. Beato C, Beccari AR, Cavazzoni C, Lorenzi S, Costantino G (2013) Use of experimental design to optimize docking performance: the case of ligendock, the docking module of liGen, a new de novo design program. *J Chem Inf Model*. <https://doi.org/10.1021/ci400079k>
 122. Wu MY, Dai DQ, Yan H (2012) PRL-dock: protein-ligand docking based on hydrogen bond matching and probabilistic relaxation labeling. *Proteins* 80(9):2137–2153
 123. Handoko SD, Ouyang X, Su CTT, Kwok CK, Ong YS (2012) QuickVina: accelerating AutoDock Vina using gradient-based heuristics for global optimization. *IEEE/ACM Trans Comput Biol Bioinform* 9 (5):1266–1272
 124. Hoffer L, Renaud J-P, Horvath D (2013) In silico fragment-based drug discovery: setup and validation of a fragment-to-lead computational protocol using S4MPLE. *J Chem Inf Model* 53(4):836–851
 125. Brylinski M, Feinstein WP (2013) eFindSite: improved prediction of ligand binding sites in protein models using meta-threading, machine learning and auxiliary ligands. *J Comput Aided Mol Des* 27(6):551–567
 126. Schmidtke P, Bidon-Chanal A, Luque FJ, Barril X (2011) MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics* 27 (23):3276–3285
 127. Oliveira SH, Ferraz FA, Honorato RV, Xavier-Neto J, Sobreira TJ, de Oliveira PS (2014) KVFinder: steered identification of protein cavities as a PyMOL plugin. *BMC Bioinformatics* 15(1):197
 128. Till MS, Ullmann GM (2010) McVol-A program for calculating protein volumes and identifying cavities by a Monte Carlo algorithm. *J Mol Model* 16(3):419–429
 129. Weisel M, Proschak E, Schneider G (2007) PocketPicker: analysis of ligand binding-sites with shape descriptors. *Chem Cent J* 1(1):7
 130. Durrant JD, Votapka L, Sørensen J, Amaro RE (2014) POVME 2.0: an enhanced tool for determining pocket shape and volume characteristics. *J Chem Theory Comput* 10 (11):5047–5056
 131. Kalid O, Warshaviak DT, Shechter S, Sherman W, Shacham S (2012) Consensus Induced Fit Docking (cIFD): methodology, validation, and application to the discovery of novel Crm1 inhibitors. *J Comput Aided Mol Des* 26(11):1217–1228
 132. Craig IR, Essex JW, Spiegel K (2010) Ensemble docking into multiple crystallographically derived protein structures: an evaluation based on the statistical analysis of enrichments. *J Chem Inf Model* 50(4):511–524
 133. Korb O, Olsson TS, Bowden SJ, Hall RJ, Verdonk ML, Liebeschuetz JW et al (2012) Potential and limitations of ensemble docking. *J Chem Inf Model* 52(5):1262–1274
 134. Xu M, Lill MA (2011) Utilizing experimental data for reducing ensemble size in flexible-protein docking. *J Chem Inf Model* 52 (1):187–198
 135. Sinko W, Lindert S, McCammon JA (2013) Accounting for receptor flexibility and enhanced sampling methods in computer-aided drug design. *Chem Biol Drug Des* 81 (1):41–49
 136. Wu J, Chen C, Hong B (2012) A GPU-based approach to accelerate computational protein-DNA docking. *Comput Sci Eng* 14 (3):20–29
 137. Scarpazza DP, Ierardi DJ, Lerer AK, Mackenzie KM, Pan AC, Bank JA et al (eds) (2013) Extending the generality of molecular dynamics simulations on a special-purpose machine. In: 2013 I.E. 27th international symposium on parallel & distributed processing (IPDPS). IEEE
 138. De Vivo M, Cavalli A (2017) Recent advances in dynamic docking for drug discovery. *Wiley Interdiscip Rev Comput Mol Sci* 7(6):e1320
 139. Whalen KL, Chang KM, Spies MA (2011) Hybrid steered molecular dynamics-docking: an efficient solution to the problem of ranking inhibitor affinities against a flexible drug target. *Mol Inform* 30(5):459–471

140. Prajapati R, Sangamwar AT (2014) Translocation mechanism of P-glycoprotein and conformational changes occurring at drug-binding site: insights from multi-targeted molecular dynamics. *Biochim Biophys Acta* 1838(11):2882–2898
141. Bohari MH, Sastry GN (2012) FDA approved drugs complexed to their targets: evaluating pose prediction accuracy of docking protocols. *J Mol Model* 18(9):4263–4274
142. Houston DR, Walkinshaw MD (2013) Consensus docking: improving the reliability of docking in a virtual screening context. *J Chem Inf Model* 53(2):384–390
143. Yuriev E, Agostino M, Ramsland PA (2011) Challenges and advances in computational docking: 2009 in review. *J Mol Recognit* 24(2):149–164
144. Yuriev E, Ramsland PA (2013) Latest developments in molecular docking: 2010–2011 in review. *J Mol Recognit* 26(5):215–239
145. Huang S-Y, Zou X (2010) Inclusion of solvation and entropy in the knowledge-based scoring function for protein–ligand interactions. *J Chem Inf Model* 50(2):262–273
146. Ghasemi JB, Abdolmaleki A, Shiri F (2016) Molecular docking challenges and limitations. In: Applied case studies and solutions in molecular docking-based drug design. IGI Global, Hershey, pp 56–80
147. Clemente JC, Govindasamy L, Madabushi A, Fisher SZ, Moose RE, Yowell CA et al (2006) Structure of the aspartic protease plasmepsin 4 from the malarial parasite *Plasmodium malariae* bound to an allophenylnorstatine-based inhibitor. *Acta Crystallogr D Biol Crystallogr* 62(3):246–252
148. Scotti L, Mendonca Junior FJ, Ishiki HM, Ribeiro FF, Singla RK, Barbosa Filho JM et al (2017) Docking studies for multi-target drugs. *Curr Drug Targets* 18(5):592–604
149. Azam F, Amer AM, Abulifa AR, Elzwawi MM (2014) Ginger components as new leads for the design and development of novel multi-targeted anti-Alzheimer's drugs: a computational investigation. *Drug Des Devel Ther* 8:2045
150. Ehrlich P (1909) Über den jetzigen Stand der Chemotherapie. *Eur J Inorg Chem* 42(1):17–47
151. Wermuth C, Ganellin C, Lindberg P, Mitscher L (1998) Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl Chem* 70(5):1129–1143
152. Seidel T, Bryant SD, Ibis G, Poli G, Langer T (2017) Tutorials in chemoinformatics. 3D pharmacophore modeling techniques in computer-aided molecular design using LigandScout. Wiley, New York. <https://doi.org/10.1002/9781119161110.ch20>
153. Chan DSH, Lee HM, Yang F, Che CM, Wong CC, Abagyan R et al (2010) Structure-based discovery of natural-product-like TNF- α inhibitors. *Angew Chem Int Ed* 49(16):2860–2864
154. Pirhadi S, Shiri F, Ghasemi JB (2013) Methods and applications of structure based pharmacophores in drug discovery. *Curr Top Med Chem* 13(9):1036–1047
155. Schneider G, Neidhart W, Giller T, Schmid G (1999) “Scaffold-Hopping” by topological pharmacophore search: a contribution to virtual screening. *Angew Chem Int Ed* 38(19):2894–2896
156. Langer T, Hoffmann R, Bryant S, Lesur B (2009) Hit finding: towards ‘smarter’ approaches. *Curr Opin Pharmacol* 9(5):589–593
157. Ma XH, Shi Z, Tan C, Jiang Y, Go ML, Low BC et al (2010) In-Silico approaches to multi-target drug discovery. *Pharm Res* 27(5):739–749
158. Haupt VJ, Daminelli S, Schroeder M (2013) Drug promiscuity in PDB: protein binding site similarity is key. *PLoS One* 8(6):e65894
159. Günther S, Senger C, Michalsky E, Goede A, Preissner R (2006) Representation of target-bound drugs by computed conformers: implications for conformational libraries. *BMC Bioinformatics* 7(1):293
160. Moser D, Wisniewska JM, Hahn S, Achenbach J, Buscató E, Klingler F-M et al (2012) Dual-target virtual screening by pharmacophore elucidation and molecular shape filtering. *ACS Med Chem Lett* 3(2):155–158
161. Hsu K-C, Cheng W-C, Chen Y-F, Wang H-J, Li L-T, Wang W-C et al (2012) Core site-moiety maps reveal inhibitors and binding mechanisms of orthologous proteins by screening compound libraries. *PLoS One* 7(2):e32142
162. Nussinov R, Tsai C-J, Csermely P (2011) Allo-network drugs: harnessing allostery in cellular networks. *Trends Pharmacol Sci* 32(12):686–693
163. Klumpp M (2016) Non-stoichiometric inhibition in integrated lead finding—a literature review. *Expert Opin Drug Discov* 11(2):149–162

164. Baell J, Walters MA (2014) Chemical con artists foil drug discovery. *Nature* 513 (7519):481
165. Erlanson DA (2015) Learning from PAINful lessons. *J Med Chem* 58(5):2088–2090
166. Baell JB, Holloway GA (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem* 53(7):2719–2740
167. Zorn JA, Wolan DW, Agard NJ, Wells JA (2012) Fibrils colocalize caspase-3 with procaspase-3 to foster maturation. *J Biol Chem* 287(40):33781–33795
168. Sassano MF, Doak AK, Roth BL, Shoichet BK (2013) Colloidal aggregation causes inhibition of G protein-coupled receptors. *J Med Chem* 56(6):2406–2414
169. Lin H, Sassano MF, Roth BL, Shoichet BK (2013) A pharmacological organization of G protein-coupled receptors. *Nat Methods* 10 (2):140
170. Coan KE, Maltby DA, Burlingame AL, Shoichet BK (2009) Promiscuous aggregate-based inhibitors promote enzyme unfolding. *J Med Chem* 52(7):2067–2075
171. Sterling T, Irwin JJ (2015) ZINC 15—ligand discovery for everyone. *J Chem Inf Model* 55 (11):2324–2337
172. Morphy R, Kay C, Rankovic Z (2004) From magic bullets to designed multiple ligands. *Drug Discov Today* 9(15):641–651
173. Peperidou A, Kapoukranidou D, Kontogiorgis C, Hadjipavlou-Litina D (2014) Multitarget molecular hybrids of cinnamic acids. *Molecules* 19(12):20197–20226
174. Bérubé G (2016) An overview of molecular hybrids in drug discovery. *Expert Opin Drug Discov* 11:281–305

Part II

Computational Multi-Target Drug Design: Literature Reviews



Multitarget Drug Design for Neurodegenerative Diseases

Marco Catto, Daniela Trisciuzzi, Domenico Alberga,
Giuseppe Felice Mangiatordi, and Orazio Nicolotti

Abstract

The quest for new pharmacological treatments of neurodegenerative diseases (NDs) still remains a priority for researchers and caregivers. Being inherently multifactorial, NDs benefited of the paradigm shift from “one-drug-one-target” to “one-drug-more-target” which is typical of the so-called multitarget approach, whose ultimate aim is that of providing a wider pharmacological spectrum to single molecular entities. A multitarget drug should encompass the basic molecular features necessary for an effective interaction with each desired biological target. In this respect, different drug design strategies, mostly inspired by ligand-based and target-based approaches, have been envisaged to achieve this goal. Indeed, huge efforts have been addressed in recent years to harmonically integrate the amount of different (bio)chemical information in the attempt to derive reliable predictive multitarget models. An overview of multitarget computational methods as well as of some successful applications to NDs will be the focus of this chapter.

Keywords Computational chemistry, Computer-aided drug design, Fragment-based drug design, Ligand-based approaches, Molecular docking, Multitarget drug design, Neurodegenerative diseases, Structure-based approaches

1 Introduction

Alzheimer’s (AD) and Parkinson’s (PD) are typical neurodegenerative diseases (NDs) still suffering from inadequate pharmacological treatments, which are mostly based on symptomatic rather than on disease-modifying drugs. The multifactorial nature of NDs, i.e., either from genetic or sporadic causes, contributed in the last years to the consolidation of the so-called multitarget approach in designing new molecular entities. Based on serendipity, this idea stems from the observation that some unwanted side effects typical of drugs could be instead exploited to address off-target desired activities, which could concur to fight and alleviate multifactorial diseases. At the molecular level, a multitarget activity can be pursued by rationally designing ligands, which can simultaneously bias different biological targets.

The multitarget approach for NDs represents still an unprecedented opportunity for researchers since no multitarget agent has reached the market so far. The only drug in advanced clinical trials

(phase III) is ladostigil, a multitarget inhibitor of cholinesterases and monoamine oxidases licensed by Avraham Pharmaceuticals for the treatment of mild cognitive impairment in AD [1]. As a result, the actual advantage of a multitarget approach, compared to a traditional drug cocktail, still needs a validated proof of concept, at least in the field of NDs. Multitarget drug design (MTDD) is rooted on the in-depth knowledge of the molecular interactions of known drugs with the binding site(s) of each target. A naïve approach in the design of multitarget molecules is the building-brick approach, which implies the assembling of two known molecules, each one separately biasing with high affinity a single biological target [2]. Normally, such chimeric structures are obtained by linking the two moieties through a spacer, or by fusing scaffolds of the two parent highly active compounds to obtain a new molecular hybrid. Indeed, making bulkier compounds provided with higher volume and molecular weight could spark some unforeseeable ADME attrition with a higher risk of an undesired lipophilic/hydrophilic balance, poor aqueous solubility, and increased cellular clearance through efflux systems such as P-gp. Mostly related with *druglikeness*, these issues represent a serious concern for the successful discovery and development of multitarget compounds.

MTDD is structured in two basic steps. The first one aims at dissecting highly active compounds engaging specific interactions with different targets to detect their main molecular determinants necessary for such interactions. The second one aims at properly assembling the identified molecular determinants into desired molecular hybrids whose further molecular optimization is however sustained by adequate druglikeness. To this respect, the large availability of high-quality ligand-protein complex crystallographic data and of supportive qualitative and quantitative structure-activity relationship studies has speeded up the development and successful application of *in silico* methods specifically tailored for MTDD.

Computational methods for MTDD derive essentially from the two traditional approaches followed by medicinal chemists, i.e., ligand-based and target (or structure)-based design (Fig. 1). However, the boundary between these two approaches is not so neat after the recent advent of chemoinformatic tools enabling their combined use, especially for the design of new multitarget ligands. In this respect, a key role is still the study of ADMET properties whose place has been anticipated in the early design stage to minimize expensive late failures on the drug discovery and development pipeline [3].

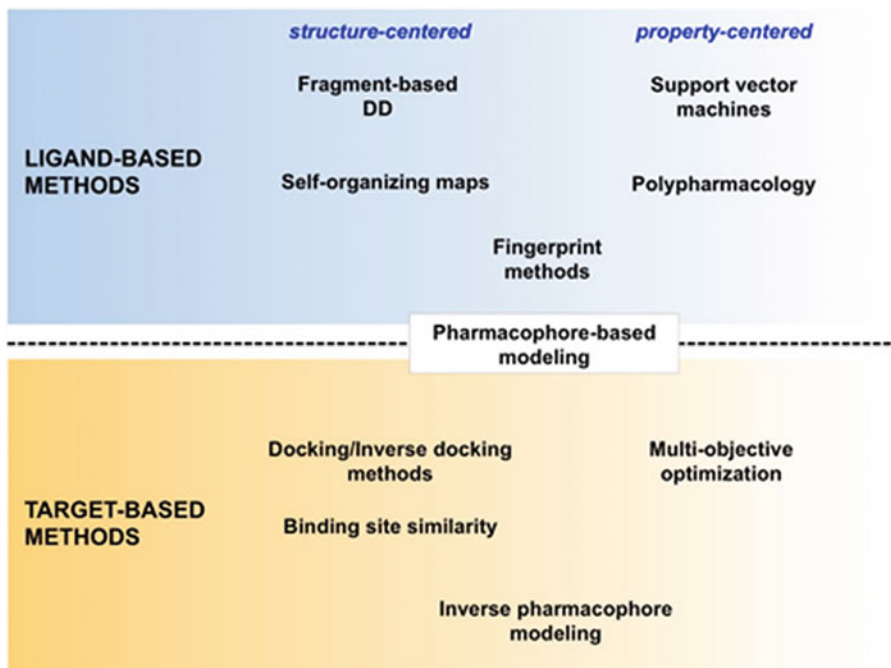


Fig. 1 Scheme of computational methods for MTDD

2 Ligand-Based Methods

2.1 Fragment-Based MTDD

Basically, the fragment-based approach consists of (1) the identification of structural alerts responsible for bioactivity within a pool of known ligands of a given specific target, (2) the detection of sub-structural patterns common to the desired targets, and (3) the incremental growth of fragments to form extended structures able to fit the active site of each target.

At each step, the computer-assisted detection of bioactive fragments, design, and refinement of the growing structure should ensure a rational worksheet where collecting compounds are first virtually screened, then prioritized for chemical preparation (or purchased from public/proprietary libraries), and finally tested.

In this respect, some successful examples of computing packages able to drive the growth of fragments taking into account both chemical feasibility and druglikeness of the evolved structures are LigBuilder [4], SYNOPSIS [5], and DOGS [6].

In the first step, common substructures responsible for affinity on a single target are retrieved from the superposition and comparison of the chemical structures of known ligands of the target. For an efficient identification of fragments, several computational tools have been developed, including alignments based on self-organizing maps (SOM) [7], docking studies [8], and quantification of ligand

efficiency [9]. Taken together, these methods represent a low-cost but very valuable complement to powerful and time-demanding experimental techniques based on biophysical screening methods, which include nuclear magnetic resonance [10], X-ray crystallography [11], surface plasmon resonance [12], and isothermal titration calorimetry [13].

The endless search for unveiling relevant structural features common to diverse biological targets, as well as the efforts toward the molecular decoration of suitable hits, resulted in a number of chemical scaffolds spanning a high grade of affinity for several targets, whence the definition of *privileged substructures* for MTDD. These studies have been carried out on the basis of designing-in strategies [2], taking into account the structural requirements of the active sites involved, and the druglike feature of the designed molecule. An example of versatile scaffold for chemical decoration is represented by coumarin, whose multitarget activity has been extensively studied [14–16]. Starting from previously described selective inhibitors of monoamine oxidase (MAO) [17] and acetylcholinesterase (AChE) [18], the structural requirements for a multitarget activity were merged into newly prepared molecules displaying good dual AChE-MAO inhibition and favorable ADMET parameters (Fig. 2).

Computational methods may involve the evaluation of binding site similarity, through the inspection of enzyme pockets displaying possible interaction points with a given molecule [19].

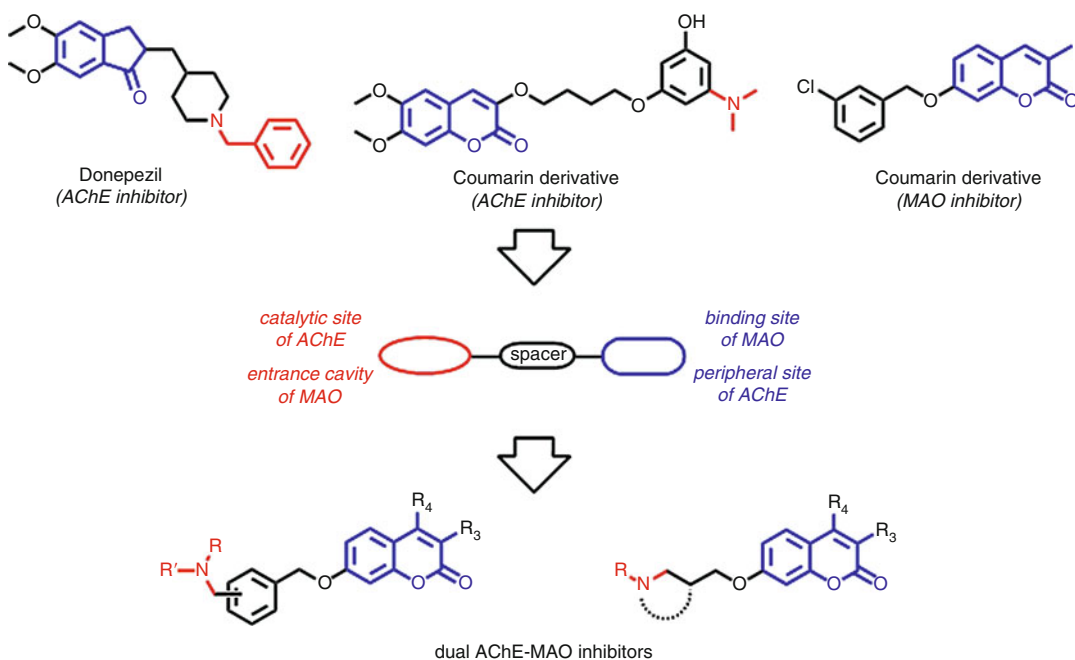


Fig. 2 Designing-in strategy for multitarget coumarin-based inhibitors

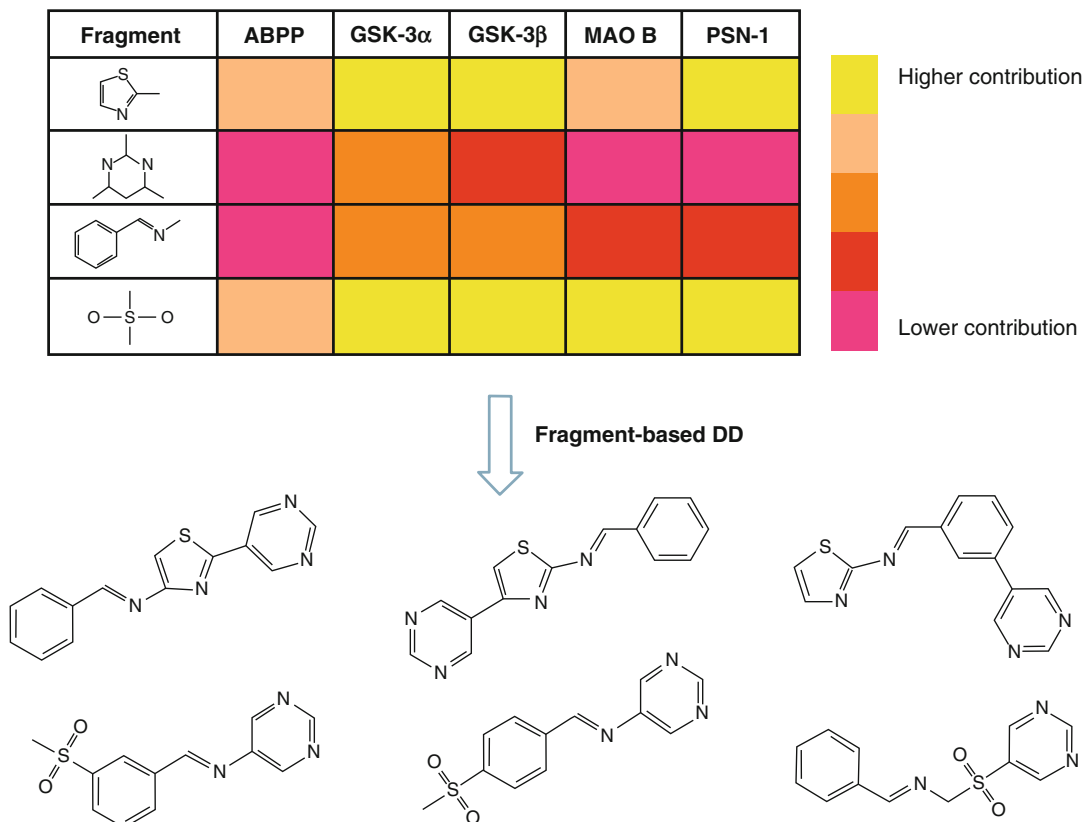


Fig. 3 Design of multitarget ligands from highly recurrent chemical fragments (adapted from ref. 21)

For instance, the SOM approach has allowed the successful alignment of the receptor maps of 5-lipoxygenase and soluble epoxide hydrolase, leading to a multi-SOM model for dual acting fragments. The structural enrichment of one of these core scaffolds was of utmost importance in discovering a multitarget ligand for both enzymes [20]. By a different approach, five targets involved in AD (amyloid β -A4 protein, glycogen synthase kinase-3, glycogen synthase kinase-3 beta, monoamine oxidase B, and presenilin-1) were explored with a library of ca. 500 compounds, and 20 fragments were recognized as highly recurrent [21]. Each fragment contributed in a different extent (i.e., giving higher or lower contribution) to the affinity on each target (Fig. 3). By combining several fragments, possibly showing the highest number of positive contributions, authors designed a pool of virtual molecules, all of them predicted as highly active by the multitarget model.

2.2 SOM-Based Methods

The lack of therapeutic alternatives for the treatment of NDs continuously claims for new molecular entities or, alternatively, for drug repurposing of known drugs. A pursued method for prediction of ND targets for non-ND molecular libraries is the

SOM consensus for macromolecular targets. By combining the SOM prediction with consensus scoring and statistical analysis, Reker and coll [22] developed an *in silico* technique of SOM-based prediction of drug equivalence relationships (SPiDER), able to disclose ND off-target features from known drugs. Application of SPiDER tool to few case studies revealed the high liability of such combination method in predicting off-target activities of known drugs and identifying possible targets of *de novo* designed molecules.

2.3 Machine Learning Methods

Prediction of druglike properties may be performed by means of *in silico* methods evaluating the physicochemical descriptors shared by known classes of compounds/drugs. Since the descriptor space may be very large, particularly in MTDD, there is an increasing need for high-performing computational methods. A statistically reliable filtering is operated by machine learning methods, among which support vector machines (SVMs) allow an iterative selection of the best descriptors for a given target, starting from known ligands, thus improving the predictivity of the computational method.

By applying a recursive feature elimination based on SVM (RFE-SVM) approach, using public databases of drugs used or not used in ND therapy (ND, non-ND drugs), Shahid and coworkers [23] disclosed the main physicochemical descriptors featured by ND drugs and able to discriminate them from non-ND drugs. It is worth of note that among the best-scored descriptors, the most used lipophilicity parameters such as $\log P$ and polar surface area (PSA) were present. The RFE-SVM model was then applied on a data set of non-ND drugs, and several of them were disclosed as potential ND drugs. Indeed, the method may also be applied to the *de novo* design of molecular libraries endowed with potential ND multitarget activity.

2.4 Polypharmacology Prediction

The high-throughput techniques (chemical genetics, next-generation sequencing) boosted by the post-genomic era made available an impressive number of biological data and continuously offer new druggable targets for screening campaigns. The opportunity in exploiting chemical databases or repositioning old drugs can be matched by means of computational tools able to manage such huge amount of data [24].

Predictive models of polypharmacology have been obtained since the early establishment of the multitarget approach by applying classical quantitative structure-activity relationship (QSAR) studies. Basically, QSAR describes biological activity of a defined set of ligands as a function of measurable physicochemical descriptors. Translation of QSAR to MTDD involves the QSAR modeling for each target activity, the analysis of QSAR models in an integrated framework, and then the design of novel multitarget

supposing that similar molecules should bind same targets. The recognition of a similarity pattern may be achieved by considering molecular fingerprints associated to the structures of the query compound and of the ligand set [28]. Fingerprints are encoded as binary strings and compared by means of opportune similarity criteria: the higher is the superposition of fingerprints, the higher the possibility of a molecule to result in a ligand for the targets considered.

3 Target-Based Methods

3.1 *Molecular Docking*

Docking methods are greatly useful in MTDD, provided that the retrieved energy scores of ligand-receptor complexes are comprehensively evaluated for all the targets involved. Molecular dynamics simulations may often strengthen the observations obtained by docking studies [29]. Of course, the increasing availability of crystallographic data for enzymes involved in NDs makes docking-based virtual screening an attractive strategy to find out new multi-target agents. Hit compounds are virtually screened for their interaction (evaluated in terms of posing or binding energy) with the main anchoring points of the active site of each target and ranked on the basis of their predicted affinity. Top-scored hits predicted for each target may be clustered on the basis of their structural similarity or other consensus approaches and finally submitted to the bioassay phase [30, 31].

The MTDD could also benefit of the so-called inverse docking approach, which is conceptually more prone to the exploration of multiple binding sites than classical docking. While docking ranks molecules on the basis of their affinity for a specific target, inverse docking allows the exploration of the interactions of a given molecule with many targets, which are ranked according to their complementarity to the molecule. It is evident that inverse docking requires high calculation performances and that scoring functions needed for targets are not always interchangeable with those used for small molecules. Nevertheless, the improvement of machine potencies and the accurate selection and refinement of scoring functions have led in the last decade to successful research projects in NDs and other diseases [32].

3.2 *Binding Site Similarity*

Target prediction may also be performed by disclosing similarities in binding sites of different proteins. When proteins differ in sequence and biological function but share a certain grade of similarity of their active sites, they would possibly interact with the same ligands. This binding site similarity approach preliminarily finds out the common geometrical patterns of each target and then

derives the pharmacophore fingerprints required for ligand-receptor comparison. Such procedure works well with the use of classical pharmacophore-deriving programs such as Fingerprint for Ligands and Proteins (FLAP) algorithm [33].

3.3 Inverse Pharmacophore Modeling

Pharmacophore models may be generated through target-based methods by the inspection of target-ligand complexes and by the mapping of the interaction site. Several chemoinformatic tools, for instance, ZINCPharmer [34] and CAVITY [35], perform ligand-binding site detection and geometry-based analysis and predict the potential of a target to be recognized and bound by a small molecule (*ligandability*) or more precisely by a lead to be developed as a drug (*druggability*). Pharmacophores are generated on the basis of robust physicochemical parameters, i.e., hydrogen bond donor/acceptor sites, positive/negative charges, and hydrophobicity. PharmMapper, a well-trained computational tool, available online, compares the fit scores of different pharmacophore models with a query compound and generates a list of possible targets on the basis of a probability-based ranking score [36].

4 Combination Methods

Examples of virtual screening combining structure-based and ligand-based techniques give a glance on their potential in MTDD. A common procedure would combine the main structural information, obtained by docking studies on target proteins, with the common pharmacophore mapping derived from a set of known ligands for each target. By applying this procedure, the docking exploration of the active sites of leukotriene A₄ hydrolase and of phospholipase A₂, two targets involved in neuroinflammation, was integrated and combined with a common pharmacophore model of affinity. A pool of molecules retrieved from a public chemical database was found having multitarget activity on the two targets [37].

By a different procedure, a multitarget pharmacophore analysis of three main targets for AD, i.e., acetylcholinesterase (AChE), beta-secretase (BACE), and beta-amyloid protein (A β), was combined with the virtual screening of chemical databases, and hit molecules were optimized in order to improve the affinity for each target [38]. In the first step, the virtual screening via molecular docking on AChE and BACE and molecular dynamics on the aggregating sequence of A β allowed to define the molecular requirements for an efficient binding. A pool of molecules was then synthesized and tested, leading to the discovery of pleiotropic agents with converging inhibitory activities on AChE, BACE, and A β aggregation (Fig. 5).

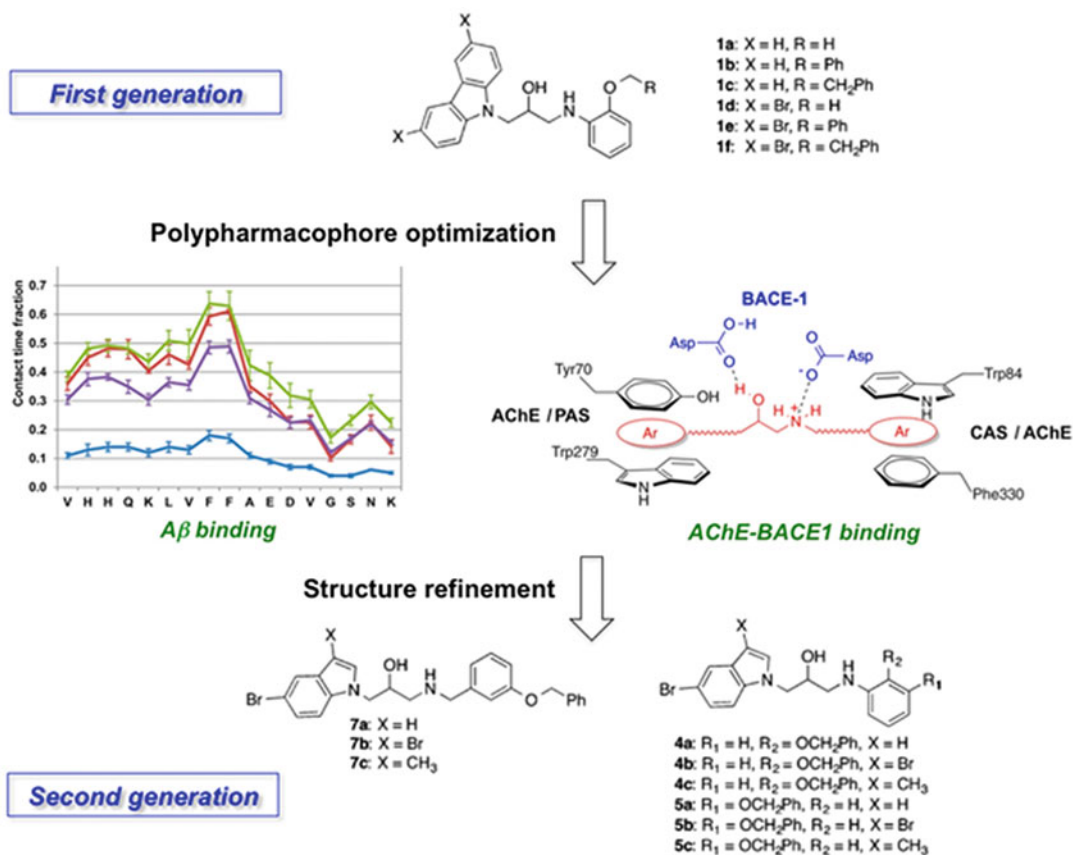


Fig. 5 Combined hit to lead optimization of multitarget inhibitors (adapted from ref. [38], © 2015 American Chemical Society)

The consolidation of MTDD has improved the development of rational drug design methodologies. While in the past researchers focused their efforts in optimizing a single property (*objective*) at a time in single-target molecules, a *paradigm shift* introduced in more recent years tends to privilege, since the early development stages, the overall druglikeness of molecules by means of the simultaneous optimization of several, often conflicting, objectives. The concept of multi-objective optimization (MOO) has emerged as a powerful methodology to this task, allowing a holistic approach to the computer-aided drug design [39–41]. MOO may involve not only the most relevant physicochemical parameters, useful for QSAR, docking, and other virtual screening protocols, but also the ADMET properties and even the synthetic feasibility. Such versatility makes MOO very prone to MTDD, interpreting the multi-*objective* as multi-*target* optimization of the structural requisites for a good correspondence to all the pharmacophores expressed by each target.

5 Conclusions and Perspectives

MTDD represents a powerful approach to the design of novel therapeutic weapons for NDs. Yet, the hybrid molecules assembled by rationally combining diverse pharmacophoric features could present severe issues in their druggability, especially depending from an increased molecular weight and decreased aqueous solubility. In this respect, the recent progress in early stage ADME studies and computational toxicology [42, 43] will contribute to alleviate these concerns of MTDD. However, the way toward a real-life multitarget therapy still needs a proof of its validity from scientists. This makes this challenge highly attractive but explains the disengagement of Big Pharma in MTDD especially for NDs. However, we are confident that an informed use of chem-bioinformatic approach can be of help in bridging the gap still existing between MTDD and the bench of ND patients.

References

1. Weinreb O, Amit T, Bar-Am O, Youdim MBH (2012) Ladostigil: a novel multimodal neuroprotective drug with cholinesterase and brain-selective monoamine oxidase inhibitory activities for Alzheimer's disease treatment. *Curr Drug Targets* 13:483–494
2. Morphy JR, Harris CJ (2012) Designing multi-target drugs. RSC Publishing, London. <https://doi.org/10.1039/9781849734912>
3. Morphy R, Rankovic Z (2006) The physico-chemical challenges of designing multiple ligands. *J Med Chem* 49:4961–4970
4. Yuan Y, Pei J, Lai L (2011) LigBuilder 2: a practical de novo drug design approach. *J Chem Inf Model* 51:1083–1091
5. Vinkers HM, de Jonge MR, Daeyaert FFD, Heeres J, Koymans LMH, van Lenthe JH et al (2003) SYNOPSIS: SYNthesize and OPTimize System in Silico. *J Med Chem* 46:2765–2773
6. Hartenfeller M, Zettl H, Walter M, Rupp M, Reisen F, Proschak E et al (2012) DOGS: reaction-driven de novo design of bioactive compounds. *PLoS Comput Biol* 8:e1002380
7. Digles D, Ecker GF (2011) Self-organizing maps for in silico screening and data visualization. *Mol Inform* 30:838–846
8. Wang T, Wu M-B, Chen Z-J, Chen H, Lin J-P, Yang L-R (2015) Fragment-based drug discovery and molecular docking in drug design. *Curr Pharm Biotechnol* 16:11–25
9. Cavalluzzi MM, Mangiatordi GF, Nicolotti O, Lentini G (2017) Ligand efficiency metrics in drug discovery: the pros and cons from a practical perspective. *Expert Opin Drug Discov* 12:1087–1104
10. Singh M, Tam B, Akabayov B (2018) NMR-fragment based virtual screening: a brief overview. *Molecules* 23:233
11. Hartshorn MJ, Murray CW, Cleasby A, Frederickson M, Tickle IJ, Jhoti H (2005) Fragment-based lead discovery using X-ray crystallography. *J Med Chem* 48:403–413
12. Retra K, Irth H, van Muijlwijk-Koezen JE (2010) Surface plasmon resonance biosensor analysis as a useful tool in FBDD. *Drug Discov Today Technol* 7:e181–e187
13. Recht MI, Nienaber V, Torres FE (2016) Fragment-based screening for enzyme inhibitors using calorimetry. *Methods Enzymol* 567:47–69
14. Farina R, Pisani L, Catto M et al (2015) Structure-based design and optimization of multitarget-directed 2H-chromen-2-one derivatives as potent inhibitors of monoamine oxidase B and cholinesterases. *J Med Chem* 58:5561–5578
15. Pisani L, Farina R, Soto-Otero R, Denora N, Mangiatordi GF, Nicolotti O et al (2016) Searching for multi-targeting neurotherapeutics against Alzheimer's: discovery of potent AChE-MAO B inhibitors through the

- decoration of the 2H-chromen-2-one structural motif. *Molecules* 21:362
16. Pisani L, Farina R, Catto M et al (2016) Exploring basic tail modifications of coumarin-based dual acetylcholinesterase-monoamine oxidase B inhibitors: identification of water-soluble, brain-permeant neuroprotective multitarget agents. *J Med Chem* 59:6791–6806
 17. Catto M, Nicolotti O, Leonetti F, Carotti A, Favia AD, Soto-Otero R et al (2006) Structural insights into monoamine oxidase inhibitory potency and selectivity of 7-substituted coumarins from ligand- and target-based approaches. *J Med Chem* 49:4912–4925
 18. Pisani L, Catto M, Giangreco I, Leonetti F, Nicolotti O, Stefanachi A et al (2010) Design, synthesis and biological evaluation of coumarin derivatives tethered to an edrophonium-like fragment as highly potent and selective dual binding site acetylcholinesterase inhibitors. *ChemMedChem* 5:1616–1630
 19. Milletti F, Vulpetti A (2010) Predicting polypharmacology by binding site similarity: from kinases to the protein universe. *J Chem Inf Model* 50:1418–1431
 20. Achenbach J, Klingler F-M, Blöcher R et al (2013) Exploring the chemical space of multitarget ligands using aligned self-organizing maps. *ACS Med Chem Lett* 4:1169–1172
 21. Speck-Planche A, Kleandrova VV, Luan F, Cordeiro MNDS (2013) Multi-target inhibitors for proteins associated with Alzheimer: in silico discovery using fragment-based descriptors. *Curr Alzheimer Res* 10:117–124
 22. Reker D, Rodrigues T, Schneider P, Schneider G (2014) Identifying the macromolecular targets of de novo-designed chemical entities through self-organizing map consensus. *Proc Natl Acad Sci U S A* 111:4067–4072
 23. Shahid M, Shahzad Cheema M, Klenner A, Younesi E, Hofmann-Apitius M (2013) SVM based descriptor selection and classification of neurodegenerative disease drugs for pharmacological modeling. *Mol Inform* 32:241–249
 24. Lauria A, Bonsignore R, Bartolotta R, Perricone U, Martorana A, Gentile C (2016) Drugs polypharmacology by in silico methods: new opportunities in drug discovery. *Curr Pharm Des* 22:3073–3081
 25. Pang X-C, Kang D, Fang J-S, Zhao Y, Xu L-J, Lian W-W et al (2018) Network pharmacology-based analysis of Chinese herbal Naodesheng formula for application to Alzheimer's disease. *Chin J Nat Med* 16:53–62
 26. Adane L, Bharatam PV, Sharma V (2010) A common feature-based 3D-pharmacophore model generation and virtual screening: identification of potential PfDHFR inhibitors. *J Enzyme Inhib Med Chem* 25:635–645
 27. Evans DA, Doman TN, Thorner DA, Bodkin MJ (2007) 3D QSAR methods: phase and catalyst compared. *J Chem Inf Model* 47:1248–1257
 28. Bender A, Young DW, Jenkins JL, Serrano M, Mikhailov D, Clemons PA et al (2007) Chemogenomic data analysis: prediction of small-molecule targets and the advent of biological fingerprint. *Comb Chem High Throughput Screen* 10:719–731
 29. Mangiatordi GF, Alberga D, Pisani L, Gadaleta D, Trisciuzzi D, Farina R et al (2017) A rational approach to elucidate human monoamine oxidase molecular selectivity. *Eur J Pharm Sci* 101:90–99
 30. Nikolic K, Mavridis L, Djikic T, Vucicevic J, Agbaba D, Yelecki K et al (2016) Drug design for CNS diseases: polypharmacological profiling of compounds using cheminformatic, 3D-QSAR and virtual screening methodologies. *Front Neurosci* 10:265
 31. Zhang W, Pei J, Lai L (2017) Computational multitarget drug design. *J Chem Inf Model* 57:403–412
 32. Chaudhari R, Tan Z, Huang B, Zhang S (2017) Computational polypharmacology: a new paradigm for drug discovery. *Expert Opin Drug Discov* 12:279–291
 33. Cross S, Baroni M, Carosati E, Benedetti P, Clementi S (2010) FLAP: GRID molecular interaction fields in virtual screening. Validation using the DUD data set. *J Chem Inf Model* 50:1442–1450
 34. Koes DR, Camacho CJ (2012) ZINCPharmer: pharmacophore search of the ZINC database. *Nucleic Acids Res* 40:W409–W414
 35. Yuan Y, Pei J, Lai L (2013) Binding site detection and druggability prediction of protein targets for structure-based drug design. *Curr Pharm Des* 19:2326–2333
 36. Wang X, Shen Y, Wang S, Li S, Zhang W, Liu X et al (2017) PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res* 45:W356–W360
 37. Wei D, Jiang X, Zhou L, Chen J, Chen Z, He C, Yang K et al (2008) Discovery of multitarget inhibitors by combining molecular docking with common pharmacophore matching. *J Med Chem* 51:7882–7888
 38. Domínguez JL, Fernández-Nieto F, Castro M et al (2015) Computer-aided structure-based design of multitarget leads for Alzheimer's disease. *J Chem Inf Model* 55:135–148

39. Nicolotti O, Gillet VJ, Fleming PJ, Green DVS (2002) Multiobjective optimization in quantitative structure-activity relationships: deriving accurate and interpretable QSARs. *J Med Chem* 45:5069–5080
40. Nicolotti O, Giangreco I, Miscioscia TF, Carotti A (2009) Improving quantitative structure-activity relationships through multi-objective optimization. *J Chem Inf Model* 49:2290–2302
41. Nicolotti O, Giangreco I, Introcaso A, Leonetti F, Stefanachi A, Carotti A (2011) Strategies of multi-objective optimization in drug discovery and development. *Expert Opin Drug Discov* 6:871–884
42. Mangiatordi GF, Alberga D, Altomare CD et al (2016) Mind the Gap! A journey towards computational toxicology. *Mol Inform* 35:294–308
43. Nicolotti O, Benfenati E, Carotti A, Gadaleta D, Gissi A, Mangiatordi GF et al (2014) REACH and in silico methods: an attractive opportunity for medicinal chemists. *Drug Discov Today* 19:1757–1768



Molecular Docking Studies in Multitarget Antitubercular Drug Discovery

Jéssika de Oliveira Viana, Marcus T. Scotti, and Luciana Scotti

Abstract

Tuberculosis, caused by *Mycobacterium tuberculosis*, is an infectious disease with high levels of mortality worldwide, currently with approximately 6.3 million new cases per year that often present resistance to both first- and second-line drugs. These high rates of incidence are due to several factors including bacterial resistance, AIDS cases, and latent tuberculosis that can reoccur in the patient. Among methods used in the search for new tuberculosis drugs are in silico or CADD (*computer-aided drug design*) studies, which are increasingly being employed in industry and universities. They investigate molecular interactions in order to understand both the structural characteristics of compounds and their activities through virtual manipulation of their three-dimensional (3D) molecular structures, as is the case with molecular docking. Such analyses allow extraction of information and characteristics relevant to compound activity, as well as to predict potential application. In our studies, we discovered antitubercular activity in various derivatives: thiophenes, sulfonamides, chalcones, nitroimidazoles, benzimidazoles, peptides, and quinolones with action in several specific *M. tuberculosis* enzymes. For each derivative, multitarget activity was evaluated in molecular docking studies to select promising compounds with activity(s) against tuberculosis. This chapter will present and discuss molecular docking studies within the bacillus complex, the pharmacological potential of multitarget compounds, and new promising drug candidates with high levels of specificity.

Keywords Antituberculosis, Docking studies, In silico, Multitarget compounds, Tuberculosis disease

1 Introduction

Tuberculosis (TB) is an infectious disease caused by the *Mycobacterium tuberculosis* complex (MTBC) which includes *Mycobacterium tuberculosis* (Mtb), *Mycobacterium bovis*, *Mycobacterium africanum*, and *Mycobacterium canettii* [1, 2]. It is one of the most lethal diseases in the world, accounting for 6.3 million new cases per year, most of which are from Asia, Africa, and the Western Pacific, with China and India alone accounting for 40% of cases. Various public awareness programs, such as the Directly Observed Treatment Short course (DOTS), and the End TB strategy have helped to reduce TB mortality by 47% in the last decade, and within 2030, they aim to reduce TB deaths by up to 90%, and the incidence of TB by up to 80% [3]. Every year, WHO publishes comprehensive

reports on tuberculosis, ensuring updated disease assessments as well as reporting on progress in care and prevention measures.

The increase in tuberculosis cases in recent years has been due to three main factors. The first factor is the increase in the susceptibility of patients with Acquired Immune Deficiency Syndrome (AIDS); co-infection increases the risk of disease by 100 times [4] and contributes to 26% of AIDS deaths in the world [5]. The second factor is increased resistance in Mycobacteria strains, presenting cross-resistance cases for up to nine drugs. Multidrug resistance (MDR-TB) and extensively drug-resistant (XDR-TB) tuberculosis have brought a return to “Global Emergency” status for the disease [6]. The third factor is the already great virulence of tuberculosis, due to the dormant state abilities of the bacillus which can be reactivated at any time.

One of the major challenges in managing tuberculosis occurs when the bacterium is in a state of slow growth or dormancy, its bacterial life cycle. Most drug treatments prioritize active bacillus growth stages, making dormant bacteria both more resistant and less susceptible [7]. It is this persistent state of slow growth or dormancy that allows the bacillus to develop resistance to various types of drugs [8, 9]. In a survey conducted in 2016, there were close to 600,000 new rifampicin resistance cases (RRTB), and 490,000 multidrug resistant (MDR-TB) cases [10–12].

Tuberculosis usually occurs in the lungs, spreading through the air through coughing, sneezing, or spitting, but it also affects other parts of the body [13]. The emergence of MTB strains that are resistant to first-line drugs, (isoniazid, rifampicin, pyrazinamide, ethambutol, rifapentine, and rifabutin), and with growing resistance (Fig. 1), to second-line drugs, (streptomycin, amikacin, kanamycin, capreomycin, and fluoroquinolones), is alarming [14].

Treatment of MTB infection is difficult to follow and usually requires patients to undergo 6–9 months of chemotherapy, or 18–24 months for patients with resistant bacteria. For resistant cases, several drugs are used, but unexpected effects and reduced treatment success have been reported. A cocktail of drugs can generate many side effects, including toxicity [15].

This *M. tuberculosis* resurgence, characterized by a growing number of drug-resistant strains, further emphasizes the need for planning and development of new antituberculosis drugs with activities that decrease resistance mechanisms. Since they block more than one target, drugs with multitarget activity have become key compounds in the search for and development of more effective, promising solutions. CADD presents several experimental advantages: (1) lower expenditure of time and money, using virtual screening through computational studies in structural optimization [16, 17]; (2) greater safety, reducing drug interactions [18]; (3) small and simple to administer doses can be optimized [19]; and (4) effective avoidance of resistance in target proteins [20].

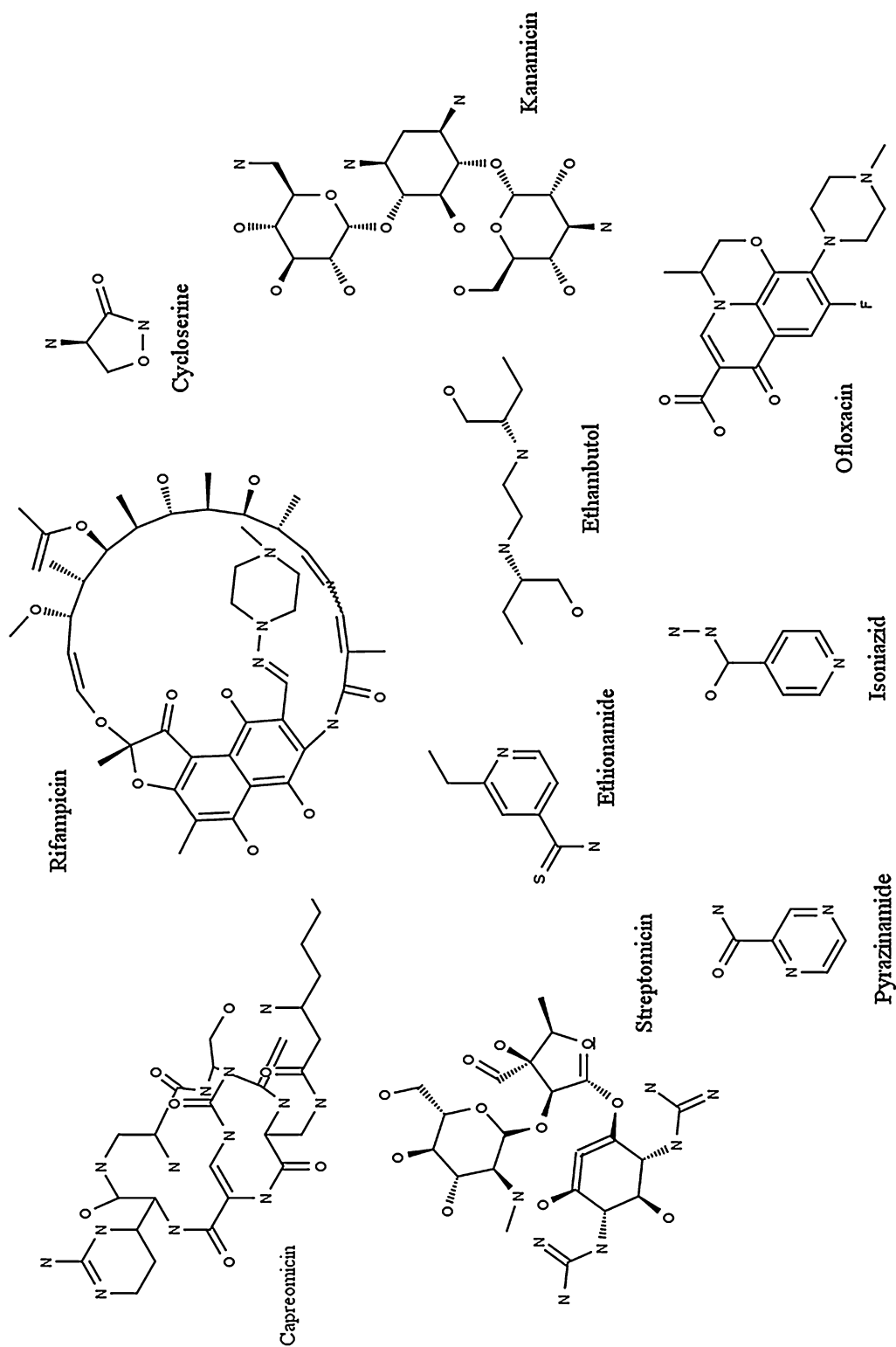


Fig. 1 First- and second-line antitubercular drugs used in therapy

New studies in the field of molecular biology have contributed to the discovery of both resistance mechanisms and new targets. This has been essential to an expansion of investigations into many anti-TB drugs [21–23]. Various protein structure databases are now available including the PDB—Protein Data Bank (www.rcsb.org) [24], which currently contains about 2145 deposited *M. tuberculosis* structures. Recognizing three-dimensional structures is crucial to identify active sites, and potential interactions between drugs and receptors can be used for molecular coupling and in silico screening studies [25].

Strategies that increase drug discovery efficiencies are highly valued and reduce expenditures of both time and money (Fig. 2). Among these strategies, CADD offers advantages that go beyond economics and data simplification; it also accelerates the discovery of compounds with therapeutic action [26]. In addition, it is possible to reconcile CADD studies with experimental studies to coordinate and model data, allowing more directed and immediate information flow.

This chapter will report on molecular docking studies in multi-target compounds with antituberculosis activity. In addition to investigating mechanisms of action and pharmacological potential, the chapter will report on new drug candidate target–ligand binding interactions that may help combat *M. tuberculosis* resistance.

CADD studies include Molecular Docking, a computational technique that seeks to ascertain the interaction affinity of a molecule (ligand) with the active site of its macromolecular (receptor) target. The complex formed involves two characteristics that need to be studied: Fitting, which refers to conformational sampling and orientation of the ligand within the active site of the receptor (Fig. 3); and Score, which correlates the best pose, that is, the best conformation and orientation of a given binder, as classified according to energy value [27].

2 Targets

Through sequencing of the *M. tuberculosis* genome, the fight against tuberculosis has expanded its identification of the proteins essential for mycobacterial survival and thus also antituberculosis compounds and their mechanisms of action [28]. With sequencing, several databases important for drug development have emerged enabling recognition of possible mechanisms of action to block mycobacterial biosynthetic pathways and lead to mycobacterial death [29]. The following will present some of the numerous therapeutic targets, validated in the literature that push development of new candidate compounds for antituberculosis drugs: they involve proteins of the mycobacterial cell wall, metabolism, and the nucleus.

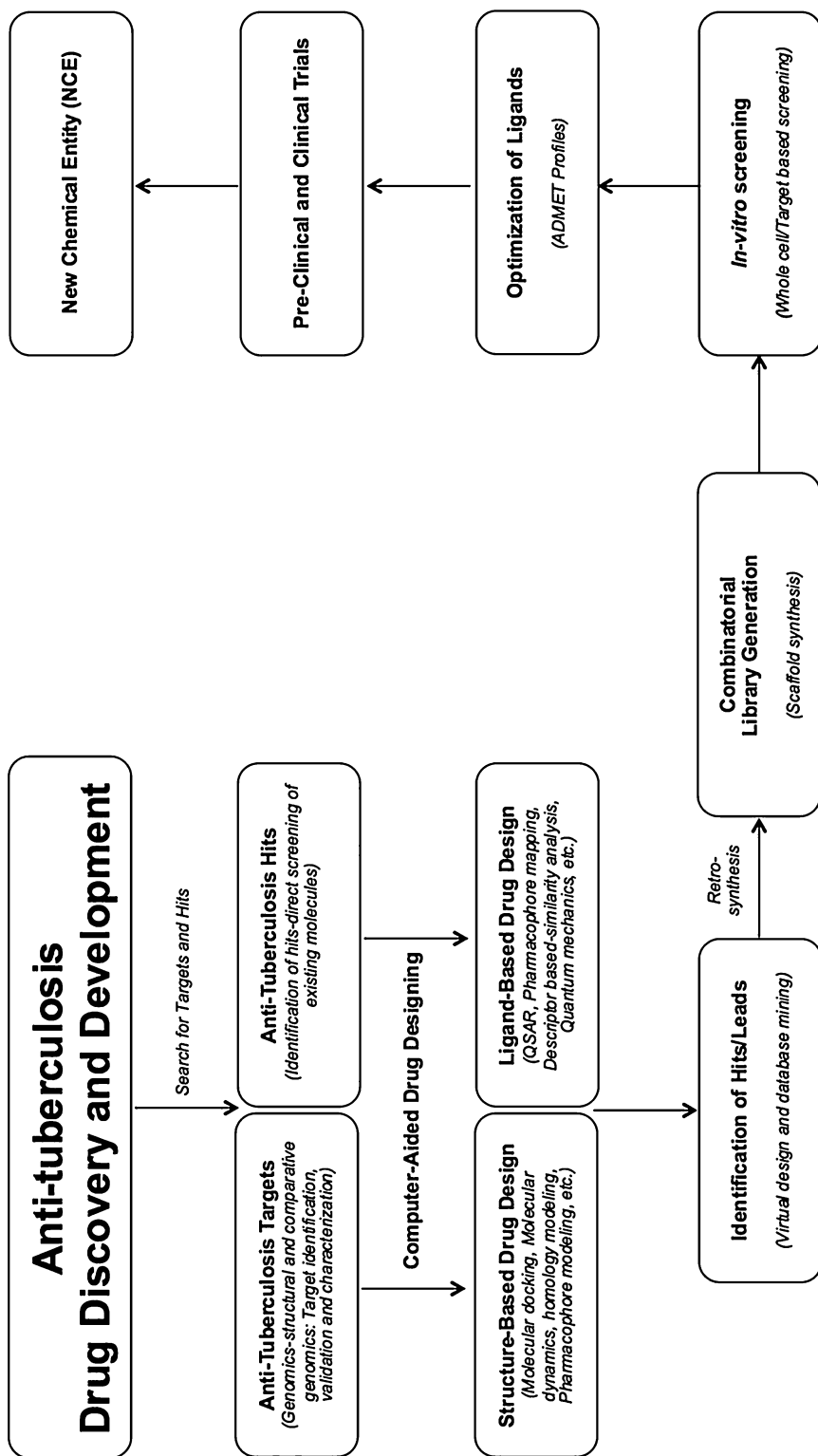


Fig. 2 In silico development and discovery route for new drug candidates against tuberculosis (Chetty et al. [26])

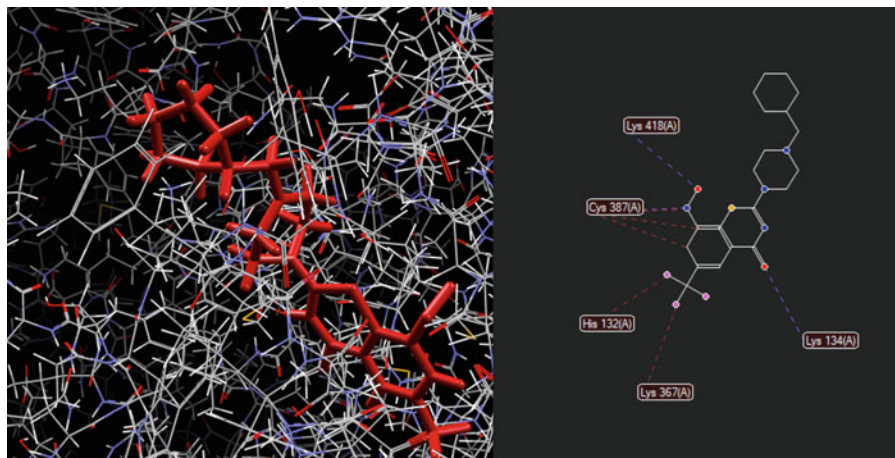


Fig. 3 Molecular docking of the drug candidate piperazinobenzothiazinone derivative PBTZ169 with the enzyme DprE1 (ID PDB 4KW5) through Molegro Virtual Docker 6.0. The Molecular Docking was performed using the setup of the software: GRID resolution 0.30, MolDock Score, ten runs, and five poses. Red lines: steric interactions; blue lines: hydrogen bonds

2.1 DprE1

Blocking of mycobacteria cell wall formation has been frequently used to develop and improve anti-TB drugs. The *M. tuberculosis* cell wall presents six external lipid layers: polysaccharides (arabinogalactan), lipoarabinomannan (LAM), peptidoglycan, mannoside phosphatidylinositol, a plasma membrane, and mycolic acid [30]. The composition of the mycobacteria cell wall is highly hydrophobic, favoring protection against external agents, including pathogens, the immune system, and the action of many antibiotics, such as isoniazid, ethambutol, ethionamide, D-cyclo-serine, PA-824, OPC67683, BTZ, and SQ109 [31].

The flavoenzyme decaprenyl phosphoryl- β -D-ribose-2'-epimerase, also known as DprE1 is necessary for formation of the *M. tuberculosis* cell wall. Within the protein machinery a reaction complex forms (Fig. 4) and acts together with DprE2 to catalyze the epimerization of decaprenyl phosphoribose (DPR) to decaprenyl-phospho-D-arabinofuranose (DPA) through an intermediate formation of decaprenyl-phosphoryl-2-keto-ribose (DPX). In the mycobacterium cell wall, DPX is the only synthesis precursor of arabinogalactan and lipoarabinomannan (LAM) [32].

In comparison with the species *Mycobacterium smegmatis*, structural homology studies of the DprE1 protein have been performed finding 83% similarity between their configurations [33, 34]. In PDB, the DprE1 enzyme crystalline structural form has been described as BTZ043 and TCA1, two benzothiazinone derivatives [35]. Descriptions of such crystallized structures have contributed to a better understanding of the mechanisms of action and inhibition of these compounds. Tiwari et al. [36] showed that BTZ, one of the most promising inhibitors currently studied, can

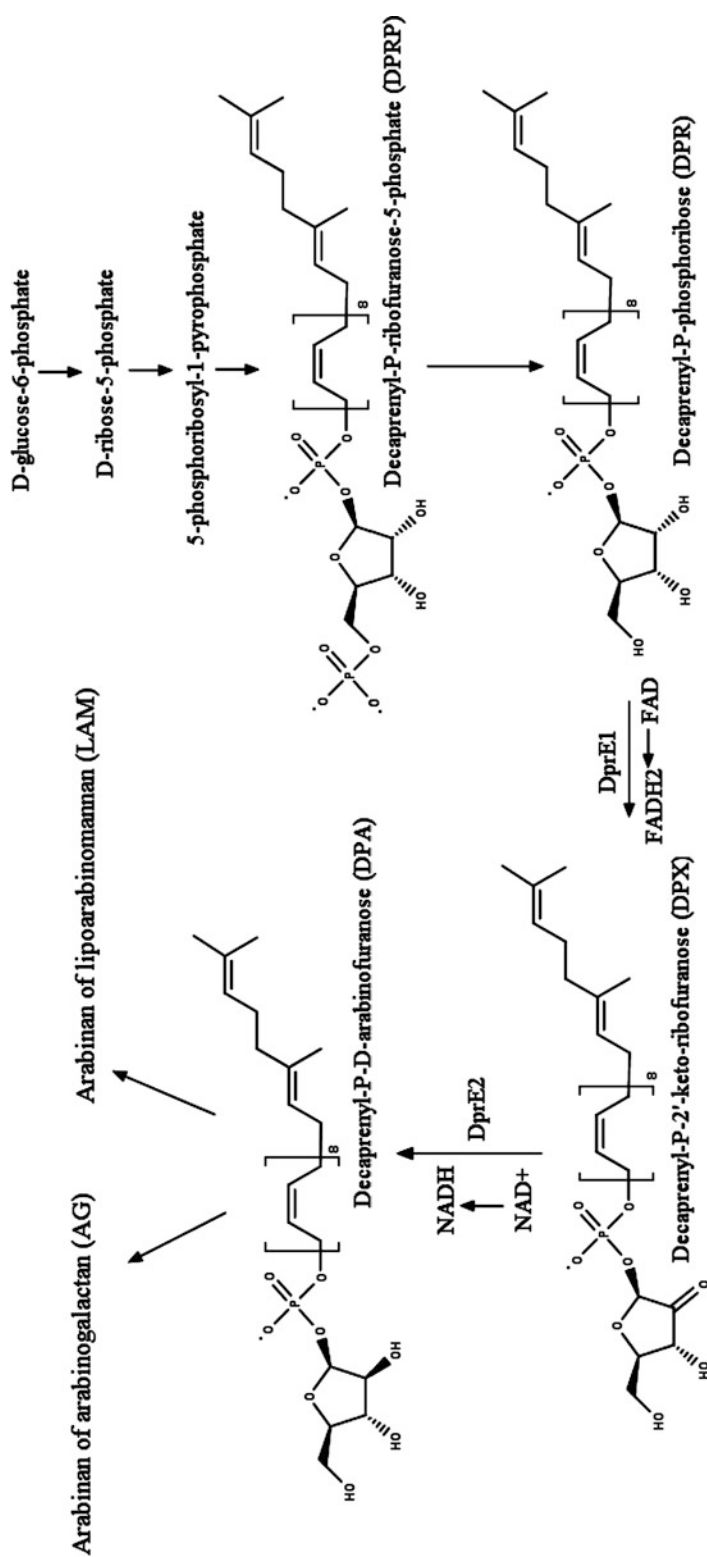


Fig. 4 Metabolic pathway involving the enzyme DprE1 for the synthesis of decaprenylphosphoryl arabinose which is responsible for the formation of the cell wall in mycobacteria

bind to the active site through cysteine. Thiolates can induce non-enzymatic nitro group reductions through addition to aromatic carbons in these compounds. BTZ compounds act as a substrate, converting the DprE1 enzyme to its reduced form through a nitro-reduction mechanism and forming a nitro-derivative in the DprE1 active site [37–39]. This presents a novel mechanism for DprE1 nitro-aromatic inhibitors.

2.2 Mur Ligase

Proteins belonging to the Mur ligase family catalyze non-ribosomal peptide bonds to form peptidoglycans [40–42]. For this reason, in the search for new promising TB targets, they attract attention [43, 44]. ATP-dependent murine ligase proteins play a key role in cell wall biosynthesis and present high specificity for their substrate.

The proteins share only a 37% structural and functional similarity with *Escherichia coli* proteins, which makes it difficult to determine their exact functions. However, what we do know is that these proteins are monomeric, presenting three conserved catalytic domains which vary according to ligase type (MurC, MurD, MurE, and MurF). These have been characterized as nucleotide-binding domains for ATP binding and substrate activation [45].

In peptidoglycan (PG) biosynthesis, the MurC biochemical pathway starts with conversion of UDP-GlcNAc to UDP-MurNAc by enolpyruvyl transferase-MurA, and a flavin-dependent reductase MurB (Fig. 5). The MurC protein then adds the amino acid L-Ala to form UDP-MurNAc-L-Ala [47]. The MurD protein catalyzes an amide bond between D-Glu and UDP-MurNAc-L-Ala to form UDP-MurNAc-L-Ala-D-Glu. MurE acts through addition of a third highly specific peptide residue that involves maintaining the integrity of peptidoglycans in the cell wall; if not maintained, cellular lysis occurs [48]. Finally, MurF adds the dipeptide D-Ala-D-Ala to UDP-MurNAc-L-Ala-D-Glu-(mA2pm/L-lys) to form the cytoplasmic PG precursor: UDP-MurNAc-L-Ala-D-Glu-(mA2pm/L-lys)-D-Ala-D-Ala [46]. This dipeptide is an essential factor for PG formation, such that its binding energizes the binding of the glycan chain [49].

2.3 Deazaflavin-Dependent Nitroreductase

Due to its enzymatic activity, the enzyme Rv3547 is named Deazaflavin-dependent nitroreductase (Ddn). Protein studies have reported that Ddn may be a peripheral membrane protein [50], but complete substrate data and functions remain unclear, since it presents low sequence homology with similar proteins of the same function [51]. Homologs do exist in several species and are commonly found in PA-824-resistant Actinobacteria [52]. These proteins form as soluble aggregates making crystal formation rare and structural elucidation difficult.

Thus, despite numerous studies reported in the literature, it has been difficult to obtain enough full-length Ddn crystals or NMR spectra. What is known is that this protein has 151 amino acids with

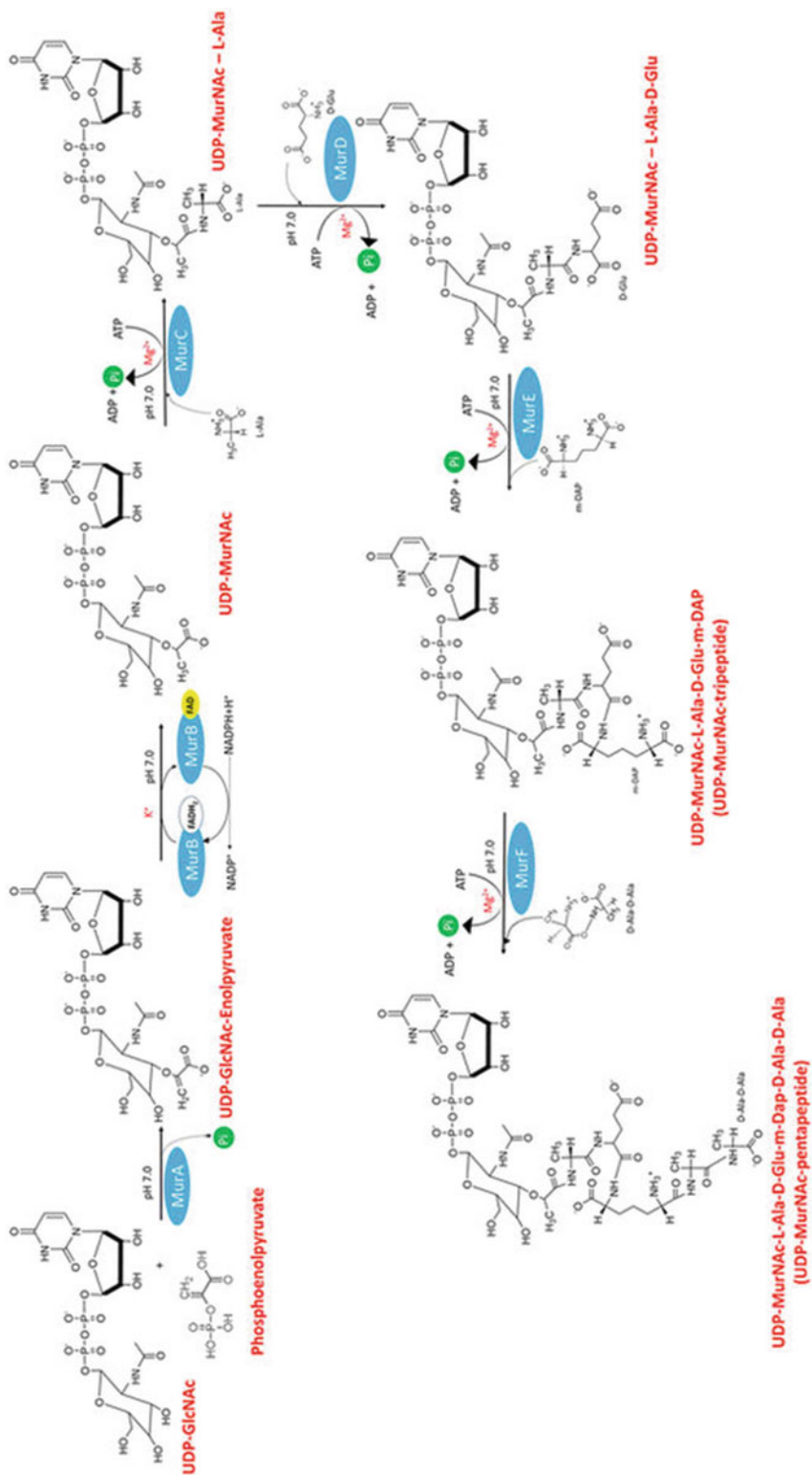


Fig. 5 Biochemical pathway of the Mur ligase enzymes, represented on the basis of in vitro activities (Eniyan et al. [46])

only one cysteine residue (in region 149), that is, the existence of only one intermolecular disulfide bond throughout the structure. However, its biochemical pathway is still unknown. Yet, in the presence of nitroimidazole, PA-824, which is already in Phase II clinical studies, has antituberculosis activity. As reported in the literature, its action on mycobacteria is promoted through enzymatic activation provided by Ddn in *M. tuberculosis*. Thus, it catalyzes PA-824 reduction, resulting in a product that releases intracellular reactive nitrogen species (RNS) (Fig. 6), compromising protein and lipid synthesis while promoting antimycobacteria activity [53].

Homologous protein sequences have been reported for actinobacteria; however, there are few confirmations of cellular function for this target. In *M. tuberculosis*, Ddn was first seen in a study of membrane proteins, which may potentially have associated functions [54]. The genes Rv1261c, Rv1558, and Rv3178 are homologues of Ddn present in the Mtb bacillus. Together with orthologs present in other species of mycobacteria, they form a class of nitroreductases, yet without delineation of their physiological mechanisms.

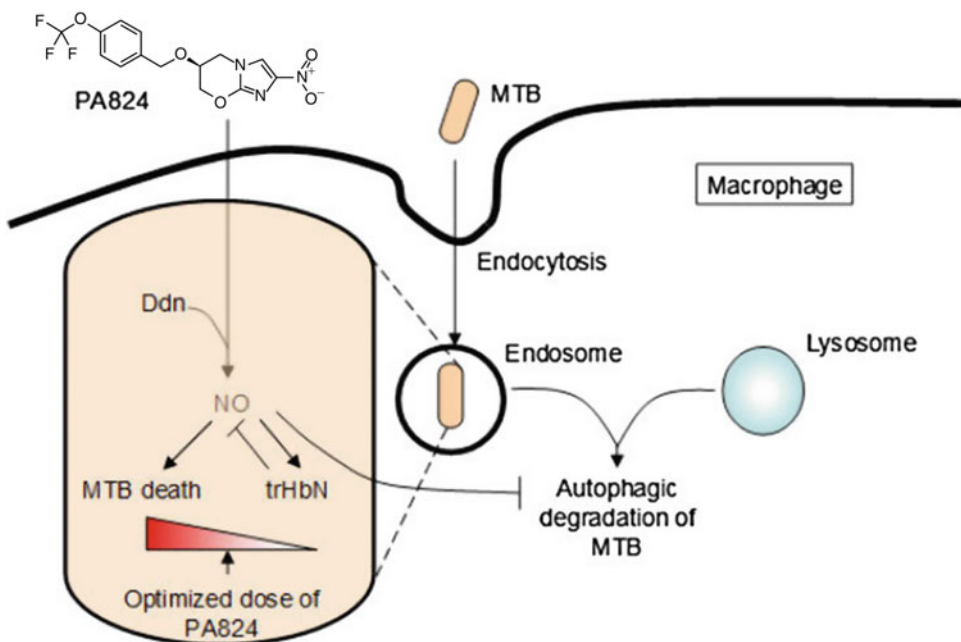


Fig. 6 Representation of the biochemical pathway associated with PA-824 and the enzyme Ddn, whose activity favors NO production and consequent *Mycobacterium tuberculosis* cell death (Somasundaram et al. [53])

2.4 *N*-Acetylglucosamine-1-Phosphate Uridyltransferase

The protein *N*-acetylglucosamine-1-phosphate uridyltransferase (GlmU) is known as a bifunctional enzyme that catalyzes the formation of UDP-*N*-acetylglucosamine (UDP-GlcNAc), which in bacteria is closely involved in the peptidoglycan pathway, and lipopolysaccharides [55]. The reaction occurs as GlmU catalyzes acetyl-CoA group transfer to glucosamine-1-phosphate (Fig. 7), towards the final formation of *N*-acetylglucosamine-1-phosphate (GlcNAc-1-P).

For being an important precursor for lipopolysaccharide biosynthetic pathways, this protein becomes an attractive target for new mycobacterial agents. Structurally, GlmU is a 49-kDa cytoplasmic protein with a dimer/trimeric solution [56], and maximal activity against magnesium and phosphate ions. The enzyme is delineated by two domains: the first is the N-terminus of pyrophosphorylase, which presents similar identity to other pyrophosphorylase enzymes, and the second is the acetyltransferase domain with hexa-peptide repeats, as also reported in other acetyltransferases [57]. In homology studies, GlmU presented sequence identities of 69% with *Haemophilus influenzae* [58] and 42% with *Bacillus subtilis* [59]. Other studies performed with GlmU have reported

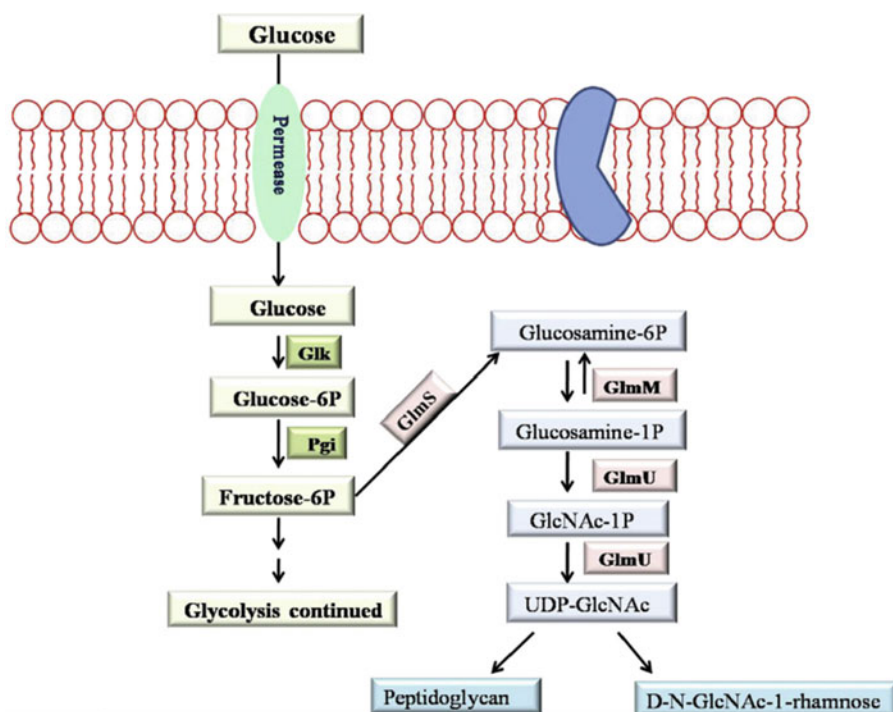


Fig. 7 Graphical representation of the Glucose pathway in production of peptidoglycan by enzymes belonging to the glucosamine-phosphate family: *Glk* glucokinase, *Pgi* phosphoglucose isomerase, *GlnS* glucosamine-6-phosphate synthase, *GlnM* phosphoglucosamine mutase, *GlmU* glucosamine-1-phosphate acetyltransferase, and *N*-acetylglucosamine-1-phosphate uridyltransferase (Rani et al. [55])

inhibition with specific thiol reagents, which indicates that cysteine residues may take part in the acetyl transfer.

In *M. tuberculosis*, mtGlmU acts as a functional enzyme, involved in synthesis of GlcNAc-1-P via acetylation of Glc-1-P in the C-terminal domain also known as the acetyltransferase domain [60]. Upon product formation, it is carried to the N-terminal domain, a portion of the uridylyltransferase site which joins with UTP for production of UDP-GlcNAc, which itself acts upon formation of peptidoglycans in the mycobacterial cell wall [61].

2.5 Pks13

PKS proteins are part of a group of multifunctional enzymes that act in the biosynthesis of innumerable natural compounds known as polyketides. This biosynthesis resembles that of fatty acids (Fig. 8), and involves decarboxylative condensation of two long-chain fatty acid derivatives; the very long (C40–C60) meromycoloyl-AMP and a shorter (C24–C26) 2-carboxyacyl-CoA [62, 63]. For these steps, several PKS protein domains are involved: an acyltransferase (AT) domain (with selection and transferring action for the extender unit); a domain of the ACP carrier protein (acyl carrier protein) activated by phosphopantetheine which transfers its unit; and a third domain called β -ketoacyl synthase (KS) which involves

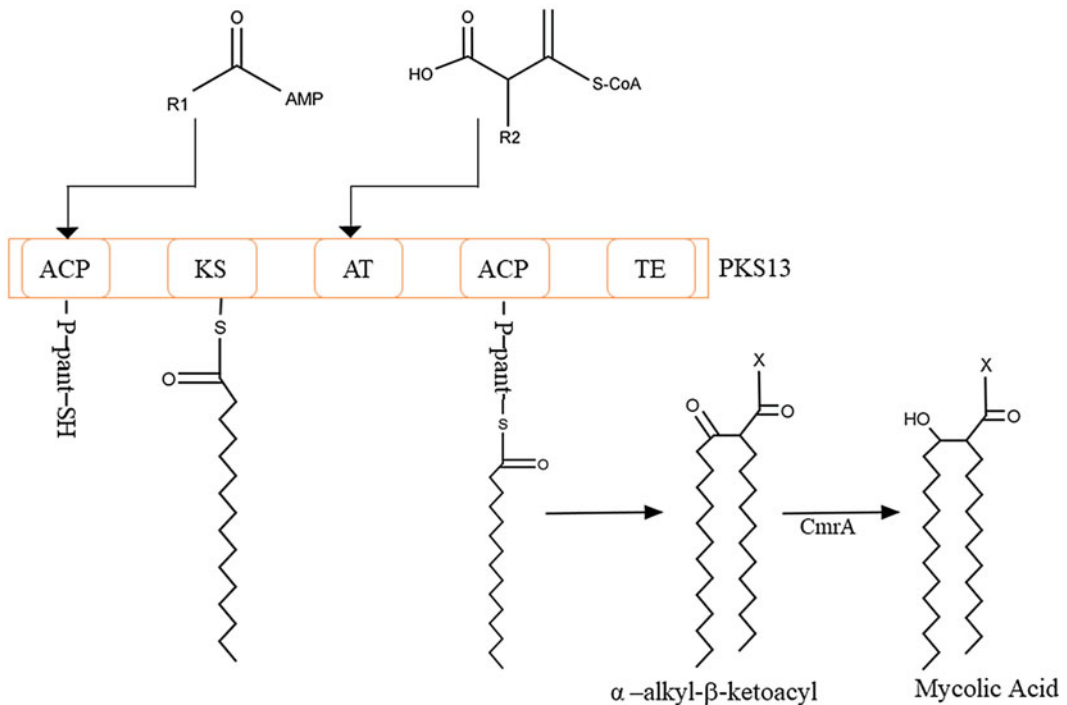


Fig. 8 Biochemical pathway involving action of the Pks13 protein for synthesis of mycolic acids in mycobacteria

the condensation of starting and extension units, forming the β -ketoacyl intermediate [64, 65].

It is important to note that all of these domains belong to a single polypeptide chain, and studies concerning structural elucidation of the PKS type I enzyme group are difficult to characterize because of the intrinsic flexibility of the protein. At this point, several structure and function relationships for these enzymes have been proposed based on FAS (fatty acid synthase) type I enzymes via KS and AT units [66, 67]. PKS13 is intriguing precisely because it is highly specific to its substrates, and performs only one condensation cycle between its two unusually long substrates.

In *M. tuberculosis* studies, the genome sequence encoding PKS type I contains about 20 PKS genes, and among them is PKS13 which acts on the final assembly of mycolic acid, a structural component of mycobacteria. Structurally, MtbS PKS13 has 1733 amino acid residues in a molecular mass of 186,446 Da, with three obligate PKS domains; in addition to an ACP in the N-terminal region, and a thioesterase (TE) domain in the C-terminal. These five domains are bound by 30–200 amino acid residues. In summary, the AT domain is responsible for loading the carboxyacyl-CoA extension unit, which is subsequently transferred to the ACP domain of the C-terminal region. The meromycolyl chain is activated and carried by the FadD32 enzyme, from which it is transferred to the KS domain of the PKS13 enzyme. This domain acts on meromycolyl and carboxy-acyl chain condensations in the production of an α -alkyl β -keto thioester linked to the C-terminal ACP domain. Several current studies point to the use of this protein as a new target when planning antituberculosis compounds [68].

2.6 MtFabH

The mycobacterial wall includes various macromolecular structures (enzymes involved in biosynthetic pathways) that are targets for development of antimycobacterial compounds. Fatty acid synthases (FAS) present in mycobacteria may be grouped into two types [69]. Type I synthases, which include multifunctional polypeptides with chain elongation functions, are commonly found in both yeast and mammals. Type II synthases whose chain elongation reactions are performed by other classes of enzymes and initiated by β -ketoacyl-ACP synthase III (ecFabH), which is also responsible for promoting condensation of acetyl-CoA with malonyl-ACP [70], are found in bacteria and plants.

The purpose of the mtFabH enzyme reaction mechanism is mycolic acid biosynthesis and fatty acid elongation. The fatty acids synthesized by Type I synthases are removed by means of a transacylase, producing acyl-CoA. The free acyl-CoA is used to promote synthesis of glycerophospholipids and lipids, but it can also be used by the enzyme mtFabH (Fig. 9), where further elongation processes, promoted by type II synthases, occur in mycolytic acid production units in the mycobacterial cell wall [71].

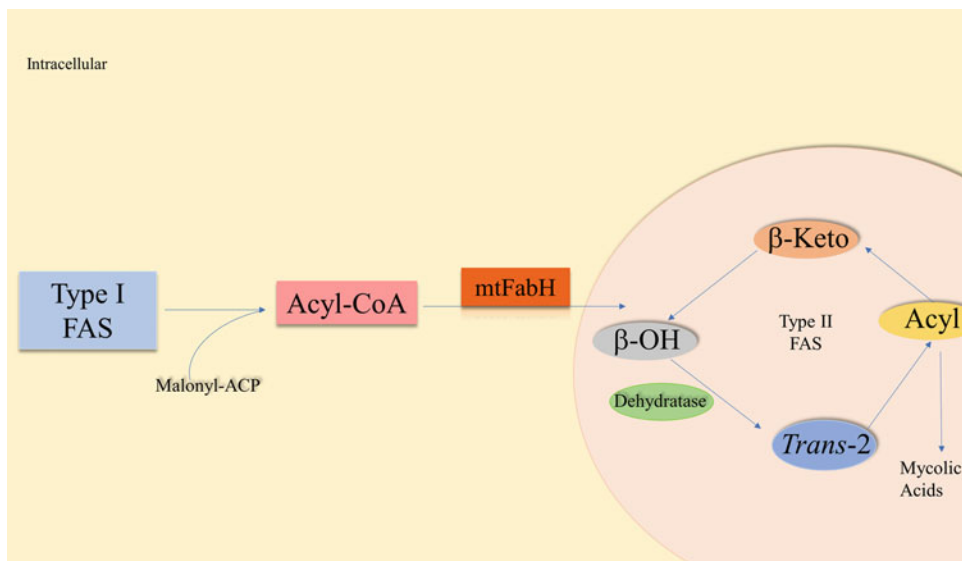


Fig. 9 *Mycobacterium tuberculosis* FabH enzyme fatty acid elongation pathway

By kinship or homology, FabH enzyme reaction mechanisms present in *M. tuberculosis* can occur in the bacillus genome. A 46.8% identity is seen with the *E. coli* FabH enzyme, and 42.9% with *B. subtilis* FabH2 [72]. This degree of similarity is found mainly in the residues binding the CoA portion, and indicates clear reaction mechanism similarities with mtFabH.

2.7 Pantothenate Synthase

For synthesis of coenzyme A and other acyl carrier proteins in energy and fatty acid metabolism, and still other cellular processes, pantothenate (also known as vitamin B5) is a precursor compound [73]. Both plants and microorganisms synthesize it; yet, only mammals obtain this nutrient through the diet. Its synthetic route thus offers many effective target enzymes for developing drugs with antimicrobial activity.

The biosynthetic pathway of pantothenate involves four stages, catalyzed by the PanB, PanC, PanD, and PanE genes [74]. These genes encode the pantothenate synthetase (PS) enzyme which catalyzes the last step in pantothenate biosynthesis (Fig. 10) involving condensation of pantoate, and ATP with alanine to form pantothenate [75].

In kinetic *M. tuberculosis* pantothenate synthetase (mtPS) studies, it was observed that this enzymatic action is performed in two steps: formation of an ATP-pantoate intermediate and pantoyl adenylate, and soon afterwards a nucleophilic attack on β-alanine to form pantothenate and AMP [76].

Proteins constructed from the PanC gene of various organisms, such as *E. coli* [77], in plants [78], *Fusarium oxysporum* [79] and *M. tuberculosis* [80], have been characterized. In these studies, the

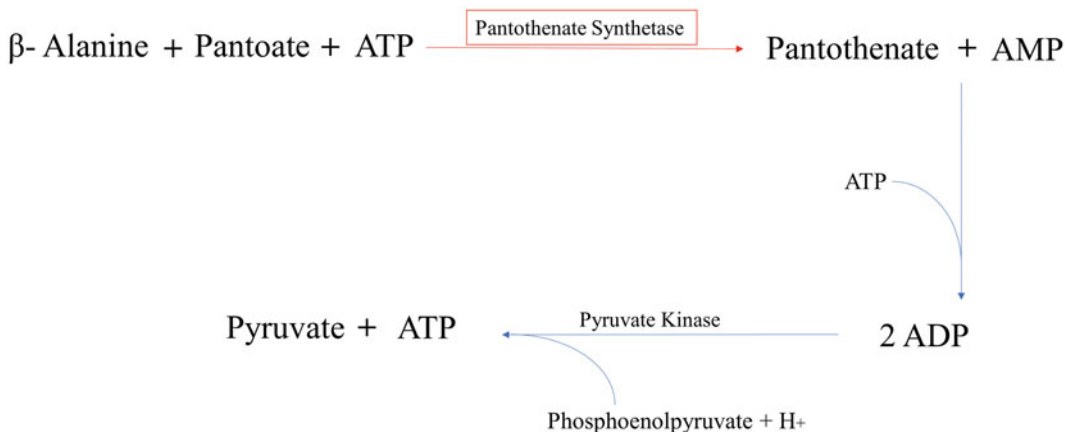


Fig. 10 Schematic diagram representing the action of Pantothenate synthetase in *M. tuberculosis* production of pantothenate

formation of dimers in solution is reported. The crystallized structure of the *M. tuberculosis* PS enzyme (mtPS) evidences the presence of folds, as seen in the *Escherichia coli* PS (ecPS) enzyme, and also presents substrate complexes that confirm the binding position of ATP and pantoate, as also reported for the ecPS enzyme. However, differences were found in the domains of each subunit, which presented closed conformations, a characteristic not found in *E. coli*. The structure also has a flexible region, which forms close to the active site cavity, and acts as a gateway to the mtPS enzyme [80]. Analysis of the crystalline structure of PS, complexed with the reaction intermediate, depicts strong interactions of pantoate adenylate with the active site, allowing for stability of the already highly reactive compound [81]. Several studies have reported pantothenate synthetase inhibition using new analogs of pantoate adenylate reaction intermediates, which can be used as agents against tuberculosis [82].

2.8 Ptp A and B

Protein tyrosine phosphatases (PTPs) are a part of a large family of approximately 100 types of enzymes, presenting great diversity in cellular processes, and being essential for phosphate hydrolysis of tyrosine residues (Fig. 11) present in proteins [84]. These are characterized as proteins present in many tissues, and belonging to a group of low molecular weight tyrosine phosphatases which have also been identified in other bacteria [85, 86]. PTPs contain, in addition to a structure composed of 240 amino acids and a precisely conserved domain, a consensus sequence at their catalytic sites. In these catalytic sites, a pattern of cysteine and arginine provides an essential contribution to enzymatic action [87, 88].

During the infectious process, bacteria present many tyrosine phosphatases [89]. Among them is low molecular weight PtpA, also present in *M. tuberculosis* and belonging to the tyrosine

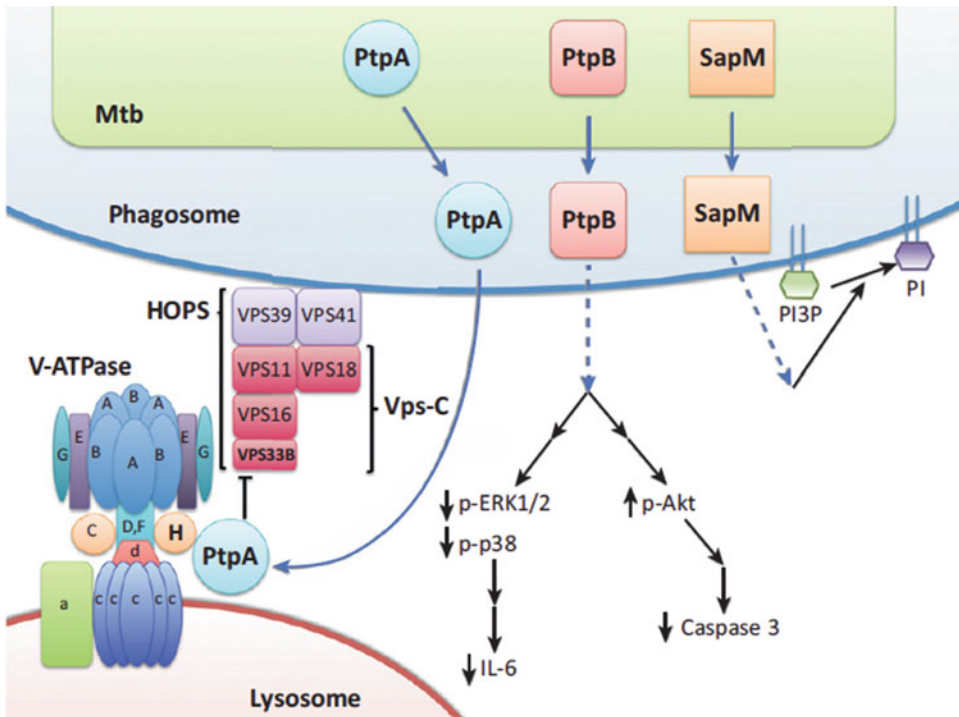


Fig. 11 Biochemical representation of the enzymatic action of PtpA and PtpB phosphatases in *M. tuberculosis* (Wong et al. [83])

phosphatases [83, 90], and it also mediates host pathogenicity. Homology studies were performed for PtpA, and a close similarity to that of *Streptomyces coelicolor* was found, where it hydrolyzes free phosphotyrosine residues. This provided an evidence of its involvement in regulation of sulfur amino acid metabolism [91]. Detailed mechanisms concerning *M. tuberculosis* PtpA are essential for understanding the bacillus virulence [90].

2.9 *InhA*

M. tuberculosis contains mycolic acids formed by chain α -alkyl, β -hydroxy fatty acids (60–90 carbons) [92], and because it is a critical biosynthetic pathway many studies have regarded the enzyme *InhA* as a possible target for the new antimycobacterial agents [93]. The biosynthetic pathway is catalyzed by several monofunctional enzymes from the FAS II group. These differ from FAS I present in mammalian cells, where all enzymatic activities are assembled into one or two multifunctional polypeptides. Such differences present in bacteria and mammalian cells are essential for design of novel inhibitors that act with increased activity, selectivity, and low toxicity. FAS I products are transferred to the FAS II pathway, i.e., C16–C26 fatty acids, where they are elongated to very long-chain C56 [94].

In *M. tuberculosis*, the elongation process is a key catalytic process for producing mycolic acid precursors and depends on the InhA enzyme. The InhA gene product is characterized as an NADH-dependent enoyl-acyl carrier protein reductase that acts in reducing 2-trans enoyl chains. These possess at least 12 carbon atoms [95, 96] and are the last step in the biosynthetic pathway of fatty acid elongation (Fig. 12). Several studies have reported the production of this InhA protein in bacteria [98, 99]. It is involved in the production of very long-chain fatty acids, mycolic acids, and is an essential component in the formation of the mycobacteria cell wall.

The InhA enzyme belongs to the short-chain dehydrogenase/reductase (SDR) family. In aqueous solution, the InhA crystal is a homotetramer structure with three perpendicular twofold symmetry axes that cross the center of the molecule [100]. Each enzyme subunit is expressed in a domain with a central nucleus containing a Rossmann fold as an NADH binding support. In the inner region, several α - and β -strand helices extend from the Rossmann fold, known as the substrate binding loop to create a slit for insertion

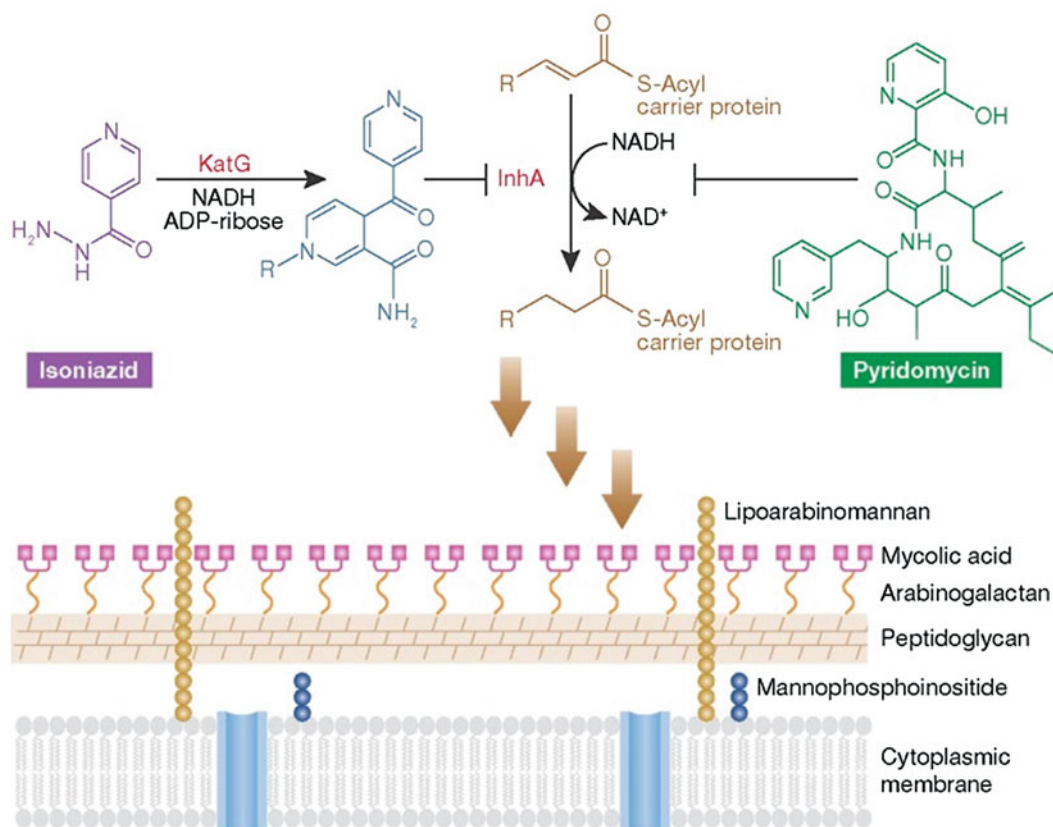


Fig. 12 Schematic representation of enoyl acyl carrier protein reductase InhA inhibition in the *M. tuberculosis* mycolic acid pathway (Wright [97])

of the fatty acid substrate. The extension contains two helices which form one side of the fatty acid cleavage slit.

Isonicotinic acid hydrazide (INH) and ethionamide inhibit this enzyme, though in *M. tuberculosis* several *InhA* gene mutations have been associated with resistance to isoniazid [98]. However, these resistance cases are found in prodrug activation processes (i.e., the catalase-peroxidase enzyme KatG) [97, 101], and in only one enzyme, enoyl-ACP reductase, while other catalytic processes of the FAS II pathway are conserved. The enzyme *InhA* therefore still remains a target in the search for new drugs.

2.10 Cytochrome P450 Monooxygenase

The cytochrome P450 monooxygenase (P450) belongs to the family of heme-containing monooxygenase enzymes [102]. They are responsible for many metabolic roles, such as detoxification and synthesis of steroids in mammals [103], and are also found in other organisms. P450 catalyzes other oxidative mechanisms, but its main activity is the oxygen transfer reaction, catalyzing reduction of oxygen reversibly bound to iron in P450 [104]. This reaction removes an oxygen atom from oxygen (O_2), and promotes its transfer to an organic substrate with parallel formation of a water molecule (Fig. 13).

This above process is essential for the survival of microorganisms, including *M. tuberculosis*; it regulates several other biochemical pathways that are essential to bacillus pathogenicity. Due to its good solubility, the enzymatic structure has been reported in several structural elucidation studies [106]. Although studies in *E. coli*

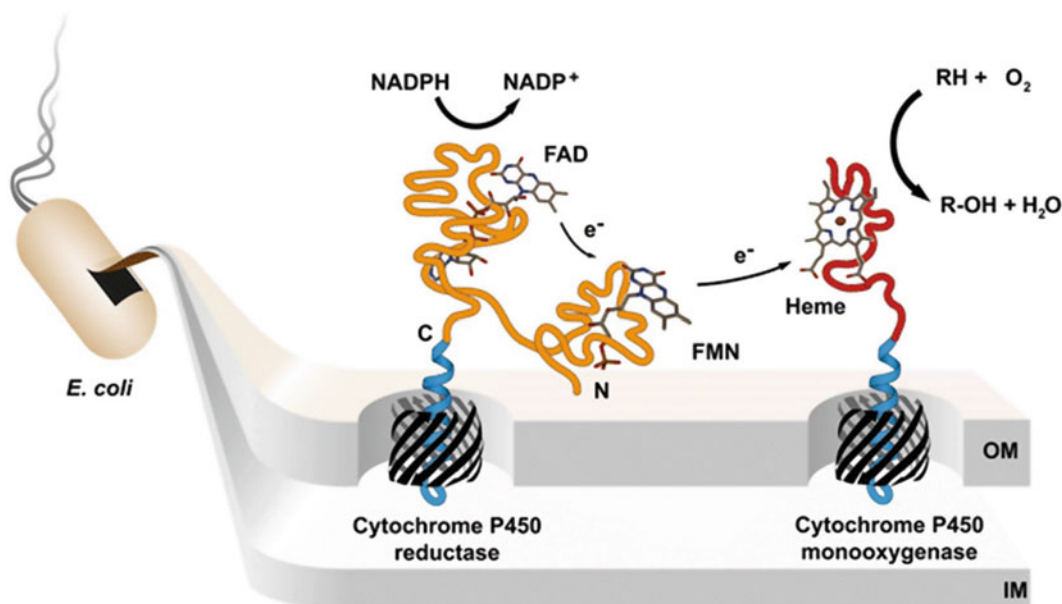


Fig. 13 Monooxygenase enzymatic activity in oxidative catalysis of the internal bacteria membrane (Quehl et al. [105])

have found P450 [105], the genomes of *S. coelicolor* and *M. tuberculosis* reveal 18–20 encoded P450 genes [107]. Thus, since various *M. tuberculosis* genes involved in the lipid metabolic system (with approximately 250 enzymes) have been processed, studies have focused on anti-P450 mycobacterial target strategies.

2.11 Pyridoxine-5-phosphate

The *pdxH* gene present in *E. coli* produces a monofunctional flavo-protein that can use either pyridoxine-5-phosphate (PNP) or pyridoxamine-5-phosphate (PMP) substrates [108, 109]. PNP oxidase acts as a key enzyme in the biosynthesis of the coenzyme pyridoxal-5-phosphate. This enzyme is present in several organisms, such as the human brain [110], *E. coli* [111], and *M. tuberculosis* [112]. The protein contains 203 amino acid residues.

In *E. coli*, the enzyme pyridoxine-5-phosphate oxidase (PNPOx) catalyzes oxidation of PNP to PLP during biosynthesis of vitamin B6, which itself also functions as a critical constituent of PLP. The enzyme catalyzes oxidation of either a C4 alcohol group or an amino group of the substrates PNP or PMP to an aldehyde, and forms PLP (Fig. 14). A hydrogen atom exits C4 during oxidation and a pair of electrons is carried towards the flavin mononucleotide (FMN) bond. PNPOx is the only flavoprotein oxidase that oxidizes an amine or primary alcohol to form the same product [113].

The protein is homodimeric with disulfide bridges connecting two monomeric portions through salt-bridge interactions. Each subunit is responsible for binding to a pyridoxal-5-phosphate molecule. This enzyme requires the presence of a cofactor, called FMN, a deep cloven composition of peptide subunits retained by hydrogen bonds with the protein. After formation of the binding substrate binding complex with the molecular receptor, it is converted to PLP. Amino acids present in the target are responsible for the PLP-hydrogen bonds constructing a cover on the active site, in closed conformation [114].

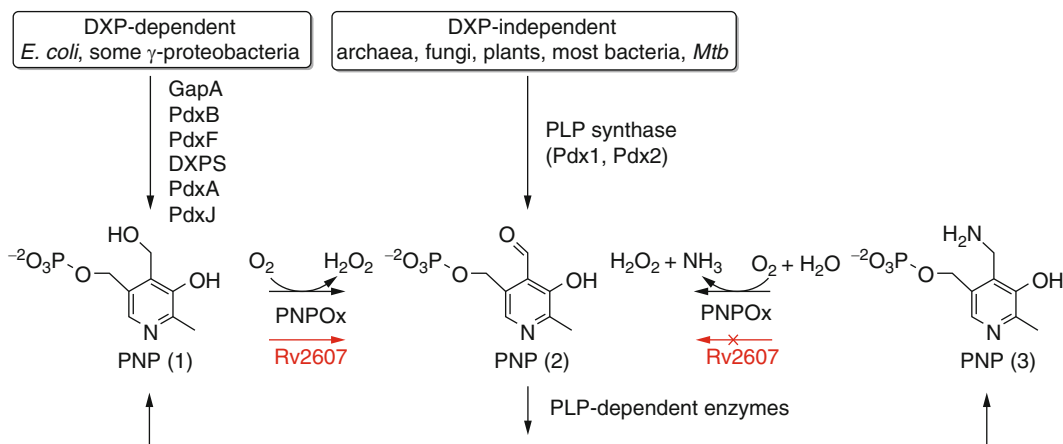


Fig. 14 Schematic of PLP formation route through the pyridoxine 5 phosphate enzyme (Mashalidis et al. [113])

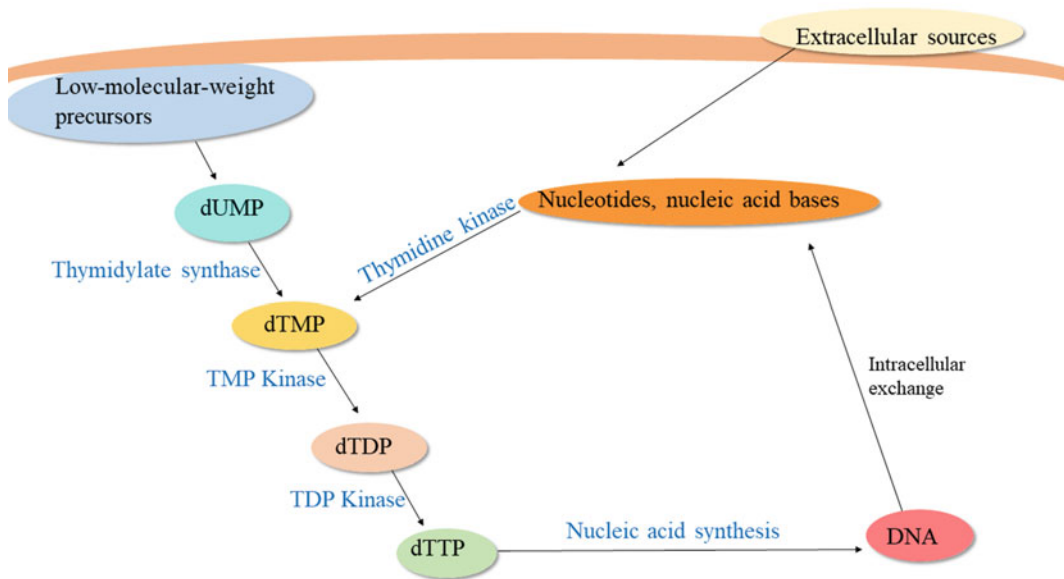


Fig. 15 Schematic view of the thymidine metabolic pathway in *M. tuberculosis*

2.12 Thymidine Monophosphate Kinase

For *M. tuberculosis* nucleotide metabolism, the enzyme thymidine monophosphate kinase (mtTMPK) is essential, its function is catalysis of reversible thymidine monophosphate (dTMP) phosphorylation to thymidine diphosphate (dTDP) [115]. The TMPK enzyme is known to catalyze the transfer of γ -phosphate ATP (Fig. 15) to thymidine monophosphate (dTMP) under the action of Mg^{2+} , to produce thymidine diphosphate (dTDP) and ADP. The presence of a tyrosine residue (Tyr103) near position 2 allows the enzyme to discriminate between 2-deoxynucleotides and ribonucleotides, and thus transform only the former [116].

Low structural similarities with the enzyme in humans make it an attractive target for planning of new agents against tuberculosis that block mycobacterial DNA synthesis. The enzyme is similar to the other TMPK isoenzymes, but in homology studies, it presents only 26% identity with *E. coli* isoenzymes, and even less (22–25%) with humans [117].

2.13 Isocitrate Lyase

Several glycolytic enzymes reported in the literature can induce the growth of *M. tuberculosis* in rats [118]. Many microorganisms can survive in the presence of acetate or fatty acids as the only source of carbon used to produce glyoxylate for cellular machinery. This involves the enzymes isocitrate lyase and malate synthase [119]. In this process, isocitrate breaks down into succinate and glyoxylate, and glyoxylate then condenses with acetyl-CoA to form malate. Glyoxylate also recycles the loss of two carbon dioxides from the tricarboxylic acid (TCA) cycle. During the TCA cycle, formation of acetyl-CoA requires the presence of isocitrate lyase.

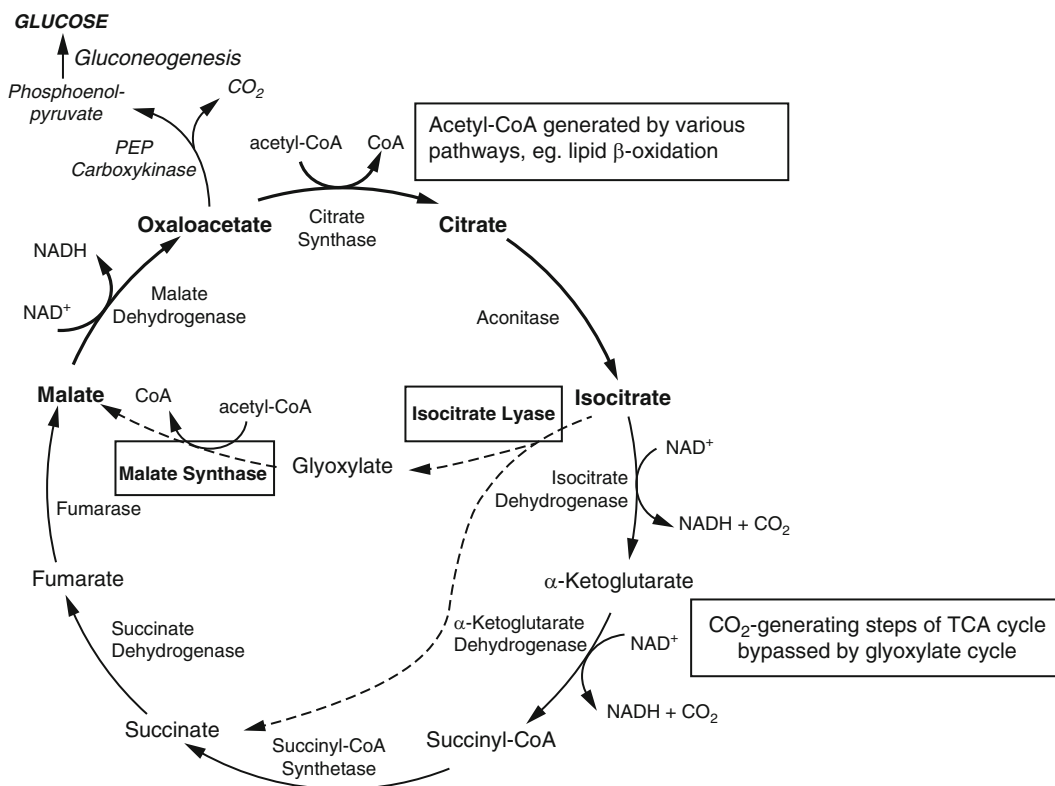


Fig. 16 Schematic view of the TCA cycle with the action of isocitrate lyase in the formation of glyoxylate (Cheah et al. [120])

Isocitrate lyase (Fig. 16) competes with isocitrate dehydrogenase for the isocitrate substrate [120].

The activity of isocitrate lyase is continuously reported for *M. tuberculosis*. The glyoxylate pathway in *M. tuberculosis* presents two genes in the production of isocitrate lyase [121]. One of them, *icl1*, encodes the same enzyme present in other eubacteria, while the second gene, called *icl2*, encodes a protein that shows greater homology with eukaryotic isocitrate lyase. Although both exhibit similar in vitro activity characteristics, *icl1* was shown to be more active than *icl2* [122], being positively regulated by Mtb.

2.14 Glutamine Synthetase

Glutamine synthetase (GS) has been reported in several studies in the literature as a major new target in the pathogenic pathway of *M. tuberculosis* [123]. The enzyme is secreted by pathogenic mycobacteria and essentially acts to influence the phagosome ammonia level of the host [124], and consequently on physiological phagosome conditions which influence poly-L-glutamic acid/glutamine synthesis [125]. GS, a dodecane formed from identical subunits is responsible for nitrogen production, catalyzing L-glutamate, ammonia, and ATP for L-glutamine synthesis (Fig. 17).

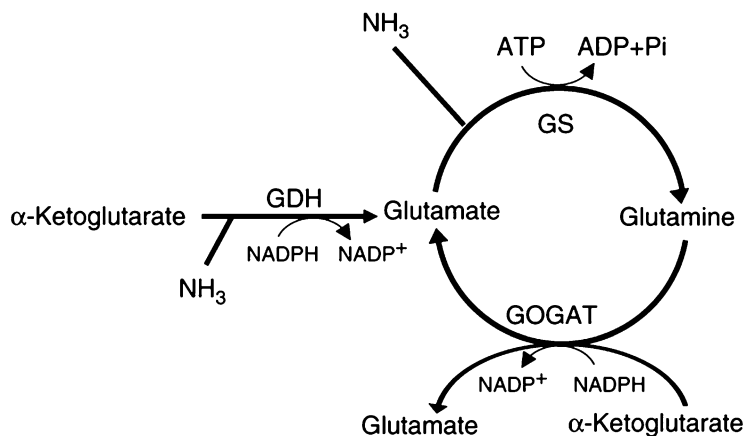


Fig. 17 Biochemical pathway of the glutamine synthetase regulation in *Escherichia coli*; assimilation of ammonia (Yuan et al. [124])

The mtGS enzyme (molecular mass of 53.6 kDa) is identical to that found in *M. tuberculosis*, and it can also be found in *M. smegmatis*, and yet it differs because in *M. smegmatis* it is exported to the outside of the cell. In *M. tuberculosis*, it remains within the cell. Studies suggest that this enzyme can be found in the early stages of mycobacteria growth, but in Mtb it appears as the only means of assimilating ammonia [126]. Genetically, the gene that determines glutamine synthetase (mtGS) expression presents only a single copy, being next to the *glnE* gene which encodes adenylyl transferase in *E. coli*, and is next to the *dsIII* gene which encodes another glutamine synthetase-like protein [127].

2.15 Shikimate Kinase

The shikimate pathway comprises seven steps to convert phenolpyruvate and erythrose 4-phosphate to chorismate, which itself is responsible for derivations in other synthesis, such as vitamins E and K, folic acid, ubiquinone, and aromatic α -amino acids [128]. The shikimate kinase (SK) enzyme, responsible for the fifth step of the pathway in the transfer of phosphoryl from ATP to shikimate [129], forms shikimate-3-phosphate (Fig. 18). This pathway has been reported to be essential for plants, fungi, algae, bacteria, and parasites [130], but its existence in mammals has not been reported; it is thus a promising target for the development of antimycobacterial agents.

Currently, there are crystallized shikimate kinase structures from *Mycobacterium tuberculosis* SK (mtSK). In Mtb, it belongs to the nucleoside family of monophosphate kinases. The core domain contains five stranded parallel β -sheets and a P-loop, which has five stranded parallel β -sheets surrounded by eight helices forming the nucleotide-binding site; the LID domain, which closes on the active site, and contains binding residues with ATP; and the

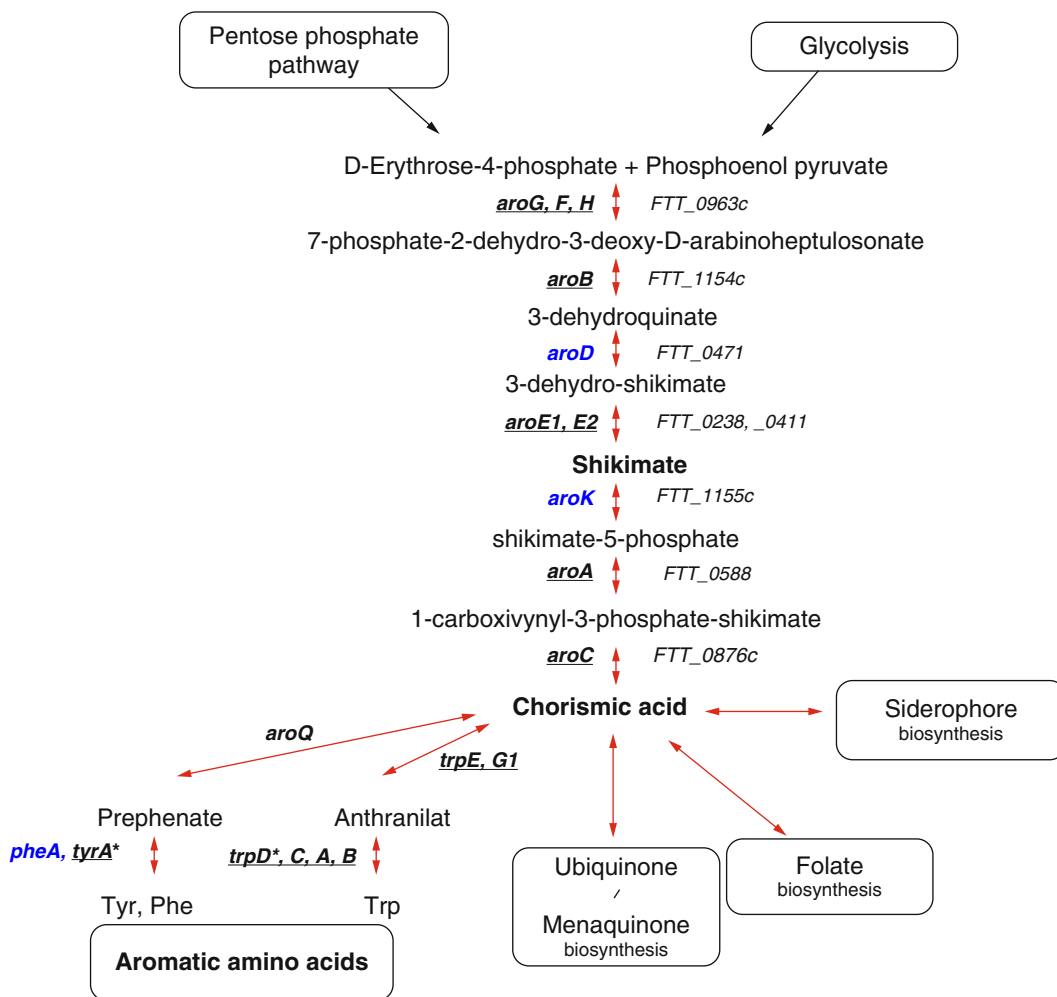


Fig. 18 Schematic representation of the shikimate pathway, with representation of the genes involved in the formation of the enzymes acting on the pathway (Meibom and Charbit [129])

shikimate binding domain, generally hydrophobic on the surface, which is essential to development of various enzyme inhibitors for antituberculosis activity [131].

The existence of a fourth domain has been reported, with a reduced core domain, due to changes in compartments resulting from binding of the linker to the enzyme, so that exact shikimate and ATP positioning occurs in the enzymatic process. ATP binding to the enzyme promotes rotation of the NB domain to the SB site, and shikimate binding promotes the rotation of the SB domain to the NB region [132].

2.16 CFP10-ESAT6

The CFP-10 and ESAT-6 proteins belong to a large family of proteins present in mycobacteria [133]. They have coiled-coil motif characteristics [134] and contain salt bridges that stabilize

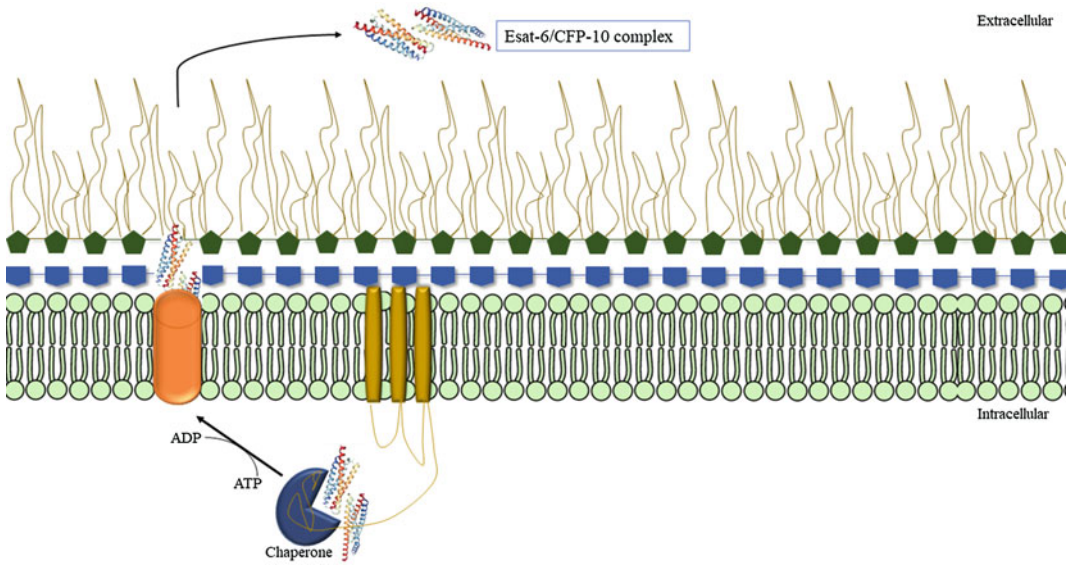


Fig. 19 Illustration of the ESAT-6 and CFP-10 secretion complex in the extracellular medium of *M. tuberculosis*

the complex [135]. The activity of this enzymatic complex involves cell signaling in lysing processes and formation of pores in the host cell membrane [136].

Approximately, 11 protein pairs are found in Mycobacteria, encoded by paired genes [137]. CFP-10 is known as a 10-kDa *M. tuberculosis* secreted antigen consisting of a 1:1 heterodimeric complex with a 6-kDa ESAT-6 protein. The two proteins are considered interdependent and act to inhibit production of reactive oxygen species (ROS), as well as regulating ROS production in bacterial lipopolysaccharides [138]. The protein pair form reveals two similar helix–turn–helix hairpin structures (Fig. 19) formed by the individual proteins, which lie antiparallel to each other, and form a four-helix bundle [135].

The genes of these enzymes are known as Rv3874 and RV3875, respectively, representing CFP-10 (with 100 residues) and ESAT-6 (with 95 residues). Both proteins are found in low levels in *M. tuberculosis* and *M. bovis* [139], despite their highly conserved status in the Mycobacteria genus. Studies have reported that the loss of one of these proteins in the mycobacteria causes reductions in tuberculosis virulence, an essential for new antituberculosis agents [140].

2.17 DNA Gyrase

DNA gyrase is a topoisomerase that acts in DNA supercoiling in a reaction that depends on ATP hydrolysis [141, 142]. Topoisomerase has two forms and depends on catalysis of one or both DNA strands (Fig. 20). Its structure is hetero-tetrameric, having two subunits of each type: A and B, which are commonly known as A2B2 [143, 144].

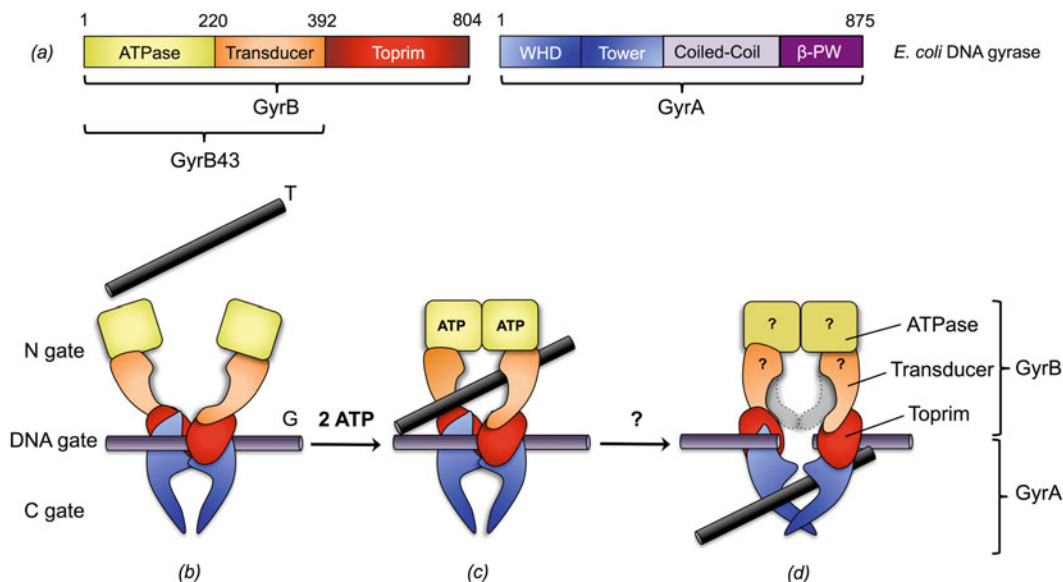


Fig. 20 Structural model of the enzyme DNA gyrase, with a representation of the binding sites, and their subunits A and B (Stanger et al. [143])

DNA gyrase is considered a promising target for the treatment of tuberculosis since several compounds such as fluoroquinolone inhibit subunit A of the enzyme, whereas the drug Novobiocin, an amino-coumarin derivative, inhibits subunit B. Chopra et al. [145] sequenced GyrB genes and found 99.9% homology between them, suggesting protein conservation, and a promising candidate for drug development.

2.18 Dihydrofolate Reductase (DHFR)

The enzyme dihydrofolate reductase acts by reducing 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF) with the aid of an NADPH cofactor [146, 147]. THF is the precursor responsible for supplying methyl groups for the synthesis of thymidylate, purine, methionine, serine, and glycine nucleotides required for DNA, RNA, and protein synthesis (Fig. 21). Dihydrofolate reductase (DHFR) receptor inhibition, as presented by THF analog compounds, freezes DNA synthesis and causes cell death. The bacteria fail to produce sufficient tetrahydrofolate which is used as a coenzyme in the production of purines necessary for DNA synthesis and cell reproduction. Human cells can use exogenous folate for metabolism, while bacteria depend on endogenous production (Fig. 4). Pathway failures do not normally affect human cells but are fatal to many bacteria [148], and this pathway has become an attractive target in the design and development of new drugs [149].

The DHFR structure has a central β sheet surrounded by four α -helices. The B sheet contains seven parallel chains and one anti-parallel chain. In homology studies, it was found that DHFR in humans has about 26% identity with mtDHFR. Crucial differences

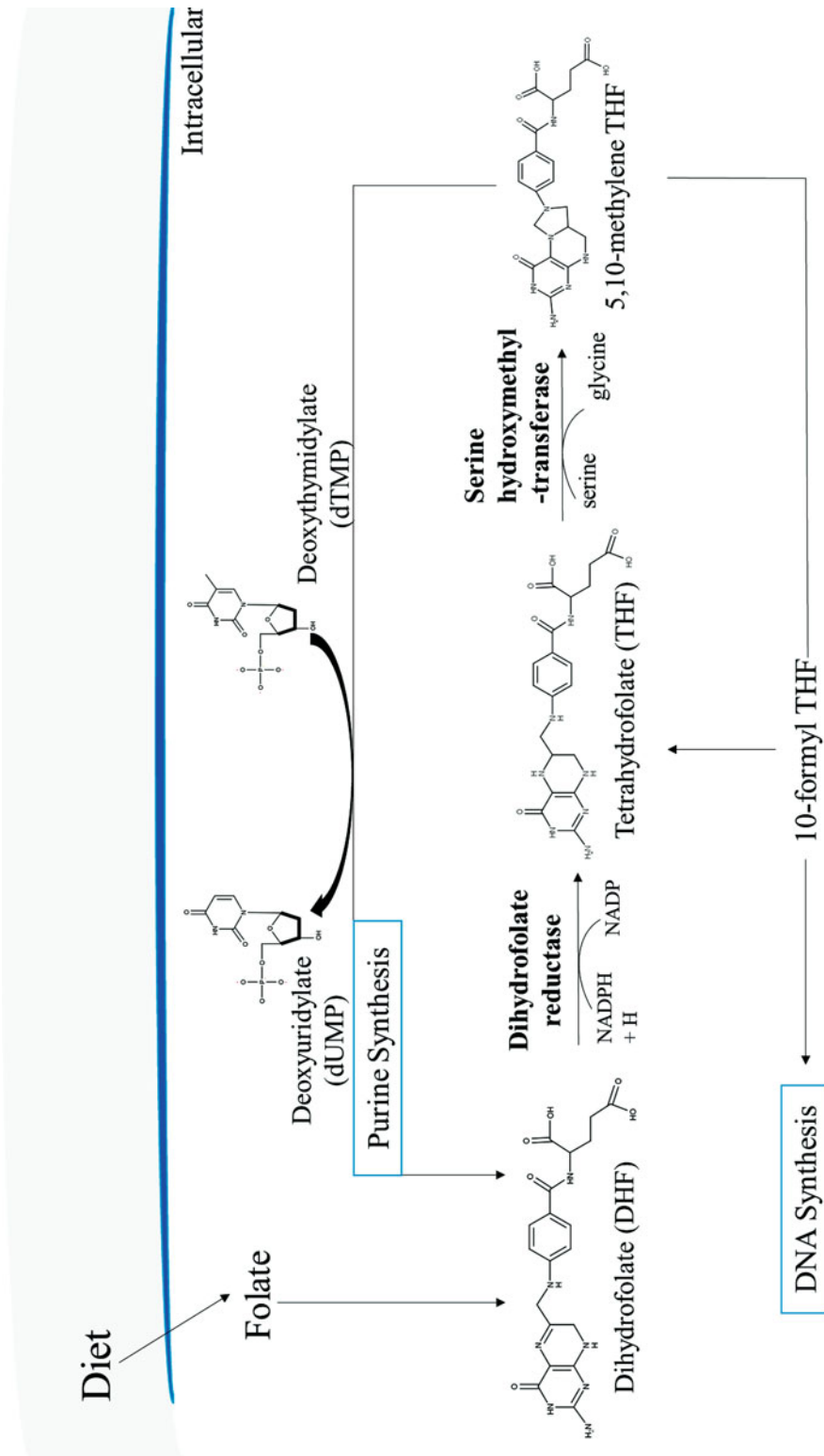


Fig. 21 Schematic DNA biosynthetic pathway and the action of the DHFR enzyme

are important in the active site specificity and can decrease cytotoxic effects [150, 151]. The differences in NADPH binding sites and active inhibitor sites is that Mtb has a glycerol molecule close to the active site coupled to Trp22, Asp27, and Gln28 residues, in humans the binding site has a glycerol islet with three hydrophobic residues Leu22, Pro26, and Phe31. Trimethoprim (TMP) and pyrimethamine (PYR) have antimycobacterial and antiprotozoal activity [152, 153] inhibiting DHFR, but TMP has the lower affinity.

3 Multitarget Compounds

Multitarget therapies are considered highly effective in treatment of multifactor diseases, especially those with high resistance. Molecular modeling studies are able to predict possible mechanisms of action for these compounds based on their physicochemical characteristics as well as measure their interactions in the ligand–receptor complex [154]. The benefits of using these compounds and their therapeutic value have been demonstrated, and the value of docking studies to assure faster planning and development of new compounds that circumvent resistance has also been demonstrated [155]. Herein will be reported several classes of compounds that act on more than one target, their interactions with active sites, and critical residues favoring greater activity.

3.1 Thiophenes

Synthesis of thiophene analogs has attracted considerable attention in pharmaceutical research. In drug design and development, heterocyclic ring substitutions are now a routine strategy [156]. Thiophenes and their derivatives belong to the aromatic heterocyclics group and are important structural fragments in many pharmaceuticals and chemical compounds [157]. The thiophene ring is a cyclic hydrocarbon; molecular formula C_4H_4S ; a hetero-aromatic compound. In the field of medicinal chemistry, benzo[*b*]thiophenes are important as biologically active heterocyclic molecules [158]; remarkably, they unite many biological activities [159].

In studies carried out by Lu et al. [160], several derivatives of 2-acylated and 2-alkylated amino-5-(4-(benzyloxy)phenyl)thiophene-3-carboxylic acid were evaluated for their antituberculosis activity. They exhibited MICs of 1.9 and 7.7 μM , low toxicity, and moderate activity; and when subjected to analysis in extensively drug-resistant strains (XDR-TB) yielded MICs of 12 and 16 μM . In order to predict the mechanisms of action involved in the antituberculosis activity, four derivative compounds (10d, 12h, 12k, and 15) were submitted to molecular docking analysis with the homologous *E. coli* FabH enzyme. The enzyme shows several key interactions in the 2,6-dichlorophobic group complex, with interaction in Cys112, His244, and Asn274 residues. The acid group forms an ionic interaction with the arginine residues (at the top of

the active site tunnel); with the 6-chloropiperonyl group, which interacts with residues close to the active site; and the indole moiety which act on alignment of the binding elements. These interactions serve as a basis for associating the activity of these compounds with mtFabH, and they were found to have such activity due to the acid group on the thiophene ring that forms an ionic interaction with Arg36 in the active site portion, and also due to the presence of hydrogen interactions between the carbonyl oxygen and hydrogen of the amino group with residue Arg249.

A second target reported in the literature for thiophene derivatives is the enzyme Pks13 [161]. In that study, in vitro activity inhibiting loading of fatty acyl-AMP in Pks13 was observed, thus inhibiting the biosynthesis of mycolic acids in *M. tuberculosis*, this yielded MIC values ranging from 0.5 to 78 μM . In order to design the in vitro studies, molecular docking studies of the thiophene derivatives with the enzyme mtPKS13 were performed. In that study, binding sites within the PKS13 protein domain were identified, separating residues Ser55 and Phe79, and confirming validation of the thiophene derivatives while characterizing their therapeutic role in inhibiting this Mtb biosynthetic pathway.

In the studies by Mahajan et al. [162], synthesis of several benzo [*b*]thiophene derivatives yielded compounds with in vitro activity and MICs in the range of 2.73–22.86 $\mu\text{g}/\text{mL}$ (*M. tuberculosis*) and 0.60 and 0.61 $\mu\text{g}/\text{mL}$ (*M. bovis*) both in dormancy, in addition to presenting low cytotoxicity for HeLa, Panc-1, and THP-1 cells. In this study, molecular docking studies elucidated the antituberculosis activity mechanism of the thiophene compounds. A good fit was predicted for thiophene linkers to the active site of the DprE1 (Decaprenyl phosphoryl- β -D-ribose-2'-epimerase) enzyme from its crystalline structure, with energy variance of only -9.198 for the most active compound (7a), to -6.995 for the less active compound (7e). The compounds appeared to have similar interactions with the active site residues; they had a number of interactions with the most active compound, with strong pi bond interactions observed for the His 132 space, and hydrogen bonds with Lys418 and Gly117. This demonstrated their high affinity for the enzyme.

3.2 Sulfonamides

Sulfonamides are a class of compounds having the functional group $-\text{SO}_2\text{NH}-$ in their chemical structure. The class constitutes the structural basis of a variety of drugs, present in many biologically active compounds, including antimicrobial drugs [163, 164], antiviral drugs [165], anticancer drugs, and anti-inflammatories [166]. In addition to their broad spectrum, they are among the most widely used antibacterial agents in the world, mainly because of their low cost, low toxicity, and excellent activity against common bacterial diseases [167].

In studies conducted by Naidu et al. [168], a series of 33 (thirty-three) new 6-(piperazin-1-yl) phenanthridine amides

and sulfonamides were evaluated for anti-Mtb activity, and presented MIC₅₀ activities of between 1.56 and 50 µg/mL. From these, two of the most active compounds were evaluated through molecular docking activity in ATPase domains of *M. tuberculosis* DNA Gyrase B. In these analyses, strong hydrophobic interactions were seen with Val98, Val99, Ile84, Pro85, Ala113, and Tyr114, in addition to certain specific interactions such as hydrogen bonding with Lys108 of compound 6d, a hydrophobic interaction with Ala87, and (for compound 7d) π - π stacking with Arg141.

Another study by Oliveira et al. [169] was carried out using sulfonamide derivatives for inhibition of the Mtb protein tyrosine phosphatase B (mtPtpB). The six sulfonamide and eight sulfonylhydrazone acyclic amide compounds presented excellent activity with IC₅₀ values ranging from 2.5 to 15 µM. The study also investigated molecular docking simulations for the PtpB enzyme with the four most active compounds. These simulations produced similar interactions which contributed significantly to the stability of the complexes. For the most favorable compound, the sulfonyl group presented hydrogen and dipole interactions, but the presence of *p*-bromo and *p*-nitrophenyl substituents overlapped amino acid Leu227. The observed substituents also exhibited favorable interactions in various amino acid residues, which may have contributed to increases or decreases in the potency of the compounds tested.

3.3 Chalcones

Chalcone, or 1,3-diphenyl-2E-propene-1-one, belongs to a class of natural products that occurs throughout the Plant kingdom. Benzylideneacetophenone is a precursor through a union of two aromatic nuclei united by a three-carbon α,β -unsaturated carbonyl bridge. The chalcones are precursors of flavonoids and isoflavonoids, and have attracted the attention of medicinal chemists synthesizing active derivatives [170], and been recorded in the literature as antioxidants [171], and antituberculosis agents [172] among others.

In 2010, Mascarello et al. [172] reported antituberculosis activity for five synthetic chalcones on protein tyrosine phosphatase A (mtPtpA), through competitive inhibition, and IC₅₀ values ranging from 7.5 to 55 µM. In this study, predictive molecular docking (Fig. 22) was performed which revealed key interaction residues of the complex, where methoxy groups in the ring interact with the residues Arg17, His49, and Thr12 in the active site of the enzyme. In addition to these interactions, substitutions on the phenyl ring by the 2-naphthyl grouping were seen to be subject to hydrophobic interactions with the PtpA Trp47 residue.

Yavad et al. [173] studied synthetic chalcone derivatives (3.5–30 µg/mL) through in vitro tests against *M. tuberculosis*. Quantitative Structure–Activity Relationship (QSAR) was studied, generating a model to correlate the physical–chemical properties of

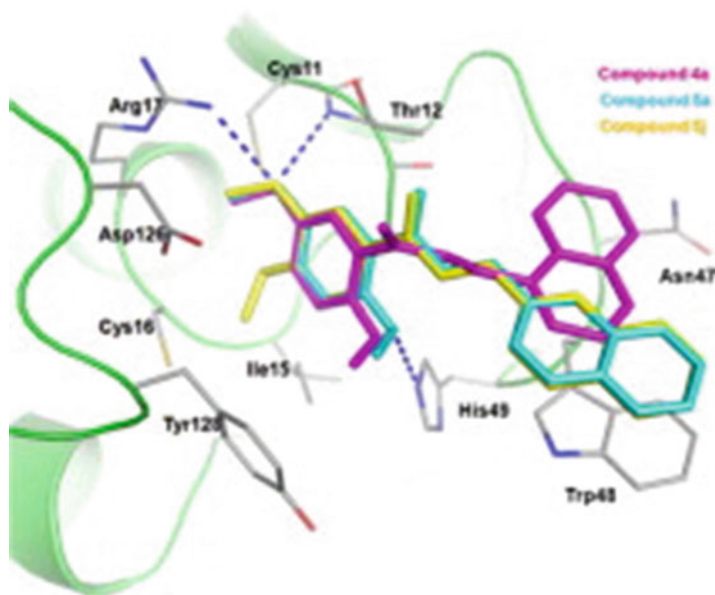


Fig. 22 Binder–receptor complex showing the interactions of the amino acid residues with the portions of the chalcone derivatives reported in the literature. Docking of a chalcone derivative complexed with the enzyme mtPtpA

chalcones with their antitubercular activity. Through regression curve analysis, it was observed that the lowest unoccupied molecular orbital energy (eV), together with the amine, hydroxyl, and methyl group counts, positively correlated with activity. In addition, the molecules that were found to be the most active underwent molecular docking studies to explore their receptor affinity mechanisms. The enzyme InhA commonly known as *M. tuberculosis* enoyl reductase (ENR) was used; chalcones present similar residues to those found with the standard drug, isoniazid (INZ): aliphatic interactions with residues Gly93 and Ile200, sulfur grouping with Met159 and Met206, the cyclic portion with Pro191, and the aromatic portion with Tyr146, try156, and Phe203. The anchoring study on mechanism of action showed high-affinity binding of the active derivatives, thus characterizing them as possible drug candidates against tuberculosis.

Gond et al. [174] analyzed 11 synthetic chalcone derivatives in molecular docking simulations to investigate possible interactions with the *M. tuberculosis* DHFR enzyme. The interactions obtained from the analysis were the same interactions as presented by methotrexate, with key residues Ala9, Ile7, Glu30, Ser59, Tyr121, and Val 115 involved.

3.4 Nitroimidazoles

Nitroimidazoles are compounds with widely reported antibacterial and antiprotozoal actions [175]. The majority of their activities are due to drug bio-reduction processes within the pathogen. The

activity of nitroimidazoles against Mtb has already been reported in the literature [176], and the activity of nitroimidazoles varies according to the location of the nitro group.

In studies conducted by Sharma et al. [177], profiles of compounds with antituberculosis activity were analyzed. The compounds were obtained from the ZINC, DRUG BANK, and PUBCHEM databases and divided into known first-line compounds and natural antituberculosis compounds. In the screening of these compounds, ZINC00004165 (5-[3-(2-nitroimidazol-1-yl) propyl] phenanthridine) presented the highest interaction with cytochrome P450 monooxygenase, presenting an interaction energy superior to drugs such as Pyrazinamide, Rifampicin, Isoniazid, Streptomycin, and Ethambutol reported in the literature.

Gupta et al. [178] analyzed a series of 2-nitroimidazooxazine derivatives subjected to QSAR studies. Their physicochemical properties were calculated and submitted to statistical analysis. The compounds were subjected to molecular docking analysis with deazaflavin-dependent nitroreductase (Ddn), and exhibited key interactions with nitro and Tyr65 residues and hydrogen interactions with the Tyr133 residue, exhibiting similar interactions with a co-crystallized linker positioned at the active site of the enzyme.

Somasundaram et al. [179] studied the activity observed at a concentration of 12.5 $\mu\text{g}/\text{mL}$ of a nitroimidazole (PA-824) activated by Ddn, which catalyzes reduction and release of lethal reactive nitrogen. To further understand its mechanisms of action, the compound was subjected to docking studies in the crystalline structures of wild-type and mutated Ddn (Fig. 23a). In these analyses, two hydrogen interactions were observed and key interactions (involving inhibition effectiveness) occurred between the nitro portion of the compound and Met87, and oxygen with Try88 of the active site. A hydrogen interaction with the Glu83 residue was indicated in wild-type Ddn, but not in mutated Ddn.

In studies conducted by Kumar and Jaleel [180], the activity of PA-824, a nitroimidazole that acts against antihypoxia activity, and replication of active or latent forms of *M. tuberculosis* were analyzed through bioinformatics. The activity of this molecule has already been reported in the literature, but its reaction mechanism has not been elucidated. The basis of the article was prediction of PA-824's activity on the PNPOx enzyme (Fig. 23b). Several PA-824 interactions were reported, including residue Ile51 with 5H bond interactions between the compound and the enzyme, yet residue Lys57 was the best pocket site with a free energy binding of -6.17 kcal/mol and six H bond interactions between the target protein and compound, making it the best pocket site for a potential inhibitor.

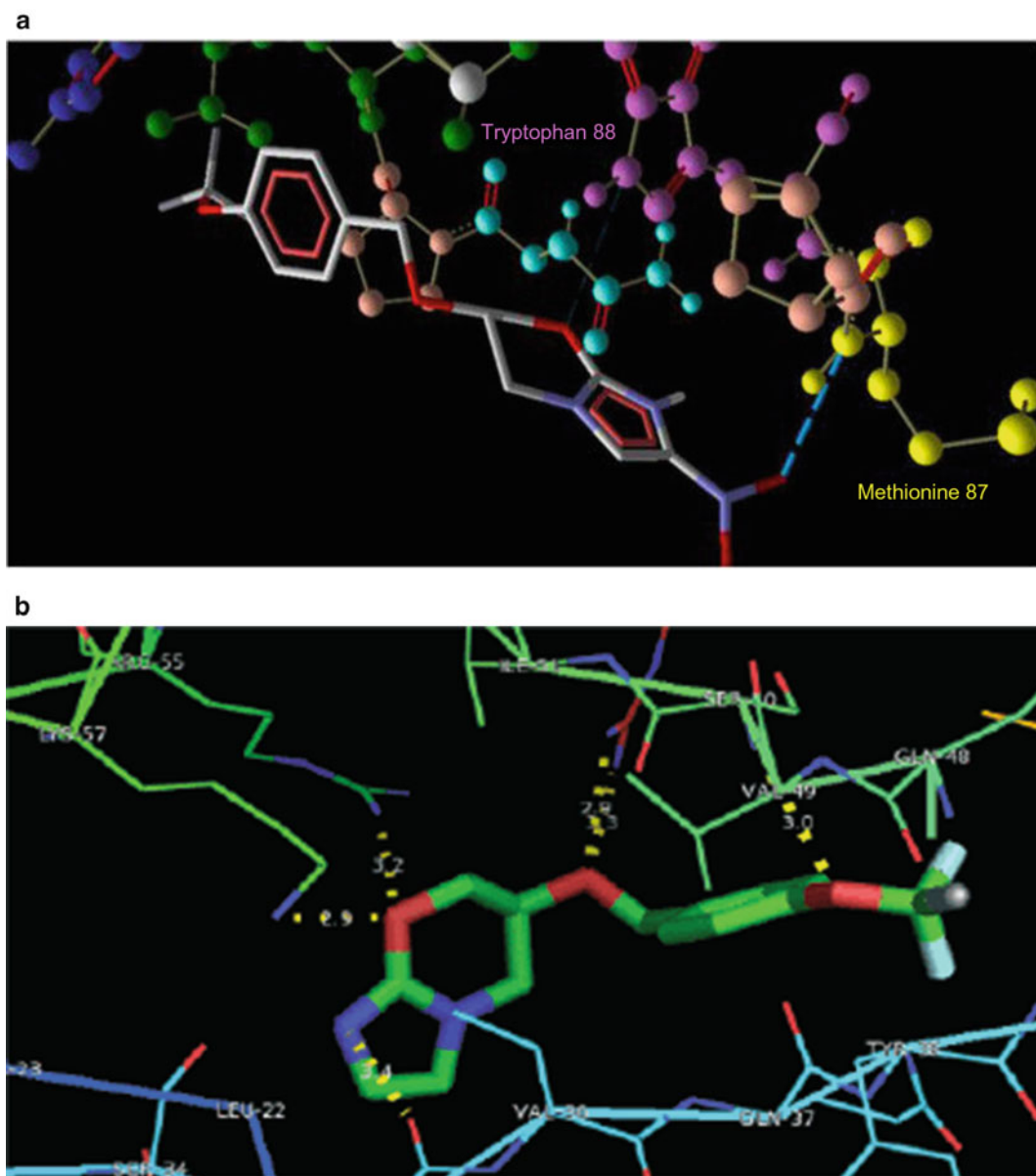


Fig. 23 Binder–receptor complex showing interactions between amino acid residues and nitroimidazole derivative portions. **(a)** Candidate PA-824 interacting with the enzyme Ddn. **(b)** Candidate PA-824 complexed with a pyridoxine-5-phosphate oxidase (PNPOx) enzyme

3.5 Benzimidazoles

Benzimidazoles or imidazolines are five-membered heterocyclic systems containing an imino group which is fused from amino acids [181]. These compounds present several therapeutic activities [182].

Saleshier and collaborators [183] in a study of new antimycobacterial agents synthesized a series of pyran-substituted

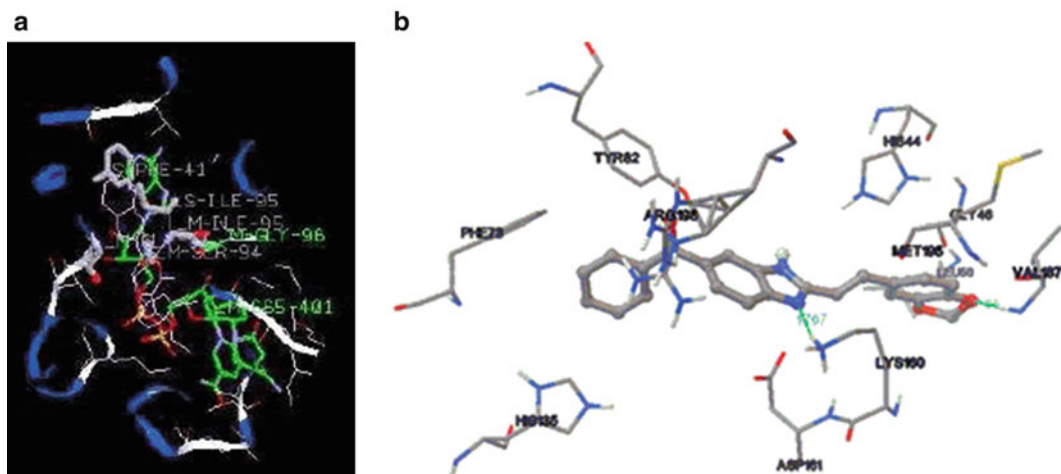


Fig. 24 Binder–receptor complex showing amino acid residue interactions with benzimidazole derivative portions. (a) Docking of pyran-substituted benzimidazoles derivatives complexed with the Mtb InhA enzyme. (b) 2-Heterostyrylbenzimidazole derivative inhibiting the Mtb Pantothenate synthetase enzyme

benzimidazoles and evaluated their activity against tuberculosis in concentrations of 1–100 mcg/mL. The molecular docking studies for the enzyme enoyl reductase (InhA) of Mtb (Fig. 24a) indicated good binding affinity to the receptor, and key interactions with the residues of Gly96, Ile95, Phe41, and Ser94 in agreement with a demonstration of in vitro activity.

Soni et al. [184] screen investigated 500,000 compounds from the Asinex database. Nine inhibitors of *N*-acetylglucosamine-1-phosphate uridylyltransferases (GmlU) from *M. tuberculosis* underwent predictive analyses through molecular docking. Key interactions with Arg19, Thr 89, Gly 151, Glu 166, Asn 181, and Asn 239 residues were presented, and the pyrimidine moiety presented hydrophobic interactions with Leu12, Ala14, Val55, Pro86, Leu87, and Ala92 residues.

The *M. tuberculosis* DHFR enzyme is another target for benzimidazolic inhibition as reported by Priyadarsini et al. [185]. At a concentration of 0.01 $\mu\text{g}/\text{mL}$, the researchers found the tested derivatives to be moderately active when compared to methotrexate. Through molecular docking studies, interaction between benzimidazoles and the mtDHFR enzyme was predicted for key residues: Ile94, Arg60, Phe31, Leu57, Ile20, and Pro51, thus characterizing antituberculosis activity.

Aanandhi et al. [186] synthesized new benzimidazole derivatives as well as performed molecular docking analyses with the enzymatic complex CFP10-ESAT6. As a result, interactions with Ser34, Glu33, Gly37, and Gln40 residues present on the active site of the compound were reported, and have been shown to be essential for inhibition activity.

Anguru and collaborators [187] reported synthesis of 2-heterostyrylbenzimidazole derivatives evaluated for their antituberculosis activity at less than 16 $\mu\text{g}/\text{mL}$. In docking studies, interaction of pantothenate synthetase (Fig. 24b) with the amino acids Lys160 and Val187 in all compounds was verified, and is essential for inhibition activity.

3.6 Peptides

Peptides are signaling molecules that bind to specific receptors on the cell membrane to elicit intracellular effects. Peptides consist of four structural parts: α -helical, β -stranded, β -hairpin, or loop and extended. They are highly specific and are generally safe and effective in humans [188]. The search for new peptides is essentially based on their biochemical affinity to their target macromolecules and their promotion of desired effects [189, 190].

Considering structural similarities between *M. tuberculosis* DHFR enzymes and human DHFR, a molecular docking study of tripeptide inhibitors with the (DHFR) enzyme was carried out by Kumar et al. [191]. The peptide designed in the study obtained potency of up to 6 \times higher than methotrexate, a crystallized PDB inhibitor, and was up to 120 \times more selective for the *M. tuberculosis* DHFR enzyme as compared to the selective for human DHFR enzyme. The tripeptide can be considered a lead candidate compound for new antituberculosis drugs.

Another cyclic peptide was designed in the studies of Chandra et al. [192], from the β -lactam structure through intramolecular ring opening transamidation, and this presented micromolar activities against the PtpA enzyme of *M. tuberculosis*. In this same work, molecular docking studies (Fig. 25a) of compound binding to the active site of the enzyme demonstrated induced flexibility in the PTP loop region, which already has a catalytic function in the interaction of hydrogen bonding of the cyclic terminal with amino acid residues Cys11 and Arg17. In addition, key interactions with Gly125 residues, and interactions of the *N*-phenyl moiety with Trp48 and Thr12, have been reported to act in strengthening the existing bond.

In the studies conducted by Yang et al. [193], a library of compounds led to the identification of the compound actinomycin D (ActD) as an inhibitor of the enzyme pantothenate synthetase of *M. tuberculosis* (mtPS), with activity at 250.72 μM . A virtual screening was performed based on the structure of ActD in which the existence of two compounds with 10 \times activity greater than that registered for ActD was evidenced. When conducting a molecular docking study (Fig. 25b) with the mtPS enzyme, the cyclopeptide portion was shown to be essential for inhibition, involving hydrogen interactions with the key residues of amino acids Try82, His135, Lys160, and Gln164.

Another study evaluated the structure of the enzyme thymidine monophosphate kinase in *M. tuberculosis* and in humans to

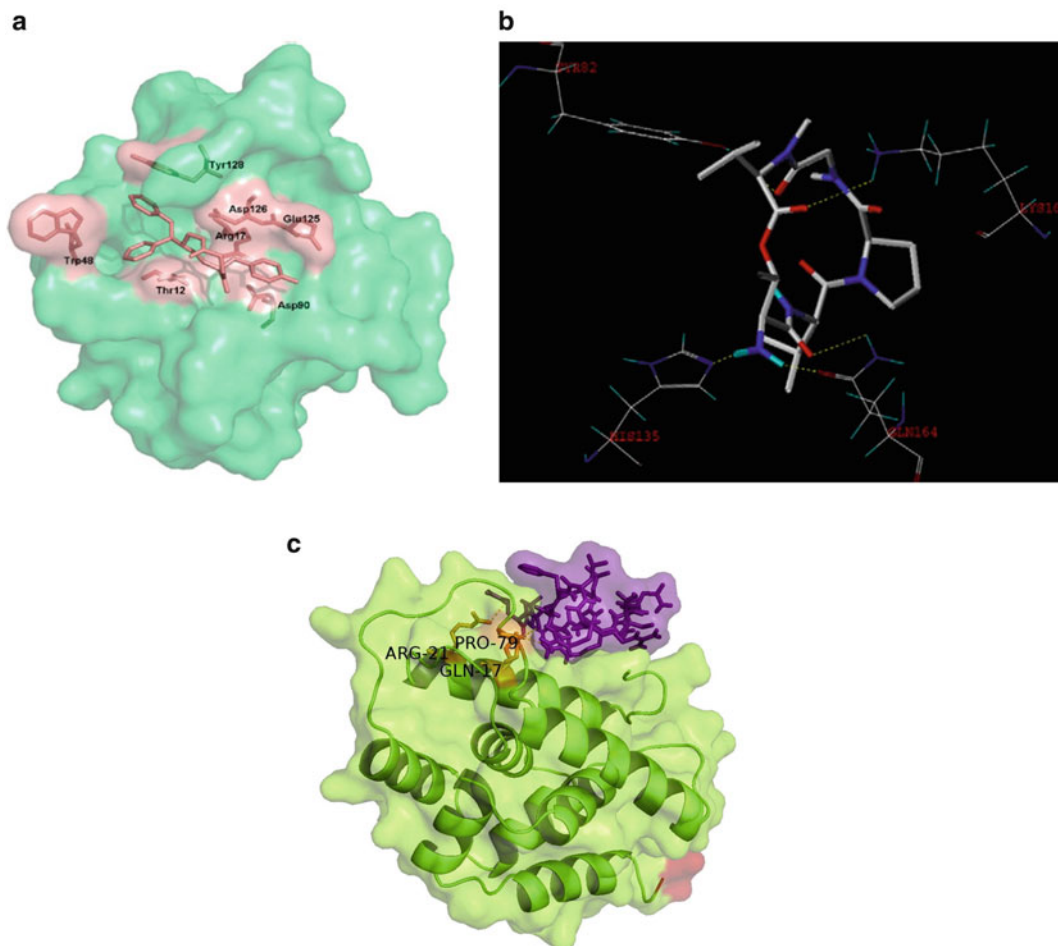


Fig. 25 Binder–receptor complex showing amino acid residue interactions with portions of peptides. (a) Cyclic peptide in the active site of the *M. tuberculosis* enzyme PtpA. (b) ActD inhibiting a pantothenate synthetase enzyme. (c) Lidamycin, an antimycobacterial cyclic peptide, inhibiting an ATP-dependent protease enzyme

investigate the efficacy of new tripeptide inhibitors [194]. By means of docking studies, it was seen that the highest affinity was due to aromatic double bond interactions with residues Tyr103 and Tyr165, interactions of the naphosultam ring with the Tyr39 residue, and guanidine portion hydrogen interactions with the Arg95 residue. The tripeptide prediction resulted in $33,000\times$ greater affinity for mtTMPK, thus highlighting the compound for further research and a possible new class of Mtb selective compounds.

Another study reported on 29 heptapeptide inhibitors of *M. tuberculosis*, specific to the enzyme isocitrate lyase (ICL). Prediction of this activity occurred during molecular docking studies. Three of these inhibitors demonstrated IC_{50} of $126\ \mu\text{M}$ and are involved in enzyme cavity binding generating conformational changes, and suppressing catalytic activity through non-competitive inhibition. Thus,

it was found that the interaction of the Gln119 residue in the active site is essential for peptide inhibitory activity [195].

Perumal et al. [196] studied interactions of a group of endogenous 10–50 amino acid peptides with voluminous amounts of hydrophobic residues. The affinity of these peptides for glutamine synthetase and RNA polymerase proteins was determined; compounds (AIA)—I and II were shown to be highly potent against both enzymes, as well as presenting higher interaction energies than standard drugs, such as rifampicin and isoniazid.

As another antituberculosis drug candidate [197], a dipeptide inhibitor based on the shikimate kinase (SK) structure, with binding affinity predicted at 5.5 η M, and a bond similar to that found for the shikimate substrate showed a high number of interactions: hydrogen with Asp34, Arg58, and Arg136, and its carboxylic group interacting with the Gly81 residue of the enzyme and residues Arg117, Pro118, and Leu119 from the LID domain.

In studies conducted by Gavriš et al. [198], lidamycin, an antimycobacterial cyclic peptide which plays a role in growing and dormant forms of *M. tuberculosis*, was studied. In molecular docking of the peptide with the ATP-Dependent Protease enzyme (Fig. 25c), the presence of Gln17 residues was found; this is key to formation of an inhibitory complex and may explain its antimycobacterial activity.

3.7 Quinolones

Due to their broad spectrum antibacterial activities, quinolones are an important class of compounds in medicinal chemistry [199]. With the various structural modifications reported in the literature, it is seen that changes in carbon positions 6 and 8 present more potent antibacterial activity [200]; yet, fluoroquinolones are effective against a wider spectrum of bacteria [201, 202].

Guzman et al. [203] reported on *N*-methyl-2-alkenyl-4-quinolones that act on the *M. tuberculosis* MurE ligase enzyme yielding IC₅₀ activity of 40–200 μ M. In addition, molecular docking analyses were performed on the target, which presented weak bonds with only one interaction coming from the nitrile group of the quinolone with the hydroxyl group present in residue Thr176. Due to the extra space existing in the cavity of the active site, development of new derivatives is possible.

Cunha et al. [204] reported on gatifloxacin analog interactions and activity against the *M. tuberculosis* DNA gyrase enzyme. Bioinformatic studies verified six compounds as having interactions with subunits A and B of the enzyme. The study reported that key steric interactions were present at hydrogen bonds of amino acids Lys49 and Asn172 and with amino acid Ser91; interactions of the bicyclic ring at the active site region of the enzyme were also observed.

Maddela and Makula [205] synthesized 22 isatin–quinoline hybrids and submitted them to molecular docking studies with the enzyme enoyl ACP reductase (InhA). The most active

compound presented hydrophobic interactions with residues Phe97, Gly96, and Met199, in addition to presenting critical interactions with residues Tyr158 and NAD500 in the active site of the receptor. The results confirmed observed in vitro activity of 0.09 μM against Mtb, verifying isatin–quinoline as a promising antituberculosis compound.

Minovski et al. [206] constructed combinatorial libraries with several analogous compounds, a total of 53,871 6-fluoroquinolone structures which passed through a molecular filter using neural networks (NN). After virtual screening, the compounds were submitted to chemometrics analysis and molecular modeling in the A subunit of the enzyme DNA gyrase. They presented interactions with Ser79, Arg117, Arg456, and Glu475 considered key interactions for GyrA enzyme activity. Levofloxacin was used as a standard.

4 Conclusion

The multitarget drug is a key that can open multiple locks. Research in multifunctional compounds can follow two paths: comprehensive experimental analyses, or computer-aided rational drug design; and screening for and identifying potential targets, with optimization; both avoid higher expenses.

In this chapter, we have presented molecular docking studies as a method of multitarget compound discovery in the treatment of tuberculosis disease. Based on this review, the following interactions of the ligand–receptor complex of compounds to their enzymes were observed: (1) thiophenes with activities in FabH, Pks13, and DprE1 targets, (2) sulfonamides with activities in DNA gyrase and PtpB targets, (3) chalcones with activities in the enzymes PtpA, InhA, and DHFR, (4) nitroimidazoles with activities in Cytochrome P450 monooxygenase, Ddn, and Pyridoxine 5'-phosphate oxidase, (5) benzimidazoles with activities in the enzymes InhA, GmlU, DHFR, CFPI0-ESAT6, and Pantothenate synthetase, (6) peptides with activities in the enzymes DHFR, PtpA, Pantothenate synthetase, TMPK, isocitrate lyase, Glutamine synthetase, Shikimate kinase, and ATP-dependent protease, and (7) quinolone derivatives with activities in MurE ligase, DNA gyrase, and InhA targets. The use of in silico methods presented the action of these compounds in greater detail, as well as predicting possible interactions with their enzymes. Such methods accelerate and favor the development of drugs with activities in more than one receptor. Development of more specific and potent compounds for such varied targets is of great interest to the public health industry and may well reduce the current high number of *M. tuberculosis* resistance cases.

Acknowledgments

The authors acknowledge support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

References

1. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, Soolingen D, Cole ST (2002) A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci U S A* 99 (6):3684–3689
2. Sensi P, Grass IGG (1996) Antimycobacterial agents. In: Wolff ME (ed) *Burger's medicinal chemistry and drug discovery*, 5th edn. Wiley, New York
3. World Health Organization (2017) Global TB Rep. 2017. <http://apps.who.int/iris/bitstream/10665/259366/1/9789241565516-eng.pdf?ua=1>. Accessed Aug 2017
4. Junior ALR, Netto AR, Castilho EA (2014) Spatial distribution of the human development index, HIV infection and AIDS-Tuberculosis comorbidity: Brazil, 1982–2007. *Rev Bras Epidemiol* 17:204–215
5. World Health Organization (2016) Global TB report. <http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf>. Accessed Aug 2017
6. The World Health Organization (WHO) (2015) Global tuberculosis report. http://www.who.int/tb/publications/global_report/en/. Accessed Apr 2017
7. Hatfull GF, Jacobs WR Jr (2014) *Molecular genetics of mycobacteria*. American Society for Microbiology Press, Washington, DC
8. Zhang HN, Xu ZW, Jiang HW, Wu FL, He X, Liu Y, Guo SJ, Li Y, Bi LJ, Deng JY, Zhang XE, Tao SC (2017) Cyclic di-GMP regulates *Mycobacterium tuberculosis* resistance to ethionamide. *Sci Rep* 7(1):5820–5832
9. Joshi RR, Barchha A, Khedkar VM, Pissurlenkar RRS, Sarkar S, Sarkar D, Joshi RR, Joshi RA, Shah AK, Coutinho EC (2015) Targeting dormant tuberculosis bacilli: results for molecules with a novel pyrimidone scaffold. *Chem Biol Drug Des* 85(2):201–207
10. Parida SK, Axelsson-Robertson R, Rao MV, Singh N, Master I, Lutckii A, Keshavjee S, Anersson J, Zumla A, Maeurer M (2015) Totally drug-resistant tuberculosis and adjunct therapies. *J Intern Med* 277 (4):388–405
11. Sotgiu G, Centis R, D'Ambrosio L, Alffenaar JWC, Anger HA, Caminero JA, Castiglia P, Lorenzo S, Ferrara G, Koh WJ, Schecter GF, Shim TS, Singla R, Skrahina A, Spanevello A, Udwardia ZF, Villar M, Zampogna E, Zellweger JP, Zumla A, Migliori GB (2012) Efficacy, safety and tolerability of linezolid containing regimens in treating MDR-TB and XDR-TB: systematic review and meta-analysis. *Eur Respir J* 40(6):1430–1442
12. Udwardia ZF, Amale RA, Ajbani KK, Rodrigues CS (2012) Nomenclature of drug-resistant tuberculosis. *Lancet Infect Dis* 13 (11):917
13. Yew WW, Leung CC (2008) Update in tuberculosis 2007. *Am J Respir Crit Care Med* 177 (5):479–485
14. Goldman RC, Plumley KV, Laughon BE (2007) The evolution of extensively drug resistant tuberculosis (XDR-TB): history, status and issues for global control. *Infect Disord Drug Targets* 7(2):73–91
15. Espinal MA (2003) The global situation of MDR-TB. *Tuberculosis* 83(1):44–51
16. Yildirim MA, Goh KI, Cusick ME, Barabasi AL, Vidal M (2007) Drug-target network. *Nat Biotechnol* 25(10):1119–1126
17. Lavecchia A, Di Giovanni C (2013) Virtual screening strategies in drug discovery: a critical review. *Curr Med Chem* 20 (23):2839–2860
18. Pei J, Yin N, Ma X, Lai L (2014) Systems biology brings new dimensions for structure-based drug design. *J Am Chem Soc* 136:11556–11565
19. Morphy R, Rankovic Z (2005) Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem* 48:6523–6543

20. Zhang W, Pei J, Lai L (2017) Computational multitarget drug design. *J Chem Inf Model* 57:403–412
21. McBryde ES, Meehan MT, Doan TN, Ragonnet R, Marais BJ, Guernier V, Trauer JM (2017) The risk of global epidemic replacement with drug-resistant *Mycobacterium tuberculosis* strains. *Int J Infect Dis* 56:14–20
22. Ramaswamy SV, Amin AG, Göksel S, Stager CE, Dou SJ, Sahly HE, Moghazeh SL, Kreiswirth BN, Musser JM (2000) Molecular genetic analysis of nucleotide polymorphisms associated with ethambutol resistance in human isolates of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 44:326–336
23. Bhatt JD, Chudasama CJ, Patel KD (2015) Pyrazole clubbed triazolo [1, 5- α] pyrimidine hybrids as an anti-tubercular agent: synthesis, *in vitro* screening and molecular docking study. *Bioorg Med Chem* 23(24):7711–7716
24. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
25. Lou Z, Zhang X (2010) Protein targets for structure-based anti *Mycobacterium tuberculosis* drug discovery. *Protein Cell* 1(5):435–442
26. Chetty S, Ramesh M, Pillay AS, Soliman MES (2017) Recent advancements in the development of anti-tuberculosis drugs. *Bioorg Med Chem Lett* 27(3):370–386
27. Yuriev E, Agostino M, Ramsland PA (2011) Challenges and advances in computational docking: 2009 in review. *J Mol Recognit* 24(2):149–164
28. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, Tekaija F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393(6685):537–544
29. Ananthan S, Faaleolea ER, Goldman RC, Hobrath JV, Kwong CD, Laughon BE, Mad-dry JA, Mehta A, Rasmussen L, Reynolds RC, Secrist JA, Shindo N, Showe DN, Sosa MI, Sunling WJ, White EL (2009) High-throughput screening for inhibitors of *Mycobacterium tuberculosis* H37Rv. *Tuberculosis* 89:335–353
30. Foo CSY, Lechartier B, Kolly GS, Röttger SB, Neres J, Rybniker J, Lupien A, Sala C, Piton J, Cole ST (2016) Characterization of DprE1-mediated benzothiazinone resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 60(11):6451–6459
31. New Drugs TB. <https://www.newtdrugs.org/pipeline/compounds>. Accessed 15 Sept 2017
32. Mikusova K, Huang H, Yagi T, Holsters M, Vereecke D, Haeze WD, Scherman MS, Brennan PJ, Mcneil MR, Crick DC (2005) Decaprenylphosphoryl arabinofuranose, the donor of the D-arabinofuranosyl residues of *Mycobacterium tuberculosis*, is formed via a two-step epimerization of decaprenylphosphoryl ribose. *J Bacteriol* 187(23):8020–8025
33. Batt SM, Jabeen T, Bhowruth V, Quill L, Lund PA, Eggeling L, Alderwick LJ, Fütterer K, Besra GS (2012) Structural basis of inhibition of *Mycobacterium tuberculosis* DprE1 by benzothiazinone inhibitors. *Proc Natl Acad Sci U S A* 109(28):11354–11359
34. Neres J, Pojer F, Molteni E, Chiarelli LR, Dhar N, Boy-röttger S, Buroni S, Fullam E, Degiacomi G, Lucarelli AP, Read RJ, Zanoni G, Edmondson DE, Rossi E, Pasca MR, Mckinney JD, Dyson PJ, Riccardi G, Mattevi A, Cole ST, Binda C (2012) Structural basis for benzothiazinone-mediated killing of *Mycobacterium tuberculosis*. *Sci Transl Med* 4(150):121–150
35. Gao C, Ye TH, Wang NY, Zeng XX, Zhang LD, Xiong Y, You XY, Xia Y, Peng CT, Zuo WQ, Wei Y, Yu LT (2013) Synthesis and structure–activity relationships evaluation of benzothiazinone derivatives as potential anti-tubercular agents. *Bioorg Med Chem Lett* 23(17):4919–4922
36. Tiwari R, Miller PA, Chiarelli LR, Mori G, Sarkan M, Centarova I, Cho S, Mikusova K, Franzblau SG, Oliver AG, Miller MJ (2016) Design, syntheses, and anti-TB activity of 1,3-benzothiazinone azide and click chemistry products inspired by BTZ043. *Med Chem Lett* 7:266–270
37. Riccardi G, Pasca MR, Chiarelli LR, Manina G, Mattevi A, Binda C (2012) The DprE1 enzyme, one of the most vulnerable targets of *Mycobacterium tuberculosis*. *Appl Microbiol Biotechnol* 97(20):8841–8848
38. Trefzer C, Gonzalez MR, Hinner MJ, Schneider P, Makarov V, Cole ST, Johnsson K (2010) Benzothiazinones: prodrugs that covalently modify the decaprenylphosphoryl- β -D-ribose 2'-epimerase DprE1 of *Mycobacterium*

- tuberculosis*. J Am Chem Soc 132:13663–13665
39. Trefzer C (2012) DprE1 as a drug target from *Mycobacterium tuberculosis*. EPFL, Lausanne
 40. Bouhss A, Mengin-Lecreulx D, Blanot D, van Heijenoort J, Parquet C (1997) Invariant amino acids in the Mur peptide synthetases of bacterial peptidoglycan synthesis and their modification by site-directed mutagenesis in the UDP-MurNAc:L-alanine ligase from *Escherichia coli*. Biochemistry 36:11556–11563
 41. Eveland SS, Pompliano DL, Anderson MS (1997) Conditionally lethal *Escherichia coli* murein mutants contain point defects that map to regions conserved among murein and folyl poly-gamma-glutamate ligases: identification of a ligase superfamily. Biochemistry 36:6223–6229
 42. Walsh AW, Falk PJ, Thanassi J, Discotto L, Pucci MJ, Ho HT (1999) Comparison of the D-glutamate adding enzymes from selected gram-positive and gram-negative bacteria. J Bacteriol 181:5395–5401
 43. Zoeiby A, Sanschagrín F, Levesque RC (2003) Structure and function of the Mur enzymes: development of novel inhibitors. Mol Microbiol 47:1–12
 44. Smith CA (2006) Structure, function and dynamics in the mur family of bacterial cell wall ligases. J Mol Biol 362:640–655
 45. Munshi T, Gupta A, Evangelopoulos D, Guzman JD, Gibbons S, Keep NH, Bhakta S (2013) Characterisation of ATP-dependent Mur ligases involved in the biogenesis of cell wall peptidoglycan in *Mycobacterium tuberculosis*. PLoS One 8(3):60143
 46. Eniyan K, Kumar A, Rayasam GV, Perdih A, Bajpai U (2016) Development of a one-pot assay for screening and identification of Mur pathway inhibitors in *Mycobacterium tuberculosis*. Sci Rep 6:35134–35146
 47. Deva T, Baker EN, Squire CJ, Smith CA (2006) Structure of *Escherichia coli* UDP-N-acetylmuramoyl:L-alanine ligase (MurC). Acta Crystallogr D Biol Crystallogr 62:1466–1474
 48. Basavannacharya C, Robertson G, Munshi T, Keep NH, Bhakta S (2010) ATP-dependent MurE ligase in *Mycobacterium tuberculosis*: biochemical and structural characterisation. Tuberculosis 90(1):16–24
 49. Silhavy TJ, Kahne D, Walker S (2010) The bacterial cell envelope. Cold Spring Harb Perspect Biol 2:1–17
 50. Sinha S, Kosalai K, Arora S, Namane A, Sharma P, Gaikwad AN, Brodin P, Cole ST (2005) Immunogenic membrane-associated proteins of *Mycobacterium tuberculosis* revealed by proteomics. Microbiology 151:2411–2419
 51. Singh R, Manjunatha U, Boshoff HI, Ha YH, Niyomrattanakit P, Ledwidge R, Dowd CS, Lee IY, Kim P, Zhang L, Kang S, Keller TH, Jiricek J, Barry CE (2008) PA-824 kills non-replicating *Mycobacterium tuberculosis* by intracellular NO release. Science 322:1392–1395
 52. Manjunatha U, Boshoff HI, Barry CE (2009) The mechanism of action of PA-824: novel insights from transcriptional profiling. Commun Integr Biol 2:215–218
 53. Somasundaram S, Anand RS, Venkatesan P, Paramasivan CN (2013) Bactericidal activity of PA-824 against *Mycobacterium tuberculosis* under anaerobic conditions and computational analysis of its novel analogues against mutant Ddn receptor. Microbiology 13(1):1392–1395
 54. Kim P, Zhang L, Manjunatha UH, Singh R, Patel S, Jiricek J, Keller TH, Boshoff HI, Barry CE III, Dowd CS (2009) Structure–activity relationships of antitubercular nitroimidazoles. Structural features associated with aerobic and anaerobic activities of 4- and 5-nitroimidazoles. J Med Chem 52:1317–1328
 55. Rani C, Mehra R, Sharma R, Chib R, Wazir P, Nargotra A, Khan IA (2015) High-throughput screen identifies small molecule inhibitors targeting acetyltransferase activity of *Mycobacterium tuberculosis* GlmU. Tuberculosis 95(6):664–677
 56. Mengin-Lecreulx D, van Heijenoort J (1994) Copurification of glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase activities of *Escherichia coli*: characterization of the *glmU* gene product as a bifunctional enzyme catalyzing two subsequent steps in the pathway for UDP-N-acetylglucosamine synthesis. J Bacteriol 176:5788–5795
 57. Zhang Z, Bulloch EM, Bunker RD, Baker EN, Squire CJ (2009) Structure and function of GlmU from *Mycobacterium tuberculosis*. Acta Crystallogr D Biol Crystallogr 65:275–283
 58. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, McKenney K, Sutton G, FitzHugh W, Fields C, Gocayne JD, Scott J, Shirley R, Liu LI, Glodek A, Kelley JM, Weidman JF, Phillips CA, Springgs T, Hedblom E, Cotton MD, Utterback TR, Hanna MC, Nguyen DT, Saudek DM, Brandon RC, Fine LD, Fritchman

- JL, Fuhrmann JL, Geoghager NSM, Gnehm CL, McDonald LA, Small KV, Fraser CM, Smith HO, Venter JC (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* RD. *Science* 269:496–512
59. Hove-Jensen B (1992) Identification of *tms-26* as an allele of the *gcd* gene, which encodes *N*-acetylglucosamine-1-phosphate uridyltransferase in *Bacillus subtilis*. *J Bacteriol* 174:6852–6856
60. Olsen LR, Vetting MW, Roderick SL (2007) Structure of the *E. coli* bifunctional GlmU acetyltransferase active site with substrates and products. *Protein Sci* 16:1230–1235
61. Sharma R, Lambu MR, Jamwal U, Rani C, Chib R, Wazir P, Mukherjee D, Chaubey A, Kan IA (2016) *Escherichia coli* *N*-acetylglucosamine-1-phosphate-uridyltransferase/glycosamine-1-phosphate-acetyltransferase (GlmU) inhibitory activity of terreic acid isolated from *Aspergillus terreus*. *J Biomol Screen* 21(4):342–353
62. Gavaldà S, Léger M, van der Rest B, Stella A, Bardou F, Montrozier H, Chalut C, Burlet-Schiltz O, Marrakchi H, Daffé M, Quémar A (2009) The Pks13/FadD32 crosstalk for the biosynthesis of mycolic acids in *Mycobacterium tuberculosis*. *J Biol Chem* 284:19255–19264
63. Maier T, Jenni S, Ban N (2006) Architecture of mammalian fatty acid synthase at 4.5 Å resolution. *Science* 311:1258–1262
64. Khosla C (2009) Structures and mechanisms of polyketide synthases. *J Org Chem* 74:6416–6420
65. Portevin D, de Sousa-D'Auria C, Houssin C, Grimaldi C, Chami M, Daffé M, Guillhot C (2004) A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. *Proc Natl Acad Sci U S A* 101(1):314–319
66. Khosla C, Tang Y, Chen AY, Schnarr NA, Cane DE (2007) Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu Rev Biochem* 76:195–221
67. Tang Y, Chen AY, Kim CY, Cane DE, Khosla C (2007) Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. *Chem Biol* 14:931–943
68. Fernandes GFS, Chin CM, Santos JL (2017) Potenciais alvos moleculares para o desenvolvimento de novos fármacos antituberculose. *Quim Nova* 40(5):572–585
69. Bloch K (2006) Control mechanisms for fatty acid synthesis in *Mycobacterium smegmatis*. Wiley, New York
70. Jackowski S, Rock CO (1987) Acetoacetyl-acyl carrier protein synthase, a potential regulator of fatty acid biosynthesis in bacteria. *J Biol Chem* 262(16):7927–7931
71. Tsay JT, Oh W, Larson TJ, Jackowski S, Rock CO (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia coli* K-12. *J Biol Chem* 267(10):6807–6814
72. Choi KH, Kremer L, Besra G, Rock CO (2000) Identification and substrate specificity of β -ketoacyl (acyl carrier protein) synthase III (*mtFabH*) from *Mycobacterium tuberculosis*. *J Biol Chem* 276(36):28201–28207
73. Jackowski S (1996) Biosynthesis of pantothenic acid and coenzyme A in *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology Press, Washington, DC
74. Merkel WK, Nichols BP (1996) Characterization and sequence of the *Escherichia coli* panBCD gene cluster. *Microbiol Lett* 143:247–252
75. Sledz P, Silvestre L, Hung AW, Ciulli A, Blundell TL, Abell C (2010) Optimization of the interligand overhauser effect for fragment linking: application to inhibitor discovery against *Mycobacterium tuberculosis* pantothenate synthetase. *J Am Chem Soc* 132:4544–4545
76. Zheng R, Blanchard JS (2001) Steady-state and pre-steady-state kinetic analysis of *Mycobacterium tuberculosis* pantothenate synthetase. *Biochemistry* 40:12904–12912
77. von Delft F, Lewendon A, Dhanaraj V, Blundell TL, Abell C, Smith AG (2001) The crystal structure of *E. coli* pantothenate synthetase confirms it as a member of the cytidyltransferase superfamily. *Structure* 9:439–450
78. Genschel U, Powell CA, Abell C, Smith AG (1999) The final step of pantothenate biosynthesis in higher plants: cloning and characterization of pantothenate synthetase from *Lotus japonicus* and *Oryza sativa* (rice). *Biochem J* 341:669–678
79. Espinosa AP, Arjona TR, Rubio MR (2001) Pantothenate synthetase from *Fusarium oxysporum* f. sp. *lycopersici* is induced by α -tomatine. *Mol Genet Genomics* 265(5):922–929
80. Wang S, Eisenberg D (2003) Crystal structures of a pantothenate synthetase from *M. tuberculosis* and its complexes with substrates and a reaction intermediate. *Protein Sci* 12:1097–1108
81. Ciulli A, Scott DE, Ando M, Reyes F, Saldanha SA, Tuck KL, Chirgadze DY, Blundell

- TL, Abell C (2008) Inhibition of *Mycobacterium tuberculosis* pantothenate synthetase by analogues of the reaction intermediate. *Chembiochem* 9(16):2606–2611
82. Tuck KL, Saldanha SA, Birch LM, Smith AG, Abell C (2006) The design and synthesis of inhibitors of pantothenate synthetase. *Org Biomol Chem* 4(19):3598–3610
 83. Wong D, Chao JD, Av-Gay Y (2013) *Mycobacterium tuberculosis*-secreted phosphatases: from pathogenesis to targets for TB drug development. *Trends Microbiol* 21(2):100–109
 84. Zhang ZY (2002) Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development. *Annu Rev Pharmacol Toxicol* 42(1):209–234
 85. Kennelly PJ, Potts M (1999) Life among the primitives: protein O-phosphatases in prokaryotes. *Front Biosci* 4:372–385
 86. Shi L, Potts M, Kennelly PJ (1998) The serine, threonine, and/or tyrosine-specific protein kinases and protein phosphatases of prokaryotic organisms: a family portrait. *Microbiol Rev* 22(4):229–253
 87. Huyter G, Liu S, Kelly J, Moffat J, Payette P, Kennedy B, Tsaprailis G, Gresse MJ, Ramachandran C (1997) Mechanism of inhibition of protein-tyrosine phosphatases by vanadate and pervanadate. *J Biol Chem* 272(2):843–851
 88. Tracey AS (2000) Hydroxamido vanadates: aqueous chemistry and function in protein tyrosine phosphatases and cell cultures. *J Inorg Biochem* 80(1):11–16
 89. Bach H, Sun J, Hmama Z, Av-Gay Y (2006) *Mycobacterium avium* ssp paratuberculosis PtpA is an endogenous tyrosine phosphatase secreted during infection. *Infect Immun* 74:6540–6546
 90. Cowley SC, Babakaiff R, Av-Gay Y (2002) Expression and localization of the *Mycobacterium tuberculosis* protein tyrosine phosphatase PtpA. *Res Microbiol* 153(4):233–241
 91. Li Y, Strohl WR (1996) Cloning, purification, and properties of a phosphotyrosine protein phosphatase from *Streptomyces coelicolor* A3(2). *J Bacteriol* 178(1):136–142
 92. Takayama K, Wang C, Besra GS (2005) Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin Microbiol Rev* 18:81–101
 93. Campbell JW, Cronan JE Jr (2001) Bacterial fatty acid biosynthesis: targets for antibacterial drug discovery. *Annu Rev Microbiol* 55:305–332
 94. White SW, Zheng J, Zhang YM, Rock CO (2005) The structural biology of type II fatty acid biosynthesis. *Annu Rev Biochem* 74:791–831
 95. Quemard A, Sacchettini JC, Dessen A, Vilchezes C, Bittman R, Jacobs WR Jr, Blanchard JS (1995) Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. *Biochemistry* 34:8235–8241
 96. Dessen A, Quemard A, Blanchard JS, Jacobs WR Jr, Sacchettini JC (1995) Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*. *Science* 267:1638–1641
 97. Wright GD (2012) Back to the future: a new ‘old’ lead for tuberculosis. *EMBO Mol Med* 4:1029–1031
 98. Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um KS, Wilson T, Collins D, Lisle G, Jacobs WR Jr (1994) *InhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science* 263:227–230
 99. Zabinski RF, Blanchard JS (1997) The requirement for manganese and oxygen in the isoniazid-dependent inactivation of *Mycobacterium tuberculosis* enoyl reductase. *J Am Chem Soc* 119:2331–2332
 100. Rozwarski DA, Vilchèze C, Sugantino M, Bittman R, Sacchettini JC (1999) Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, *InhA*, in complex with NAD⁺ and a C16 fatty acyl substrate. *J Biol Chem* 274(22):15582–15589
 101. Musser JM, Kapur V, Williams DL, Kreiswirth BN, Van Soolingen D, Van Embden JD (1996) Characterization of the catalase-peroxidase gene (*katG*) and *inhA* locus in isoniazid-resistant and-susceptible strains of *Mycobacterium tuberculosis* by automated DNA sequencing: restricted array of mutations associated with drug resistance. *J Infect Dis* 173(1):196–202
 102. McLean KJ, Sabri M, Marshall KR, Lawson RJ, Lewis DG, Clift D, Balding PR, Dunford AJ, Warman AJ, McVey JP, Quin AM, Sutcliffe MJ, Scrutton NS, Munro AW (2005) Biodiversity of cytochrome P450 redox systems. *Biochem Soc Trans* 33:796–801
 103. Guengerich FP (2001) Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem Res Toxicol* 14(6):611–650
 104. Reichhart DW, Feyereisen R (2000) Cytochromes P450: a success story. *Genome Biol* 1(6):3001–3009

105. Quehl P, Hollender J, Schüürmann J, Brossette T, Maas R, Jose J (2016) Co-expression of active human cytochrome P450 1A2 and cytochrome P450 reductase on the cell surface of *Escherichia coli*. *Microb Cell Factories* 15(1):26–51
106. Munro AW, Lindsay JG (1996) Bacterial cytochromes P-450. *Mol Microbiol* 20(6):1115–1125
107. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O’Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, Hopwood DA (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2). *Nature* 417(6885):141–147
108. Notheis C, Drewke C, Leistner E (1997) Purification and characterization of the pyridoxol-5-phosphate:oxygen oxidoreductase (deaminating) from *Escherichia coli*. *Biochim Biophys Acta* 1247(2):265–271
109. Zhao G, Winkler ME (1995) Kinetic limitation and cellular amount of pyridoxine (pyridoxamine) 5'-phosphate oxidase of *Escherichia coli* K-12. *J Bacteriol* 177:883–891
110. Kwok F, Churchich JE (1980) Interaction between pyridoxal kinase and pyridoxine-5'-P oxidase, two enzymes involved in the metabolism of vitamin B6. *J Biol Chem* 255:882–887
111. Salvo M, Yang E, Zhao G, Winkler ME, Schirch V (1998) Expression, purification, and characterization of recombinant *Escherichia coli* pyridoxine 5'-phosphate oxidase. *Protein Expr Purif* 13:349–356
112. Pédelacq JD, Rho BS, Kim CY, Waldo GS, Lakin TP, Segelke BW, Rupp B, Hung LW, Kim S, Terwilliger TC (2006) Crystal structure of a putative pyridoxine 5'-phosphate oxidase (Rv2607) from *Mycobacterium tuberculosis*. *Proteins* 62(3):563–569
113. Mashalidis EH, Mukherjee T, Sledz P, Vinkovic DM, Boshoff H, Abell C, Barry CE (2011) Rv2607 from *Mycobacterium tuberculosis* is a pyridoxine 5'-phosphate oxidase with unusual substrate specificity. *PLoS One* 6(11):27643–27650
114. Sang Y, Barbosa JM, Wu H, Locy RD, Singh NK (2007) Identification of a pyridoxine (pyridoxamine) 5'-phosphate oxidase from *Arabidopsis thaliana*. *FEBS Lett* 581(3):344–348
115. Lehmann HM, Chaffotte A, Pochet S, Labesse G (2001) Thymidylate kinase of *Mycobacterium tuberculosis*: a chimera sharing properties common to eukaryotic and bacterial enzymes. *Protein Sci* 10:1195–1205
116. Sierra IL, Lehmann HM, Gilles AM, Bâzcu O, Delarue M (2001) X-ray structure of TMP kinase from *Mycobacterium tuberculosis* complexed with TMP at 1.95 Å resolution. *J Mol Biol* 311(1):87–100
117. Shmalenyuk ER, Kochetkov SN, Alexandrova LA (2013) Novel inhibitors of *Mycobacterium tuberculosis* growth based on modified pyrimidine nucleosides and their analogues. *Russ Chem Rev* 82(9):896–915
118. Shukla H, Kumar V, Singh AK, Singh N, Kashif M, Siddiqi MI, Krishnan MY, Akhtar MS (2015) Insight into the structural flexibility and function of *Mycobacterium tuberculosis* isocitrate lyase. *Biochimie* 110:73–80
119. Bhusal RP, Bashiri G, Kwai BXC, Sperry J, Leung IKH (2017) Targeting isocitrate lyase for the treatment of latent tuberculosis. *Drug Discov Today* 22(7):1008–1016
120. Cheah HL, Lim V, Sandai D (2014) Inhibitors of the Glyoxylate cycle enzyme ICL1 in *Candida albicans* for potential use as antifungal agents. *PLoS One* 9(4):95951–95959
121. Fleischmann RD, Alland D, Eisen JA, Carpenter L, White O, Peterson J, DeBoy R, Dodson R, Gwinn M, Haft D, Hickey E, Kolonay JF, Nelson WC, Umayam LA, Ermolaeva M, Salzberg SL, Delcher A, Utterback T, Weidman J, Khouri H, Gill J, Mikula A, Bishai W, Jacobs WR, Venter JC, Fraser CM (2002) Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J Bacteriol* 184(19):5479–5490
122. Bentrup KHZ, Miczak A, Swenson DL, Russell DG (1999) Characterization of activity and expression of Isocitrate Lyase in *Mycobacterium avium* and *Mycobacterium tuberculosis*. *J Bacteriol* 181(23):7161–7167
123. Odell LR, Nilsson MT, Gising J, Lagerlund O, Muthas D, Nordqvist A, Karlen A, Larhed M (2009) Functionalized 3-amino-imidazo [1, 2- α] pyridines: a novel class of drug-like *Mycobacterium tuberculosis* glutamine synthetase inhibitors. *Bioorg Med Chem Lett* 19(16):4790–4793

124. Yuan J, Doucette CD, Fowler WU, Feng XJ, Piazza M, Wingreen HNS, Rabinowitz JD (2009) Metabolomics-driven quantitative analysis of ammonia assimilation in *E. coli*. *Mol Syst Biol* 5:302–318
125. Harth G, Clemens DL, Horwitz MA (1994) Glutamine synthetase of *Mycobacterium tuberculosis*: extracellular release and characterization of its enzymatic activity. *Proc Natl Acad Sci U S A* 91(20):9342–9346
126. Tullius MV, Harth G, Horwitz MA (2003) Glutamine synthetase GlnA1 is essential for growth of *Mycobacterium tuberculosis* in human THP-1 macrophages and Guinea pigs. *Infect Immun* 71(7):3927–3936
127. Tullius MV, Harth G, Horwitz MA (2001) High extracellular levels of *Mycobacterium tuberculosis* glutamine synthetase and superoxide dismutase in actively growing cultures are due to high expression and extracellular stability rather than to a protein-specific export mechanism. *Infect Immun* 69(10):6348–6363
128. Herrmann KM, Weaver LM (1999) The shikimate pathway. *Annu Rev Plant Biol* 50(1):473–503
129. Meibom KL, Charbit A (2010) *Francisella tularensis* metabolism and its relation to virulence. *Front Microbiol* 1:140–153
130. Roberts F, Roberts CW, Johnson JJ, Kyle DE, Krell T, Coggins JR, Coombs GH, Milhous WK, Tzipori S, Ferguson DJ, Chakrabarti D, McLeod R (1998) Evidence for the shikimate pathway in apicomplexan parasites. *Nature* 393(6687):801–805
131. Ducati RG, Basso LA, Santos DS (2007) Mycobacterial shikimate pathway enzymes as targets for drug design. *Curr Drug Targets* 8:423–435
132. Hartmann MD, Bourenkov GP, Oberschall A, Strizhov N, Bartunik HD (2006) Mechanism of phosphoryl transfer catalyzed by shikimate kinase from *Mycobacterium tuberculosis*. *J Mol Biol* 364(3):411–423
133. Pallen MJ (2002) The ESAT-6/WXG100 superfamily – and a new Grampositive secretion system? *Trends Microbiol* 10:209–212
134. Brodin P, Jonge MI, Majlessi L, Leclerc C, Nilges M, Cole ST, Brosch R (2005) Functional analysis of early secreted antigenic target-6, the dominant T-cell antigen of *Mycobacterium tuberculosis*, reveals key residues involved in secretion, complex formation, virulence, and immunogenicity. *J Biol Chem* 280(40):33953–33959
135. Renshaw PS, Lightbody KL, Veverka V, Muskett FW, Kelly G, Frenkiel TA, Gordon SV, Hewinson RG, Burke B, Norman J, Williamson RA, Carr MD (2005) Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. *EMBO J* 24:2491–2498
136. Guinn KM, Hickey MJ, Mathur SK, Zakei KL, Grotzke JE, Lewinson DM, Smith S, Sherman DR (2004) Individual RDI-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* 51(2):359–370
137. Renshaw PS, Veverka V, Kelly G, Frenkiel TA, Williamson RA, Gordon SV, Hewinson RG, Carr MD (2004) Letter to the editor: Sequence-specific assignment and secondary structure determination of the 195-residue complex formed by the *Mycobacterium tuberculosis* proteins CFP-10 and ESAT-6. *J Biol NMR* 30:225–226
138. Meher AK, Bal NC, Chary KV, Arora A (2006) *Mycobacterium tuberculosis* H37Rv ESAT-6-CFP-10 complex formation confers thermodynamic and biochemical stability. *FEBS J* 273(7):1445–1462
139. Berthet FX, Rasmussen PB, Rosenkrands I, Andersen P, Gicquel B (1998) A *Mycobacterium tuberculosis* operon encoding ESAT6 and a novel low-molecular-mass culture filtrate protein (CFP-10). *Microbiology* 144(11):3195–3203
140. Wards BJ, De Lisle GW, Collins DM (2000) An ESAT-6 knockout mutant of *Mycobacterium bovis* produced by homologous recombination will contribute to the development of a live tuberculosis vaccine. *Tuber Lung Dis* 80(5):185–189
141. Nollmann M, Crisona NJ, Arimondo PB (2007) Thirty years of *Escherichia coli* DNA gyrase: from in vivo function to single molecule mechanism. *Biochimie* 89:490–499
142. Bates AD, Maxwell A (2007) Energy coupling in type II topoisomerases: why do they hydrolyze ATP? *Biochemistry* 46:7929–7941
143. Stanger FV, Dehio C, Schirmer T (2014) Structure of the N-terminal Gyrase B fragment in complex with ADPPi reveals rigid-body motion induced by ATP hydrolysis. *PLoS One* 9(9):107289–107302
144. Champoux JJ (2001) DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem* 70:369–413
145. Chopra S, Matsuyama K, Tran T, Malerich JP, Wan B, Franzblau SG, Lun S, Guo H, Maiga MC, Bishai WR, Madrid PB (2012) Evaluation of gyrase B as a drug target in *Mycobacterium tuberculosis*. *J Antimicrob Chemother* 67(2):415–421

146. Mugumbate G, Abrahams KA, Cox JAG, Papadatos G, Westen G, Lelièvre J, Calus ST, Loman NJ, Ballell L, Barros D, Overington JP, Besra GS (2015) Mycobacterial dihydrofolate reductase inhibitors identified using chemogenomic methods and *in vitro* validation. *PLoS One* 10(2):121492
147. Feng J, Goswami S, Howell EE (2008) R67, the other dihydrofolate reductase: rational design of an alternate active site configuration. *Biochemistry* 47:555–565
148. Argyrou A, Vetting MW, Aladegbami B, Blanchard JS (2006) *Mycobacterium tuberculosis* dihydrofolate reductase is a target for isoniazid. *Nat Struct Mol Biol* 13:408–413
149. El-Subbagh HI, Hassan GS, El-Messery SM, Al-Rashood ST, Al-Omary FA, Abulfadl YS, Shabayek MI (2014) Nonclassical antifolates: part 5. Benzodiazepine analogs as a new class of DHFR inhibitors: synthesis, antitumor testing and molecular modeling study. *Eur J Med Chem* 74:234–245
150. Li R, Sirawaraporn R, Chitnumsub P, Sirawaraporn W, Wooden J, Athappilly F, Turley S, Hol WG (2000) Three-dimensional structure of *M. tuberculosis* dihydrofolate reductase reveals opportunities for the design of novel tuberculosis drugs. *J Mol Biol* 295:307–323
151. Cody V, Galitsky N, Luft JR, Pangborn W, Rosowsky A, Blakley RL (1997) Comparison of two independent crystal structures of human dihydrofolate reductase ternary complexes reduced with nicotinamide adenine dinucleotide phosphate and the very tight binding inhibitor PT523. *Biochemistry* 36:4399–4411
152. White EL, Ross LJ, Cunningham A, Escuyer V (2004) Cloning, expression and characterization of *Mycobacterium tuberculosis* dihydrofolate reductase. *FEMS Microbiol Lett* 232:101–105
153. Sardarian A, Douglas KT, Read M, Sims PFG, Hyde JE, Chitnumsub P, Sirawaraporn R, Sirawaraporn W (2003) Pyrimethamine analogues as strong inhibitors of double and quadruple mutants of dihydrofolate reductase in human malaria parasites. *Org Biomol Chem* 1:960–964
154. Morphy JR (2012) The challenges of multi-target lead optimization. In: Morphy JR, Harris CJ (eds) *Designing multi-target drugs*. The Royal Society of Chemistry, London
155. Abdolmaleki A, Ghasemi JB, Ghasemi F (2017) Computer aided drug design for multi-target drug design: SAR/QSAR, molecular docking and pharmacophore methods. *Curr Drug Targets* 18(5):556–575
156. Mohareb RM, El-Sayed NNE, Abdelaziz MA (2013) The Knoevenagel reactions of pregnenolone with cyanomethylene reagents: synthesis of thiophene, thieno[2,3-b]pyridine, thieno[3,2-d]isoxazole derivatives of pregnenolone and their *in vitro* cytotoxicity towards tumor and normal cell lines. *Steroids* 78:1209–1219
157. Meotti FC, Silva DO, Santos ARS, Zeni G, Rocha JBT, Nogueira CW (2003) Thiophenes and furans derivatives: a new class of potential pharmacological agents. *Environ Toxicol Pharmacol* 15(1):37–44
158. Ferreira AP, da Silva JLF, Duarte MT, da Piedade MFM, Robalo MP, Harjivan SG, Marzano C, Gandin V, Marques MM (2009) Synthesis and characterization of new organometallic benzo[*b*]thiophene derivatives with potential antitumor properties. *Organometallics* 28(18):5412–5423
159. Khalil AM, Berghot MA, Gouda MA (2009) Synthesis and antibacterial activity of some new thiazole and thiophene derivatives. *Eur J Med Chem* 44(11):4434–4440
160. Lu X, Wan B, Franzblau SG, You Q (2011) Design, synthesis and anti-tubercular evaluation of new 2-acylated and 2-alkylated amino-5-(4-(benzyloxy) phenyl) thiophene-3-carboxylic acid derivatives. Part 1. *Eur J Med Chem* 46(9):3551–3563
161. Wilson R, Kumar P, Parashar V, Vilchère C, Veyron-Churlet R, Freundlich JS, Barnes SW, Walker JR, Szymonifka J, Marchiano E, Shenai S, Colangeli R, Jacobs WR, Neiditch MB, Kremer L, Alland D (2013) Antituberculosis thiophenes define a requirement for Pks13 in mycolic acid biosynthesis. *Nat Chem Biol* 9(8):499–506
162. Mahajan PS, Nikam MD, Nawale LU, Khedkar VM, Sarkar D, Gill CH (2016) Synthesis and antitubercular activity of new benzo[*b*]thiophenes. *ACS Med Chem Lett* 7:751–756
163. Leitans J, Sprudza A, Tanc M, Vozny I, Zalubovskis R, Tars K, Supuran CT (2013) 5-Substituted-(1, 2, 3-triazol-4-yl) thiophene-2-sulfonamides strongly inhibit human carbonic anhydrases I, II, IX and XII: solution and X-ray crystallographic studies. *Bioorg Med Chem* 21(17):5130–5138
164. Wang ZC, Qin YJ, Wang PF, Yang YA, Wen Q, Zhang X, Qiu HY, Duan YT, Wang YT, Sang YL, Zhu HL (2013) Sulfonamides containing coumarin moieties selectively and potently inhibit carbonic anhydrases II and IX: design, synthesis, inhibitory activity and 3D-QSAR analysis. *Eur J Med Chem* 66:1–11

165. Supuran CT, Casini A, Scozzafava A (2003) Protease inhibitors of the sulfonamide type: anticancer, antiinflammatory, and antiviral agents. *Med Res Rev* 5:535–558
166. Scozzafava A (2003) Anticancer and antiviral sulfonamides. *Curr Med Chem* 10:925–953
167. Özbek N, Katircioglu H, Karacan N, Baykal T (2007) Synthesis, characterization and antimicrobial activity of new aliphatic sulfonamide. *Bioorg Med Chem* 15 (15):5105–5109
168. Naidu KM, Nagesh HN, Singh M, Sriram D, Yogeeswari P, Sekhar KVG (2015) Novel amide and sulfonamide derivatives of 6-(piperazin-1-yl) phenanthridine as potent *Mycobacterium tuberculosis* H37Rv inhibitors. *Eur J Med Chem* 92:415–426
169. Oliveira KN, Chiaradia LD, Martins PGA, Mascarello A, Cordeiro MNS, Guido RVC, Andricopulo AD, Yunes RA, Nunes RJ, Vernal J, Terenzi H (2011) Sulfonylhydrazones of cyclic imides derivatives as potent inhibitors of the *Mycobacterium tuberculosis* protein tyrosine phosphatase B (PtpB). *MedChemComm* 2(6):500–504
170. Elias DW, Beazely MA, Kandepu NM (1999) Bioactivities of chalcones. *Curr Med Chem* 6 (12):1125
171. Aoki N, Muko M, Ohta E, Ohta S (2008) C-geranylated chalcones from the stems of *Angelica keiskei* with superoxide-scavenging activity. *J Nat Prod* 71(7):1308–1310
172. Mascarello A, Chiaradia LD, Vernal J, Villarino A, Guido RV, Perizzolo P, Poirier V, Wong D, Martins PGA, Nunes RJ, Yunes RA, Andricopulo AD, Gay YA, Terenzi H (2010) Inhibition of *Mycobacterium tuberculosis* tyrosine phosphatase PtpA by synthetic chalcones: kinetics, molecular modeling, toxicity and effect on growth. *Bioorg Med Chem* 18(11):3783–3789
173. Yadav DK, Ahmad I, Shukla A, Khan F, Negi AS, Gupta A (2014) QSAR and docking studies on chalcone derivatives for antitubercular activity against *M. tuberculosis* H37Rv. *J Chemom* 28(6):499–507
174. Gond DS, Meshram RJ, Jadhav SG, Wadhwa G, Gacche RN (2013) *In silico* screening of chalcone derivatives as potential inhibitors of dihydrofolate reductase: assessment using molecular docking, paired potential and molecular hydrophobic potential studies. *Drug Invent Today* 5(3):182–191
175. Carroll MW, Jeon D, Mountz JM, Lee JD, Jeong YJ, Zia N, Lee M, Lee J, Via LE, Lee S, Eum SY, Lee SJ, Goldfeder LC, Cai Y, Jin B, Kim Y, Oh T, Chen RY, Dodd LE, Gu W, Dartois V, Park SK, Kim CT, Barry CE III, Cho SN (2013) Efficacy and safety of metro-nidazole for pulmonary multidrug-resistant tuberculosis. *Antimicrob Agents Chemother* 57:3903–3909
176. Mukherjee T, Boshoff H (2011) Nitroimidazoles for the treatment of TB: past, present and future. *Future Med Chem* 3:1427–1454
177. Sharma A, Subbias KK, Robine O, Chaturvedi I, Nigam A, Sharma N, Chaudhary PP (2012) Computational finding of potential inhibitor for cytochrome P450 mono-oxygenases enzyme of *Mycobacterium tuberculosis*. *Bioinformatics* 8(19):931
178. Gupta N, Vyas VK, Patel BD, Ghate M (2017) Design of 2-nitroimidazooxazine derivatives as deazaflavin-dependent nitroreductase (Ddn) activators as antimycobacterial agents based on 3D QSAR, HQSAR, and docking study with *in silico* prediction of activity and toxicity. *Comput Life Sci* 1–15
179. Somasundaram S, Anand RS, Venkatesan P, Paramasivan CN (2013) Bactericidal activity of PA-824 against *Mycobacterium tuberculosis* under anaerobic conditions and computational analysis of its novel analogues against mutant Ddn receptor. *BMC Microbiol* 13 (1):218
180. Kumar SM, Jaleel UCA (2016) Molecular docking studies of PA-824 with pyridoxine 5'-phosphate oxidase. *Biol Med* 8(2):1–4
181. Keri RS, Hiremathad A, Budagumpi S, Nagaraja BM (2015) Comprehensive review in current developments of benzimidazole-based medicinal chemistry. *Chem Biol Drug Des* 86(1):19–65
182. Grassi A, Ippen J, Bruno M, Thomas G (1991) BAY P 1455, a thiazolylaminobenzimidazole derivative with gastroprotective properties in the rat. *Eur J Pharmacol* 195 (2):251–259
183. Saleshier FM, Divakar MC (2011) Design, docking and synthesis of some 6-benzimidazolyl pyrans and screening of their anti tubercular activity. *Eur J Exp Biol* 1(2):150–159
184. Soni V, Suryadevara P, Sriram D, Kumar S, Nandicoori VK, Yogeeswari P, OSDD Consortium (2015) Structure-based design of diverse inhibitors of *Mycobacterium tuberculosis* N-acetylglucosamine-1-phosphate uridyltransferase: combined molecular docking, dynamic simulation, and biological activity. *J Mol Model* 21(7):174
185. Priyadarsini R, Tharani CB, Niraimathi V (2012) Docking, synthesis, characterisation

- of certain substituted benzimidazole pyrimidines as potent DHFR inhibitors with antimycobacterial activity. *J Pharm Res* 5 (11):5173–5177
186. Aanandhi MV, Chandran M, Sujatha R, Nishanthi M, Vijayakumar B (2012) Comparative study on conventional and microwave assisted synthesis and molecular docking studies of mannich bases of benzimidazoles and their derivatives. *Int J Biol Pharm Res* 3 (4):1–9
187. Anguru MR, Taduri AK, Bhoomireddy RD, Jojula M, Gunda SK (2017) Novel drug targets for *Mycobacterium tuberculosis*: 2-heterostyrylbenzimidazoles as inhibitors of cell wall protein synthesis. *Chem Central J* 11 (1):68
188. Otvos L (2008) Peptide-based drug design: here and now. *Methods Mol Biol* 494:1–8
189. Hancock R, Bertrand HC, Tsujita T, Naz S, El-Bakry A, Laoruchupong J, Hayes JD, Wells G (2012) Peptide inhibitors of the Keap1-Nrf2 protein-protein interaction. *Free Radic Biol Med* 52:444–451
190. Shi W, Ma H, Duan Y, Aubart K, Fang Y, Zonis R, Yang L, Hu W (2011) Design, synthesis and antibacterial activity of 3-methylenepyrrolidine formyl hydroxyamino derivatives as novel peptide deformylase inhibitors. *Bioorg Med Chem Lett* 21:1060–1063
191. Kumar M, Vijaykrishnan R, Rao GS (2010) *In silico* structure-based design of a novel class of potent and selective small peptide inhibitor of *Mycobacterium tuberculosis* dihydrofolate reductase, a potential target for anti-TB drug discovery. *Mol Divers* 14(3):595–604
192. Chandra K, Dutta D, Das AK, Basak A (2010) Design, synthesis and inhibition activity of novel cyclic peptides against protein tyrosine phosphatase A from *Mycobacterium tuberculosis*. *Bioorg Med Chem* 18(23):8365–8373
193. Yang Y, Gao P, Liu Y, Ji X, Gan M, Guan Y, Hao X, Li Z, Xiao C (2011) A discovery of novel *Mycobacterium tuberculosis* pantothenate synthetase inhibitors based on the molecular mechanism of actinomycin D inhibition. *Bioorg Med Chem Lett* 21 (13):3943–3946
194. Kumar M, Sharma S, Srinivasan A, Singh TP, Kaur P (2011) Structure-based *in silico* rational design of a selective peptide inhibitor for thymidine monophosphate kinase of *Mycobacterium tuberculosis*. *J Mol Model* 17 (5):1173–1182
195. Liu X, Zang Y, Sun B, Yin Y (2014) Optimization of phage heptapeptide library-screening process for developing inhibitors of the isocitrate lyase homologue from *Mycobacterium tuberculosis*. *Med Chem Res* 23 (5):2543–2553
196. Perumal P, Pandey VP, Parasuraman P (2014) Docking studies on antimicrobial peptides related to apidaecinia and human histatin against glutamine synthetase and RNA polymerase in *Mycobacterium tuberculosis*. *Asian J Pharm Clin Res* 7(5):195–201
197. Kumar M, Verma S, Sharma S, Srinivasan A, Singh TP, Kaur P (2010) Structure-based *In silico* design of a high-affinity dipeptide inhibitor for novel protein drug target shikimate kinase of *Mycobacterium tuberculosis*. *Chem Biol Drug Des* 76(3):277–284
198. Gavrish E, Sit CS, Cao S, Kandror O, Spoering A, Peoples A, Ling L, Fetterman A, Hughes D, Bissel A, Torrey H, Akopian T, Mueller A, Epstein S, Goldberg A, Clardy J, Lewis K (2014) Lassomycin, a ribosomally synthesized cyclic peptide, kills *Mycobacterium tuberculosis* by targeting the ATP-dependent protease ClpC1P1P2. *Chem Biol* 21(4):509–518
199. Hu W, Lin JP, Song LR, Long YQ (2015) Direct synthesis of 2-aryl-4-quinolones via transition-metal-free intramolecular oxidative C (sp³)-H/C(sp³)-H coupling. *Org Lett* 17:1268–1271
200. Mitscher LA (2005) Bacterial topoisomerase inhibitors: quinolone and pyridone antibacterial agents. *Chem Rev* 105:559–592
201. Guo X, Liu ML, Guo HY, Wang YC, Wang JX (2011) Synthesis and *in vitro* antibacterial activity of 7-(3-amino-6,7-dihydro-2-methyl-2H-pyrazolo[4,3-c]pyridin-5(4H)-yl) fluoroquinolone derivatives. *Molecules* 16:2626–2635
202. Aldred KJ, Kerns RJ, Osheroff N (2014) Mechanism of quinolone action and resistance. *Biochemistry* 53:1565–1574
203. Guzman JD, Wube A, Evangelopoulos D, Gupta A, Hüfner A, Basavannacharya C, Rahman M, Thomaschitz C, Bauer R, McHugh TD, Nobeli I, Prieto JM, Gibbons S, Bucar F, Bhakta S (2011) Interaction of N-methyl-2-alkenyl-4-quinolones with ATP-dependent MurE ligase of *Mycobacterium tuberculosis*: antibacterial activity, molecular docking and inhibition kinetics. *J Antimicrob Chemother* 66(8):1766–1772
204. Cunha EF, Barbosa EF, Oliveira AA, Ramalho TC (2010) Molecular modeling of *Mycobacterium tuberculosis* DNA gyrase and its molecular docking study with gatifloxacin inhibitors. *J Biomol Struct Dyn* 27 (5):619–625

205. Maddela S, Makula A (2016) Design, synthesis and docking study of some novel isatin-quinoline hybrids as potential antitubercular agents. *Anti-Infect Agents* 14(1):53–62
206. Minovski N, Perdih A, Solmajer T (2012) Combinatorially-generated library of 6-fluoroquinolone analogs as potential novel antitubercular agents: a chemometric and molecular modeling assessment. *J Mol Model* 18(5):1735–1753



Advanced Chemometric Modeling Approaches for the Design of Multitarget Drugs Against Neurodegenerative Diseases

Amit Kumar Halder, Ana S. Moura, and M. Natália D. S. Cordeiro

Abstract

Neurodegenerative diseases (ND), a major worldwide health problem, present a multifactorial nature. This implies that a multitargeted therapy approach can be considered more effective in such cases when comparing with “one drug-one target” based therapies. Multitarget drugs interact simultaneously with two or more therapeutic targets, thus acting synergistically to improve the disease conditions. This chapter discusses the recent advances in chemometric techniques in multitarget anti-ND drug design. After a brief introduction to the most relevant pathophysiological aspects of some common neurodegenerative diseases, it analyses not only pathophysiology versus therapeutic targets but also conventional versus novel chemometric techniques within such context. The emergence of novel and various chemometric techniques undoubtedly contributed to the design of multitarget-directed ligands (MTDLs) over the last decade, laying emphasis on the sound prospective for future therapeutics regarding diseases such as Alzheimer’s and Parkinson’s disease.

Keywords Chemometrics, Multitarget-directed ligands (MTDLs), Multitargeted therapies, Neurodegenerative diseases, QSAR

1 Introduction

The wide spectrum of neurological disorders, characterized by gradual loss of neuronal integrity and functions until neuronal demise, encompasses the neurodegenerative disorders (NDs), such as Alzheimer’s disease (AD), Parkinson’s disease (PD), or Huntington’s disease (HD), whose several symptoms include cognitive decline, dementia or even breathing problems, and loss of motor functions, compromising life itself as well as its quality [1–3]. Further, as the human life expectancies have increased, and at the same time neurodegeneration has become more common as age progresses, NDs are increasing their incidences in alarming proportions. In fact, in 2016, the World Health Organization (WHO) estimated that 47.5 million people are suffering from dementia worldwide and this number is expected to increase to around 75 million by 2030 and 135 million by 2050 [4–6].

Current therapies of NDs are based on symptomatic treatments, i.e., relief from symptoms, and, as such, there is an emerging and urgent need for the development of new ND therapeutics [7]. Recent years saw a surge of new therapeutic targets for NDs taking into account the complex multifactorial etiologies, with the development of multitarget-directed ligands (MTDLs) being considered as a promising strategy to achieve therapeutic success [8–13].

Rational drug discovery is a complex, exhaustive, and expensive process, and computer-aided drug design (CADD)-based techniques are frequently utilized in both industrial and academic platforms to assist the traditional experimental drug discovery [14–17]. Chemometric techniques are frequently used in CADD. Lavine et al. considered that “A chemometric based approach to scientific problem solving attempts to explore the implications of data so that hypotheses, i.e., models of the data, are developed with greater awareness of reality” [18–20]. Chemometrics itself presents branches and methodologies which are frequently used in drug design, such as chemical database storage and recovery, quantitative structure-activity relationships (QSAR), or virtual screening [18, 21].

In this chapter, focus is made on the recent development of chemometric modeling approaches for the design of MTDLs against different NDs, without overlooking the neurodegeneration context and the most interesting issues regarding the subject, either considering a health sciences or computational point of views. The chapter comprises the following main themes: (1) pathological characteristics of and currently available treatment options for common NDs, as presented in Subheading 2; (2) multitarget drugs for NDs, dealt with in Subheading 3; (3) chemometric methods in polypharmacology, discussed in Subheading 4; and (4) chemometric modeling approaches for design of anti-ND MTDLs, in Subheading 5, which presents the state-of-the-art analysis of the main subject of the chapter. A final section, Subheading 6, presents the future directions of the research in this area and some important conclusions. It should be noted that themes (1), (2), and (3) provide the necessary background for understanding not only the current state of in silico methods in NDs and MTDLs context but chemometrics in general, while the theme (4) represents the core of the discussion.

2 Pathological Characteristics and Current Available Treatment Options for Common NDs

In order to fully grasp the intention and scope of application of the chemometric techniques and models presented in this chapter, a brief introduction to the main pathologic features and a summarized survey to the current available treatments for common NDs are made in this section. It should be noted that several NDs share

Table 1
Common neurodegenerative diseases and main pathologic features

Neurodegenerative disease	Pathologic features
Alzheimer's disease (AD)	<ul style="list-style-type: none"> • Neuron loss <ul style="list-style-type: none"> – Most affected brain regions: (1) cortex, (2) hippocampus, (3) basal forebrain, and (4) brain stem • Depositions of extracellular plaques <ul style="list-style-type: none"> – Cause: β-amyloid peptide formation • Intracellular neurofibrillary tangles <ul style="list-style-type: none"> – Cause: hyper phosphorylated amyloid tau protein formation in the CNS
Parkinson's disease (PD)	<ul style="list-style-type: none"> • Neuron loss <ul style="list-style-type: none"> – Most affected regions: (1) substantial nigra, (2) cortex, (3) locus ceruleus, and (4) raphe nuclei • Most common cause: mutations in LRRK2 and parkin
Huntington's disease (HD)	<ul style="list-style-type: none"> • Cause: elongation of trinucleotide CAG on the short arm of chromosome 4p16.3 in the Huntingtin gene (HD or IT15 gene)
Amyotrophic lateral sclerosis (ALS)	<ul style="list-style-type: none"> • Most common cause: mutation of the gene encoding the SOD • Pathogenesis factors: (1) glutamate excitotoxicity, (2) mitochondrial abnormalities, (3) impaired axonal structure, and (4) free radical-mediated oxidative stress
Prion disease (PD)	<ul style="list-style-type: none"> • Cause: mutations in prion gene <ul style="list-style-type: none"> – Conformationally modified prion protein damages the neuronal cells of cortex, thalamus, brain stem, and cerebellum of mammalian CNS • Can be transmitted through infected brain tissue, or surgical tools

common pathologic features, such as the loss of neuronal cells or the deposition and spread of pathogenic misfolded protein aggregates, the latter due to spontaneous or inherited mutations and a main causative factor in neurodegeneration cases [22, 23].

In Table 1, the main pathologic factors for common NDs are displayed. In the first row is presented Alzheimer's disease (AD), which is reported as the sixth leading cause of death in the United States and related to progressive deterioration of cognition, memory, and executive task and speech performance, affecting not only approximately 44 million people worldwide but also presenting 70% death rate increase over the past decade [24, 25]. Among its pathological characteristics, there are depositions of extracellular plaques due to β -amyloid peptide formation, derived from the cleavage of the amyloid precursor protein (APP), and intracellular neurofibrillary tangles caused by the hyper phosphorylated amyloid tau protein formation in the central nervous system (CNS) [26, 27]. Also interestingly, some studies demonstrated the relation between generic mutations and onsets of familial forms of ADs [28, 29].

Sharing an association with gene mutation with ADs, the second row presents main pathologic aspects of Parkinson's disease (PD).

In fact, alongside with the formation of Lewi bodies, i.e., cytoplasmic aggregates formed by α -synuclein, in the midbrain and other monoaminergic neurons in the brain stem and consequent effects, recessive PDs can also be caused by mutations in the genes encoding parkin, DJ-1, or PINK1 [26, 30–33].

The fatal Huntington's disease (HD) is caused by the elongation and repetition on the short arm of chromosome 4p16.3 in the Huntingtin gene (HD or IT15 gene) of the trinucleotide CAG, a building stone of DNA, resulting in Huntingtin proteins with elongated polyglutamine stretches and a symptomology of irregular/involuntary muscular movements, cognitive decline, psychiatric disturbances, and dementia [34, 35]. HD is considered to be spread when polyglutamine stretches exceed the critical length of 37 glutamine units [36].

Another progressive fatal ND is amyotrophic lateral sclerosis (ALS), although unlike the previous cases, it does not present cognitive impairment on its symptoms [37]. The most common cause of ALS is a mutation of the gene encoding the superoxide dismutase (SOD), an enzyme that reduces superoxides formed during cellular metabolism [38]. To conclude the common NDs discussed in this section, we also mention prion disease (PD), characterized by conformationally modified prion protein damaging the neuronal cells [39]. Further, mutations in prion gene can cause a specific group of ND diseases, such as kuru, Creutzfeldt-Jakob disease (CJD), Gerstmann-Straussler-Scheinker (GSS), and fatal familial disorder, which unlike other NDs can be transmitted through infected brain tissue or surgical tools [40].

The abovementioned NDs and all the remaining NDs that are unmentioned in Table 1 have now no absolute cure. The therapies are solely based on symptomatic relief, aiming at the improvement on patient life quality, albeit how small and/or enduring. Among the mentioned NDs, Parkinson's disease seems to have the most effective treatment options, which have some measure of significance on improving the patient's motor symptoms, namely, by promoting dopaminergic functions in the substantia nigra [41].

The prominent status of PD among available therapies in NDs scenario does not exclude the existence of other interesting and promising approved therapies, namely, for Alzheimer's disease. In fact, the Food and Drug Administration (FDA) approved for AD acetylcholinesterase (AChE) inhibitors (donepezil, galantamine, and rivastigmine) and NMDA receptor antagonist memantine, which is believed to protect neurons from excitotoxicity [42, 43]. It should be noted that though the β -amyloid and tau proteins are the prime targets for disease-modifying therapies in AD, at least 100 different classes of therapeutic agents are now in different phases of clinical trials for future AD therapy [43, 44].

We briefly mention here that the most successful therapy for ALS impacting the disease's progression is the neuroprotective agent riluzole, which extends the life span of patients by a few months, while

the tetrabenazine is the only FDA-approved drug for the treatment of HD-related chorea, though drugs such as amantadine (dopamine agonist) and riluzole are also used in different cases [45, 46].

3 Multitarget Drugs for NDs

Though the typical drug design “one molecule-one target” strategy, i.e., the development of small molecules capable of modulating the function of a disease-responsible single biological target, is accountable for several successful drug discoveries, it may not deliver satisfactory outcomes in NDs, and this is due to the multifactorial pathological nature of these diseases [12, 47]. As an example, the drug impact on a particular biological target can be nullified due to emergence of alternative biological pathways [12, 48]. This situation may be resolved by polypharmacology-based research, and two of the most frequently followed polypharmacological approaches are (a) multiple-medication therapy (MMT), where different formulations containing more than one drugs are prescribed in the same time, and (b) multiple-drug medication (MDM), in which the same formulation contains multiple drugs acting in different mechanisms.

A third proposed polypharmacological approach is based on the concept of multitarget-directed ligands (MTDL) [49, 50]. In this strategy, a single drug is aimed at interacting with multiple biological targets, associated in a specific disease, in order to promote synergistic therapeutic response. The MTDL has several advantages over MMT and MDM, such as ease of administration and formulation development as well as higher patient compliance [51–53]. Though there are many different terminologies in literature, the term multitarget-directed ligand (MTDL) is used in this chapter and henceforth when in context of NDs chemometric methods. Although the design of MTDL is comparatively simpler for the structurally homologous proteins, the development of MTDL is undoubtedly more challenging than single-target drugs due to increased risk of toxicity and poor pharmacokinetic profiles [50]. Furthermore, target proteins of NDs are polygenic and diverse in nature. It also poses a significant challenge to develop MTDLs for the treatment of NDs [54, 55]. Nevertheless, there are promising cases, such as the AD treatments with a dual orexin receptor antagonist suvorexant (MK-4305) or the dual cholinesterase-monoamine oxidase B (ChE-MAO_B) inhibitor (anti-AD agent ladostigil), and also in the PD treatment, the dual adenosine A₂A/MAO_B inhibitor has possibilities in AD treatment, while another promising MTDL memquin shows free-radical scavenging action, inhibition of AChE activity, as well as inhibition of A β aggregation [44, 55–57].

4 Chemometric Methods in Polypharmacology

As previously stated (*vide* Subheading 1) and in accordance to the International Chemometric Society (ICS), the MTDL methods led to the popular chemoinformatic approaches for large-scale efficient computational screening within the scope of recognition of new ligand-protein interactions [21, 58, 59]. Focusing further on the analysis of these methods, we can divide them broadly into two areas: (a) ligand-based drug design (LBDD) and (b) structure-based drug design (SBDD).

LBDD is based on the chemical, structural, and functional knowledge of the ligands interacting with the target of interest. Different quantitative structure-activity relationship (QSAR)-based approaches as well as ligand-based pharmacophore mapping are examples of LBDD approaches. On the other hand, SBDD relies on the three-dimensional structures of the biomolecular targets obtained through X-ray crystallography or nuclear magnetic resonance (NMR) or homology models. As such, structure-based pharmacophore mapping, molecular docking, and molecular dynamics (MD) simulations are examples of some frequently used SBDD techniques. Combined LBDD and SBDD approaches may improve the overall reliability of computational predictions [17, 60].

Several chemometric techniques in the following subheadings are presented (i.e., Subheadings 4.1 Conventional 2D-QSAR and 3D-QSAR, 4.2 Multitarget QSAR modeling, 4.3 Similarity searches, 4.4 Pharmacophore mapping approaches, 4.5 Structure-based analyses, and 4.6 Chemical databases and web-based biological target searching tools).

Virtual screening (VS) is a computational strategy that identifies the chemical molecules expected to interact with a specific biological target. The chemicals are generally retrieved from large diverse chemical libraries, i.e., databases. A good VS process is expected to maintain a balance between high screening speed and high accuracy, especially since its *in silico* competitors, the structure-based methods such as molecular docking, are computationally more time-consuming [61, 62]. Moreover, in order to reduce the chemical search space, compounds with poor ADMET profiles or less prospective therapeutically properties may be discarded from the screening databases, allowing subsequent application of Lipinski's rule of drug-likeness and computational ADMET predictivity methods for screening datasets [62, 63].

4.1 Conventional 2D-QSAR and 3D-QSAR

Quantitative structure-activity relationship (QSAR) calculations are frequently employed to model the biological and physical properties of the compounds, its purpose being to obtain a well-defined mathematical relationship between the biological activity and molecular descriptors, i.e., the descriptive parameters of molecular chemical or physicochemical properties [19, 64, 65]. As MTDL

methods are based on compound-specific properties, either chemical, structural, physicochemical, topological, or biological, and within the context of general *in silico* methodologies, it is important to define formally a “descriptor.”

In chemometrics, “the molecular descriptor is the final result of a logic and mathematical procedure which transforms chemical information encoded within a symbolic representation of a molecule into a useful number or the result of some standardized experiment,” i.e., it is the mathematical expression of a given compound attribute formulated in order to be used in an *in silico* model [66].

As examples of descriptors, we can mention the molar volume, which is a steric descriptor; the Hammett constant, within the category of electronic descriptors; or the energy of highest occupied molecular orbital, representing a quantum chemical descriptor. In Table 2 are displayed some of the currently available commercial and noncommercial software solutions for the calculations of molecular descriptors.

Following this stage, regression or classification mathematical models are then derived using different statistical machine learning tools, such as multiple linear regression (MLR), partial least squares (PLS), linear discriminant analyses (LDA), support vector machines (SVM), artificial neural networks (ANN), deep neural networks, genetic algorithms, random forest (RF), Bayesian modeling, or others [67, 68].

Table 2
List of some common softwares for descriptor calculation in 2D-QSAR analyses

Software name	Descriptor numbers	Home page
ADAPT	260	http://research.chem.psu.edu/pcjgroup/adapt.html
ADMET Predictor	297	http://www.simulations-plus.com/
ADRIANA	1244	https://www.mn-am.com/software/adrianacode/index.html
CODESSA	1500	http://www.codessa-pro.com/
DRAGON	5270	http://www.taletе.mi.it/products/dragon_description.htm
ISIDA FRAGMENTOR	NA	http://infochim.u-strasbg.fr/spip.php?rubrique49
MARVIN Beans	500	http://www.chemaxon.com/
MOE	300	http://www.chemcomp.com/
MOLCONN-Z	40	http://www.edusoft-lc.com/molconn/
MOLGEN-QSPR	700	http://www.molgen.de/?src=documents/molgenqspr.html
PaDEL-Descriptor	850	http://padel.nus.edu.sg/software/padeldescriptor/
PowerMV	1000	https://www.niss.org/research/software/powermv
Sarchitect	1050	http://strandls.com/sarchitect/index.html

It should be noted that the reliability of the QSAR models depends on the dataset size and diversity, robust feature selection technique, and validation methods. Therefore, after construction of the QSAR models, suitable internal and external validation tools should be applied, respectively, on the training/modeling set and on the test, validation, or prediction set to justify statistical robustness of the model. Possible internal validation includes the leave-one-out and leave-many-out cross-validation techniques. Moreover, all QSAR model must have a well-defined applicability domain, i.e., a chemical space under which the prediction is expected to be reliable [19].

3D-QSAR correlates the biological activities of a set of bioactive target molecules with the non-covalent interactions, and the results are represented by contour maps displaying favorable and unfavorable non-covalent interactions surrounding the molecules. Examples of alignment-dependent 3D-QSAR techniques, where proper alignment of 3D molecular structures and appropriate template structure is mandatory, include comparative molecular field analysis (CoMFA), comparative molecular similarity indices analysis (CoMSIA), self-organizing molecular field analysis (SOMFA), *k*-nearest neighbor-molecular field analysis (kNN-MFA), topomer CoMFA, and Open3DQSAR [69–74]. Grid-independent descriptor (GRIND) is an example of an alignment-independent QSAR technique [75–77].

As far as design of MTDL is concerned, multiple 2D-/3D-QSAR models may be developed separately based on various datasets with different targets of interest, and afterward they may be recruited separately to screen a whole plethora of chemical libraries. Compounds with higher predicted activity values in multiple models may be considered as multipotent virtual hits, and in fact these may be refined in a posterior phase with other chemometric modeling tools, in order to validate the results [78].

4.2 Multitarget QSAR Modeling

Another promising QSAR method in multitarget drug design is multitarget QSAR (mt-QSAR) [79, 80]. While conventional QSAR modeling depends on only one type of assay conditions and biological target, mt-QSAR allows incorporation of the assayed compounds against different biological targets in a single dataset. Further, this incorporation is possible for several assay conditions, such as information, target mapping, level of curation, or reliability of the assay [79]. The mt-QSAR approaches therefore expand both the chemical-biological spaces and the applicability domain of the chemometric models, and a schematic presentation of mt-QSAR approach is depicted in Fig. 1 [80].

In particular, input descriptors are modified based on the characteristics of the individual system. One example of such modification is the moving average approach (MAA) [81]. The following equation illustrates the MAA applied for the creation of new

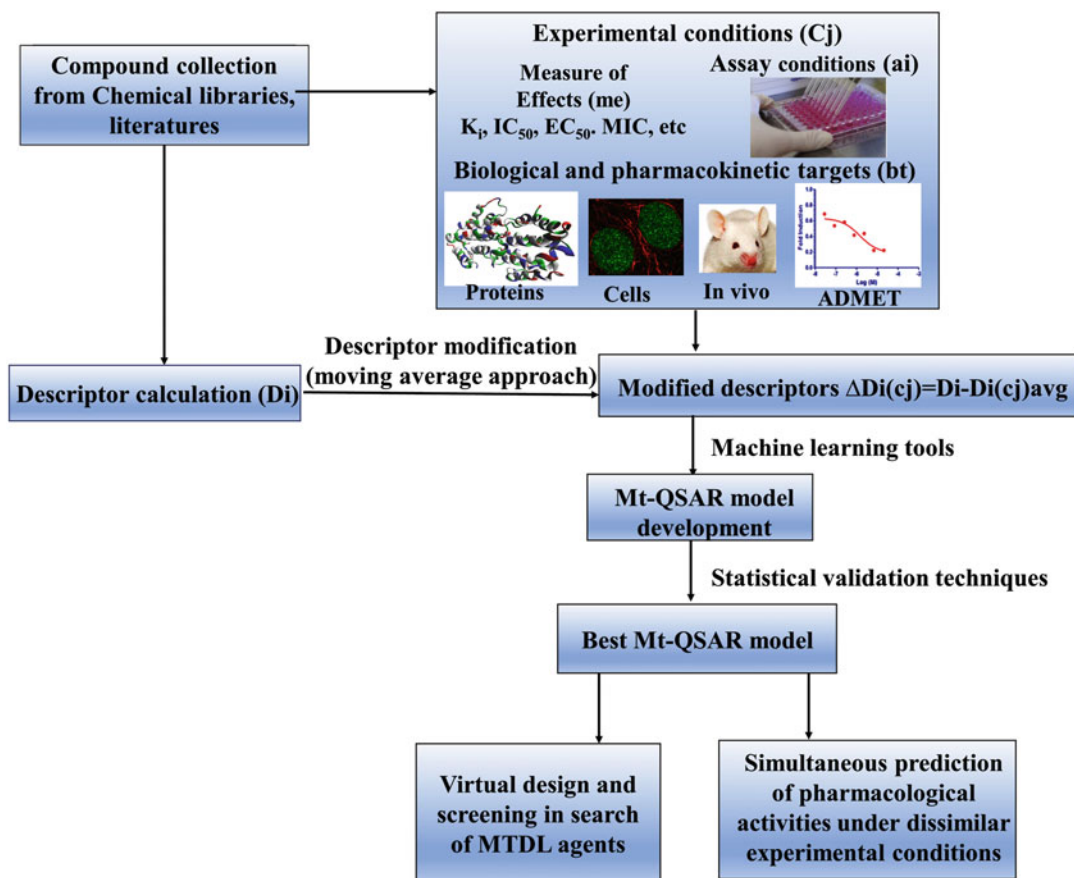


Fig. 1 Schematic presentation of the mt-QSAR approach

molecular descriptors $[\Delta D_i(c_j)]$ from originally calculated descriptors (D_i).

$$\Delta D_i(c_j) = D_i - D_i(c_j)_{\text{avg}}$$

The essential element c_j defines various experimental conditions (me , bt , ai , etc.), and $D_i(c_j)_{\text{avg}}$ characterizes average descriptor values calculated for the active compounds in each set tested under the same experimental condition (c_j) (*vide* Fig. 1).

Furthermore, not only different machine learning tools such as LDA, ANN, or SVM are applied to develop the classification models, but a recent proposal for a regression-based mt-QSAR provides a context for the development of regression models alongside a novel modular neural network technique [80, 82]. The capability of mt-QSAR for simultaneous prediction multiple biological systems, assay parameters, and conditions makes it invaluable as a highly efficient tool for multitarget drug design [81–84].

4.3 Similarity Searches

Based on the assumption that structurally similar compounds frequently have similar biological activity and bind with same group of proteins, chemical similarity search approaches are effective in identifying alternate targets for existing compounds, and thus they may assist in both polypharmacology and drug repositioning [85, 86].

Fingerprint-based similarity search is widely used for the target predictions, and it generally represents the compounds by 2D fingerprints, whereas the similarity is calculated by similarity indices such as Tanimoto coefficients (Tc) [87].

Similarity ensemble approach (SEA) is another promising approach which quantitatively relates receptors to each other on the basis of the chemical similarity among their ligands [88]. Apart from structural fingerprints, molecular shape is another important parameter for measuring similarity among ligands. Finally, shape-based techniques may identify similarity between structurally unrelated compounds, using 3D descriptors presenting shape, chirality, and charges to compare chemical compounds [85, 89].

4.4 Pharmacophore Mapping Approaches

Pharmacophores are defined as minimum 3D structural features, or attributes, required for eliciting biological response, and pharmacophore models are frequently used for the screening of chemical databases to find potent virtual hits for biological targets. The process of developing pharmacophores is called pharmacophore mapping, and examples of important pharmacophores include hydrogen bond acceptor/donor or aromatic rings [62, 90].

Ligand-based pharmacophores are generated on a set of diverse compounds active against the same biological target, with the quality of ligand-based pharmacophores depending on several parameters, such as the conformation generation techniques, molecular alignments, or ligand flexibility, among others [90, 91].

Though structure-based pharmacophores are derived from receptor-bound conformations of the ligands, apoprotein structures are also used for pharmacophore mapping, with active protein amino acid residues as basis for complementary pharmacophores development.

Pharmacophore mapping may be efficiently used for the development of multipotent targets if pharmacophores were developed on more than one target. Compounds achieving good fit values in different pharmacophores for different targets (ligand or structure-based) may be considered as primary virtual hits for MTDL [92]. Examples of some frequently used well-known pharmacophore mapping tools are provided in Table 3. Also, PharmMapper is a web-based platform that uses inverse structure-based pharmacophore matching techniques of the query molecules for biomacromolecular target prediction in order to understand polypharmacological profiles of the chemical structures [93].

Table 3
List of some common ligand- and structure-based pharmacophore mapping and their characteristics

Software	Type ^a	Characteristics
GALAHAD [94]	L	Atom-based alignment and quantitative modeling analyses
ALADDIN [95]	L	Steric, geometric, substructure-based model, open for academic use
DANTE [96]	L	Noncommercial, SAR-based modeling technique
DISCO [97]	L	Concord and Confort based conformation search, clique detection algorithm-based alignment
E-Pharmacophore [98]	S	Energy-based feature generation in ligand on protein complex
GASP [99]	L	Atom-based alignment and genetic algorithm-based analyses
HipHop [100]	L	Feature-based alignment and quantitative modeling analyses
HypoGen [101]	L	Feature-based alignment and QSAR
LigandScout [102]	S	Energy-based feature generation in ligand on protein complex
MOE [103]	L	Property-based alignment, manual pharmacophore mapping
MPHIL [104]	L	Atom-based alignment, GA-based analyses with clique detection
PHASE [105]	L	ConfGen conformational analyses, feature-based alignment, QSAR

^aL ligand-based, S structure-based

4.5 Structure-Based Analyses

Frequently used for virtual screening of large chemical databases, molecular docking is the most popular and frequently used *in silico* structure-based tool, helping hypothesizing the bioactive conformation of the ligands through intermolecular interactions between receptor and ligand. The performance of each docking program depends as much on the protein and ligand preparation, information about the binding site, and conformation generation techniques as on docking algorithms [95, 106, 107]. Different docking tools may differ on the basis of conformation generation techniques, search algorithms, and scoring functions. Some frequently used docking tools are provided in Table 4.

Information about the receptor binding sites is vital for structure-based analysis. In the absence of binding site information, as in the case of newly reported proteins, blind docking may be performed, with the grid extended as to cover the whole protein [96]. Additionally, *in silico* binding site identification tools, such as SiteMap, FTMAP, fpocket, or others, may help in locating possible enzyme active sites, doing so through identifying the largest concave protein pocket possessing hydrophobic, hydrogen bond acceptor and hydrogen bond donor sites [117–120].

Since proteins adopt multiple conformational states, only one docking conformation fails to account for all possible binding states. Ensemble docking tends to solve this issue by allowing the docking of the ligands against multiple rigid receptor conformations. Therefore,

Table 4
List of some common molecular docking tools and their characteristics

Docking program	Characteristics
AutoDock [108]	Genetic algorithm (GA)-based search, free energy-based scoring
Dock [109]	Anchor-and-Grow-based docking program, GA-/MC-based search
FlexX [110]	Incremental reconstruction search, modified Bohm's scoring
FRED [111]	Exhaustive search, Gaussian scoring function
GLIDE [112]	Hybrid exhaustive search-based docking program, GlideScore scoring
GOLD [113]	GA-based search and Goldscore, ChemScore, ASP, and ChemPLP scoring
LigandFit [114]	Monte Carlo (MC)-based search, LigScore, PLP, PMF scoring
Surflex [115]	Incremental reconstruction search, idealized active site ligand, Bohm's scoring
VLifeDock [116]	GA-based search, PLP score, Xcscore

conformational ensemble leads to improved virtual screening outcomes [85, 121].

It is noteworthy that MTDLs often have low-binding affinities against multiple proteins due to weak interactions, which renders structure-based design more challenging [122]. Thus, rather than simply depending on higher docking scores, or fitting values, the explanations of structure-based tools are dependent on higher docking scores and pharmacophore fitting values, underlining how important is examining dynamic stabilities of the ligand-protein complexes. Further, long-time molecular dynamics (MD) simulations allow conformational flexibilities of both proteins and ligands. Since MD simulation provides both high- and low-populated conformational states, the overall interpretability of the results is increased [123].

Inverse docking is a comparatively novel technique where a single ligand is docked against multiple biological targets. On the basis of ligand-receptor interactions, the targets are ranked. INVDOCK was the first reported web-based inverse docking program that selects potential protein targets based on competitive binding analysis and ligand-protein interaction energy [124]. Two other web-based inverse docking tools are idTarget and TarFisDock [125, 126].

In addition, binding site similarity-based search, grounded on the assumption that structurally similar proteins interact with structurally similar compounds, is frequently used for the target prediction and may contribute in the development of multitarget drugs [127]. IsoMIF Finder and BioGPS27 are two such tools that implement the fingerprints for ligands and proteins (FLAP) algorithm, which generates pharmacophoric features from GRID molecular interaction fields (GRID-MIFs), and these pharmacophoric features are then utilized for searching of find ligand-ligand, ligand-receptor, and receptor-receptor similarities [127–129].

Table 5
Publicly accessible web-based target prediction tools [127, 129]

Program	Similarity method	Website
ChemMapper	SHAFTS molecular shape fit	http://lilab.ecust.edu.cn/chemmapper/
ChemProt	Molecular fingerprints	http://potentia.cbs.dtu.dk/ChemProt
DRAR-CPI	Docking	https://cpi.bio-x.cn/drar/
HitPick	Molecular fingerprints	http://mips.helmholtz-muenchen.de/proj/hitpick
Polypharmacology browser	Molecular fingerprints	www.gdb.unibe.ch
Similarity ensemble approach	Molecular fingerprints	http://sea.bkslab.org
SuperPred	Molecular fingerprints	http://prediction.charite.de/index.php
SwissTargetPrediction	Fingerprints/electro-shape descriptors	www.swisstargetprediction.ch
TarFisDock	Docking	http://www.dddc.ac.cn/tarfisdock/
TargetHunter	Molecular fingerprints	www.cbiligand.org/TargetHunter

4.6 Chemical Databases and Web-Based Biological Target Searching Tools

Chemometrics and its feasibility and success for polypharmacology-based research depend considerably on the availability of chemical and biological data. For this reason, computational chemists rely on different web-based databases for fast and focused collections of data and information. Some examples of such databases include chemogenomic, drug target, toxicogenomic, disease-specific target, pathway analysis, and metabolomic databases, as well as protein data and information repositories [130–143].

Similar to online chemical databases, different web-based target fishing tools may provide excellent opportunities for finding new biological targets for chemical compounds. A list of a few important and publicly accessible web-based target prediction tools is provided in Table 5.

5 Chemometric Modeling for Design of Anti-ND MTDLs

In the former section, polypharmacological research via chemometric methods were reviewed, and the discussion comprised a wide scope of areas, techniques, and subjects, from conventional 2D/3D-QSAR to pharmacophore mapping approaches and even web-based biological target searching tools. In this section, the focus will be an in-depth analysis on different in silico chemometric techniques within the specific context of anti-ND MTDLs development.

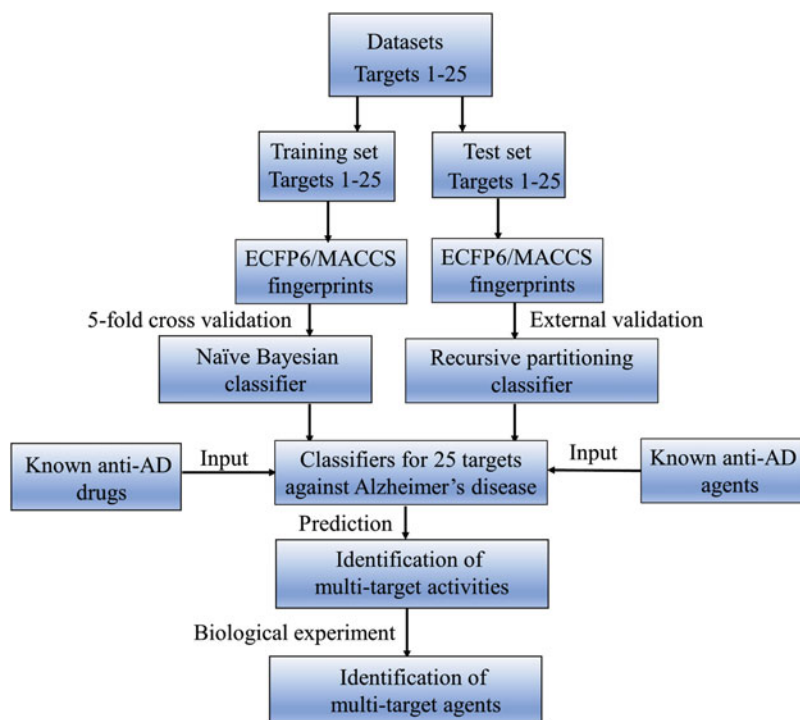


Fig. 2 The multi-QSAR workflow of the procedure adopted by Fang et al. [78] to identify multitarget compounds

5.1 Design of Anti-ND MTDLs: 2D-QSAR and mt-QSAR Approaches

Though in the last few years were presented several conventional 2D-QSAR models which considered various biological targets of AD, PD HD, ALS, as well as neuroprotective agents, none of them were intended for modeling new multipotent compounds [144–164]. In 2015, Fang et al. attempted building 100 binary classifier models using two machine learning approaches, Naive Bayesian (NB) and recursive partitioning (RP), and the workflow of the procedure adopted by them is outlined in Fig. 2 [78].

In a brief summary, multiple QSAR models were developed with a large number of compounds collected from Binding Database with biological activity against 25 AD-related targets, and each molecule was represented by two types of fingerprint descriptors: (1) ECFP6, a circular topological extended-connectivity fingerprint designed for molecular characterization, similarity search, and structure-activity modeling, and (2) MACCS, whose fingerprints are 2D substructure descriptors encoding atoms, atom types, rings, and bond information about the molecule [135, 165, 166].

The best predictive models were used to evaluate six approved anti-AD drugs and 19 known anti-AD agents in an attempt to identify multitarget activity compounds. Subsequent in vitro experimental validation confirmed predictivity of the developed models, exemplified by the compliance between the in silico prediction and

the properties of one compound, which exhibited submicromolar inhibitory and antagonistic potential toward AChE and histamine 3 receptor (H_3R), respectively [78]. In conclusion, this investigation justified how reliable multiple QSAR models are for identifying multipotent lead molecules.

Another very interesting work was reported by Besnard and colleagues, presenting the development of multiple category Laplacian-modified naïve Bayesian models with ECFP6 fingerprints for 784 biological targets [167]. The investigators started with the chemical structure of the AChE inhibitor donepezil, aiming at first to improve activity toward dopaminergic D_2 receptor. Secondly, they intended to ensure good blood-brain barrier (BBB) permeability. After applying classification models, some novel analogues were synthesized, and all of these showed substantial D_2 receptor affinities in experimental assays. The most potent compound possessed BBB penetration, although it also showed affinity toward anti-target α_1 -adrenergic receptors. Therefore, further investigation is ensued, aiming to increase selectivity toward D_2 receptor using classification models. This particular work is an important example to demonstrate how *in silico* modeling may be used to tailor multipotent activity, toxicity, and ADME properties of the compounds [167].

Speck-Planche et al. reported mt-QSAR modeling with a dataset containing 483 compounds with 1244 cases of inhibitory activity against at least one of five proteins related to AD, i.e., amyloid β -A4 protein (ABPP), glycogen synthase kinase-3 α (GSK3 α), GSK3 β , MAO_B, and presenilin-1 (PSEN1) [168]. The activities of the compounds were classified into two categorical outputs depending on specific cutoff IC_{50} values for each target. Spectral moment topological descriptors of different orders, μ_k , were calculated for each molecule, as well as the average value of each spectral moment, $av\mu_k$, for all active against the same classification variable compounds. Further, the values obtained from deviations of μ_k and $av\mu_k$, $d\mu_k$, were also determined, introducing variables characteristics of each system. Finally, linear discriminant analysis (LDA) was used to develop a 14 descriptors mt-QSAR model with accuracies of approximately 95% in both the training and validation sets. The novel mt-QSAR model was used for the automatic extraction of fragments responsible for the inhibitory activity against these five AD-related proteins [168].

Using a similar mt-QSAR approach, another modeling study was reported by Luan et al. on neuroprotective agents, considering 2217 compounds with 4915 bioactivity data collected from ChEMBL database [130, 169, 170]. This highly complex data contained 338 types of biological assays depending on the assay techniques, measurement of effects, and biological targets (protein and cell lines). The elaboration of the LDA mt-QSAR model used the moving average technique along with the help of spectral moment descriptors. The

best model, developed with five descriptors, achieved an accuracy of 89.5% and 89.1% for the training and the validation sets, respectively. To further understand the practical utility of the work, the neuroprotective effects of novel synthesized 1,3-rasagiline derivatives were evaluated by three different assays, with two compounds showing considerable neuroprotective activities in multi-assays. Simultaneously, the mt-QSAR model was also used to predict the probable responses of the new compounds in 559 unique tests, which served to confirm the experimental findings.

A similar strategy was adopted in a neuroprotection-related investigation, where the development of a mt-QSAR model departed from a dataset containing 8309 samples with bioactivities toward more than 20 target proteins [170]. The best model showed more than 80% overall accuracy. Eight new synthesized rasagiline derivatives were biologically tested to confirm their neuroprotective actions. Interestingly, the developed mt-QSAR model predicted that brain inducible nitric oxide synthase (iNOS) enzyme is the most probable target for these compounds.

To conclude this subheading, it is worth mentioning another mt-QSAR modeling work on neuroprotective agents more recently reported by Romero Duran et al. [171]. In this work, multiple outcomes for more than 30 measures of neurotoxic/neuroprotective effects in more than 400 different assay conditions were calculated, and some 1,2-rasagiline derivatives were synthesized, including the most active compound of those evidencing a higher predictive potency against the 5-hydroxytryptamine 3 receptor (5HT₃R) [171].

5.2 Design of Anti-ND MTDLs: 3D-QSAR

Within 3D-QSAR research toward novel anti-ND MTDLs development, an interesting case study was on the derivatives of donepezil, PF9601N, ASS234, and donepezil-pyridyl, which were all investigated in the ND field of research (Fig. 3 presents the structures of the compounds) [172, 173]. Multipotent ASS234, designed as a hybrid of AChE inhibitor donepezil and MAO_B inhibitor PF9601N, was found to be a promising multitarget agent with inhibitory potentials toward MAO_A, MAO_B, AChE, and butyrylcholinesterase (BuChE) enzymes. Additionally, it was demonstrated to have anti-apoptotic and antioxidant properties as well as satisfactory BBB permeability.

The bioactivity data of different analogues and derivatives of donepezil, PF9601N, clorgyline, and ASS234 against MAO_A, MAO_B, AChE, and BuChE enzymes were separately used by Bautista-Aguilera and co-workers to develop 3D-QSAR models using the Pentacle program [172, 174]. These generated 3D-QSAR models were shown to have a moderate to high internal and external predictivity, though they also revealed some crucial information regarding structural requirements of these derivatives [172]. For example, the essential role of the propargylamine moiety for MAO_A and MAO_B inhibition was highlighted.

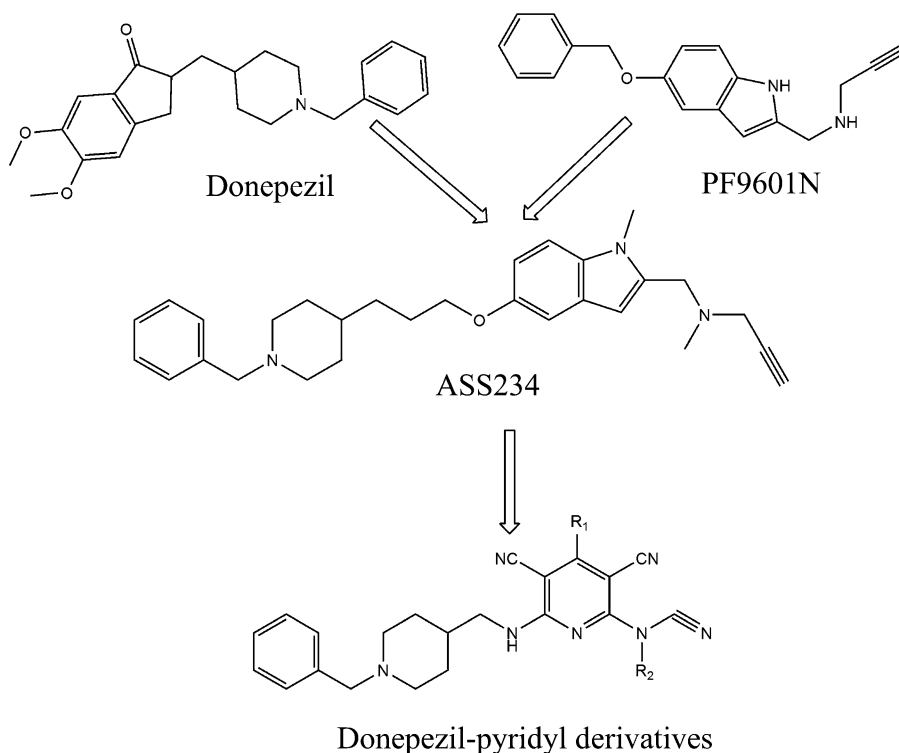


Fig. 3 Structures of donepezil, PF9601N, ASS234, and donepezil-pyridyl derivatives investigated by 3D-QSAR analysis

Similarly, bulky substituents at *p/m*-positions of the benzyl moiety evidenced enhancement of the MAO_A inhibitory potential at the expense of MAO_B inhibitory activity. The 3D-QSAR model also predicted that an electron withdrawing substituent in the benzyl and piperidinyl moieties might increase AChE inhibiting activity, whereas the presence of bulky substituent at *p*-position of the benzyl moiety could lower BuChE inhibiting activity. The oxygen bridge formed unfavorable interactions with the benzyl moiety in both MAO subtypes, whereas this bridge has a positive influence on ChE subtypes. Therefore, substitution of the oxygen in the bridge with other atoms indeed could increase both MAO_A and MAO_B inhibiting activities [172].

Departing from the structure of ASS234, the same researchers then designed a new set of donepezil-pyridyl hybrid derivatives (*vide* Fig. 3), which were synthesized and biologically evaluated [173]. They followed the study with 3D-QSAR analysis, performed with a larger dataset ($N = 37$), containing these newly synthesized donepezil-pyridyl derivatives as well as different analogues and derivatives of donepezil, PF9601N, clorgyline, and ASS234, thus providing additional information about SAR of these compounds. It is worth mentioning here that overall statistical predictabilities of

these models were increased, probably due to the expansion of the dataset size.

The inferences of the 3D-QSAR models are (a) substitution of the terminal hydrogen of propargylamine moiety with bulky groups and replacement of the N-Me methyl with hydrogen of the indole moiety may selectively enhance MAO_A inhibiting activity, (b) bulky groups in *m/p*-positions of the benzyl moiety may lower MAO_A inhibiting activity, (c) electron donating substituents in the propargylamine or indole moiety and bulky substituents on N-atom of the indole moiety may selectively increase MAO_B inhibition activity, (d) electron withdrawing bulky groups at *m/p*-positions of the benzyl moiety may selectively increase MAO_B inhibiting activity, (e) extension of the N-alkyl bridge could selectively strongly increase AChE inhibitory potential, and (f) in donepezil-pyridyl hybrids, phenyl substituent in the pyridine ring may selectively increase AChE inhibiting activity, though substitution at *m*-position of this phenyl ring decreases AChE inhibitory activity [173]. These investigations substantiate that 3D-QSAR modeling may be successfully implemented for the design of multipotent inhibitors [172, 173].

More recently, Nikolic et al. investigated multitarget ligands interacting with some important biological targets for the treatment of neurological conditions, including AD and PD [175, 176]. In one of these studies, 35 amino-quinoline derivatives with inhibitory data against histamine 3 receptor (H₃R) and histamine N-methyltransferase (HMT) were considered for the alignment-independent 3D-QSAR analysis, and the generated QSAR models presented high and moderate statistical predictivity for H₃R and HMT datasets [175]. 3D-QSAR analysis revealed that the presence of specific lipophilic or steric components is important for H₃R inhibition. Further, hydrogen bond acceptor features were found in the certain regions of H₃R pharmacophore, notwithstanding that the same regions showed hydrogen bond donor features for HMT pharmacophore. Moreover, longer optimal distance between hydrogen bond donor and steric hot spots in the H₃R pharmacophore distinguished it from the HMT pharmacophore [175].

In another work, Nikolic et al. resorted to alignment-independent 3D-QSAR models for tackling the activity against three different groups of multipotent inhibitors [176]. The first set of ligands consisted of 17 compounds with MAO_A and MAO_B inhibitory activities, while the second group included 67 derivatives with polypharmacology against MAO_A, MAO_B, AChE, and BuChE. The final and third group contained multipotent histamine H₃R/HMT antagonists. Using the Parzen-Rosenblatt kernel density estimation method, probabilistic models were developed for both primary targets and off-targets using data collected from the ChEMBL and DrugBank databases [130, 136]. The cheminformatic-based target identifications complied with four 3D-QSAR models developed on H₃R, D₁R, D₂R, and 5HT_{2A}R. Also, *in vitro* 5HT_{1A}

and 5HT_{2A} receptor binding assays corroborated these theoretical predictions. As such, several other multitarget ligands were chosen for further investigation of their possible additional beneficial pharmacological activities [176].

5.3 Design of Anti-ND MTDLs: Pharmacophore Mapping

With regard to pharmacophore mapping for the design of novel anti-ND MTDLs, there are some interesting and worth noting published works. Huang et al. developed a ligand-based pharmacophore with BACE-1 inhibitors. This pharmacophore was used for screening a chemical library containing N-phenyl-1-arylamide and N-phenylbenzene sulfonamide derivatives and allowed them to identify a lead molecule with moderate inhibitory potential against BACE-1 enzyme ($IC_{50} = 18.33 \mu\text{M}$) (*vide* Fig. 4) [177].

In another investigation, this lead molecule was hypothetically combined with metal chelator LR-20 to design some hybrid 1,3-diphenylurea derivatives, displayed in Fig. 4 [178]. Screening of these hypothetical derivatives with the abovementioned BACE-1 pharmacophore model identified some potential virtual hits. Subsequent syntheses and biological assays of these hits yielded multipotent BACE-1 inhibitor/metal chelator agents. Though the most potent derivative showed slightly reduced BACE-1 inhibition ($IC_{50} = 27.85 \mu\text{M}$), its copper and iron chelating properties made it a promising multitarget lead molecule.

This pharmacophore model was used in other research to screen a chemical database of quinoxaline derivatives, and some top hits were subjected to molecular docking into AChE protein structures to identify some virtual hits with possible multipotent activities [179]. Syntheses of these hits and their biological analyses

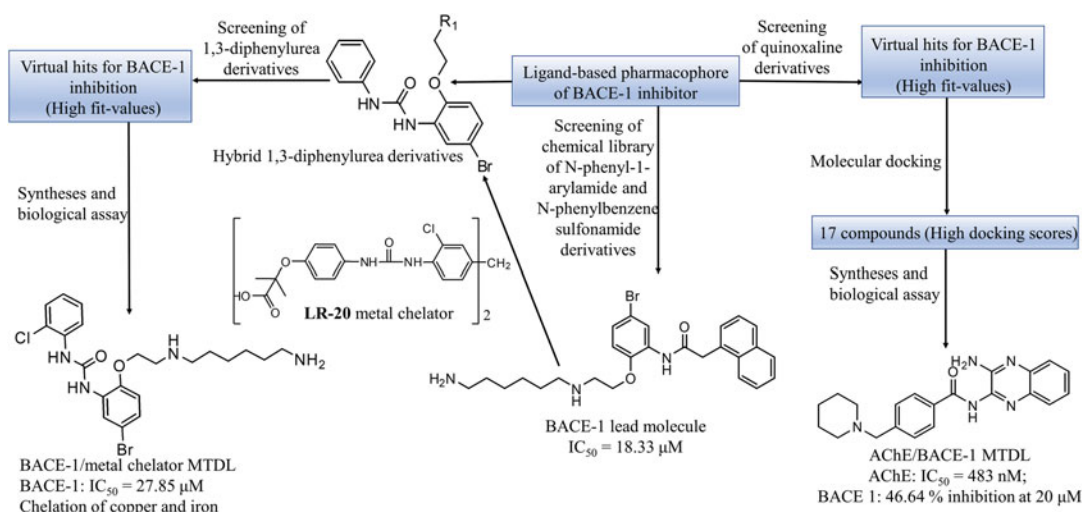


Fig. 4 Design of MTDLs for AChE/BACE-1 inhibition and BACE-1 inhibitor/metal chelator using ligand-based pharmacophore modeling

confirmed the inhibitory potency of some of the virtual hits against both AChE ($IC_{50} = 483$ nM) and BACE-1 (46.64% inhibition at 20 μ M) (*vide* Fig. 4).

In 2017, Xie et al. developed a highly predictive ligand-based pharmacophore model with 41 structurally diverse carbamate derivatives, divided into 25 training and 16 test set molecules. This five-feature pharmacophore model was subsequently utilized to identify (–)-meptazinol phenylcarbamate as a potent nanomolar AChE inhibitor with anti-amyloidogenic properties [180].

In a recently reported multi-chemometric analysis, Bhayye et al. developed statistically significant pharmacophore and hologram QSAR (HQSAR) models with some previously reported dual A_{2A} antagonists/MAO_B inhibitors as anti-PD agents [181]. Pharmacophore mapping revealed the importance of heterocyclic rings, hydrogen bond-forming polar groups, hydrophobic core residues, and substitution of the hydrophobic cores. In addition, use of a fragment-based HQSAR technique allowed to point out important favorable and unfavorable fragments for higher biological activity toward A_{2A} antagonism and MAO_B inhibition [182]. Therefore, fragments that lead to positive contributions for both bioactivities may help in increasing multitarget properties [181].

5.4 Design of Anti-ND MTDLs: Structure-Based Analyses

Molecular docking has been proved as one of the most essential approaches for the design of the dual-binding AChE inhibitors. One interesting aspect of AChE enzyme is that it may itself serve as an excellent target for multifunctional drug development due to “cholinergic” and “non-cholinergic” roles that are regulated from two distinct binding sites, catalytic site (CAS) and peripheral anionic site (PAS) of the enzyme, and, in fact, interactions of A β at the PAS of AChE lead to the conformational changes of this peptide that facilitates its aggregation [183, 184]. Therefore, dual-binding inhibitors which simultaneously interact with both CAS and PAS sites of AChE are likely to be more effective against AD than compounds interacting with either of these binding sites [184, 185].

Interestingly, the binding modes of the dual-binding AChE inhibitors at CAS and PAS sites of the enzyme are frequently analyzed and monitored to understand the mechanism of action as well as to modify the lead molecules [178, 186, 187]. As an example, Xie et al. developed a tacrine-coumarin hybrid molecule, labeled **1** in Fig. 5, considering both the MAO inhibitory potential of coumarin and the AChE inhibitory potential of tacrine [188].

Xie et al. also combined the structures of compound **1** and donepezil to obtain another hybrid molecule (*vide* **2**, Fig. 5), and the performed molecular docking analysis indicated that compound **1** binds only with the CAS site, whereas compound **2** interacts simultaneously with CAS and PAS binding sites of AChE [186]. Therefore, though compound **2** is less potent than compound **1**, the former may have higher possibility to inhibit A β aggregation.

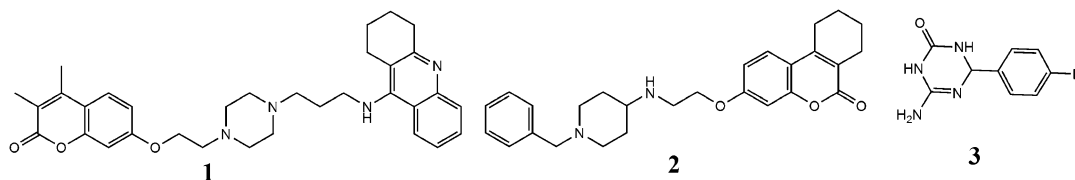


Fig. 5 Chemical structures of some anti-ND MTDLs investigated by molecular docking analysis

It should be noted that the application of docking-based analysis is not limited to structural studies of already synthesized multi-potent molecules, as it may also help identifying novel MTDLs. A molecular docking-aided investigation reported by Prati et al. demonstrated successful design of multitarget triazinones as BACE-1 and GSK-3 β inhibitors [189]. These two enzymes share only 19% sequence identity and are therefore divergent in nature, and, as such, finding dual-targeting agents for these two enzymes is challenging. Guanidine and cyclic amide fragments to target BACE-1 and GSK-3 β , respectively, were considered to intuitively design of 4-substituted triazone scaffold. Considering this scaffold as a starting point, different derivatives were docked in these two enzymes separately to identify the most suitable dual inhibitors. Subsequent syntheses and biological screening led to identification of one triazone derivative (*vide* 3, Fig. 5) with well-balanced in vitro potencies against these two enzymes, which demonstrated neuroprotective action along with good BBB permeability.

In another recently reported investigation, systematic molecular docking analyses were performed with thiazolylhydrazone derivatives at the binding sites of AChE, BuChE, MAO_A, and MAO_B to identify multitarget AChE/MAO_B inhibitors as anti-PD agents [190]. A common scaffold, 4-(phenyl)-thiazol-2-ylhydrazone (4, Fig. 6), showed docking interactions with all these four enzymes. However, docking orientation of compound 4 also indicated unwanted irreversible inhibition of MAOs due to covalent binding, and in order to avoid such interaction, hydrogen atoms of terminal hydrazone were replaced with methyl groups to produce compound 5 (*vide* Fig. 6). Additionally, molecular docking recognized possible interactions of the phenyl groups of these compounds with electron-rich tryptophan aromatic residues of AChE. Therefore, to promote π -stacking interactions, nitrophenyl derivatives were taken into considerations (*vide* 6, Fig. 6). Subsequent quantum mechanical dihedral energy profile analysis indicated *m*-nitro derivatives as most suitable multitarget agents due to its coplanarity (*vide* 7, Fig. 6). Multitarget activities of these compounds were established through syntheses and experimental bioassays, indicating that docking analysis may take vital roles in taking minor but crucial decision during lead modification.

Similarly, docked structures may be subjected to MD simulation analyses to understand the dynamic behavior of the complex

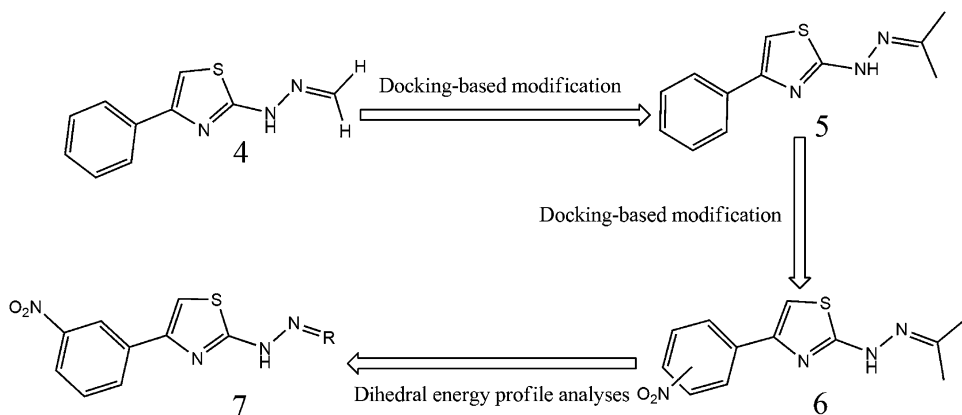


Fig. 6 Stepwise design of multitarget ChE/MAO inhibitors through a molecular docking approach

and the most probable binding patterns of the ligands. Marco-Contelles et al. synthesized some multipotent tacripyridine derivatives as racemic mixtures, and their biological activities were determined against human AChE enzyme [191]. Subsequently MD simulations and binding energy analysis were performed for separate enantiomers of the most potent derivative, revealing that the (R)-enantiomer is the major contributing enantiomer for the anti-ChE activity.

Inverse docking program InvDock was used to identify possible targets for icariin, a natural compound isolated from *Epimedium brevicornum* [124, 192]. The docking results indicated that the anti-AD property of this compound may be mainly related to three factors: attenuation of hyperphosphorylation of tau protein, regulation of Ca²⁺ homeostasis, and anti-inflammation.

5.5 Design of Anti-ND MTDLs: Web-Based Platform for Polypharmacology

In this section, we discuss some relevant studies for the design of anti-ND MTDLs through polypharmacological web-based platform strategies. Some years ago, Liu et al. developed an AD domain-specific chemogenomic knowledgebase named AlzPlatform to support polypharmacology-based drug discovery for AD [193, 194]. The platform assembled 928 AD-related genes, 320 AD-related proteins, and around 200 compounds that were FDA approved or entered clinical trial. Furthermore, this platform was generated with 405,188 chemical compounds with 1,023,137 reported bioactivity data. AlzPlatform is implemented within the molecular database prototype CBID, a MySQL database, an apache web server, and the Open Babel chemical toolbox [195–198]. Additionally, this platform uses Target Hunter program for off-target identification of chemical structures by molecular fingerprint-based similarity searches [199].

Alternatively, the HTDocking program is implemented for docking-based off-target identification of chemical structures [200].

This platform also provides BBB predictor program to understand blood-brain barrier permeation of the compounds [194, 201].

Very recently, AlzhCPI has been introduced as another web-based platform to predict chemical-protein interaction with the help of naive Bayesian and recursive partitioning-based mt-QSAR analyses [202, 203]. This platform was developed based on 204 binary classifiers to predict the chemical-protein interactions for 54 key targets related to AD using the mt-QSAR method. Both AlzPlatform and AlzhCPI may serve as important web-based platforms for the design of multitarget compounds of AD.

6 Conclusive Remarks and Future Directions

It is estimated that probability of success of a novel CNS molecule to become a clinical agent is as low as 3%, making the need for improving the rate of feasibility of development of new CNS drug a paramount and urgent task [96]. Among different CNS agents, anti-ND agents are of special significance due to the highly complex nature of these diseases. Multitarget drug design is one of the best strategies to cope with these complex multifactorial disorders, and, as such, several research groups are currently involved in investigations related to the development of MTDLs for different NDs.

Most of these investigations are focused on a few limited NDs such as AD and PD, as research on other disorders such as HD or ALS have rarely been reported. However, biological targets of these NDs, especially for symptomatic treatments, are often found to be common, and therefore, a major number of recently reported investigations related with design of anti-ND MTDLs adopted molecular hybridization strategies. Nevertheless, alternative rational design methods including *in silico* chemometric techniques were proved to be equally promising.

It is interesting to find that applications of computational chemometric techniques steadily increased in recent years, indicating the growing interest among the scientific community to apply these techniques on multitarget anti-ND agents design.

Proteochemometric modeling (PCM) is a comparatively novel chemometric method that aims at predicting bioactivity more generally, by integrating descriptors of both the ligand and protein target along with cross-term interactions [204]. Despite tremendous potential, proteochemometric modeling has been less utilized so far in MTDL design. Furthermore, network polypharmacology may serve as a better option for new targets of multifactorial diseases [205]. In fact, recently, Chen et al. introduced a noncommercial, Internet-accessible database called the Multiple Target Ligand Database (MTLD) to facilitate the MTDL-based design [206]. Finally, chemometric techniques are not only helpful for the design of MTDLs, as they are also predicting ADMET profiles of the

molecules, in order to save considerable time and efforts. Currently available in silico ADMET models may greatly contribute to the knowledge of screening approaches in the early stages of drug discovery and the development process [207].

Acknowledgments

This work is supported by Fundação para a Ciência e a Tecnologia (FCT/MEC) through national funds and co-financed by FEDER, under the partnership agreement PT2020 (Projects UID/QUI/50006/2013 and POCI/01/0145/FEDER/007265). To all financing sources, the authors are greatly indebted.

References

1. Gitler AD, Dhillon P, Shorter J (2017) Neurodegenerative disease: models, mechanisms, and a new hope. *Dis Model Mech* 10(5):499–502
2. Yacoubian TA (2017) Chapter 1 – neurodegenerative disorders: why do we need new therapies? A2 – Adejare, Adeboye. In: *Drug discovery approaches for the treatment of neurodegenerative disorders*. Academic Press, New York, pp 1–16
3. Lardenoije R, Iatrou A, Kenis G, Kompotis K, Steinbusch HW, Mastroeni D, Coleman P, Lemere CA, Hof PR, van den Hove DL, Rutten BP (2015) The epigenetics of aging and neurodegeneration. *Prog Neurobiol* 131:21–64
4. Coleby R (2017) Medicinal chemistry advances in neurodegenerative disease therapy: part 2. *Future Med Chem* 9(10):951–952
5. Arevalo-Villalobos JI, Rosales-Mendoza S, Zarazua S (2017) Immunotherapies for neurodegenerative diseases: current status and potential of plant-made biopharmaceuticals. *Expert Rev Vaccines* 16(2):151–159
6. <http://www.alz.co.uk/research/statistics>
7. Pen AE, Jensen UB (2017) Current status of treating neurodegenerative disease with induced pluripotent stem cells. *Acta Neurol Scand* 135(1):57–72
8. Khanam H, Ali A, Asif M, Shamsuzzaman (2016) Neurodegenerative diseases linked to misfolded proteins and their therapeutic approaches: a review. *Eur J Med Chem* 124:1121–1141
9. Cummings J, Aisen PS, DuBois B, Frolich L, Jack CR Jr, Jones RW, Morris JC, Raskin J, Dowsett SA, Scheltens P (2016) Drug development in Alzheimer's disease: the path to 2025. *Alzheimers Res Ther* 8:39
10. Mathis S, Couratier P, Julian A, Vallat JM, Corcia P, Le Masson G (2017) Management and therapeutic perspectives in amyotrophic lateral sclerosis. *Expert Rev Neurother* 17(3):263–276
11. Anighoro A, Bajorath J, Rastelli G (2014) Polypharmacology: challenges and opportunities in drug discovery. *J Med Chem* 57(19):7874–7887
12. Cavalli A, Bolognesi ML, Minarini A, Rosini M, Tumiatti V, Recanatini M, Melchiorre C (2008) Multi-target-directed ligands to combat neurodegenerative diseases. *J Med Chem* 51(3):347–372
13. Lauria A, Bonsignore R, Bartolotta R, Perricone U, Martorana A, Gentile C (2016) Drugs polypharmacology by in silico methods: new opportunities in drug discovery. *Curr Pharm Des* 22(21):3073–3081
14. Satyanarayanan SD (2011) Drug design and discovery: methods and protocols. In: *Methods in molecular biology*, vol 716. Humana Press, New York
15. Zheng Y (2012) Rational drug design: methods and protocols. In: *Method in molecular biology*, vol 928. Humana Press, Springer, New York
16. Bajorath J (2015) Pushing the boundaries of computational approaches: special focus issue on computational chemistry and computer-aided drug discovery. *Future Med Chem* 7(18):2415–2417
17. Bajorath J (2016) Computational chemistry and computer-aided drug discovery: part II. *Future Med Chem* 8(15):1799–1800
18. Jamal S, Grover A (2017) Cheminformatics approaches in modern drug discovery. In:

- Grover A (ed) Drug design: principles and applications. Springer Singapore, Singapore, pp 135–148
19. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, Dearden J, Gramatica P, Martin YC, Todeschini R, Consonni V, Kuz'min VE, Cramer R, Benigni R, Yang C, Rathman J, Terfloth L, Gasteiger J, Richard A, Tropsha A (2014) QSAR modeling: where have you been? Where are you going to? *J Med Chem* 57(12):4977–5010
 20. Lavine BK (2005) Chemometrics and chemoinformatics, ACS symposium series, vol vol 894. American Chemical Society, Washington, DC
 21. Begam BF, Kumar JS (2012) A study on cheminformatics and its applications on modern drug discovery. *Procedia Eng* 38(suppl C):1264–1275
 22. Dugger BN, Dickson DW (2017) Pathology of neurodegenerative diseases. *Cold Spring Harb Perspect Biol* 9(7):a028035
 23. Labzin LI, Heneka MT, Latz E (2017) Innate immunity and neurodegeneration. *Annu Rev Med* 69:437–449
 24. Chetelat G, Villemagne VL, Pike KE, Ellis KA, Bourgeat P, Jones G, O'Keefe GJ, Salvado O, Szoceke C, Martins RN, Ames D, Masters CL, Rowe CC, Australian Imaging B, Lifestyle Study of Ageing Research G (2011) Independent contribution of temporal beta-amyloid deposition to memory decline in the pre-dementia phase of Alzheimer's disease. *Brain* 134(Pt 3):798–807
 25. Possin KL (2010) Visual spatial cognition in neurodegenerative disease. *Neurocase* 16(6):466–487
 26. Ross CA, Poirier MA (2004) Protein aggregation and neurodegenerative disease. *Nat Med* 10:S10
 27. Jacobsen KT, Iverfeldt K (2009) Amyloid precursor protein and its homologues: a family of proteolysis-dependent receptors. *Cell Mol Life Sci* 66(14):2299–2318
 28. Larner AJ (2013) Presenilin-1 mutations in Alzheimer's disease: an update on genotype-phenotype relationships. *J Alzheimers Dis* 37(4):653–659
 29. Adeniji AO, Adams PW, Mody VV (2017) Chapter 7 – amyloid β hypothesis in the development of therapeutic agents for Alzheimer's disease A2 – Adejare, Adeboye. In: Drug discovery approaches for the treatment of neurodegenerative disorders. Academic Press, New York, pp 109–143
 30. Kalia LV, Lang AE (2015) Parkinson's disease. *Lancet* 386(9996):896–912
 31. Ferreira M, Massano J (2017) An updated review of Parkinson's disease genetics and clinicopathological correlations. *Acta Neurol Scand* 135(3):273–284
 32. Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M (1997) Alpha-synuclein in Lewy bodies. *Nature* 388(6645):839–840
 33. Dawson TM, Dawson VL (2003) Rare genetic mutations shed light on the pathogenesis of Parkinson disease. *J Clin Invest* 111(2):145–151
 34. Roos RA (2010) Huntington's disease: a clinical review. *Orphanet J Rare Dis* 5:40
 35. Ross CA, Tabrizi SJ (2011) Huntington's disease: from molecular pathogenesis to clinical treatment. *Lancet Neurol* 10(1):83–98
 36. Moncke-Buchner E, Reich S, Mucke M, Reuter M, Messer W, Wanker EE, Kruger DH (2002) Counting CAG repeats in the Huntington's disease gene by restriction endonuclease EcoP15I cleavage. *Nucleic Acids Res* 30(16):e83
 37. Witgert M, Salamone AR, Strutt AM, Jawaid A, Massman PJ, Bradshaw M, Mosnik D, Appel SH, Schulz PE (2010) Frontal-lobe mediated behavioral dysfunction in amyotrophic lateral sclerosis. *Eur J Neurol* 17(1):103–110
 38. Zarei S, Carr K, Reiley L, Diaz K, Guerra O, Altamirano PF, Pagani W, Lodin D, Orozco G, China A (2015) A comprehensive review of amyotrophic lateral sclerosis. *Surg Neurol Int* 6:171
 39. Kovacs GG, Budka H (2008) Prion diseases: from protein to cell pathology. *Am J Pathol* 172(3):555–565
 40. Piccardo P, Manson JC, King D, Ghetti B, Barron RM (2007) Accumulation of prion protein in the brain that is not associated with transmissible disease. *Proc Natl Acad Sci U S A* 104(11):4712–4717
 41. Stayte S, Vissel B (2014) Advances in non-dopaminergic treatments for Parkinson's disease. *Front Neurosci* 8:113
 42. Casey DA, Antimisiaris D, O'Brien J (2010) Drugs for Alzheimer's disease: are they effective? *P T* 35(4):208–211
 43. Yiannopoulou KG, Papageorgiou SG (2013) Current and future treatments for Alzheimer's disease. *Ther Adv Neurol Disord* 6(1):19–33
 44. Cummings J, Lee G, Mortsdorf T, Ritter A, Zhong K (2017) Alzheimer's disease drug

- development pipeline: 2017. *Alzheimers Dement (N Y)* 3(3):367–384
45. Miller RG, Jackson CE, Kasarskis EJ, England JD, Forshew D, Johnston W, Kalra S, Katz JS, Mitsumoto H, Rosenfeld J, Shoesmith C, Strong MJ, Woolley SC, Quality Standards Subcommittee of the American Academy of N (2009) Practice parameter update: the care of the patient with amyotrophic lateral sclerosis: drug, nutritional, and respiratory therapies (an evidence-based review): report of the quality standards Subcommittee of the American Academy of Neurology. *Neurology* 73(15):1218–1226
 46. Paleacu D (2007) Tetrabenazine in the treatment of Huntington's disease. *Neuropsychiatr Dis Treat* 3(5):545–551
 47. Mencher SK, Wang LG (2005) Promiscuous drugs compared to selective drugs (promiscuity can be a virtue). *BMC Clin Pharmacol* 5:3
 48. Lu JJ, Pan W, Hu YJ, Wang YT (2012) Multitarget drugs: the trend of drug research and development. *PLoS One* 7(6):e40262
 49. Iyengar R (2013) Complex diseases require complex therapies. *EMBO Rep* 14(12):1039–1042
 50. Morphy R, Rankovic Z (2005) Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem* 48(21):6523–6543
 51. Bolognesi ML (2013) Polypharmacology in a single drug: multitarget drugs. *Curr Med Chem* 20(13):1639–1645
 52. Keith CT, Borisy AA, Stockwell BR (2005) Multicomponent therapeutics for networked systems. *Nat Rev Drug Discov* 4(1):71–78
 53. Smid P, Coolen HK, Keizer HG, van Hes R, de Moes JP, den Hartog AP, Stork B, Plekkenpol RH, Niemann LC, Stroemer CN, Tulp MT, van Stuijvenberg HH, McCreary AC, Hesselink MB, Herremans AH, Kruse CG (2005) Synthesis, structure-activity relationships, and biological properties of 1-heteroaryl-4-[omega-(1H-indol-3-yl)alkyl] piperazines, novel potential antipsychotics combining potent dopamine D2 receptor antagonism with potent serotonin reuptake inhibition. *J Med Chem* 48(22):6855–6869
 54. Tsuji S (2010) Genetics of neurodegenerative diseases: insights from high-throughput resequencing. *Hum Mol Genet* 19(R1):R65–R70
 55. Van der Schyf CJ (2011) The use of multitarget drugs in the treatment of neurodegenerative diseases. *Expert Rev Clin Pharmacol* 4(3):293–298
 56. Weinreb O, Amit T, Bar-Am O, Yogev-Falach M, Youdim MB (2008) The neuroprotective mechanism of action of the multimodal drug ladostigil. *Front Biosci* 13:5131–5137
 57. Prati F, Uliassi E, Bolognesi ML (2014) Two diseases, one approach: multitarget drug discovery in Alzheimer's and neglected tropical diseases. *MedChemComm* 5(7):853–861
 58. Beebe KR, Pell RJ, Seasholtz MB (1998) *Chemometrics: A practical guide*. John Wiley & Sons, Inc., New York
 59. Gad HA, El-Ahmady SH, Abou-Shoer MI, Al-Azizi MM (2013) Application of chemometrics in authentication of herbal medicines: a review. *Phytochem Anal* 24(1):1–24
 60. Huang H-J, Yu HW, Chen C-Y, Hsu C-H, Chen H-Y, Lee K-J, Tsai F-J, Chen CY-C (2010) Current developments of computer-aided drug design. *J Taiwan Inst Chem Eng* 41(6):623–635
 61. Christofferson AJ, Huang N (2012) How to benchmark methods for structure-based virtual screening of large compound libraries. *Methods Mol Biol* 819:187–195
 62. Halder AK, Saha A, Jha T (2013) The role of 3D pharmacophore mapping based virtual screening for identification of novel anticancer agents: an overview. *Curr Top Med Chem* 13(9):1098–1126
 63. Lipinski CA (2004) Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol* 1(4):337–341
 64. Hansch C, Maloney PP, Fujita T, Muir RM (1962) Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. *Nature* 194:178
 65. Cramer RD (2012) The inevitable QSAR renaissance. *J Comput Aided Mol Des* 26(1):35–38
 66. Todeschini R, Consonni V (2000) *Handbook of molecular descriptors*, vol 11. Wiley VCH, Weinheim
 67. Quintero FA, Patel SJ, Muñoz F, Sam Manman M (2012) Review of existing QSAR/QSPR models developed for properties used in hazardous chemicals classification system. *Ind Eng Chem Res* 51(49):16101–16115
 68. Lewis RA, Wood D (2014) Modern 2D QSAR for drug discovery. *Wiley Interdiscip Rev Comput Mol Sci* 4(6):505–522
 69. Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 110(18):5959–5967
 70. Klebe G, Abraham U (1999) Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. *J Comput Aided Mol Des* 13(1):1–10

71. Robinson DD, Winn PJ, Lyne PD, Richards WG (1999) Self-organizing molecular field analysis: a tool for structure-activity studies. *J Med Chem* 42(4):573–583
72. Ajmani S, Jadhav K, Kulkarni SA (2006) Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. *J Chem Inf Model* 46(1):24–31
73. Cramer RD (2003) Topomer CoMFA: a design methodology for rapid lead optimization. *J Med Chem* 46(3):374–388
74. Tosco P, Balle T (2011) Open3DQSAR: a new open-source software aimed at high-throughput chemometric analysis of molecular interaction fields. *J Mol Model* 17(1):201–208
75. Hechinger M, Leonhard K, Marquardt W (2012) What is wrong with quantitative structure-property relations models based on three-dimensional descriptors? *J Chem Inf Model* 52(8):1984–1993
76. Wang T, Wu MB, Lin JP, Yang LR (2015) Quantitative structure-activity relationship: promising advances in drug discovery platforms. *Expert Opin Drug Discov* 10(12):1283–1300
77. Pastor M, Cruciani G, McLay I, Pickett S, Clementi S (2000) GRIND-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors. *J Med Chem* 43(17):3233–3243
78. Fang J, Li Y, Liu R, Pang X, Li C, Yang R, He Y, Lian W, Liu A-L, Du G-H (2015) Discovery of multitarget-directed ligands against Alzheimer's disease through systematic prediction of chemical-protein interactions. *J Chem Inf Model* 55(1):149–164
79. Speck-Planche A, Cordeiro MNDS (2017) Advanced in silico approaches for drug discovery: mining information from multiple biological and chemical data through mtk-QSBER and pt-QSPR strategies. *Curr Med Chem* 24(16):1687–1704
80. Speck-Planche A, Cordeiro MNDS (2015) Multitasking models for quantitative structure-biological effect relationships: current status and future perspectives to speed up drug discovery. *Expert Opin Drug Discov* 10(3):245–256
81. Speck-Planche A, Kleandrova VV, Ruso JM, Cordeiro MNDS (2016) First multitarget chemo-bioinformatic model to enable the discovery of antibacterial peptides against multiple gram-positive pathogens. *J Chem Inf Model* 56(3):588–598
82. Antanasijevic D, Antanasijevic J, Trisovic N, Uscumlic G, Pocajt V (2017) From classification to regression multitasking QSAR modeling using a novel modular neural network: simultaneous prediction of anticonvulsant activity and neurotoxicity of succinimides. *Mol Pharm* 14(12):4476–4484
83. Casanola-Martin GM, Le-Thi-Thu H, Perez-Gimenez F, Marrero-Ponce Y, Merino-Sanjuan M, Abad C, Gonzalez-Diaz H (2016) Multi-output model with box-Jenkins operators of quadratic indices for prediction of malaria and cancer inhibitors targeting ubiquitin-proteasome pathway (UPP) proteins. *Curr Protein Pept Sci* 17(3):220–227
84. Casanola-Martin GM, Le-Thi-Thu H, Perez-Gimenez F, Marrero-Ponce Y, Merino-Sanjuan M, Abad C, Gonzalez-Diaz H (2015) Multi-output model with box-Jenkins operators of linear indices to predict multi-target inhibitors of ubiquitin-proteasome pathway. *Mol Divers* 19(2):347–356
85. Wathieu H, Issa NT, Byers SW, Dakshana-murthy S (2016) Harnessing polypharmacology with computer-aided drug design and systems biology. *Curr Pharm Des* 22(21):3097–3108
86. Matter H (1997) Selecting optimally diverse compounds from structure databases: a validation study of two-dimensional and three-dimensional molecular descriptors. *J Med Chem* 40(8):1219–1229
87. Bajusz D, Rácz A, Héberger K (2015) Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J Cheminform* 7(1):20
88. Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK (2007) Relating protein pharmacology by ligand chemistry. *Nat Biotechnol* 25(2):197–206
89. Huang T, Mi H, Lin CY, Zhao L, Zhong LL, Liu FB, Zhang G, Lu AP, Bian ZX, for MG (2017) MOST: most-similar ligand based approach to target prediction. *BMC Bioinformatics* 18(1):165
90. Yang SY (2010) Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov Today* 15(11–12):444–450
91. Mason JS, Good AC, Martin EJ (2001) 3-D pharmacophores in drug discovery. *Curr Pharm Des* 7(7):567–597
92. Halder AK, Mallick S, Shikha D, Saha A, Saha KD, Jha T (2015) Design of dual MMP-2/HDAC-8 inhibitors by pharmacophore mapping, molecular docking, synthesis and biological activity. *RSC Adv* 5(88):72373–72386
93. <http://59.78.96.61/pharmmapper/>

94. Richmond NJ, Abrams CA, Wolohan PR, Abrahamian E, Willett P, Clark RD (2006) GALAHAD: I. Pharmacophore identification by hypermolecular alignment of ligands in 3D. *J Comput Aided Mol Des* 20(9):567–587
95. Van Drie JH, Weininger D, Martin YC (1989) ALADDIN: an integrated tool for computer-assisted molecular design and pharmacophore recognition from geometric, steric, and substructure searching of three-dimensional molecular structures. *J Comput Aided Mol Des* 3(3):225–251
96. Van Drie JH (1996) An inequality for 3D database searching and its use in evaluating the treatment of conformational flexibility. *J Comput Aided Mol Des* 10(6):623–630
97. Martin YC, Bures MG, Danaher EA, DeLazzer J, Lico I, Pavlik PA (1993) A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. *J Comput Aided Mol Des* 7(1):83–102
98. Salam NK, Nuti R, Sherman W (2009) Novel method for generating structure-based pharmacophores using energetic analysis. *J Chem Inf Model* 49(10):2356–2368
99. Jones G, Willett P, Glen RC (1995) A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. *J Comput Aided Mol Des* 9(6):532–549
100. Barnum D, Greene J, Smellie A, Sprague P (1996) Identification of common functional configurations among molecules. *J Chem Inf Comput Sci* 36(3):563–571
101. Li H, Sutter J, Hoffman R (2000) HypoGen: an automated system for generating 3D predictive pharmacophore models. In: Guner O (ed) *Pharmacophore perception, development and use in drug design*. International University Line, La Jolla, CA
102. Wolber G, Langer T (2005) LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J Chem Inf Model* 45(1):160–169
103. Molecular Operating Environment (MOE). Chemical Computing Group. www.chemcomp.com
104. Holliday JD, Willett P (1997) Using a genetic algorithm to identify common structural features in sets of ligands. *J Mol Graph Model* 15(4):221–232
105. Dixon SL, Smondryev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA (2006) PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: I. Methodology and preliminary results. *J Comput Aided Mol Des* 20(10–11):647–671
106. Chen YC (2015) Beware of docking! *Trends Pharmacol Sci* 36(2):78–95
107. Gabel J, Desaphy J, Rognan D (2014) Beware of machine learning-based scoring functions on the danger of developing black boxes. *J Chem Inf Model* 54(10):2807–2815
108. Koebel MR, Schmadeke G, Posner RG, Sirimulla S (2016) AutoDock VinaXB: implementation of XBSF, new empirical halogen bond scoring function, into AutoDock Vina. *J Cheminform* 8:27
109. Lang PT, Brozell SR, Mukherjee S, Pettersen EF, Meng EC, Thomas V, Rizzo RC, Case DA, James TL, Kuntz ID (2009) DOCK 6: combining techniques to model RNA-small molecule complexes. *RNA* 15(6):1219–1230
110. Rarey M, Kramer B, Lengauer T, Klebe G (1996) A fast flexible docking method using an incremental construction algorithm. *J Mol Biol* 261(3):470–489
111. McGann M (2012) FRED and HYBRID docking performance on standardized datasets. *J Comput Aided Mol Des* 26(8):897–906
112. Schrodinger release 2017-1: Glide. Schrodinger, LLC, New York, NY, 2017
113. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD (2003) Improved protein-ligand docking using GOLD. *Proteins* 52(4):609–623
114. Venkatachalam CM, Jiang X, Oldfield T, Waldman M (2003) LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J Mol Graph Model* 21(4):289–307
115. Jain AN (2007) Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. *J Comput Aided Mol Des* 21(5):281–306
116. VLifeMDS: molecular design suite. VLife Sciences Technologies Pvt. Ltd., Pune, India, 2010
117. <https://www.schrodinger.com/sitemap>
118. Ngan CH, Bohnuud T, Mottarella SE, Beglov D, Villar EA, Hall DR, Kozakov D, Vajda S (2012) FTMAP: extended protein mapping with user-selected probe molecules. *Nucleic Acids Res* 40(web server issue):W271–W275
119. Lionta E, Spyrou G, Vassilatis DK, Cournia Z (2014) Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem* 14(16):1923–1938

120. Le Guilloux V, Schmidtke P, Tuffery P (2009) Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* 10:168
121. Amaro RE, Li WW (2010) Emerging methods for ensemble-based virtual screening. *Curr Top Med Chem* 10(1):3–13
122. Csermely P, Agoston V, Pongor S (2005) The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol Sci* 26(4):178–182
123. Durrant JD, McCammon JA (2011) Molecular dynamics simulations and drug discovery. *BMC Biol* 9:71
124. <http://bidd.nus.edu.sg/group/software/invdock.htm>
125. Wang JC, Chu PY, Chen CM, Lin JH (2012) idTarget: a web server for identifying protein targets of small chemical molecules with robust scoring functions and a divide-and-conquer docking approach. *Nucleic Acids Res* 40(web server issue):W393–W399
126. Chen YZ, Ung CY (2002) Computer automated prediction of potential therapeutic and toxicity protein targets of bioactive compounds from Chinese medicinal plants. *Am J Chin Med* 30(1):139–154
127. Chaudhari R, Tan Z, Huang B, Zhang S (2017) Computational polypharmacology: a new paradigm for drug discovery. *Expert Opin Drug Discov* 12(3):279–291
128. Chartier M, Adriansen E, Najmanovich R (2016) IsoMIF finder: online detection of binding site molecular interaction field similarities. *Bioinformatics* 32(4):621–623
129. Awale M, Reymond JL (2017) The polypharmacology browser: a web-based multi-fingerprint target prediction tool using ChEMBL bioactivity data. *J Cheminform* 9:11
130. <http://www.ebi.ac.uk/chembl/db>
131. <http://pubchem.ncbi.nlm.nih.gov>
132. <http://insilico.charite.de/supertarget/>
133. <http://r2d2drug.org/DMC.aspx>
134. <http://bidd.nus.edu.sg/group/cjttd>
135. <https://www.bindingdb.org/>
136. <http://www.drugbank.ca>
137. <http://www.t3db.org>
138. <http://thomsonreuters.com/metacore>
139. <http://www.biocarta.com>
140. <http://www.pantherdb.org>
141. <http://www.hmdb.ca>
142. <https://www.rcsb.org/>
143. <http://www.uniprot.org/>
144. Bitam S, Hamadache M, Hanini S (2017) QSAR model for prediction of the therapeutic potency of N-benzylpiperidine derivatives as AChE inhibitors. *SAR QSAR Environ Res* 28(6):471–489
145. Wong KY, Mercader AG, Saavedra LM, Honarparvar B, Romanelli GP, Duchowicz PR (2014) QSAR analysis on tacrine-related acetylcholinesterase inhibitors. *J Biomed Sci* 21:84
146. Vats C, Dhanjal JK, Goyal S, Bharadvaja N, Grover A (2015) Computational design of novel flavonoid analogues as potential AChE inhibitors: analysis using group-based QSAR, molecular docking and molecular dynamics simulations. *Struct Chem* 26(2):467–476
147. Mahmoodabadi N, Ajloo D (2016) QSAR, docking, and molecular dynamic studies on the polyphenolic as inhibitors of β -amyloid aggregation. *Med Chem Res* 25(10):2104–2118
148. Ambure P, Roy K (2016) Understanding the structural requirements of cyclic sulfone hydroxyethylamines as hBACE1 inhibitors against A[small beta] plaques in Alzheimer's disease: a predictive QSAR approach. *RSC Adv* 6(34):28171–28186
149. Toropova MA, Toropov AA, Raska I Jr, Raskova M (2015) Searching therapeutic agents for treatment of Alzheimer disease using the Monte Carlo method. *Comput Biol Med* 64:148–154
150. Niu B, Zhao M, Su Q, Zhang M, Lv W, Chen Q, Chen F, Chu D, Du D, Zhang Y (2017) 2D-SAR and 3D-QSAR analyses for acetylcholinesterase inhibitors. *Mol Divers* 21(2):413–426
151. Ambure P, Roy K (2014) Exploring structural requirements of leads for improving activity and selectivity against CDK5/p25 in Alzheimer's disease: an in silico approach. *RSC Adv* 4(13):6702–6709
152. Jain P, Jadhav HR (2013) Quantitative structure activity relationship analysis of aminoimidazoles as BACE-I inhibitors. *Med Chem Res* 22(4):1740–1746
153. Cai C, Wu Q, Luo Y, Ma H, Shen J, Zhang Y, Yang L, Chen Y, Wen Z, Wang Q (2017) In silico prediction of ROCK II inhibitors by different classification approaches. *Mol Divers* 21(4):791–807
154. Gharaghani S, Khayamian T, Ebrahimi M (2013) Molecular dynamics simulation study and molecular docking descriptors in structure-based QSAR on acetylcholinesterase (AChE) inhibitors. *SAR QSAR Environ Res* 24(9):773–794
155. Helguera AM, Perez-Garrido A, Gaspar A, Reis J, Cagide F, Vina D, Cordeiro MNDS, Borges F (2013) Combining QSAR

- classification models for predictive modeling of human monoamine oxidase inhibitors. *Eur J Med Chem* 59:75–90
156. Lu P, Wei X, Zhang R, Yuan Y, Gong Z (2011) Prediction of the binding affinities of adenosine A2A receptor antagonists based on the heuristic method and support vector machine. *Med Chem Res* 20(8):1220–1228
 157. Karolidis DA, Agatonovic-Kustrin S, Morton DW (2010) Artificial neural network (ANN) based modelling for D1 like and D2 like dopamine receptor affinity and selectivity. *Med Chem* 6(5):259–270
 158. Speck-Planche A, Kleandrova VV (2012) QSAR and molecular docking techniques for the discovery of potent monoamine oxidase B inhibitors: computer-aided generation of new rasagiline bioisosteres. *Curr Top Med Chem* 12(16):1734–1747
 159. Kahn I, Lomaka A, Karelson M (2014) Topological fingerprints as an aid in finding structural patterns for LRRK2 inhibition. *Mol Inform* 33(4):269–275
 160. Zambre VP, Hambarde VA, Petkar NN, Patel CN, Sawant SD (2015) Structural investigations by in silico modeling for designing NR2B subunit selective NMDA receptor antagonists. *RSC Adv* 5(30):23922–23940
 161. Amin SA, Adhikari N, Jha T, Gayen S (2016) First molecular modeling report on novel arylpyrimidine kynurenine monooxygenase inhibitors through multi-QSAR analysis against Huntington's disease: a proposal to chemists! *Bioorg Med Chem Lett* 26(23):5712–5718
 162. Joshi K, Goyal S, Grover S, Jamal S, Singh A, Dhar P, Grover A (2016) Novel group-based QSAR and combinatorial design of CK-1delta inhibitors as neuroprotective agents. *BMC Bioinformatics* 17(suppl 19):515
 163. Amirhamzeh A, Vosoughi M, Shafiee A, Amini M (2013) Synthesis and docking study of diaryl-isothiazole and 1,2,3-thiadiazole derivatives as potential neuroprotective agents. *Med Chem Res* 22(3):1212–1223
 164. Fang J, Pang X, Yan R, Lian W, Li C, Wang Q, Liu A-L, Du G-H (2016) Discovery of neuroprotective compounds by machine learning approaches. *RSC Adv* 6(12):9857–9871
 165. Durant JL, Leland BA, Henry DR, Nourse JG (2002) Reoptimization of MDL keys for use in drug discovery. *J Chem Inf Comput Sci* 42(6):1273–1280
 166. Rogers D, Hahn M (2010) Extended-connectivity fingerprints. *J Chem Inf Model* 50(5):742–754
 167. Besnard J, Ruda GF, Setola V, Abecassis K, Rodriguiz RM, Huang XP, Norval S, Sassano MF, Shin AI, Webster LA, Simeons FR, Stojanovski L, Prat A, Seidah NG, Constam DB, Bickerton GR, Read KD, Wetsel WC, Gilbert IH, Roth BL, Hopkins AL (2012) Automated design of ligands to polypharmacological profiles. *Nature* 492(7428):215–220
 168. Speck-Planche A, Kleandrova VV, Luan F, Cordeiro MNDS (2013) Multi-target inhibitors for proteins associated with Alzheimer: in silico discovery using fragment-based descriptors. *Curr Alzheimer Res* 10(2):117–124
 169. Luan F, Cordeiro MNDS, Alonso N, Garcia-Mera X, Caamano O, Romero-Duran FJ, Yanez M, Gonzalez-Diaz H (2013) TOPS-MODE model of multiplexing neuroprotective effects of drugs and experimental-theoretic study of new 1,3-rasagiline derivatives potentially useful in neurodegenerative diseases. *Bioorg Med Chem* 21(7):1870–1879
 170. Alonso N, Caamano O, Romero-Duran FJ, Luan F, Cordeiro MNDS, Yanez M, Gonzalez-Diaz H, Garcia-Mera X (2013) Model for high-throughput screening of multitarget drugs in chemical neurosciences: synthesis, assay, and theoretic study of rasagiline carbamates. *ACS Chem Neurosci* 4(10):1393–1403
 171. Romero Duran FJ, Alonso N, Caamano O, Garcia-Mera X, Yanez M, Prado-Prado FJ, Gonzalez-Diaz H (2014) Prediction of multi-target networks of neuroprotective compounds with entropy indices and synthesis, assay, and theoretical study of new asymmetric 1,2-rasagiline carbamates. *Int J Mol Sci* 15(9):17035–17064
 172. Bautista-Aguilera OM, Esteban G, Bolea I, Nikolic K, Agbaba D, Moraleda I, Iriepa I, Samadi A, Soriano E, Unzeta M, Marco-Contelles J (2014) Design, synthesis, pharmacological evaluation, QSAR analysis, molecular modeling and ADMET of novel donepezil-indolyl hybrids as multipotent cholinesterase/monoamine oxidase inhibitors for the potential treatment of Alzheimer's disease. *Eur J Med Chem* 75:82–95
 173. Bautista-Aguilera OM, Esteban G, Chioua M, Nikolic K, Agbaba D, Moraleda I, Iriepa I, Soriano E, Samadi A, Unzeta M, Marco-Contelles J (2014) Multipotent cholinesterase/monoamine oxidase inhibitors for the treatment of Alzheimer's disease: design, synthesis, biochemical evaluation, ADMET, molecular modeling, and QSAR analysis of

- novel donepezil-pyridyl hybrids. *Drug Des Devel Ther* 8:1893–1910
174. http://www.moldiscovery.com/soft_pentacle.php
175. Nikolic K, Agbaba D, Stark H (2015) Pharmacophore modeling, drug design and virtual screening on multi-targeting procognitive agents approaching histaminergic pathways. *J Taiwan Inst Chem Eng* 46(suppl C):15–29
176. Nikolic K, Mavridis L, Bautista-Aguilera OM, Marco-Contelles J, Stark H, do Carmo Carreiras M, Rossi I, Massarelli P, Agbaba D, Ramsay RR, Mitchell JB (2015) Predicting targets of compounds against neurological diseases using cheminformatic methodology. *J Comput Aided Mol Des* 29(2):183–198
177. Huang W, Yu H, Sheng R, Li J, Hu Y (2008) Identification of pharmacophore model, synthesis and biological evaluation of N-phenyl-1-arylamide and N-phenylbenzenesulfonamide derivatives as BACE 1 inhibitors. *Bioorg Med Chem* 16(24):10190–10197
178. Huang W, Lv D, Yu H, Sheng R, Kim SC, Wu P, Luo K, Li J, Hu Y (2010) Dual-target-directed 1,3-diphenylurea derivatives: BACE 1 inhibitor and metal chelator against Alzheimer's disease. *Bioorg Med Chem* 18(15):5610–5615
179. Huang W, Tang L, Shi Y, Huang S, Xu L, Sheng R, Wu P, Li J, Zhou N, Hu Y (2011) Searching for the multi-target-directed ligands against Alzheimer's disease: discovery of quinoxaline-based hybrid compounds with AChE, H(3)R and BACE 1 inhibitory activities. *Bioorg Med Chem* 19(23):7158–7167
180. Xie Q, Zheng Z, Shao B, Fu W, Xia Z, Li W, Sun J, Zheng W, Zhang W, Sheng W, Zhang Q, Chen H, Wang H, Qiu Z (2017) Pharmacophore-based design and discovery of (–)-meptazinol carbamates as dual modulators of cholinesterase and amyloidogenesis. *J Enzyme Inhib Med Chem* 32(1):659–671
181. Bhayye SS, Roy K, Saha A (2016) Pharmacophore generation, atom-based 3D-QSAR, HQSAR and activity cliff analyses of benzothiazine and deaxanthine derivatives as dual A2A antagonists/MAOB inhibitors. *SAR QSAR Environ Res*:1–20
182. Heritage TW, Lowis DR (1999) Molecular hologram QSAR. In: *Rational drug design, ACS symposium series, vol vol 719*. American Chemical Society, Washington, DC, pp 212–225
183. Inestrosa NC, Alvarez A, Calderon F (1996) Acetylcholinesterase is a senile plaque component that promotes assembly of amyloid beta-peptide into Alzheimer's filaments. *Mol Psychiatry* 1(5):359–361
184. Bartolini M, Bertucci C, Cavrini V, Andrisano V (2003) beta-Amyloid aggregation induced by human acetylcholinesterase: inhibition studies. *Biochem Pharmacol* 65(3):407–416
185. Ismaili L, Refouvelet B, Benchekroun M, Brogi S, Brindisi M, Gemma S, Campiani G, Filipic S, Agbaba D, Esteban G, Unzeta M, Nikolic K, Butini S, Marco-Contelles J (2017) Multitarget compounds bearing tacrine- and donepezil-like structural and functional motifs for the potential treatment of Alzheimer's disease. *Prog Neurobiol* 151:4–34
186. Xie SS, Lan JS, Wang X, Wang ZM, Jiang N, Li F, Wu JJ, Wang J, Kong LY (2016) Design, synthesis and biological evaluation of novel donepezil-coumarin hybrids as multi-target agents for the treatment of Alzheimer's disease. *Bioorg Med Chem* 24(7):1528–1539
187. Li SY, Jiang N, Xie SS, Wang KD, Wang XB, Kong LY (2014) Design, synthesis and evaluation of novel tacrine-rhein hybrids as multi-functional agents for the treatment of Alzheimer's disease. *Org Biomol Chem* 12(5):801–814
188. Xie SS, Wang X, Jiang N, Yu W, Wang KD, Lan JS, Li ZR, Kong LY (2015) Multi-target tacrine-coumarin hybrids: cholinesterase and monoamine oxidase B inhibition properties against Alzheimer's disease. *Eur J Med Chem* 95:153–165
189. Prati F, De Simone A, Bisignano P, Armirotti A, Summa M, Pizzirani D, Scarpelli R, Perez DI, Andrisano V, Perez-Castillo A, Monti B, Massenzio F, Polito L, Racchi M, Favia AD, Bottegoni G, Martinez A, Bolognesi ML, Cavalli A (2015) Multitarget drug discovery for Alzheimer's disease: triazinones as BACE-1 and GSK-3beta inhibitors. *Angew Chem Int Ed Engl* 54(5):1578–1582
190. Carradori S, Ortuso F, Petzer A, Bagetta D, De Monte C, Secci D, De Vita D, Guglielmi P, Zengin G, Aktumsek A, Alcaro S, Petzer JP (2018) Design, synthesis and biochemical evaluation of novel multitarget inhibitors as potential anti-Parkinson agents. *Eur J Med Chem* 143:1543–1552
191. Marco-Contelles J, Leon R, de los Rios C, Samadi A, Bartolini M, Andrisano V, Huertas O, Barril X, Luque FJ, Rodriguez-Franco MI, Lopez B, Lopez MG, Garcia AG, Carreiras Mdo C, Villarroya M (2009)

- Tacipyridines, the first tacrine-dihydropyridine hybrids, as multitarget-directed ligands for the treatment of Alzheimer's disease. *J Med Chem* 52(9):2724–2732
192. Cui Z, Sheng Z, Yan X, Cao Z, Tang K (2016) In silico insight into potential anti-Alzheimer's disease mechanisms of icariin. *Int J Mol Sci* 17(1). <https://doi.org/10.3390/ijms17010113>
 193. <http://www.cbligand.org/AD/>
 194. Liu H, Wang L, Lv M, Pei R, Li P, Pei Z, Wang Y, Su W, Xie XQ (2014) AlzPlatform: an Alzheimer's disease domain-specific chemogenomics knowledgebase for polypharmacology and target identification research. *J Chem Inf Model* 54(4):1050–1060
 195. www.CBLIgand.org/CBID
 196. <http://www.mysql.com>
 197. <http://www.apache.org/>
 198. http://openbabel.org/wiki/Main_Page
 199. www.cbligand.org/TargetHunter
 200. <http://www.cbligand.org/HTDocking/>
 201. <http://www.cbligand.org/BBB/>
 202. <http://rcidm.org/AlzhCPI>
 203. Fang J, Wang L, Li Y, Lian W, Pang X, Wang H, Yuan D, Wang Q, Liu AL, Du GH (2017) AlzhCPI: a knowledge base for predicting chemical-protein interactions towards Alzheimer's disease. *PLoS One* 12(5): e0178347
 204. Qiu T, Qiu J, Feng J, Wu D, Yang Y, Tang K, Cao Z, Zhu R (2017) The recent progress in proteochemometric modelling: focusing on target descriptors, cross-term descriptors and application scope. *Brief Bioinform* 18(1):125–136
 205. Barneh F, Jafari M, Mirzaie M (2016) Updates on drug-target network; facilitating polypharmacology and data integration by growth of DrugBank database. *Brief Bioinform* 17(6):1070–1080
 206. Chen C, He Y, Wu J, Zhou J (2015) Creation of a free, internet-accessible database: the multiple target ligand database. *J Cheminform* 7:14
 207. Alqahtani S (2017) In silico ADME-Tox modeling: progress and prospects. *Expert Opin Drug Metab Toxicol* 13(11):1147–1158



Computational Studies on Natural Products for the Development of Multi-target Drugs

Veronika Temml and Daniela Schuster

Abstract

Secondary plant metabolites represent “privileged structures” in drug development; they frequently interact with multiple protein targets within the body. For example, the anti-inflammatory natural product resveratrol from red wine has been shown to be active on over ten targets. Computational methods allow us to tackle the complexity of plant extracts, which often contain multiple active structures, which are in turn interacting with multiple targets. Virtual screening-based target fishing with pharmacophore modeling can help to identify protein targets, and docking simulations can be employed to propose a binding mechanism. Computational methods also play an important role in the analysis of plant extracts. Dereplication databases can be used to compare mass spectra of new extracts to a database of literature data to identify already known natural products. Activity networks of plant constituents help to understand the effect of extracts on specific pathologies and help to determine the active principles. We provide an overview, over the currently used computational methods in natural product research.

Keywords Activity networks, Dereplication, Molecular docking, Multi-target inhibitors, Natural products, Pharmacophore modeling, Polypharmacology, Virtual screening

1 Introduction

1.1 *Natural Products as Multi-target Leads*

While plant-based remedies constitute the roots of pharmaceutical sciences, drug development has largely shifted to explore and optimize synthetic lead structures. For a long time, the “one drug-one target” paradigm was dominant in pharmaceutical research, sidelining natural products, because of their unspecific modes of action. Over the last decade, however, this paradigm is crumbling more and more, as it turns out that even among successful synthetic drugs, many actually interact with several proteins aside from the intended target [1]. These off-target effects, long regarded mainly as causes for side effects, can actually improve a drug’s efficacy and pave the way for drug repurposing, where new applications for known drugs are found. One prominent example of this process is acetylsalicylic acid (ASS, aspirin), derived from the natural product salicylic acid. It is commonly used as an analgesic and antipyretic. The anti-inflammatory properties of ASS are attributed to its irreversible inhibition of cyclooxygenases 1 and 2, which blocks

prostaglandin formation [2]. Over the years, however, ASS has also been found to interact with multiple other targets, such as 20- α -hydroxysteroid dehydrogenase [3], 5'-AMP-activated protein kinase [4], and the endothelin-1 receptor [5].

The multi-target approach is also much discussed in cancer treatment, where multi-kinase oncolytic drugs aim to impede the development of treatment resistance in cancer cells [6].

With the reinvigorated interest in multi-target drugs, natural products are once more heading toward the center stage of drug development. Due to similarities in the genetic and biochemical makeup of plant and animal life, secondary plant metabolites often represent so-called privileged structures for interaction with animal proteins [7]. They frequently feature structural motifs and scaffolds that evolved as ligand-protein binding pharmacophores [8]. This also predisposes natural products for interaction with multiple targets, leading to the complex mechanisms of action that natural product researchers so often encounter.

A typical representative of compounds displaying substantial polypharmacological effects is resveratrol, the active constituent from red wine, which exhibits activity on nuclear factor κ B, cyclooxygenases, 5-lipoxygenase, silent information regulator 2, -AMP-activated protein kinase, and others [9]. It is likely the combination of all these effects that leads to the well-documented anti-inflammatory and antiaging properties of resveratrol. Curcumin is another example of a natural product that is hailed to have positive effect on a whole range of pathologies and has been shown to be effective in vitro on so many targets that it fills whole books [10].

Resveratrol and curcumin are particularly thoroughly investigated natural products, but there is a wealth of active natural products, whose mechanisms of action remain to be elucidated.

The 2015 Nobel Prize was awarded for the discovery of artemisinin, a potent natural antimalarial compound, whose mechanism of action is still under investigation and intense discussion. Recently it has been shown that activation by heme leads to a cornucopia of interactions with multiple proteins [11].

Libraries of natural product-like compounds are created to profit from their beneficial mechanism of action while accessing the optimization potential provided by synthetic chemistry [12]. Natural products are also a valuable source of inspiration for fragment-based drug discovery, as their ring systems and functional groups often display target affinity and are therefore particularly suited for fragment growth.

The search for multi-target drugs is a constant walk on a tight-rope, because while it is becoming clear that hitting multiple targets can be an advantage for a drug, it of course also remains true that interactions with other targets in the body can also cause side effects. Aside from creating unwanted effects, a drug's promiscuity might also hamper its efficacy [13]. Compounds that show significant

activity in an in-vitro setting may lose their activity in the organism, because the compound never reaches the intended targets.

Some combinations of targets are medicinally useful, while others do not really add any beneficial effect. The search for so-called master key drugs that bind to a number of targets that add a desirable effect and avoid off-targets causing adverse effects [14] is therefore quite challenging and unlikely to be successful by testing based on trial and error. Computational methods play an important role in rationalizing the available data and pointing the search into the right direction.

1.2 Natural Product Drug Discovery and Computer Aided Methods

In the development of synthetic drugs, the starting point is usually a lead structure discovered to have a specific activity in high-throughput screening. Natural product research by contrast typically starts with a plant that is in use for specific pathologies in often ancient local folk medicine. In the bio-guided fractionation approach, different extracts of a plant are tested for a specific biological activity, often not target specific but more general (e.g., anti-inflammatory). The active extracts are then further fractionated to enrich the active compounds [15]. The isolation of active natural products is at the end of this process, and sometimes this laborious process leads to the re-isolation of active constituents, which are already known from other plants. For new active natural products, the molecular targets often remain unknown.

Identifying and isolating individual compounds from these complex mixtures is a time- and material-consuming task. Computational methods that help to predict constituents with a high likelihood for activity or that can help to elucidate the molecular targets of natural products are therefore highly valuable in natural product research. While experimental tests are usually limited to a small set of targets that are available for biological tests, virtual screening allows us to consider a larger space of known targets to get a more complete picture of a compound's putative polypharmacology.

Several methods, such as molecular fingerprinting, pharmacophore modeling, and docking, are used to rationalize activities and to predict likely targets. Chemogenomics aims to comprehensively chart the associations of all possible ligands with all possible targets [16]. Dereplication can be used to compare mass spectrometry (MS) and NMR data from the extracts to databases of known natural products to focus the fractionation on novel structures [17]. Molecular networking takes this approach one step further and connects the structural data from MS and NMR with activity data from databases, to predict putative targets for the whole extract [18].

We aim to give an overview over recent applications of computational methods in natural product research with special regards to elucidating their polypharmacological profiles.

2 Computational Methods in Natural Product Drug Development

2.1 Pharmacophore Modeling

In the search for multi-target inhibitors, pharmacophores can be used in two fashions [19]. Either the activity on different targets is related to different parts of the molecule that fit a ligand binding motif of different targets. On the other hand, two similar binding sites can possess similar structural requirements that are mapped by a single molecule; this is likely if two protein targets have the same natural ligand (Fig. 1).

Collections of pharmacophore models, such as the Pharma DB (available in Discovery Studio (<http://accelrys.com/products/colaborative-science/biovia-discovery-studio/>), the intel:ligand pharmacophore DB (<http://www.inteligand.com/pharmdb/>) [20, 21], a collection of structure-based pharmacophore models available in the online tool PharmMapper [22, 23], or pharmacophore models created by individual research groups, can be used for parallel screening. This allows researchers to predict potential protein targets for their active compounds.

This approach was successfully applied in natural product research in several published cases: Rollinger et al. investigated 16 secondary metabolites isolated from *Ruta graveolens*, which were screened against 2208 pharmacophore models. For three targets, acetylcholinesterase (AChE), the human rhinovirus (HRV) coat protein, and the cannabinoid receptor type-2 (CB2), the virtual hits were biologically tested. The study revealed arborinine as a dual inhibitor of AChE and HRV coat protein [24].

Leoligin, the major lignan from *Leontopodium alpinum*, was predicted by a pharmacophore-based virtual screening to bind to the cholesteryl ester transfer protein (CETP). This activity was

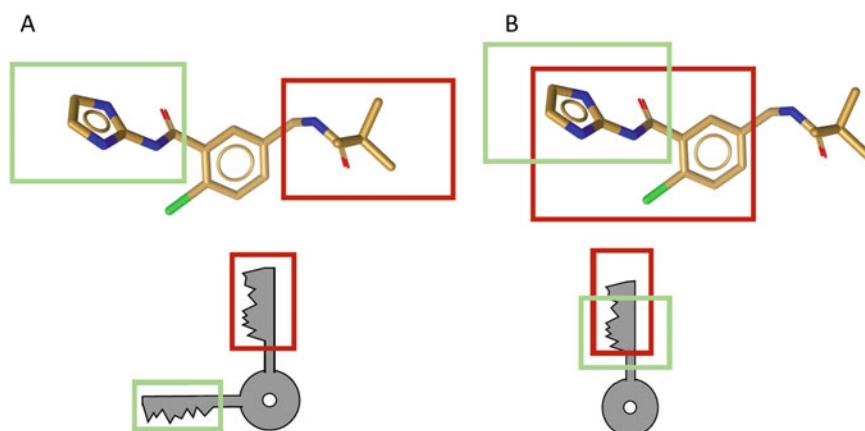


Fig. 1 (a) Two separate parts of the molecule bind to the two different targets (green and red); in the lock key analogy, this is represented by a key with different teeth, and (b) the binding motifs for two targets are the same or overlap represented by a key that fits two locks

experimentally confirmed [25]. A ligand-based pharmacophore model for 3-hydroxy-3-methyl-glutaryl-(HMG)-CoA reductase also predicted leoligin activity that was experimentally confirmed. Leoligin as a dual CETP and HMG-CoA reductase ligand reduces LDL cholesterol levels and improves LDL/HDL and LDL/total cholesterol ratios [26].

Pharmacophore modeling is also used to predict adverse effects for plants in medical use. For example, natural product human ether-a-go-go-related gene channel blockers were identified in a study by Kratz et al. [27].

2.2 Molecular Docking Simulations

Molecular docking is widely used in natural product research, predominantly in retrospective to elucidate the mechanism of action of active plant constituents. In the search for multi-target ligands, the interaction patterns observed in docking poses can be compared to each other to find common or complementary binding features.

The natural product embelin was found to be a potent inhibitor of 5-lipoxygenase (5-LO) and microsomal prostaglandin E₂ synthase-1 (mPGES-1). Binding modes for both enzymes were suggested in docking simulations [28]. Over the recent years, several dual inhibitors of 5-LO and mPGES-1 of natural origin have been discovered [29–31], suggesting that the binding sites of these two enzymes share several common features.

If the molecular targets are preselected, docking can also be used to predict potential dual (or multi-target) inhibitors. Park et al. used this method to predict dual inhibitors of wild-type stem cell factor receptor (c-KIT) and its most abundant gain-of-function mutant (D816V). Four dual inhibitors of natural origin were identified, which were active in the micromolar to submicromolar range [32]. This illustrates the value of computational strategies to anticipate mutations in target proteins.

Analogous to parallel screening, there are also docking tools available, which propose targets for natural products (this strategy is sometimes called “reverse molecular docking”) (*see* Table 1). The tool Selnergy™ docks compounds into more than 7000 targets and proposes the most promising hits [33]. The natural coumarin meranzin was screened against 400 proteins, and the 30 best-fitting targets were proposed. Out of these, three targets were selected for experimental validation: COX 1 and 2 and the peroxisome proliferator-activated receptor γ (PPAR γ). Meranzin was shown to bind to all three selected targets [34].

2.3 Similarity Search Methods

Several online tools allow users to search for targets based on 2D and 3D similarity measures, comparing a query molecule to known ligands. The SwissTargetPrediction webserver compares a query molecule to a library of 280,000 compounds active on more than 2000 targets of five different organisms [37].

Table 1
Pharmacophore modeling and docking identifying multi-target activities of natural products

Natural products	Targets	Method	Software	Reference
400 compounds from African medicinal plants	Multiple cancer targets	Pharmacophore modeling	LigandScout	[35]
Arborinine	AChE, HRV	Pharmacophore modeling	Discovery Studio/ LigandScout	[24]
Leoligin	CTP, HMG, COA reductase	Pharmacophore modeling	Discovery Studio/ LigandScout	[25, 26]
Embelin	5-LO, mPGES-1	Docking	Gold/Glide	[28]
Two compounds from the InterBioScreen DB	MMP-2, HDAC-6	Pharmacophore modeling/docking	Discovery Studio	[36]
Four compounds from the InterBioScreen DB	c-KIT and mutant	Docking	AutoDock	[32]
Meranzin	PPAR γ , COX1/2	Docking	Selnergy	[34]
Vitamin E metabolites	mPGES-1/5-LO	Docking	Gold	[30, 31]

The SPiDER target prediction software moves away from direct molecular comparison to a combination of pharmacophore and physicochemical descriptors, ideally enabling scaffold hopping and the discovery of novel chemical entities [38]. This is valuable in natural product research, since active natural products often differ in their physicochemical properties and structural motifs from synthetic drugs; therefore judging them solely by their similarity to them often leads to false-negative results (*see* part 3 on challenges).

This approach was further evolved to the chemically advanced template search (CATS) [39] and successfully applied to identify a set of pyrrolopyrazines as multi-target antimalarial agents [40].

In a recently published study by Schneider et al., this ligand-based target prediction method, based on self-organizing feature maps, was further refined with the Target Inference GENERator (TIGER) software, which assumes that similar pharmacophore patterns overlap for ligand binding in a specific binding site. Its scoring function therefore relies on the rank order of isofunctional ligands in the local neighborhood of the query compound. (\pm)-Marinopyrrole A was identified as a potent glucocorticoid, cholecystokinin, and orexin receptor antagonist with this method [41].

2.4 Machine Learning

In this approach different descriptors are calculated for the molecules and correlated, e.g., with their activity on a specific target. The more data so-called neural networks are fed, the more accurately they can predict activities for novel compounds.

PASS Online provides a tool predicting activities with a Bayesian neural network based on molecular fragment descriptors [42]. This approach was successfully used to predict oncolytic activity among

marine sponge alkaloids. St. John's wort is known to have an activating effect on cytochrome P450 (CYP), leading to often undesired interactions with other drugs. A set of 93 individual constituents of St. John's wort were screened for modulation of CYPs in PASS, and several virtual hits were experimentally confirmed to have a modulating effect [43].

2.5 Computational Binding Site Comparison

Computational binding site comparison is getting renewed attention as another valuable aid in the search for multi-target inhibitors, allowing researchers to identify proteins that share binding features and have a high probability of attracting similar ligands. There are several software tools available to compare binding sites in a geometry-based approach, e.g., CavBase [44], SiteEngine [45], and TrixP [46]. Lately also LigandScout (www.inteligand.com) implemented a binding site analysis tool. Alternatives are descriptor-based tools that assign fingerprint-like descriptors to the binding sites, such as FLAP [47] and FuzCav [48]. A comprehensive review on available methods is given by Ehrt et al. [49].

In a strategy published by Dekker et al., protein structure similarity clustering was used to select natural products as a starting point for potential multi-target leads [50]. In one study the enzymes Cdc25A phosphatase, AChE, 11 β HSD1, and 11 β HSD2 were clustered and investigated for structural resemblance. The natural sesterterpene dysidiolide was used as a base for derivatives active on the multiple targets [51].

2.6 Analysis of Polypharmacological Networks

The work with plant extracts confronts researchers with complex mixtures of many ingredients active on a wide network of targets. Systems (or network) pharmacology is an emerging concept that is especially interesting for natural product research, while polypharmacology leaves the one drug–one target principle to move on to a one drug-multiple target approach; systems pharmacology moves on to the ambitious catchphrase “one treatment-one network” aiming at target networks instead of isolated targets [52] (Fig. 2).

Several case studies in the last years used the connectivity map (CMap) data. This CMap relies on the CMap data resource, which is publicly available from the Broad Institute at MIT and Harvard, and provides the genome-wide gene expression profiles of over 5000 small molecules [53].

The underlying hypothesis is that cellular signatures representing systematic perturbation are connected with genetic (protein function) and pharmacologic (small molecules) perturbagens. Similar network signatures represent a connection (e.g., between a small molecule and its protein target). Researchers can record gene expression profiles for their samples and compare them to the database with online tools. CMap data was used in several instances to elucidate a mechanism of action for natural products. Pristimerin, a triterpenoid, isolated from *Celastrus* and *Maytenus*

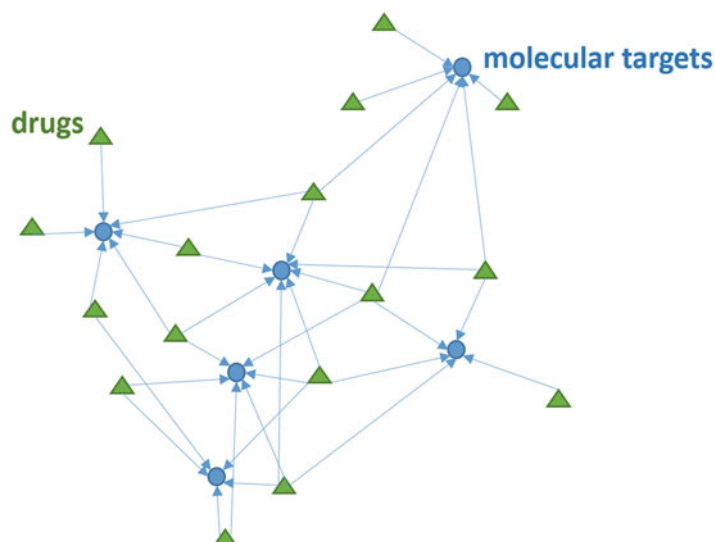


Fig. 2 Schematic drug-target connectivity map: the different drugs (green triangles), in case of NPs often multiple components of the extracts, are connected to their molecular targets (blue circles), forming an overview over likely activity patterns for a mixture

spp., which induces selective myeloma cell apoptosis, was used to treat cells. Their transcriptional response was compared to the CMap data, which lead the researchers to several molecular targets. Pristimerin was shown to inhibit IKK phosphorylation of I κ B and to act as a proteasome inhibitor [54].

Genistein, a major isoflavone from soybean, which is known to inhibit tumor cell growth [55] was analyzed with CMap. Kibble et al. extracted the transcriptional response profiles for genistein from the CMap database and compared them with the online tool MANTRA [56] to other small molecules. Known targets for molecules close to genistein on the connectivity map included COX 2, which was confirmed experimentally as a target for genistein and histone deacetylases (HDACs), which was confirmed by literature data [57].

2.7 De-replication

Tackling the problem of the chemical complexity of the biological matrices in which natural products and their metabolites are found, bioinformatics is working hand in hand with analytic profiling to predict the bioactivity potential of extracts. Computer-assisted dereplication strategies have been developed to avoid the problem of isolating already known structures over the course of bio-guided isolation. The spectral data of the extracts can be compared to databases to quickly identify known chemotypes [58].

The compound activity mapping platform offers the possibility to predict the identities and modes of actions of bioactive constituents of complex natural product mixtures directly from primary screening data [59].

Olivon et al. employed a massive multi-informational molecular network approach to select extracts from a Euphorbiaceae extract library and to predict active and non-active metabolites from *Bocquillonia nervosa* and *Neoguillauminia cleopatra* against the Wnt signaling pathway and chikungunya virus replication. All data acquired on chemical (LC-MS, UHPLC-HRMS), biological (assay data), and taxonomical (comparison with natural product databases) level was combined to form a multiinformative map, using an online workflow from Global Natural Products Social Molecular Networking (GNPS, <https://gnps.ucsd.edu>) [60].

These valuable data organizing tools aid researches in dealing with the immense complexity of plant extracts.

3 Notes

3.1 Physicochemical Differences Between NPs and Synthetic Drugs

As we saw in the previously described computational approaches currently used in natural product research, many of them rely on training the models with previously known active compounds with a high affinity for the target protein (*see* Subheadings 2.1–2.4). Most of the data found in, e.g., biological activity databases, such as the ChEMBL [61] mainly refers to synthetic molecules, simply because there is more activity data available for them. However this means that most models are trained predominantly with compounds that differ significantly in their physicochemical properties from natural products [62]. To name just a few general differences, natural products tend to have a higher molecular weight; they contain more saturated bonds and are therefore more flexible and contain also more stereocenters. In addition they have a lower nitrogen content and a higher oxygen content, and they favor aliphatic over aromatic rings. It has been shown that natural products cover different areas of the chemical space [63].

A computational filter just based on a simple concept like Lipinski's "rule of five" [64] would fail to predict a large portion of natural product drugs, because they do not share the properties of the average drug-like compound. But also more complex models can fail to correctly predict active natural products as hits, if they were trained exclusively on synthetic compounds. It is therefore key to consider natural products already in the model generation.

Awareness of the differences between natural products and synthetic drugs is also advised for similarity-based methods. If there are no natural products in the reference database, it is less likely that they will be returned as hits, depending on the search criteria. For example, an algorithm searching for similar ring

systems as observed in synthetic drugs will inevitably overlook complex ring system like they are found in many terpenoids. When applying machine learning or similarity-based methods to screen for natural products, it is therefore crucial to inspect suitability of the underlying dataset.

3.2 Natural Products and PAINS

Since they were first described in 2010 [65], the discussion about pan assay interference compounds (PAINS) has been ongoing in the drug development community. The concept of PAINS was originally conceived by observing the results of a synthetic high-throughput screening library and finding that there were about 10% of promiscuous compound classes that occurred as frequent hitters in the employed biological assays but turned out not to possess a dose-dependent activity in later experiments and could therefore not progress to the later stages of drug development. From this dataset, several structural motifs with a high likelihood of assay interference were extracted.

Especially for researchers in natural product-based drug development, the notion that several very prevalent structural motifs such as quinones or catechols should be excluded from target-based drug development was a hard blow. Of course scientists working in the field were quick to point out all the clinically active natural products containing these substructures and their immense pharmaceutical value.

In reaction to this controversy, Baell published a review, focusing on the consequences of PAINS for natural product research [66]. He stresses that the PAINS philosophy is especially relevant in hit to lead discovery. Usually compounds with a good activity in a target-based high-throughput screening are then progressed to tests in cell-based assays. For PAINS it is often observed that they lose their activity in the cell-based assay and they are discarded. Either they were falsely identified as actives, due to interference with the target-based assay, or they are acting on multiple targets and have difficulties reaching the intended target in cell-based environment. However, many natural drugs, which structurally fall in the PAINS category (such as e.g., curcumin, resveratrol, or thymoquinone), have been discovered by observing *in vivo* efficacy in relatively low doses, not in *in vitro* assays. This shows that the PAINS approach is only partly applicable to natural product drug discovery.

Recently a study showed that PAINS are for the most part not extraordinarily promiscuous and should not be disregarded prematurely [67].

There are computational filters available that point out PAINS compounds to the researcher (<http://www.cbligand.org/PAINS/>). While it would be premature to exclude a PAINS compound from further investigation, by default, it is extremely advisable to check

them for possible interference with the assay and to be aware of their potential to give false-positive readouts.

Efforts have also been made to define criteria for problematic natural products and to provide filter tools for them [68].

4 Future Outlook

Due to evolution nature offers an enormous wealth of compounds, many of which are multi-target ligands. Many structures and activities of natural products have yet to be elucidated. Additionally natural product fragments have a high potential as an inspiration for multi-target drug discovery and lead optimization [69]. With every new part in a multicomponent system, the complexity increases and becomes more difficult to grasp.

Computational methods help to organize, rationalize, and analyze the sometimes overwhelming amounts of data available and generated. Activity network maps allow researchers to gain a glimpse at the bigger picture behind complex mixtures of different multi-target active natural products.

Filter tools can alert us to potentially harmful structural motifs and preemptively point out potential assay-interfering compounds.

Target fishing methods can be employed to identify individual molecular targets of active natural products, and computational visualization and simulations allow us to investigate the ligand-target interaction mechanism that underlies the individual activities of a compound. This also offers a base for rational multi-target optimization.

Overall, computational methods can be a powerful assistant in meeting the challenges of leaving the one drug-one target paradigm behind and opening up for the possibilities and pitfalls of multi-target drug design.

References

1. Reddy A, Zhang S (2013) Polypharmacology: drug discovery for the future. *Expert Rev Clin Phar* 6(1). <https://doi.org/10.1586/ecp.12.74>
2. Vane J (1971) Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. *Nat New Biol* 231:232. <https://doi.org/10.1038/newbio231232a0>
3. Urmi D, Vincenzo C, Roland P et al (2007) A salicylic acid-based analogue discovered from virtual screening as a potent inhibitor of human 20 alpha-hydroxysteroid dehydrogenase. *Med Chem* 3(6):546–550. <https://doi.org/10.2174/157340607782360399>
4. Din F, Valanciute A, Houde V et al (2012) Aspirin inhibits mTOR signaling, activates AMP-activated protein kinase, and induces autophagy in colorectal cancer cells. *Gastroenterology* 142(7):1504–1515.e1503. <https://doi.org/10.1053/j.gastro.2012.02.050>
5. Talbodec A, Berkane N, Blandin V et al (2000) Aspirin and sodium salicylate inhibit endothelin ETA receptors by an allosteric type of mechanism. *Mol Pharmacol* 57(4):797–804. <https://doi.org/10.1124/mol.57.4.797>
6. Jeong W, Doroshow J, Kummar S et al (2013) US FDA approved oral kinase inhibitors for the treatment of malignancies. *Curr Probl Cancer*

- 37(3):110–144. <https://doi.org/10.1016/j.currprobcancer.2013.06.001>
7. Koch M, Waldmann H (2005) Natural product-derived compound libraries and protein structure similarity as guiding principles for the discovery of drug candidates. In: Kubinyi H (ed) *Chemogenomics in drug discovery*. Wiley, New York. <https://doi.org/10.1002/3527603948.ch14>
 8. Koch M, Schuffenhauer A, Scheck M et al (2005) Charting biologically relevant chemical space: a structural classification of natural products (SCONP). *Proc Natl Acad Sci U S A* 102(48):17272–17277. <https://doi.org/10.1073/pnas.0503647102>
 9. Kulkarni S, Cantó C (2015) The molecular targets of resveratrol. *Biochim Biophys Acta* 1852(6):1114–1123. <https://doi.org/10.1016/j.bbadis.2014.10.005>
 10. Aggarwal B, Surh Y, Shishodia S (eds) (2013) *The molecular targets and therapeutic uses of curcumin in health and disease*. Springer, Heidelberg
 11. Wang J, Zhang C, Chia W (2015) Haem-activated promiscuous targeting of artemisinin in *Plasmodium falciparum*. *Nat Commun* 6:10111. <https://doi.org/10.1038/ncomms10111>
 12. Pascolutti M, Quinn R (2014) Natural products as lead structures: chemical transformations to create lead-like libraries. *Drug Discov Today* 19(3):215–221. <https://doi.org/10.1016/j.drudis.2013.10.013>
 13. Hu Y, Bajorath J (2014) Monitoring drug promiscuity over time. *F1000Res* 3:218. <https://doi.org/10.12688/f1000research.5250.2>
 14. Méndez-Lucio O, Naveja J, Vite-Caritino H et al (2016) F.D. One drug for multiple targets: a computational perspective. *J Mex Chem Soc* 60:168–181
 15. Weller M (2012) A unifying review of bioassay-guided fractionation, effect-directed analysis and related techniques. *Sensors (Basel)* 12(7):9181
 16. Jacoby E (2011) Computational chemogenomics. *Wires Comput Mol Sci* 1(1):57–67. <https://doi.org/10.1002/wcms.11>
 17. Oettl S, Hubert J, Nuzillard J et al (2014) Dereplication of depsides from the lichen pseudovernia furfuracea by centrifugal partition chromatography combined to ¹³C nuclear magnetic resonance pattern recognition. *Anal Chim Acta* 846:60–67. <https://doi.org/10.1016/j.aca.2014.07.009>
 18. Allard P, Péresse T, Bisson J et al (2016) Integration of molecular networking and in-silico MS/MS fragmentation for natural products dereplication. *Anal Chem* 88(6):3317–3323. <https://doi.org/10.1021/acs.analchem.5b04804>
 19. Morphy R, Rankovic (2005) Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem* 48(21):6523–6543. <https://doi.org/10.1021/jm058225d>
 20. Steindl T, Schuster D, Laggner C et al (2006) Parallel screening: a novel concept in pharmacophore modeling and virtual screening. *J Chem Inf Model* 46(5):2146–2157. <https://doi.org/10.1021/ci6002043>
 21. Steindl T, Schuster D, Wolber G et al (2006) High-throughput structure-based pharmacophore modelling as a basis for successful parallel virtual screening. *J Comput Aided Mol Des* 20(12):703–715. <https://doi.org/10.1007/s10822-006-9066-y>
 22. Liu X, Ouyang S, Yu B et al (2010) PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach. *Nucleic Acids Res* 38(2):609–614. <https://doi.org/10.1093/nar/gkq300>
 23. Wang X, Shen Y, Wang S (2017) PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res* 45:356–360. <https://doi.org/10.1093/nar/gkx374>
 24. Rollinger J, Schuster D, Danzl B et al (2009) In silico target fishing for rationalized ligand discovery exemplified on constituents of *Ruta graveolens*. *Planta Med* 75(3):195–204. <https://doi.org/10.1055/s-0028-1088397>
 25. Duwensee K, Schwaiger S, Tancevski I et al (2011) Leoligin, the major lignan from edelweiss, activates cholesteryl ester transfer protein. *Atherosclerosis* 219(1):109–115. <https://doi.org/10.1016/j.atherosclerosis.2011.07.023>
 26. Scharinger B, Messner B, Türkcan A et al (2016) Leoligin, the major lignan from edelweiss, inhibits 3-hydroxy-3-methyl-glutaryl-CoA reductase and reduces cholesterol levels in ApoE^{-/-} mice. *J Mol Cell Cardiol* 99:35–46. <https://doi.org/10.1016/j.yjmcc.2016.08.003>
 27. Kratz J, Mair C, Oettl S et al (2016) hERG channel blocking ipecac alkaloids identified by combined in silico – in vitro screening. *Planta*

- Med 82(11):1009–1015. <https://doi.org/10.1055/s-0042-105572>
28. Schaible A, Traber H, Temml V et al (2013) Potent inhibition of human 5-lipoxygenase and microsomal prostaglandin E2 synthase-1 by the anti-carcinogenic and anti-inflammatory agent embelin. *Biochem Pharmacol* 86(4):476–486. <https://doi.org/10.1016/j.bcp.2013.04.015>
 29. Reker D, Perna A, Rodrigues T et al (2014) Revealing the macromolecular targets of complex natural products. *Nat Chem* 6:1072. <https://doi.org/10.1038/nchem.2095>
 30. Alsabil K, Suor-Cherer S, Koeberle A et al (2016) Semisynthetic and natural garcinoic acid isoforms as new mPGES-1 inhibitors. *Planta Med* 82(11):1110–1116. <https://doi.org/10.1055/s-0042-108739>
 31. Pein H, Helesbeux J-J, Garscha U et al (2018) Endogenous metabolites of vitamin E limit inflammation by targeting 5-lipoxygenase. *Nat Commun* (in revision)
 32. Park H, Lee S, Hong S (2016) Discovery of dual inhibitors for wild type and D816V mutant of c-KIT kinase through virtual and biochemical screening of natural products. *J Nat Prod* 79(2):293–299. <https://doi.org/10.1021/acs.jnatprod.5b00851>
 33. Quoc-Tuan D, Isabelle R, Patrice A et al (2005) Reverse pharmacognosy: application of selnergy, a new tool for lead discovery. The example of epsilon viniferin. *Curr Drug Discov Technol* 2(3):161–167. <https://doi.org/10.2174/1570163054866873>
 34. Do Q-T, Lamy C, Renimel I et al (2007) Reverse pharmacognosy: identifying biological properties for plants by means of their molecule constituents: application to meranzin. *Planta Med* 73(12):1235–1240. <https://doi.org/10.1055/s-2007-990216>
 35. Ntie-Kang F, Simoben C, Karaman B et al (2016) Pharmacophore modeling and in silico toxicity assessment of potential anticancer agents from African medicinal plants. *Drug Des Devel Ther* 10:2137–2154. <https://doi.org/10.2147/DDDT.S108118>
 36. Wang Y, Yang L, Hou J et al (2018) Hierarchical virtual screening of the dual MMP-2/HDAC-6 inhibitors from natural products based on pharmacophore models and molecular docking. *J Biomol Struct Dyn*:1–59. <https://doi.org/10.1080/07391102.2018.1434833>
 37. Gfeller D, Grosdidier A, Wirth M et al (2014) SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucleic Acids Res* 42(1):32–38. <https://doi.org/10.1093/nar/gku293>
 38. Reker D, Rodrigues T, Schneider P et al (2014) Identifying the macromolecular targets of de novo-designed chemical entities through self-organizing map consensus. *Proc Natl Acad Sci U S A* 111(11):4067–4072. <https://doi.org/10.1073/pnas.1320001111>
 39. Reutlinger M, Koch C, Reker D et al (2013) Chemically advanced template search (CATS) for scaffold-hopping and prospective target prediction for ‘orphan’ molecules. *Mol Inform* 32(2):133–138. <https://doi.org/10.1002/minf.201200141>
 40. Reker D, Seet M, Pillong M et al (2014) Deorphaning pyrrolopyrazines as potent multi-target antimalarial agents. *Angew Chem Int Ed Engl* 53(27):7079–7084. <https://doi.org/10.1002/anie.201311162>
 41. Schneider P, Schneider G (2017) De-orphaning the marine natural product (±)-marinopyrrole A by computational target prediction and biochemical validation. *Chem Commun* 53:2272–2274. <https://doi.org/10.1039/C6CC09693J>
 42. Lagunin A, Stepanchikova A, Filimonov D et al (2000) PASS: prediction of activity spectra for biologically active substances. *Bioinformatics* 16(8):747–748. <https://doi.org/10.1093/bioinformatics/16.8.747>
 43. Lagunin A, Filipov D, Poroikov V (2010) Multi-targeted natural products evaluation based on biological activity prediction with PASS. *Curr Pharm Des* 16(15):1703–1717. <https://doi.org/10.2174/138161210791164063>
 44. Schmitt S, Kuhn D, Klebe G (2002) A new method to detect related function among proteins independent of sequence and fold homology. *J Mol Biol* 323(2):387–406. [https://doi.org/10.1016/S0022-2836\(02\)00811-2](https://doi.org/10.1016/S0022-2836(02)00811-2)
 45. Shulman-Peleg A, Nussinov R, Wolfson H (2005) SiteEngines: recognition and comparison of binding sites and protein–protein interfaces. *Nucleic Acids Res* 33(2):337–341. <https://doi.org/10.1093/nar/gki482>
 46. von Behren M, Volkamer A, Henzler A et al (2013) Fast protein binding site comparison via an index-based screening technology. *J Chem Inf Model* 53(2):411–422. <https://doi.org/10.1021/ci300469h>
 47. Baroni M, Cruciani G, Sciabola S et al (2007) A common reference framework for analyzing/ comparing proteins and ligands. Fingerprints for ligands and proteins (FLAP): theory and application. *J Chem Inf Model* 47

- (2):279–294. <https://doi.org/10.1021/ci600253e>
48. Weill N, Rognan D (2010) Alignment-free ultra-high-throughput comparison of drug-gable protein–ligand binding sites. *J Chem Inf Model* 50(1):123–135. <https://doi.org/10.1021/ci900349y>
49. Ehrt C, Brinkjost T, Koch O (2016) Impact of binding site comparisons on medicinal chemistry and rational molecular design. *J Med Chem* 59(9):4121–4151. <https://doi.org/10.1021/acs.jmedchem.6b00078>
50. Dekker F, Koch M, Waldmann H (2005) Protein structure similarity clustering (PSSC) and natural product structure as inspiration sources for drug development and chemical genomics. *Curr Opin Chem Biol* 9(3):232–239. <https://doi.org/10.1016/j.cbpa.2005.03.003>
51. Koch M, Wittenberg L-O, Basu S et al (2004) Compound library development guided by protein structure similarity clustering and natural product structure. *Proc Natl Acad Sci U S A* 101(48):16721–16726. <https://doi.org/10.1073/pnas.0404719101>
52. Vicini P, van der Graaf P (2013) Systems pharmacology for drug discovery and development: paradigm shift or flash in the pan? *Clin Pharmacol Ther* 93(5):379–381. <https://doi.org/10.1038/clpt.2013.40>
53. Lamb J, Crawford E, Peck D et al (2006) The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313(5795):1929–1935. <https://doi.org/10.1126/science.1132939>
54. Tiedemann R, Schmidt J, Keats J et al (2009) Identification of a potent natural triterpenoid inhibitor of proteasome chymotrypsin-like activity and NF- κ B with antimyeloma activity in vitro and in vivo. *Blood* 113(17):4027–4037. <https://doi.org/10.1182/blood-2008-09-179796>
55. Banerjee S, Li Y, Wang Z et al (2008) Multi-targeted therapy of cancer by genistein. *Cancer Lett* 269(2):226–242. <https://doi.org/10.1016/j.canlet.2008.03.052>
56. Carrella D, Napolitano F, Rispoli R et al (2014) Mantra 2.0: an online collaborative resource for drug mode of action and repurposing by network analysis. *Bioinformatics* 30(12):1787–1788. <https://doi.org/10.1093/bioinformatics/btu058>
57. Kibble M, Saarinen N, Tang J et al (2015) Network pharmacology applications to map the unexplored target space and therapeutic potential of natural products. *Nat Prod Rep* 32(8):1249–1266. <https://doi.org/10.1039/C5NP00005J>
58. Hubert J, Nuzillard J-M, Renault J-H (2017) Dereplication strategies in natural product research: how many tools and methodologies behind the same concept? *Phytochem Rev* 16(1):55–95. <https://doi.org/10.1007/s11101-015-9448-7>
59. Kurita K, Glassey E, Linington R (2015) Integration of high-content screening and untargeted metabolomics for comprehensive functional annotation of natural product libraries. *Proc Natl Acad Sci U S A* 112(39):11999–12004. <https://doi.org/10.1073/pnas.1507743112>
60. Olivon F, Allard P-M, Koval A et al (2017) Bioactive natural products prioritization using massive multi-informational molecular networks. *ACS Chem Biol* 12(10):2644–2651. <https://doi.org/10.1021/acscchembio.7b00413>
61. Bento A, Gaulton A, Hersey A et al (2014) The ChEMBL bioactivity database: an update. *Nucleic Acids Res* 42(1):1083–1090. <https://doi.org/10.1093/nar/gkt1031>
62. Feher M, Schmidt J (2003) Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inf Comput Sci* 43(1):218–227. <https://doi.org/10.1021/ci0200467>
63. Rosén J, Gottfries J, Muresan S et al (2009) Novel chemical space exploration via natural products. *J Med Chem* 52(7):1953–1962. <https://doi.org/10.1021/jm801514w>
64. Lipinski C, Lombardo F, Dominy B et al (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development. *Adv Drug Deliv Rev* 46(1):3–26. [https://doi.org/10.1016/S0169-409X\(00\)00129-0](https://doi.org/10.1016/S0169-409X(00)00129-0)
65. Baell J, Holloway G (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem* 53(7):2719–2740. <https://doi.org/10.1021/jm901137j>
66. Baell J (2016) Feeling nature’s PAINS: natural products, natural product drugs, and pan assay interference compounds (PAINS). *J Nat Prod* 79(3):616–628. <https://doi.org/10.1021/acs.jnatprod.5b00947>
67. Jasial S, Hu Y, Bajorath J (2017) How frequently are pan-assay interference compounds active? Large-scale analysis of screening data

- reveals diverse activity profiles, low global hit frequency, and many consistently inactive compounds. *J Med Chem* 60(9):3879–3886. <https://doi.org/10.1021/acs.jmedchem.7b00154>
68. Bisson J, McAlpine J, Friesen J et al (2016) Can invalid bioactives undermine natural product-based drug discovery? *J Med Chem* 59(5):1671–1690. <https://doi.org/10.1021/acs.jmedchem.5b01009>
69. Rodrigues T, Reker D, Schneider P et al (2016) Counting on natural products for drug design. *Nat Chem* 8:531. <https://doi.org/10.1038/nchem.2479>



Computational Design of Multitarget Drugs Against Alzheimer's Disease

Sotirios Katsamakos and Dimitra Hadjipavlou-Litina

Abstract

In the present review, the authors provide the basic background about the molecular targets implicated in the pathogenesis of Alzheimer's disease. Furthermore, the authors review structure–activity relationships (SAR), 2D- and 3D-quantitative structure–activity relationships (QSAR), as well as other computational modeling studies performed on multitarget agents for Alzheimer's disease.

The information provided includes chemical structures of multitarget agents and/or of hybrids acting on several molecular target enzymes implicated in the Alzheimer's disease pathogenesis and information for the used computational techniques. This should be useful in the development of new multitarget drugs with clinical applicability in Alzheimer's disease.

Keywords Alzheimer's disease, Docking, Modeling, Multitarget agents, QSAR

1 Introduction

Human brain is a very complex organ and the ion channels endow the neurons to generate action potentials used to signal other neurons. Abnormalities in neuroregulation lead to deficit in attention and learning, and mood disorders. Alzheimer disease (AD) is a chronic neurodegenerative disorder that slowly destroys neurons leading to serious cognitive disability [1]. The number of AD patients will be more than 10.4 million in the America and there will be an increment in the 80+ segment of the society at an alarming rate. AD is the progressive neurodegenerative disease of aging, leading to senile dementia including progressive memory loss with difficulty in performing daily activities, lack of coordination, social withdrawal, vision problems, and poor judgment. Physicians' therapeutic recommendations include the use of major six classes of drugs: acetylcholinesterase inhibitors (AChEI), *N*-methyl-*D*-aspartate (NMDA) receptor antagonists, monoamine oxidase inhibitors (MAOI), antioxidants, metal chelators, and anti-inflammatory drugs [1].

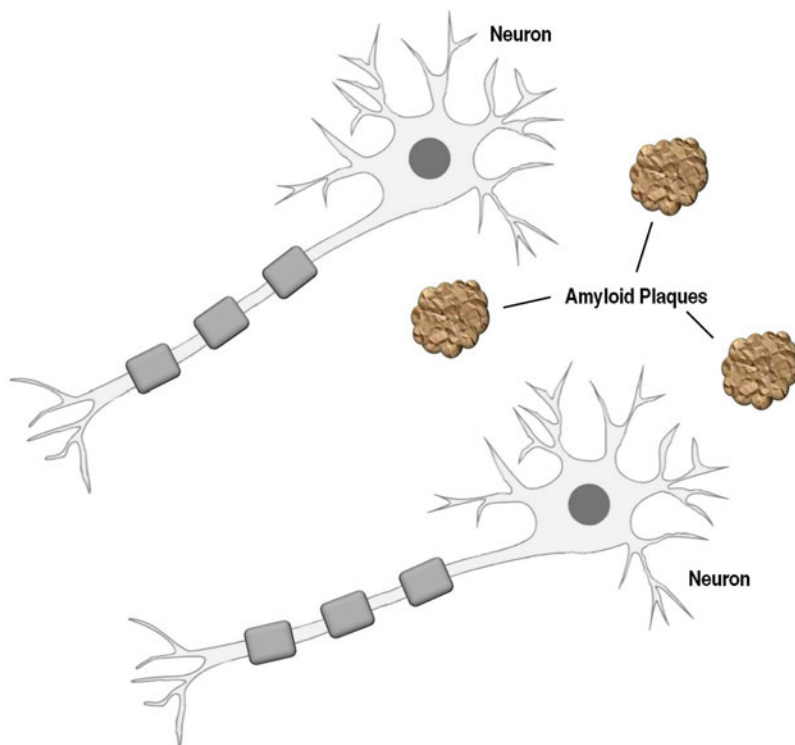


Fig. 1 The pathogenesis of the Alzheimer's disease

Several review papers and book chapters have been written describing the molecular biology and the pathogenesis of the AD (Fig. 1) [2–5]. Alzheimer's disease is a complex neurodegenerative disorder. So far, the therapeutic paradigm “one-compound-one-target” has failed to cure the disease. The multitarget-directed ligand (MTDL) approach has gained increasing attention by many research groups, for the rational design of new drug candidates, developing a variety of hybrid compounds acting simultaneously on diverse biological targets [6]. This review aims to show some recent advances and examples of the exploitation of the MTDL approach in the rational design of novel drug candidate prototypes for the treatment of AD and examples of “one-compound-one-target.”

2 Virtual Screening Approaches

The current scenario of drug discovery has undergone a drastic change, and the latest pharmaceutical research aims to develop new therapeutic entities characterized by selectivity and specificity with the implementations of *in silico* and “omics” technologies and 2D and 3D quantum and docking studies. Computational biology

approaches have proved to be reliable tools for the selection of novel targets and therapeutic ligands. Molecular docking is a key tool in drug design and development to propose structural hypotheses of how the ligands bind with the target for lead optimization.

It is a fact nowadays that virtual screening approaches offer a quicker, broader, and more efficient way to identify bioactive small molecules than the typical high throughput screening (HTS). Hit compounds' results have increased over 1% or more instead the typical $\leq 0.1\%$ rate of HTS [7]. Though being a complex effort, concerning the amount of data that ought to be analyzed each time. Currently, besides commercial libraries (i.e., Enamine, Mcule, ChemBridge, etc.), there are numerous freely obtained sources of compound libraries that include several millions of small molecules in different formats like SMILES, CDX, MOL2, SDF, and much more, as ZINC¹² [8, 9] and ZINC¹⁵ [10] databases, PubChem database [11], ChemSpider database [12], and the ChEMBL_23 database [13]. However, the duplication of compound occurrence within these databases cannot be excluded since they are individual initiatives. This exponential increase on data availability goes in parallel to the increasing structural data of proteins being published on the PDB [14]. Today, there are 136,594 structures deposited that can be utilized *in virtual* screening processes. Alternatively in the absence of structural data, the generation of a homology model is needed to follow this approach. Furthermore, several online tools exist for homology model production, presented by the Swiss Institute of Bioinformatics [15].

Generally, the computer-aided drug design pipeline consists of several steps as presented in Fig. 2.

As mentioned above, the initial library volume may vary to several millions. Hence, the most important step in the process is their 2D filtering based on desired properties. During this step, several filters may be applied as: (1) physicochemical properties [16, 17], (2) PAINS [18], and (3) structural features [19].

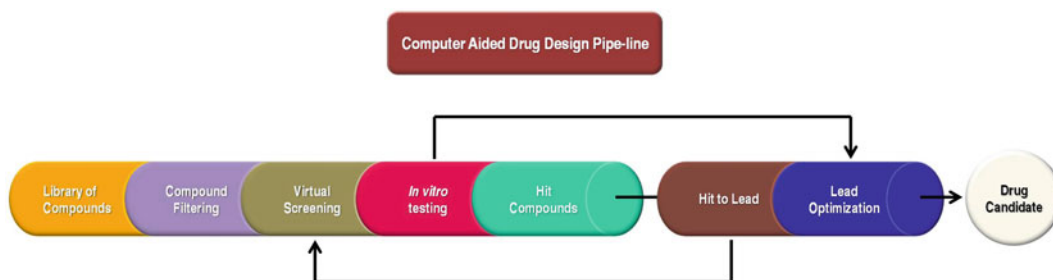


Fig. 2 The computer-aided drug design pipeline

3 Molecular Known Targets for the Treatment of AD and Active “One-Compound-One-Target”/Multitarget Agents Against AD

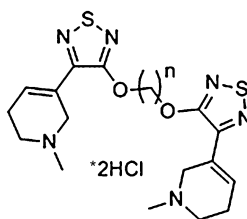
This chapter summarizes the known drug targets of AD, in vivo active agents against AD, state-of-the-art docking studies done in AD, and docking studies of multitarget agents with particular emphasis on AD.

3.1 Acetylcholine Receptors

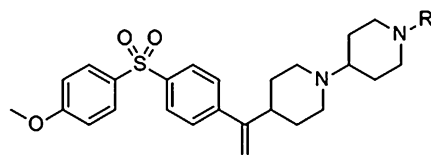
One of the consistent findings in brains of AD patients is the declining level of acetylcholine (ACh), the muscarinic endonative agonist-mediated cognitive effects by stimulating the postsynaptic muscarinic M1 receptors in central nervous system. Since muscarinic M2 subtype receptor antagonists could increase ACh release and improve learning ability through its blockade of the presynaptic M2 autoreceptors in the brain, the development of muscarinic M2 subtype receptor antagonists becomes one of the cholinergic approaches to AD by improving ACh level [20].

In 2000, Messer et al. [20] continuing their efforts on the development of muscarinic M1 receptor inhibitors designed bivalent analogues (Scheme 1) of xanomeline. The authors utilized molecular modeling to design the synthesized compounds over a homology model of the muscarinic M1 receptor as it was previously described originally from Nordvall and Hacksell [21] and replicated later by Messer et al. [22]. The biological study of the bivalent xanomeline derivatives indicated a much higher affinity than the parent compound, towards both M1 and M2 receptors. Although the mode of activity for these compounds has not been clearly elucidated, the authors speculate a dual receptor activity over a kind of a formed dimmer and/or a nearby region of the active site.

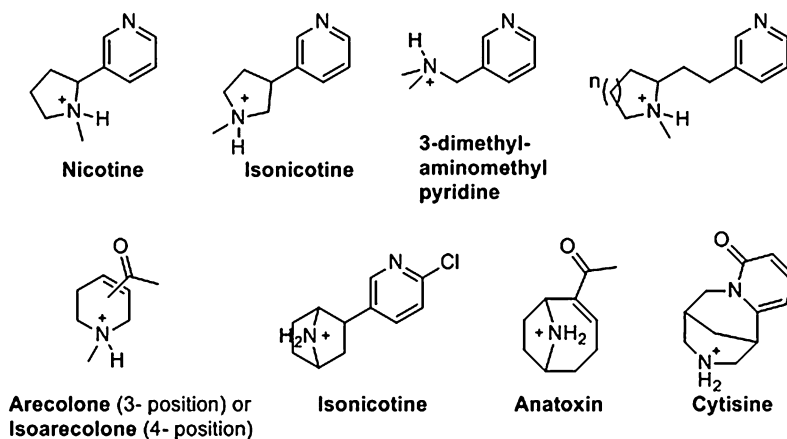
Niu and coworkers [23] used a 3D mathematic analytical method resembling a comparative molecular field analysis (CoMFA) [24] in combination with quantitative structure–selectivity relationship (QSSR) on both muscarinic subtypes (M2/M1), in order to design M2 receptor antagonists. Simultaneously, they worked on a series of piperidinyl piperidine derivatives known as muscarinic M2 antagonists (Scheme 2) synthesized by Wang et al. [25]. All the produced models showed high predictive abilities and statistical significance.



Scheme 1 Synthesized bivalent xanomeline derivatives



Scheme 2 The general structure of used piperidyl piperidine derivatives



Scheme 3 General lead structures of compounds included in the study

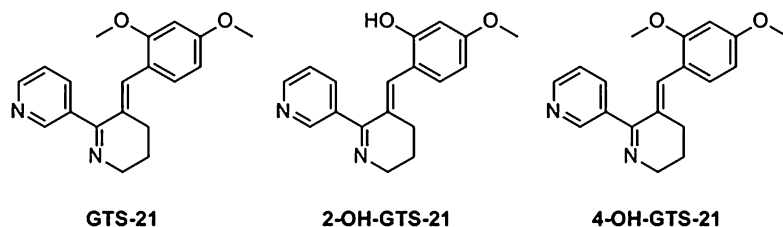
3.2 Neuronal Nicotinic Acetylcholine Receptor

Ongoing clinical trials and preclinical models in nonhuman primates have validated the neuronal nicotinic acetylcholine receptor (nAChRs) as promising drug targets for various CNS pathological conditions including Parkinson's disease, L-dopa-induced dyskinesia, Alzheimer's disease, addiction, ADHD, schizophrenia, and major depressive disorders.

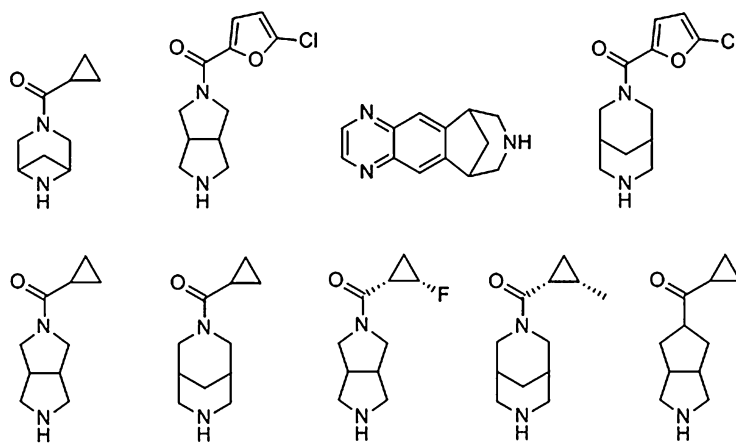
nAChRs are transmembrane ligand-gated ion-channel proteins resulting from the pentameric assembly of 16 human subunit gene products. The $\alpha 4\beta 2$ nAChRs are believed to represent one of the most abundant isoforms in mammalian brain. Several isoforms containing different stoichiometry of $\alpha 4$ and $\beta 2$ (the high- and low-sensitivity receptors) in addition to subtypes containing the $\alpha 5$, or the $\alpha 6$ and $\beta 3$ have expanded the diversity of the $\alpha 4\beta 2$ targets.

In 2002, Nicolotti et al. [26] reviewed in terms of quantitative structure–activity relationship (QSAR) and CoMFA, an extensive and diverse collection of 270 compounds on nAChR (Scheme 3).

Wei and coworkers [27] produced a homology model of $\alpha 7$ nAChR based on the AChBP structure 1I9B [28], and the ligand binding site was considered to be between the two subunits of the protein. The docked compounds were GTS-21 and derivatives (2-OH-GTS21 and 4-OH-GTS-21, Scheme 4). The results of



Scheme 4 The chemical structure of GTS-21 and its derivatives



Scheme 5 Diazabicyclic amide analogues

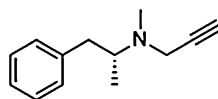
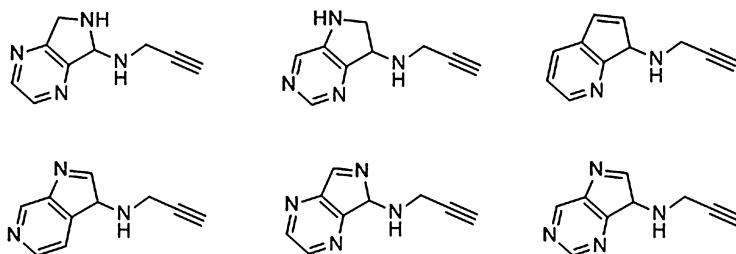
their study were in accordance with the previously reported experimental data.

In 2013, Combo et al. [29] performed a docking analysis on several diazabicyclic amide analogues (Scheme 5) on $\alpha 4\beta 2$ nAChR, using a produced homology model based on *Aplysia* AChBP with the entry 2BYQ [30]. The researchers through their *in silico* study highlighted an alternative binding mode for the molecules, by the formation of residue interactions from the amine group.

3.3 Monoamine Oxidases

Monoamine oxidase (MAO) is a flavin adenine dinucleotide (FAD)-containing enzyme that catalyzes the degradation of biogenic and xenobiotic amines. Two isoforms, namely MAO-A and MAO-B, have been characterized by their amino acid sequence, tissue distribution, substrate specificity, and inhibitor sensitivity. These isoforms present structural differences in the binding sites as revealed by high-resolution X-ray structures.

MAO-A, which preferentially degrades serotonin, adrenaline, and noradrenaline, is irreversibly inhibited by clorgyline. Monoamine oxidase B (MAO-B) constitutes one of the key targets for the development of new neuroprotective agents in both anti-AD and anti-PD chemotherapies. MAO-B is responsible for the oxidative deamination

**Selegiline****Fig. 3** Structure of selegiline**Scheme 6** Chemical structures of the proposed rasagiline bioisosteres

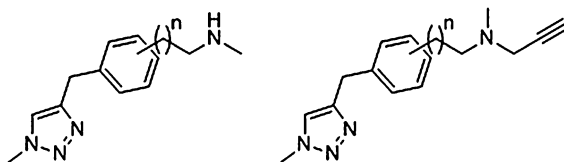
of phenylethylamine and benzylamine. It is irreversibly inhibited by (R)-deprenyl (selegiline) (Fig. 3).

In particular, a key structural feature is the replacement of the pair Phe208/Ile335 in MAO-A by Ile199/Tyr326 in MAO-B, leading to the distinction between “substrate” and “entrance” sites in MAO-B [31]. The replacement of Ile180/Asn181 and Val210 in MAO-A by Leu17/Cys172 and Thr201 in MAO-B are additional differences in the binding sites, which may modulate the selective inhibition by certain MAO inhibitors.

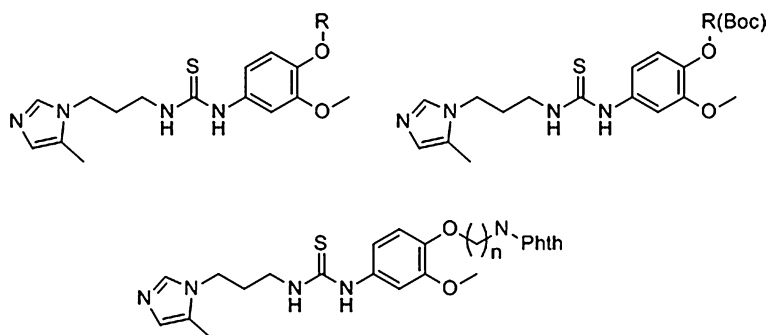
The neuroprotection exerted by MAO inhibitors is correlated also with the prevention of the formation of neurotoxic species, which may lead to neuronal damage, and from the antiapoptotic properties of the propargylamine group present in some MAO inhibitors.

In 2012, Speck-Planche and Kleandrova [32] analyzed MAO-B inhibitors with diverse structures (i.e., coumarins, chromones, and flavones derivatives), in terms of 2D QSAR and artificial neuronal networks (ANNs). Based on the precursor compound rasagiline, they suggested several new bioisosteres (Scheme 6).

Di Petro et al. [33] reported their research on three generations of triazole-based compounds (Scheme 7) tested as dual MAO-A and MAO-B inhibitors. The *in vitro* results varied from high- to low-micromolar range, IC_{50} 0.5–553.5 μM for human MAO-A and 0.6–373.9 μM for human MAO-B. The selectivity index of the compounds was in favor of MAO-B and therefore the researchers docked this series over two MAO-B crystal structures complexed with inhibitor ASS234 (PDB entry 4CRT) [34] and deprenyl (PDB entry 2BYB) [35], respectively. Tested compounds exhibited good blood–brain barrier permeation.



Scheme 7 Designed triazole derivatives



Scheme 8 Glutaminyl cyclase inhibitors

3.4 Glutaminyl Cyclase

Glutaminyl cyclase (GC) has been implicated in the formation of toxic amyloid plaques and thus may participate in the pathogenesis of Alzheimer's disease. An *in vitro* structure–activity relationship (SAR) study [36] identified several excellent QC inhibitors demonstrating 5- to 40-fold increases in potency compared to the reference inhibitor. Potent A β -lowering effect was achieved by incorporating an additional binding region into a previously established pharmacophoric model [36], resulting in strong interactions with the carboxylate group of Glu327 in the GC binding site.

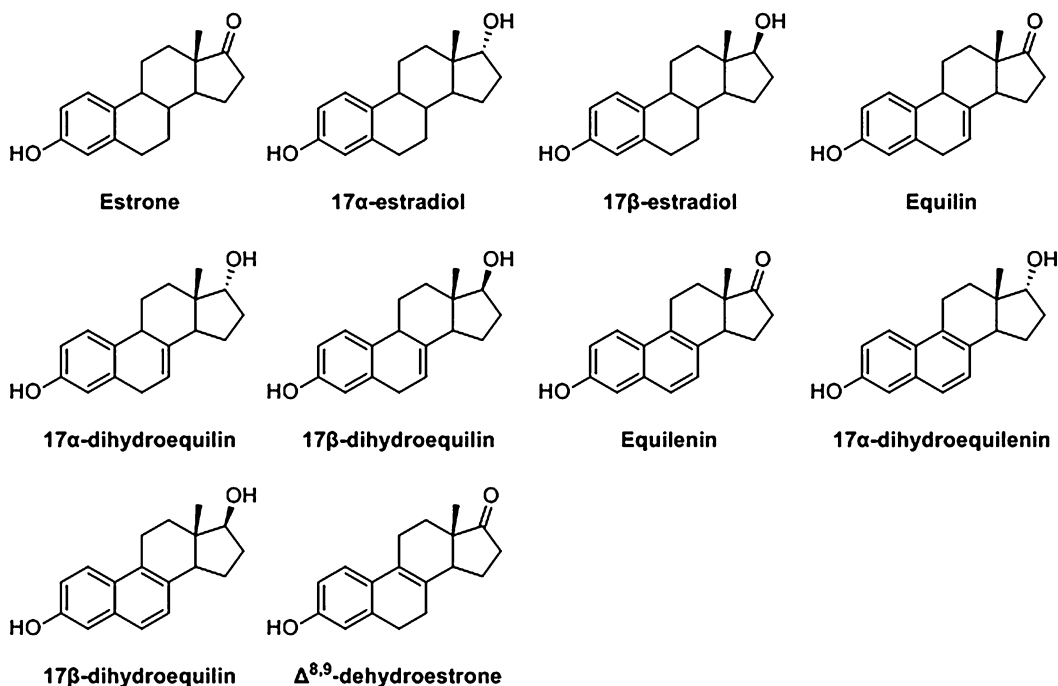
Hoang and coworkers [36] designed a new series of thiourea compounds (Scheme 8) that potentially act as multifunctional agents against AD by targeting glutaminyl cyclase and the formation of A β fibrils. Their design strategy included molecular modeling over the known crystal structure of human GC with the ascending number 3PBB [37] over the protein databank. All of the compounds exhibited activity in the low- to sub-nanomolar range with values between 0.7 and 41.2 nM. The most active compound presented optimum activity both *in vitro* and *in vivo*, while significantly inhibited aggregate formation. This activity was in alignment with docking results of the derivatives for the GC active site and its interacting regions A-, B-, C-, and D-, respectively.

3.5 Estrogen Receptor

Computer-aided modeling analyses were used to investigate the potential correlation of the molecular mechanisms that conferred estrogen neuroprotection with estrogen interactions with the

estrogen receptor (ER). Cultured basal forebrain neurons were exposed to either β -amyloid or excitotoxic glutamate with or without pretreatment with estrogens followed by neuroprotection analyses. Three indicators of neuroprotection were used to assess neuroprotective efficacy, LDH release, intracellular ATP level, and MTT formazan formation. Results of these analyses indicate that the estrogens, 17α -estradiol, 17β -estradiol, equilin, 17α -dihydroequilin, equilenin, 17α -dihydroequilenin, 17β -dihydroequilenin, and $\Delta^{8,9}$ -dehydroestrone were each significantly neuroprotective in reducing neuronal plasma membrane damage induced by glutamate excitotoxicity. 17β -Estradiol and $\Delta^{8,9}$ -dehydroestrone were effective in protecting neurons against β -amyloid₂₅₋₃₅-induced intracellular ATP decline. Co-administration of two from 17β -estradiol, equilin and $\Delta^{8,9}$ -dehydroestrone, exerted greater neuroprotective efficacy than individual estrogens.

Zhao et al. [38] studied the potential neuroprotective activity of estrogen compounds (Scheme 9). Selected estrogens have been docked by the researchers over the estrogen receptor subunit α (ER α /PDB entry 1ERE [39]) and were found to be directly correlated with their overall biological effects.



Scheme 9 Estrogen compounds studied for their neuroprotective activity

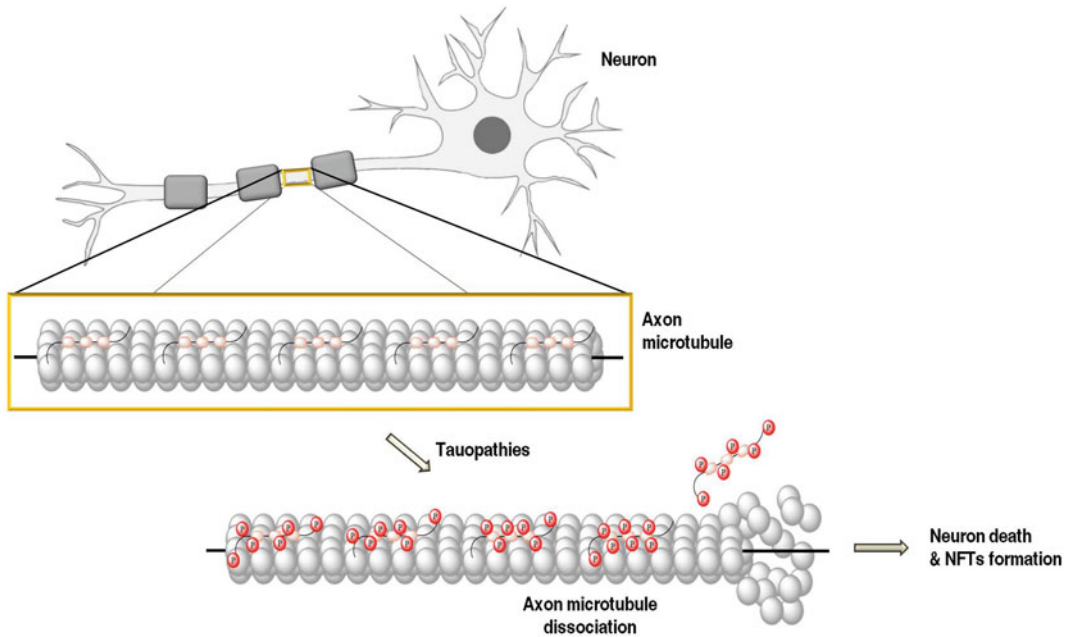
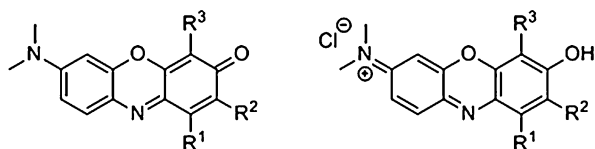


Fig. 4 The hyperphosphorylation of Tau and the formation of NFT

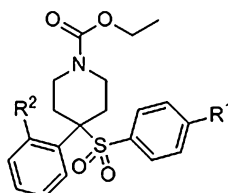
3.6 DKK1–LRP6 Complex Formation

Wnt signaling is found to contribute to AD pathophysiology when deregulated, while activation of the pathway is neuroprotective [40]. According to Hanger and Noble, the kinase activity of GSK3b is inhibited following the activation of Wnt signaling, implicating this pathway in the hyperphosphorylation of Tau and the formation of NFTs (Fig. 4) [41]. Wnt signaling involves the interaction of Frizzled (FRZ) and LRP5/6 transducing to GSK3b with a subsequent stabilization of β -catenin, which in turn enters the nucleus and forms a complex with TCF/LEFs to activate transcription of target genes [42]. The canonical Wnt pathway is negatively modulated by the extracellular protein DKK1, which binds to LRPs inhibiting Wnt by preventing their interaction with FRZ receptor [43]. DKK1 is a chemokine overexpressed in brains of AD patients and in AD transgenic mouse models [44, 45]. Hence, disruption of the LRPs and DKK1 complex formation could provide a way to regulate Wnt/ β -catenin signaling, offering an attractive therapeutic approach for treating AD. In 2010, LRP5–DKK1 interaction has been known to be targeted by monoclonal antibodies [46].

The last 2 years, Sarli’s group [47, 48] designed and synthesized NCI8642 derivatives (Scheme 10) targeting the interruption of DKK1/LRP6 complex formation as an alternative approach to treat AD through the Wnt activated pathway. The docking experiments were performed on the DKK1–LRP6 interface of the available crystal structure of the complex, entry 3S8V [49]. All the



Scheme 10 General structures of synthesized compounds



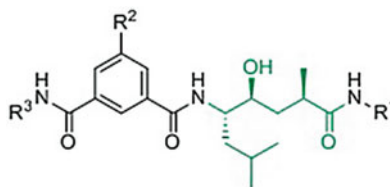
Scheme 11 Overview of the piperidine-based structure of synthesized derivatives

derivatives were tested on their ability to inhibit the complex formation with IC₅₀ at the low-micromolar range, similarly or better to the parent compound NCI8642. Moreover, the compounds exhibited increased β -catenin expression and reduction of PGJ2 induced tau phosphorylation indicating that Wnt canonical pathway activation is occurring.

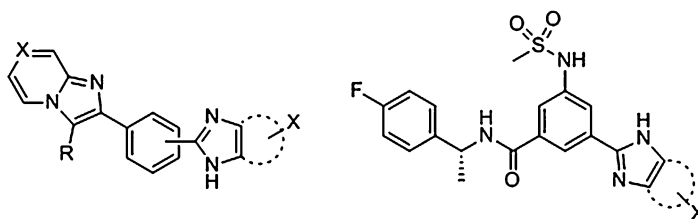
3.7 Secretases

Amyloid plaques represent one of the central pathological hallmarks of AD. Plaques mainly consist of A β peptides which are formed from the proteolysis of APP [50]. APP is located on the cell membrane as a single pass protein with the N-terminus located extracellularly, while its C-terminus is located in the cytoplasm. Several isoforms of APP exist ranging from 695 to 770 amino acids long. In neurons, the most common isoform is the 695 amino acid APP [51]. Cleavage of APP is performed by secretases α -, β -, and γ -. Among them, α -secretases do not promote the formation of A β , while the other two secretases (β - and γ -) generate amyloidogenic peptides [52]. β -Secretase cleaving enzyme 1 (BACE1) cleaves APP at a unique site, while γ -secretase cleaves APP at the C-terminus site at positions 40 and 42 leading to the formation of A β ₁₋₄₀ and A β ₁₋₄₂, respectively [53].

Secretase inhibition was a research target from the early twenty-first century. Thus, Gundersen and coworkers [54] developed a series of multitargeting compounds over γ -secretase inhibition, with the use of 3D-mapping simulations over previously reported inhibitors, combining also prevention of A β aggregate formation. Among the 3D shape similarity search applied, they identified the molecule that is illustrated in Scheme 11. The reported activity data suggest that their approach was successful but with no significant increase in γ -secretase inhibition nor A β aggregate formation.



Scheme 12 General structure of studied derivatives with the HE framework highlighted in green color



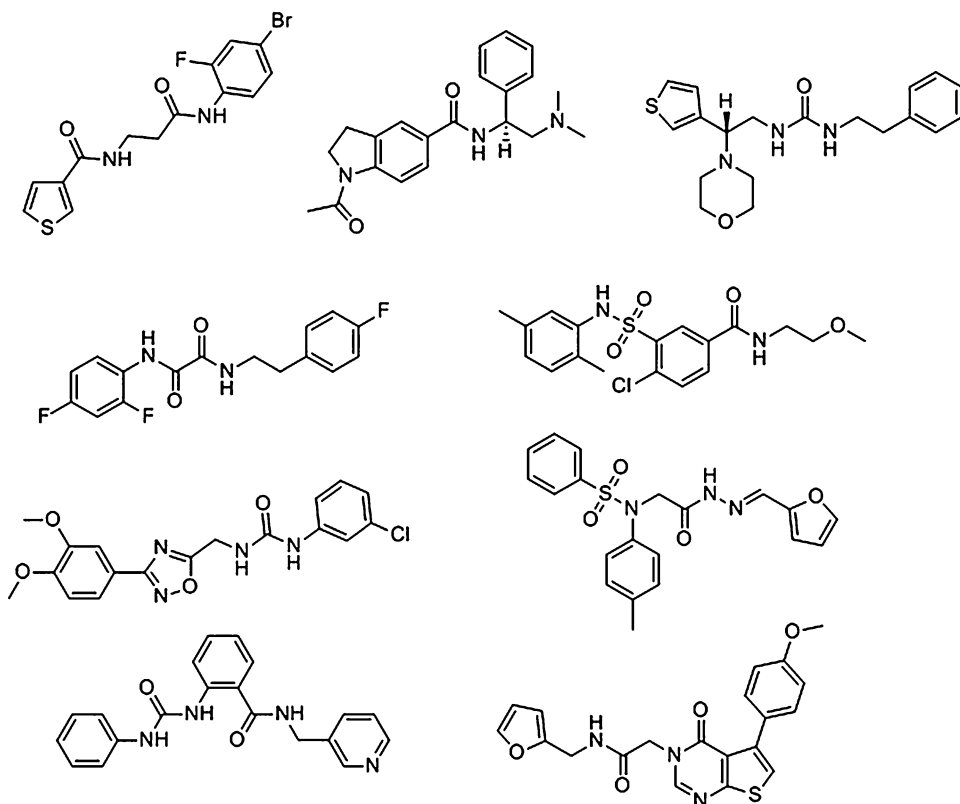
Scheme 13 General structures of imidazopyridine-based compounds

In 2009, Zhu et al. [55] synthesized a series of compounds bearing a hydroxyethylene framework (Scheme 12) as β -secretase (BACE-1) inhibitors. In their design, they incorporated docking experiments using the crystal structure 1TQF [56] that was available through the PDB and was co-crystallized with a non-peptide inhibitor. The most active compounds were found fivefold more active than previous inhibitors [57] with IC_{50} of 0.010 and 0.031 $\mu\text{mol/L}$, respectively.

Al-Tel and coworkers [58], in 2010, identified two highly selective classes of imidazopyridine-based inhibitors (Scheme 13) for β -secretase (BACE-1 and BACE-2). In their design, they utilized the available X-ray co-crystallized structure of BACE-1 protein with a small molecule having the entry code 2B8L [59]. The most active compound displayed an IC_{50} of 18 nM inhibition of BACE-1 with high affinity ($KI = 17$ nM), followed by an EC_{50} of 37 nM. Overall, the ligand efficiency value for the most active compound was 1.7 kJ/mol and it was found 204-fold more selective for BACE-1 compared to BACE-2.

Recently, Ajmani and coworkers [60] reported a QSAR study of approximately 233 known γ -secretase inhibitors, drawn from the ChEMBL database, in order to fill gap of the absence on a crystal structure for the protein. They used a combination of partial least squares (PLS) regression and neural networks (NN) to obtain key descriptors responsible for secretase inhibition.

In the work of Semighini et al. [61], several *in silico* techniques such as virtual screening, ADME and ADMET predictions, and similarity search were incorporated in order to be designed novel BACE-1 inhibitors. Pharmacophore-based approaches were created

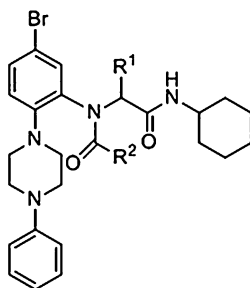
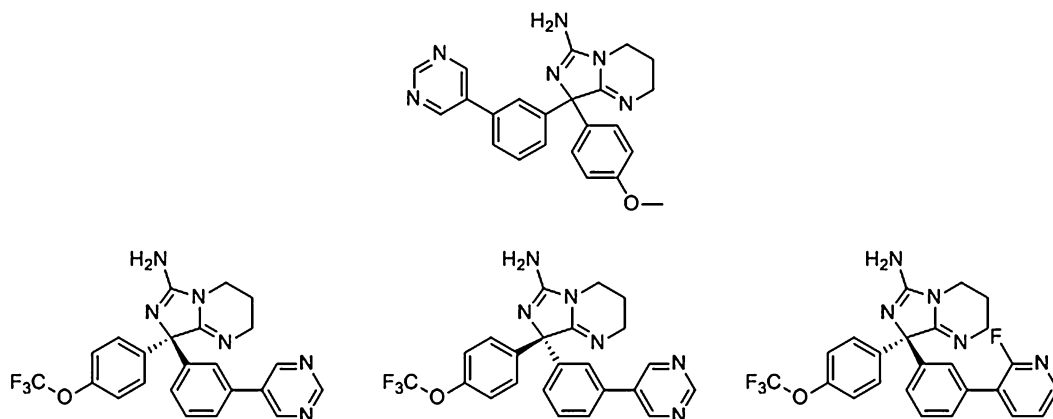


Scheme 14 Structures of the identified compounds

by the combination of crystal structures, with entries 2VJ7 [62], 2VNM [63], 2WF1 [64], 2WF2 [64], 2WF3 [64], 2WF0 [65], 2VNN [13], 2WEZ [65], and 2VIJ [66]. The physicochemical profiles for all the compounds were performed. The results led to ten promising candidate compounds (Scheme 14), three of which were on the low- μM range.

Edraki et al. [67] in 2015 have developed a novel series of piperazine-based derivatives (Scheme 15) as β -secretase (BACE-1) inhibitors. The molecular docking was performed in the BACE-1 active site of the 2ZJM crystal structure available on PDB. The most active compound showed superior protein inhibition at 10 and 40 μM , presumably through the molecule interactions with residues Asp228 and Thr72 within the active site.

In 2016, Zheng and Wu [68] reviewed all recent progress of CADD techniques regarding β - and γ -secretase inhibitors development. In the same year, Hernandez-Rodriguez et al. [69] studied extensively the binding mode of four known selective BACE-1 inhibitors (Scheme 16) against BACE-1, BACE-2, and Cathepsin D due to their high homology. The used crystal structures were entry 2QP8 [70] for BACE1, entry 2EWY [71] for BACE2, and entry 1LYW [72] for CTSD, respectively.

**Scheme 15** Piperazine-based derivatives**Scheme 16** Studied α -secretase inhibitors

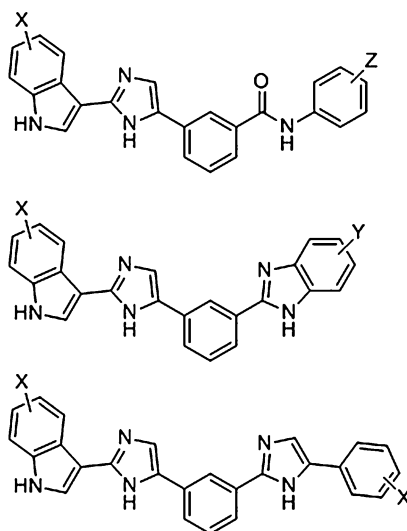
Tarazi et al. [73] identified some potent isophthalic acid derivatives (Scheme 17) as potent β -secretase inhibitors. Crystal structures employed in docking experiments were withdrawn from the PDB under the entry codes 1TQF [56], 2B8L [59], 2QZL [74], and 3EXO [75] for human β -secretase. Most active compounds exhibited a low-nanomolar range for BACE-1. Moreover, molecular descriptor values for the compounds were calculated in order to extract SARs.

3.8 $A\beta$ Fibrils

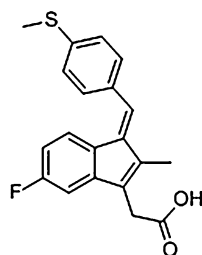
Although the mechanism of small molecules interaction with $A\beta$ fibrils still remains elusive, Prade et al. [76] in 2015 used PDB entries 2LMN [77] and 2LMP [77]. Solid-state NMR-obtained structures of $A\beta_{1-40}$ were used to study the binding mode of sulindac sulfide (Scheme 18), suggesting that the tested compound seems to bind in a specific way.

3.9 Cholinesterases

Acetylcholinesterase (AChE, E.C. 3.1.1.7) predominates in the healthy brain, playing a key role in cholinergic neurotransmission within the autonomic and somatic nervous system. It is a serine protease that hydrolyses the carboxylic ester of neurotransmitter



Scheme 17 General structures of synthesized compounds



Scheme 18 Chemical structure of sulindac sulfide

acetylcholine (ACh), expressed in nervous tissue, neuromuscular junctions, plasma, and red blood cells [78–80]. The active site of the enzyme is located in the bottom of a deep narrow gorge characterized by several subsites: (1) the anionic site, where the interaction with ACh occurs, (2) the esteratic site (ES), that contains three residues of the catalytic triad, (3) the oxyanion hole, (4) the acyl pocket, which confers substrate selectivity, and (5) the peripheral anionic site (PAS) which is located approximately 15°Å from the CAS [81, 82].

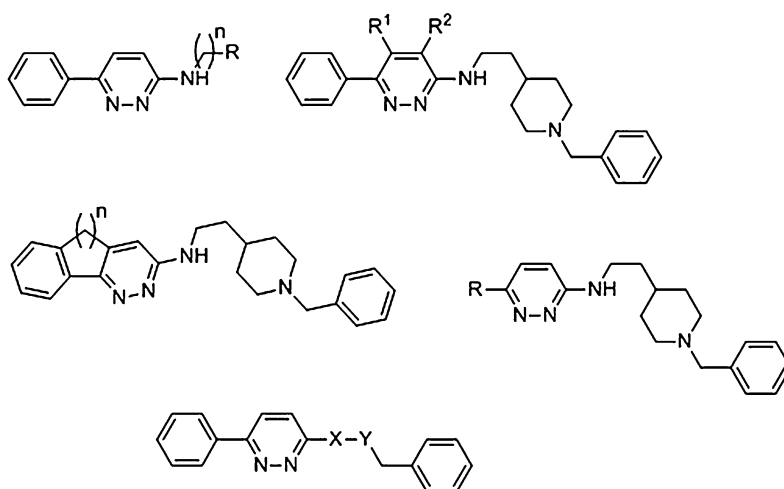
The appearance of AD is accompanied by extensive degenerative changes, resulting in ACh depletion and cholinergic hypofunction, which contribute to the progressive memory deficit and cognitive decline. Thus, AChE inhibitors represent a rational approach in AD pharmacotherapy restoring the synaptic levels of ACh and thus treating the symptoms caused by cholinergic imbalance.

In addition, AChE presents secondary noncholinergic functions attributed to the PAS active site, co-localizing with A β deposits, and directly promotes A β assembly and aggregation into insoluble plaques [83]. Butyrylcholinesterase (BChE) is considered to play a minor role in the regulation of synaptic ACh levels.

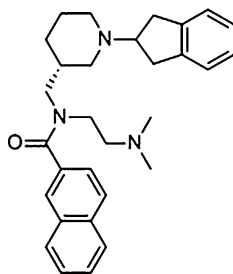
These two enzymes differ in substrate specificity, kinetics, and activity in different brain regions. BChE inhibition may also be considered a valid approach to restore cholinergic function in AD [83, 84]. However, its role in the regulation of cholinergic transmission in humans is not yet fully understood, and AChE remains the main target within this hypothesis.

In 1997, Recanatini et al. [85] performed an extensive QSAR analyses on 13 different series of derivatives in order to identify important physicochemical properties governing their inhibitory activity against acetylcholinesterase. Among the analyzed classes, physostigmine analogues, 1,2,3,4-tetrahydroacridines, and benzylamines were included. The results pointed that hydrophobicity, electronic effects, and steric factors were equally implicated.

Few years later, Sippl et al. [86] described the construction, validation, and application of a 3D QSAR model in combination with docking for the development of novel AChE inhibitors. The PDB structures used were 1ACL co-crystallized with decamethonium [87], 2ACK with edrophonium [88], 1ACJ with tacrine [87], and 1VOT with huperzine A [89]. Generally, they were able to produce a very good predictive model, since a number of designed inhibitors (Scheme 19) were synthesized and tested for their biological activity and the results were validated. The compounds seemed to interact with the cation-p subsite and the peripheral site in AChE.

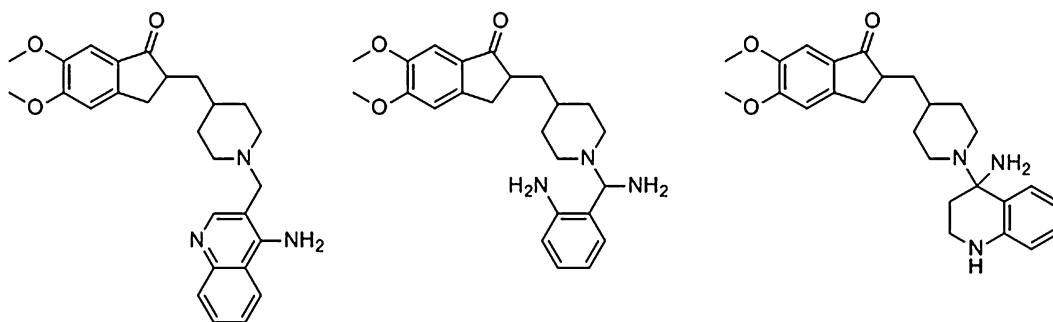


Scheme 19 General chemical structures of aminopyridazine derivatives reported



(*R*)-*N*-((1-(2,3-dihydro-1*H*-inden-2-yl)piperidin-3-yl)methyl)-*N*-(2-(dimethylamino)ethyl)-2-naphthamide

Scheme 20 General structure of naphthamides

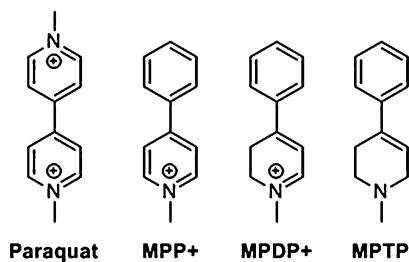


Scheme 21 Donepezil–tacrine-based hybrid derivatives

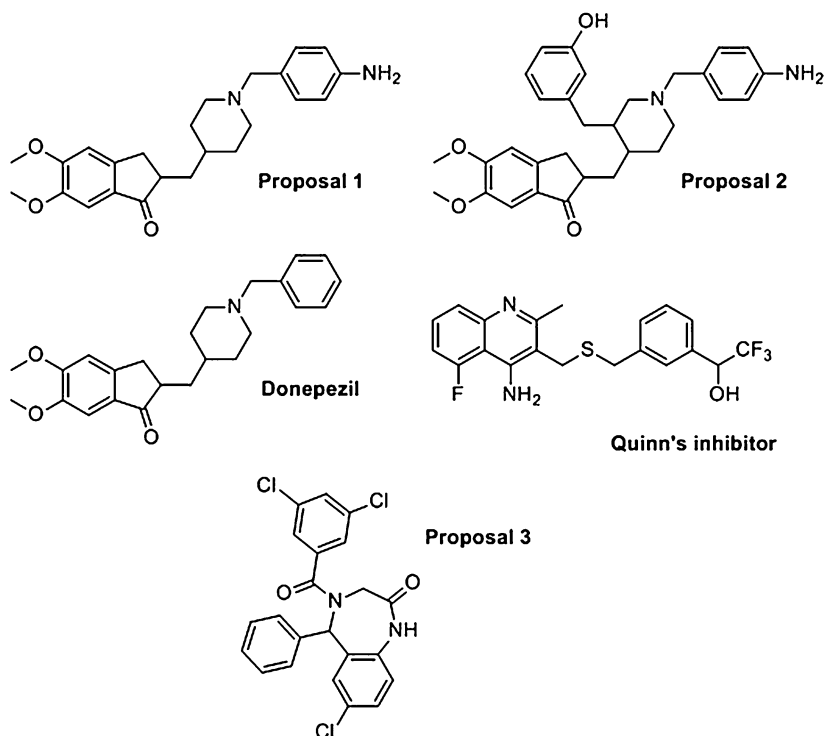
Finally, Kosak et al. [90] in continuation to their previous studies further optimized their previous inhibitors [19] (Scheme 20) driving the potency to the picomolar range. This particular compound co-crystallized with BChE under the PDB entry 5NN0 [90].

In 2006, da Silva and coworkers [91] with the use of docking produced new hybrid molecules (Scheme 21) combining in one entity the donepezil and tacrine structures, in order to produce AChE inhibitors capable of keeping the original biological interactions of both molecules. The AChE–donepezil complex with PDB entry 1EVE was used in the X-ray study [92]. Herein, the researchers predicted the ADMET properties of the hybrid molecules in order to make a drug-like alternative suggestion for treatment.

In 2007, Alcaro and coworkers [93] performed a study on known herbicide analogues like paraquat, MPP⁺, MPDP⁺, and MPTP (Scheme 22). Since these compounds were damaging neurons, the researchers considered to test them as AChE and BChE inhibitors. Their approach included molecular modelling over 1ACJ PDB structure [87] and an enzymatic study against the two enzymes. As a reference compound, they used tacrine. The reported docking results supported a binding mode similar to



Scheme 22 Chemical structures of known herbicides

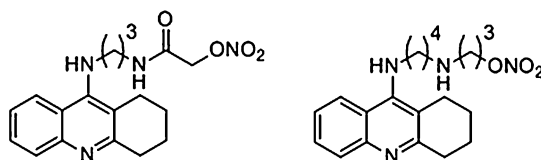


Scheme 23 Structures of the newly proposed and known inhibitors

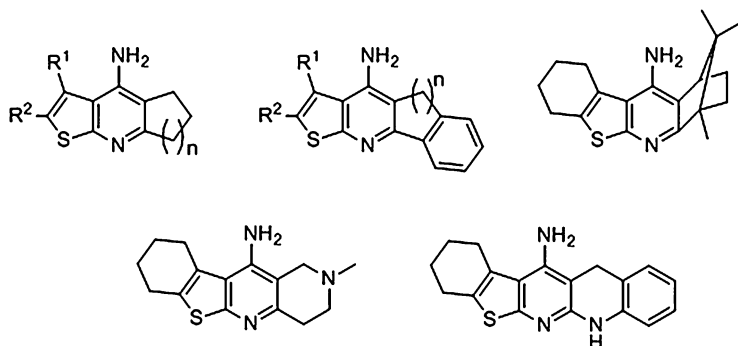
tacrine, whereas the SARs underlined the significant role of the electrostatic properties.

Da Silva et al. [94] reported the computer-aided drug design of three new AChE inhibitors (Scheme 23) following molecular modeling techniques. The structure of choice was the AChE–donepezil complex with PDB entry 1EVE [92]. The results were in accordance to calculated ADMET. Compounds 1 and 3 (proposal-1 and 3) proposed to be used as interesting AChE inhibitors for the treatment of AD.

Fang et al. [95] designed a new series of tacrine-based NO donor hybrid compounds (Scheme 24) as ChE multitarget



Scheme 24 General structures of novel NO-donating tacrine derivatives



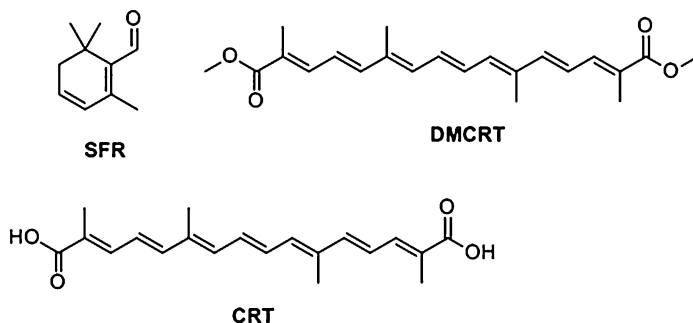
Scheme 25 General chemical structures of the fused thiophene derivatives

inhibitors. For the docking experiments, they used crystal structure 2CKM [96] to proceed AChE binding. All the synthesized compounds showed significant *in vitro* results with their inhibition being in the low-nM range. Specifically, the most active compounds showed AChE inhibitory activities between 5.6 and 9.1 nM being five- to sixfold more potent than the reference compound tacrine. BChE activity remained also on the low-nanomolar range with IC_{50} varying from 7.2 to 18.1 nM. These compounds seemed to occupy both CAS and PAS domains simultaneously and they show a nonselective inhibitory activity.

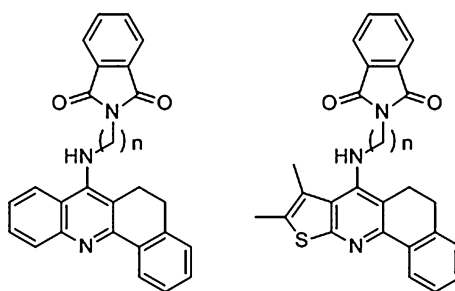
In 2010, Badran et al. [97] synthesized some novel thiophene-fused derivatives (Scheme 25) to replace the tacrine scaffold and tested them as AChE inhibitors. The *in vitro* results showed moderate activity in comparison to tacrine. Hence, a molecular modeling study was also performed to the X-ray structure of AChE with the entry number 1ACJ [87] in order to rationalize the obtained biological results in terms of SAR.

In 2012, emerged a study on the multifunctional properties of saffron natural products (Scheme 26) as anti-AD agents [98]. Four compounds of saffron CRT, DMCRT, and SFR were studied as AChE inhibitors with IC_{50} values of 96.33, 107.1, and 21.09 μ M, respectively. These natural products were also docked against AChE with PDB entry of 1ACJ [87].

El-Malah and coworkers [99] were another team of researchers who published some novel tacrine analogues (Scheme 27) as AChE inhibitors in 2014. Their approach included docking experiments



Scheme 26 Saffron natural compounds studied

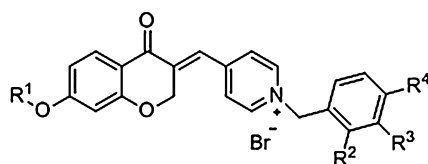


Scheme 27 General structures of tacrine analogues

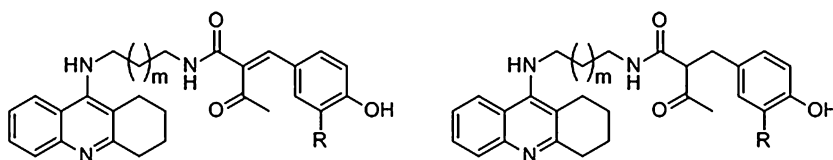
over AChE crystal structure with code 1ACJ [87]. Almost all of the studied compounds exhibited significant *in vitro* activity with some compounds being more potent than tacrine, which have been used as an internal standard.

In 2014, Brus et al. [19] performed drug design based on virtual screening and developed a novel class of BChE selective inhibitor. The used X-ray structure was human BChE in complex with a choline molecule and entry 1P0M [100]. Among 40 resulted hits, the *in vitro* results revealed a highly potent nanomolar inhibitor (IC_{50} 21.3 nM) as a racemate, with the pure (+) enantiomeric compound having an IC_{50} 13.4 nM and K_i 2.7 nM. A crystal structure of this compound over human BChE was also obtained and was available over the Protein Data Bank with code 4TPK [19]. Additionally, a significant inhibition of β fibrils was observed (61.7%).

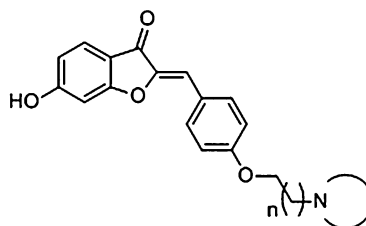
In the same year, another study for AChE inhibition on novel series of chroman-4-one derivatives was published from the group of Arab et al. [101] (Scheme 28). The docking study was performed over AChE PDB entry 1EVE [92]. The most active compound exhibited $IC_{50} = 0.048$ mM. The *in silico* study was found in full compliance with the *in vitro* results.



Scheme 28 Chroman-4-one derivative's general structure



Scheme 29 Tacrine-curcumin hybrid derivatives



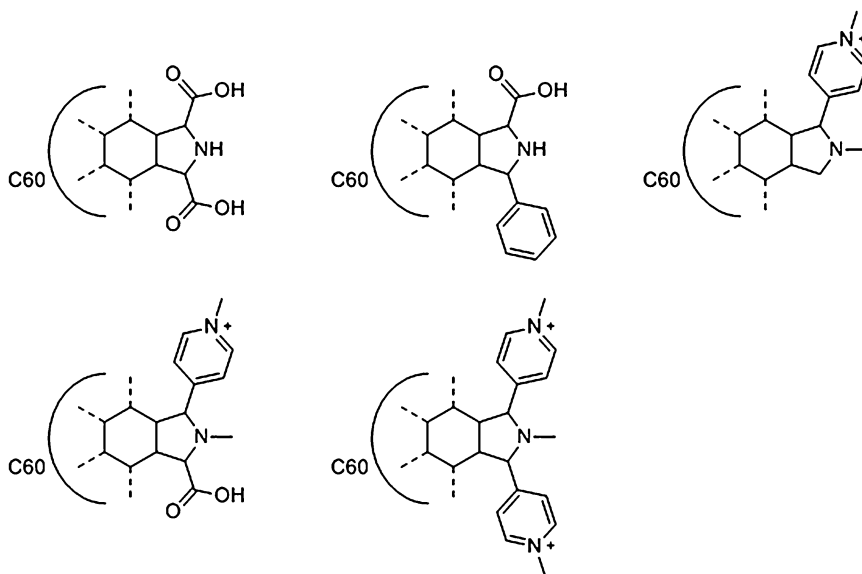
Scheme 30 General structure of benzofuranone derivatives

Liu et al. [102] in 2017 synthesized several tacrine-curcumin hybrid compounds (Scheme 29) as multitarget agents against ChEs. Molecular modeling was performed over PDB entry 2CKM [96] for AChE and 4BDS [103] for BChE. The biological results showed that these hybrids were good cholinesterase inhibitors, compared to the standard used drug (tacrine).

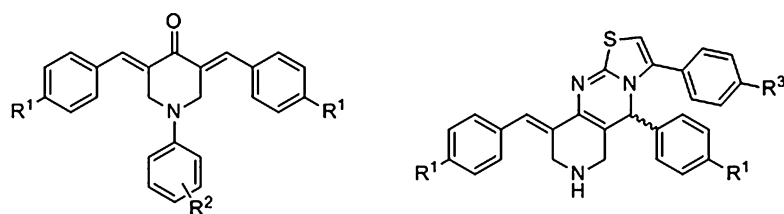
Mehrabi et al. [104] developed benzofuran-3-ones derivatives (Scheme 30) as dual ChE inhibitors. X-ray data utilized in the docking were AChE with PDB entry 1EVE [92] and BChE with PDB entry 1P0I [100]. The *in vitro* inhibitory results for AChE/BChE revealed the presence of very active derivative with $IC_{50} = 0.045 \mu\text{M}$ against AChE. Docking study confirmed these results and provided the possible binding conformation.

Da Silva Goncalves et al. [105] in 2016 proposed some fullerene (C60) derivatives (Scheme 31) as AChE inhibitors. Docking of these compounds was performed using AChE from the Protein Data Bank with the entry code 2X8B [106]. In that work, the results show that the ligands can bind to FASII region of the AChE.

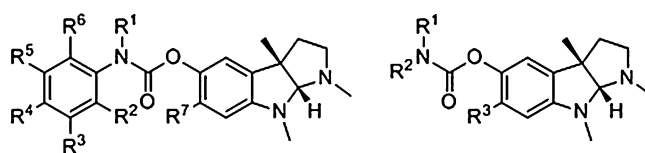
Basiri et al. [107] developed two series of piperidine derivatives (Scheme 32) as AChE inhibitors. Docking was performed with the



Scheme 31 Chemical structures of the C60 derivatives



Scheme 32 Piperidine-based inhibitors

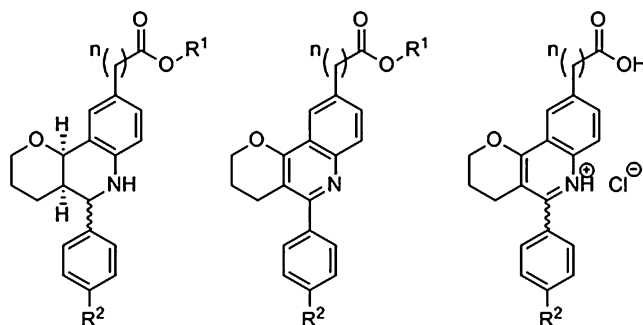


Scheme 33 General structure of the synthesized derivatives

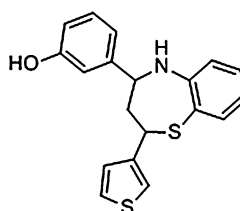
structure 1EVE [92]. The most active compounds displayed IC_{50} at the sub-micromolar range.

4 Multitarget Agents

In 2001, Yu et al. [108] performed a molecular modeling-based approach to synthesize a novel series (Scheme 33) of AChE and BChE inhibitors. Quantification of their results confirmed their



Scheme 34 Synthesized hybrid derivatives



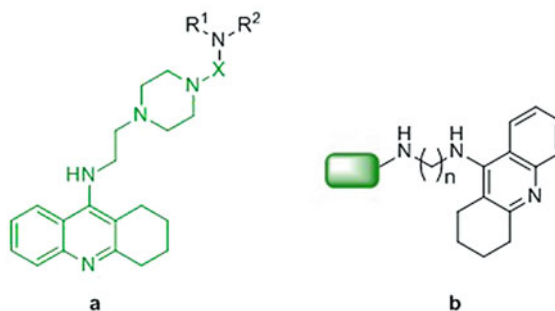
Scheme 35 Chemical structure of the new benzothiazepine studied

binding mode within the enzymes whereas their structural features were also defined.

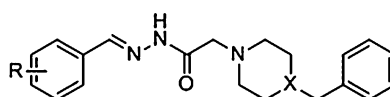
In 2009, Camps and coworkers [109] developed a series of multitarget compounds (Scheme 34) as AChE, BChE, AChE-induced, and self-induced A β fibril aggregate inhibitors and β -secretase (BACE-1) inhibitors. Their design included docking of the compounds against X-ray crystallographic structures of AChE, entries utilized here were 1Q83 [110], 1Q84 [110], 1ODC [96], 1ZGB [111], 1ZGC [111], 2CKM [96], and 2CMF [96]. The reported hybrid compounds showed great potency and selectivity against human AChE activity. Simultaneously, they exhibited significant *in vitro* inhibition of both A β aggregation and BACE-1 and a great ability to enter the central nervous system. The new hybrid compounds bear a 6-chlorotacrine and pyrano[3,2-*c*]-quinoline moiety in their structure and the modeling showed that they reside in both CAS and PAS at the same time.

Ul-Haq and coworkers [112] reported a molecular modeling study of a novel multitarget benzothiazepine (Scheme 35) acting as a dual ChE inhibitor. They docked the studied compound in AChE 1ACL [87] and BChE 1P0P [100].

Hamulakova and coworkers [113] in 2012 developed new tacrine derivatives (Scheme 36) as multifunctional ChE inhibitors, based on docking simulations over the human proteins 1B41/



Scheme 36 Tacrine derivative's general structure

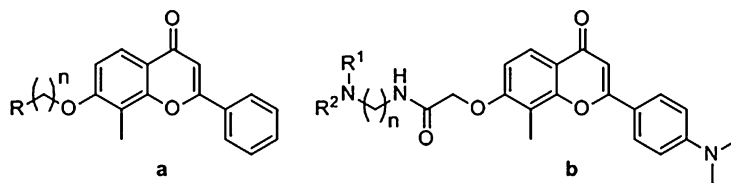


Scheme 37 General structure of the derivatives

AChE [114] and 1P0I/BChE [100]. Mainly, the studied derivatives exhibited an increased inhibition towards both AChE and BChE with IC_{50} values in the low-nanomolar range, compared to tacrine and 7-MEOTA that were used as reference standards. The most active compounds showed IC_{50} values of 4.49 and 4.97 nM, respectively, on AChE with a noticeable selectivity. On the contrary, the most active BChE inhibitor presented an IC_{50} value of 33.7 nM exhibiting also a good selectivity towards the protein. The docking results showed that the structures interact both in CAS and PAS sites.

Makhaeva et al. [115] in the same year made an extensive study combining QSAR approaches to analyze and predict inhibitory properties of 58 O-phosphorylated oximes against AChE, BChE, NTE, and CaE. The following QSAR techniques were also used: Molecular Field Topology Analysis (MFTA), Comparative Molecular Similarity Index Analysis (CoMSIA), and molecular modelling. For the simulations, mouse AChE 2HA3 [116] and human BChE 1P0M [100] crystal data were used. Moreover, due to their interest as ChE inhibitors, the researchers applied MFTA to design a library of some novel inhibitors.

In 2013, Ozturan Ozer et al. [117] designed and tested a series of *N'*-2-(4-benzylpiperidin-/piperazin-1-yl)acylhydrazone derivatives (Scheme 37) as multitarget ligands against both cholinesterases along with anti- $A\beta$ aggregates activity. The design of all derivatives was based on docking simulations performed on the human proteins 1B41/AChE [114] and 1P0I/BChE [100]. Although, due to the fact that the selected proteins did not have any ligands co-crystallized, they superimposed these structures over protein entry 1EVE/AChE [92] co-crystallized with the known

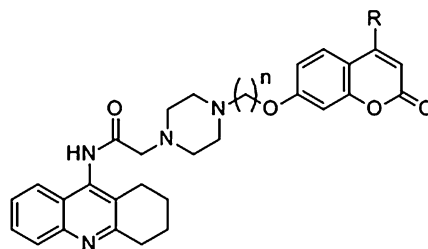


Scheme 38 Developed flavonoid derivatives' general structures as multitarget AD agents

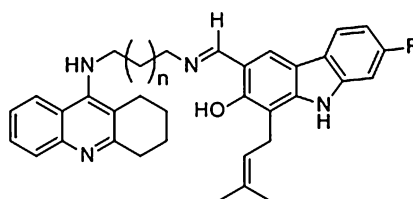
drug Aricept [118]. Moreover, for all the designed molecules were calculated Lipinski's rule of five [16, 17] properties, along with the polar surface area indicator in an attempt to predict their blood-brain barrier penetration. The in vitro results showed no selectivity towards the targeted proteins, and no greater potency in comparison to donepezil (IC_{50} 23.1 nM/human AChE and IC_{50} 7.4 μ M/equine BChE). Their activities varied from 53.1 to 88.5 μ M for AChE and 48.8 to 98.8 μ M for BChE. The calculated binding constants (K_i) for some derivatives showed similar values in the low- μ M range. On the contrary, all the derivatives significantly inhibit A β aggregation compared to rifampicin [119] which was used as a reference for the assay. The docking studies indicated that both PAS and CAS interactions of human AChE were present, supporting a multisite binding mode for the compounds.

Li and coworkers [120] reported multitarget ligands against ChEs and A β aggregate inhibition based on the flavone moiety (Scheme 38, structure a). Their approach also contained docking experiments over known crystal data of ChEs, using 1EVE/AChE [92] 1P0I/BChE [100], respectively. These compounds showed greater potency and selectivity for AChE over BChE, combined with a good inhibition against A β fibril formation. All the compounds present in vitro low inhibitory potency in the sub-micromolar range. However, the best compound exhibited a 20-fold and a 2-fold drop in activity compared to the used control substances, galanthamine and tacrine, respectively. Additionally, these derivatives were designed to chelate metals. The most active compound presented similar activity towards Cu^{2+} and Fe^{2+} . The present study revealed that compounds targeted both CAS and PAS regions of AChE, a factor that was directly associated with the alkyl chain length.

Luo et al. in 2013 [121] developed a series of multitarget flavone derivatives (structure b) (Scheme 38), as inhibitors of ChEs and A β aggregates. The design of the new molecules was based on molecular modeling techniques over the torpedo 1ACJ/AChE [87] and the human 1P0I/BChE [100]. All compounds presented in vitro equipotent inhibitory activities (on the low-micromolar region with IC_{50} values varying from 1.83 to 33.20 μ M and 0.82 to 11.45 μ M, respectively) on both AChE and BChE. Rivastigmine was used as the reference compound [122]. A significant inhibition of A β aggregates was also recorded



Scheme 39 Tacrine–coumarin hybrid derivative's general structure

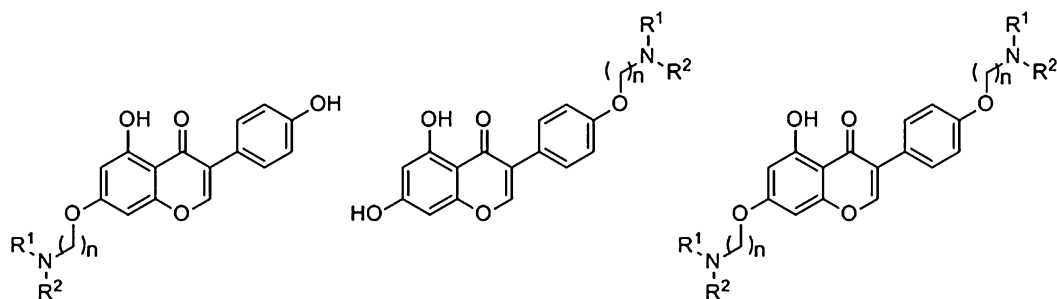


Scheme 40 Tacrine–carbazole hybrid molecule's general structure

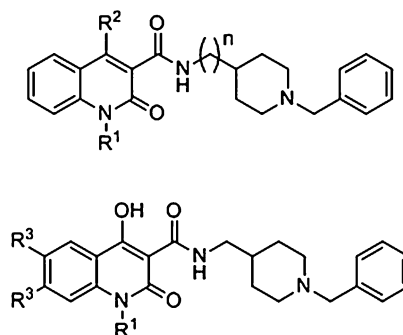
for the mentioned compounds. Moreover, some of the compounds exhibited a radical scavenging ability. The activity profile along with the modeling study suggests that these compounds can target both the CAS and PAS regions of AChE.

A group from China [123] published their work on tacrine–coumarin hybrid molecules (Scheme 39) as an alternative multi-target approach for the treatment of AD. They focused on a more combined activity against ChEs and A β aggregation. The design of the molecules was based on docking over AChE protein with entry number 2CKM [96], which was co-crystallized with a carbon chain linked bis-tacrine ligand. All the compounds exhibited significant inhibition upon the abovementioned targets. The best compound showed IC₅₀ value of 0.092 μ M for AChE, 0.234 μ M for BChE, and 67.8% inhibition of fibril formation at a 20- μ M concentration. The binding mode of derivatives revealed a mixed-type inhibition occupying simultaneously CAS, PAS, and mid-gorge sites of AChE.

Thiratmatrakul et al. [124] in 2014 developed a series of new tacrine–carbazole hybrids (Scheme 40) as potential multifunctional anti-Alzheimer agents with ChE inhibitory and radical scavenging activities. The binding interactions of these derivatives were also studied utilizing AChE with the PDB entry 2CEK [125]. All developed compounds showed high inhibitory activity on AChE with IC₅₀ values 0.48–1.03 μ M and good selectivity index over AChE/BChE. Molecular modeling studies revealed that these tacrine–carbazole hybrids interacted simultaneously with CAS and PAS sites of AChE. Some derivatives showed scavenging activity; while considering their neuroprotection in total, these derivatives could



Scheme 41 Genistein derivatives



Scheme 42 General structures of the quinolinecarboxamide derivatives

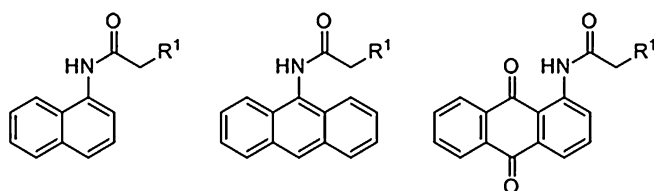
reduce neuronal death induced by oxidative stress and A β fibrils formation.

Qiang et al. [126] developed a series of genistein derivatives with various carbon spacer-linkers (Scheme 41) as multifunctional agents for AD. Their docking experiments were performed using crystal data of AChE in presence with donepezil and code 1EVE [92] in the Protein Data Bank. In vitro results exhibited good AChE activity combined with moderate-to-good anti-oxidative activity. The most potent compound was on the range of sub-micromolar values (IC_{50} 0.09 ± 0.02 μ M). Also, several of the studied compounds exhibited significant inhibition of A β aggregation with metal-chelating properties. The binding mode analysis for these derivatives suggested the typical occupation of CAS and PAS domains.

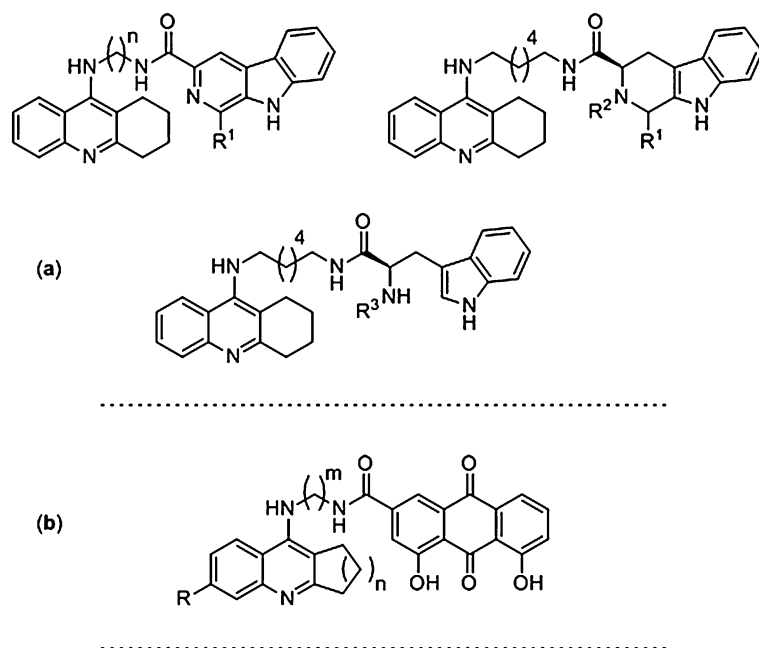
Pudlo and coworkers [127] designed, synthesized, and evaluated a series of quinolinecarboxamide derivatives (Scheme 42) as new multifunctional AChE inhibitors and radical scavengers. Molecular modeling was performed over AChE structures obtained from Protein Data Bank with accession codes 4EY7 [128] and 1EVE [92], respectively. The in vitro results showed good AChE inhibition with good selectivity index over BChE and high radical scavenging. Most active compounds were found to inhibit AChE in the sub-micromolar range.

The work of Hong et al. [129] in 2014 was focused on the development of multitarget compounds on both ChEs and thus reported three types of aromatic–polyamine conjugates (Scheme 43). Derivatives were subjected to a docking analysis to clarify their binding modes and to derive SAR. The procedure utilized AChE structure with code entry 1ACJ [87]. Anthraquinone conjugates were the most potent inhibitors for AChE with IC_{50} values 1.5–11.13 μM , while anthracene conjugates proved to be highly selective towards BChE accompanied by IC_{50} values ranging from low- to high-nanomolar range (0.016–0.657 μM). Kinetic and modeling analysis revealed that the compounds under study occupy both CAS and PAS domains in ChEs.

Continuing their efforts, Lan et al. [130] reported the design of tacrine–carboline hybrids (Scheme 44a) as multitarget agents aiming ChEs, $A\beta$ aggregation, and metal chelation. Their design



Scheme 43 Polyamine conjugate structures



Scheme 44 Tacrine–carboline hybrid derivatives (a), tacrine–rhein hybrid derivatives (b), tacrine-related analogues

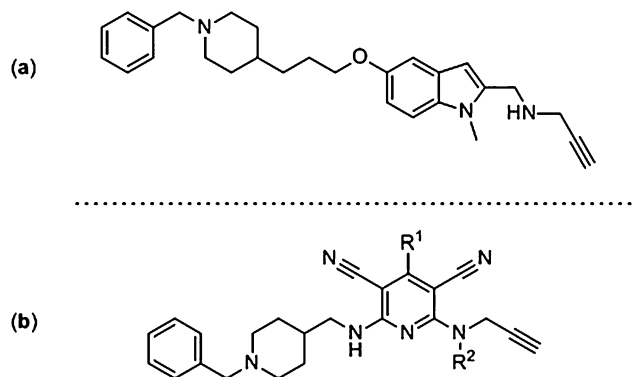
strategy was based on docking experiments of the AChE PDB entry 2CKM [96] and the BChE PDB entry 1P0I [100]. In vitro results showed for most of the compounds significant inhibition of AChE, BChE, and self-induced A β aggregation, Cu²⁺-induced A β _{1–42} aggregation, and metal chelation. The most active compound of all exhibited IC₅₀ 21.6 nM for eeAChE, 63.2 nM for hAChE, and 39.8 nM for BChE, 65.8% inhibition on A β aggregate formation at a 20- μ M concentration combined with antioxidant activity. The study indicated that most active derivative was mixed-type inhibitor simultaneous binding to CAS and PAS sites of AChE.

The group of Li and coworkers [131] reported in 2014 a series of tacrine–rhein hybrids (Scheme 44b) as novel multitarget inhibitors for ChEs. Crystallographic data of AChE complexed with bis (7)-tacrine (PDB entry 2CKM [96]) was utilized in their design efforts. In vitro results of the most active compound showed a fivefold drop when compared to reference compound tacrine for AChE inhibition and low-nanomolar potency for BChE inhibition (IC₅₀ 200 nM). The mode of activity was the simultaneous typical binding at the catalytic and peripheral sites of AChE. Additionally, the most active compound showed a moderate inhibition of A β aggregate formation and metal-chelating properties.

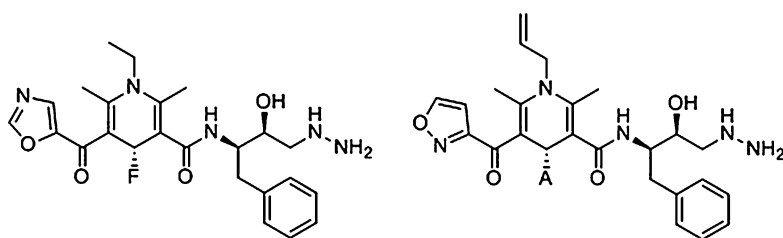
Stoddard and coworkers [132] based on marine metabolites developed a new series of multitarget compounds aiming inhibition of AChE and A β aggregate formation. Their work was involved over seven classes of known marine metabolites by combining experimental data with docking results. Several PDB entries of AChE were utilized in this study, including 1ACL [87], 1ACJ [87], 1DX6 [133], and 1EVE [92].

Bautista-Aguilera et al. [134] reported the design, synthesis, and pharmacological evaluation of donepezil–indolyl-based derivatives as multitarget agents able to inhibit simultaneously ChEs and MAO. In their theoretical studies, they used 3D-QSAR to define pharmacophores able to inhibit enzymes MAO A/B, AChE, and BChE. The in silico approaches included also the docking of derivatives against the selected enzymes with X-ray structures 1C2B [135] for AChE, 2PM8 [136] for BChE, 2Z5X [137] for MAO-A, and 2V5Z [138] for MAO-B. The most active compound propargylamine (Scheme 45a) exhibited a nanomolar inhibition for mostly all enzymes (MAO-A IC₅₀ = 5.5 \pm 1.4 nM, MAO-B IC₅₀ = 150 \pm 31 nM, AChE IC₅₀ = 190 \pm 10 nM, and BChE IC₅₀ = 830 \pm 160 nM inhibitor).

Continuing their efforts, Bautista-Aguilera et al. [139] reported in the same year a new generation of multitarget compounds (Scheme 45b) as inhibitors of ChEs and MAO. The in silico approach was the same as mentioned above, combining 3D-QSAR for identifying pharmacophores and docking of derivatives against the selected enzymes with crystal data of 1C2B [135] for AChE, 2PM8 [136] for BChE, 2Z5X [137] for MAO-A, and 2V5Z [138]



Scheme 45 General chemical structure of the propargylamine derivative (a) and donepezil-pyridyl hybrids (b)

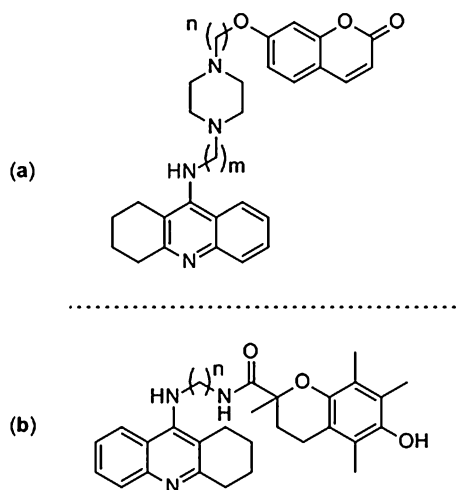


Scheme 46 Chemical structure of the compounds EDC and FDC, respectively

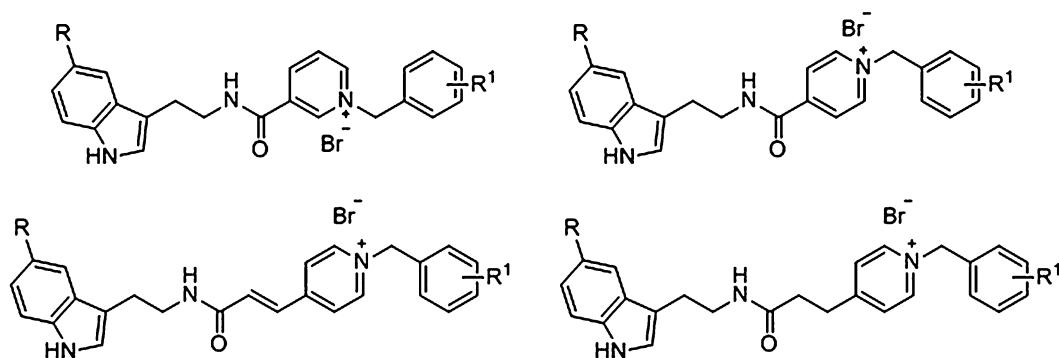
for MAO-B. For the most potent inhibitor, the *in vitro* results showed $IC_{50} = 1.1 \pm 0.3$ nM for AChE, $IC_{50} = 600 \pm 80$ nM for BChE, and $IC_{50} = 3.950 \pm 940$ nM for MAO-B.

Goyal and coworkers [140] in 2014 designed multitarget dihydropyridine derivatives as AChE and β -secretase (BACE-1) inhibitors. Their approach included a highly predictive group-based QSAR (GQSAR) model followed by molecular docking. The X-ray structures used were AChE with PDB code 4M0E [141] and BACE-1 with PDB code 2B8L [59]. The present study gave two highly potent dual inhibitors (EDC and FDC) as depicted in Scheme 46.

In 2015, Xie et al. [142] developed a series of tacrine-coumarin hybrid compounds (Scheme 47a) as multitarget agents against AChE, BChE, and MAO-B inhibitors. In the docking experiments, the researchers used structures PDB code 4EY7 [128] for AChE and PDB code 2V61 [138] hMAO-B. The *in vitro* results of the hybrids indicated that most of them were potent inhibitors against the tested enzymes, and the most active one exhibited $IC_{50} = 33.63$ nM for eeAChE and $IC_{50} = 16.11$ nM for hAChE, $IC_{50} = 80.72$ nM for eqBChE and $IC_{50} = 112.72$ nM for hBChE, and $IC_{50} = 0.24$ μ M for hMAO-B. The binding analysis indicated a



Scheme 47 General chemical structure of the synthesized tacrine–coumarin hybrids (a) and tacrine–trolox hybrids (b)

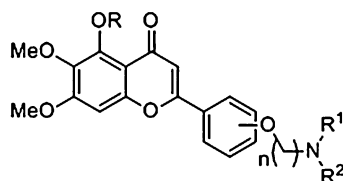


Scheme 48 General structures of the benzylpyridinium hybrids

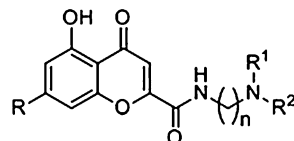
mixed-type inhibitory activity at catalytic, peripheral, and mid-gorge sites of AChE.

Another tacrine-based hybrid series of compounds (Scheme 47b) have been presented by the group of Xie et al. [143] as novel multifunctional hybrids showing ChEs inhibition combined with antioxidant activity. X-ray structure of the human AChE with PDB entry 4EY7 [128] was utilized for the *in silico* study. The most active *in vitro* compound showed an $IC_{50} = 9.8$ nM for eeAChE and $IC_{50} = 23.5$ nM for hAChE, $IC_{50} = 22.2$ nM for eqBChE and $IC_{50} = 20.5$ nM for hBChE. All the compounds identified as mixed-type inhibitors.

The group of Luo et al. [144] in 2015 reported the design and synthesis of a novel series of melatonin-based benzylpyridinium bromides (Scheme 48) as multitarget agents presenting ChEs inhibition, antioxidant, and neuroprotective activities. The *in silico*



Scheme 49 Scutellarein derivative's general structure

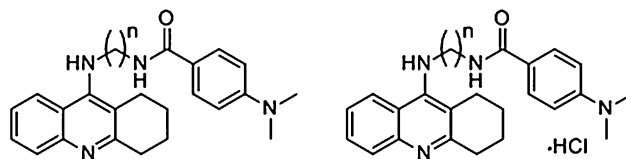


Scheme 50 General structure of the synthesized chromone derivatives

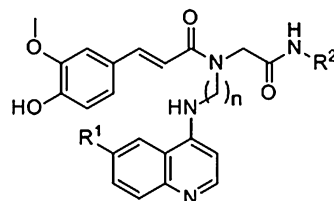
simulations were performed over X-ray structure of AChE with code 1EVE [92]. The *in vitro* biological investigation showed that the majority of the compounds were potent inhibitors against AChE and BChE with good antioxidant activity. The most active of them displayed an $IC_{50} = 0.11 \mu\text{M}$ for AChE, $IC_{50} = 1.1 \mu\text{M}$ for BChE and good antioxidant activity. The mode of binding followed the usual mixed-type.

Sang et al. [145] designed and synthesized a series of scutellarein-based derivatives as multifunctional agents targeting AChE and A β aggregate inhibition along with metal-chelating properties (Scheme 49). Docking was used for the design of the molecules utilizing the X-ray structure 1EVE [92]. Most of the compounds exhibited good activity against the selected targets. The best one demonstrated significant metal-chelating properties, moderate acetylcholinesterase (AChE) inhibition and antioxidant activity, combined excellent inhibitory effects on self-induced and Cu^{2+} -induced A β_{1-42} aggregation, human AChE-induced A β_{1-40} aggregation, and disassembled Cu^{2+} -induced aggregation of the well-structured A β_{1-42} fibrils. The binding mode of the compounds was typically situated at both CAS and PAS domains.

Liu et al. [146] in 2015 designed a series of multifunctional chromone derivatives (Scheme 50) as AChE inhibitors with antioxidant, metal-chelating properties, and A β aggregation ability. The binding mode and the interactions of compounds were performed over AChE PDB entry 1EVE [92]. The most active compound showed AChE $IC_{50} = 0.07 \pm 0.01 \mu\text{M}$. The analysis showed a mixed-type inhibition. An excellent self-induced A β aggregation inhibition was observed (63.0%) as well as a good Cu^{2+} -induced A β aggregation inhibition (55.6%) and metal-chelating ability.



Scheme 51 Tacrine hybrid derivatives



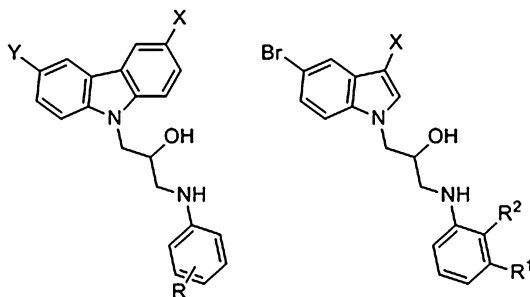
Scheme 52 Tacrine-ferulic acid hybrid derivatives

Bajda and coworkers [147] designed a novel series of tacrine derivatives (Scheme 51) as multitarget inhibitors against ChEs and A β aggregate formation. The utilized crystal data of ChEs were entry 2CKM for AChE [96] and 1P0I for BChE [100]. The derivatives were proved very potent inhibitors of both AChE and BChE with IC₅₀ values from sub-nanomolar to the nanomolar range higher than rivastigmine and tacrine that were used as reference compounds in the biological testing. Inhibition of A β fibrils formation was also significant.

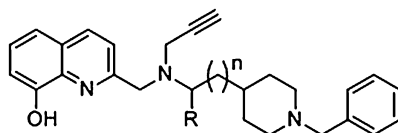
Benchekroun et al. [148] designed some new tacrine-ferulic acid hybrids (Scheme 52) as ChEs and A β aggregate inhibitors with antioxidant activity. They used AChE PDB entry 4EY7 [128] and BChE PDB entry 4BDS for the X-ray docking studies [103]. The most active compound showed moderate and selective BChE inhibition with IC₅₀ = 68.2 \pm 3.9 nm, good BBB permeation, strong antioxidant activity, high A β ₁₋₄₂ aggregate formation (65.6%), and good neuroprotection.

Dominguez et al. [149] designed some novel multitarget ligands for AChE and BACE1 inhibition (Scheme 53). The docking experiments were performed using the X-ray structure of AChE entry 1ODC [96] and BACE1 entry 1FKN [150]. Besides the robust predictions on designed compounds, the *in silico* protocol revealed insights on the flap opening in BACE1. The *in vitro* results of most active candidates show high ligand efficiency displaying a multitarget behavior on the amyloid cascade and cholinergic activity.

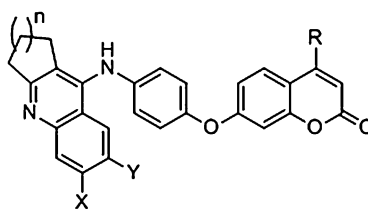
In 2016, Wu and coworkers [151] developed some donepezil-based multitarget chelators (Scheme 54) as inhibitors of ChEs, MAO-A, and MAO-B. The design was based on their docking against AChE PDB entry 1B41 [114], MAO-A PDB entry 2Z5X



Scheme 53 General chemical structure



Scheme 54 General chemical structure of the donepezil-like derivatives

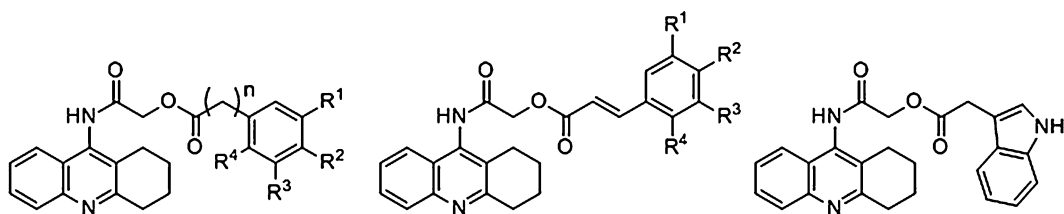


Scheme 55 Novel chromenone hybrid's general structure

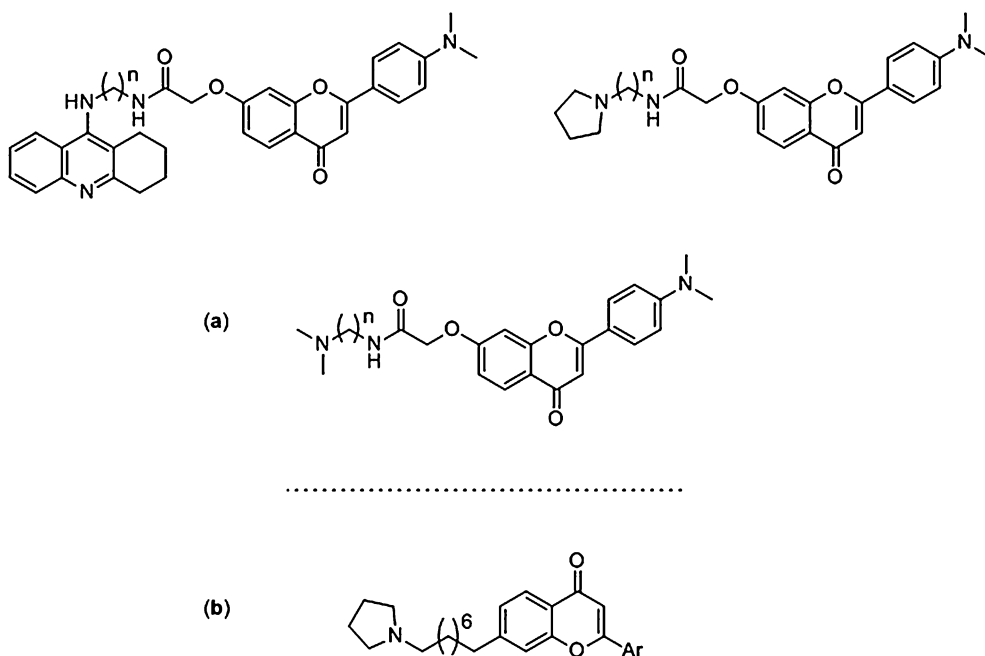
[137], and MAO-B PDB entry 2V5Z [138]. The *in vitro* results exhibited high potencies for AChE and BChE inhibition with IC_{50} values in the nanomolar range, followed by a micromolar selective MAO-A inhibition. Moreover, compounds showed metal-chelating activity and moderate antioxidant ability.

Najafi et al. [152] developed chromenone hybrid derivatives (Scheme 55) and evaluated them as inhibitors of ChEs and BACE-1. The *in silico* design included docking against the X-ray data of AChE PDB entry 2CMF [96]. The most active compounds showed IC_{50} values comparable to the internal standard (rivastigmine). In addition, they exhibited satisfactory inhibition of BACE-1 and neuroprotective activity. Herein, also the binding mode of the inhibitors showed both CAS and PAS occupancy.

Zhang et al. [153] designed new tacrine hybrids (Scheme 56) as multitarget compounds against ChEs and A β fibril formation. In order to evaluate their inhibitory binding modes, they docked them against TcAChE PDB entry 2CMF [96] and hBChE PDB entry 4BDS [103]. All the designed molecules presented IC_{50} values at



Scheme 56 Tacrine hybrid derivatives' structures

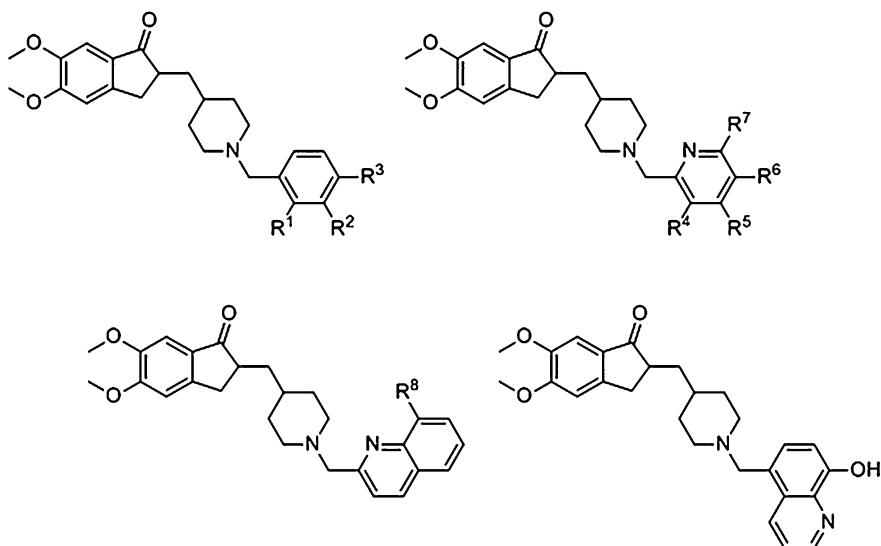


Scheme 57 Flavonoid-based hybrid derivatives (a) and flavonoid derivatives (b)

the nanomolar range, with the best displaying high selectivity and $IC_{50} = 5.63$ nM for AChE. Also, the exhibited inhibition of A β aggregate formation was 51.81 nM.

Luo et al. [154] on the contrary developed a new series of flavonoid-based derivatives (Scheme 57a) as multitarget agents for ChEs, self-induced A β aggregates, and antioxidant activity. The crystal structures of AChE code 1ACJ [87] and the BChE code 1POI [100] were utilized in their *in silico* experiments in order to determine their mode of activity, showing the typical CAS and PAS interaction in AChE. Almost all of the compounds exhibited potent AChE and BChE inhibition along with potent self-induced A β fibril formation inhibition and radical scavenging activity.

Continuing their efforts, Luo et al. [155] designed novel flavonoid derivatives (Scheme 57b) as dual ChEs inhibitors. They used AChE PDB entry 1ACJ [87] to extract binding mode features. The



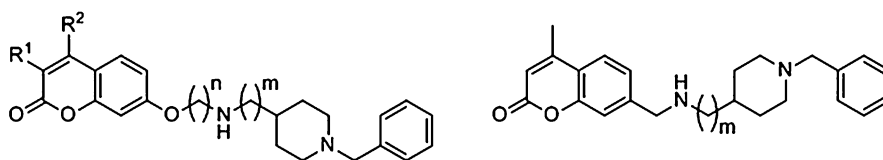
Scheme 58 Donepezil hybrid derivatives

in vitro results showed that most of the synthesized compounds exhibited potent AChE and BChE inhibitory activities at the low- and sub-micromolar range. The most active compound exhibited $IC_{50} = 0.64 \mu\text{M}$ against AChE and $IC_{50} = 0.42 \mu\text{M}$ for BChE, with the same interactions on the protein.

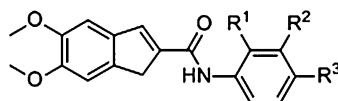
Wang and coworkers [156] synthesized a series of donepezil derivatives (Scheme 58) as multifunctional AChE and BChE inhibitors, metal chelators, and antioxidant agents and inhibitors of AChE- and Cu^{2+} -induced A β inhibition of fibril formation. Binding modes were analyzed by docking using AChE PDB entry 4EY7 [128] and BChE PDB entry 4XII [157]. These derivatives displayed good antioxidant, A β aggregation inhibition as well as good calculated in silico ADMET properties. The most active derivative presented a low-nanomolar inhibition with $IC_{50} = 85 \text{ nM}$ for eeAChE and $IC_{50} = 73 \text{ nM}$ for hAChE.

Xie and coworkers [158] developed some novel coumarin-based hybrid compounds (Scheme 59) as ChEs and MAO-B inhibitors. Docking approaches were used for binding mode prediction over AChE PDB entry 4EY7 [128] and MAO-B PDB entry 2V61 [138]. The most active compound exhibited inhibition of *Electrophorus electricus* AChE (eeAChE) $IC_{50} = 0.87 \mu\text{M}$ and of equine BChE (eqBChE) $IC_{50} = 0.93 \mu\text{M}$. Another compound presented a good and balanced inhibition with $IC_{50} = 1.37 \mu\text{M}$ for hAChE, $IC_{50} = 1.98 \mu\text{M}$ for hBChE, and $IC_{50} = 2.62 \mu\text{M}$ for hMAO-B. The mode of activity was a typical mixed-type with simultaneous binding to CAS, PAS, and mid-gorge site of the AChE.

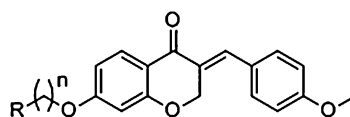
Koca et al. [159] designed some novel indene derivatives (Scheme 60) as multifunctional ChEs and A β aggregate inhibitors.



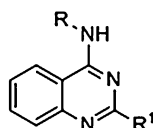
Scheme 59 Coumarin hybrid derivatives



Scheme 60 Indene derivatives



Scheme 61 Homoisoflavonoid derivatives

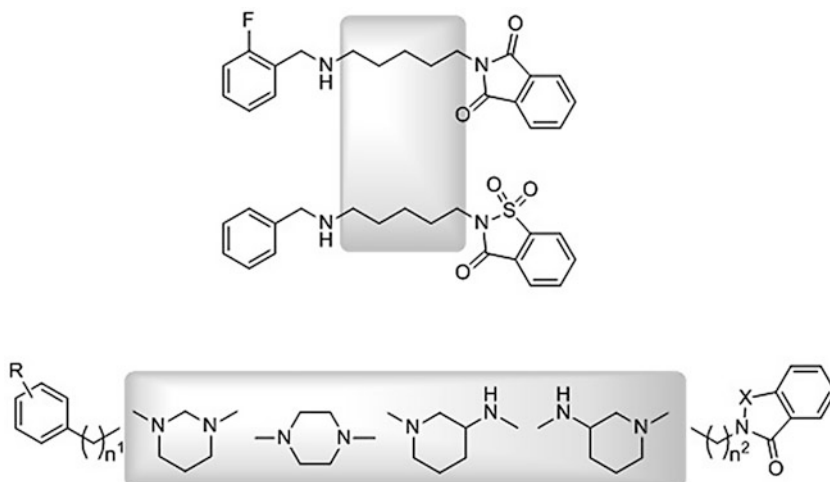


Scheme 62 General chemical structure of quinazoline hybrids

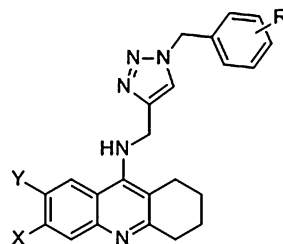
In silico simulations were based on X-ray structures of AChE entry 1EVE [92] and BChE entry 1P0I [100]. The compounds revealed higher inhibitory potency against BChE and most of them presented also remarkable A β fibril formation. The mode of inhibition was noncompetitive and the interactions were on CAS and PAS regions.

Wang and coworkers [160] published in 2016 another study on multitarget flavonoid hybrids (Scheme 61) as dual AChE and MAO-B inhibitors. Docking simulations were based on X-ray structures PDB entry 2V60 [138] for MAO-B and 2CMF [96] for ACh. The most active compound showed a balanced inhibitory activity with IC₅₀ values of 3.94 for AChE and 3.44 μ M MAO-B, revealing a mixed-type inhibition for the proteins as confirmed also from the obtained in silico data.

Mohamed et al. [161] designed 2,4-disubstituted quinazoline compounds (Scheme 62) as multitarget agents presenting ChEs inhibition and antioxidant activity. Molecular modeling studies included X-ray structures AChE PDB entry 1B41 [114], BChE PDB entry 1P0I [100], and NMR solution structure of A β fibrils



Scheme 63 Donepezil derivatives

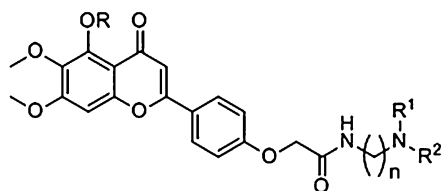


Scheme 64 General structure of tacrine hybrid derivatives

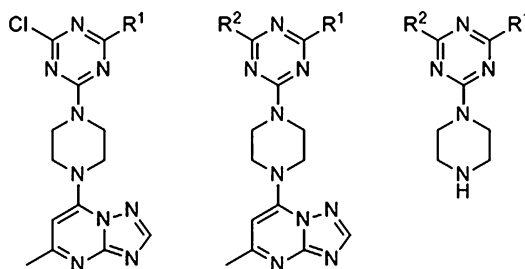
with PDB entry 2LMN [77]. The *in vitro* inhibition results for AChE and BChE enzymes as IC_{50} values were ranged from 1.6 to 30.5 μM , whereas the $A\beta$ fibril formation was high with $IC_{50} = 270 \text{ nM}$ –16.7 μM .

In 2017, Panek et al. [162] developed phthalimide and saccharin derivatives (Scheme 63) as multifunctional inhibitors against ChEs, β -secretase, and $A\beta$. Their binding modes were investigated over the structures AChE PDB entry 1EVE [92] and BACE-1 PDB entry 4D8C [163]. For the majority of the compounds, the AChE inhibition was in the low-micromolar region 0.83–19.18 μM . On the other hand, BACE-1 inhibition ranged from 26.71 to 61.42% at 50 μM concentration. Binding mode of the compounds was the same, noncompetitive, interacting with both CAS and PAS sites.

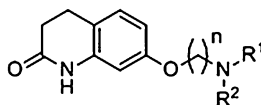
Najafi et al. [164] developed new tacrine-based hybrids (Scheme 64) as potent dual ChEs inhibitors. The utilized structures in this study were 2CMF for AChE [96] and 1POI for BChE [100]. Most of the tested compounds showed good *in vitro* inhibition on both AChE and BChE. The most active exhibited AChE inhibition of $IC_{50} = 0.521 \mu\text{M}$ and $IC_{50} = 0.055 \mu\text{M}$.



Scheme 65 General structure of scutellarein derivatives



Scheme 66 Triazine hybrid derivatives



Scheme 67 Quinoline hybrid derivatives

Sang et al. [165] developed multitarget scutellarein derivatives (Scheme 65) as ChEs inhibitors, metal chelators, antioxidant agents, and for their ability to influence the A β aggregation and disaggregation. The structure of AChE PDB entry 1EVE [92] was used for simulations. The most active compound demonstrated excellent AChE inhibition, moderate A β inhibitory effects, metal-chelating, and antioxidant activity.

Jameel and coworkers [166] developed some multitarget ligands as triazine–triazolopyrimidine hybrids (Scheme 66). Molecular modeling was performed over the X-ray structures of AChE with PDB entry 1EVE [92] and BChE with PDB entry 4TPK [19]. The best compounds showed encouraging inhibitory activity against AChE with IC₅₀ values 0.065–0.092 μ M, and selectivity in comparison to BChE.

Sang et al. [167] developed a novel series of quinoline-based (Scheme 67) multitarget molecules as ChEs and MAOs inhibitors. The utilized structures were AChE PDB entry 1EVE [92] and BChE PDB entry 4TPK [19]. The most active compound demonstrated an AChE and BChE inhibition with IC₅₀ values of 0.56 μ M and 2.3 μ M, respectively, while MAO-A IC₅₀ = 0.3 \pm 0.001 μ M and MAO-B IC₅₀ = 1.4 \pm 0.01 μ M inhibition.

5 Conclusion

This chapter summarizes the diverse receptors or enzymatic systems which are implicated in the pathogenesis of AD, in vitro active agents against AD, state-of-the-art docking studies done in AD targets, and docking studies of multitarget agents with particular emphasis on AD. Recent advances and examples of the exploitation of MTDL approach in the rational design of novel drug candidate prototypes for the treatment of AD and examples of “one-compound-one-target” were described. Alzheimer’s disease is a complex neurodegenerative disorder thus, so far, the therapeutic paradigm “one-compound-one-target” has failed to cure the disease. The QSAR results revealed that some important molecular characteristics should significantly affect ligands’ binding affinities, e.g., hydrophobic and electrostatic interactions.

In general, from the presented computational results it can be concluded that the MTDL approach can offer in the design of multifunctional agents targeting simultaneously more than one of the following enzymes: AChE, BChE, MAO, and BACE. Most of the examples describe inhibitors of both cholinesterases. In some cases, molecules with combined inhibitory activities against AChE, BChE, MAO or AChE, MAO, or AChE, BChE and BACE were recorded. In most of the cases, the multitarget molecules exhibited some other activities as A β aggregates inhibition, metal-chelating, and scavenging activities. Among the studied multitarget agents, a number of hybrids were also computationally studied. These groups were synthesized by the combination of known anti-Alzheimer’s drugs, especially tacrine and donepezil or coumarin.

The developed models are useful tools to predict the binding activities for new multifunctional small anti-AD agents for which ADMET properties should be also predicted.

References

1. Mitra A, Dey B (2013) Therapeutic interventions in Alzheimer disease. In: Neurodegenerative diseases. InTech, London
2. Silva T, Reis J, Teixeira J, Borges F (2014) Alzheimer’s disease, enzyme targets and drug discovery struggles: from natural products to drug prototypes. *Ageing Res Rev* 15:116–145. <https://doi.org/10.1016/j.arr.2014.03.008>
3. Cheng X, Zhang L, Lian Y-J (2015) Molecular targets in Alzheimer’s disease: from pathogenesis to therapeutics. *Biomed Res Int* 2015:760758
4. Grill JD, Cummings JL (2010) Novel targets for Alzheimer’s disease treatment. *Expert Rev Neurother* 10(5):711
5. Kumar A, Nisha CM, Silakari C, Sharma I, Anusha K, Gupta N, Nair P, Tripathi T (2016) Current and novel therapeutic molecules and targets in Alzheimer’s disease. *J Formos Med Assoc* 115(1):3–10. <https://doi.org/10.1016/j.jfma.2015.04.001>
6. Nicolotti O, Giangreco I, Introcaso A, Leonetti F, Stefanachi A, Carotti A (2011) Strategies of multi-objective optimization in drug discovery and development. *Expert Opin Drug Discov* 6(9):871–884. <https://doi.org/10.1517/17460441.2011.588696>
7. Dobi K, Hajdu I, Flachner B, Fabo G, Szaszko M, Bogнар M, Magyar C, Simon I, Szisz D, Lorincz Z, Cseh S, Dorman G (2014) Combination of 2D/3D ligand-

- based similarity search in rapid virtual screening from multimillion compound repositories. Selection and biological evaluation of potential PDE4 and PDE5 inhibitors. *Molecules* 19(6):7008–7039. <https://doi.org/10.3390/molecules19067008>
8. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG (2012) ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model* 52(7):1757–1768. <https://doi.org/10.1021/ci3001277>
 9. Irwin JJ, Shoichet BK (2005) ZINC—a free database of commercially available compounds for virtual screening. *J Chem Inf Model* 45(1):177–182. <https://doi.org/10.1021/ci049714+>
 10. Sterling T, Irwin JJ (2015) ZINC 15—ligand discovery for everyone. *J Chem Inf Model* 55(11):2324–2337. <https://doi.org/10.1021/acs.jcim.5b00559>
 11. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2016) PubChem substance and compound databases. *Nucleic Acids Res* 44(D1):D1202–D1213. <https://doi.org/10.1093/nar/gkv951>
 12. ChempSpider (2017) <http://www.chemspider.com/>
 13. ChEMBLdb (2017) Release 23. <https://www.ebi.ac.uk/chembl/downloads>
 14. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* 28(1):235–242. <https://doi.org/10.1093/nar/28.1.235>
 15. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, Stockinger H (2012) ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res* 40(web server issue):W597–W603. <https://doi.org/10.1093/nar/gks400>
 16. Lipinski CA (2004) Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol* 1(4):337–341. <https://doi.org/10.1016/j.ddtec.2004.11.007>
 17. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 46(1–3):3–26. [https://doi.org/10.1016/S0169-409X\(00\)00129-0](https://doi.org/10.1016/S0169-409X(00)00129-0)
 18. Baell JB, Holloway GA (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem* 53(7):2719–2740. <https://doi.org/10.1021/jm901137j>
 19. Brus B, Kosak U, Turk S, Pislari A, Coquelle N, Kos J, Stojan J, Colletier JP, Gobec S (2014) Discovery, biological evaluation, and crystal structure of a novel nanomolar selective butyrylcholinesterase inhibitor. *J Med Chem* 57(19):8167–8179. <https://doi.org/10.1021/jm501195e>
 20. Messer WS Jr, Rajeswaran WG, Cao Y, Zhang HJ, el-Assadi AA, Dockery C, Liske J, O'Brien J, Williams FE, Huang XP, Wroblewski ME, Nagy PI, Peseckis SM (2000) Design and development of selective muscarinic agonists for the treatment of Alzheimer's disease: characterization of tetrahydropyrimidine derivatives and development of new approaches for improved affinity and selectivity for M1 receptors. *Pharm Acta Helv* 74(2–3):135–140. [https://doi.org/10.1016/S0031-6865\(99\)00026-6](https://doi.org/10.1016/S0031-6865(99)00026-6)
 21. Nordvall G, Hacksell U (1993) Binding-site modeling of the muscarinic m1 receptor: a combination of homology-based and indirect approaches. *J Med Chem* 36(8):967–976. <https://doi.org/10.1021/jm00060a003>
 22. Messer WS Jr, Abuh YF, Liu Y, Periyasamy S, Ngur DO, Edgar MA, El-Assadi AA, Sbeih S, Dunbar PG, Roknich S, Rho T, Fang Z, Ojo B, Zhang H, Huzl JJ 3rd, Nagy PI (1997) Synthesis and biological characterization of 1,4,5,6-tetrahydropyrimidine and 2-amino-3,4,5,6-tetrahydropyrimidine derivatives as selective m1 agonists. *J Med Chem* 40(8):1230–1246. <https://doi.org/10.1021/jm960467d>
 23. Niu YY, Yang LM, Deng KM, Yao JH, Zhu L, Chen CY, Zhang M, Zhou JE, Shen TX, Chen HZ, Lu Y (2007) Quantitative structure-selectivity relationship for M2 selectivity between M1 and M2 of piperidinyl piperidine derivatives as muscarinic antagonists. *Bioorg Med Chem Lett* 17(8):2260–2266. <https://doi.org/10.1016/j.bmcl.2007.01.058>
 24. Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 110(18):5959–5967. <https://doi.org/10.1021/ja00226a005>

25. Wang Y, Chackalamannil S, Hu Z, Clader JW, Greenlee W, Billard W, Binch H, Crosby G, Ruperto V, Duffy RA, McQuade R, Lachowicz JE (2000) Design and synthesis of piperidinyl piperidine analogues as potent and selective M2 muscarinic receptor antagonists. *Bioorg Med Chem Lett* 10(20):2247–2250. [https://doi.org/10.1016/s0960-894x\(00\)00457-1](https://doi.org/10.1016/s0960-894x(00)00457-1)
26. Nicolotti O, Pellegrini-Calace M, Altomar C, Carotti A, Carrieri A, Sanz F (2002) Ligands of neuronal nicotinic acetylcholine receptor (nAChR): inferences from the Hansch and 3-D quantitative structure-activity relationship (QSAR) models. *Curr Med Chem* 9(1):1–29. <https://doi.org/10.2174/0929867023371463>
27. Wei DQ, Sirois S, Du QS, Arias HR, Chou KC (2005) Theoretical studies of Alzheimer's disease drug candidate 3-[(2,4-dimethoxy)benzylidene]-anabaseine (GTS-21) and its derivatives. *Biochem Biophys Res Commun* 338(2):1059–1064. <https://doi.org/10.1016/j.bbrc.2005.10.047>
28. Brejc K, van Dijk WJ, Klaassen RV, Schuurmans M, van Der Oost J, Smit AB, Sixma TK (2001) Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* 411(6835):269–276. <https://doi.org/10.1038/35077011>
29. Kombo DC, Mazurov AA, Strachan JP, Bencherif M (2013) Computational studies of novel carbonyl-containing diazabicyclic ligands interacting with alpha4beta2 nicotinic acetylcholine receptor (nAChR) reveal alternative binding modes. *Bioorg Med Chem Lett* 23(18):5105–5113. <https://doi.org/10.1016/j.bmcl.2013.07.028>
30. Hansen SB, Sulzenbacher G, Huxford T, Marchot P, Taylor P, Bourne Y (2005) Structures of Aplysia AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *EMBO J* 24(20):3635–3646. <https://doi.org/10.1038/sj.emboj.7600828>
31. Catto M, Nicolotti O, Leonetti F, Carotti A, Favia AD, Soto-Otero R, Méndez-Alvarez E, Carotti A (2006) Structural insights into monoamine oxidase inhibitory potency and selectivity of 7-substituted coumarins from ligand-and target-based approaches. *J Med Chem* 49(16):4912–4925
32. Speck-Planche A, Kleandrova V (2012) QSAR and molecular docking techniques for the discovery of potent monoamine oxidase B inhibitors: computer-aided generation of new rasagiline bioisosteres. *Curr Top Med Chem* 12(16):1734–1747. <https://doi.org/10.2174/1568026611209061734>
33. Di Pietro O, Alencar N, Esteban G, Viayna E, Szalaj N, Vazquez J, Juarez-Jimenez J, Sola I, Perez B, Sole M, Unzeta M, Munoz-Torrero D, Luque FJ (2016) Design, synthesis and biological evaluation of N-methyl-N-[(1,2,3-triazol-4-yl)alkyl]propargylamines as novel monoamine oxidase B inhibitors. *Bioorg Med Chem* 24(20):4835–4854. <https://doi.org/10.1016/j.bmc.2016.06.045>
34. Esteban G, Allan J, Samadi A, Mattevi A, Unzeta M, Marco-Contelles J, Binda C, Ramsay RR (2014) Kinetic and structural analysis of the irreversible inhibition of human monoamine oxidases by ASS234, a multi-target compound designed for use in Alzheimer's disease. *Biochim Biophys Acta* 1844(6):1104–1110. <https://doi.org/10.1016/j.bbapap.2014.03.006>
35. De Colibus L, Li M, Binda C, Lustig A, Edmondson DE, Mattevi A (2005) Three-dimensional structure of human monoamine oxidase A (MAO A): relation to the structures of rat MAO A and human MAO B. *Proc Natl Acad Sci U S A* 102(36):12684–12689. <https://doi.org/10.1073/pnas.0505975102>
36. Hoang VH, Tran PT, Cui M, Ngo VT, Ann J, Park J, Lee J, Choi K, Cho H, Kim H, Ha HJ, Hong HS, Choi S, Kim YH (2017) Discovery of potent human glutaminyl cyclase inhibitors as anti-Alzheimer's agents based on rational design. *J Med Chem* 60(6):2573–2590. <https://doi.org/10.1021/acs.jmedchem.7b00098>
37. Huang KF, Liaw SS, Huang WL, Chia CY, Lo YC, Chen YL, Wang AH (2011) Structures of human Golgi-resident glutaminyl cyclase and its complexes with inhibitors reveal a large loop movement upon inhibitor binding. *J Biol Chem* 286(14):12439–12449. <https://doi.org/10.1074/jbc.M110.208595>
38. Zhao L, Brinton RD (2006) Select estrogens within the complex formulation of conjugated equine estrogens (Premarin®) are protective against neurodegenerative insults: implications for a composition of estrogen therapy to promote neuronal function and prevent Alzheimer's disease. *BMC Neurosci* 7(1):24
39. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engstrom O, Ohman L, Greene GL, Gustafsson JA, Carlquist M (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* 389(6652):753–758. <https://doi.org/10.1038/39645>

40. Cerpa W, Godoy JA, Alfaro I, Farias GG, Metcalfe MJ, Fuentealba R, Bonansco C, Inestrosa NC (2008) Wnt-7a modulates the synaptic vesicle cycle and synaptic transmission in hippocampal neurons. *J Biol Chem* 283(9):5918–5927. <https://doi.org/10.1074/jbc.M705943200>
41. Hanger DP, Noble W (2011) Functional implications of glycogen synthase kinase-3-mediated tau phosphorylation. *Int J Alzheimers Dis* 2011:352805. <https://doi.org/10.4061/2011/352805>
42. Cadigan KM, Waterman ML (2012) TCF/LEFs and Wnt signaling in the nucleus. *Cold Spring Harb Perspect Biol* 4(11). <https://doi.org/10.1101/cshperspect.a007906>
43. Zorn AM (2001) Wnt signalling: antagonistic Dickkops. *Curr Biol* 11(15):R592–R595. [https://doi.org/10.1016/S0960-9822\(01\)00360-8](https://doi.org/10.1016/S0960-9822(01)00360-8)
44. Rosi MC, Luccarini I, Grossi C, Fiorentini A, Spillantini MG, Prisco A, Scali C, Gianfriddo M, Caricasole A, Terstappen GC, Casamenti F (2010) Increased Dickkopf-1 expression in transgenic mouse models of neurodegenerative disease. *J Neurochem* 112(6):1539–1551. <https://doi.org/10.1111/j.1471-4159.2009.06566.x>
45. Caricasole A, Copani A, Caraci F, Aronica E, Rozemuller AJ, Caruso A, Storto M, Gaviraghi G, Terstappen GC, Nicoletti F (2004) Induction of Dickkopf-1, a negative modulator of the Wnt pathway, is associated with neuronal degeneration in Alzheimer's brain. *J Neurosci* 24(26):6021–6027. <https://doi.org/10.1523/JNEUROSCI.1381-04.2004>
46. Glantschnig H, Hampton RA, Lu P, Zhao JZ, Vitelli S, Huang L, Haytko P, Cusick T, Ireland C, Jarantow SW, Ernst R, Wei N, Nantermet P, Scott KR, Fisher JE, Talamo F, Orsatti L, Reszka AA, Sandhu P, Kimmel D, Flores O, Strohl W, An Z, Wang F (2010) Generation and selection of novel fully human monoclonal antibodies that neutralize Dickkopf-1 (DKK1) inhibitory function in vitro and increase bone mass in vivo. *J Biol Chem* 285(51):40135–40147. <https://doi.org/10.1074/jbc.M110.166892>
47. Mpousis S, Thysiadis S, Avramidis N, Katsamakas S, Efthimiopoulos S, Sarli V (2016) Synthesis and evaluation of galloyanine dyes as potential agents for the treatment of Alzheimer's disease and related neurodegenerative tauopathies. *Eur J Med Chem* 108:28–38. <https://doi.org/10.1016/j.ejmech.2015.11.024>
48. Thysiadis S, Mpousis S, Avramidis N, Katsamakas S, Balomenos A, Remelli R, Efthimiopoulos S, Sarli V (2016) Discovery of novel phenoxazinone derivatives as DKK1/LRP6 interaction inhibitors: synthesis, biological evaluation and structure-activity relationships. *Bioorg Med Chem* 24(5):1014–1022. <https://doi.org/10.1016/j.bmc.2016.01.025>
49. Cheng Z, Biechele T, Wei Z, Morrone S, Moon RT, Wang L, Xu W (2011) Crystal structures of the extracellular domain of LRP6 and its complex with DKK1. *Nat Struct Mol Biol* 18(11):1204–1210 <http://www.nature.com/nsmb/journal/v18/n11/abs/nsmb.2139.html#supplementary-information>
50. Kang J, Lemaire HG, Unterbeck A, Salbaum JM, Masters CL, Grzeschik KH, Multhaup G, Beyreuther K, Muller-Hill B (1987) The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature* 325(6106):733–736. <https://doi.org/10.1038/325733a0>
51. Mattson MP (2004) Pathways towards and away from Alzheimer's disease. *Nature* 430(7000):631–639. <https://doi.org/10.1038/nature02621>
52. Citron M, Teplow DB, Selkoe DJ (1995) Generation of amyloid β protein from its precursor is sequence specific. *Neuron* 14(3):661–670
53. Zhang YW, Thompson R, Zhang H, Xu H (2011) APP processing in Alzheimer's disease. *Mol Brain* 4(1):3. <https://doi.org/10.1186/1756-6606-4-3>
54. Gundersen E, Fan K, Haas K, Hury D, Steven Jacobsen J, Kreft A, Martone R, Mayer S, Sonnenberg-Reines J, Sun SC, Zhou H (2005) Molecular-modeling based design, synthesis, and activity of substituted piperidines as gamma-secretase inhibitors. *Bioorg Med Chem Lett* 15(7):1891–1894. <https://doi.org/10.1016/j.bmcl.2005.02.006>
55. Zhu YP, Xiao K, Yu HP, Ma LP, Xiong B, Zhang HY, Wang X, Li JY, Li J, Shen JK (2009) Discovery of potent beta-secretase (bace-1) inhibitors by the synthesis of isophthalamide-containing hybrids. *Acta Pharmacol Sin* 30(2):259–269. <https://doi.org/10.1038/aps.2008.26>
56. Coburn CA, Stachel SJ, Li YM, Rush DM, Steele TG, Chen-Dodson E, Holloway MK, Xu M, Huang Q, Lai MT, DiMuzio J, Crouthamel MC, Shi XP, Sardana V, Chen Z, Munshi S, Kuo L, Makara GM, Annis DA, Tadikonda PK, Nash HM, Vacca JP, Wang T (2004) Identification of a small molecule nonpeptide active site beta-secretase

- inhibitor that displays a nontraditional binding mode for aspartyl proteases. *J Med Chem* 47(25):6117–6119. <https://doi.org/10.1021/jm049388p>
57. Stachel SJ, Coburn CA, Steele TG, Jones KG, Loutzenhiser EF, Gregro AR, Rajapakse HA, Lai MT, Crouthamel MC, Xu M, Tugusheva K, Lineberger JE, Pietrak BL, Espeseth AS, Shi XP, Chen-Dodson E, Holloway MK, Munshi S, Simon AJ, Kuo L, Vacca JP (2004) Structure-based design of potent and selective cell-permeable inhibitors of human beta-secretase (BACE-1). *J Med Chem* 47(26):6447–6450. <https://doi.org/10.1021/jm049379g>
 58. Al-Tel TH, Semreen MH, Al-Qawasmeh RA, Schmidt MF, El-Awadi R, Ardah M, Zaarour R, Rao SN, El-Agnaf O (2011) Design, synthesis, and qualitative structure–activity evaluations of novel β -Secretase inhibitors as potential Alzheimer’s drug leads. *J Med Chem* 54(24):8373–8385
 59. Stachel SJ, Coburn CA, Steele TG, Crouthamel MC, Pietrak BL, Lai MT, Holloway MK, Munshi SK, Graham SL, Vacca JP (2006) Conformationally biased P3 amide replacements of beta-secretase inhibitors. *Bioorg Med Chem Lett* 16(3):641–644. <https://doi.org/10.1016/j.bmcl.2005.10.032>
 60. Ajmani S, Janardhan S, Viswanadhan VN (2013) Toward a general predictive QSAR model for gamma-secretase inhibitors. *Mol Divers* 17(3):421–434. <https://doi.org/10.1007/s11030-013-9441-2>
 61. Semighini EP (2015) In silico design of beta-secretase inhibitors in Alzheimer’s disease. *Chem Biol Drug Des* 86(3):284–290. <https://doi.org/10.1111/cbdd.12492>
 62. Clarke B, Demont E, Dingwall C, Dunsdon R, Faller A, Hawkins J, Hussain I, MacPherson D, Maile G, Matico R, Milner P, Mosley J, Naylor A, O’Brien A, Redshaw S, Riddell D, Rowland P, Soleil V, Smith KJ, Stanway S, Stemp G, Sweitzer S, Theobald P, Vesey D, Walter DS, Ward J, Wayne G (2008) BACE-1 inhibitors part 2: identification of hydroxy ethylamines (HEAs) with reduced peptidic character. *Bioorg Med Chem Lett* 18(3):1017–1021. <https://doi.org/10.1016/j.bmcl.2007.12.019>
 63. Charrier N, Clarke B, Cutler L, Demont E, Dingwall C, Dunsdon R, East P, Hawkins J, Howes C, Hussain I, Jeffrey P, Maile G, Matico R, Mosley J, Naylor A, O’Brien A, Redshaw S, Rowland P, Soleil V, Smith KJ, Sweitzer S, Theobald P, Vesey D, Walter DS, Wayne G (2008) Second generation of hydroxyethylamine BACE-1 inhibitors: optimizing potency and oral bioavailability. *J Med Chem* 51(11):3313–3317. <https://doi.org/10.1021/jm800138h>
 64. Charrier N, Clarke B, Demont E, Dingwall C, Dunsdon R, Hawkins J, Hubbard J, Hussain I, Maile G, Matico R, Mosley J, Naylor A, O’Brien A, Redshaw S, Rowland P, Soleil V, Smith KJ, Sweitzer S, Theobald P, Vesey D, Walter DS, Wayne G (2009) Second generation of BACE-1 inhibitors. Part 2: Optimisation of the non-prime side substituent. *Bioorg Med Chem Lett* 19(13):3669–3673. <https://doi.org/10.1016/j.bmcl.2009.03.150>
 65. Charrier N, Clarke B, Cutler L, Demont E, Dingwall C, Dunsdon R, Hawkins J, Howes C, Hubbard J, Hussain I, Maile G, Matico R, Mosley J, Naylor A, O’Brien A, Redshaw S, Rowland P, Soleil V, Smith KJ, Sweitzer S, Theobald P, Vesey D, Walter DS, Wayne G (2009) Second generation of BACE-1 inhibitors. Part 1: The need for improved pharmacokinetics. *Bioorg Med Chem Lett* 19(13):3664–3668. <https://doi.org/10.1016/j.bmcl.2009.03.165>
 66. Beswick P, Charrier N, Clarke B, Demont E, Dingwall C, Dunsdon R, Faller A, Gleave R, Hawkins J, Hussain I, Johnson CN, MacPherson D, Maile G, Matico R, Milner P, Mosley J, Naylor A, O’Brien A, Redshaw S, Riddell D, Rowland P, Skidmore J, Soleil V, Smith KJ, Stanway S, Stemp G, Stuart A, Sweitzer S, Theobald P, Vesey D, Walter DS, Ward J, Wayne G (2008) BACE-1 inhibitors part 3: identification of hydroxy ethylamines (HEAs) with nanomolar potency in cells. *Bioorg Med Chem Lett* 18(3):1022–1026. <https://doi.org/10.1016/j.bmcl.2007.12.020>
 67. Edraki N, Firuzi O, Fatahi Y, Mahdavi M, Asadi M, Emami S, Divsalar K, Miri R, Iraj A, Khoshneviszadeh M (2015) N-(2-(Piperazin-1-yl) phenyl) arylamide derivatives as β -secretase (BACE1) inhibitors: simple synthesis by Ugi four-component reaction and biological evaluation. *Arch Pharm* 348(5):330–337
 68. Zeng H, Wu X (2016) Alzheimer’s disease drug development based on computer-aided drug design. *Eur J Med Chem* 121:851–863. <https://doi.org/10.1016/j.ejmech.2015.08.039>
 69. Hernandez-Rodriguez M, Correa-Basurto J, Gutierrez A, Vitorica J, Rosales-Hernandez MC (2016) Asp32 and Asp228 determine the selective inhibition of BACE1 as shown by docking and molecular dynamics simulations. *Eur J Med Chem* 124:1142–1154.

- <https://doi.org/10.1016/j.ejmech.2016.08.028>
70. Iserloh U, Wu Y, Cumming JN, Pan J, Wang LY, Stamford AW, Kennedy ME, Kuvelkar R, Chen X, Parker EM, Strickland C, Voigt J (2008) Potent pyrrolidine- and piperidine-based BACE-1 inhibitors. *Bioorg Med Chem Lett* 18(1):414–417. <https://doi.org/10.1016/j.bmcl.2007.10.116>
 71. Ostermann N, Eder J, Eidhoff U, Zink F, Hassiepen U, Worpenberg S, Maibaum J, Simic O, Hommel U, Gerhartz B (2006) Crystal structure of human BACE2 in complex with a hydroxyethylamine transition-state inhibitor. *J Mol Biol* 355(2):249–261. <https://doi.org/10.1016/j.jmb.2005.10.027>
 72. Lee AY, Gulnik SV, Erickson JW (1998) Conformational switching in an aspartic proteinase. *Nat Struct Biol* 5(10):866–871. <https://doi.org/10.1038/2306>
 73. Tarazi H, Odeh RA, Al-Qawasmeh R, Yousef IA, Voelter W, Al-Tel TH (2017) Design, synthesis and SAR analysis of potent BACE1 inhibitors: possible lead drug candidates for Alzheimer's disease. *Eur J Med Chem* 125:1213–1224. <https://doi.org/10.1016/j.ejmech.2016.11.021>
 74. Coburn CA, Stachel SJ, Jones KG, Steele TG, Rush DM, DiMuzio J, Pietrak BL, Lai MT, Huang Q, Lineberger J, Jin L, Munshi S, Katharine Holloway M, Espeseth A, Simon A, Hazuda D, Graham SL, Vacca JP (2006) BACE-1 inhibition by a series of psi [CH₂NH] reduced amide isosteres. *Bioorg Med Chem Lett* 16(14):3635–3638. <https://doi.org/10.1016/j.bmcl.2006.04.076>
 75. Steele TG, Hills ID, Nomland AA, de Leon P, Allison T, McGaughey G, Colussi D, Tugusheva K, Haugabook SJ, Espeseth AS, Zuck P, Graham SL, Stachel SJ (2009) Identification of a small molecule beta-secretase inhibitor that binds without catalytic aspartate engagement. *Bioorg Med Chem Lett* 19(1):17–20. <https://doi.org/10.1016/j.bmcl.2008.11.027>
 76. Prade E, Bittner HJ, Sarkar R, Lopez Del Amo JM, Althoff-Ospelt G, Multhaup G, Hildebrand PW, Reif B (2015) Structural mechanism of the interaction of Alzheimer disease Abeta fibrils with the non-steroidal anti-inflammatory drug (NSAID) sulindac sulfide. *J Biol Chem* 290(48):28737–28745. <https://doi.org/10.1074/jbc.M115.675215>
 77. Paravastu AK, Leapman RD, Yau WM, Tycko R (2008) Molecular structural basis for polymorphism in Alzheimer's beta-amyloid fibrils. *Proc Natl Acad Sci U S A* 105(47):18349–18354. <https://doi.org/10.1073/pnas.0806270105>
 78. de Almeida JP, Saldanha C (2010) Nonneuronal cholinergic system in human erythrocytes: biological role and clinical relevance. *J Membr Biol* 234(3):227–234. <https://doi.org/10.1007/s00232-010-9250-9>
 79. Brimijoin S (1983) Molecular forms of acetylcholinesterase in brain, nerve and muscle: nature, localization and dynamics. *Prog Neurobiol* 21(4):291–322
 80. Heller M, Hanahan DJ (1972) Human erythrocyte membrane bound enzyme acetylcholinesterase. *Biochim Biophys Acta* 255(1):251–272
 81. Dvir H, Silman I, Harel M, Rosenberry TL, Sussman JL (2010) Acetylcholinesterase: from 3D structure to function. *Chem Biol Interact* 187(1–3):10–22. <https://doi.org/10.1016/j.cbi.2010.01.042>
 82. Silman I, Sussman JL (2008) Acetylcholinesterase: how is structure related to function? *Chem Biol Interact* 175(1–3):3–10
 83. Lane RM, Kivipelto M, Greig NH (2004) Acetylcholinesterase and its inhibition in Alzheimer disease. *Clin Neuropharmacol* 27(3):141–149
 84. Lane RM, Potkin SG, Enz A (2006) Targeting acetylcholinesterase and butyrylcholinesterase in dementia. *Int J Neuropsychopharmacol* 9(1):101–124. <https://doi.org/10.1017/S1461145705005833>
 85. Recanatini M, Cavalli A, Hansch C (1997) A comparative QSAR analysis of acetylcholinesterase inhibitors currently studied for the treatment of Alzheimer's disease. *Chem Biol Interact* 105(3):199–228. [https://doi.org/10.1016/S0009-2797\(97\)00047-1](https://doi.org/10.1016/S0009-2797(97)00047-1)
 86. Sippl W, Contreras JM, Parrot I, Rival YM, Wermuth CG (2001) Structure-based 3D QSAR and design of novel acetylcholinesterase inhibitors. *J Comput Aided Mol Des* 15(5):395–410
 87. Harel M, Schalk I, Ehret-Sabatier L, Bouet F, Goeldner M, Hirth C, Axelsen PH, Silman I, Sussman JL (1993) Quaternary ligand binding to aromatic residues in the active-site gorge of acetylcholinesterase. *Proc Natl Acad Sci U S A* 90(19):9031–9035
 88. Ravelli RB, Raves ML, Ren Z, Bourgeois D, Roth M, Kroon J, Silman I, Sussman JL (1998) Static Laue diffraction studies on acetylcholinesterase. *Acta Crystallogr D Biol Crystallogr* 54(Pt 6 Pt 2):1359–1366. <https://doi.org/10.1107/s0907444998005277>
 89. Raves ML, Harel M, Pang Y-P, Silman I, Kozikowski AP, Sussman JL (1997) Structure

- of acetylcholinesterase complexed with the nootropic alkaloid, (–)-huperzine A. *Nat Struct Mol Biol* 4(1):57–63. <https://doi.org/10.1038/nsb0197-57>
90. Kosak U, Brus B, Knez D, Zakelj S, Trontelj J, Pisljar A, Sink R, Jukic M, Zivin M, Podkova A, Nachon F, Brazzolotto X, Stojan J, Kos J, Coquelle N, Salat K, Colletier JP, Gobec S (2017) The magic of crystal structure-based inhibitor optimization: development of a butyrylcholinesterase inhibitor with picomolar affinity and in vivo activity. *J Med Chem*. <https://doi.org/10.1021/acs.jmedchem.7b01086>
 91. da Silva CH, Campo VL, Carvalho I, Taft CA (2006) Molecular modeling, docking and ADMET studies applied to the design of a novel hybrid for treatment of Alzheimer's disease. *J Mol Graph Model* 25(2):169–175. <https://doi.org/10.1016/j.jmgm.2005.12.002>
 92. Kryger G, Silman I, Sussman JL (1999) Structure of acetylcholinesterase complexed with E2020 (Aricept (R)): implications for the design of new anti-Alzheimer drugs. *Structure* 7(3):297–307. [https://doi.org/10.1016/S0969-2126\(99\)80040-9](https://doi.org/10.1016/S0969-2126(99)80040-9)
 93. Alcaro S, Arcone R, Vecchio I, Ortuso F, Gallelli A, Pasceri R, Procopio A, Iannone M (2007) Molecular modelling and enzymatic studies of acetylcholinesterase and butyrylcholinesterase recognition with paraquat and related compounds. *SAR QSAR Environ Res* 18(5–6):595–602. <https://doi.org/10.1080/10629360701428433>
 94. da Silva CH, Carvalho I, Taft CA (2007) Virtual screening, molecular interaction field, molecular dynamics, docking, density functional, and ADMET properties of novel AChE inhibitors in Alzheimer's disease. *J Biomol Struct Dyn* 24(6):515–524. <https://doi.org/10.1080/07391102.2007.10507140>
 95. Fang L, Appenroth D, Decker M, Kiehntopf M, Lupp A, Peng SX, Fleck C, Zhang YH, Lehmann JC (2008) NO-donating tacrine hybrid compounds improve scopolamine-induced cognition impairment and show less hepatotoxicity. *J Med Chem* 51(24):7666–7669. <https://doi.org/10.1021/jm801131a>
 96. Rydberg EH, Brumshtein B, Greenblatt HM, Wong DM, Shaya D, Williams LD, Carlier PR, Pang YP, Silman I, Sussman JL (2006) Complexes of alkylene-linked tacrine dimers with *Torpedo californica* acetylcholinesterase: binding of Bis5-tacrine produces a dramatic rearrangement in the active-site gorge. *J Med Chem* 49(18):5491–5500. <https://doi.org/10.1021/jm060164b>
 97. Badran MM, Abdel Hakeem M, Abuel-Maaty SM, El-Malah A, Abdel Salam RM (2010) Design, synthesis, and molecular-modeling study of aminothienopyridine analogues of tacrine for Alzheimer's disease. *Arch Pharm (Weinheim)* 343(10):590–601. <https://doi.org/10.1002/ardp.200900226>
 98. Geromichalos GD, Lamari FN, Papandreou MA, Trafalis DT, Margarity M, Papageorgiou A, Sinakos Z (2012) Saffron as a source of novel acetylcholinesterase inhibitors: molecular docking and in vitro enzymatic studies. *J Agric Food Chem* 60(24):6131–6138. <https://doi.org/10.1021/jf300589c>
 99. El-Malah A, Gedawy EM, Kassab AE, Salam RMA (2014) Novel tacrine analogs as potential cholinesterase inhibitors in Alzheimer's disease. *Arch Pharm* 347(2):96–103
 100. Nicolet Y, Lockridge O, Masson P, Fontecilla-Camps JC, Nachon F (2003) Crystal structure of human butyrylcholinesterase and of its complexes with substrate and products. *J Biol Chem* 278(42):41141–41147. <https://doi.org/10.1074/jbc.M210241200>
 101. Arab S, Sadat-Ebrahimi SE, Mohammadi-Khanaposhtani M, Moradi A, Nadri H, Mahdavi M, Moghimi S, Asadi M, Firoozpour L, Pirali-Hamedani M, Shafiee A, Foroumadi A (2015) Synthesis and evaluation of chroman-4-one linked to N-benzyl pyridinium derivatives as new acetylcholinesterase inhibitors. *Arch Pharm* 348(9):643–649. <https://doi.org/10.1002/ardp.201500149>
 102. Liu Z, Fang L, Zhang H, Gou S, Chen L (2017) Design, synthesis and biological evaluation of multifunctional tacrine-curcumin hybrids as new cholinesterase inhibitors with metal ions-chelating and neuroprotective property. *Bioorg Med Chem* 25(8):2387–2398. <https://doi.org/10.1016/j.bmc.2017.02.049>
 103. Nachon F, Carletti E, Ronco C, Trovaslet M, Nicolet Y, Jean L, Renard PY (2013) Crystal structures of human cholinesterases in complex with huprine W and tacrine: elements of specificity for anti-Alzheimer's drugs targeting acetyl- and butyryl-cholinesterase. *Biochem J* 453(3):393–399. <https://doi.org/10.1042/BJ20130013>
 104. Mehrabi F, Pourshojaei Y, Moradi A, Sharifzadeh M, Khosravani L, Sabourian R, Rahmani-Nezhad S, Mohammadi-Khanaposhtani M, Mahdavi M, Asadipour A,

- Rahimi HR, Moghimi S, Foroumadi A (2017) Design, synthesis, molecular modeling and anticholinesterase activity of benzylidene-benzofuran-3-ones containing cyclic amine side chain. *Future Med Chem* 9 (7):659–671. <https://doi.org/10.4155/fmc-2016-0237>
105. da Silva Goncalves A, Franca TC, Vital de Oliveira O (2016) Computational studies of acetylcholinesterase complexed with fullerene derivatives: a new insight for Alzheimer disease treatment. *J Biomol Struct Dyn* 34 (6):1307–1316. <https://doi.org/10.1080/07391102.2015.1077345>
106. Carletti E, Colletier JP, Dupeux F, Trovaslet M, Masson P, Nachon F (2010) Structural evidence that human acetylcholinesterase is inhibited by tabun ages through O-dealkylation. *J Med Chem* 53 (10):4002–4008. <https://doi.org/10.1021/jm901853b>
107. Basiri A, Xiao M, McCarthy A, Dutta D, Byrareddy SN, Conda-Sheridan M (2017) Design and synthesis of new piperidone grafted acetylcholinesterase inhibitors. *Bioorg Med Chem Lett* 27(2):228–231. <https://doi.org/10.1016/j.bmcl.2016.11.065>
108. Yu Q, Holloway HW, Flippen-Anderson JL, Hoffman B, Brossi A, Greig NH (2001) Methyl analogues of the experimental Alzheimer drug phenserine: synthesis and structure/activity relationships for acetyl- and butyrylcholinesterase inhibitory action. *J Med Chem* 44(24):4062–4071
109. Camps P, Formosa X, Galdeano C, Munoz-Torrero D, Ramirez L, Gómez E, Isambert N, Lavilla R, Badia A, Clos MV (2009) Pyrano [3, 2-c] quinoline-6-chlorotacrine hybrids as a novel family of acetylcholinesterase- and β -amyloid-directed anti-Alzheimer compounds. *J Med Chem* 52(17):5365–5379
110. Bourne Y, Kolb HC, Radic Z, Sharpless KB, Taylor P, Marchot P (2004) Freeze-frame inhibitor captures acetylcholinesterase in a unique conformation. *Proc Natl Acad Sci U S A* 101(6):1449–1454. <https://doi.org/10.1073/pnas.0308206100>
111. Haviv H, Wong DM, Greenblatt HM, Carlier PR, Pang YP, Silman I, Sussman JL (2005) Crystal packing mediates enantioselective ligand recognition at the peripheral site of acetylcholinesterase. *J Am Chem Soc* 127 (31):11029–11036. <https://doi.org/10.1021/ja051765f>
112. Ul-Haq Z, Khan W, Kalsoom S, Ansari FL (2010) In silico modeling of the specific inhibitory potential of thiophene-2,3-dihydro-1,5-benzothiazepine against BChE in the formation of beta-amyloid plaques associated with Alzheimer's disease. *Theor Biol Med Model* 7(1):22. <https://doi.org/10.1186/1742-4682-7-22>
113. Hamulakova S, Janovec L, Hrabínova M, Kristian P, Kuca K, Banasova M, Imrich J (2012) Synthesis, design and biological evaluation of novel highly potent tacrine congeners for the treatment of Alzheimer's disease. *Eur J Med Chem* 55:23–31. <https://doi.org/10.1016/j.ejmech.2012.06.051>
114. Kryger G, Harel M, Giles K, Tokar L, Velan B, Lazar A, Kronman C, Barak D, Ariel N, Shafferman A, Silman I, Sussman JL (2000) Structures of recombinant native and E202Q mutant human acetylcholinesterase complexed with the snake-venom toxin fasciculín-II. *Acta Crystallogr D Biol Crystallogr* 56(Pt 11):1385–1394. <https://doi.org/10.1107/S0907444900010659>
115. Makhaeva GF, Radchenko EV, Baskin II, Palyulin VA, Richardson RJ, Zefirov NS (2012) Combined QSAR studies of inhibitor properties of O-phosphorylated oximes toward serine esterases involved in neurotoxicity, drug metabolism and Alzheimer's disease. *SAR QSAR Environ Res* 23 (7–8):627–647. <https://doi.org/10.1080/1062936X.2012.679690>
116. Bourne Y, Radic Z, Sulzenbacher G, Kim E, Taylor P, Marchot P (2006) Substrate and product trafficking through the active center gorge of acetylcholinesterase analyzed by crystallography and equilibrium binding. *J Biol Chem* 281(39):29256–29267. <https://doi.org/10.1074/jbc.M603018200>
117. Ozturan Ozer E, Tan OU, Ozadali K, Kucukkilinc T, Balkan A, Ucar G (2013) Synthesis, molecular modeling and evaluation of novel N'-2-(4-benzylpiperidin-/piperazin-1-yl)acylhydrazone derivatives as dual inhibitors for cholinesterases and A β aggregation. *Bioorg Med Chem Lett* 23(2):440–443. <https://doi.org/10.1016/j.bmcl.2012.11.064>
118. Geldmacher DS (2004) Donepezil (Aricept®) for treatment of Alzheimer's disease and other dementing conditions. *Expert Rev Neurother* 4(1):5–16
119. Maggi N, Pasqualucci CR, Ballotta R, Sensi P (1966) Rifampicin: a new orally active rifamycin. *Chemotherapy* 11(5):285–292. <https://doi.org/10.1159/000220462>
120. Li RS, Wang XB, Hu XJ, Kong LY (2013) Design, synthesis and evaluation of flavonoid derivatives as potential multifunctional acetylcholinesterase inhibitors against Alzheimer's disease. *Bioorg Med Chem Lett* 23

- (9):2636–2641. <https://doi.org/10.1016/j.bmc.2013.02.095>
121. Luo W, Su YB, Hong C, Tian RG, Su LP, Wang YQ, Li Y, Yue JJ, Wang CJ (2013) Design, synthesis and evaluation of novel 4-dimethylamine flavonoid derivatives as potential multi-functional anti-Alzheimer agents. *Bioorg Med Chem* 21 (23):7275–7282. <https://doi.org/10.1016/j.bmc.2013.09.061>
 122. Birks J, Grimley Evans J, Iakovidou V, Tsolaki M, Holt F (2000) Rivastigmine for Alzheimer's disease. *Cochrane Database Syst Rev* (4):CD001191
 123. Xie SS, Wang XB, Li JY, Yang L, Kong LY (2013) Design, synthesis and evaluation of novel tacrine-coumarin hybrids as multifunctional cholinesterase inhibitors against Alzheimer's disease. *Eur J Med Chem* 64:540–553. <https://doi.org/10.1016/j.ejmech.2013.03.051>
 124. Thiratmatrakul S, Yenjai C, Waiwut P, Vajragupta O, Reubroycharoen P, Tohda M, Boonyarat C (2014) Synthesis, biological evaluation and molecular modeling study of novel tacrine-carbazole hybrids as potential multifunctional agents for the treatment of Alzheimer's disease. *Eur J Med Chem* 75:21–30. <https://doi.org/10.1016/j.ejmech.2014.01.020>
 125. Colletier JP, Sanson B, Nachon F, Gabellieri E, Fattorusso C, Campiani G, Weik M (2006) Conformational flexibility in the peripheral site of *Torpedo californica* acetylcholinesterase revealed by the complex structure with a bifunctional inhibitor. *J Am Chem Soc* 128(14):4526–4527. <https://doi.org/10.1021/ja058683b>
 126. Qiang X, Sang Z, Yuan W, Li Y, Liu Q, Bai P, Shi Y, Ang W, Tan Z, Deng Y (2014) Design, synthesis and evaluation of genistein-O-alkylbenzylamines as potential multifunctional agents for the treatment of Alzheimer's disease. *Eur J Med Chem* 76:314–331. <https://doi.org/10.1016/j.ejmech.2014.02.045>
 127. Pudlo M, Luzet V, Ismaili L, Tomassoli I, Iutzler A, Refouvelet B (2014) Quinolone-benzylpiperidine derivatives as novel acetylcholinesterase inhibitor and antioxidant hybrids for Alzheimer disease. *Bioorg Med Chem* 22(8):2496–2507. <https://doi.org/10.1016/j.bmc.2014.02.046>
 128. Cheung J, Rudolph MJ, Burshteyn F, Cassidy MS, Gary EN, Love J, Franklin MC, Height JJ (2012) Structures of human acetylcholinesterase in complex with pharmacologically important ligands. *J Med Chem* 55 (22):10282–10286. <https://doi.org/10.1021/jm300871x>
 129. Hong C, Luo W, Yao D, Su YB, Zhang X, Tian RG, Wang CJ (2014) Novel aromatic-polyamine conjugates as cholinesterase inhibitors with notable selectivity toward butyrylcholinesterase. *Bioorg Med Chem* 22 (12):3213–3219. <https://doi.org/10.1016/j.bmc.2014.03.045>
 130. Lan JS, Xie SS, Li SY, Pan LF, Wang XB, Kong LY (2014) Design, synthesis and evaluation of novel tacrine-(beta-carboline) hybrids as multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem* 22(21):6089–6104. <https://doi.org/10.1016/j.bmc.2014.08.035>
 131. Li S-Y, Jiang N, Xie S-S, Wang KD, Wang X-B, Kong L-Y (2014) Design, synthesis and evaluation of novel tacrine-rhein hybrids as multifunctional agents for the treatment of Alzheimer's disease. *Org Biomol Chem* 12 (5):801–814
 132. Stoddard SV, Hamann MT, Wadkins RM (2014) Insights and ideas garnered from marine metabolites for development of dual-function acetylcholinesterase and amyloid-beta aggregation inhibitors. *Mar Drugs* 12 (4):2114–2131. <https://doi.org/10.3390/md12042114>
 133. Greenblatt HM, Kryger G, Lewis T, Silman I, Sussman JL (1999) Structure of acetylcholinesterase complexed with (–)-galanthamine at 2.3 Å resolution. *FEBS Lett* 463 (3):321–326. [https://doi.org/10.1016/s0014-5793\(99\)01637-3](https://doi.org/10.1016/s0014-5793(99)01637-3)
 134. Bautista-Aguilera OM, Esteban G, Bolea I, Nikolic K, Agbaba D, Moraleta I, Iriepa I, Samadi A, Soriano E, Unzeta M, Marco-Contelles J (2014) Design, synthesis, pharmacological evaluation, QSAR analysis, molecular modeling and ADMET of novel donepezil-indolyl hybrids as multipotent cholinesterase/monoamine oxidase inhibitors for the potential treatment of Alzheimer's disease. *Eur J Med Chem* 75:82–95. <https://doi.org/10.1016/j.ejmech.2013.12.028>
 135. Bourne Y, Grassi J, Bougis PE, Marchot P (1999) Conformational flexibility of the acetylcholinesterase tetramer suggested by X-ray crystallography. *J Biol Chem* 274 (43):30370–30376. <https://doi.org/10.1074/jbc.274.43.30370>
 136. Ngamelue MN, Homma K, Lockridge O, Asojo OA (2007) Crystallization and X-ray structure of full-length recombinant human butyrylcholinesterase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 63

- (Pt 9):723–727. <https://doi.org/10.1107/S1744309107037335>
137. Son SY, Ma J, Kondou Y, Yoshimura M, Yamashita E, Tsukihara T (2008) Structure of human monoamine oxidase A at 2.2-Å resolution: the control of opening the entry for substrates/inhibitors. *Proc Natl Acad Sci U S A* 105(15):5739–5744. <https://doi.org/10.1073/pnas.0710626105>
138. Binda C, Wang J, Pisani L, Caccia C, Carotti A, Salvati P, Edmondson DE, Mattevi A (2007) Structures of human monoamine oxidase B complexes with selective noncovalent inhibitors: safinamide and coumarin analogs. *J Med Chem* 50(23):5848–5852. <https://doi.org/10.1021/jm070677y>
139. Bautista-Aguilera OM, Esteban G, Chioua M, Nikolic K, Agbada D, Moraleda I, Iriepa I, Soriano E, Samadi A, Unzeta M, Marco-Contelles J (2014) Multipotent cholinesterase/monoamine oxidase inhibitors for the treatment of Alzheimer's disease: design, synthesis, biochemical evaluation, ADMET, molecular modeling, and QSAR analysis of novel donepezil-pyridyl hybrids. *Drug Des Devel Ther* 8:1893–1910. <https://doi.org/10.2147/DDDT.S69258>
140. Goyal M, Dhanjal JK, Goyal S, Tyagi C, Hamid R, Grover A (2014) Development of dual inhibitors against Alzheimer's disease using fragment-based QSAR and molecular docking. *Biomed Res Int* 2014:979606. <https://doi.org/10.1155/2014/979606>
141. Cheung J, Gary EN, Shiomi K, Rosenberry TL (2013) Structures of human acetylcholinesterase bound to dihydrotanshinone I and territrein B show peripheral site flexibility. *ACS Med Chem Lett* 4(11):1091–1096. <https://doi.org/10.1021/ml400304w>
142. Xie SS, Wang X, Jiang N, Yu W, Wang KD, Lan JS, Li ZR, Kong LY (2015) Multi-target tacrine-coumarin hybrids: cholinesterase and monoamine oxidase B inhibition properties against Alzheimer's disease. *Eur J Med Chem* 95:153–165. <https://doi.org/10.1016/j.ejmech.2015.03.040>
143. Xie SS, Lan JS, Wang XB, Jiang N, Dong G, Li ZR, Wang KD, Guo PP, Kong LY (2015) Multifunctional tacrine-trox hybrids for the treatment of Alzheimer's disease with cholinergic, antioxidant, neuroprotective and hepatoprotective properties. *Eur J Med Chem* 93:42–50. <https://doi.org/10.1016/j.ejmech.2015.01.058>
144. Luo XT, Wang CM, Liu Y, Huang ZG (2015) New multifunctional melatonin-derived benzylpyridinium bromides with potent cholinergic, antioxidant, and neuroprotective properties as innovative drugs for Alzheimer's disease. *Eur J Med Chem* 103:302–311. <https://doi.org/10.1016/j.ejmech.2015.08.052>
145. Sang Z, Qiang X, Li Y, Yuan W, Liu Q, Shi Y, Ang W, Luo Y, Tan Z, Deng Y (2015) Design, synthesis and evaluation of scutellarein-O-alkylamines as multifunctional agents for the treatment of Alzheimer's disease. *Eur J Med Chem* 94:348–366. <https://doi.org/10.1016/j.ejmech.2015.02.063>
146. Liu Q, Qiang X, Li Y, Sang Z, Tan Z, Deng Y (2015) Design, synthesis and evaluation of chromone-2-carboxamido-alkylbenzylamines as multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem* 23(5):911–923. <https://doi.org/10.1016/j.bmc.2015.01.042>
147. Bajda M, Jonczyk J, Malawska B, Czarna K, Girek M, Olszewska P, Sikora J, Mikiciuk-Olasik E, Skibinski R, Gumieniczek A, Szymanski P (2015) Synthesis, biological evaluation and molecular modeling of new tetrahydroacridine derivatives as potential multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem* 23(17):5610–5618. <https://doi.org/10.1016/j.bmc.2015.07.029>
148. Benchekroun M, Bartolini M, Egea J, Romero A, Soriano E, Pudlo M, Luzet V, Andrisano V, Jimeno ML, Lopez MG, Wehle S, Gharbi T, Refouvet B, de Andres L, Herrera-Arozamena C, Monti B, Bolognesi ML, Rodriguez-Franco MI, Decker M, Marco-Contelles J, Ismaili L (2015) Novel tacrine-grafted Ugi adducts as multipotent anti-Alzheimer drugs: a synthetic renewal in tacrine-ferulic acid hybrids. *ChemMedChem* 10(3):523–539. <https://doi.org/10.1002/cmdc.201402409>
149. Dominguez JL, Fernandez-Nieto F, Castro M, Catto M, Paleo MR, Porto S, Sardinia FJ, Brea JM, Carotti A, Villaverde MC, Sussman F (2015) Computer-aided structure-based design of multitarget leads for Alzheimer's disease. *J Chem Inf Model* 55(1):135–148. <https://doi.org/10.1021/ci500555g>
150. Hong L, Koelsch G, Lin X, Wu S, Terzyan S, Ghosh AK, Zhang XC, Tang J (2000) Structure of the protease domain of memapsin 2 (beta-secretase) complexed with inhibitor. *Science* 290(5489):150–153. <https://doi.org/10.1126/science.290.5489.150>
151. Wu MY, Esteban G, Brogi S, Shionoya M, Wang L, Campiani G, Unzeta M, Inokuchi T, Butini S, Marco-Contelles J (2016) Donepezil-like multifunctional

- agents: design, synthesis, molecular modeling and biological evaluation. *Eur J Med Chem* 121:864–879. <https://doi.org/10.1016/j.ejmech.2015.10.001>
152. Najafi Z, Saedi M, Mahdavi M, Sabourian R, Khanavi M, Tehrani MB, Moghadam FH, Edraki N, Karimpor-Razkenari E, Sharifzadeh M, Foroumadi A, Shafiee A, Akbarzadeh T (2016) Design and synthesis of novel anti-Alzheimer's agents: Acridine-chromenone and quinoline-chromenone hybrids. *Bioorg Chem* 67:84–94. <https://doi.org/10.1016/j.bioorg.2016.06.001>
153. Zhang C, Du QY, Chen LD, Wu WH, Liao SY, Yu LH, Liang XT (2016) Design, synthesis and evaluation of novel tacrine-multialkoxybenzene hybrids as multi-targeted compounds against Alzheimer's disease. *Eur J Med Chem* 116:200–209. <https://doi.org/10.1016/j.ejmech.2016.03.077>
154. Luo W, Wang T, Hong C, Yang YC, Chen Y, Cen J, Xie SQ, Wang CJ (2016) Design, synthesis and evaluation of 4-dimethylamine flavonoid derivatives as potential multifunctional anti-Alzheimer agents. *Eur J Med Chem* 122:17–26. <https://doi.org/10.1016/j.ejmech.2016.06.022>
155. Luo W, Chen Y, Wang T, Hong C, Chang LP, Chang CC, Yang YC, Xie SQ, Wang CJ (2016) Design, synthesis and evaluation of novel 7-aminoalkyl-substituted flavonoid derivatives with improved cholinesterase inhibitory activities. *Bioorg Med Chem* 24(4):672–680. <https://doi.org/10.1016/j.bmc.2015.12.031>
156. Wang ZM, Cai P, Liu QH, Xu DQ, Yang XL, Wu JJ, Kong LY, Wang XB (2016) Rational modification of donepezil as multifunctional acetylcholinesterase inhibitors for the treatment of Alzheimer's disease. *Eur J Med Chem* 123:282–297. <https://doi.org/10.1016/j.ejmech.2016.07.052>
157. Knez D, Brus B, Coquelle N, Sosic I, Sink R, Brazzolotto X, Mravljak J, Colletier JP, Gobec S (2015) Structure-based development of nitroxoline derivatives as potential multifunctional anti-Alzheimer agents. *Bioorg Med Chem* 23(15):4442–4452. <https://doi.org/10.1016/j.bmc.2015.06.010>
158. Xie SS, Lan JS, Wang X, Wang ZM, Jiang N, Li F, Wu JJ, Wang J, Kong LY (2016) Design, synthesis and biological evaluation of novel donepezil-coumarin hybrids as multi-target agents for the treatment of Alzheimer's disease. *Bioorg Med Chem* 24(7):1528–1539. <https://doi.org/10.1016/j.bmc.2016.02.023>
159. Koca M, Yerdelen KO, Anil B, Kasap Z, Sevindik H, Ozyurek I, Gunesacar G, Turkaydin K (2016) Design, synthesis and biological activity of 1H-indene-2-carboxamides as multi-targeted anti-Alzheimer agents. *J Enzyme Inhib Med Chem* 31(sup2):13–23
160. Wang Y, Sun Y, Guo Y, Wang Z, Huang L, Li X (2016) Dual functional cholinesterase and MAO inhibitors for the treatment of Alzheimer's disease: synthesis, pharmacological analysis and molecular modeling of homoiso-flavonoid derivatives. *J Enzyme Inhib Med Chem* 31(3):389–397. <https://doi.org/10.3109/14756366.2015.1024675>
161. Mohamed T, Rao PPN (2017) 2,4-Disubstituted quinazolines as amyloid- β aggregation inhibitors with dual cholinesterase inhibition and antioxidant properties: development and structure-activity relationship (SAR) studies. *Eur J Med Chem* 126:823–843. <https://doi.org/10.1016/j.ejmech.2016.12.005>
162. Panek D, Wieckowska A, Wichur T, Bajda M, Godyn J, Jonczyk J, Mika K, Janockova J, Soukup O, Knez D, Korabecny J, Gobec S, Malawska B (2017) Design, synthesis and biological evaluation of new phthalimide and saccharin derivatives with alicyclic amines targeting cholinesterases, beta-secretase and amyloid beta aggregation. *Eur J Med Chem* 125:676–695. <https://doi.org/10.1016/j.ejmech.2016.09.078>
163. Rueeger H, Lueoend R, Rogel O, Rondeau JM, Mobitz H, Machauer R, Jacobson L, Staufenbiel M, Desrayaud S, Neumann U (2012) Discovery of cyclic sulfone hydroxyethylamines as potent and selective beta-site APP-cleaving enzyme 1 (BACE1) inhibitors: structure-based design and in vivo reduction of amyloid beta-peptides. *J Med Chem* 55(7):3364–3386. <https://doi.org/10.1021/jm300069y>
164. Najafi Z, Mahdavi M, Saedi M, Karimpour-Razkenari E, Asatouri R, Vafadarnejad F, Moghadam FH, Khanavi M, Sharifzadeh M, Akbarzadeh T (2017) Novel tacrine-1,2,3-triazole hybrids: in vitro, in vivo biological evaluation and docking study of cholinesterase inhibitors. *Eur J Med Chem* 125:1200–1212. <https://doi.org/10.1016/j.ejmech.2016.11.008>
165. Sang Z, Qiang X, Li Y, Xu R, Cao Z, Song Q, Wang T, Zhang X, Liu H, Tan Z, Deng Y (2017) Design, synthesis and evaluation of scutellarein-O-acetamidoalkylbenzylamines as potential multifunctional agents for the treatment of Alzheimer's disease. *Eur J Med*

- Chem 135:307–323. <https://doi.org/10.1016/j.ejmech.2017.04.054>
166. Jameel E, Meena P, Maqbool M, Kumar J, Ahmed W, Mumtazuddin S, Tiwari M, Hoda N, Jayaram B (2017) Rational design, synthesis and biological screening of triazine-triazolopyrimidine hybrids as multitarget anti-Alzheimer agents. *Eur J Med Chem* 136:36–51. <https://doi.org/10.1016/j.ejmech.2017.04.064>
167. Sang Z, Pan W, Wang K, Ma Q, Yu L, Liu W (2017) Design, synthesis and biological evaluation of 3,4-dihydro-2(1H)-quinoline-O-alkylamine derivatives as new multipotent cholinesterase/monoamine oxidase inhibitors for the treatment of Alzheimer's disease. *Bioorg Med Chem* 25(12):3006–3017. <https://doi.org/10.1016/j.bmc.2017.03.070>



Design of Multi-target Directed Ligands as a Modern Approach for the Development of Innovative Drug Candidates for Alzheimer's Disease

Cindy Juliet Cristancho Ortiz, Matheus de Freitas Silva, Vanessa Silva Gontijo, Flávia Pereira Dias Viegas, Kris Simone Tranches Dias, and Claudio Viegas Jr.

Abstract

Alzheimer's disease (AD) is a complex neurodegenerative disorder with a multi-faceted pathogenesis. So far, the therapeutic paradigm “one-compound-one-target” has failed and despite enormous efforts to elucidate the pathophysiology of AD, the disease is still incurable, with all current medicines only being capable to slow up its progress and ameliorate the quality of life of the patients. The multiple factors involved in AD include amyloid aggregation to form insoluble neurotoxic plaques of A β , hyperphosphorylation of tau protein, oxidative stress, calcium imbalance, mitochondrial dysfunction, deterioration of synaptic transmission, and neuronal loss. These factors together accentuate changes in the central nervous system (CNS) homeostasis, starting a complex process of interconnected physiological damage, leading to cognitive and memory impairment and neuronal death. A recent approach for the rational design of new drug candidates, also called multi-target directed ligand (MTDL) approach, has gained increasing attention by many research groups, which have developed a variety of hybrid compounds acting simultaneously on diverse biological targets. In this chapter, we aimed to show some recent advances during the last decade and examples of the exploitation of MTDL approach in the rational design of novel drug candidate prototypes for the treatment of AD.

Keywords Alzheimer's disease, Multifunctional drugs, Multi-target directed drugs, Multi-target drugs, Neurodegenerative disorders, Rational drug design

1 Introduction

Alzheimer's disease (AD) is one of the greatest challenges of current research for new drugs [1–3]. Similar to cancer, diabetes, rheumatoid arthritis, and other chronic inflammatory and neurodegenerative diseases, AD is characterized by multiple factors involving physiological, biochemical, and chemical mediators operating concurrently with, caused by the same or different pathways [4, 5].

In recent years, advances in biochemistry, neuropharmacology, and other biological fields have been responsible for new insights in

the high complexity multifactorial pathophysiological hallmarks of AD. Thus, the development of innovative and effective multi-target directed drugs for AD represents a new paradigm in the discovery and drug design. Until then, the question that guided the planning of a new drug was “What is the best target for the treatment or prevention of a disease?” which supported, since 1990s, the origin of the reductionist strategy. By this concept, a drug should have its action limited to individual genes, a single protein or enzyme, and a specific ligand to a target, following the current paradigm “one gene, one target, one drug” [1, 4, 6].

The limitations imposed by the low therapeutic efficacy against most of these diseases, coupled with the apparent reduction in the approval of new bioactive chemical entities, despite significant investments from the Pharmaceutical Industry, have reinforced the need to look up new strategies for planning and discovery of more effective and safer drugs.

In this context, polypharmacology seems to be a real and plausible possibility, based in three possible therapeutic strategies: the first would be the use in combination of multiple active ingredients, like the associations of natural ingredients used for centuries in Traditional Chinese Medicine; or cocktails of drugs, such as used in the treatment of the acquired immunodeficiency syndrome (AIDS) and cancer. Another alternative would be a combination of more than one active ingredient in the same formulation. In a third and more recent approach, it would be a single molecule assembling structural requirements to allow molecular recognition of more than one target simultaneously, featuring a multi-target or multifunctional drug [4, 5]. When multiple targets related to the same pathology are associated with distinct biochemical cascades, the active ligand can be also called as a symbiotic compound [4].

2 Socioeconomic Impact and Prevalence of Alzheimer’s Disease

With the increase in average life expectancy worldwide, neurodegenerative disorders, such as AD, have attracted a great attention, especially in more developed countries. As a progressive, disabling, and incurable disease, patients die within an average of 10 years after its installation and first symptoms [7]. The age is its main risk factor, with a prevalence of 0.7% among individuals 60–65 years of age and about 40% in the age groups above 90 years [8, 9]. In 2012, about 5.4 million of American people were affected by AD. In spite of very imprecise estimates available, it is believed that by 2050, over 16 million Americans will have AD, with an annual cost that could exceed US\$ 1.1 trillion [10]. Other estimates point to a global epidemic of AD, reaching 26 million people by 2050, if a cure is not yet discovered and available [3, 11].

The increase in life quality, as a natural consequence of recent advances in medicine and technological development, has been reflected directly in the increased longevity worldwide. The United Nations (UN) believes that the “age of aging” began in 1975 and this trend is expected to extend until 2025 [11, 12]. According to the Brazilian Institute of Geography and Statistics, Europe occupies the first position in the proportion of the population above 60 years (19.8%), followed by North America (16.3%) [13].

3 Multifactorial Aspects of Alzheimer’s Disease

AD is a degenerative disorder of the central nervous system (CNS) and is one of the most common dementia among the population above age 65. AD is characterized by a progressive loss in memory, cognition, motor and functional capacity, gradually undermining social behavior, and ability to perform routine tasks, such as feeding, personal hygiene, and interpersonal relationships. Although the etiology of AD is not completely understood, it is well established that operation of multiple interconnected factors (Fig. 1) is related to the installation, development, and evolution of the disease.

AD patients present lower levels of the neurotransmitter acetylcholine (ACh) in the synaptic cleft and an impaired cholinergic transmission, resulting in learning and memory dysfunction. These related events originated the proposal of the “Cholinergic Hypothesis” for AD, stating that one way to enhance cholinergic neurotransmission is to inhibit an enzyme responsible for the metabolic breakdown of ACh [14].

Apparently, physiological deregulation in some brain regions, the origin of which being not entirely understood, is responsible for an overproduction of amyloid peptide (A β). Then, β - and γ -secretase enzymes abnormally cleave A β , producing insoluble fragments of 39–43 amino acid residues. Fragments A β _{1–42}, even at low concentrations, seem to be more prone to oligomerization and formation of insoluble neurotoxic aggregates, the so-called amyloid plaques, and are the main therapeutic target advocated by the “Amyloid Hypothesis” [5, 6, 15].

A secondary event due to neurotoxicity of amyloid plaques is the hyperphosphorylation of the neuronal microtubule constitutive tau protein. This abnormal phosphorylation of tau leads to structural collapse of microtubules and the consequent release of tau protein fragments, which take the form of insoluble fibrillar coil, depositing intracellular as neurofibrillary tangles [6, 10, 16–18].

These two biochemical factors acting together lead to accentuate changes in the CNS homeostasis and function, causing numerous cellular processes and organelles activation, acting as starting points of a complex process of interconnected physiological

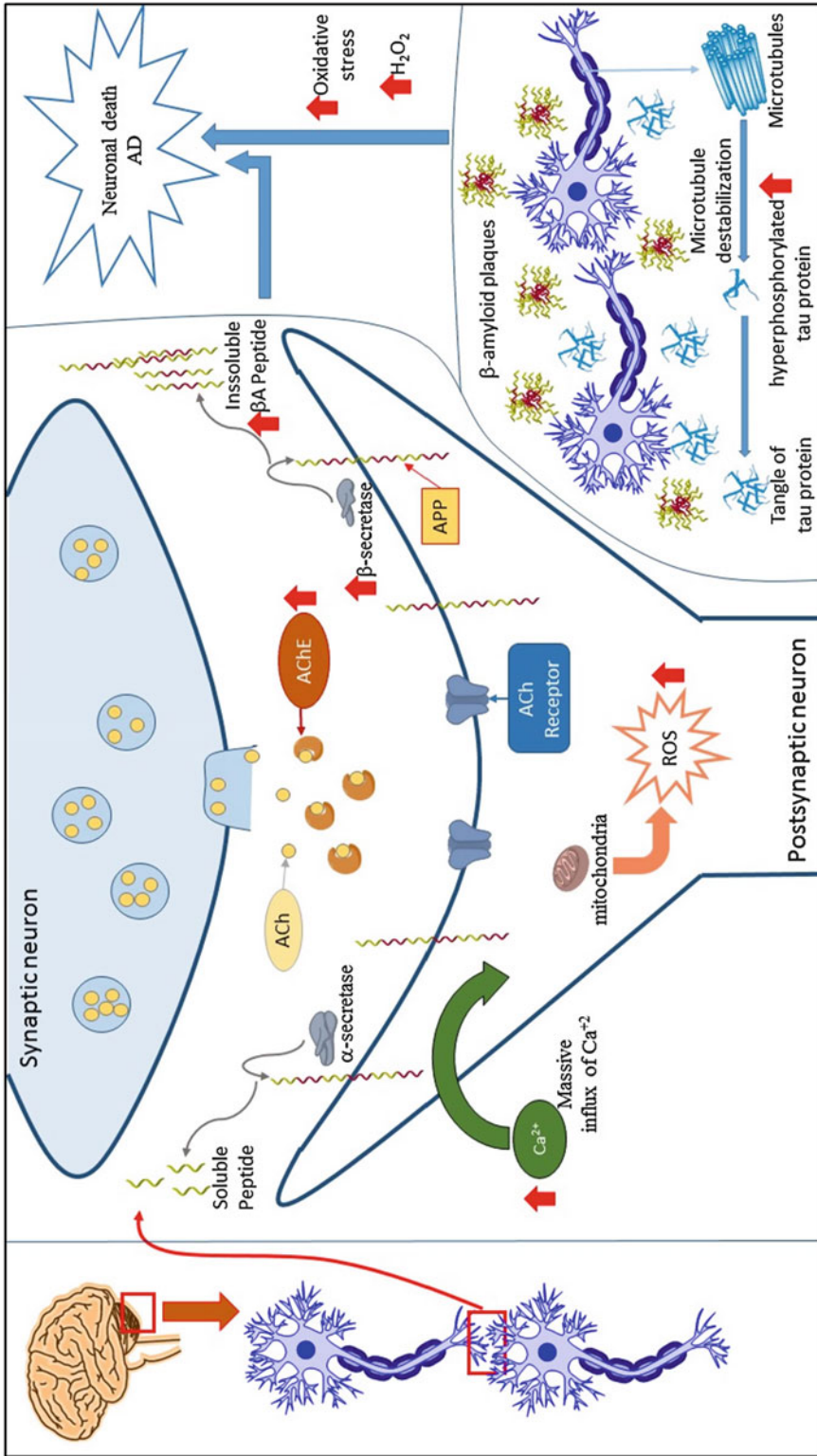


Fig. 1 Multifactorial pathophysiology of Alzheimer's disease

damage, also leading to neuronal death [17]. Although most studies have emphasized the neurotoxicity of amyloid plaques, recent evidences indicate that an increase in the concentration of soluble amyloid oligomers generated by overproduction of A β (Fig. 1) can be also responsible for increasing neuronal injury [6, 19].

Mitochondria are the major intracellular targets of soluble A β (sA β) oligomers. In excess, sA β eventually interferes with normal function of mitochondria, causing overproduction of reactive oxygen species (ROS), inhibition of cellular respiration, ATP production, and damage in mitochondrial structure [19, 20]. It seems that the deleterious effects of sA β on the mitochondria are the result of changes in homeostasis and intracellular Ca $^{2+}$ signaling, as evidenced by the induction of massive influx of Ca $^{2+}$ in cultured neurons, causing its concentration increasing in mitochondria and neuronal apoptosis (Fig. 1) [19]. Accumulation of Ca $^{2+}$ in mitochondria causes opening of the mitochondrial permeability transition pore (mPTP), a wide mitochondrial membrane channel that allows the passage of large molecules bidirectional uncontrolled, resulting in disintegration of organelles and functional structure [19].

Recent studies have shown that the dephosphorylation of neurofibrillary tangles of tau protein was able to restore its ability to bind to microtubule neurons, indicating that the kinetic mechanisms that regulate the phosphorylation/dephosphorylation process are altered in AD [21]. The nature of the protein kinases, phosphatases, and tau sites involved in these lesions was recently unveiled, suggesting that activation of protein phosphatase phosphoserine or phosphothreonine (PP-2a) or inhibition of glycogen synthase kinase-3 β (GSK-3 β), and protein cyclin-dependent kinase 5 (CDK5) may be required for inhibiting the degeneration caused by neurofibrils in AD [21].

In recent years, many other hypotheses have been proposed to explain the complexity and multifactorial pathogenesis of AD, including oxidative stress, disruption of homeostasis by metal ions, and neuroinflammation [17, 22, 23]. Currently, oxidative stress is considered one of the major causative factors of AD, unifying a number of other sequential or individual pathophysiological events. Oxidative damage in the brain of AD patients is a result of excessive production of free radicals induced by A β , functional alteration in mitochondria, inadequacy in energy supply, production of inflammatory mediators, and alteration of antioxidant defenses (Fig. 1) [24, 25]. Modulation of cellular oxidative processes is closely related to the redox properties of some metals. Change in concentration of such ions can lead to oxidative stress and increased ROS production. Copper (Cu $^{2+}$), zinc (Zn $^{2+}$), and other metal ions influence the processes of protein aggregation, a critical step in many neurodegenerative diseases. In the case of AD, amyloid precursor protein (APP) and A β are able to form

complexes with and reduce Cu^{2+} , which forms a high affinity complex with $\text{A}\beta$, promoting their aggregation.

Furthermore, *in vitro* studies showed that $\text{A}\beta$ neurotoxicity is dependent on the catalytic generation of H_2O_2 , which is enhanced by the presence of $\text{A}\beta\text{-Cu}^{2+}$ complexes. In addition, Cu^{2+} , Zn^{2+} , and Fe^{3+} are present in amyloid plaques in brains with AD, which can be dissolved by the action of metal-chelating substances [6, 18, 26].

The production of $\text{A}\beta$ also depends on the bioavailability of cholesterol in nerve cells, since the activity balance of the α - and β -secretases (BACE-1) is related to the lipid composition of cells. High concentrations of cholesterol into the cells lead to an increase in amyloidogenic APP process by β -secretase. Lower levels of cholesterol stimulates APP processing physiological by action of enzyme α -secretase. The hypothesis that the control of plasmatic cholesterol levels would be beneficial for treating AD has been demonstrated by using anticholesterolemic drugs, such as statins, which act as HMG-CoA reductase inhibitors [18].

Finally, deposition of $\text{A}\beta$ fragments and neurofibrils, coupled with the uncontrolled production of ROS, is crucial for the installation of a neuroinflammatory process, with the same complexity observed in peripheral tissues. The scope and relevance of this process in the establishment and development of chronic DA have been demonstrated in several recent studies in the literature [27–30]. Among all the brain cells, microglia appears to have fundamental importance in CNS inflammation associated with AD. These cells could be activated by $\text{A}\beta$, and then they modulate the production of cytokines, chemokines, and neurotoxins that are highly neurotoxic, contributing to neuronal degeneration [17, 27–30].

Given the variety of factors associated with the onset, progress, and severity of AD, increasing their degree of pathophysiological complexity, and associated to the inefficiency of the current therapeutic arsenal available, it becomes unavoidable to adopt a new concept for the rational design of new drugs against DA (Fig. 1). In this context, drug candidate prototypes with dual mode of action were the first attempts to look up ligands recognized by more than one molecular target or more than one site on the same macromolecular target. Currently, a new strategy of multi-target directed ligands (MTDL) is gaining special attention in the scientific community, which has been seeking molecular hybridization, a tool for designing new molecular patterns. These molecular hybrids could lead to the identification of new bioactive chemical entities with selective affinity for multiple targets, preferably in different biochemical cascades. Therefore, these innovative ligands could play a singular role in the advance of a broadly and more efficient therapy, and perhaps, in the cure of AD.

4 Recent Multi-target Directed Drug Candidates Designed for the Treatment of AD

This new approach, considering multifunctional drugs or ligands directed at multiple targets associated with the same disease (symbiotic drugs), has gained special importance and introduced a new thinking approach in the design of new drug candidates for AD. Molecular hybridization of pharmacophoric subunits of different bioactive molecules is the main tool for structural planning and has provided the discovery of numerous ligands with multiple properties, including antioxidant, neuroprotective, metal chelation, anti-inflammatory, anti-A β aggregation, and cholinesterase and secretase inhibitory activities. Therefore, a set of other potential therapeutic targets have been studied for simultaneous intervention, seeking for more efficacy in relieving the symptoms and slowing the progression of AD, and aiming for its definitive control and cure [1, 2, 31–33].

Since 2005, the literature has shown several results from applying this innovative approach in drug design. Drugs such as donepezil, tacrine, and rivastigmine [34, 35] have been used as structural models for molecular hybridization with bioactive synthetic and natural products such as curcumine [15, 36], berberine [37, 38], and 8-hydroxyquinolines [39], among others in the search for new chemical entities with multiple bioactive properties useful for the treatment of AD.

Memoquin (**2**, Fig. 2), a 1,4-benzoquinone-polyamine hybrid compound, was reported by Cavalli and co-workers as a promising drug prototype for AD. The structural design of compound **2** was based on a polyamine core, derived from coproctamine (**1**), a dimeric agent with anticholinesterasic and antimuscarinic. Aiming to add to these properties, the ability to neutralize ROS and neuroprotection, the authors proposed the insertion of a 1,4-benzoquinone, from coenzyme Q10 (CoQ10), a fragment with a recognized potent mitochondrial antioxidant property and a protector of the hippocampus neurons against A β _{1–40} induced neurotoxicity. The inner polymethylene chain was replaced by the benzoquinone nucleus. The insertion of the benzoquinone nucleus contributes to increase the neuroprotective activity, because a hydrophobic and planar π system is generated, which is able in principle to bind A β [36, 40].

Memoquin (**2**) exhibited ability to inhibit acetylcholinesterase (AChE) ($IC_{50} = 1.55$ nM), almost 15-fold more potent than the reference compound (coproctamine, **1**). Memoquin (**2**) also inhibited AChE-induced A β _{1–40} aggregation with an IC_{50} of 28.3 μ M and the self-induced A β _{1–42} aggregation with an IC_{50} of 5.93 μ M. In addition, compound **2** is also capable to inhibit BACE-1 (β secretase) activity in the range of nanomolar concentrations ($IC_{50} = 108$ nM) and showed antioxidant properties, reducing

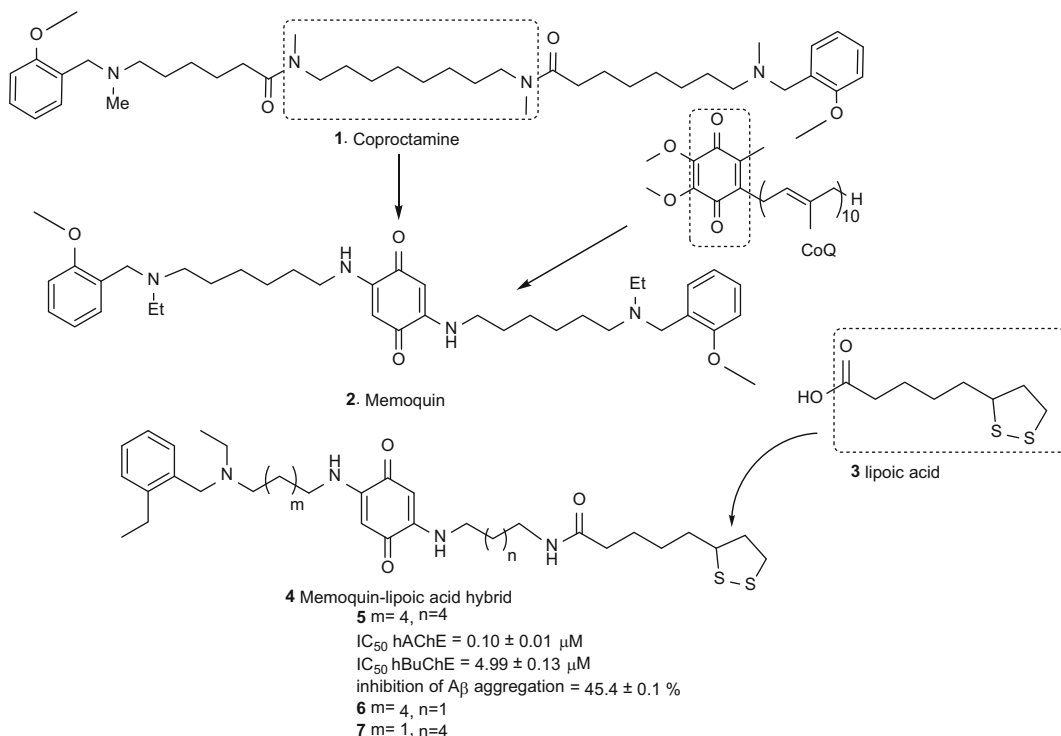


Fig. 2 Design strategy for memoquin (2) and their derivatives (5–7)

the formation of free radicals by 44.1%. It was demonstrated that the 1,4-benzoquinone moiety can be reduced *in vivo* to the 1,4-dihydroquinone form, increasing its antioxidant potential and scavenging properties. Finally, *in vivo* assays showed the potential of memoquin to restore the cholinergic deficit and successfully reduce $A\beta$ expression and accumulation, also reducing hyperphosphorylation of tau protein [40, 41].

Considering all these multi-target profiles of memoquin (2), Bolognesi and co-workers reported in 2009 the design of new hybrid lipoic acid–memoquin (4–7) derivatives inspired on the memoquin structure and lipoic acid structure which is a potent antioxidant. In addition to the observed biological properties of memoquin, compound 5 (IC_{50} human AChE (hAChE) = $0.10 \pm 0.01 \mu\text{M}$, IC_{50} hBuChE = $4.99 \pm 0.13 \mu\text{M}$, and inhibition of $A\beta$ aggregation = $45.4 \pm 0.1\%$) also exhibited a good neuroprotective effect against oxidative stress, greater than the reference control lipoic acid (3) [41].

Curcumin (8, Fig. 3) is the major constituent of *Curcuma longa*, species that is used in Indian Traditional Medicine that exhibit various neuroprotective properties. This compound has been used as structural model molecular to generate new hybrids with memoquin (2). These hybrids exhibit effects $A\beta$ -aggregation inhibitory, neuroprotective and AChE and butyrylcholinesterase

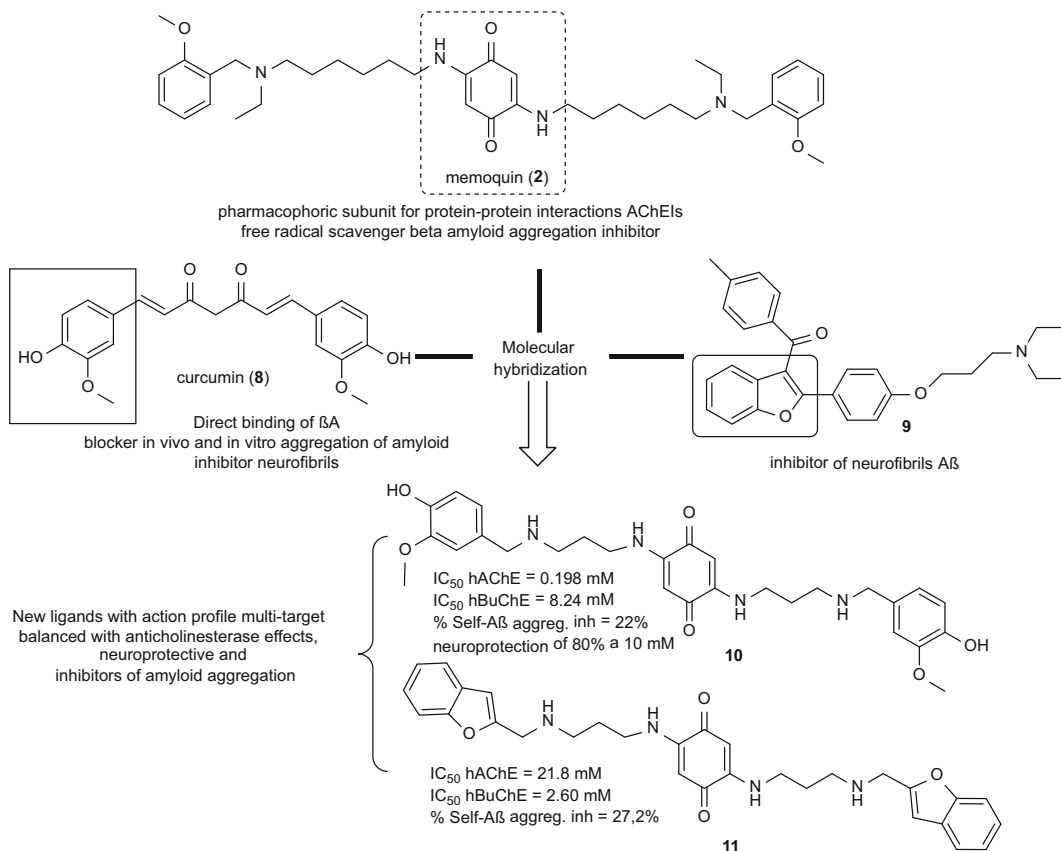


Fig. 3 Compounds **10** and **11** designed as new analogues of memoquin (**2**) with neuroprotective properties, anti-amyloid, and selective acetylcholinesterase (AChE) and butyrylcholinesterase (BuChE) inhibitory activities

(BuChE) inhibitory properties. Among the novel memoquin hybrid derivatives, compounds **10** and **11** (Fig. 3) showed the best multiple-target action profile with selective inhibitory activities of AChE and BuChE, with modulatory effects on the amyloid aggregation and inhibition of neurofibrils formation [42].

5 Multi-target Directed Ligands Inspired by Galanthamine

According to the studies of Simoni and co-workers, a series of hybrid derivatives (**14**) of galanthamine (**12**), an AChE inhibitor and memantine (**13**), a noncompetitive NMDA receptor antagonist, showed potent AChE inhibitory activities with multiple neuroprotective effects in vitro (Fig. 4). Some of these hybrid compounds were also capable to prevent glutamate-induced neurotoxicity by moderately blocking glutamate receptor NMDA subtype and neuronal cell death via NMDA receptor-dependent effect at nanomolar concentrations (e.g., memagal (**15**), IC_{50}

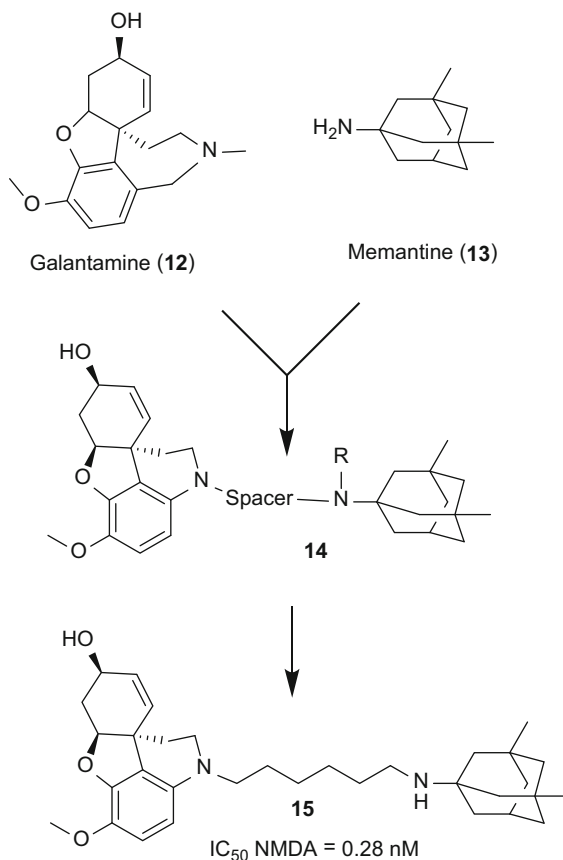


Fig. 4 Chemical structure of galanthamine–memantine hybrid compounds (14) with dual AChE/NR2B effect

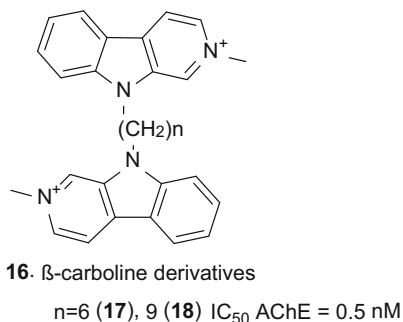


Fig. 5 Chemical structure of new bivalent β -carboline derivatives 17 and 18

NMDA = 0.28 nM). These molecules were the first examples reported of dual AChE/NR2B drug prototype candidates [43].

A new series of bivalent β -carboline derivatives (16–18) was synthesized by Rook and co-workers based on previous studies that disclosed monovalent β -carbolines as potent dual AChE/BuChE inhibitors (Fig. 5). During an investigative study for the ability of

these β -carboline derivatives to interact with multiple targets, the researchers also found that several compounds were also potent NMDA receptor blockers, in addition to the potent inhibitory properties of both AChE/BuChE enzymes. A structure–activity study revealed that the inhibitory activity of quaternary N9-bivalent β -carbolines was increased compared to their monovalent analogues. Compound **18** showed the higher AChE inhibitory activity (IC_{50} AChE = 0.5 nM) and proved to be a potent inhibitor of the transient induced-glutamate Ca^{+2} with an IC_{50} = 1.4 μ M, which is fourfold more potent than memantine (**13**) (IC_{50} = 5.6 μ M) in the same assay conditions [44].

Rizzo and co-workers reported a new series of hybrid molecules based on the structure of the AChE inhibitor *N*-methyl-*N*-benzylamine (**20**), and compound SKF-64346 (**19**), a benzofuran derivative with good inhibitory properties of A β -fibril formation, using a heptyloxy moiety as a linker subunit. All compounds evaluated have shown to be good AChE and A β -fibril formation inhibitors, reducing the A β neurotoxicity. Compound **21** showed an IC_{50} of 38.1 μ M in BuChE inhibition, being 1.06-fold more selective towards BuChE than AChE (IC_{50} AChE = 40.7 μ M). Furthermore, this compound was also able to inhibit A β self-aggregation with an IC_{50} of 12.5 μ M and showed a neuroprotective effect of 58% at 30 mM [45]. Some years later, aiming to improve its pharmacological profile the researchers modified the original structure of compound **21** by variation of the spacer subunit between the 2-arylbenzofuran and the *N*-methyl-*N*-benzylamine moieties. They proposed a variation in the length of the spacer subunit, modification in the substituents in three positions of the benzofuran scaffold, and change of the *N*-methyl-*N*-benzylamine heptyloxy side chain from the *para* to *meta* position (Fig. 6) [45]. These structural modifications led to a new series **22** that exhibited an increase in activity for most of the modified compounds.

Among the most active compounds of the series **22**, compounds **23** (IC_{50} hAChE = 102 ± 18 μ M, IC_{50} hBuChE = 0.40 ± 0.07 μ M) and **24** (IC_{50} hBuChE = 0.048 ± 0.008 μ M) exhibited the best BuChE inhibitory properties, with carbamate **25** (IC_{50} hAChE = 0.34 ± 0.03 μ M, IC_{50} hBuChE = 0.88 ± 0.10 μ M) also showing higher dual inhibitory activity of AChE and BuChE in comparison to rivastigmine. In addition, compounds **26** (IC_{50} hAChE = 32.6 ± 11.9 μ M, IC_{50} hBuChE = 0.28 ± 0.02 μ M) and **27** (IC_{50} hAChE = 0.24 ± 0.02 μ M, IC_{50} hBuChE = 2.88 ± 0.26 μ M) also showed significant in vitro inhibitory activities on A β -fibril formation up to 81% (Fig. 7) [45, 46].

Piazza and co-workers also searching for multifunctional drug candidates for AD planned a series dual of inhibitors **30–31** of hAChE and BACE-1. The molecular design was based in a dual-

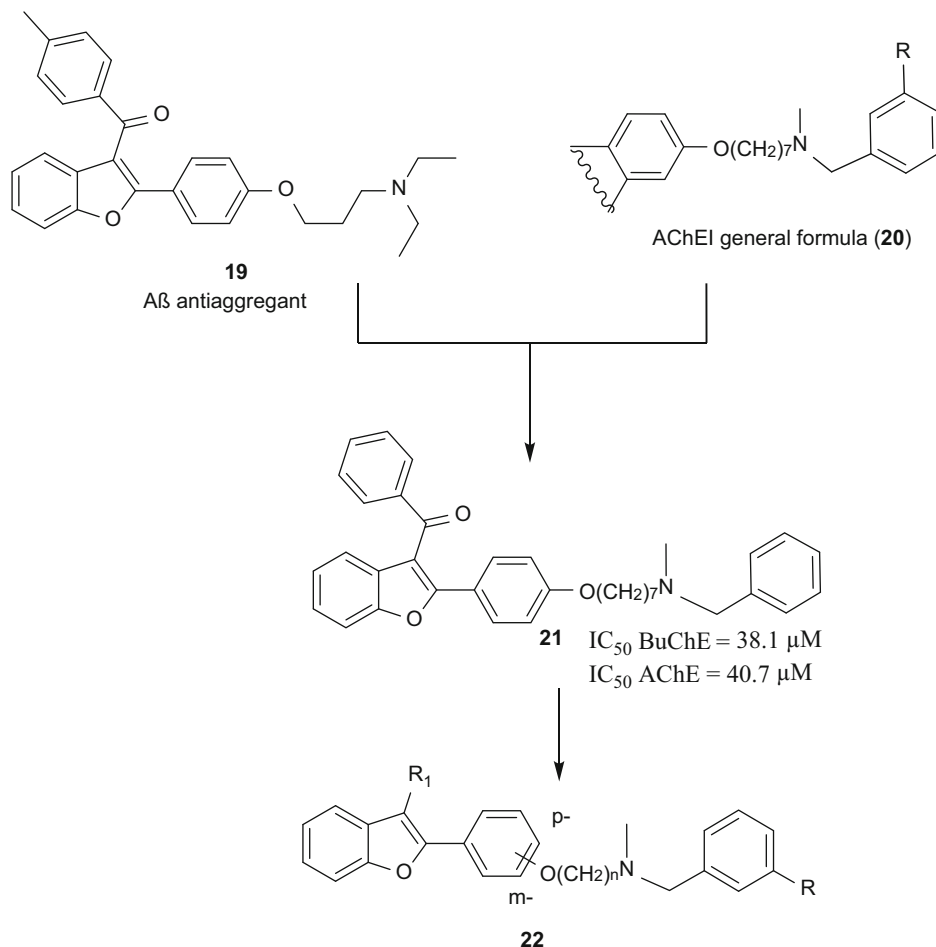


Fig. 6 Design of compound **21** by molecular hybridization of SK-64346 (**19**) and *N*-methyl-*N*-benzylamines (**20**) and its molecular derivative series **22**

binding site (DBS) AChE inhibitor, AP2238 (**28**), which is able to simultaneously interact with both the central and the peripheral anionic sites (PAS) of AChE. It was synthesized a series of compound derivatives by replacing the methoxy substituents of the coumarin moiety by an amidic chain subunit aiming to extend the activity to BACE-1 (Fig. 8). The position 6 or 7 of the coumarin moiety was also replaced by a di-halophenyl acid subunit, considering that this moiety emerged as a leitmotif in different BACE-1 inhibitors reported in the literature. All compounds were potent inhibitors, particularly compounds **30** (IC_{50} hAChE = 0.551 μ M, IC_{50} BACE-1 = 0.149 μ M) and **31** (IC_{50} hAChE = 0.181 μ M, IC_{50} BACE-1 = 0.150 μ M) that exhibited the best inhibitory profile in the target series. Biological data clearly indicated that the introduction of a halophenyl alkylamidic subunit on the scaffold of **28** allowed obtaining more potent BACE-1 inhibitors, and the

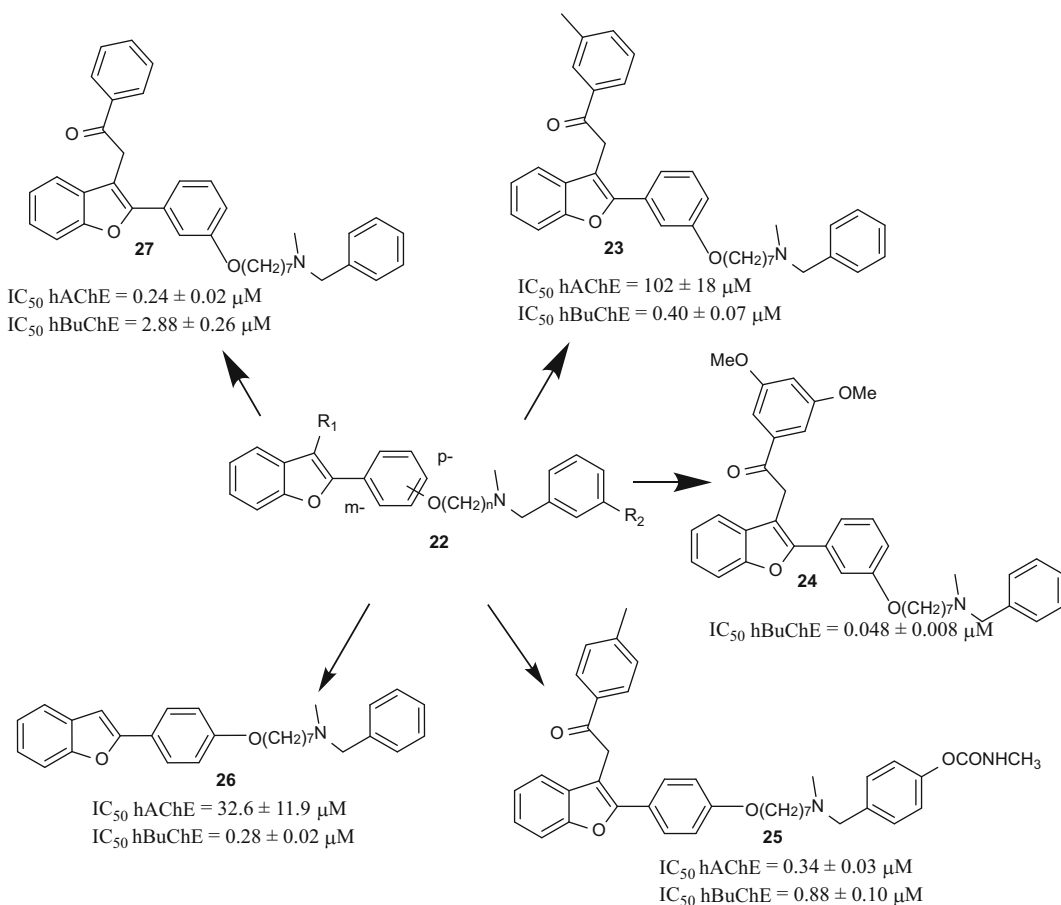


Fig. 7 Chemical structures of the improved derivatives **23–27** designed by molecular optimization of the hybrid prototype **22**

presence of substituents in the position 6 or 7 on the coumarin nucleus seems to equally contribute to the relative binding affinity of this molecular series [46].

6 Multi-target Directed Ligands Inspired by Donepezil

The donepezil (Aricept[®], **32**) arose in the 1980s as a reversible and noncompetitive inhibitor of AChE, to be the most used drug for the AD therapy [47]. Computational studies indicate that its action mechanism is derived from the *N*-benzylpiperidine and indanone subunits, which guarantee high affinity and selectivity for AChE [48]. Clinical studies revealed that the use of donepezil resulted in significant improvement in memory, concentration, language, and reasoning, without signs of toxicity; however, there was no cure [48]. We will see below (Fig. 9) that donepezil (**32**) has been widely

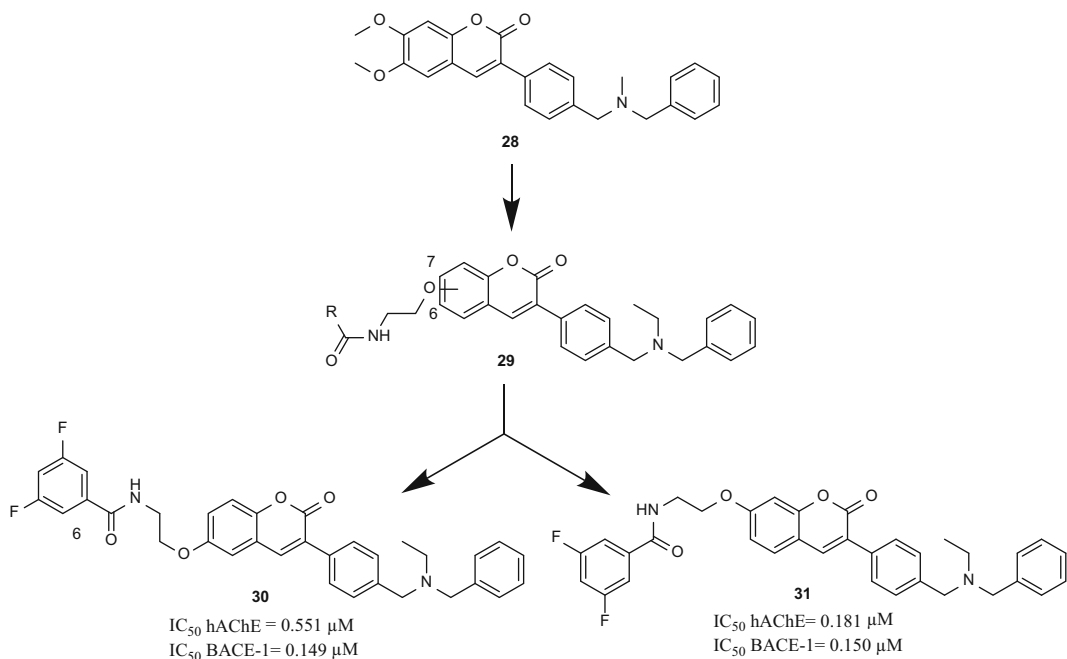


Fig. 8 Chemical structure of compounds **30** and **31**, designed as dual hybrid inhibitors of human AChE (hAChE) and β -secretase (BACE-1) based on the AChE inhibitor AP2238 (**28**)

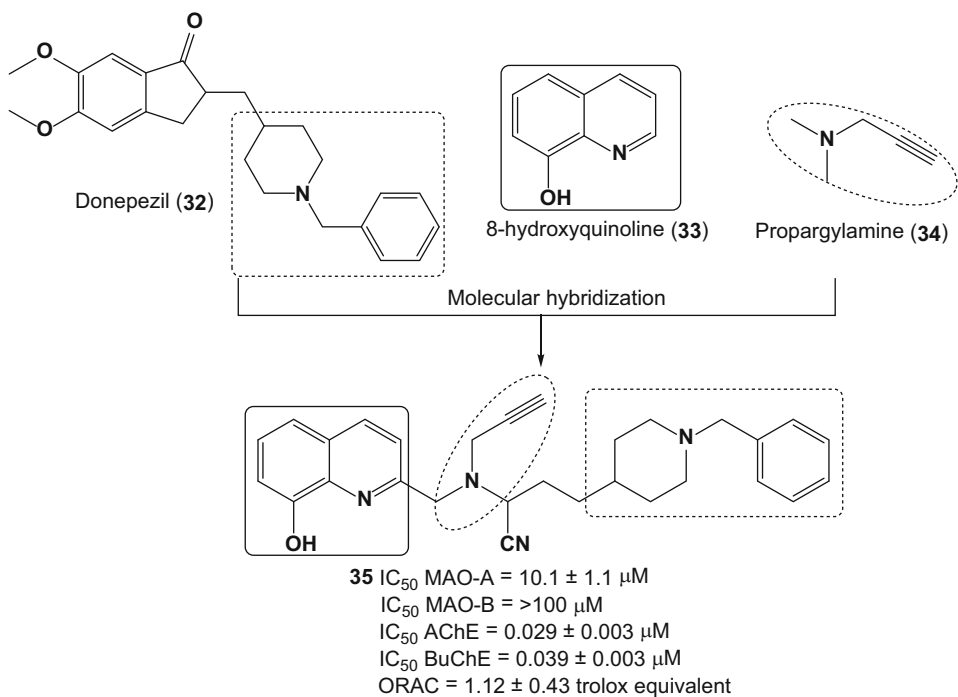


Fig. 9 Design of a new donepezil-8-hydroxyquinoline-propargylamine (**35**) hybrid compound with antioxidant, metal-chelating, and inhibition of AChE and monoamine oxidases (MAOs) properties

used as a prototype in the design of new MTDL drug candidates planned by molecular hybridization.

Wu and co-authors develop new substances capable of acting concomitantly as AChEIs, monoamine oxidase (MAO) inhibitors, and antioxidants, interacting with different targets involved in AD pathogenesis. The novel hybrid series was planned by structural combination of the *N*-benzyl-piperidine moiety from donepezil (**32**) as an AChEI pharmacophore, a metal-chelating portion represented by the 8-hydroxyquinoline system (**33**) and a propargylamine moiety (**34**) for MAO inhibitory activity (Fig. 9). In this context, six new substances were synthesized, with compound **35** which stood out from the others, showing a selective inhibition of MAO (IC_{50} MAO-A = $10.1 \pm 1.1 \mu\text{M}$ and IC_{50} MAO-B > $100 \mu\text{M}$) and interesting IC_{50} values for AChE inhibition ($IC_{50} = 0.029 \pm 0.003 \mu\text{M}$) and BuChE ($IC_{50} = 0.039 \pm 0.003 \mu\text{M}$). Then, the metal-chelating properties of compound **35** were evaluated and this has shown to be capable of chelating Zn^{+2} and Cu^{+2} . These results explain an observed dose-dependent pharmacological profile, when used in concentrations greater than $0.4 \mu\text{M}$, for inhibition of Cu-mediated H_2O_2 production. In addition, compound **35** also showed antioxidant activity (1.12 ± 0.43 trolox equivalent by ORAC assay). Finally, we can state that compound **35** represents a genuine multi-target hybrid ligand, being a strong inhibitor of ChEs and a selective inhibitor of MAO-A, with antioxidant and Cu^{+2} chelating activities [49].

The *N*-benzylpiperidine moiety, derived from donepezil (**32**), was also used as a scaffold by Pudlo and co-workers for the design of a series of donepezil–quinolones hybrid compounds (**37–38**) with antioxidant and AChE inhibitory properties (Fig. 10). Quinolone derivatives (**36**) include motifs exhibiting a wide variety of biological activities, such as ROS scavenging ability. All phenolic compounds of the target series **37–38** have shown moderate to

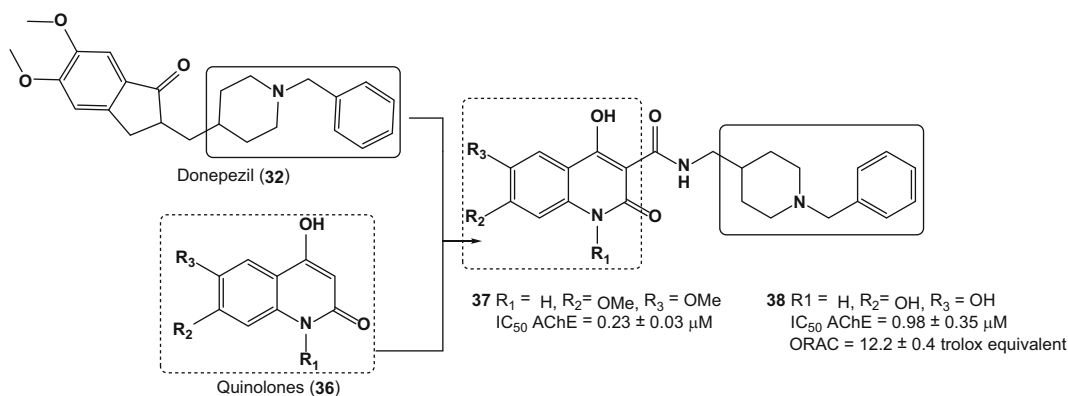


Fig. 10 Chemical structure of the multifunctional quinolone–benzylpiperidine derivatives and the most active derivatives **37** and **38**, designed as novel antioxidant and AChE inhibitors for Alzheimer’s disease

high radical scavenging activities. Compounds **37** and **38** were the most potent, with compound **37** showing an effective inhibition of AChE and high radical scavenging activity, without altering the inhibition of AChE. The catechol moiety of compound **38** provided antioxidant activity and was also involved in an additional interaction with AChE. However, despite the multifunctional properties exhibited by both compounds **37** and **38**, their AChE inhibitory activities were considered low, with IC_{50} values of $0.23 \pm 0.03 \mu\text{M}$ and $0.98 \pm 0.35 \mu\text{M}$, respectively. By contrast, these results suggested that, due to its ability in radical scavenging, quinolone moiety could be considered an interesting biophore group in the design of novel multipotent molecules for the treatment of Alzheimer's disease [50].

A new family of multi-target molecules able to interact with both AChE and BuChE as well as with MAO A and B has been synthesized by Samadi and collaborators. Rational structural pattern of these compounds considered conjunctive approach by combination of benzylpiperidine and *N*-propargylamine moieties present in the AChE inhibitor donepezil (**32**) and the MAO inhibitor PF9601N (**39**), respectively, connected by a central pyridine or naphthyridine ring system (Fig. 11). The most promising derivative **40** showed a potent inhibitory activity of AChE ($IC_{50} = 37 \text{ nM}$), but less potent than donepezil. This compound also showed a moderate inhibitory potency of MAO-A ($IC_{50} = 41 \mu\text{M}$). Moreover, molecular modeling showed that these inhibitors probably act in two binding sites of AChE, and that the length of the spacer subunit is of particular relevance in the modulation of AChE inhibition, being crucial for the dual interaction of these molecules with both catalytic anionic site (CAS) and PAS sites of the enzyme [51].

In 2011, another publication by Bolea and collaborators also reported the synthesis and pharmacological evaluation of a new

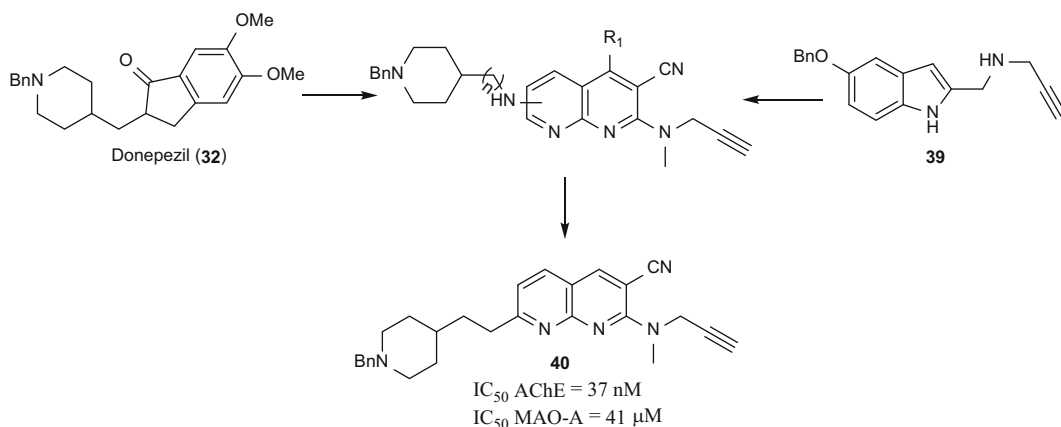


Fig. 11 Chemical structure of the most active dual AChE and MAO-A inhibitor designed by molecular hybridization of donepezil and PF9601N (**40**)

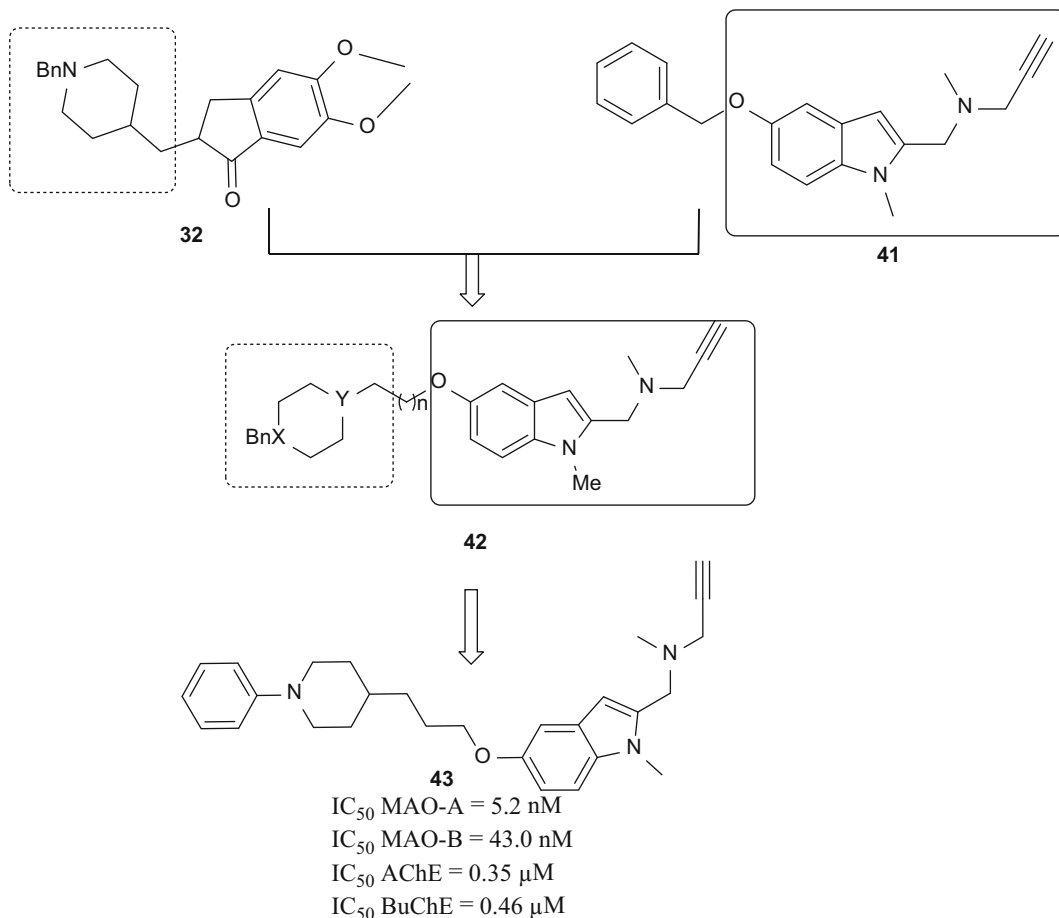


Fig. 12 Chemical structure of the lead compound **43** designed as a multi-target directed ligand (MTDL) for ChEs and MAO-A and B

family of multifunctional molecules capable of interacting simultaneously with the AD-related enzymes AChE, BuChE, MAO-A, and MAO-B. The structural architecture of the series **42** (Fig. 12) was based on the molecular hybridization of the benzylpiperidine moiety of donepezil (**32**) and the indolyl-propapylamino subunit present in the structure of the MAO inhibitor *N*-[(5benzyloxy-1-methyl-1*H*-indol-2-yl)methyl]-*N*-methylprop-2-yn-1-amine (**41**), connected by a central oxy-methylene chain. Compound **43** was identified as the most promising compound from series **42**, showing a good selectivity and high potency in the inhibition of MAO-A with an IC₅₀ = 5.2 nM (IC₅₀ MAO-B = 43.0 nM), but less potent and with poor selectivity for the AChE (IC₅₀ = 0.35 μM) and BuChE (IC₅₀ = 0.46 μM). Further molecular modeling and enzymatic kinetics studies clarified that these inhibitors act in two binding sites of AChE, which could explain the inhibitory

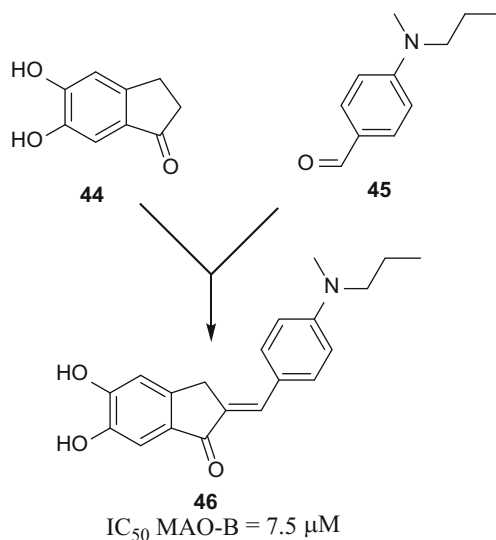


Fig. 13 Chemical structure of compound **46**, a substituted-benzylidene indanone derivative with a wide range of multi-activity profile

effect over A β -self-induced ($47.8 \pm 2.1\%$) and AChE-dependent aggregation ($32.4 \pm 7\%$) exerted compound **43** [52].

Taking the structure of donepezil (**32**) as a model, Huang and co-workers reported in 2012 a new family of derivatives based on the structure of benzylidene-indanone (**45**) pharmacophoric subunit of **32** with the insertion of set of substituted benzene ring systems (**44**, Fig. 13). In vitro evaluation showed that most of the compounds were potent inhibitors of MAO-B (IC₅₀ of 7.5–40.5 μM), besides antioxidant (ORAC-FL value of 2.75–9.37 μM) and metal chelator properties. In addition, they showed a great ability for inhibition of self-induced β -amyloid aggregation (10.5–80.1%, 20 μM). Compound **46** was identified as the most potent inhibitor agent of A β _{1–42} aggregation (80.1%) and MAO-B (IC₅₀ = 7.5 μM). Moreover, compound **42** proved to be an excellent antioxidant and metal chelator agent, being able to inhibit Cu²⁺-induced A β _{1–42} aggregation and disassembling the A β fibrils [53]. Thus compound **46** is one of the most current examples from the literature for a genuine MTDL drug prototype candidate for AD treatment.

In 2015, Guzior and co-workers reported the synthesis and pharmacologic evaluation of a new series of donepezil-based compounds endowed with inhibitory properties against cholinesterases and β -amyloid aggregation. The donepezil-based compounds **48** and **49** consisted of an isoindoline-1,3-dione fragment connected to *N*-methyl-benzylamine subunit by alkyl linkers of different lengths **47** (Fig. 14). Crystallographic studies revealed that AChE has two binding sites: CAS and PAS. AChE is associated with

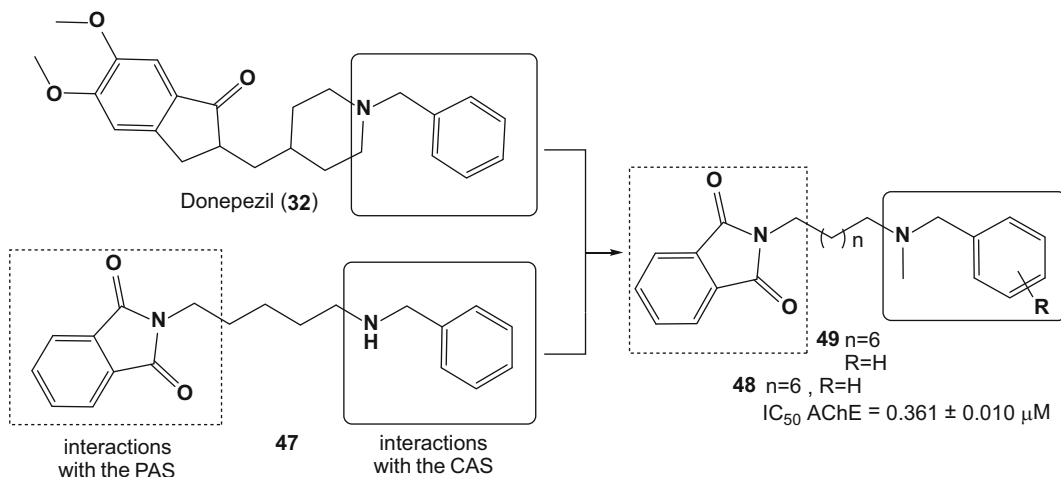


Fig. 14 Design of isoindolino-1,3-dione derivatives and structure of the most active compound **49**

induced β -amyloid aggregation. The peripheral anionic site (PAS) of AChE is the local where the interaction between the enzyme and A β apparently occurs. This enzyme interacts with the amyloid β -peptide (A β) and promotes amyloid fibril formation [54]. Dual AChEIs could have pro-cognitive effects as well disease-modifying properties by inhibiting A β aggregation in AD. These findings showed an additional role of PAS, resulting in the development of novel classes of active compounds-bivalent AChE inhibitors – that interact with CAS and PAS at the same time. Thus, a series of phthalimide–benzylamine derivatives were designed as DBS AChE inhibitors in which the *N*-benzylamine moiety could interact with the CAS of the AChE and an isoindoline-1,3-dione fragment could bind to the PAS of the enzyme. The results of pharmacological evaluation led to compound **49** (Fig. 14) as the most potent and selective hAChE inhibitor ($IC_{50} = 0.361 \pm 0.010 \mu M$) with additional properties such as A β aggregation inhibition and neuroprotective effect against A β toxicity. Kinetic studies revealed that **49** inhibited AChE in noncompetitive mode and the results from blood–brain barrier (BBB) permeability assay showed their ability to penetrate in the CNS [55].

In another approach, Więckowska and co-workers designed and synthesized new compounds as donepezil derivatives containing the *N*-benzylpiperidine moiety combined with phthalimide (**50**) or indole moieties (**51**) (e.g., **52** and **53**, Fig. 15). Most of these compounds showed micromolar-range activities towards cholinesterases and β -amyloid aggregation, along with positive results in BBB permeability assays. Derivative **52** is an example of such a compound, as it combines inhibitory activity against BuChE ($IC_{50} = 0.72 \pm 0.038 \mu M$) with A β anti-aggregation activity (72.5% inhibition at 10 μM). Moreover, these compounds

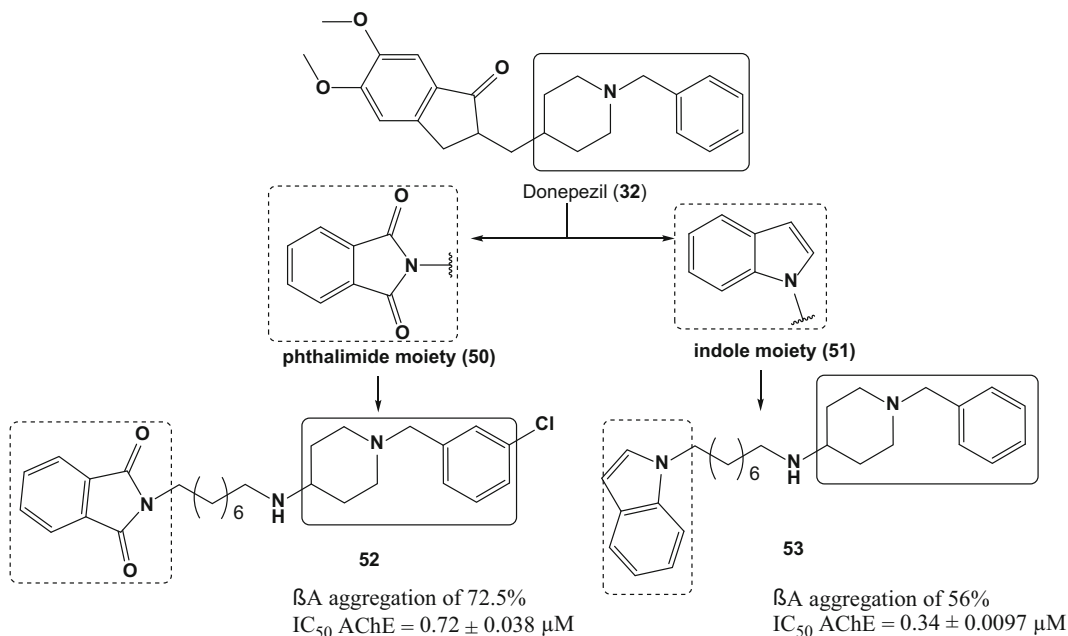


Fig. 15 Design and structure of the most active hybrid donepezil–phthalimide **52** and donepezil–indole derivative **53**

exhibited a similar effect of donepezil (**32**) in animal model of memory impairment induced by scopolamine [56].

Considering that compound M30 (**54**) had been already described as a potent selective inhibitor of MAO-A and metal chelator, its structural framework was selected as a model for molecular hybridization with the *N*-benzylpiperidine moiety present in donepezil (**32**) in the drawing of a novel family of M30–donepezil hybrid ligands (Fig. 16). These compounds were designed to act as multifunctional inhibitors of AChE, MAO-A, metal chelators, and inhibitors of A β -protein formation. Among seven synthesized compounds, compound **55** was the most active, being able to inhibit AChE ($IC_{50} = 1.8 \pm 0.1 \mu\text{M}$) and BuChE ($IC_{50} = 1.6 \pm 0.2 \mu\text{M}$) in almost the same proportion and was able to inhibit MAO-A ($IC_{50} = 6.2 \pm 0.7 \mu\text{M}$) and MAO-B ($IC_{50} = 10.2 \pm 0.9 \mu\text{M}$) with a twofold selectivity for MAO-A. This same compound also showed chelating ability for Cu^{2+} and Fe^{3+} ions, when subjected to the toxicity test with HepG2 cells showing lower toxicity than donepezil, with similar oral absorption. Thus, it can be stated that compound **55** exhibited balanced properties as AChE and MAO-A inhibitor, besides biometal-chelating ability, being an interesting ligand for future studies aiming at an improvement of its action profile for the therapeutics of AD [57].

Other two new hybrid compounds **57–59** were designed as multipotent inhibitors of ChEs and MAO based on the structure of

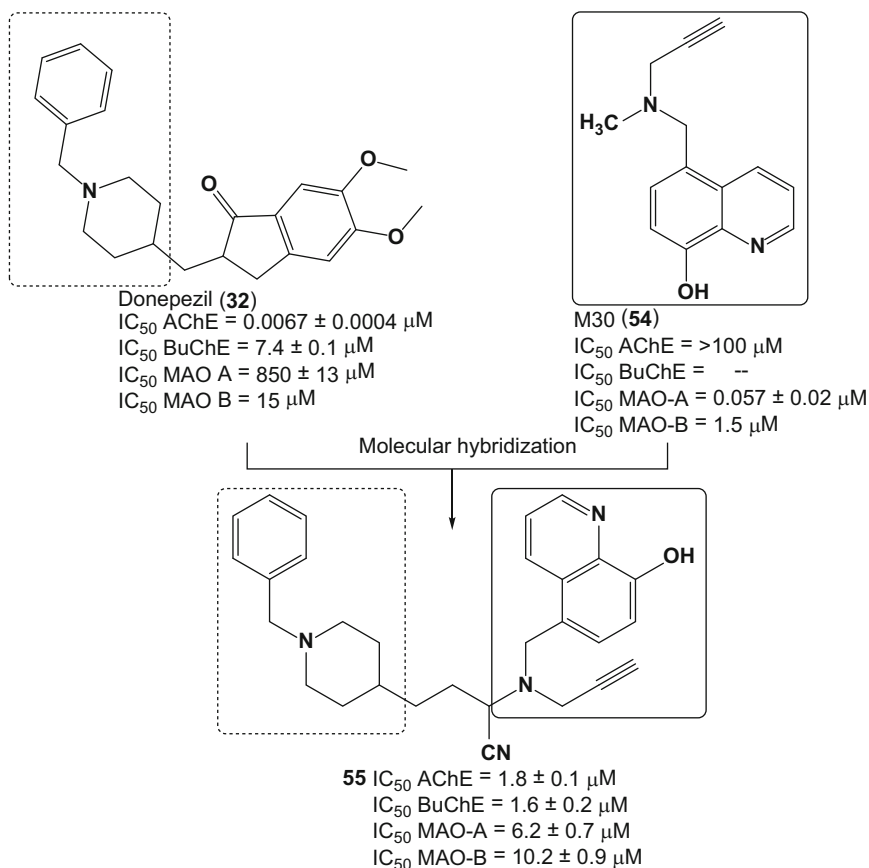


Fig. 16 Design and structure of donepezil–M30 hybrid **55** and its inhibitory data for ChEs and MAO enzymes

ASS234 (**56**), an MAO-A/B, AChE and BuChE inhibitor, donepezil (**32**), an AChE inhibitor, and PF9601N (**40**), a potent and selective MAO-B inhibitor (Fig. 17). The hybrid compound **57** was identified as a potent nanomolar-range inhibitor of MAO-A ($IC_{50} = 5.5 \pm 1.4 \text{ nM}$) and moderate inhibitor of MAO-B ($IC_{50} = 150 \pm 31 \text{ nM}$), AChE ($IC_{50} = 190 \pm 10 \text{ nM}$), and BuChE ($IC_{50} = 830 \pm 160 \text{ nM}$). Molecular modeling analysis suggests that compound **57** is a mixed-type EeAChE inhibitor, and that its linear conformation allows to span both the catalytic active site and peripheral anionic subsite, compound **57** contributing to its superior binding towards AChE respect to compound **58**. Furthermore, this compound was able to establish more interactions with the MAO-A active site compared to the MAO-B, which may indicate more tight interaction and selectivity with the former enzyme [58].

Bautista-Aguilera and collaborators described in another work a new family based on the structure of ASS234 (**56**). In that approach, the authors varied the piperidine ring substituents, showing a very completed quantitative structure–activity

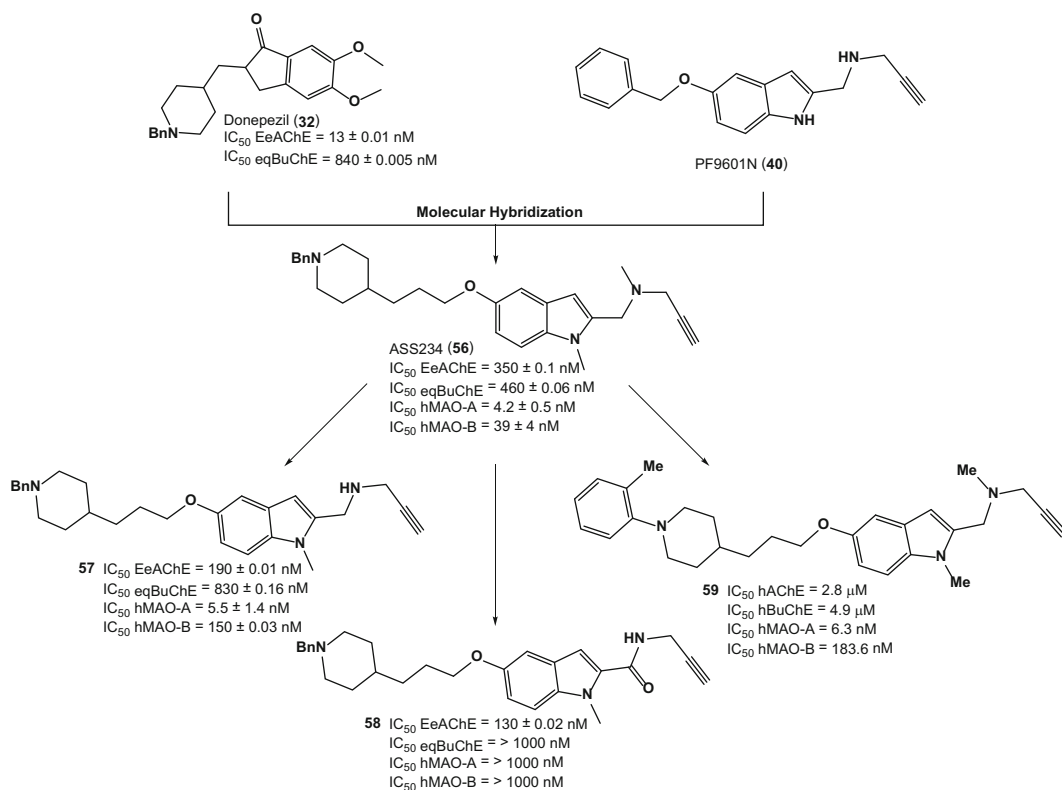


Fig. 17 Design and biological data of the donepezil–PF9601N hybrid compounds **56–59**

relationships (QSAR) study, identifying compound **59** (Fig. 17) as the lead molecule. The conclusions of the QSAR study were confirmed by the biological evaluation, where substance **59** was able to inhibit hMAO-A (6.3 ± 0.4 nM), hMAO-B (183.6 ± 7.4 nM), hAChE (2.8 ± 0.1 nM), and hBuChE (4.9 ± 0.2 nM). In that second work, Bautista-Aguilera et al. reported the design, synthesis, and biological evaluation of 19 new donepezil–indolyl hybrids as multifunctional drugs able to bind human MAO-A and ChE enzymes. According to QSAR studies, the *o*-methyl group in compound **59** improves the ligand recognition, increasing the hydrophobic interaction with hBuChE and π – π stacking interaction in hMAO-A, hMAO-B, and hAChE. The ADMET virtual analysis suggested that compound **59** has good druggable characteristics, similar to compound **40**, a prediction experimentally confirmed by in vitro BBB permeation assays [59].

A series of 5,6-dimethoxy-indanone-benzamides planned as analogs of donepezil was also drawn as multifunctional agent candidates for the treatment of AD. In the study, the aim was to evaluate a family of donepezil-like secondary amide compounds that display a potent inhibition of cholinesterases and A β , with antioxidant and metal chelation abilities. The authors synthesized

some 5,6-dimethoxy-indanone-2-carboxamide derivatives containing *ortho*-, *meta*-, and *para*-substituted secondary aromatic amines which could act as potent AChE inhibitors. The results showed that the designed compounds are likely to interact with amino acid residues in the catalytic site of AChE. Compounds **60** (IC_{50} BuChE = $2.10 \pm 0.015 \mu\text{M}$) and **61** (IC_{50} AChE = $0.08 \pm 1.813 \mu\text{M}$) (Fig. 18) were found to be the most potent inhibitors in the series, and compound **60** is the most potent antioxidant too. Compound **61** was selective for AChE, with additional good anti-aggregation activity (55.3% at $25 \mu\text{M}$) as well as a moderate radical scavenging activity. Additionally, docking studies revealed that compound **61** was able to interact with both catalytic and peripheral AChE binding sites, which justifies its high inhibitory potency and mixed-type inhibition mechanism [60].

Recently, intense research efforts in neuroscience disclosed serotonin receptors (5HT₆Rs) playing an important role in progress of neurodegenerative disorders and as a new suitable molecular target in drug development, these receptors are not only related to the cognitive aspects of AD but also to the behavioral and pathophysiological aspects of the disease. These receptors are distributed almost exclusively in the brain areas responsible for the learning process and memory [33–36]. Looking to this new perspective, inhibition of 5HT₆Rs was considered in the design of novel hybrid multifunctional AChE inhibitors **64** and **65** (Fig. 19) that were drawn based on the structure of compound (**62**), a known 5HT₆Rs antagonist, donepezil (**32**) or tacrine (**63**) as AChE inhibitors prototypes. In this way, tacrine–**63** hybrids and donepezil–**32** hybrids were synthesized and evaluated for the ability to inhibit concomitantly 5HT₆Rs, AChE, and BuChE, leading to compounds **64** (IC_{50} 5HT₆ = $2.0 \pm 0.2 \text{ nM}$, IC_{50} AChE = $12.9 \pm 0.17 \text{ nM}$, and IC_{50} BuChE = $8.2 \pm 0.23 \text{ nM}$) and **65** (IC_{50} 5HT₆ = $2.0 \pm 0.3 \text{ nM}$, 37.2% AChE inhibition at 10 nM, and IC_{50} BuChE = $2,384 \pm 52.10 \text{ nM}$) as the most active compounds in the target series, respectively. However, only

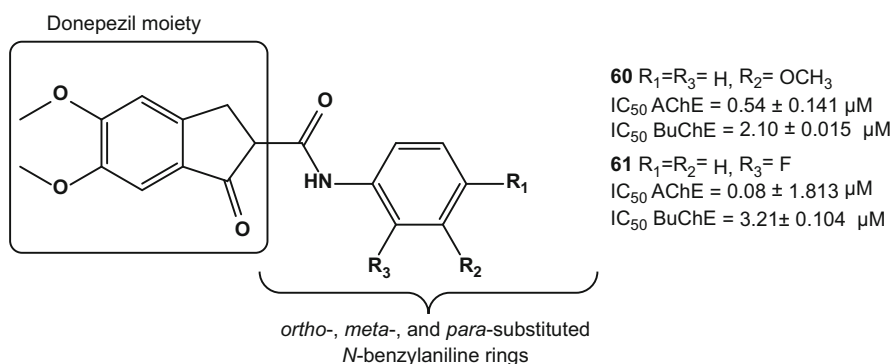


Fig. 18 Chemical structures of the most active 5,6-dimethoxy-indanone-2-carboxamide derivatives **60** and **61**

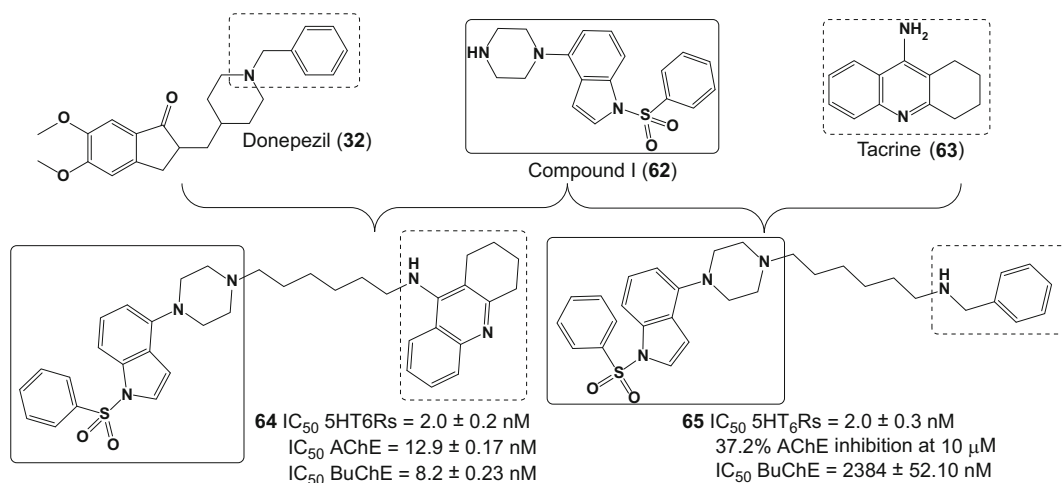


Fig. 19 Designed structure of multifunctional hybrids donepezil–62 (**64**) and tacrine–63 (**65**)

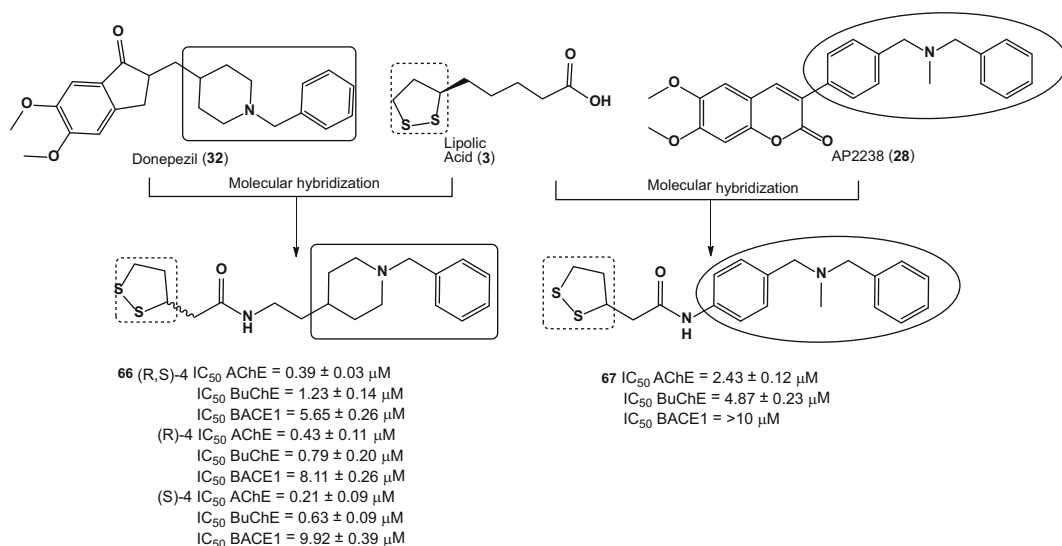


Fig. 20 Chemical structures and biological data for the most potent LA–donepezil–AP2238 hybrids **66** and **67**

compound **65** showed a promising multi-target profile, being able to act on 5HT₆Rs and AChE. Additionally, compound **65** was able to permeate *in vitro* and *in vivo* the BBB, also exhibiting a dose-dependent behavior initially, but when evaluated with a fixed dose, it was verified that its activity increases with time. Finally, the authors considered that compound **65** has a multi-target behavior, being a potent AChE, BuChE, and 5HT₆Rs inhibitor and an interesting prototype for the development of new drug candidate for AD treatment [61].

Lipoic acid (LA, **3**, Fig. 20) is a naturally occurring substance known for its therapeutic potential as a direct scavenger of ROS,

with fast absorption and fast tissue distribution. It is able to quench free radicals in aqueous and lipid phases, chelate bimetals, and regenerate other biogenic antioxidants [62]. Data from the literature suggest that LA could also play remarkable role in neuroprotection, and cognitive enhancing effects related to AD pathology [63–65]. Donepezil (**32**) and AP2238 [46] (**28**), as described above, in spite of being AChE inhibitors, are also a potent σ -1 receptor agonist and BACE1 inhibitor, respectively. These compounds were taken as molecular prototypes for the designing of two new families of LA–donepezil and LA–AP2238 hybrid compounds (Fig. 20). Six different molecules were synthesized, four LA–donepezil and two LA–AP2238 hybrids, with some of them separated in their enantiomers and tested separately in order to compare their results with those of the racemate mixtures. Among the target substances, compounds **66** and **67** were highlighted, showing the best results in inhibition of AChE. The racemic form of **66** showed a selective AChE inhibitory activity with IC_{50} AChE = $0.39 \pm 0.03 \mu\text{M}$, IC_{50} BuChE = $1.23 \pm 0.14 \mu\text{M}$, and IC_{50} BACE1 = $5.65 \pm 0.26 \mu\text{M}$, with the *S*-isomer showed to be the eutomer (IC_{50} AChE = $0.21 \pm 0.09 \mu\text{M}$, IC_{50} BuChE = $0.63 \pm 0.09 \mu\text{M}$, and IC_{50} BACE1 = $9.92 \pm 0.39 \mu\text{M}$), whereas the distomer *R*-**66** showed to be twofold and 1.4-fold less potent and selective towards AChE (IC_{50} = $0.43 \pm 0.11 \mu\text{M}$) and BuChE (IC_{50} = $0.79 \pm 0.20 \mu\text{M}$) as well as on BACE1 (IC_{50} = $8.11 \pm 0.26 \mu\text{M}$). In the same assay, compound **67** showed the weaker enzymatic inhibitory potency with IC_{50} AChE = $2.43 \pm 0.12 \mu\text{M}$, IC_{50} BuChE = $4.87 \pm 0.23 \mu\text{M}$, and IC_{50} BACE1 > $10 \mu\text{M}$, in comparison to racemic **66** and its single isomers. Furthermore, compounds (*R*)-**66**, (*S*)-**66**, (*R,S*)-**66**, and **67** were subjected to the inhibition test of σ_1R and σ_2R , showing affinities of $Ki \sigma_1$ = $8.90 \pm 0.45 \text{ nM}$, $Ki \sigma_2$ = $232 \pm 27 \text{ nM}$ for (*R,S*)-**66**, $Ki \sigma_1$ = $7.56 \pm 0.98 \text{ nM}$, $Ki \sigma_2$ = $205 \pm 42 \text{ nM}$ for (*R*)-**66**, $Ki \sigma_1$ = $15.1 \pm 1.4 \text{ nM}$, $Ki \sigma_2$ = $289 \pm 51 \text{ nM}$ for (*S*)-**66** and $Ki \sigma_1$ = $21.0 \pm 2.6 \text{ nM}$, $Ki \sigma_2$ = $1,400 \pm 230 \text{ nM}$ for compound **67**. All tested substances showed selectivity for AChE and for σ -1 receptor agonist, with weak antioxidant activities, and good CNS permeability [66].

In a recent work, Dias and co-authors reported the synthesis and pharmacological evaluation of a new series of molecular hybrids feruloyl–donepezil based on the combination of the pharmacophoric *N*-benzylpiperidine subunit from donepezil (**32**) and the feruloyl subunit present in ferulic acid (**69**) and curcumin (**68**). Curcumin (**68**), an abundant natural polyphenol found in *C. longa* rhizomes, is also widely used as scaffold for the planning of new MTDLs for AD due to its potent antioxidant and anti-inflammatory properties, playing an important role in the decrease of oxidative damage, inflammation, and amyloid accumulation with an additional biometal-chelating ability. The feruloyl moiety

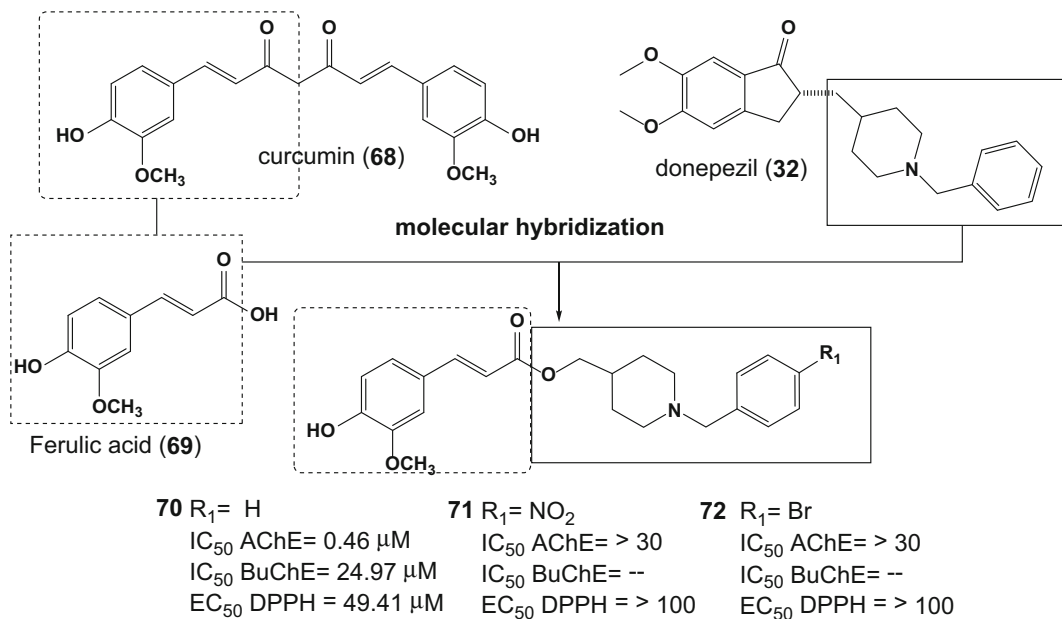


Fig. 21 Design and chemical structures of compounds **70**, **71**, and **72** as new donepezil–feruloyl hybrids with neuroprotective, metal-chelating, antioxidant, anti-inflammatory, and AChE inhibitory properties

present in curcumin (**68**) is responsible for its antioxidant activity and is also present in the ferulic acid (**69**) structure (Fig. 21), a natural compound with potent antioxidant activity. Based on these findings, a novel series of feruloyl–donepezil hybrid compounds were designed as multi-target drug candidates for the treatment of AD. In vitro results revealed potent AChE inhibitory activity for some of these compounds and all of them showed moderate antioxidant properties. Compounds **70**, **71**, and **72** (Fig. 21) were the most potent AChE inhibitors, highlighting **70** with IC₅₀ = 0.46 μM. Kinetic and molecular docking studies revealed that compounds **70** and **72** are noncompetitive inhibitors, capable of interacting with PAS of AChE. In addition, these three most promising compounds exhibited significant in vivo anti-inflammatory activity in the mice paw edema, pleurisy, and formalin-induced hyperalgesy models, in vitro metal chelation activity for Cu²⁺ and Fe²⁺, and neuroprotection of human neuronal cells against oxidative damage. Based on these data, **70** was elected as a lead compound in the series due to its best inhibitory activity of AChE, also displaying high antioxidant activities in neuronal SH-SY5Y cells, in both direct and indirect mode (activating the Keap1/Nrf2/ARE pathway), besides being a good biometal chelator and with significant in vivo anti-inflammatory activity in different animal models [67].

In another recent approach, Xu and collaborators synthesized and evaluated a novel family of donepezil–ferulic acid hybrids as MTDLs against AD (Fig. 22). In vitro results indicated that some of these molecules exhibited potent cholinesterase inhibitory

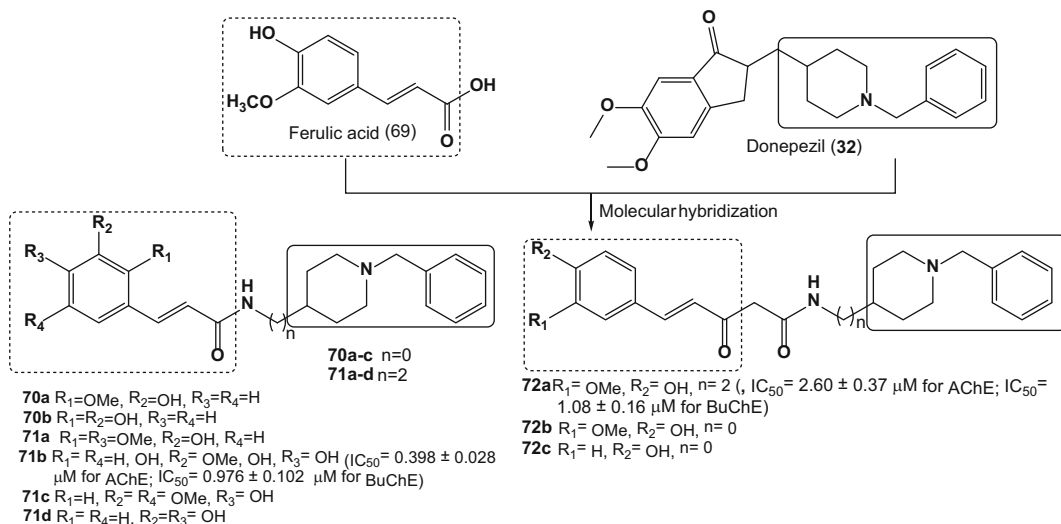


Fig. 22 Design strategy for a new series of donepezil–ferulic acid hybrid compounds **70–72**

activities, outstanding radical scavenging activities, and good neuroprotective effects on PC12 cells and could penetrate the CNS. The compounds of series **71** without hydroxyl groups exhibited better inhibition of both ChEs than compounds with hydroxyl groups. Compound **71b** bearing two methoxy groups on the R_2 and R_4 positions and one hydroxyl group on the R_3 position has better ChE inhibitory activity ($IC_{50} = 0.398 \pm 0.028 \mu M$ for AChE; $IC_{50} = 0.976 \pm 0.102 \mu M$ for BuChE). The compounds of series **72** exhibited moderate inhibitory activities towards both ChEs (**72a**, $IC_{50} = 2.60 \pm 0.37 \mu M$ for AChE; $IC_{50} = 1.08 \pm 0.16 \mu M$ for BuChE), suggesting that the β -diketone bond is not needed to induce inhibition for these analogues. Compounds **70a–c**, **71a–c**, **72a**, and **72b** had the ability to scavenge the ABTS radical with 1.41, 1.81, 1.65, 1.39, 1.78, 2.10, 0.76, and 7. Compounds **70a**, **70b**, **71a**, and **71b**, which bear a phenolic hydroxyl group or methoxy group on R_2 , demonstrated that the group at the ortho position of the phenolic hydroxyl is crucial. Moreover, the IC_{50} of compound **71b** ($IC_{50} = 24.9 \pm 0.4 \mu M$) showed that the locations of the phenolic hydroxyl and methoxy groups are nonadjustable. It can be seen that compounds **71a** ($5.16 \times 10^{-6} \text{ cm s}^{-1}$), **71b** ($7.68 \times 10^{-6} \text{ cm s}^{-1}$), and **71c** ($5.38 \times 10^{-6} \text{ cm s}^{-1}$) might be able to cross the BBB. In addition, these compounds of series **71** are being considered a potential multifunctional neuroprotective agent as a new lead candidate for the treatment of AD [68].

Similar strategy was used by Wang and co-workers on taking advantage of the neurogenic potential profile of melatonin-based hybrids, which are endowed with additional anticholinergic properties. Thus, they designed a novel series of compounds (**74–76**,

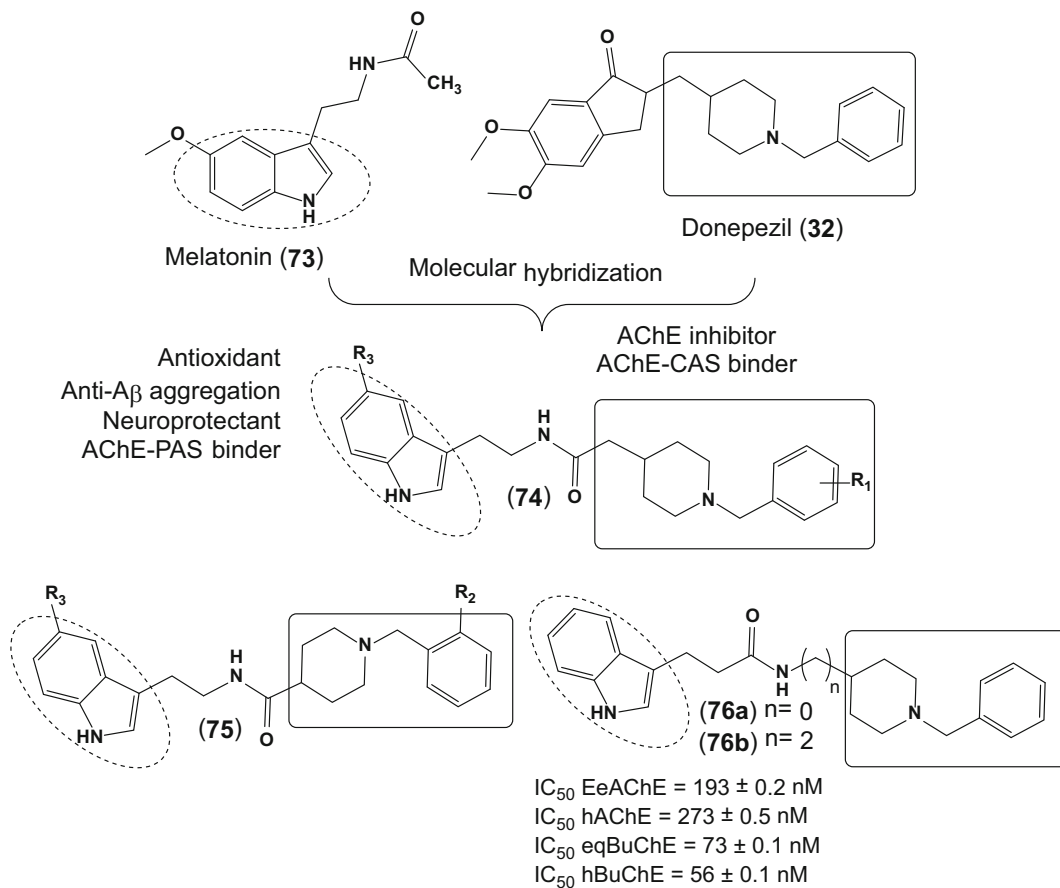


Fig. 23 Design and chemical structures of melatonin–donepezil hybrid compounds **74–76**

Fig. 23) obtained by fusing the *N*-benzylpiperidine moiety of the AChEI donepezil (**32**) and the indole subunit of the antioxidant melatonin **73**. The design was based on the anticipation that melatonin–indole subunit could ensure neuroprotective features and could also interact with the AChE-PAS via π – π aromatic stacking for its aromatic character. On the side, the protonable *N*-benzyl piperidine moiety from donepezil could be responsible for the interaction with the AChE-CAS through cation– π interaction. Therefore the new molecular scaffold could be able to a dual interaction with PAS and CAS of AChE. Biological evaluation of this series led to compound **76** with the most promising multifunctional profile, showing inhibitory activity of AChE ($IC_{50} = 193 \pm 0.2$ nM (EeAChE) and 273 ± 0.5 nM (hAChE)), with higher selectivity for BuChE ($IC_{50} = 73 \pm 0.1$ nM (eqBuChE) and 56 ± 0.1 nM (hBuChE)), along with moderate inhibition of A β_{1-42} self-aggregation (56.3% at 20 μ M), good antioxidant activity (3.28 trolox equivalent by ORAC assay), and

good biometal-chelating ability, also reducing PC12 cells death induced by oxidative stress and adequate BBB permeability [69].

Genistein (77) is a flavonoid found in soybeans and other plants that contain red clover. This substance presents several pharmacological activities, including antioxidant, anti-inflammatory activity, metal chelator, and neuroprotective against A β -protein. These characteristics make 77 a good prototype for planning anti-Alzheimer drugs. However, its inability to inhibit AChE prevents its direct use in the treatment of AD. For this reason, the *N*-benzylpiperidine fragment, that is one of the well-known pharmacophore groups of donepezil (32), of which tertiary amino group is related to adequate AChE inhibition, was used for a rational molecular hybridization with 77 aiming new genistein–donepezil hybrids designed as MTDLs with innovative structural pattern (Fig. 24). Thus, three families of hybrid compounds were prepared, with compounds 78 (IC_{50} hAChE = $5.80 \pm 0.65 \mu\text{M}$), 80 (IC_{50} hAChE = $3.88 \pm 0.64 \mu\text{M}$), and 79 (IC_{50} hAChE = $0.35 \pm 0.03 \mu\text{M}$) being the most active in each family for AChE inhibition. None of the synthesized compounds showed good antioxidant activity. Meanwhile, compound 79, in addition to a significant AChE inhibitory activity, also showed ability in Cu²⁺ chelation, with neuroprotective activity against A β -protein greater than curcumin. Moreover, despite the low antioxidant activity, compound 79 proved to be promising for further studies, as it has high ability to inhibit AChE and moderate ability to inhibit aggregation of A β -protein, both self-induced aggregation and Cu²⁺-induced aggregation [70].

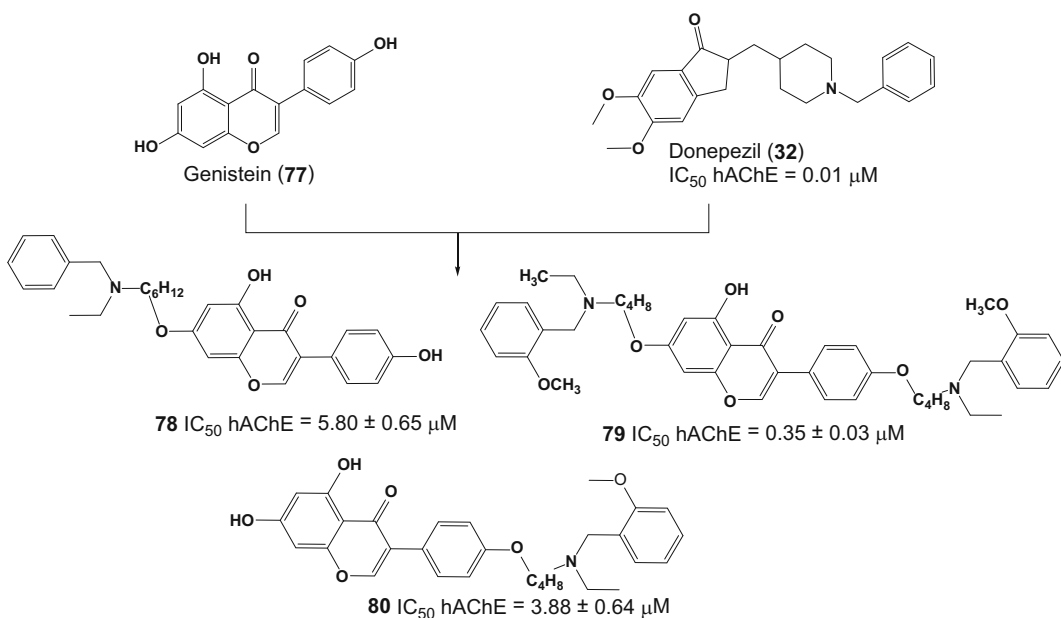


Fig. 24 Design and chemical structure of the multifunctional hybrids genistein–donepezil 78–80

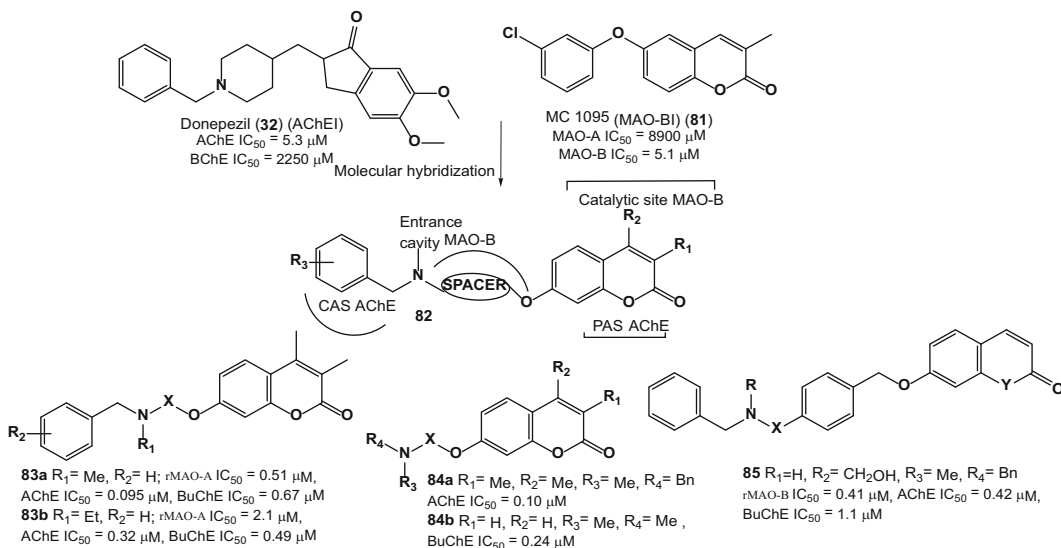


Fig. 25 Chemical structure of new derivatives aminocoumarins **83–85**

Docking-assisted hybridization strategy of aminocoumarins (**81**) and donepezil (**32**) (Fig. 25) was designed and tested by Farina and co-workers over MAO-A and B, AChE, and BuChE. The compounds (**82**) displayed from low to sub-micromolar potencies against rat MAO-B (rMAO-B) and EeAChE, whereas potencies against rat MAO-A (rMAO-A) and EeBuChE were slightly lower. Compounds with the longest linker, such as **83a** and **83b**, displayed the highest inhibitory potency on rMAO-A (IC₅₀ = 0.51 and 2.1 μM, respectively) and ChEs (IC₅₀ AChE = 0.095 and 0.32 μM, respectively, and IC₅₀ BuChE = 0.67 and 0.49 μM, respectively). The compounds with substituents in *meta* and *para* positions of the 7-benzyloxy subunit of the coumarin ring also were evaluated, with 11 substances showing significant inhibition of AChE, and compound **84a** as the most potent AChE inhibitor (IC₅₀ = 0.10 μM). In contrast, only six compounds exhibited a sub-micromolar activity over BuChE, with compound **84b** showing the highest inhibitory potency (IC₅₀ = 0.24 μM). Moreover, compound **85** showed a good inhibitory activity over the three target enzymes with IC₅₀ = 0.41, 0.42, and 1.1 μM for rMAO-B, EeAChE, and EeBuChE, respectively. Thus, some of these multipotent inhibitors, specially compound **84a**, may be considered as promising drug candidate prototypes for further preclinical studies in cognitive and neurodegenerative disease models [71].

Considering a number of recent findings that points a potential role of serotonin receptors in neurophysiology of AD [72, 73], Rochais and co-workers report the synthesis of a novel series of MTDLs displaying DBS inhibition of AChE and a partial 5-HT₄R

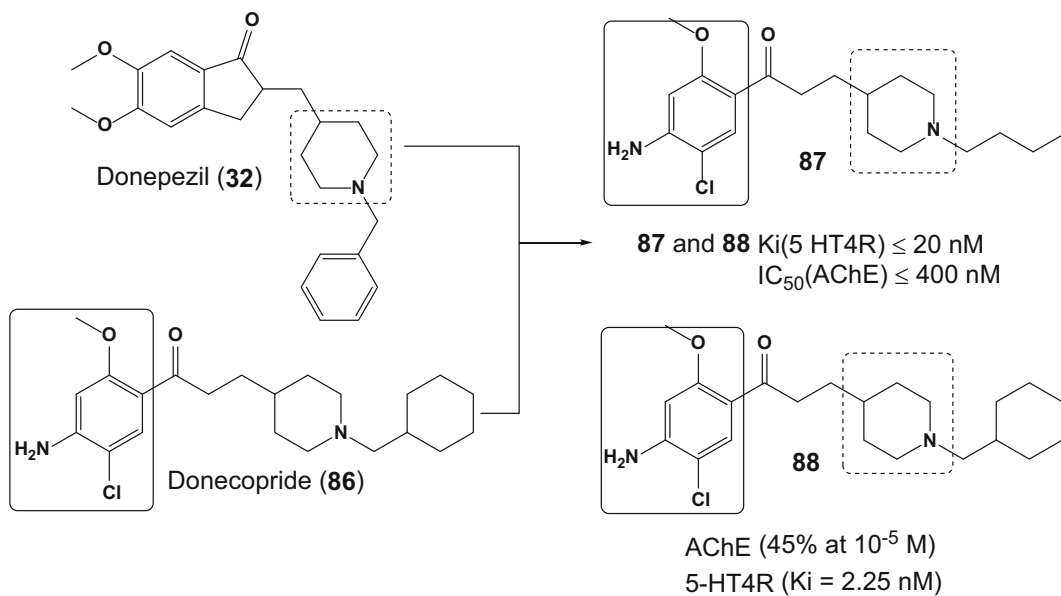


Fig. 26 Design of compounds **87** and **88** by molecular hybridization of donepezil (**32**) and donecopride (**86**)

agonist activity (Fig. 26). In spite of the AChE inhibitor donepezil (**32**) and donecopride (**86**), a novel drug candidate exhibiting, for the first time, both in vitro DBS AChE inhibitory activity and a serotonergic subtype 4 receptor (5-HT4R) partial agonist effect [74], was used as molecular model in the design approach of the new hybrid compounds **87** and **88**. The compound **87** could be considered as MTDL since they exhibit both a $K_i(5\text{-HT4R}) \leq 20\text{ nM}$ and an $IC_{50}(\text{AChE}) \leq 400\text{ nM}$. Already, the compound **88** was a potent inhibitor of AChE (45% at 10^{-5} M), but it is almost devoid of affinity for 5-HT4R ($K_i = 2.25\text{ nM}$), as is the case for DPZ. These results seem to confer to compound **87** (Fig. 26) a greater symptomatic and disease-modifying effect in relation to donecopride (**86**), thus designing it as a potentially promising drug candidate in AD [75].

Mishra and co-workers synthesized a novel series of (*E*)-2-(4-(4-(substituted) piperazin-1-yl)benzylidene)-5,6-dimethoxy-2,3-dihydro-1*H*-inden-1-one (**90–92**) based donepezil (**32**) multifunctional agents (Fig. 27). In vitro studies revealed that these compounds demonstrated moderate to good AChE and A β aggregation inhibitory activity and these derivatives are also endowed with admirable antioxidant activity. The compounds **90**, **91**, and **92** appeared as most active multifunctional agents and displayed marked AChE inhibitory, A β disaggregation, and antioxidant activity. Studies indicate that **91** and **92** showed better AChE inhibitory activity than the standard drug donepezil **32** ($IC_{50} \pm SD$ (μM) = 0.039 ± 0.002) with $IC_{50} \pm SD$ (μM) = 0.034 ± 0.002 and $IC_{50} \pm SD$ (μM) = 0.025 ± 0.001 , respectively. The compounds **90** ($IC_{50} \pm SD$ (μM) = 7.4 ± 0.10), **91** ($IC_{50} \pm SD$

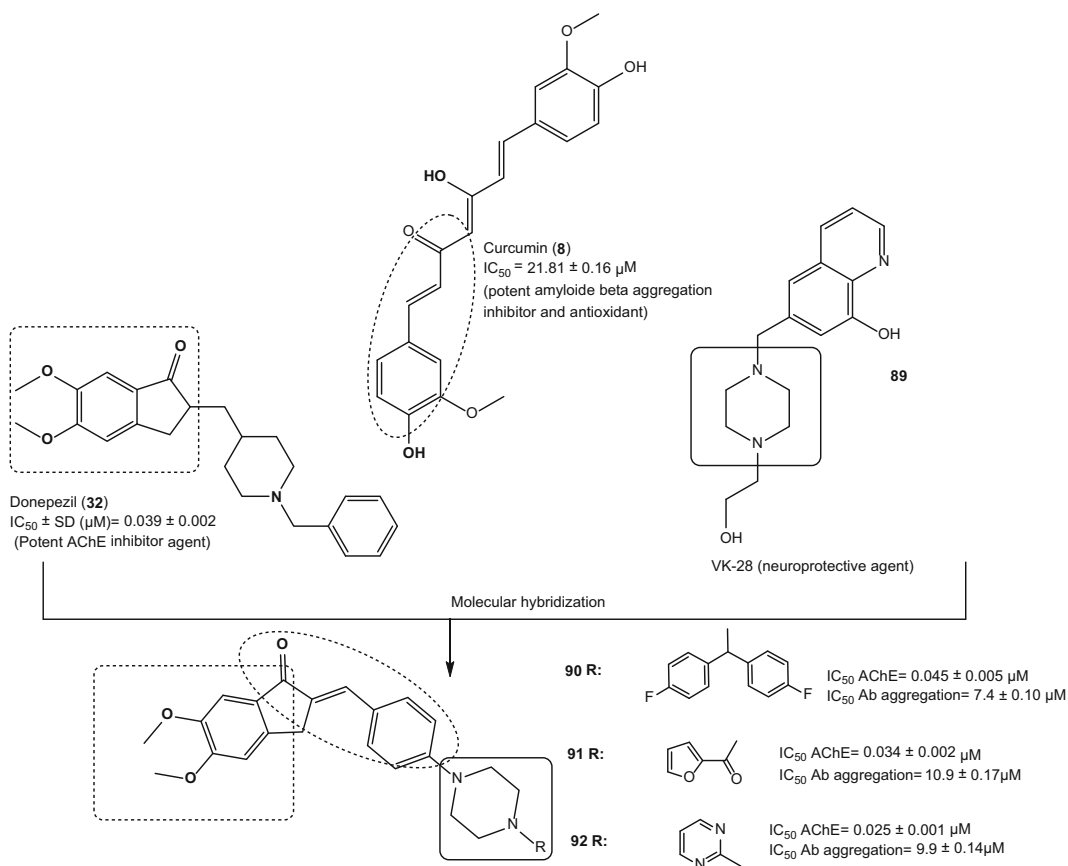


Fig. 27 Molecular framework of designed novel hybrid derivatives of donepezil-based multifunctional agents

(μM) = 10.9 ± 0.17) as well as **92** ($IC_{50} \pm SD (\mu M) = 9.9 \pm 0.14$) exhibited better $A\beta$ aggregation inhibitory activity than curcumin ($IC_{50} = 21.81 \pm 0.16 \mu M$). These three compounds successfully diminished H_2O_2 -induced oxidative stress in SH-SY5Y cells and displayed excellent neuroprotective activity against H_2O_2 as well as $A\beta$ induced toxicity in SH-SY5Y cells in a concentration dependent manner. Moreover, these derivatives did not exert any significant toxicity in neuronal SH-SY5Y cells in cytotoxicity assay [76].

7 Multi-target Directed Ligands Inspired by Tacrine

Tacrine (THA, Cognex[®], **63**) was the first cholinesterase inhibitor approved by the FDA for the treatment of AD. However, in a few years it was banned in some countries due to its high hepatotoxicity and low bioavailability. In recent years, tacrine **63** has been widely used as a scaffold for the development of new multifunctional agents [34, 77, 78].

Contelles and co-workers synthesized a new series of 4*H*-pyrano[2,3-*b*]quinolone derivatives (**93**), planned by molecular

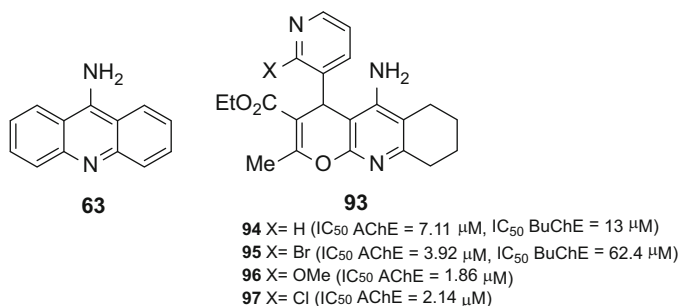


Fig. 28 Chemical structure of tacrine (**63**) and tacrine-hybrid derivatives (**94–97**)

hybridization of tacrine fused to a benzylamino-3-pyridyl ring moiety, aiming to obtain new AChEIs with an innovative molecular scaffold and to explore its pharmacophoric contributions to additional neuroprotective properties and modulatory effects on voltage-dependent Ca⁺² channels. In general, the new tacrine-hybrid compounds showed to be better selective for AChE inhibition than tacrine. It seems that the substitution at C-20 position of the 3-pyridyl ring system (compounds **95**, **96**, and **97**, Fig. 28) was deleterious for AChE inhibition as observed for derivative **94** (X=H). In contrast, compounds **95–97** showed a good neuroprotective profile, as the derivative **96** was exhibiting the best neuroprotective effect, with only 46% of cell-death suppression [79].

The dimer bis-(7)-tacrine (**102**, Fig. 29) was one of the first homodimers reported in the literature with increased AChE affinity, exhibiting a 1,000-fold higher inhibitory potency than tacrine. This best inhibitory profile is due to a dual simultaneous interaction with active and peripheral sites of AChE. Further studies demonstrated that this compound was also capable to inhibit the AChE-induced Aβ aggregation with an IC₅₀ of 41.7 μM. Because of this symbiotic profile, compound **102** has been described as an important structural model for rational design of novel MTDLs [80].

In 2007, Bolognesi and co-workers proposed modifications on the structure of bis-(7)-tacrine (**102**) with the insertion of functional groups capable of chelate metals (**100**, **101**, **103**), which are involved in the degenerative process. For this purpose, compounds **98** (BW284c51) and ambenonium (**99**) were selected as model prototypes due to their cholinesterase inhibitory activity and their singular structural feature, carrying carbonyl and oxalamide functionalities which could lead to the desired metal chelation property (Fig. 29). The multifunctional compounds maintained potent AChEIs in the nanomolar range, which showed additional inhibitory effect of AChE-induced amyloid-β aggregation. The results showed the ability for chelation of Cu²⁺ and Fe³⁺ ions and suggest that these compounds might act against AD by a chelation mechanism [35].

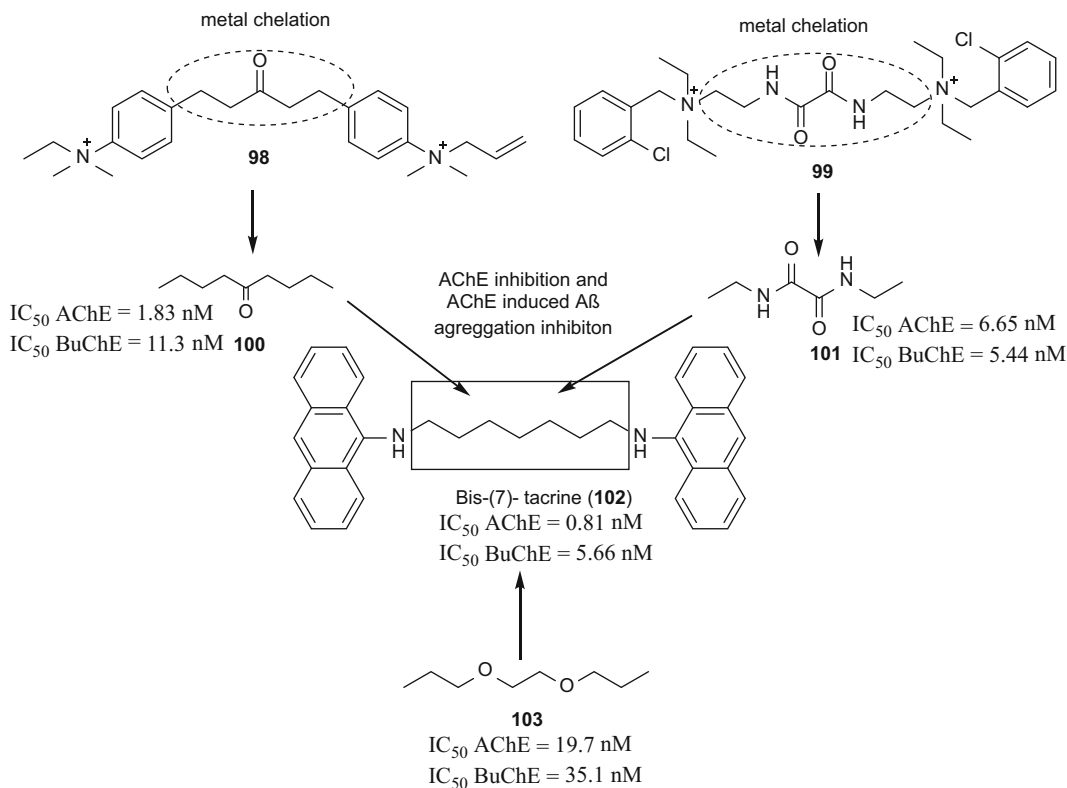


Fig. 29 Design of the new series of derivatives of bis-(7)-tacrine (**102**) with improved activities for inhibition of AChE inhibition and AChE-induced amyloid- β aggregation, and ion metal chelation

Minarini and co-workers also explored the structure of bis-(7)-tacrine (**104**), in an attempt to obtain the hybrid derivative cystamine bis-(7)-tacrine (**105**), with an insertion of the cystamine moiety **106** as a linker between two acrydine subunits of the bis-tacrine framework of **104** (Fig. 30). Cystamine (**105**) has shown important biological properties such as antioxidant and neuroprotection. This study revealed that the hybrid dimer cystamine-tacrine **105** was able to inhibit AChE (IC₅₀ = 5.04 nM), BuChE (IC₅₀ = 4.23 nM), self-A β aggregation (IC₅₀ = 24.2 μ M), and AChE-induced A β aggregation (52.6%) in the same range of the reference compound, with additional neuroprotective effect on SH-SY5Y cell line against H₂O₂-induced oxidative injury, with low toxicity [80].

Fang and co-workers reported in 2008 the synthesis and pharmacologic evaluation of a new series of tacrine-ferulic acid hybrid derivatives (**69**) (Fig. 31). They planned a new structural pattern by the connection of ferulic acid (**107**) to the structure of tacrine (**63**) via an alkylenediamine as a linker side chain. This approach was used aiming to obtain new multi-target hybrid compounds which could conjugate the AChE inhibitory activity originating from the tacrine

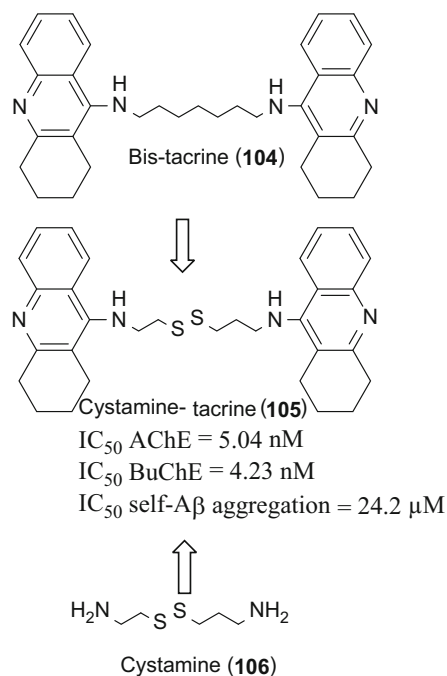


Fig. 30 Design of the multipotent cystamine–tacrine dimer (**105**) by molecular hybridization of bis-(7)-tacrine (**104**) and cystamine (**106**)

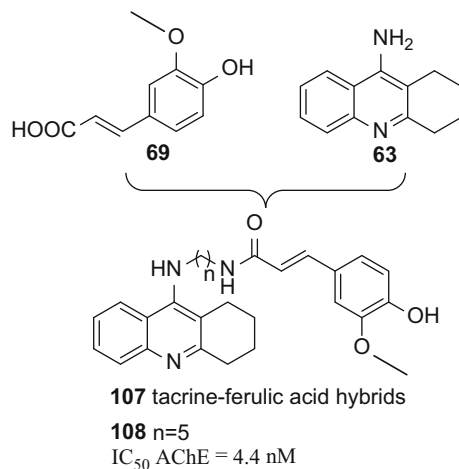


Fig. 31 Design of the new series of hybrid tacrine–ferulic acid (**107**) derivatives with dual antioxidant and IChE properties

template and the antioxidant activity from the ferulic acid moiety. The compounds showed lower antioxidant activity than ferulic acid, but all hybrids tested showed moderate to good antioxidant activity. All compounds effectively inhibited AChE in in vitro assays. Particularly, compound **108** showed a tenfold higher

AChE inhibitory activity ($IC_{50} = 4.4$ nM) than tacrine. Enzymatic kinetics studies suggest that this compound possesses high affinity binding to the PAS of AChE and, therefore, could inhibit A β -peptide fibril formation [81].

In another approach, Rosini and co-workers also used tacrine as a model for the planning of a new series of hybrid compounds (**110–113**) based on the structural feature of carvedilol (**109**). This compound showed a significant neuroprotective efficacy, probably related to its modulatory action as a low-affinity antagonist of *N*-methyl-D-aspartate (NMDAR). Thus, carbazole moiety of **109** was elected as a probable structural pharmacophoric subunit responsible by its antioxidant properties and efficient inhibitory activity of A β -fibril formation (Fig. 32). Biological evaluation revealed that all compounds showed effective AChE-inhibiting activity in a nanomolar range, being more potent than tacrine. Particularly, compounds **110–113** showed an adequate biological profile for multi-target drugs, inhibiting AChE activity and also capable of blocking in vitro A β self-aggregation, AChE-mediated A β -aggregation, NMDARs antagonistic effect, and to reduce cerebral oxidative stress [6].

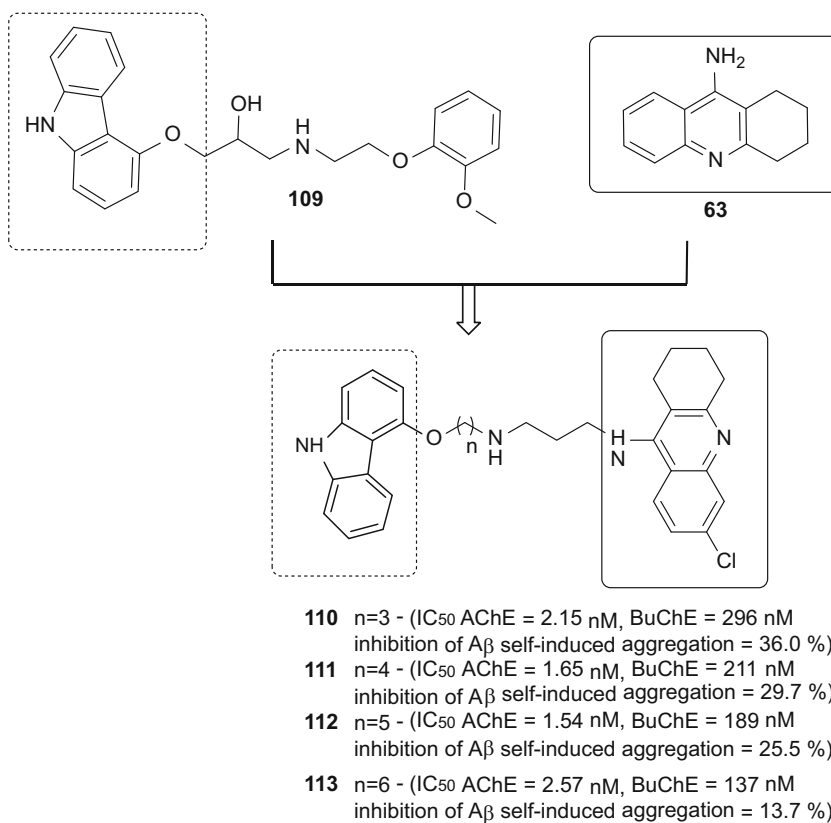


Fig. 32 Chemical structure of carvedilol (**109**) and their active derivatives **110–113**

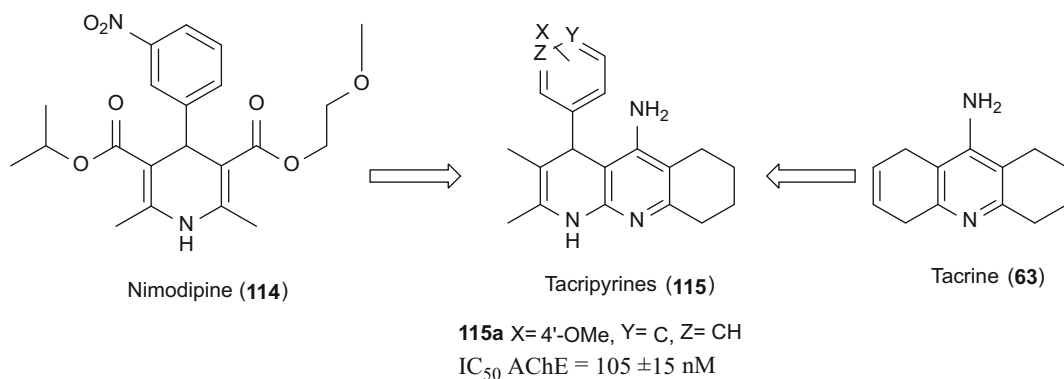


Fig. 33 Structure of tacripyrines (**40**) and the most active compound **115a** designed by molecular hybridization of tacrine (**63**) and nimodipine (**114**)

In another approach, tacripyrines (**115**) were designed by combining tacrine (**63**) with a calcium antagonist such as nimodipine (**114**) (Fig. 33). In this study, the AChE inhibitory activity of the nimodipine–tacrine hybrids **115** is potentiated, with IC_{50} in the nanomolar range for AChE. Besides acting as an inhibitor of AChE, these compounds also showed significant neuroprotective effects. In particular, compound **115a** ($IC_{50} = 105 \pm 15$ nM) was moderately active in AChE-induced and $A\beta$ -self-aggregation models (30 and 35% inhibition, respectively). Further kinetic studies and molecular modeling data suggest that compound **115a** could interact with the PAS of the AChE, exerting its inhibitory properties in noncompetitive mode. Additionally, most of the compounds showed moderate Ca^{2+} channel blocking effect and therefore the compounds are neuroprotective agents [82].

A novel series of tacrine–caffeic acid hybrids (**116**) were designed by combining the structure of caffeic acid (**117**) and tacrine (**63**) (Fig. 34) as MTDL against Alzheimer's disease. In vitro studies showed that most of the target molecules exhibited significant antioxidant activities on oxygen radical absorbance capacity (ORAC) method. The hybrid molecules showed to be good metal chelators and exhibited higher antioxidant capacity than the caffeic acid with values of 2.75–9.37 μ M. In particular, compound **118** exhibited the greatest ability in inhibition of self- or AChE-induced β -amyloid_{1–40} aggregation as well as showed potent neuroprotective effects against H_2O_2 - and glutamate-induced cell death, with low toxicity in HT22 cells. Further molecular modeling and enzymatic kinetics studies clarified that these inhibitors may act in two binding sites of AChE, which could explain the inhibitory effect exerted over $A\beta$ -aggregation [83].

In another work, Wang and collaborators reported a series of tacrine-based **63** hybrid compounds, containing an additional pharmacophoric subunit 4-phenyl-2-aminothiazole with different

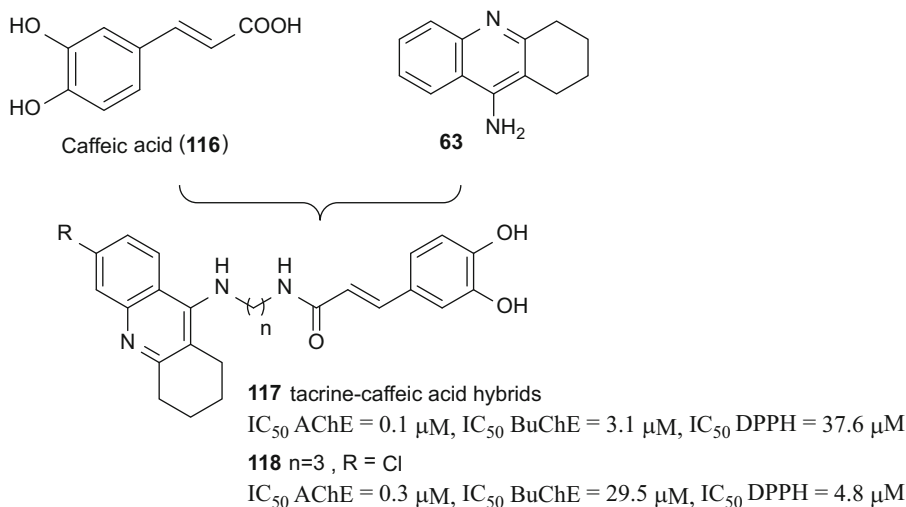


Fig. 34 Chemical structure of caffeic acid (**116**) and their hybrid derivatives **117** and **118**

spacer subunits. The pharmacophoric grouping 4-phenyl-2-aminothiazole **119** was elected based on data from the literature that suggest this functionality as responsible for inhibitory effects on tau protein aggregation and on A β self-aggregation, neuroprotection, and anti-inflammatory. All ten compounds of the series of phenylthiazole–tacrine hybrids (**120**, Fig. 35) were potent inhibitors of cholinesterases with IC_{50} values ranging from 5.78 ± 0.05 to 7.14 ± 0.01 nM for AChE, and from 5.75 ± 0.03 to 10.35 ± 0.15 nM for BuChE. Moreover, a structure–activity relationship study was conducted to study the influence on the type of middle linker and substitutions at 40-position of 4-phenyl-2-aminothiazole. The results were indicative that the length of middle linker affected the AChE inhibitory potency. Compound **121**, with the largest linker subunit exhibited, showed a greater potency in AChE inhibition ($IC_{50} = 7.14 \pm 0.01$ nM). Data from kinetic studies have shown that this compound is a mixed-type inhibitor and could bind simultaneously at the catalytic and the peripheral sites of AChE. Most of phenylthiazole–tacrine hybrids showed a good inhibitory potency on A β_{1-42} self-aggregation, although the activity was lower when compared with subunit 4-phenyl-2-aminothiazole. Additionally, compound **121** displayed a blockade effect on Ca²⁺ overload in the primary cultured cortical neurons [84].

Lan and co-workers outlined novel tacrine– β -carboline hybrid compounds (**123**), designed from the structural feature of natural β -carboline-(pyrido[3,4-b]-indoles (**124**) and tacrine (**63**) (Fig. 36). Tacrine was used for the inhibition of ChE through its binding ability to the CAS of AChE, while β -carboline (**122**) was used for its potential interaction with the PAS due to its aromatic

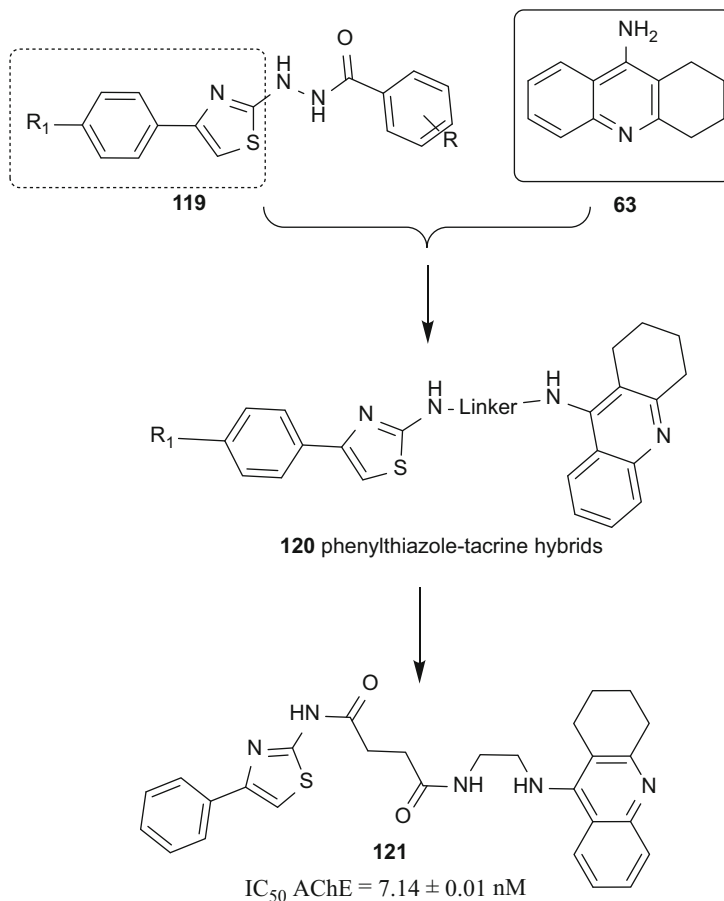


Fig. 35 Design of the phenylthiazole–tacrine hybrid compounds **120** and structure of the most active derivative **121**

character. In vitro evaluation showed that most of the target compounds exhibited significant inhibition of AChE (EcAChE and hAChE), BuChE, and self-induced A β aggregation, along with Cu²⁺-induced A β _{1–42} aggregation and metal-chelating ability. Compound **122** showed the greatest ChE inhibitory activity (IC₅₀ = 21.6 ± 0.8 nM (ecAChE), 63.2 ± 2.5 nM (hAChE) and 39.8 ± 1.6 nM (BuChE), good inhibition of A β aggregation (65.8% at 20 μ M), and good antioxidant activity (1.57 trolox equivalents). Kinetic and molecular modeling studies indicated that compound **122** was a mixed-type inhibitor, binding simultaneously to CAS and PAS of AChE. Compound **122** was also capable to reduce PC12 cells death induced by oxidative stress and penetrate the BBB in vitro [85].

Chioua and co-workers also used tacrine (**63**) as scaffold for planning a series of eight tacrine–pyranopyrazole derivatives (Fig. 37). All compounds were initially submitted to toxicity tests by using cell viability assay with Hep G2 cells (300 μ M). The results

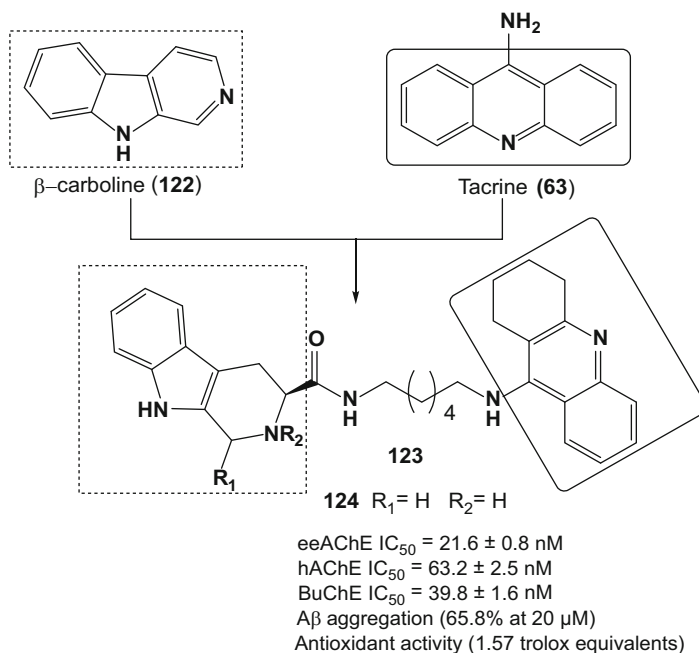


Fig. 36 Structural design of a novel series of tacrine- β -carboline hybrids (**124**) from tacrine (**63**) and natural β -carboline (**122**)

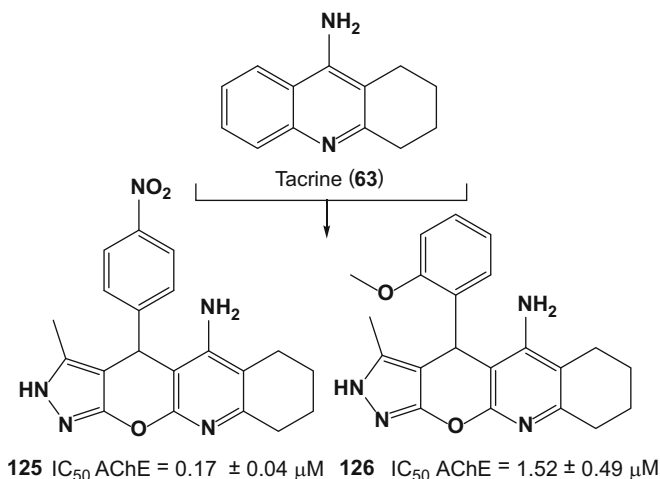


Fig. 37 Chemical structures of novel tacrine-pyranopyrazole multifunctional hybrid compounds **125** and **126**

showed that all tacrine derivatives were safe, with cell viabilities of 76–95%, in comparison to tacrine (40%). Then, AChE inhibition assay was carried out and compounds **125** ($IC_{50} = 0.17 \pm 0.04$ μ M) and **126** ($IC_{50} = 1.52 \pm 0.49$ μ M) showed the best results, with compound **125** being almost ninefold more potent than **126**, with 82% of cell viability. Both compounds were additionally tested for inhibition of A β aggregation and were able to reduce by half the

amount of A β aggregates induced by AChE. Compound **125** was evaluated for its neuroprotective activity, showing 164% of cellular viability in cortical neurons after 24 h of treatment (10 μ M). This activity showed to be dose dependent, reaching its maximum effect at 10 μ M with an $EC_{50} = 0.25 \pm 0.06 \mu$ M. Taking all these results, compound **125** was elected as the most promising tacrine–pyranopyrazole derivative, with a genuine structural framework and a singular and safe multifunctional mode of action, inhibiting AChE and AChE-induced A β aggregation, with neuroprotective activity [86].

Ferulic acid (**69**) acid was also used as molecular scaffold by Fu and co-workers in the design of a new series of tacrine–ferulic acid hybrids (Fig. 38) with expected multi-target effects in the inhibition of cholinesterases, reduction of self-induced β -amyloid (A β) aggregation, chelation of Cu^{2+} , and neuroprotection. Among all the synthesized compounds, **127** and **128** displayed the highest selectivity in inhibiting AChE over BuChE (SI = 4.087 and 1.733, respectively). Moreover, compound **128** also showed dramatic inhibition of self-A β aggregation ($37.2 \pm 0.9\%$ at 20 μ M), Cu^{2+} chelating activity, and the best activity against A β -induced neurotoxicity in neuro-2A cells [87].

Benchekroun and co-workers selected ferulic acid (**69**) and melatonin (**73**) to draw multifunctional tacrine-derived hybrid compounds for AD. Melatonin (**73**) is a molecule produced by

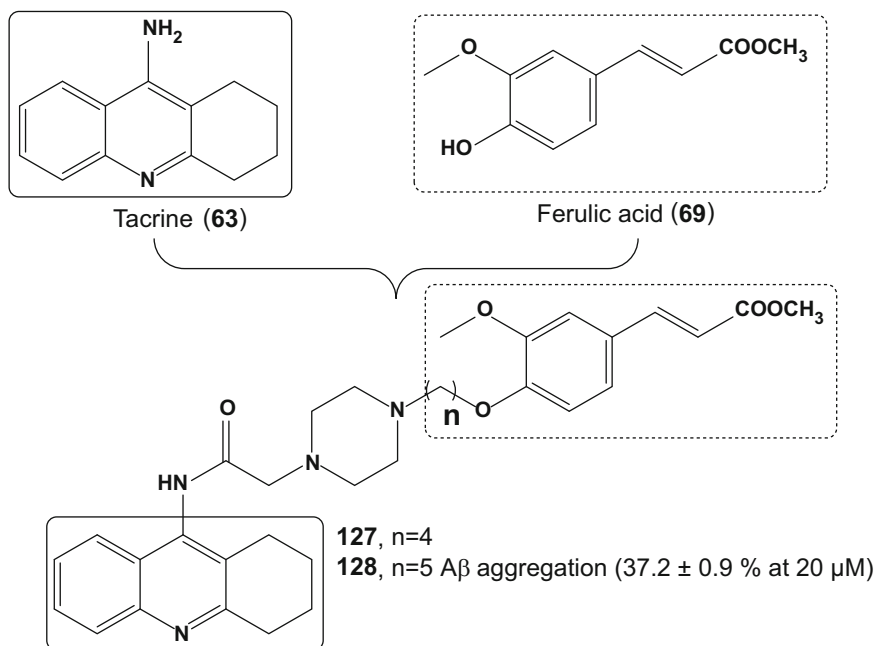


Fig. 38 Design and chemical structures of new multifunctional tacrine–ferulic acid hybrid compounds **127** and **128**

various human organs and tissues and is involved in many physiological processes, such as the modulation of endogenous antioxidants and immune system regulation. With aging, human beings have a natural decline in the levels of melatonin, which has been associated with the development of neurodegenerative diseases such as AD. Melatonin (**73**) has been shown to be capable of quenching free radicals, stimulating the biosynthesis of antioxidant enzymes, reducing the hyperphosphorylation of neurofilaments and protective activity against A β protein. Moreover, melatonin could induce the proliferation and differentiation of neural cells in the hippocampus of adult mice. Given these data, new ferulic acid–tacrine–melatonin hybrids (FATMHs) (**129**, Fig. 39) were synthesized and evaluated for their abilities in the inhibition of AChE and BuChE, toxicity towards HepG2 cells, neuroprotection in SH-SY5Y cells, Nrf2 pathway induction in AREc32 cells, BBB permeability as well as their antioxidant capacity through ORAC. Among all tested compounds, **129** was identified as a most potent human cholinesterase inhibitor (IC_{50} hAChE = $1.29 \pm 0.070 \mu\text{M}$; IC_{50} hBuChE = $0.234 \pm 0.008 \mu\text{M}$) and strong antioxidant ($9.11 \pm 0.21 \mu\text{M}$), with adequate BBB permeability. This compound also showed the best in vitro neuroprotective profile against toxic insults mediated by H_2O_2 (300 μM), A β_{1-40} and A β_{1-42}

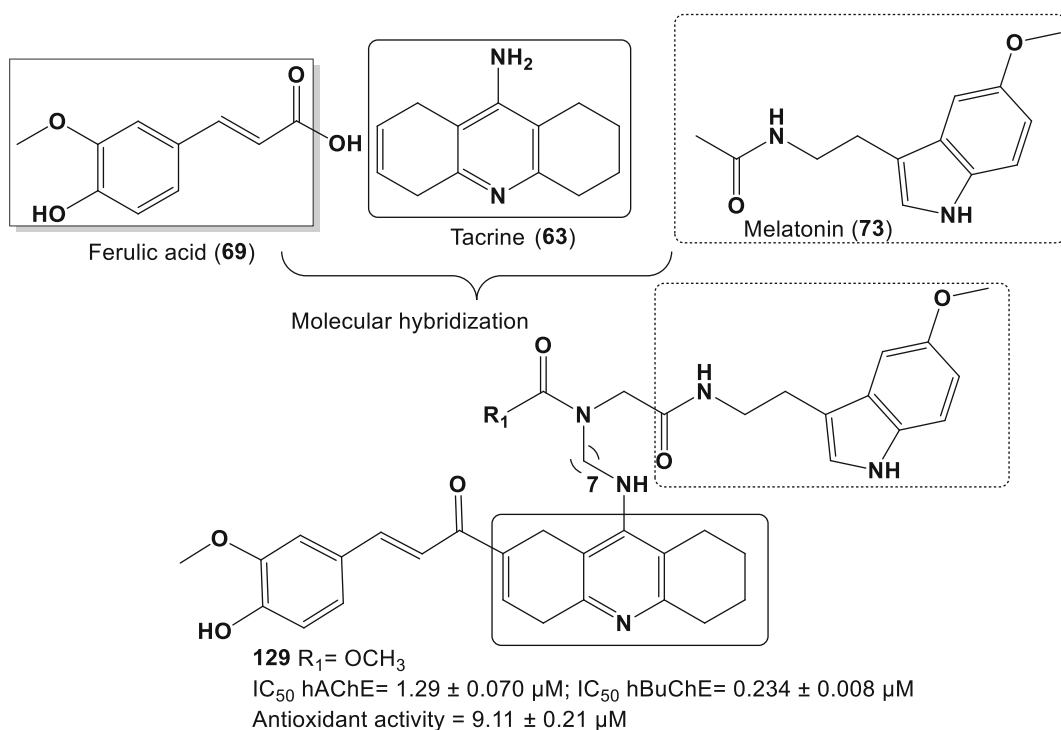


Fig. 39 Design of the new MTDL based on the structure of tacrine (**63**), ferulic acid (**69**), and melatonin (**73**)

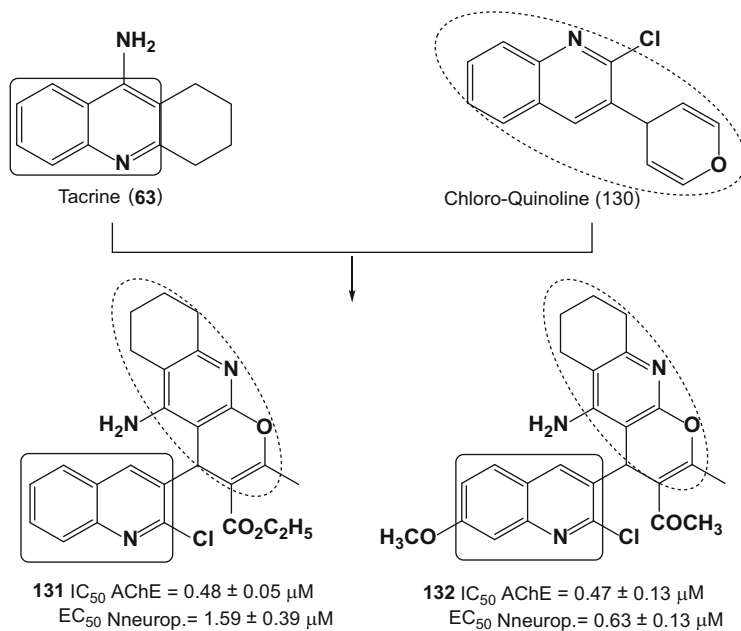


Fig. 40 Tacrine–quinoline hybrids and more potent derivatives **131** and **132**

($30 \mu\text{M}$) at $1 \mu\text{M}$, with significant induction of Nrf2 transcriptional pathway (at $3 \mu\text{M}$) in AREc32 cells [88].

Quinolines are also considered privileged heterocyclic structures present in different natural or synthetic bioactive substances used in the treatment of different pathologies, including neurodegenerative diseases. Thus, chloro-quinoline **130** was used for molecular hybridization with the structure of tacrine (**63**) to generate a new family of tacrine–quinoline hybrids, designed as AChE inhibitors and neuroprotector MTDL candidates (Fig. 40). Thirteen new substances were synthesized and initially submitted to cell viability assays with Hep G2 cells, in order to verify their safety and their hepatotoxicity. Almost all substances showed cytotoxicity lower than tacrine ($EC_{50} = 179 \mu\text{M}$) and were then evaluated for their AChE inhibitory properties. The lowest toxic compound **131** exhibited a very significant AChE inhibition (IC_{50} AChE = $0.48 \pm 0.05 \mu\text{M}$), similarly to its methoxy derivative **132** (IC_{50} AChE = $0.47 \pm 0.13 \mu\text{M}$), were identified as the most AChEI inhibitors. All molecules were also subjected to neuroprotection assays for cell death mediated by oxidative stress and, once again, compounds **131** ($EC_{50} = 1.59 \pm 0.39 \mu\text{M}$) and **132** ($EC_{50} = 0.63 \pm 0.13 \mu\text{M}$) have stood out and were considered as innovative candidates for the development of novel MTDLs for AD therapeutics [89].

Tacrine (**63**) and resveratrol (**133**) were used as molecular prototypes for molecular hybridization in the design of tacrine–resveratrol hybrids with improved biological properties

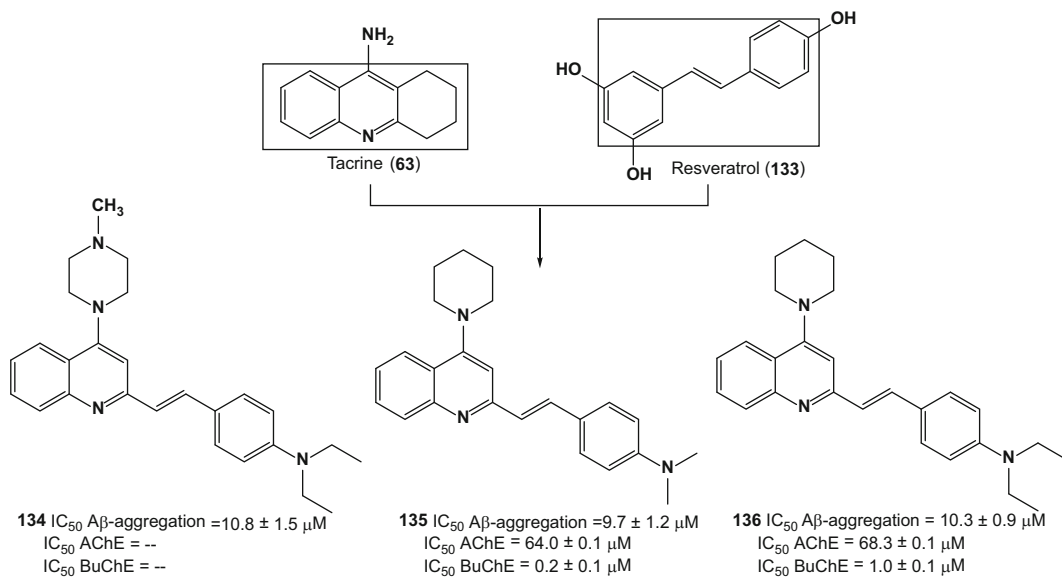


Fig. 41 Design, chemical, and biological data of tacrine–resveratrol hybrids **134–136**

against AD (Fig. 41). In this context, 36 hybrids were synthesized, and initially tested for their ability to inhibit A β aggregation and antioxidant activity. Compounds **134–136** were the most active in the inhibition of A β aggregation with IC_{50} values of $10.8 \pm 1.5 \mu\text{M}$, $9.7 \pm 1.2 \mu\text{M}$, and $10.3 \pm 0.9 \mu\text{M}$, respectively. However, the antioxidant activity of all hybrids was lower than that of resveratrol. Compounds **135** and **136** were evaluated towards ChE activity, both exhibited a moderate activity and slightly selective by BuChE with IC_{50} values of $64.0 \pm 0.1 \mu\text{M}$ (AChE), $0.2 \pm 0.1 \mu\text{M}$ (BuChE) and $68.3 \pm 0.1 \mu\text{M}$ (AChE), $1.0 \pm 0.1 \mu\text{M}$ (BuChE), respectively [90].

Coumarins are an important class of natural metabolites with a wide spectrum of biological properties, and recently they have been drawing special attention due to their biological activity related to neuronal disorder [91]. Coumarin derivatives have been reported as excellent MAO inhibitors, especially any 7-substituted coumarin [92]. The substituents at three and/or four positions may increase the selectivity to MAO-B, along with interaction with the AChE-PAS [93, 94]. Therefore, a series of hybrids of tacrine (**63**) and coumarin (**137**) was designed aiming novel dual inhibitors of AChE and MAO-B (Fig. 42). Among 20 new tacrine–coumarin hybrids, in vitro evaluation disclosed compounds **138** (IC_{50} AChE = $16.11 \pm 0.09 \text{ nM}$, IC_{50} BuChE = $112.72 \pm 0.93 \text{ nM}$, IC_{50} MAO-A = $15.07 \pm 0.88 \mu\text{M}$, and IC_{50} MAO-B = $0.24 \pm 0.01 \mu\text{M}$) and **139** (IC_{50} AChE = $24.37 \pm 0.23 \text{ nM}$, IC_{50} BuChE = $124.32 \pm 1.19 \text{ nM}$, IC_{50} MAO-A = $53.07 \pm 1.24 \mu\text{M}$, and IC_{50} MAO-B = $0.70 \pm 0.02 \mu\text{M}$) showing

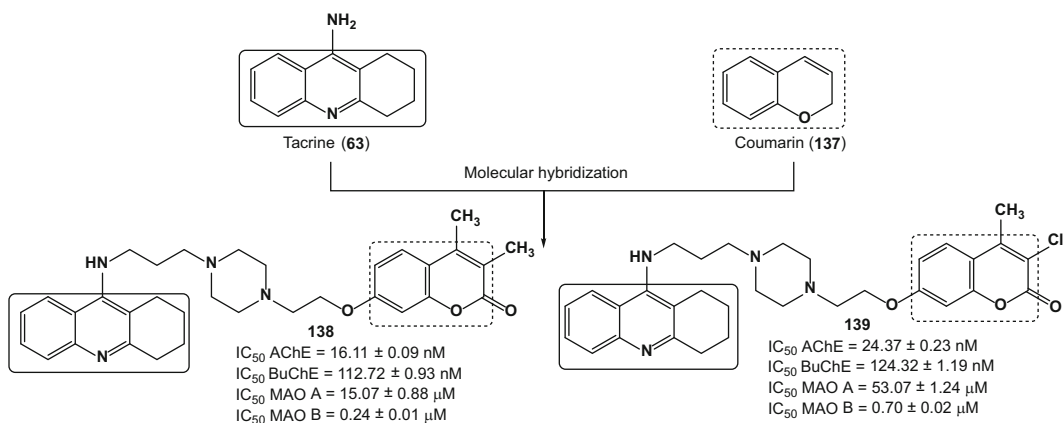


Fig. 42 Design of the dual AChE and MAO-B inhibitors **138** and **139** by molecular hybridization of tacrine (**63**) and coumarin **137**

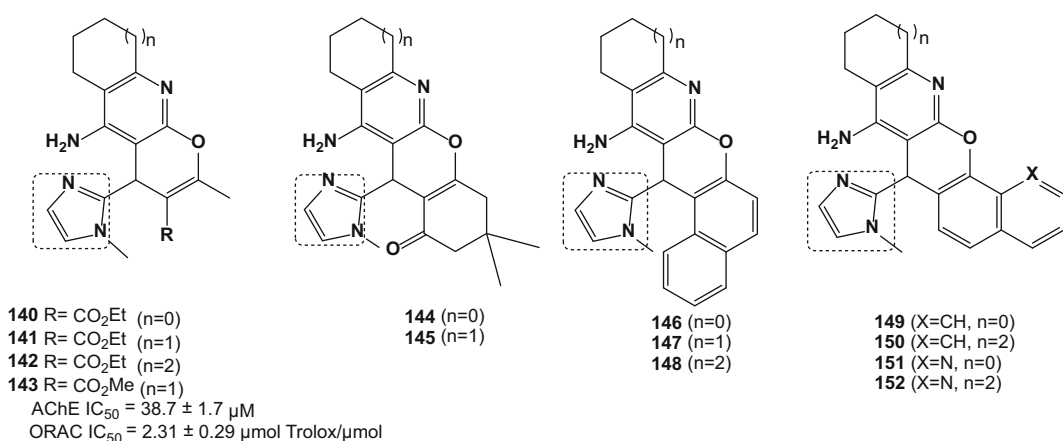


Fig. 43 Chemical structures of tacrine analogues **140–152** designed by molecular optimization of the imidazopyrroloquinoline derivative

the desired dual inhibitory profile, with good BBB permeability [95].

Boulebd and co-workers described the synthesis and in vitro biological evaluation of 13 new racemic and diversely functionalized imidazolyl-pyrano-tacrine derivatives (**140–152**) as non-hepatotoxic multipotent compounds (Fig. 43). Compound **143** was highlighted for its selective, but moderate AChE inhibitory activity (IC₅₀ = 38.7 ± 1.7 μM), and a very potent antioxidant activity on the basis of the ORAC test (2.31 ± 0.29 μmol Trolox/μmol compound). In the search for enhanced inhibitory and antioxidant properties, other tacrine analogues, such as compounds **147** and **150**, were obtained significantly more potent AChEIs than **143**, with high antioxidant activity, but, unfortunately, a high

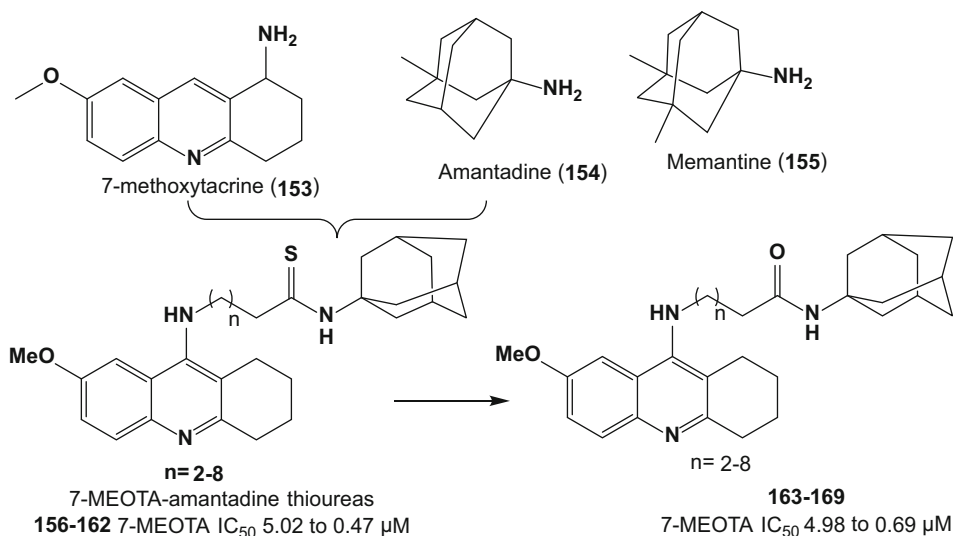


Fig. 44 Design of compounds 7-methoxytacrine (7-MEOTA)-amantadine ureas (**156–162**) and series of 7-MEOTA-amantadine thioureas (**163–169**)

hepatotoxicity comparable to tacrine **63** at high concentrations ($\geq 300 \mu\text{M}$) [96].

Spilovska and co-workers reported the synthesis and pharmacologic evaluation of a new series of cholinesterase inhibitors acting as DBS heterodimers for the management of AD. Based on the MTDL strategy by Simoni et al. [43], a series of 7-MEOTA-amantadine urea-linked derivatives were designed, synthesized, and evaluated as AChE/BuChE inhibitors (Fig. 44). All 14 7-MEOTA-amantadine derivatives were compared with 7-MEOTA-amantadine thioureas, 7-MEOTA (**153**), tacrine **63**, and galanthamine–memantine dimers. The new hybrids showed to be more potent hAChE and hBuChE inhibitors than 7-MEOTA with IC_{50} values ranging from 5.02 to 0.47 μM for thioureas (**156–162**) and from 4.98 to 0.69 μM for urea derivatives (**163–169**). In the 7-MEOTA-amantadine thioureas, series of five compounds had IC_{50} values in sub-micromolar range for hAChE. Only two derivatives (**156** and **158**) exhibited inhibition potency in sub-micromolar range for hAChE. However, compounds **156**, and **158–162** showed sub-micromolar inhibition potency for hBuChE. The best inhibitory activity was identified for compound **158** with IC_{50} of $0.11 \pm 0.02 \mu\text{M}$, bearing five methylene groups in the linker, that displayed inhibitory potency in the same magnitude order of THA for hAChE [97].

Based on the remarkable capacity of tacrine to interact in the active site of AChE, Martins and collaborators proposed a new series of tacrine-derived analogues by the insertion of a furo [2,3-*b*]quinolin-4-amine (**170**) and pyrrolo[2,3-*b*]quinolin-4-amine (**171**) subunits (Fig. 45). Pharmacological results revealed

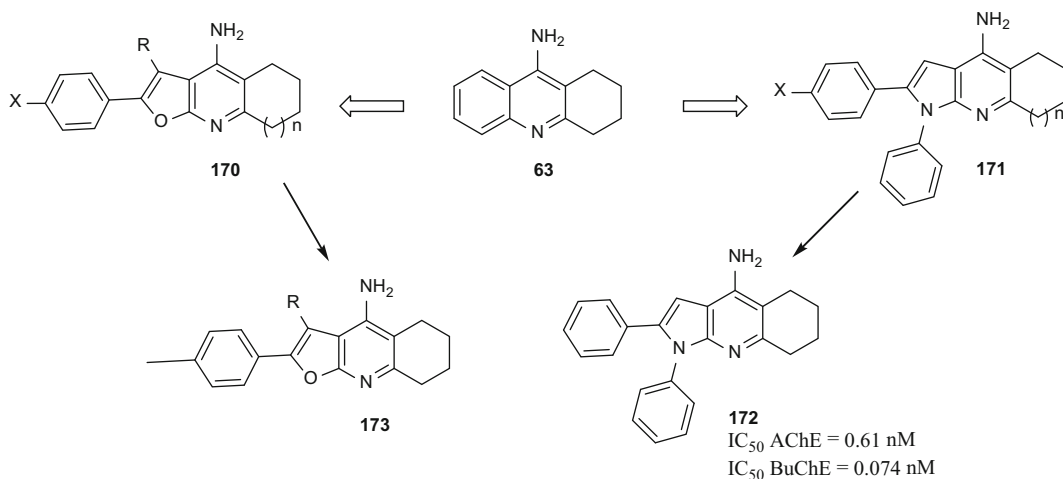


Fig. 45 Chemical structure of the compound **172**, the lead molecule forms the pyrrolo[2,3-b] quinolin-4-amine series

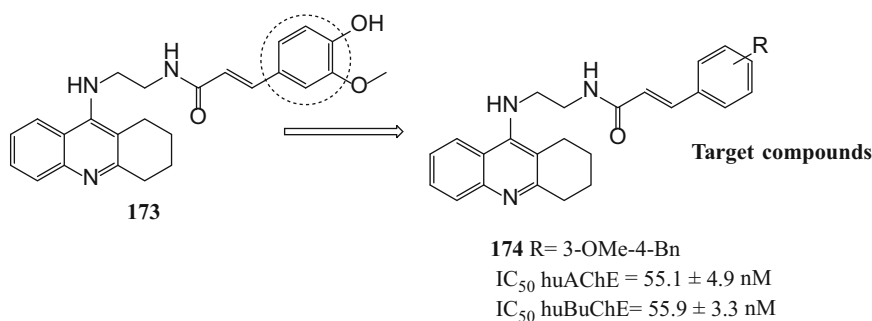


Fig. 46 Design strategy and the compound **174**, one of the most potent analogues

that the furanotacrines and pyrrolo[2,3-b]quinolin-4-amine derivatives showed selective inhibition of BuChE, in a micromolar range, but these compounds have lower potency than tacrine. The lead compound **172** exhibited the best inhibitory profile with IC_{50} AChE = 0.61 nM and IC_{50} BuChE = 0.074 nM, for AChE from *Electrophorus electricus* and BuChE from equine serum, with significant neuroprotective effects against A β -induced toxicity in concentrations below 300 nM. Molecular modelling studies have corroborated that the hybrid derivatives containing pyrrole or furan ring subunits interact with AChE and BuChE and the presence of a phenyl ring at the position 1 of the pyrrole ring is beneficial for both AChE and BuChE inhibition [98].

Chen and co-authors reported the design, synthesis, and in vitro and in vivo evaluation of tacrine-cinnamic acid hybrids as multi-target AChE and BuChE inhibitors. In this work, the structural modification of ferulic acid **69** moiety was replaced by cinnamic acid with different substitutions (Fig. 46). The target

compounds were synthesized and evaluated for their in vitro and in vivo activities related to the treatment of AD, including in vitro assays for $A\beta_{1-42}$ self-aggregation, ChE catalytic activity, cytoprotective effects against hydrogen peroxide, and antiproliferative activity in PC-12 cells. In vitro assays proved that most of the compounds effectively inhibited ChEs in the nanomolar range ($IC_{50} = 3.8\text{--}173$ nM for AChE and $IC_{50} = 13.6\text{--}170.6$ nM for BuChE) and the compound **174** was one of the most potent analogs (IC_{50} hAChE = 55.1 ± 4.9 nM, IC_{50} hBuChE = 55.9 ± 3.3 nM), which was about fourfold more active than the compound **173** against AChE [99].

Jerábek and co-workers synthesized some tacrine–resveratrol fused hybrids as a new series of anti-AD MTDLs. Their rational design was based on the combination of the structural features of the cholinesterase inhibitor tacrine with that of resveratrol **133**, which is known for its purported antioxidant and anti-neuroinflammatory activities. Compound **175** (Fig. 47) showed intriguing anti-inflammatory and immunomodulatory properties

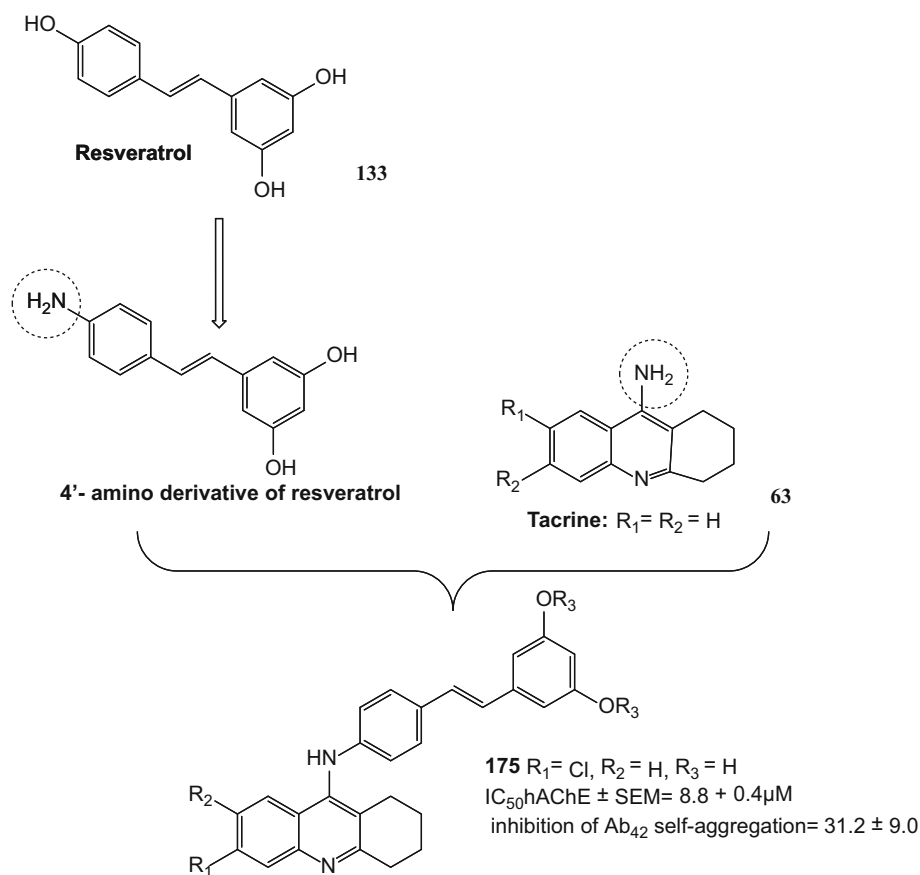


Fig. 47 Design strategy to hybrids and compound **175**

in neuronal and glial AD cell models and IC_{50} hAChE \pm SEM = $8.8 \pm 0.4 \mu\text{M}$, inhibition of $A\beta_{42}$ self-aggregation = 31.2 ± 9.0 . Importantly, the MTDL profile is accompanied by high-predicted BBB permeability, and low cytotoxicity on primary neurons [100].

Tepponou and co-authors designed and synthesized tacrine–trolox **178** and tacrine–tryptoline **179** hybrids with various linker chain lengths. Hybrid derivative **178**, containing the trolox moiety, showed moderate to high *Tc*AChE inhibition (IC_{50} = 49.31 nM), eqBuChE inhibition (IC_{50} = 4.74 nM), and free radical scavenging activities (IC_{50} = 12.67 μM). The hybrid compound containing the tryptoline moiety linked with a seven-carbon spacer to tacrine (**63**) displayed the best AChE (IC_{50} = 17.37) and BuChE inhibitory activity (IC_{50} = 3.16 nM). Novel multi-target agents that exhibit good ChE inhibition (**178** and **179**) and antioxidant (**178**) activity were identified as suitable candidates for further investigation (Fig. 48) [101].

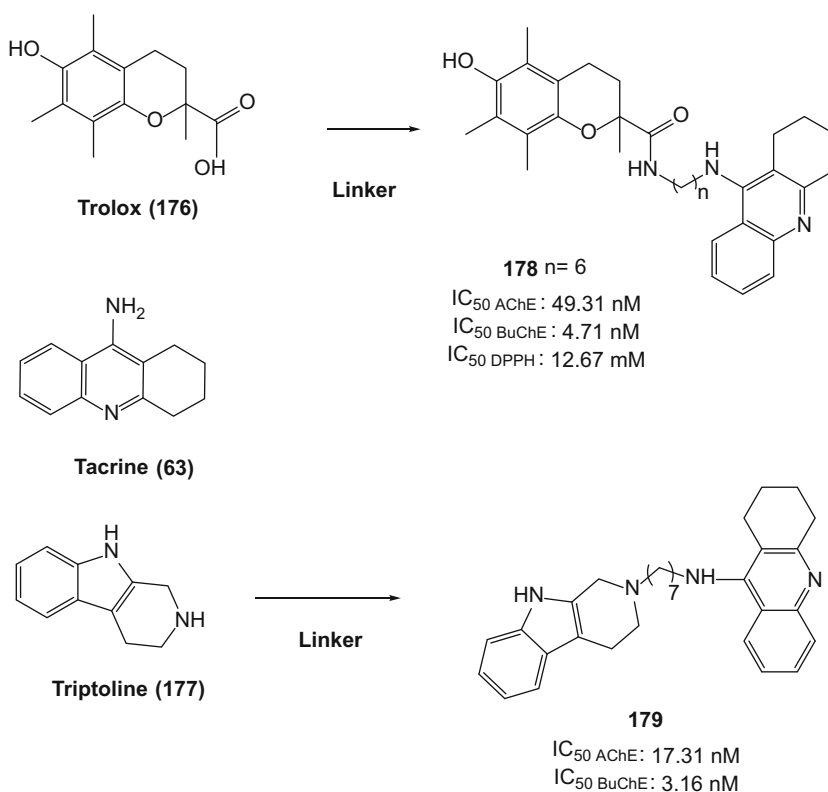


Fig. 48 Design strategy towards tacrine–trolox and tacrine–tryptoline hybrids and the compounds **178** and **179**

8 Multi-target Directed Ligands Inspired by Natural Products (NPs)

Different structural patterns derived from plants or microorganisms present AChE inhibitor activity among others, so many natural products (NP) are used as inspiration for the design of new MTDLs [102]. Whether by structural modifications or molecular hybridization we have numerous examples of MTDLs inspired in NP.

As one example of this strategy, berberine (**180**) was used as a basic skeleton for the construction of a series of hybrid compounds with molecular subunits that include melatonin (**73**) and ferulic acid (**69**, Fig. 49). The aim was to obtain new derivatives **181–184** with antioxidant properties and anti-A β aggregation, but with reduced inhibitory potency of AChE. The hybrid compound hydro-quinone **181** demonstrated the best ability for suppression of A β aggregation, plus excellent antioxidant effect and reasonable ability to inhibit AChE and BuChE. In addition, derivative **182** (Fig. 49) was identified as a potent inhibitor of AChE, with strong antioxidant properties, confirming the multi-target potential of these substances in comparison to berberine, the original prototype [38].

Melatonin (**73**) has been shown to be capable of capturing free radicals, stimulating the synthesis of antioxidant enzymes, reducing

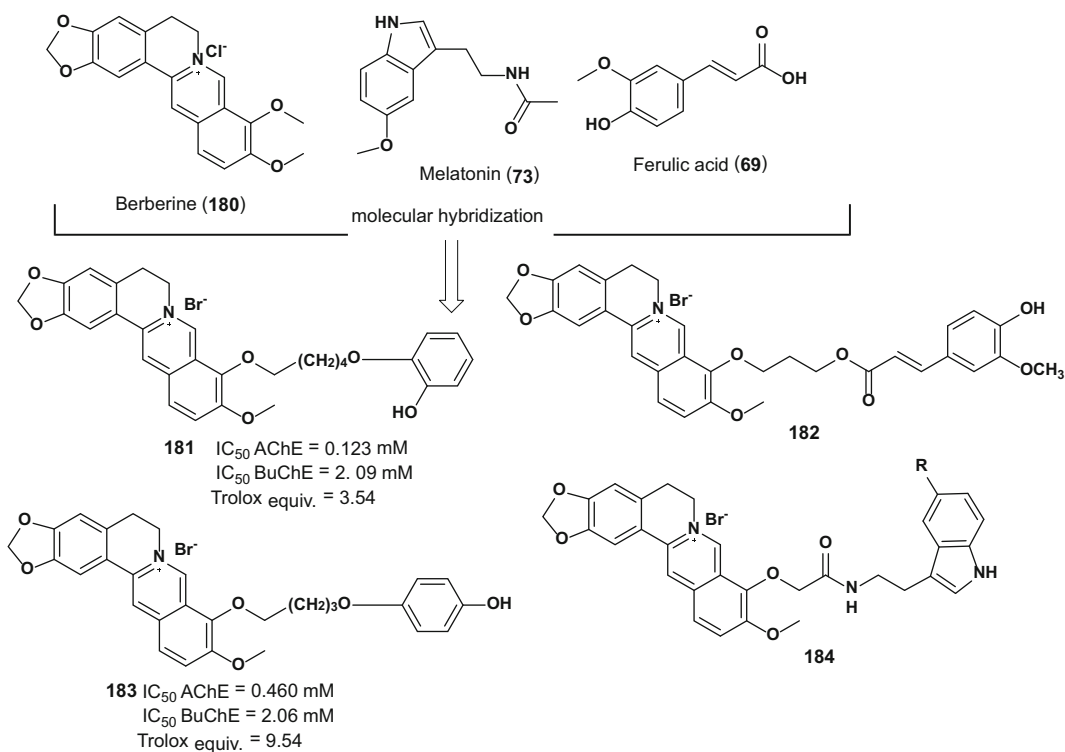


Fig. 49 Series of hybrid derivatives (**181–184**) structurally designed from berberine (**180**)

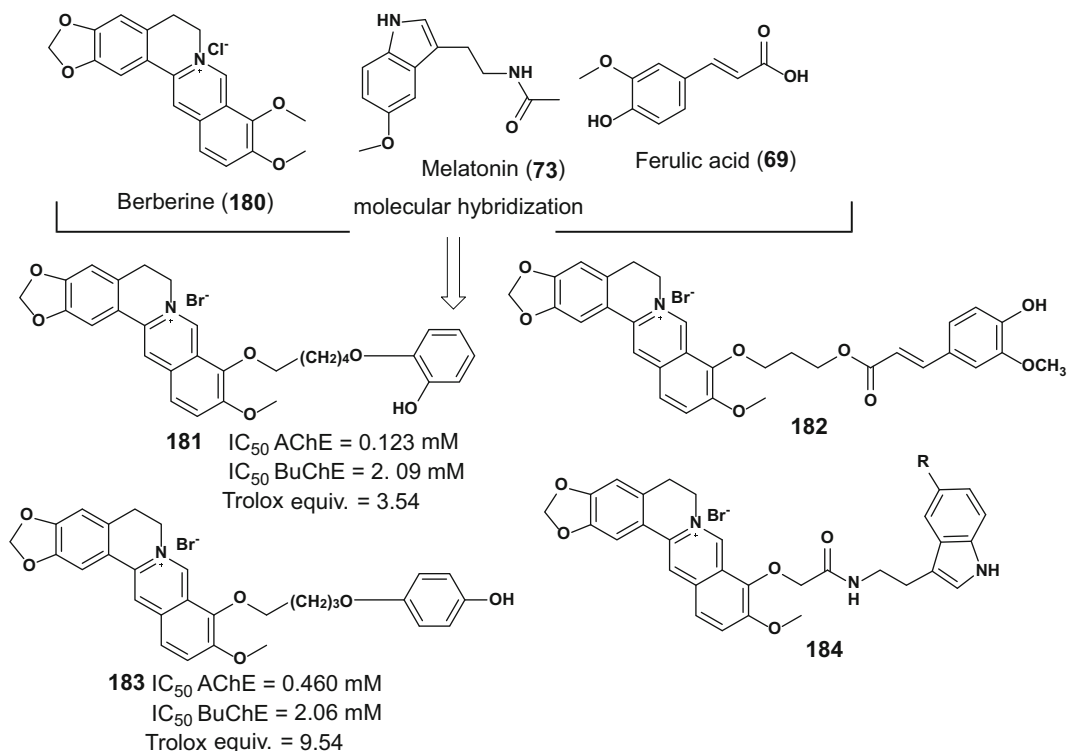


Fig. 50 Chemical structures of compounds **185–186**, designed as MTDLs melatonin–AP2238 hybrids

the neurofilaments hyperphosphorylation and having protective activity against amyloid- β ($A\beta$) protein. Moreover, melatonin was able to induce the proliferation and differentiation of neural cells in the hippocampus of adult mice. *N,N*-dibenzyl (*N*-methyl) is a protonable amine present in the structure of the known AChE inhibitor AP2238 (**28**, $IC_{50} = 0.044 \pm 0.006 \mu\text{M}$), because its interaction with the AChE-CAS has been probed [103–105]. Thus, a series of melatonin–AP2238 (Fig. 50) was planned as novel AChE inhibitors, assembling the endogenous characteristics of both bioactive molecular prototypes. Hence, 14 new substances were synthesized and evaluated, all showing anticholinesterase activities with IC_{50} values in micromolar order, but less potent than AP2238 (**28**). Permeability assay prove that all molecules were able to cross BBB, with additional antioxidant activity close to the reference value of melatonin, but with lower neuroprotective activity than melatonin. Finally, considering that methoxy group is indicated as crucial for neurogenic activity of melatonin, the methoxy-substituted hybrids **185–187** were submitted to neurogenic studies. The results evidenced that these three compounds were able to induce neurogenesis, but compound **186** excelled in ability to induce maturation of these cells. Therefore, one can expect that hybrids **185**, **187**, and specially compound **186** could

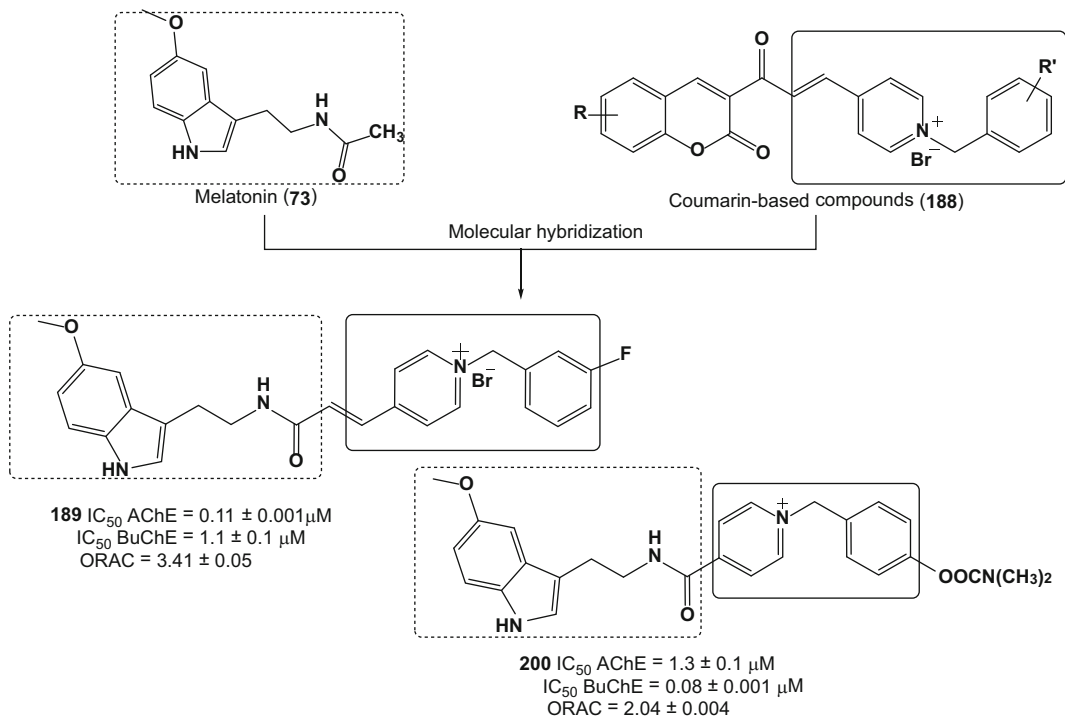


Fig. 51 Design of novel benzylopyridinium–melatonin hybrids **189** and **190** with anticholinesterase and antioxidant properties

represent innovative multifunctional drug candidates capable to repair CNS damage caused by neurodegenerative diseases and protect neuronal cells from oxidative processes [106].

Besides melatonin (**73**) as an antioxidant prototype molecule, benzylopyridinium salt (**188**) has been considered a privileged structure in the development of AChE inhibitors. In this context, benzylopyridinium–melatonin hybrids were planned with the aim of obtaining novel antioxidant-AChE inhibitor molecules. Twenty-three new substances were synthesized (Fig. 51), highlighting compounds **189** and **190** that showed high ChE inhibitory activities with IC_{50} values of $0.11 \pm 0.001 \mu\text{M}$ (AChE), $1.1 \pm 0.1 \mu\text{M}$ (BuChE) and $1.3 \pm 0.1 \mu\text{M}$ (AChE), $0.08 \pm 0.001 \mu\text{M}$ (BuChE), respectively, with pronounced antioxidant activity, especially for compounds **189** (ORAC = 3.41 ± 0.05) and **190** (ORAC = 2.04 ± 0.004). Toxicity was also evaluated in neuroblastoma cell cultures, with both compounds **189** and **190** showing similar toxicity to or less than melatonin **73** [107].

Resveratrol (**133**), a potent phenolic natural plant metabolite, with remarkable antioxidant and anti-inflammatory properties, has been recently described as an inhibitor of A β aggregation. Thus, its molecular scaffold has been also explored for assembly of innovative MTDLs with antioxidant, anti-inflammatory, and inhibitory of A β

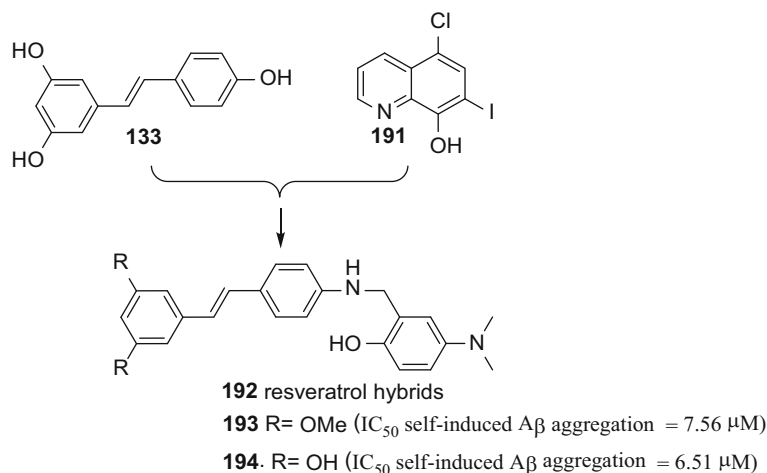


Fig. 52 Chemical structures of the very active MTDL resveratrol-derived compounds **193** and **194**

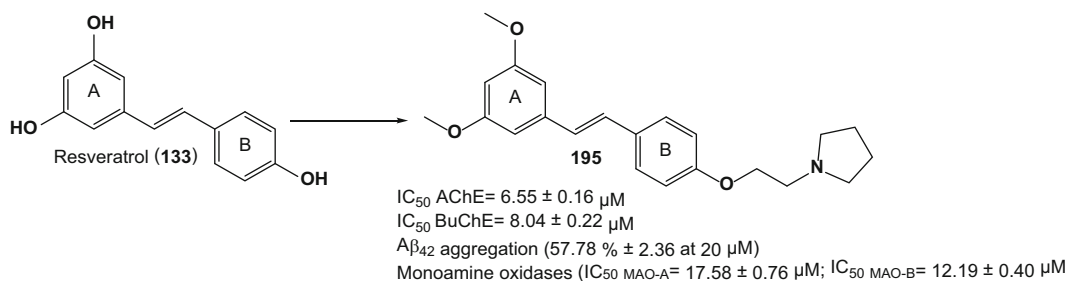


Fig. 53 Chemical structures of 3,5-dimethoxy-4'-*O*-alkylamine-resveratrol derivatives (**195**) with AChE and $A\beta$ inhibitory properties

aggregation properties. Compounds **193** and **194** were the lead compounds identified among a series of resveratrol-derived molecular hybrids (**192**), planned by the combination of the catechol pharmacophore of **83** with the well-known metal chelator clioquinol (**191**) (Fig. 52). These two hybrid compounds showed high potency in the inhibition of self-induced $A\beta$ aggregation (**193**, IC_{50} = 7.56 μ M and **194**, IC_{50} = 6.51 μ M) in comparison to resveratrol (IC_{50} = 15.11 μ M). Furthermore, antioxidant assays performed by the ORAC method showed the ability of compounds **193** and **194** in decreasing the generation of ROS species in 4.72 and 4.70 trolox equivalent, respectively. These compounds were also capable to disassembling the highly structured $A\beta$ fibrils generated by self- and Cu(II)-induced $A\beta$ aggregation and exhibited significant inhibitory effect on MAO-A and MAO-B with a moderate AChE inhibition and low neurotoxicity [108].

In 2014, Pan and co-authors have already reported a series of resveratrol (**133**) derivatives (Fig. 53), with potential therapeutical use for AD treatment. Basically, derivatives **195** were planned by

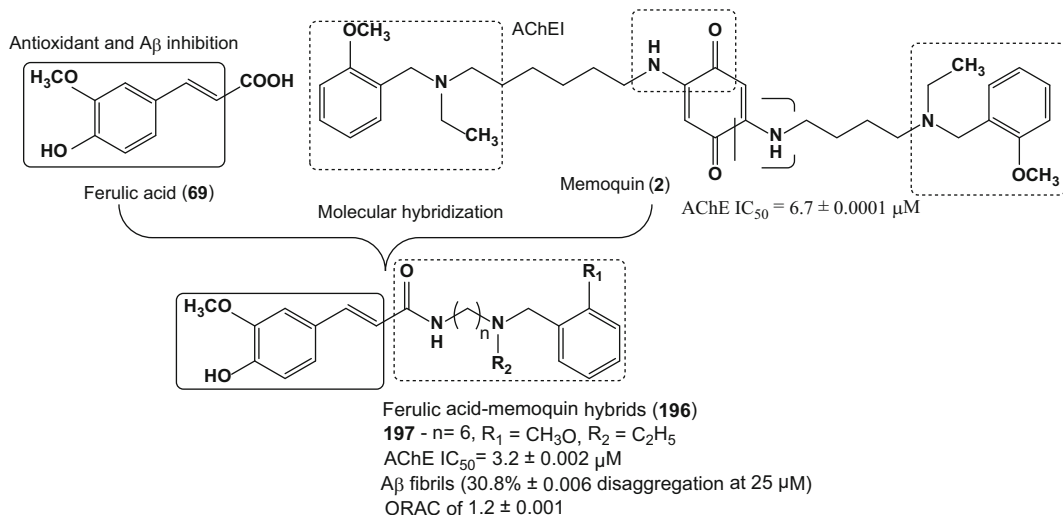


Fig. 54 Design strategy for ferulic acid–memoquin hybrids (**196**) and structure of the most active hybrid **197**

exchanged the 3,5-di-hydroxy group on the ring A of resveratrol by 3,5-di-methoxy and insertion of an alkylamino side chain attached to the oxygen on the ring B. Among them, compound **195** exhibited the best biological results, with significant inhibitory activity of cholinesterases (IC₅₀ AChE = 6.55 \pm 0.16 μ M; IC₅₀ BuChE = 8.04 \pm 0.22 μ M), A β ₄₂ aggregation (57.78% \pm 2.36 at 20 μ M), and MAOs (IC₅₀ MAO-A = 17.58 \pm 0.76 μ M; IC₅₀ MAO-B = 12.19 \pm 0.40 μ M) [109].

Memoquin (**2**) is considered an example of success, being one of the first AD multi-target drug discovery efforts. It interacts with three molecular targets involved in AD pathology: AChE, β -amyloid (A β), and BACE-1. Ferulic acid **69** has antioxidant and anti-inflammatory effects, inhibits A β fibril aggregation, and prevents A β -mediated toxicity both in vitro and in vivo. Based on these data, Pan and co-workers selected ferulic acid (**69**) to combine with different alkyl-benzylamine fragments to design a series of novel ferulic acid–memoquin hybrids (**196**, Fig. 54) that are expected to show potentially applicable multifunctional properties towards AD. All the target compounds exhibited more potent inhibitory activities than ferulic acid (**69**, IC₅₀ > 100 μ M), but lower inhibitory activities than memoquin (**2**, IC₅₀ = 6.7 \pm 0.0001 μ M). Compound **197** showed the strongest inhibitory activity with IC₅₀ = 3.2 \pm 0.002 μ M. In addition, **197** was able to disaggregate A β fibrils (30.8% \pm 0.006 disaggregation at 25 μ M), with significant antioxidant potential (ORAC of 1.2 \pm 0.001) and adequate BBB permeability in vitro. At a concentration of 10.0 μ M, compound **197** exhibited significant neuroprotective effect and cell viability (88.3 \pm 6.5%). It is important to note that compound **197** exhibited interesting multi-target ligand profile, without

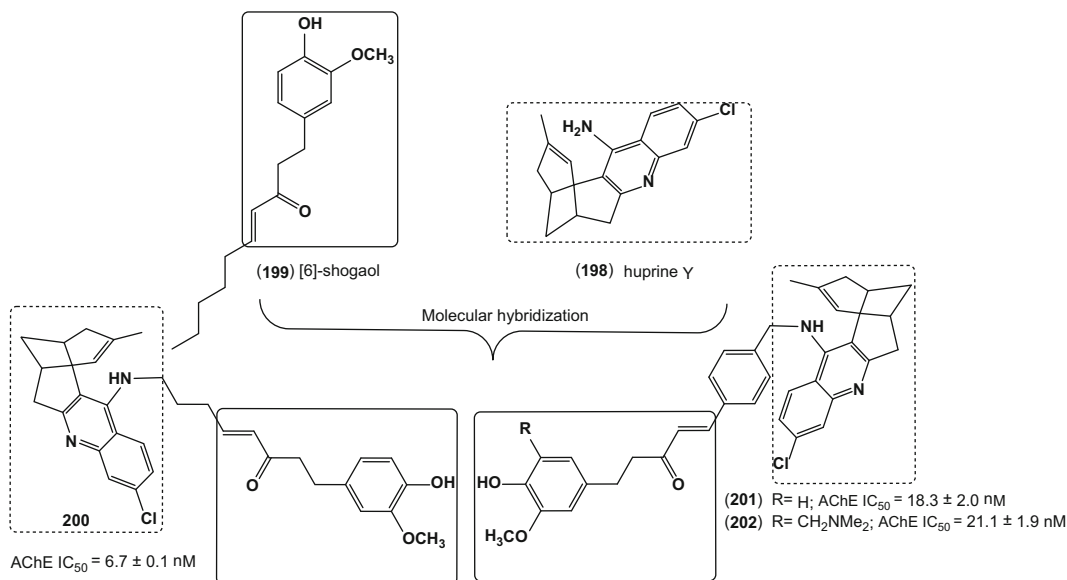


Fig. 55 Structures of the natural antioxidant shogaol, the AChE inhibitor huprine Y, and shogaol–huprine Y hybrids **200–202**

restrictions of solubility, which is a great advantage related to memoquin (**2**) that failed in preclinical trials due to its very low solubility [110].

Pérez-Areales and co-workers described the design, synthesis, and pharmacological evaluation of a short series of multi-target anti-Alzheimer hybrid compounds that combine a fragment of the highly potent AChE inhibitor huprine Y (**198**) with the 4-hydroxy-3-methoxyphenylpentenone moiety of shogaols (**199**) (Fig. 55). A series of hybrids **200–202**, differing in the nature and size of the spacer between the huprine (**198**) and shogaol (**199**) fragments, were docked in the hAChE models. All three compounds turned out to be potent hAChE and hBuChE inhibitors with IC_{50} values of 6.7 ± 0.1 , 18.3 ± 2.0 , and 21.1 ± 1.9 nM, respectively, and potent antioxidant agents, even though this hybridization strategy led to slightly decreased hAChE inhibitory activity relative to the parent huprine Y (**198**). The presence of the additional aromatic ring in the linker of hybrids **201** and **202**, which led to increased antioxidant activity, seemingly enhances their interaction with $A\beta_{42}$ and tau protein, leading to potent $A\beta_{42}$ and tau anti-aggregating activities. The $A\beta_{42}$ and tau anti-aggregating activities were in the ranges 39–71% and 35–51%, respectively, using a $10 \mu\text{M}$ concentration of the hybrids, they being clearly more potent than the parent huprine Y (**198**) and shogaol (**199**) [111].

Viayna and collaborators designed a family of rhein–huprine Y hybrids (**203**, Fig. 56) to hit several key targets for Alzheimer’s disease. All the racemic hybrids (\pm)-**204–210**, obtained from

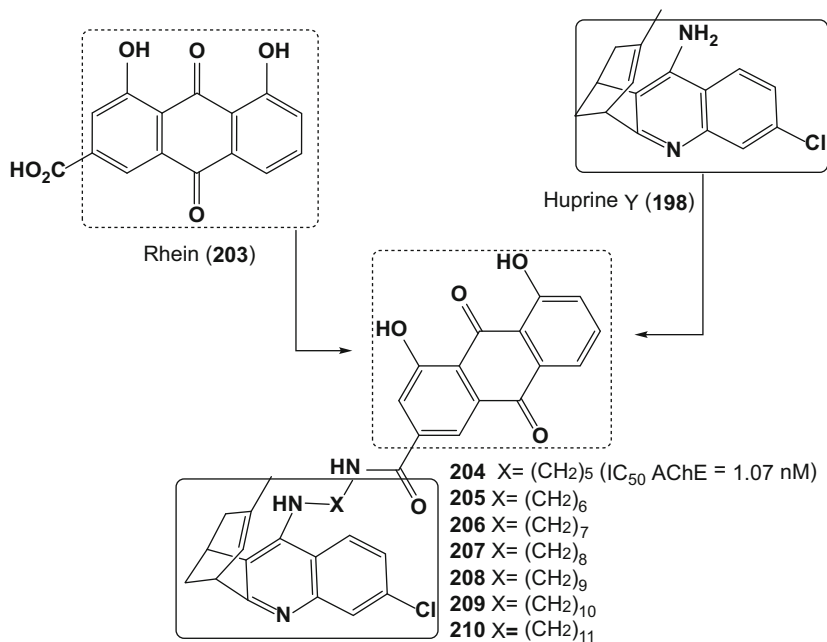


Fig. 56 Design of the new series of derivatives of rhein (**203**) and huprine Y (**198**)

racemic huprine Y (**198**), turned out to be potent inhibitors of hAChE, with IC₅₀ values in the low nanomolar range. Hybrid (±)-**204**, the most potent hAChEI of the series (IC₅₀ = 1.07 nM), was equipotent to the parent racemic huprine Y [(±)-**198**]. Additionally, the beneficial effects of (+)-**204** on synaptic integrity were apparent in the context of LTP induction. Finally, in vivo experiments with transgenic APP-PS1 mice have shown that (+)- and (–)-**204** are able to lower the levels of hippocampal total soluble Aβ and increase the levels of APP in both initial and advanced stages of this AD model, thus suggesting a reduction of APP processing, as expected from their potent BACE-1 inhibitory activity. Overall, the novel rhein–huprine hybrids (+)- and (–)-**204** emerge as very promising multi-target anti Alzheimer drug candidates with the potential to positively modify the underlying mechanisms of this disease [112].

To date, it is well known that MAO-B activity can increase up to threefolds in the AD patients compared with controls. This increase in MAO-B activity produces higher levels of H₂O₂ and oxidative free radicals, which have been correlated with the development of oxidative stress [113]. In addition, highly concentrated metal ions (e.g., Cu²⁺, Zn²⁺, and Fe³⁺) in the neuropil and plaques of the brain are closely associated with the formation of Aβ plaques and neurofibrillary tangle as well as linked to the production of ROS and oxidative stress [114, 115]. According to the studies of Huangi and co-workers, a series of hybrid derivatives (**211**) with the

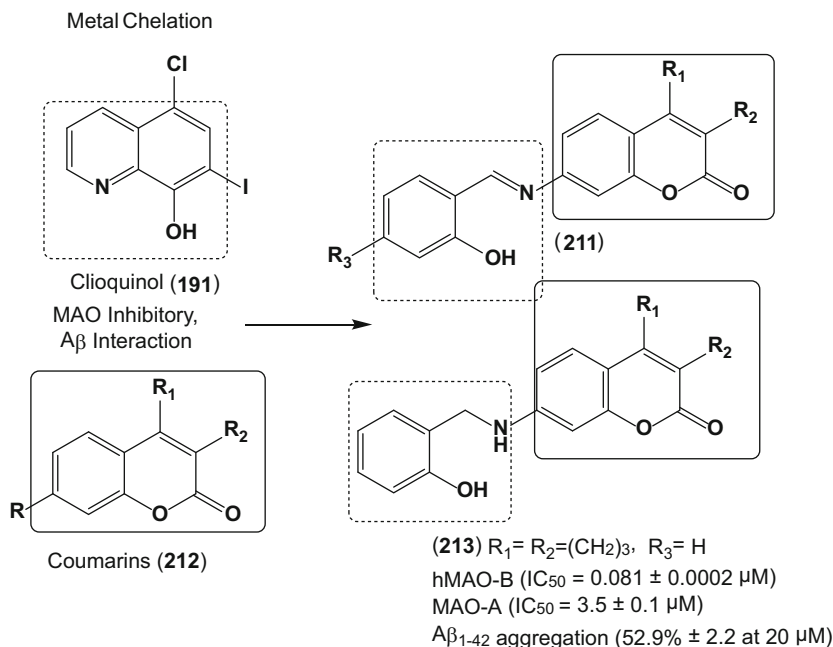


Fig. 57 Structure of coumarin (**212**) and the most active compounds **211** and **213** designed by molecular hybridization of clioquinol (**191**) and coumarins (**212**)

pharmacophore moiety of metal chelator clioquinol (CQ) (**191**) and coumarin (**212**) (Fig. 57) was rationally planned to generate coumarin derivatives with expected biometal chelation ability, along with inhibitory activity of MAO-B and A β anti-aggregation. Compound **213** showed the greatest potential to inhibit hMAO-B (IC₅₀ = 0.081 ± 0.0002 μ M) with >1.234-fold selectivity over MAO-A (IC₅₀ = 3.5 ± 0.1 μ M) as well as good inhibition of A β ₁₋₄₂ aggregation (52.9% ± 2.2 at 20 μ M), low cell toxicity in rat pheochromocytoma (PC12) and SH-SY5Y cells, and BBB permeability [116].

Aurones, 2-benzylidenebenzofuran-3(2*H*)-ones **214**, which are structural isomers of flavones, are present in vegetables and flowers and have attracted considerable attentions because they possess a wide range of bioactivities associated with neurological diseases. Thus, Li and co-workers synthesized and evaluated a series of 4-hydroxy-aurone derivatives **215**–**216** (Fig. 58) designed as potential multifunctional agents for the treatment of AD. They observed that all derivatives showed high antioxidant activities, ranging from 1.00- to 3.56-fold of Trolox, especially compound **215e** that showed an antioxidant activity 1.90-fold higher than trolox, in spite of a good inhibitory activities of self- and Cu²⁺-induced A β ₁₋₄₂ aggregation with 99.2 ± 1.1% and 84.0 ± 1.5% at 25 μ M, respectively. In addition, **215e** also showed remarkable inhibitory activities of both MAO-A and B with IC₅₀ values of

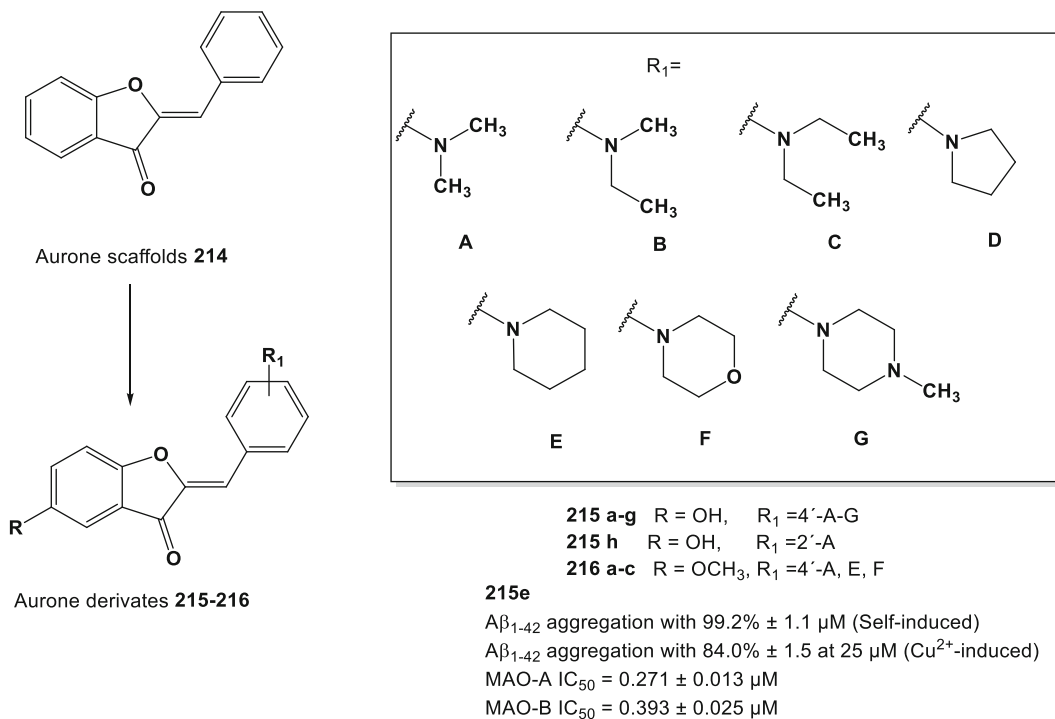


Fig. 58 Structure of multifunctional aurone derivatives **215–216** designed as multifunctional drug candidates for AD

0.271 ± 0.013 μM and 0.393 ± 0.025 μM, respectively. Furthermore, **216b** exhibited high selectivity to MAO-B over MAO-A, which may serve as potential MAO-B selective inhibitors. These lead compounds **215e** and **216b** also showed good metal-chelating properties and BBB permeability [117].

Li and co-workers reported the design, synthesis, and biological evaluation of a new series of pterostilbene–benzylamines hybrid derivatives (**218–222**, Fig. 59). Pterostilbene (**217**) is a dimethoxy derivative of resveratrol (**133**) and it was used as molecular model for planning new MTDL candidate prototypes with inhibitory activity of AChE and BuChE, along with antioxidant and Aβ aggregation inhibitory effects. The design approach was based on the connection of pterostilbene **217** with benzylamines **218a–d** using an amide functionality as spacer subunit with different chain lengths. Pharmacological evaluation revealed compound **221d** as the best selective AChE inhibitor (IC₅₀ = 0.06 ± 0.03 μM) and good inhibition of BuChE (IC₅₀ = 28.04 ± 1.71 μM). Both, inhibition kinetic analysis and molecular modeling study indicated that these compounds showed mixed-type inhibition mode, binding simultaneously to the CAS and PAS of AChE. In addition to cholinesterase inhibitory activities, these compounds showed different levels of antioxidant activity (ORAC 0.51 ± 0.03 of Trolox

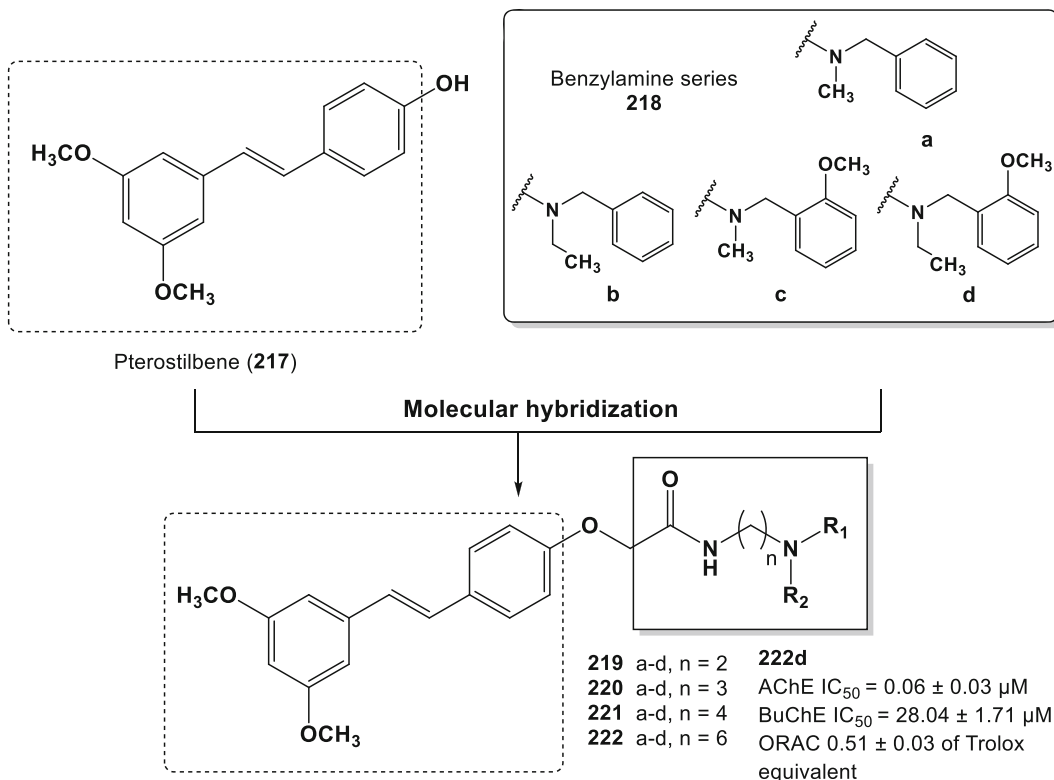


Fig. 59 Design of the pterostilbene–benzylamines hybrid multifunctional derivatives **141–144**

equivalent). SAR studies suggested that the introduction of an amide function on the side chain of the pterostilbene fragment led to significantly increased enzymatic activity as well as increasing the length of the linker may give rise to a better activity [118].

Chromone **223**, a privileged scaffold in Medicinal Chemistry, is the core fragment of several flavonoid derivatives such as flavones and isoflavones. Liu and co-workers selected chromone-2-carboxamide moiety to combine with alkyl-benzylamine fragments of different lengths (**a–h**), inspired in the structure of genistein derivatives **224**, to design a potential multifunctional series of novel chromone-2-carboxamido-alkylbenzylamine series (**225**, Fig. 60). In vitro biological evaluation showed that all the target compounds significantly inhibited AChE activity in a sub-micromolar to micromolar range, with good selectivity. Compound **226** exhibited the best inhibitory potency over rat cortex homogenate AChE (RatAChE) ($IC_{50} = 0.07 \pm 0.002 \mu M$), with 736-fold higher selectivity for AChE over BuChE. Kinetics studies revealed a mixed-type inhibition mode for this compound, being capable of accessing both the CAS and PAS of AChE. In addition, **226** exhibited a moderate anti-oxidative activity, selective biometal-chelating ability, along with excellent self-induced and good Cu^{2+} -induced

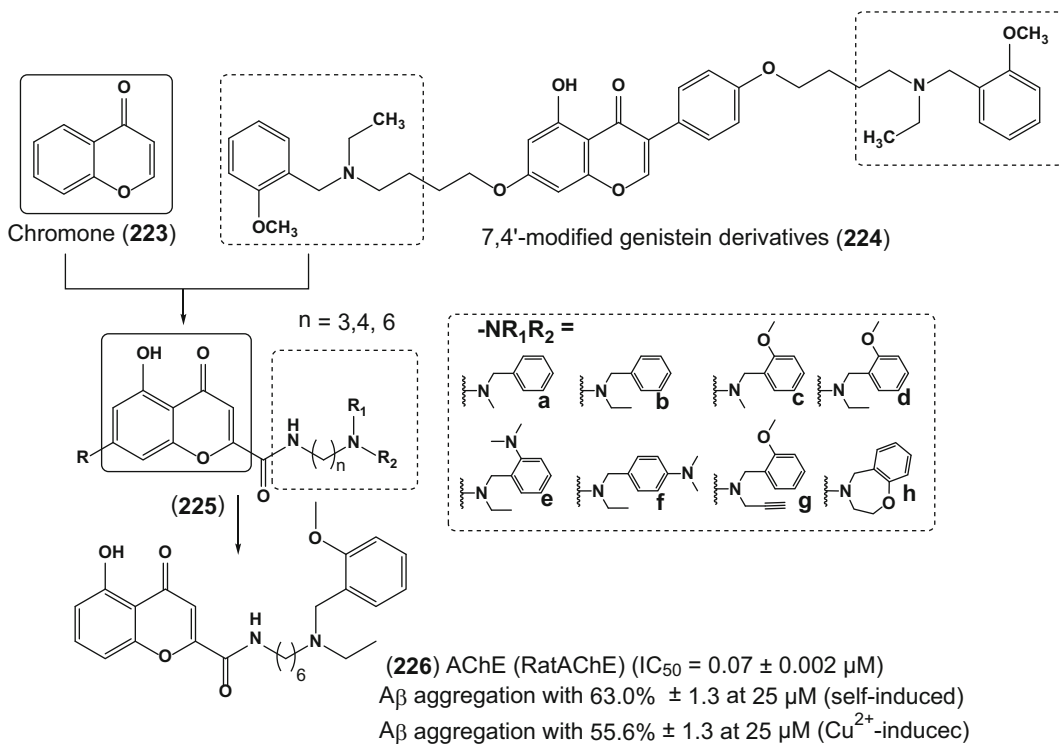


Fig. 60 Structure of chromone (**223**), and 7,4'-*O*-modified genistein derivative **224**, used as molecular prototypes in the design of chromone-2 carboxamide-alkyl-benzylamine derivatives **225** and the most active compound **226**

$A\beta$ aggregation inhibitory activity ($63.0\% \pm 1.3$ and $55.6\% \pm 1.3$, respectively) [119].

Synthetic and natural coumarins (**137**) have recently received special attention due to their wide biological employability. Some of their derivatives have been described as good AChE inhibitors, $A\beta$ -aggregation inhibitors, and neuroprotectors against oxidative damages. Thus, coumarin was used as the basic scaffold in the drawing of a series of innovative and with singular structural pattern coumarin derivatives. Nineteen new coumarin derivatives were synthesized and showed a very selective and potent AChE inhibition profile (Fig. 61). Among all these, compounds **227** ($IC_{50} = 0.016 \pm 0.0021 \mu M$), **228** ($IC_{50} = 0.003 \pm 0.0007 \mu M$), **229** ($IC_{50} = 0.012 \pm 0.0018 \mu M$), and **230** ($IC_{50} = 0.019 \pm 0.001 \mu M$) (Fig. 61) were the most prominent AChE selective inhibitors (IC_{50} BuChE $> 100 \mu M$). These most active substances were submitted to cell viability assay (MTT) and H_2O_2 -induced oxidative damage in human neuroblastoma cells to verify possible neurotoxic effects, among the four, one presented a profile similar to that of galanthamine and other three had better results, especially **228**. Taking all data set, compound **229** exhibited the

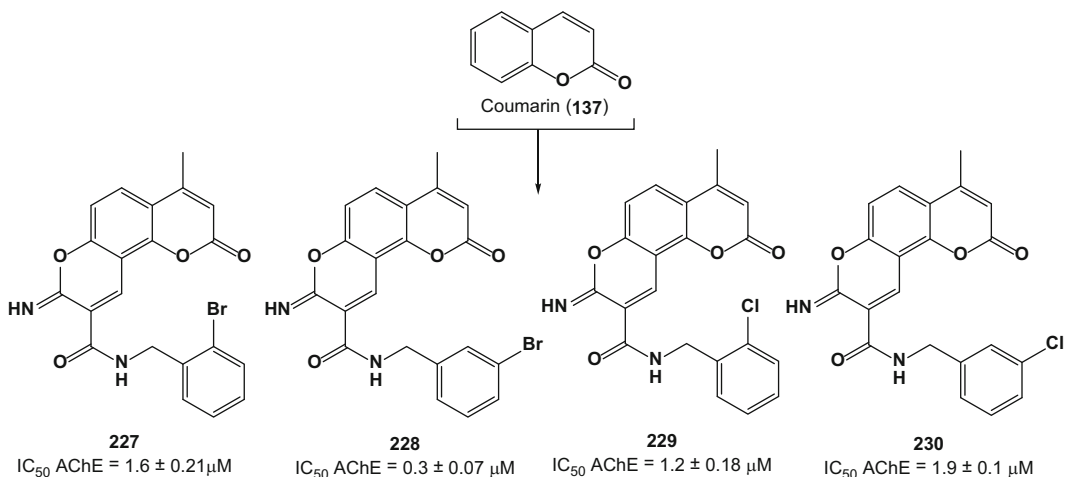


Fig. 61 Chemical structural of more potent derivatives of coumarin (**227–230**)

best MTDL profile, with lower toxicity than galanthamine, and could be considered an interesting prototype candidate in the process of drug development for AD therapeutics [120].

Lan and co-workers designed and synthesized by combining *N*-benzyl pyridinium moiety and coumarin **231** into in a single molecule, novel hybrids with ChE and MAO-B inhibitory activities. The biological screening results indicated that most of compounds displayed potent inhibitory activity for ChE and $A\beta_{1-42}$ self-aggregation, and clearly selective inhibition to MAO-B over MAO-A. The compound **232** (Fig. 62) was the most potent inhibitor for hMAO-B, and it was also a good and balanced inhibitor to ChEs and hMAO-B (0.0373 μM for EeAChE; 2.32 μM for eqBuChE; and 1.57 μM for hMAO-B). Furthermore, the active compound **232** with no toxicity on PC12 neuroblastoma cells showed good ability to inhibit $A\beta_{1-42}$ self-aggregation and cross the BBB. The results suggested that this compound is a promising multi-target candidate [121].

A series of prenylated resveratrol **133** derivatives were designed, synthesized, and biologically evaluated for inhibition of BACE1 and $A\beta$ aggregation as well as free radical scavenging and neuroprotective and neurotogenic activities, as potential novel multifunctional agents against AD by Puksasook and co-workers. Compound **233** (Fig. 63) exhibited good antioxidant activity ($IC_{50} = 41.22 \mu\text{M}$), anti- $A\beta$ aggregation ($IC_{50} = 4.78 \mu\text{M}$), and moderate anti-BACE1 inhibitory activity (23.70% at 50 μM). Moreover, this compound showed no neurotoxicity along with a greater ability to inhibit oxidative stress on P19-derived neuronal cells (50.59% cell viability at 1 nM). All results suggest that compound **233** had the greatest multifunctional activities and might be a very promising lead compound for the further development of drugs for AD [122].

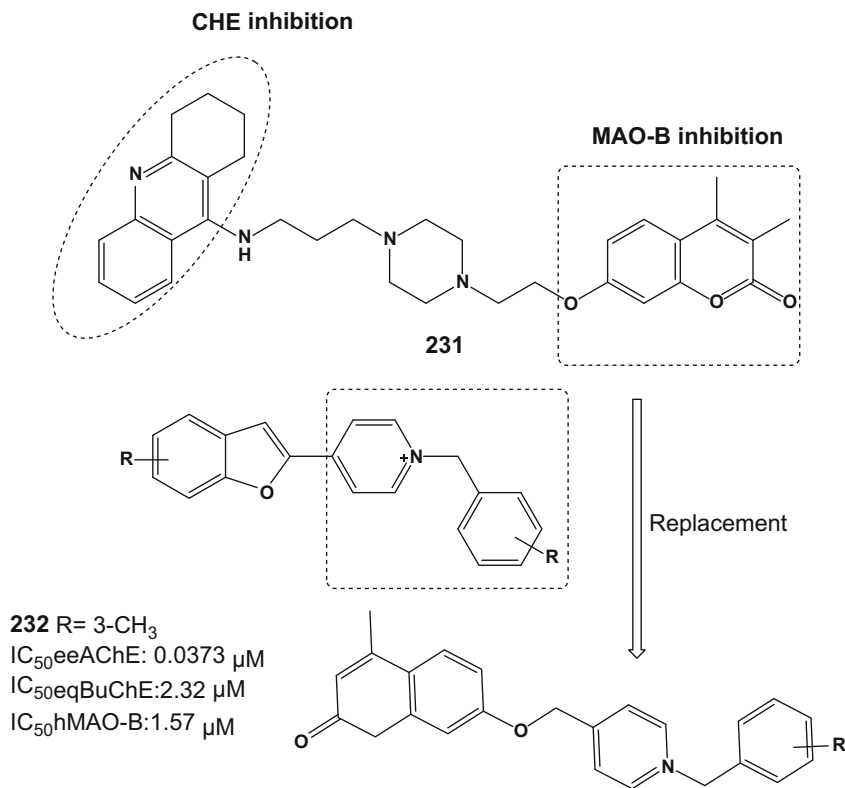


Fig. 62 Drug design strategy for MTDLs and compound **232**

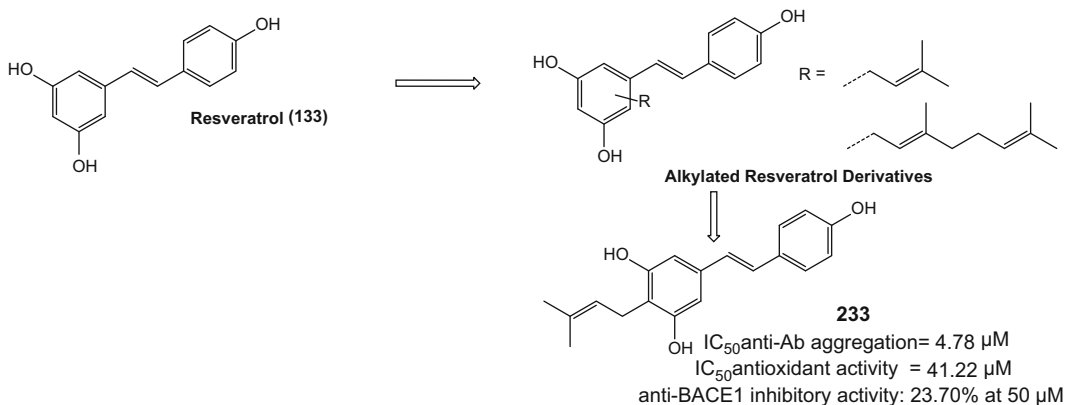


Fig. 63 Designed strategy for alkylated resveratrol derivatives and compound **233**

A series of salicyladimine derivatives were designed by Yang and co-authors as multifunctional anti-AD agents by fusing the structures of resveratrol **133**, benzyloxy **234**, and cloquinol **131** (Fig. 64). Biological assays demonstrated that some derivatives possessed significant inhibitory activities against Aβ aggregation

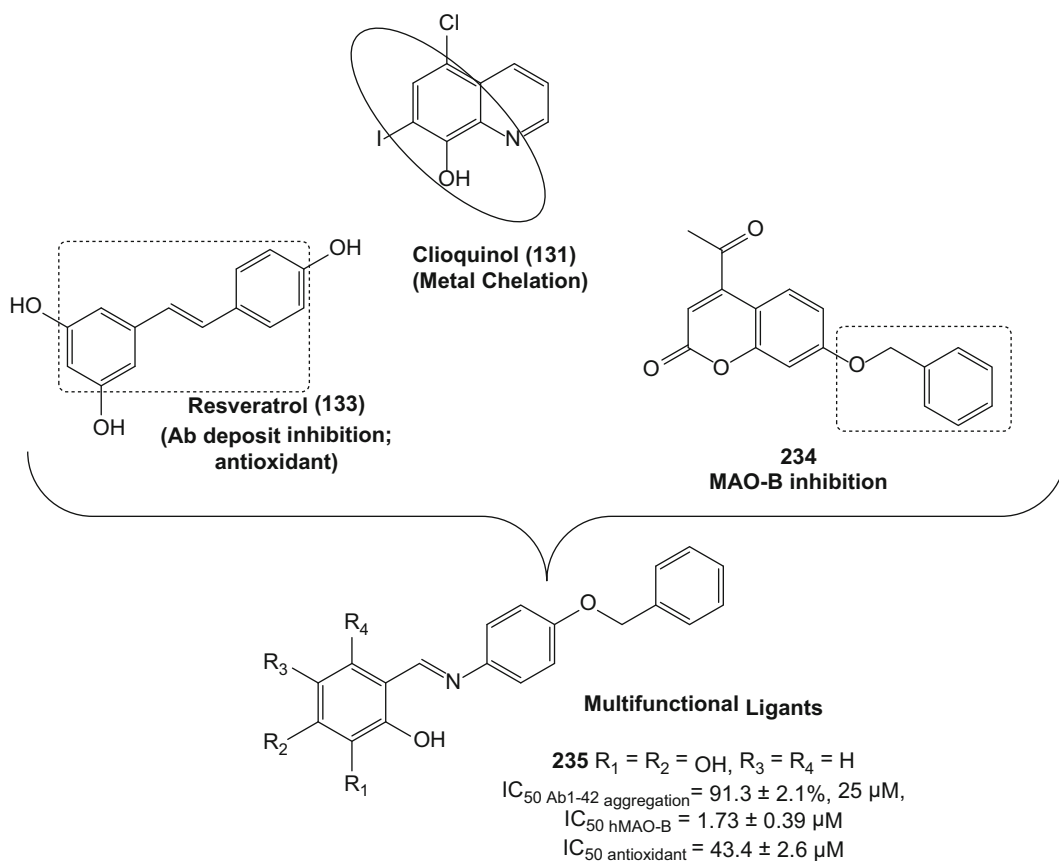


Fig. 64 Design strategy for the multifunctional salicyladimine derivatives and compound **235**

and hMAO-B as well as remarkable antioxidant effects and low cell toxicity. Compound **235** exhibited excellent potency for inhibition of self-induced Aβ₁₋₄₂ aggregation (91.3 ± 2.1%, 25 μM), antioxidant effects on the DPPH (IC₅₀ = 43.4 ± 2.6 μM) and ABTS methods (0.67 ± 0.06 trolox equivalent), inhibition of hMAO-B (IC₅₀ = 1.73 ± 0.39 μM), metal chelation, and BBB penetration. In addition, this compound showed neuroprotective effects against ROS generation, 6-OHDA-induced cell injury, H₂O₂-induced apoptosis, and a significant in vitro anti-inflammatory activity, suggesting its promising multifunctional profile [123].

9 Multi-target Directed Ligands Inspired by Other Polycyclic Structures

The aspartyl protease BACE-1 is the enzyme responsible for Aβ formation and it has become an attractive drug target for AD. In addition, an enzyme known as GSK-3β is involved in the tau hyperphosphorylation process, promoting tau detachment from the microtubules and the consequent formation of neurofibrillary

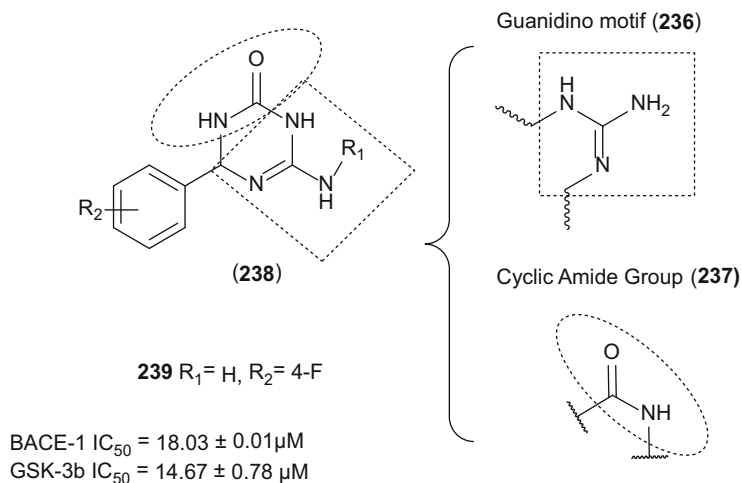


Fig. 65 Design strategy of dual BACE-1 and glycogen synthase kinase-3 β (GSK-3 β) inhibitors compounds **238** and structure of the most potent derivate **239**

tangles. In 2015, Prati and co-workers discovered a new series of 6-amino-4-phenyl-3,4-dihydro-1,3,5-triazin-2(1*H*)-ones (**238**) by hybridization of pharmacophoric features responsible for the BACE-1 and GSK-3 β activity, such as the guanidine **236** and the cyclic amide group **237** (Fig. 65), respectively. Compound **239** was the most potent derivative, inhibiting both enzymes BACE-1 and GSK-3 β with IC_{50} values of $18.03 \pm 0.01 \mu M$ and $14.67 \pm 0.78 \mu M$, respectively. Moreover, compound **239** has shown a good profile in neuroprotection, neurogenesis, good CNS permeability, and no neurotoxicity [124].

The combination of the quinoline (**238**) and triazolopyrimidine (**239**) groups has led to hybrid compounds with diverse biological effects, including neuroprotection, AChE inhibitory activity, free radical scavenging effect, and metal complexation ability. The quinoline structure was expected to interact with AChE via π - π interaction. Beside this, it was supposed to intercalate between A β sheets and was expected to enhance its disaggregation due to its planar structure. On the other hand, the triazolopyrimidine scaffold was anticipated to interact with important amino acid residues of AChE. Indeed, the triazole scaffold has an excellent previous record in the inhibition of AChE and BuChE. In this context, a number of quinoline-triazolopyrimidine hybrids (**240**) were designed as multi-target candidates for the treatment of AD, with both pharmacophoric subunits connected by a piperazine moiety (Fig. 66). Eleven new hybrids were synthesized and tested for inhibition of AChE, BuChE, self-induced A β -aggregation, AChE-induced A β -aggregation, and antioxidant properties. Among all, compound **241** stands out for its potent inhibitory activity of cholinesterases (IC_{50} AChE = $0.042 \pm 0.79 \mu M$; IC_{50}

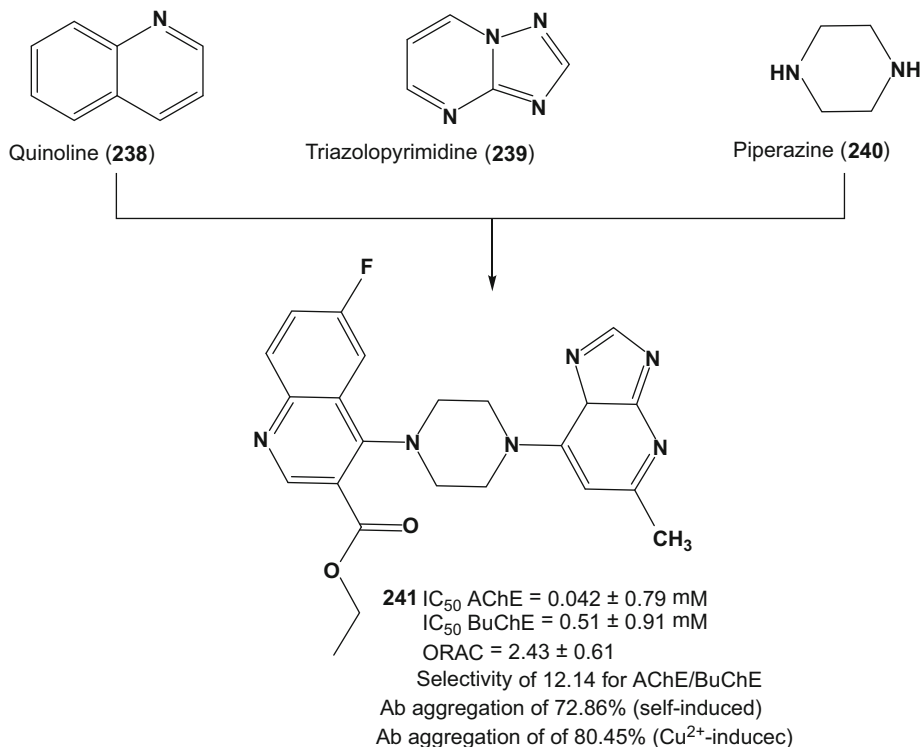


Fig. 66 Quinoline–triazolopyrimidine–piperazine hybrids and the more potent multifunctional derivative **241**

BuChE = 0.51 ± 0.91 μ M), with an AChE inhibition similar to donepezil (IC_{50} AChE = 0.038 ± 0.003 μ M and IC_{50} BuChE = 2.58 ± 0.65 μ M) and a good selectivity of 12-fold higher for AChE. In relation to inhibition of A β aggregation, once again compound **241** stood out by showing a reduction of 72.9% (at 25 μ M) in the amount of self-induced A β -aggregation and 80.45% (at 100 μ M) for AChE-induced A β -aggregation. Antioxidant activity was accessed from MTT cell viability assay in human neuroblastoma cells (SH-SY5Y), in the concentrations of 1, 5, 10, 20, and 25 μ M. The lead compound **241** showed a dose-dependent activity 2.5-fold higher than Trolox, highlighting its multi-target profile as a promise drug candidate prototype for AD treatment [125].

In another approach to pull off novel tacrine derivatives with MTDL profile, Liao and collaborators studied new family of tacrine–flavonoid hybrids designed to act as selective AChE inhibitors, antioxidants, and inhibitors of self-induced β -amyloid peptide (A β) aggregation. The structural architecture of this new series was based on the molecular hybridization of the 5,6,7-trimethoxyflavone (**242**) with 6-chlorotacrine (**243**, Fig. 67). From the target hybrids **244a–d** and **245a–d**, compound **243**, with five methylene groups into the spacer subunit between the

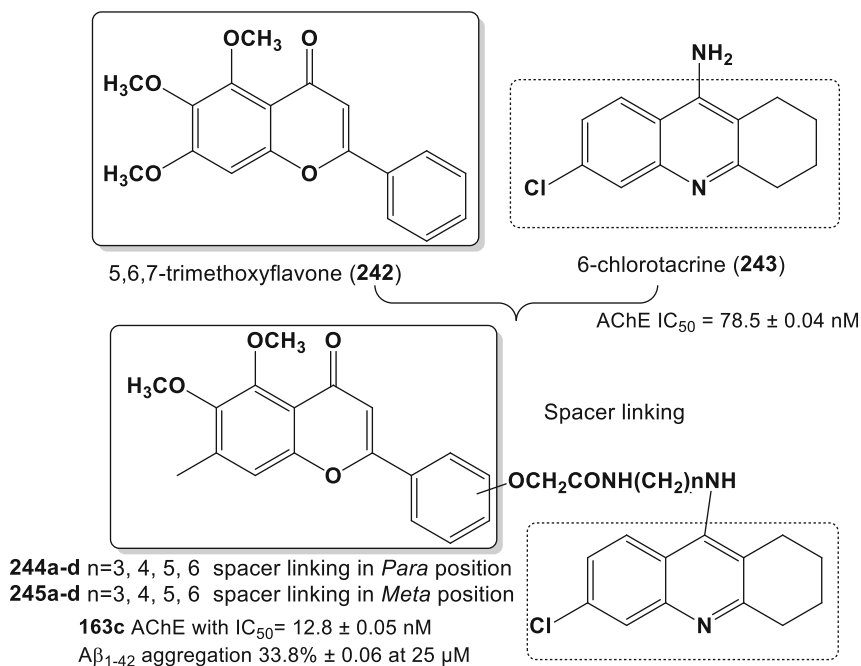


Fig. 67 Design strategy for the 5,6,7-trimethoxyflavone–6-chlorotacrine hybrids **244a–d** and **245a–d**

5,6,7-trimethoxyflavone and 6-chlorotacrine moieties, showed the strongest selective inhibitory potency for AChE with $IC_{50} = 12.8 \pm 0.05$ nM, being sixfold more potent than 6-chlorotacrine ($IC_{50} = 78.5 \pm 0.04$ nM). Furthermore, this compound also showed significant antioxidant activity, a remarkable inhibition of self-induced $A\beta_{1-42}$ aggregation ($33.8\% \pm 0.06$ at $25 \mu M$), good neuroprotective effect against H_2O_2 -induced PC12 cell injury, and BBB permeability [126].

In another recent approach, Sang and co-workers combined the structure of scutellarin (**246**) and rivastigmine (**247**) for the design of a new molecular architecture of MTDL candidates (Fig. 68). Scutellarin is a natural compound with a number of pharmacological properties related to neurological disorders, such as free radical scavenging and metal-chelating activities, anti-inflammatory effects, neuroprotective action, and inhibition of $A\beta$ fibrils. Thus, a series of scutellarin–rivastigmine (**248**) based carbamates was synthesized and evaluated for multifunctional biological properties, including AChE and BuChE inhibition, metal-chelating properties, anti-oxidative, and neuroprotective effects against H_2O_2 -induced PC12 cell injury and BBB permeability. Compounds **249c** and **250c**, containing the *N,N*-diethyl carbamate moiety, showed the most potent and selective inhibition of AChE over BuChE, with **249c** exhibiting an IC_{50} value of $0.34 \mu M$ for AChE and a 24.1-fold higher selectivity, followed by compound **250c** with a 39.7-fold higher selectivity for AChE ($IC_{50} = 0.57 \mu M$).

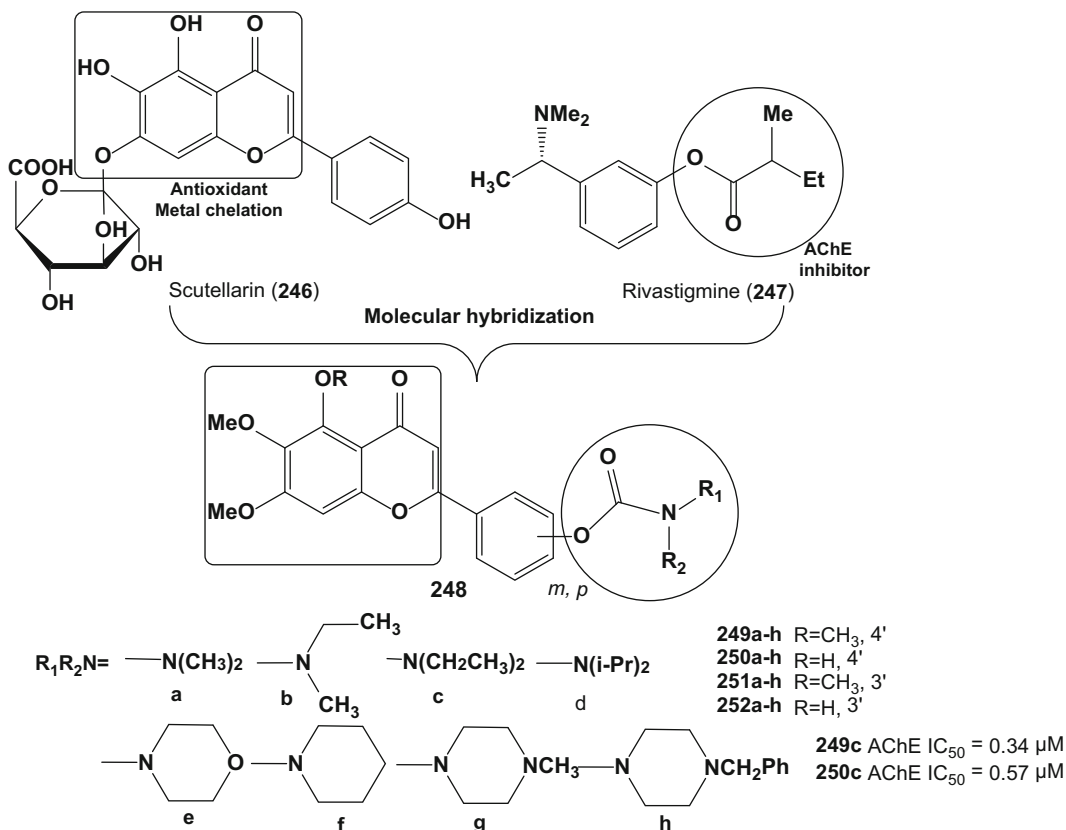


Fig. 68 Designing strategy for the new series of scutellarin–rivastigmine derivatives **249–252**

Kinetic studies demonstrated that compound **250c** could interact concomitantly with the CAS and PAS of AChE, which is consistent with molecular modeling results. Compound **250c** also exhibited a good antioxidative activity with a value 1.3-fold of Trolox, and ability for selective biometal chelation. Furthermore, compound **250c** showed neuroprotective effect against H₂O₂-induced PC12 cell injury and adequate in vitro BBB permeability [127].

BuChE has several neural and nonneural functions. Recent observations suggest that rather than selective inhibition of AChE, BuChE inhibition may be more effective for the treatment of neurodegenerative diseases such as AD [128]. Studies indicate that high levels of BuChE in the cortex are related to some DA markers such as the extracellular deposition of the A β and the aggregation of hyper-phosphorylated tau protein [129]. According to these statements, the use of nonselective ChEs inhibitors would lead to more expressed improvements than the use of selective AChE inhibitors [130]. Previously, some researchers reported that hybrids combining hydroxycinnamic acids (**253**) and ChE inhibitory pharmacophores like in rivastigmine (**247**) could have potential multi-target profile for the therapy of AD [110, 131,

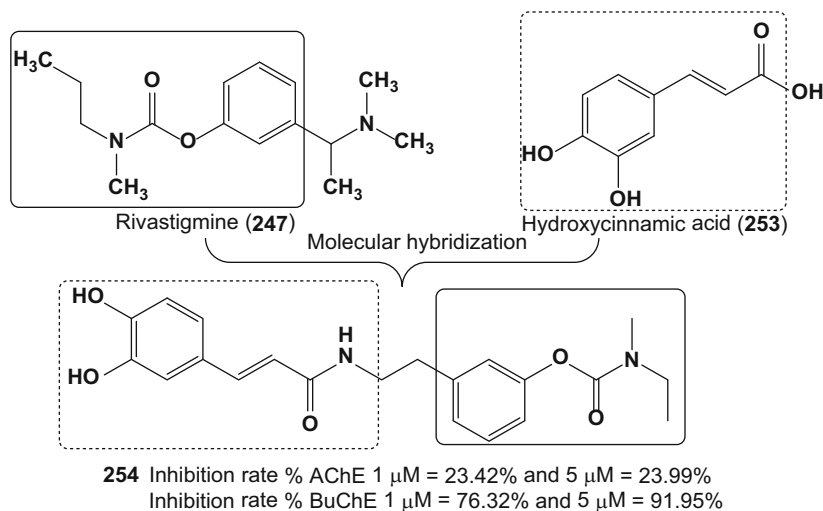


Fig. 69 Design and structure of the most potent rivastigmine–hydroxycinnamic acid hybrid derivative **254**

[132]. On this basis, novel rivastigmine–hydroxycinnamic acid hybrid compounds were planned as inhibitors of AChE, BuChE, A β aggregation, and ROS scavengers (Fig. 69). Among eight hybrids, compound **254** stood out, exhibiting moderate AChE inhibition rate (23.42% at 1 μM and 23.99% at 5 μM) and a higher and selective BuChE inhibitory activity (76.32% at 1 μM and 91.95% at 5 μM), with additional good ability to inhibit A β self-aggregation (85.3% of inhibition at 10 μM). All compounds showed good safety in the MTT cell viability assay with hippocampus neuronal cells (HT22 cells). Finally, a neuroprotection test was performed with the lead compound **254** that showed neuroprotective effects on HT22 cells against cell injury induced by glutamate and against H₂O₂-induced cell death. In spite of the moderate AChE inhibitory activity, compound **254** was able to inhibit the hydrolytic activity of ChEs, to potentially prevent A β self-aggregation, to protect HT22 cells from glutamate, and H₂O₂-induced cell death. These data support that compound **254** could be an attractive lead compound for further optimization in the drug discovery process [133].

Wang and co-authors synthesized a novel series of hybrid compounds **256a–c**, designed by inspiration in the structural scaffold of the metal chelator clioquinol (**191**) and the glutathione peroxidase (GPx) mimic ebselen (**255**) (Fig. 70). The target hybrid compounds were evaluated as potential inhibitors of A β _{1–42} aggregation, antioxidant activity, and superfast antioxidant catalysts against H₂O₂ and metal-chelating ligands. All hybrids exhibited more significant inhibition of A β _{1–42} aggregation than ebselen, clioquinol, and an association of ebselen/clioquinol, with the best results for compounds **256a** (IC₅₀ = 9.6 \pm 0.4 μM) and **256b** (IC₅₀ = 8.1 \pm 0.3 μM). Most of the target compounds

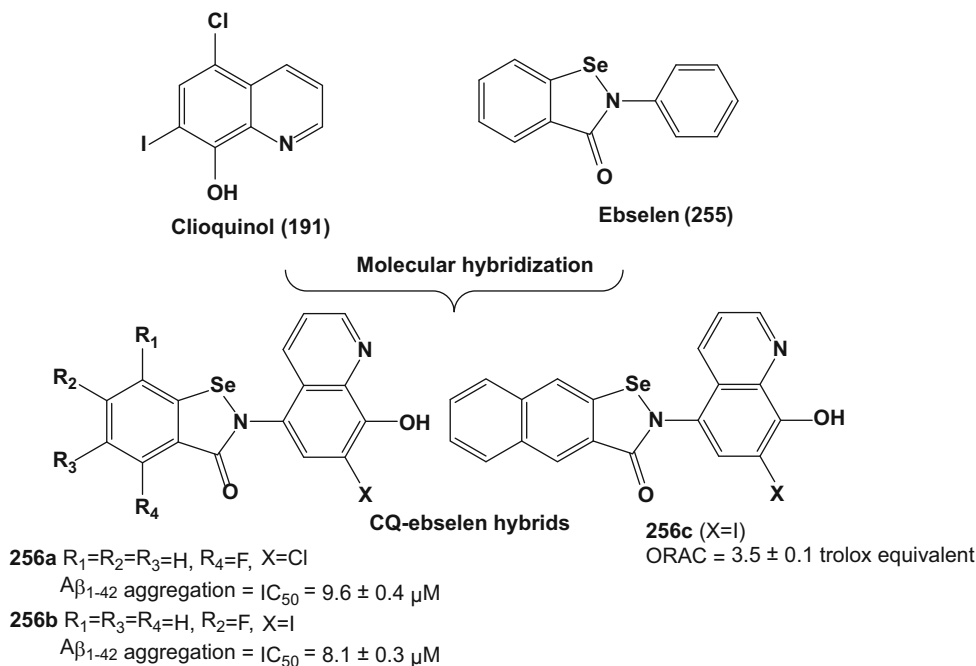


Fig. 70 Chemical structure of series compounds **256a–c** with multi-target profile designed from the prototypes ebselen (**255**) and clioquinol (**191**)

demonstrated strong antioxidant activity ranging from 0.6 to 3.5 trolox equivalents, surpassing ebselen, clioquinol, and ebselen + clioquinol. Compound **256c** exhibited the best antioxidant activity (3.5 ± 0.1 trolox equivalent) and compound **256a** the best brain permeability ($>4.7 \times 10^{-6} \text{ cm s}^{-1}$). These properties highlight the potential of these compounds, especially for **256a**, as promising candidates for the development of innovative multifunctional drug candidates for AD therapeutics [134].

The same research group exploited the structure of coumarins (**212**) in the design of a series of 3-imine-4-hydroxycoumarin derivatives (**257**) (Fig. 71). Biological evaluation for their multifunctional activities in MAO, A β ₁₋₄₂, antioxidant, and biometal chelators disclosed compounds **258** (IC₅₀ MAO-A = 0.673 ± 0.011 μ M, IC₅₀ MAO-B = 0.711 ± 0.013 μ M), **259** (IC₅₀ MAO-A = 4.97 ± 0.41 μ M, IC₅₀ MAO-B = 0.851 ± 0.047 μ M), and **260** (IC₅₀ MAO-A = 3.78 ± 0.62 μ M, IC₅₀ MAO-B = 1.32 ± 0.18 μ M) as the most potent MAO inhibitors. All compounds exhibited moderate to good potencies (20.2–82.3% at 20 μ M) in the inhibition of self-induced A β ₁₋₄₂ aggregation, compared to resveratrol (67.3 ± 3.4 μ M at 20 μ M) and curcumin (50.2% ± 5.9 at 20 μ M). Compound **258** was capable to chelate Cu²⁺ and all compounds demonstrated moderate to good antioxidant activities ranging from 0.12- to 1.57-fold of the trolox values. The best activity was observed for **260** with 1.57 trolox equivalent. In the DPPH radical scavenging method,

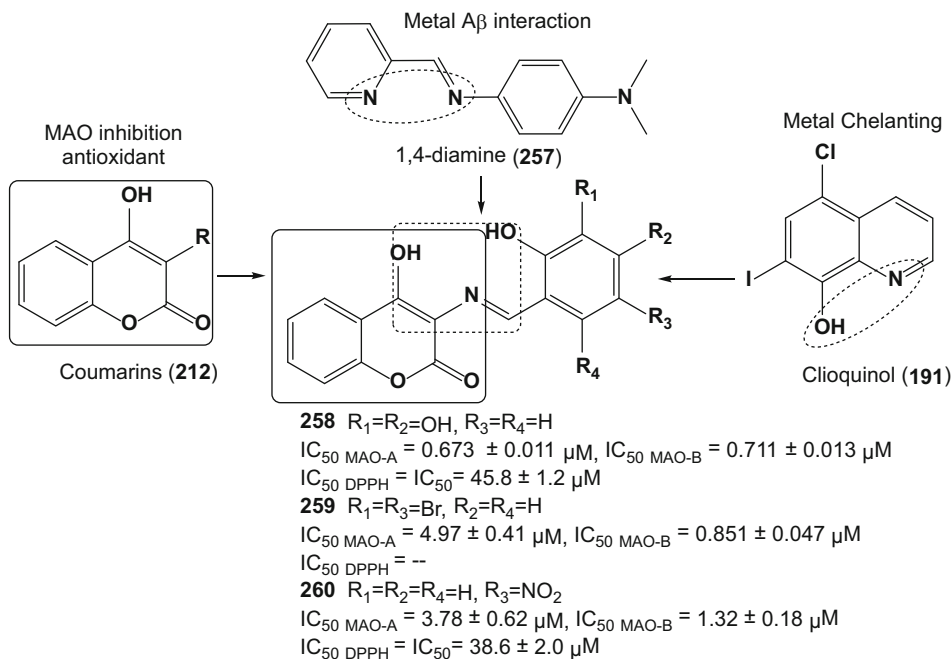


Fig. 71 Design of the new series of derivatives of coumarin–clioquinol hybrids (**258–260**) with improved activities for MAO inhibition, antioxidant, anti-aggregation of amyloid- β , and ion metal chelation

the compounds **258** ($IC_{50} = 45.8 \pm 1.2 \mu M$) and **260** ($IC_{50} = 38.6 \pm 2.0 \mu M$) showed the most potent radical scavenging activities. These results suggested that compounds **258** and **260** might be promising lead compounds with balanced properties for AD treatment [135].

In another work, Wang and co-authors also described the use of clioquinol (**191**) for molecular hybridization with moracin (**261**), leading to the novel series of hybrid compounds (**262–265**, Fig. 72). Among the four series of compounds (**262–265**), **262a** (PDE4D2 $IC_{50} = 2.31 \pm 0.32 \mu M$), **263a** (PDE4D2 $IC_{50} = 0.96 \pm 0.11 \mu M$), and **264a** (PDE4D2 $IC_{50} = 0.32 \pm 0.02 \mu M$), which bear a phenolic hydroxyl group on the R_3 position, demonstrated better activities than the other analogs in the series in the inhibition of PDE4D2. The compound **265a** (PDE4D2 $IC_{50} = 8.86 \pm 0.39 \mu M$) exhibited moderate activity. The inhibitory activities of the target compounds on $A\beta_{1-42}$ self-induced aggregation were first determined by a thioflavin T (ThT) fluorescence assay using curcumin and resveratrol as reference compounds. None of the tested compounds exhibited significant fluorescence signals under the experimental conditions. Moreover, most of the target compounds demonstrated a higher antioxidant ability compared to moracin M (**261**) and clioquinol (**191**). Compounds possessing a phenolic hydroxyl group on the benzofuran moiety, such as **264a**, exhibited significantly higher

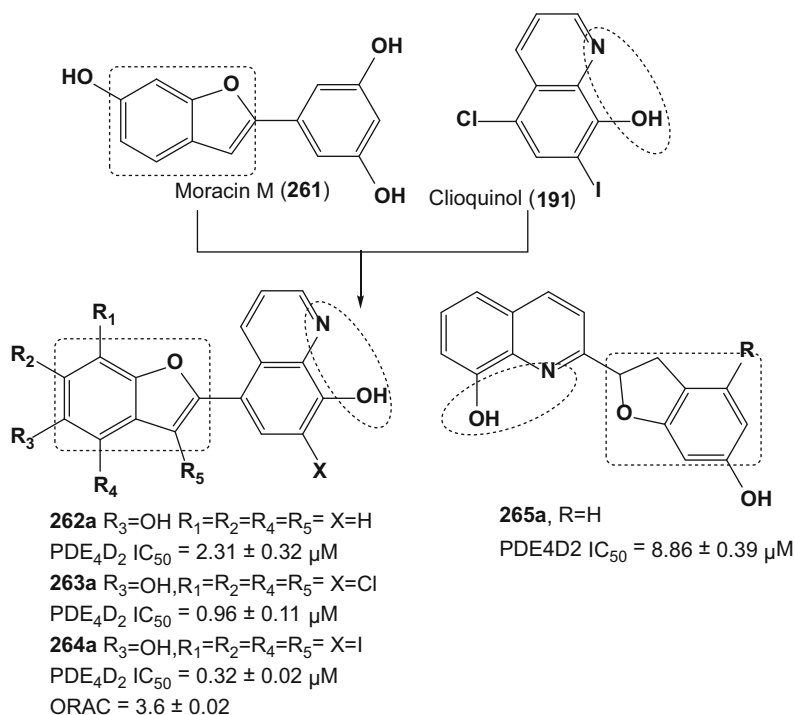


Fig. 72 Design of the new series of derivatives of moracin M (**261**) and clioquinol (**191**)

ORAC values (3.6 ± 0.02). The parallel artificial membrane permeability assay (PAMPA) indicated that most of the target compounds exhibited significant BBB permeability. Among them, compound **264a** ($19.1 \times 10^{-6} \text{ cm s}^{-1}$) exhibited the best BBB permeability with additional biometal-chelating property functions and an interesting neuroprotective effect against inflammation in microglial cells [136].

1,3,5-triazine scaffold (**266**) has been serving medicinal chemists for a long time for the development of antifungal, anticancer, and antiviral agents. Regarding neurodegenerative diseases, due to its roughly planar structure, triazine is expected to intercalate beta-amyloid sheets and to enhance the $A\beta$ disaggregation. For this reason, triazine was selected by Maqbool and co-workers as preferred scaffold in the draw of a series of new cyanopyridine-triazine (**267**) hybrids rationally designed as persuasive multifunctional agents for the treatment of AD (Fig. 73). All eight derivatives obtained were potent and selective AChE inhibitors, with compounds **268** and **269** showing the most promising activities ($IC_{50} = 0.059 \pm 0.003$ and $0.080 \pm 0.005 \mu M$, respectively), along with high inhibition of $A\beta_{1-42}$ aggregation ($IC_{50} = 10.1 \pm 0.09$ and $10.9 \pm 0.15 \mu M$, respectively). Molecular modelling studies suggested that compounds **268** and **269** have significant binding affinity with both CAS and PAS of the AChE. In

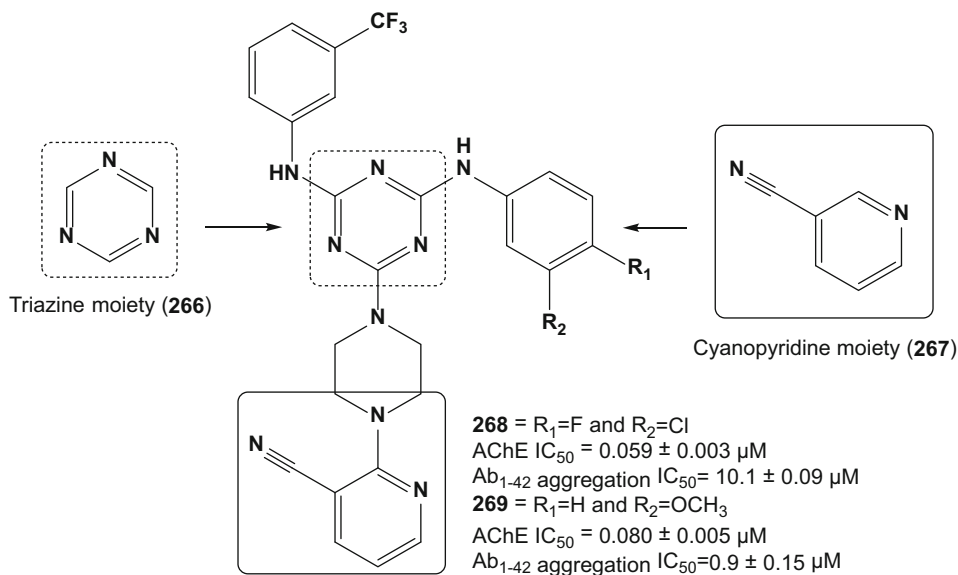


Fig. 73 Chemical structures of the very active MTDL cyanopyridine–triazine hybrids **268** and **269**

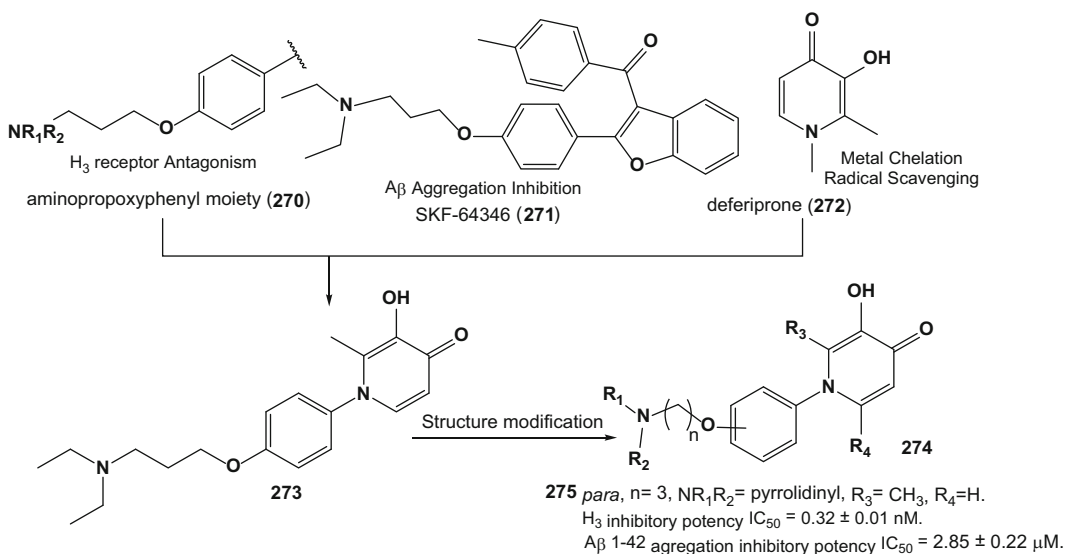


Fig. 74 Rational design of 1-phenyl-3-hydroxy-4-pyridinone derivatives (**275**) as multifunctional agents

addition, the results from neuroprotection studies indicated that these derivatives can reduce H₂O₂-induced neuronal death mediated by oxidative stress and Aβ₁₋₄₂ induced cytotoxicity [137].

Sheng and co-workers planned a series of novel 1-phenyl-3-hydroxy-4-pyridinone (**275**), as multifunctional agents for AD therapy through incorporation of 3-hydroxy-4-pyridinone moiety from deferiprone (**272**) into the general scaffold of H₃ receptor antagonists (**271**, Fig. 74). The 3-hydroxy-4-pyridinone moiety

(274) of the commercial metal chelator deferiprone was chosen due to its desirable chelating properties with high affinity for Cu^{2+} , Zn^{2+} , and Fe^{3+} ions, but low affinity for sodium, potassium, and magnesium. The generally accepted pharmacophore model of H_3 receptor antagonists (272) is composed of a “western part” containing a tertiary basic amine and a spacer, an aromatic central core, and an “eastern part” consisting of a polar group, a second basic amine, or a lipophilic residue. Next, they prioritized the introduction of an appropriate metal-chelating pharmacophore 274 according to the “eastern part” attributes to generate novel molecular hybrids with multiple functions. Interestingly, the newly designed 1-phenyl-3-hydroxy-4-pyridinone derivatives share similar structural elements to $\text{A}\beta$ aggregation inhibitor SKF-64346 (271). Thus, molecular hybridization of the molecular models 271 and 272 led to compound 273. Structural modifications on compound 273 were carried out and gave rise to series 274. Among all tested compounds, hybrid 275 displayed excellent selective H_3 receptor antagonistic activity ($\text{IC}_{50} = 0.32 \pm 0.01$ nM), efficient $\text{ABTS}\cdot$ + scavenging effect (1.54 ± 0.15), good Cu^{2+} and Fe^{3+} chelating properties, and effective inhibitory activity against self- and Cu^{2+} -induced $\text{A}\beta_{1-42}$ aggregation ($\text{IC}_{50} = 2.85 \pm 0.22$ μM). More interestingly, an *in vivo* study revealed that 275 possesses suitable PK profiles in plasma and acceptable BBB penetration behavior, highlighting this compound as a new promising and innovative drug candidate prototype for AD [138].

Wei and co-workers used the structure of oxoisoaporphine (276) to plan two series of 8- and 11-substituted-amide derivatives (277 and 278, Fig. 75) as potential inhibitors of AChE and $\text{A}\beta$ aggregation with neuroprotective properties. Oxoisoaporphine is an alkaloid and was previously isolated and identified as an important lead for anti-Alzheimer drugs [139]. Biological evaluation of the oxoisoaporphine derivatives 277a–e and 278a–e disclosed that analogues with the same length in side chain exhibited similar inhibitory potency towards AChE. Compounds 277b and 278b, with two methylene groups in the side chain, showed the strongest AChE inhibitory potency ($\text{IC}_{50} = 28.4 \pm 14$ nM and 80.8 ± 26 nM, respectively) and a significant inhibition of self-induced $\text{A}\beta_{1-42}$ aggregation potency ($\text{IC}_{50} = 74.6 \pm 1.9$ and 76.1 ± 1.2 , respectively). Moreover, these compounds showed to be able to cross the BBB to reach their targets in the CNS and a significant reduction in $\text{A}\beta$ secretion levels by human neuronal cells (SH-SY5Y), which overexpress the Swedish mutant form of human β -amyloid precursor protein (APP_{sw}) [140].

In another approach for searching innovative and effective drug candidates for AD, and based in previous research data [141], Hebda and co-workers reported the synthesis and evaluation of two novel series of phthalimide derivatives 279 and 281 (A and B, Fig. 76), designed as DBS AChE inhibitors. Compound

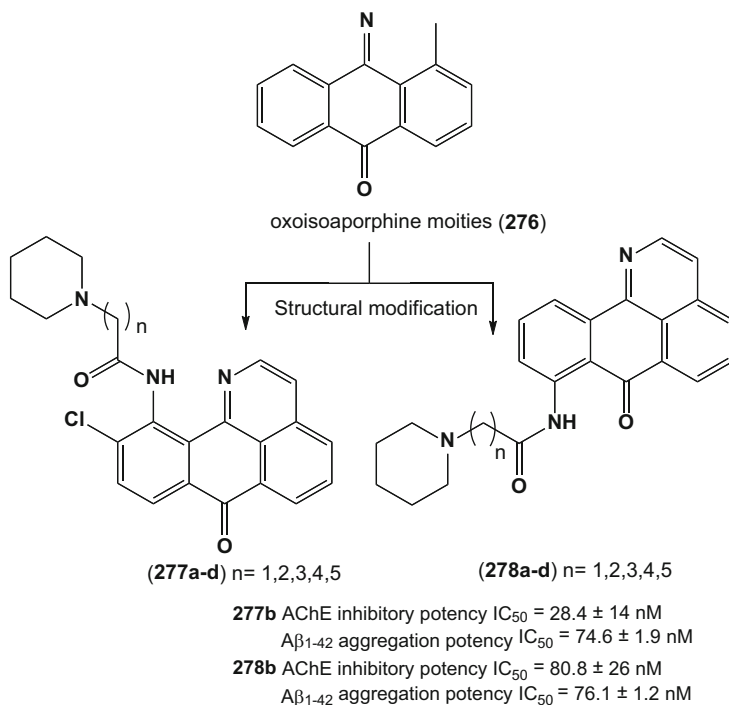


Fig. 75 Structure of 8- and 11-substituted 1-azabenzanthrone derivatives **277** and **278**, designed as oxisoaporphine analogues

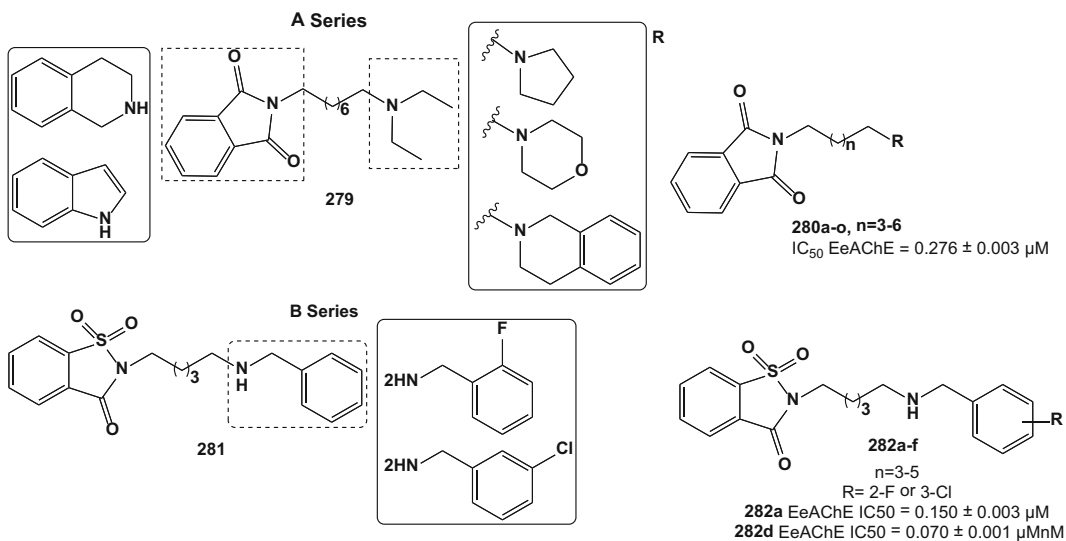


Fig. 76 Chemical structure of compounds **280** and **282**, designed as dual hybrid inhibitors of AChE and BuChE and amyloid β aggregation

280 (from series A), holding a pyrrolidine moiety, was the most potent AChE inhibitor with $IC_{50} = 0.276 \pm 0.003 \mu\text{M}$. Regarding the length of the linker, the most potent were the compounds with six to eight carbon atoms as linker length. Some compounds from series B inhibited EeAChE in the low micromolar to nanomolar range, highlighting compounds with a five-carbon atom linker. Compounds **282a** (2-fluoro derivative) and **282d** (3-chloro derivative) were the most potent AChE inhibitors with IC_{50} values of 150 and 70 nM, respectively. These two compounds were also evaluated for their ability to inhibit self-induced $A\beta_{1-42}$ aggregation, with only compound **280n** demonstrating a moderate inhibition activity of 35.8% [142].

In another strategy, considering that molecular hybridization plays a highlighted role in the design of novel MTDLs and that the search of new pharmacophores is determinant in the search of novel drug candidate prototypes, more than 600 compounds were screened by HTS (high throughput screening) for AChE inhibitory activity. The aim was to identify promising structures to be used in a structural optimization study. As a result, the pyrimidine derivative **202** (AChE inhibition of 32.15% at 40 μM) was chosen as initial prototype molecule and 15 derivatives were synthesized and evaluated for their ability to inhibit AChE (Fig. 77). Aiming to identify the optimal side chain for the desired enzyme inhibition, compound **284** was recognized as the most active of this first series (77.26% of AChE activity at 40 μM and $IC_{50} = 8.14 \pm 0.40 \mu\text{M}$). In a second step, optimization of the lateral 1,2-diazol ring in compound **284** led to additional nine new analogues, with the 1,3-diazol derivative **285** disclosing a discreet improving in AChE inhibition ($IC_{50} = 1.59 \pm 0.02 \mu\text{M}$). In sequence, modification in the central ring confirmed that 1,3-diazine was the best pharmacophore, and it was conserved in another 24 analogues with different substituents at the lateral 1,3-imidazolyl ring. After all these SAR studies, compound **287** was finally identified with a very improved

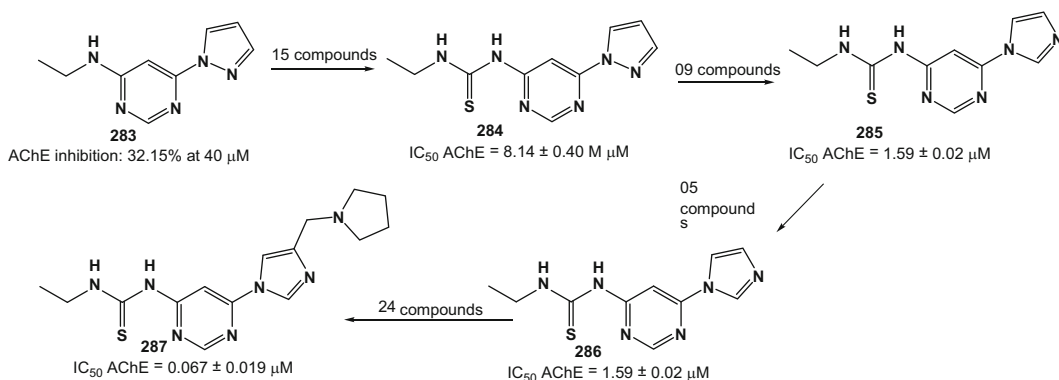


Fig. 77 Key transition compounds of the structural optimization of pyrimidine derivatives (**283–287**)

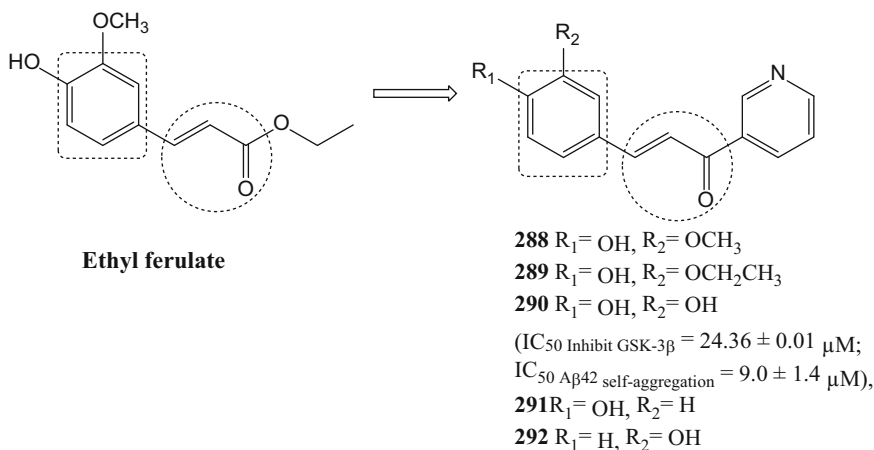


Fig. 78 Design strategy leading to derivatives **288–292**

selective AChE inhibitory activity ($\text{IC}_{50} = 0.067 \pm 0.019 \mu\text{M}$), and additional chelating ability for Cu^{2+} ions, antioxidant activity, inhibitory activity of A β -aggregation, and low toxicity in human neuroblastoma cells [143].

Simone and co-workers synthesized a series of hydroxy-substituted *trans*-cinnamoyl derivatives **288–292** as multifunctional tools in the context of DA. To confirm the multifunctional activity profile envisaged for the designed compounds (Fig. 78), they were evaluated for their ability to inhibit GSK-3 β and A β_{42} self-aggregation and to counteract reactive oxygen radical (ROS) formation. Compound **290** revealed the most promising multifunctional profile, showing ability to inhibit both GSK-3 β ($\text{IC}_{50} = 24.36 \pm 0.01 \text{ mM}$) and A β_{42} self-aggregation ($\text{IC}_{50} = 9.0 \pm 1.4 \text{ mM}$), chelate copper (II) and act as exceptionally strong radical scavenger ($k_{\text{inh}} = 6.8 \pm 0.5 \cdot 10^5 \text{ M}^{-1} \text{ s}^{-1}$) even in phosphate buffer at pH 7.4 ($k_{\text{inh}} = 3.2 \pm 0.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$). Importantly, compound **290** showed high-predicted BBB permeability, did not exert any significant cytotoxic effects in immature cortical neurons up to 50 μM and showed additional neuroprotective properties at micromolar concentration against toxic insult induced by glutamate [144].

Ozadali-Sari and co-workers described the synthesis, pharmacological evaluation (BuChE/AChE inhibition, A β anti-aggregation, and neuroprotective effects), and molecular modeling studies of 2-[4-(4-substituted piperazin-1-yl)phenyl]benzimidazole derivatives **293–300** (Fig. 79). The alkyl-substituted derivatives exhibited selective inhibition on BuChE and the compounds **294** and **296** were found to be the most potent inhibitors of BuChE with IC_{50} values of 5.18 and 5.22 μM , respectively. The compounds with an inhibitory effect on BuChE were subsequently screened for their A β anti-aggregating and neuroprotective

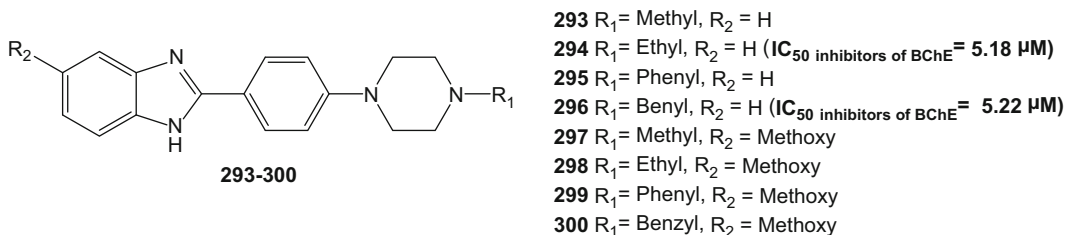


Fig. 79 Target compounds 2-[4-(4-substituted piperazin-1-yl)phenyl]benzimidazole derivatives **293–300**

activities. Compounds **293** and **294** exerted a potential neuroprotective effect against H₂O₂ with cell viability of 104% and 105%, respectively, which is similar to rifampicin (102%), and Aβ induced cytotoxicity in SH-SY5Y cells results revealed that **293**, **294**, **296**, and **298** were the most potent compounds in the series with cell viability ranging from 81 to 130% at 10 μM. The protective effect of these compounds against Aβ_{1–40} induced cell death was significantly higher compared to donepezil (60%). Collectively, compound **294** was found as the most promising compound for the development of MTDL against AD [145].

Aiming the identification of multi-targeted prototypes with multiple activities as MT-stabilizing agents and/or inhibitors of the cyclooxygenase (COX) and 5-lipoxygenase (5-LOX) pathways, Cornec and co-workers started a SAR study from a series of 1,5-diarylimidazoles with microtubule (MT)-stabilizing activity and structural similarities with known NSAIDs. Several compounds showed brain-penetrant abilities and exhibited balanced multi-targeted in vitro activity in the low μM range, with compounds **301–304** (Fig. 80) showing the most promising properties as MTDLs prototype candidates [146].

A series of *N*-propargylpiperidines with naphthalene-2-carboxamide or naphthalene-2-sulfonamide moieties was synthesized and their potential multifunctional anti-Alzheimer was investigated by Kosak and co-workers. Their rational design approach was based on piperidine-based selective human BuChE (hBuChE) inhibitors and propargylamine-based MAO inhibitors. Biological investigation disclosed all compounds as good hBuChE inhibitors, with good selectivity over AChE and adequate BBB permeability in a PAMPA. The dual BuChE (IC₅₀ = 2.60 ± 0.35 μM) and MAO-B (IC₅₀ = 53.9 ± 4.78 μM) inhibitor **307** (Fig. 81) also protected neuronal SH-SY5Y cells from toxic amyloid β-peptide species, without any cytotoxic effects [147].

A series of novel ferulic acid-*O*-alkylamines derivatives were designed as multi-target against AD by Sang and co-authors. In vitro assays displayed that all the synthesized target compounds showed impressive inhibitory activity against BuChE, with significant inhibition/disaggregation of self-induced Aβ aggregation and

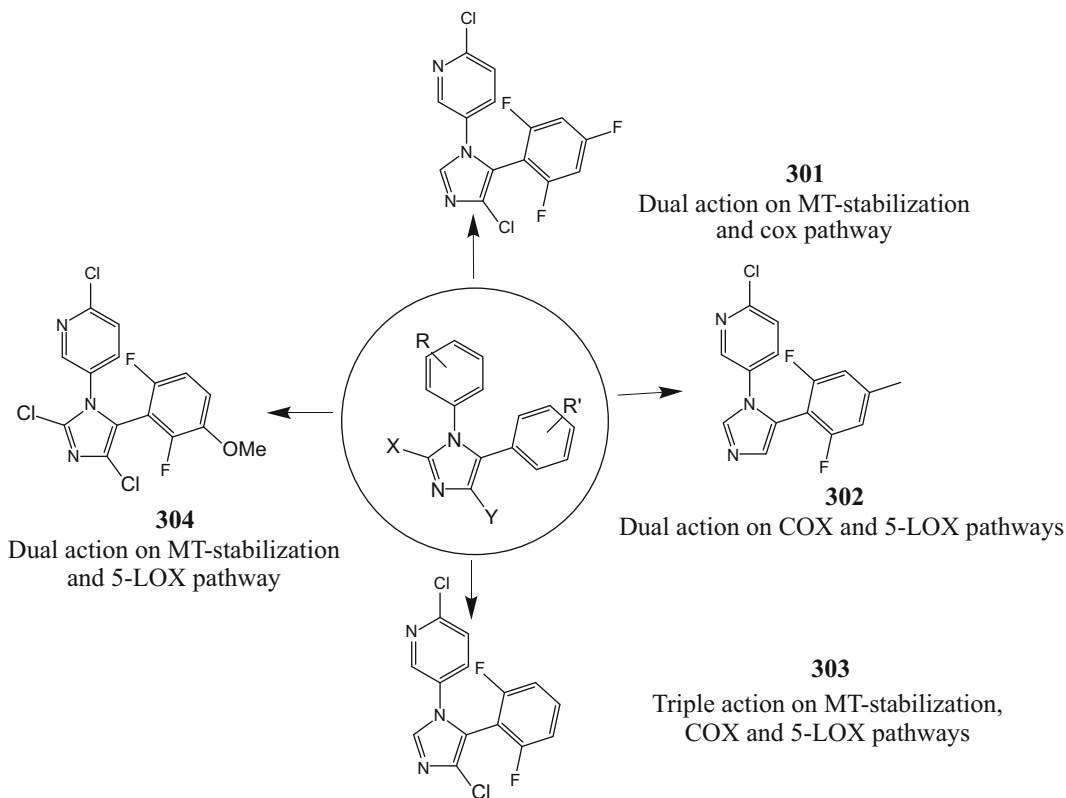


Fig. 80 Summary of SARs and representative structures exemplifying the multi-target activity profiles exhibited by the 1,5-diarylimidazoles and the most promising compounds

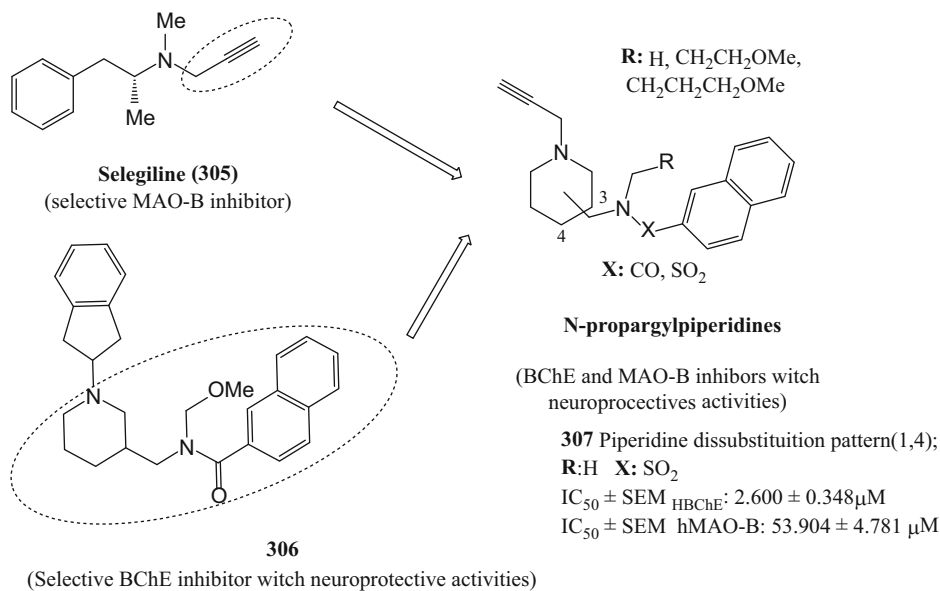


Fig. 81 Design of the *N*-propargylpiperidines and the compound 307

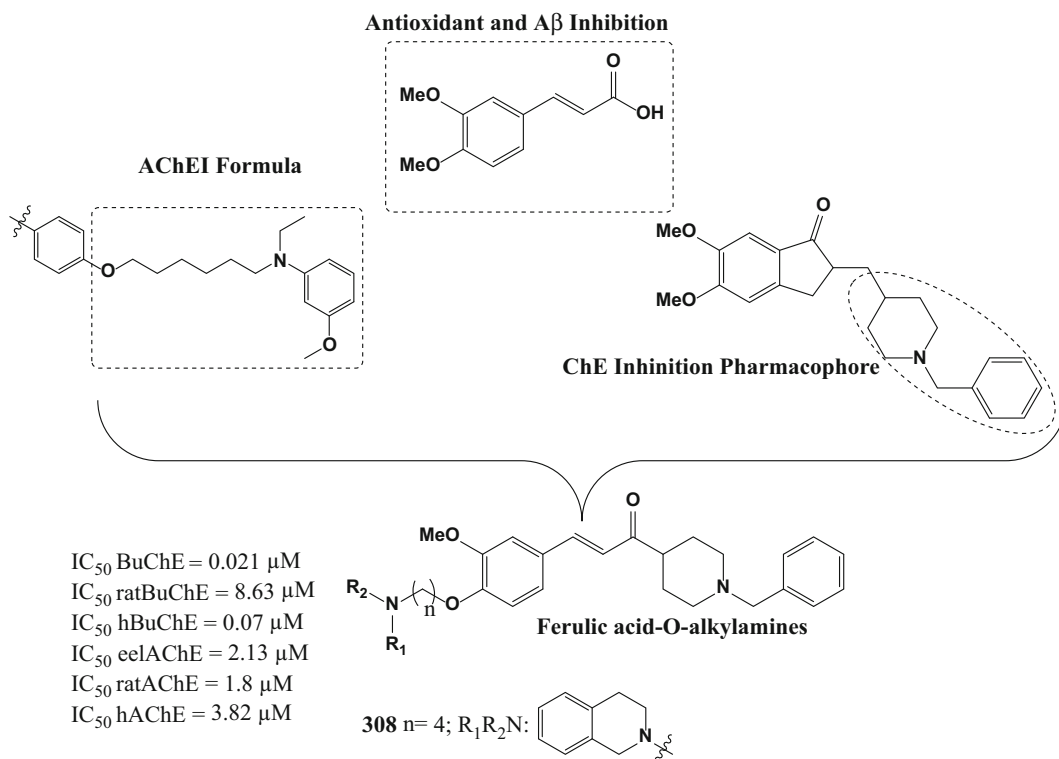


Fig. 82 Design for ferulic acid-*O*-alkylamines and compound **308**

antioxidant activity. Compound **308** (Fig. 82), one of the most potent BuChE inhibitors (IC_{50} = 0.021 μ M for EqBuChE, 8.63 μ M for ratBuChE, and 0.07 μ M for hBuChE), was found to be also a good selective AChE inhibitor (IC_{50} = 1.8 μ M for RatAChE, 2.13 μ M for EeAChE, and 3.82 μ M for hAChE). This compound also had noteworthy inhibitory effects on self-induced A β_{1-42} aggregation ($50.8 \pm 0.82\%$) and was capable to disaggregate self-induced A β_{1-42} aggregation (38.7 ± 0.65). Furthermore, compound **308** showed a modest antioxidant activity (0.55 eq of Trolox), good protective effect against H₂O₂-induced PC12 cell injury, with low toxicity and good BBB permeability in vitro. Moreover, compound **308** did not exhibit any acute toxicity in mice at doses up to 1.000 mg/kg, and in the step-down passive avoidance test this compound significantly reversed scopolamine-induced memory deficit in mice. These results set indicated compound **308** as a very promising multifunctional agent in the treatment of AD, particularly the advanced stages [148].

Mohamed and Rao designed a novel class of 2,4-disubstituted quinazoline derivatives, planned as innovative multi-targeting agents to treat AD. Biological results demonstrated the ability of several quinazoline derivatives in inhibiting both AChE and BuChE, with additional prevention of A β aggregation and

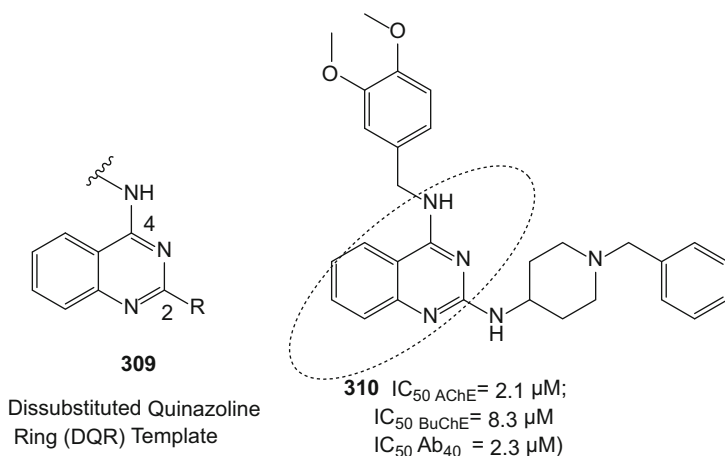


Fig. 83 Template and compound **310**, multi-targeting agent to treat Alzheimer's disease

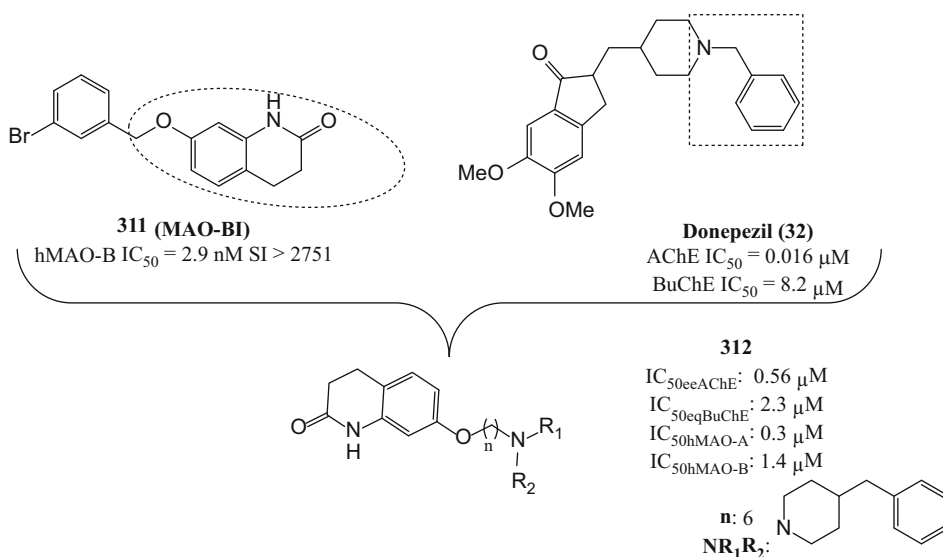
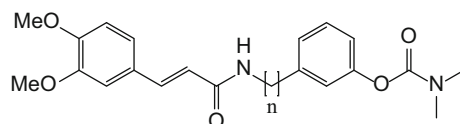


Fig. 84 Design strategy for 3,4-dihydro-2(1*H*)-quinoline-*O*-alkylamine derivatives and most promising compound **312**

antioxidant properties. Compound **310** (Fig. 83) was identified as a dual inhibitor of cholinesterases (IC_{50} AChE = 2.1 μ M; IC_{50} BuChE = 8.3 μ M) and inhibitory activity of A β aggregation (IC_{50} A β 40 = 2.3 μ M) [149].

A new series of 3,4-dihydro-2(1*H*)-quinoline-*O*-alkylamine derivatives, capable of interacting with both AChE and BuChE as well as MAO-A and B was designed by Sang and co-workers by using a conjunctive approach that combines structural fragments of JMC49 and donepezil **32** (Fig. 84). The most promising compound **312** showed potent and balanced inhibitory activities

313 $n=2$ BuChE Inhibition at 1 μM = 76% $\text{A}\beta_{1-42}$ self-aggregation inhibitory 10 μM = 85.3%**Fig. 85** Structure of compound 313 (rivastigmine–hydroxycinnamic acid hybrid)

towards AChE ($\text{IC}_{50} = 0.56 \mu\text{M}$), BuChE ($\text{IC}_{50} = 2.3 \mu\text{M}$), MAO-A ($\text{IC}_{50} = 0.3 \mu\text{M}$), and MAO-B ($\text{IC}_{50} = 1.4 \mu\text{M}$), but with low selectivity for both cholinesterase and monoaminoxidase isoforms. Furthermore, this study revealed that compound **312** could adequately cross BBB *in vitro* and abided by Lipinski's rule of five, which suggests this derivative as a potential and interesting multi-targeted active molecule for further development [150].

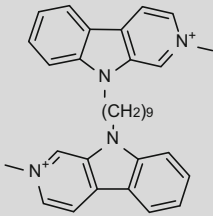
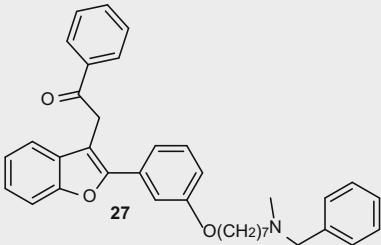
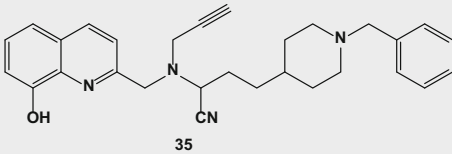
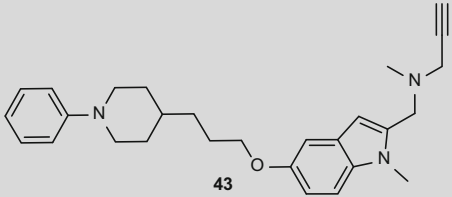
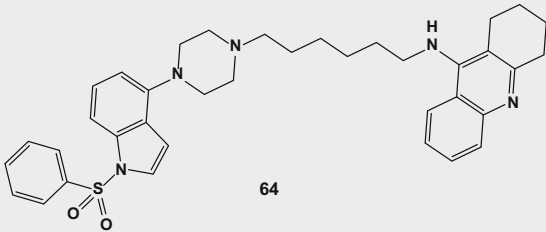
A series of rivastigmine–caffeic acid and rivastigmine–ferulic acid hybrids were designed, synthesized, and evaluated as multi-functional agents for AD *in vitro* by Chen and co-authors. In general, these new compounds exerted antioxidant neuroprotective properties, good ChE inhibitory activities, and with some compounds also inhibiting $\text{A}\beta$ aggregation. In special, compound **313** (Fig. 85) emerged as promising drug candidate endowed with neuroprotective potential, stronger inhibition (inhibition rate of 76%) against BuChE than rivastigmine at 1 μM , inhibition of $\text{A}\beta_{1-42}$ self-aggregation (85.3% at 10 μM) and copper chelation properties [133].

10 General Vision of the Main Therapeutic Strategies Explored in the Search of Multi-target Directed Ligands for AD Treatment

As discussed above, a number of research groups have explored different ways in the design and discovery of innovative drug candidate prototypes capable to represent effective advances in the development of therapeutics and, hopefully, the cure of AD. In fact, there are many active and promising ligands recently discovered, with diverse structural pattern and mechanisms of action that could represent innovation and could seduce the Pharmaceutical Industry to invest in these new alternatives.

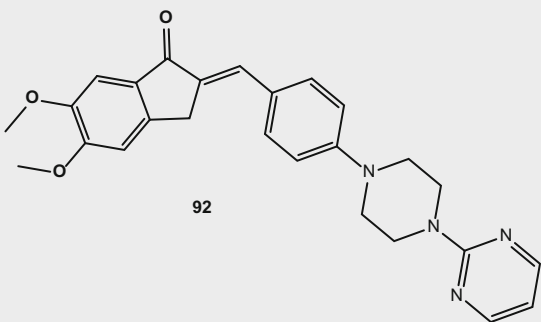
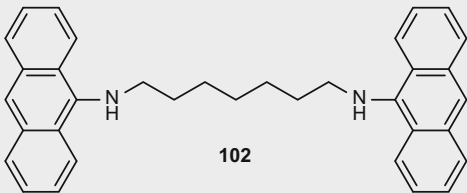
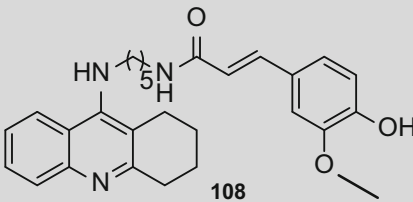
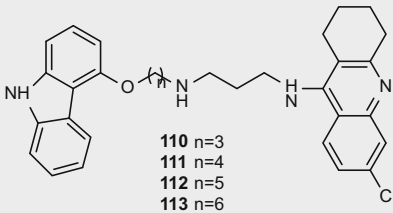
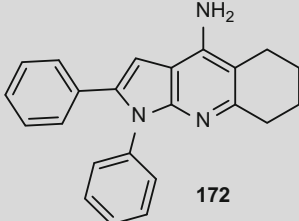
The main advances in the search of such innovative MTDLs are summarized in Table 1.

Table 1
Chemical structure and biological data of the main multi-target directed ligands (MTDLs) designed as innovative drug candidate prototypes for AD

Compound	Target/IC ₅₀	Ref.
MTDLS inspired by galanthamine		
 18. β-carboline derivatives	AChE = 0.5 nM	[44]
 27	hAChE = 24 ± 2 nM hBuChE = 2.88 ± 0.26 μM	[45] [46]
MTDLS inspired by donepezil		
 35	MAO-A = 10.1 ± 1.1 μM MAO-B = >100 μM AChE = 29 ± 3 nM BuChE = 39 ± 3 nM ORAC = 1.12 ± 0.43	[49]
 43	MAO-A = 5.2 nM MAO-B = 43.0 nM AChE = 0.35 μM BuChE = 0.46 μM	[52]
 64	5HT6Rs = 2.0 ± 0.2 nM AChE = 12.9 ± 0.17 nM BuChE = 8.2 ± 0.23 nM	[61]

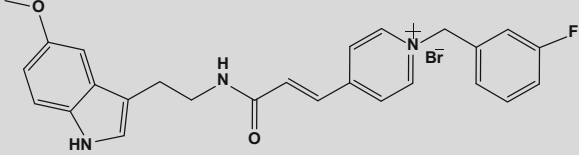
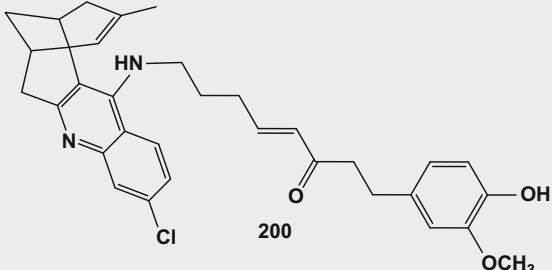
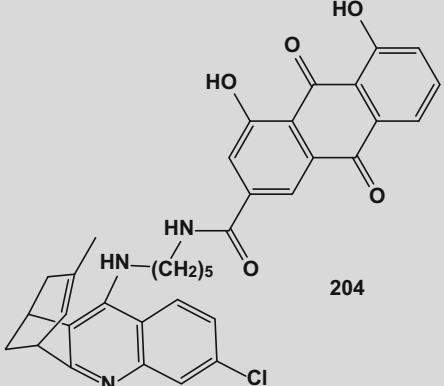
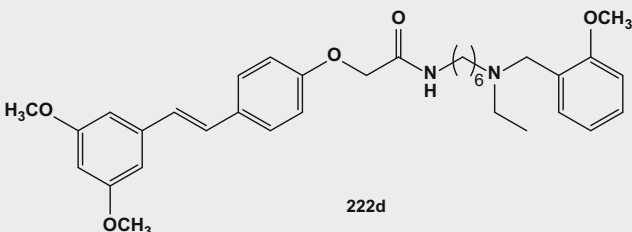
(continued)

Table 1
(continued)

Compound	Target/IC ₅₀	Ref.
 <p>92</p>	AChE = $0.025 \pm 0.001 \mu\text{M}$ [76] Ab aggregation = $9.9 \pm 0.14 \mu\text{M}$	
MTDLs inspired by tacrine		
 <p>102</p>	AChE = 0.81 nM [35] BuChE = 5.66 nM	
 <p>108</p>	AChE = 4.4 nM [81]	
 <p>110 n=3 111 n=4 112 n=5 113 n=6</p>	110 n = 3 AChE = 2.15 nM, BuChE = 296 nM 111 n = 4 AChE = 1.65 nM, BuChE = 211 nM 112 n = 5 AChE = 1.54 nM, BuChE = 189 nM 113 n = 6 AChE = 2.57 nM, BuChE = 137 nM	[6]
 <p>172</p>	AChE = 0.61 nM [98] BuChE = 0.074 nM	

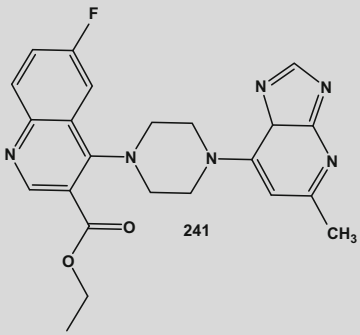
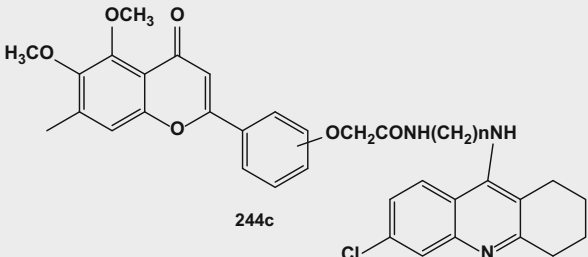
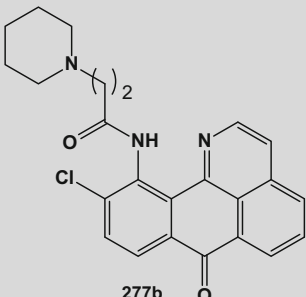
(continued)

Table 1
(continued)

Compound	Target/IC ₅₀	Ref.
MTDLs inspired by natural products (NPs)		
 <p>189</p>	AChE = $0.11 \pm 0.001 \mu\text{M}$ BuChE = $1.1 \pm 0.1 \mu\text{M}$ ORAC = 3.41 ± 0.05	[107]
 <p>200</p>	AChE = $6.7 \pm 0.1 \text{ nM}$	[111]
 <p>204</p>	AChE = 1.07 nM	[112]
 <p>222d</p>	AChE = $0.06 \pm 0.03 \mu\text{M}$ BuChE = $28.04 \pm 1.71 \mu\text{M}$ ORAC 0.51 ± 0.03	[118]

(continued)

Table 1
(continued)

Compound	Target/IC ₅₀	Ref.
MTDLs inspired by other polycyclic structures		
 <p>241</p>	AChE = 0.042 ± 0.79 mM BuChE = 0.51 ± 0.91 mM ORAC = 2.43 ± 0.61	[125]
 <p>244c</p> <p>244a-d n=3, 4, 5, 6 spacer linking in <i>Para</i> position</p>	AChE = 12.8 ± 0.05 nM	[126]
 <p>277b</p>	AChE = 28.4 ± 14 nM Aβ ₁₋₄₂ aggregation = 74.6 ± 1.9 nM	[140]

11 Concluding Remarks

The development of more effective, secure, well tolerated, with innovative mechanism of action drugs for Alzheimer's disease is still a challenge for Medicinal Chemistry. Due to the characteristic multifactorial pathogenesis of AD and other neurodegenerative disorders, the discovering and development of innovative disease-modifying drugs is a problem to be solved to overcome current clinical limitations and therapeutical inefficiency for such diseases.

Regarding this complex multifactorial pathophysiology, polypharmacology, and especially the multifunctional or MTDLs approach represents a new perspective and stated a new paradigm in the way of thinking, design, evaluate, and interpret the simultaneously modulation of diverse molecular targets, aiming to block out compensative mechanisms and alternative biochemical routes, reaching a more efficient control of the progress and, hopefully, the cure of the disease. The more recent insights about the pathophysiology of AD have proven a complex interconnected variety of deleterious events, with multiple pathways involving in its etiology and a variety of factors that are associated with the installation, progress, and severity of AD. In this way, in consonance with many medicinal chemists worldwide, we can affirm that the AD treatment will be effective only when we find druggable chemical entities capable of interacting and modulating concomitantly different targets, and blocking or activating different biochemical interconnected pathways related to the disease progression. In the last decade, efforts have been employed in the design and search for new biological chemical entities capable to direct simultaneously with different molecular targets involved in the pathogenesis of AD and many new significantly active and promise molecules have been discovered. Consequently, new drug prototype candidates have been identified and they are the starting point of a revolutionary new way of thinking drug design and some drug candidates are under preclinical evaluation. Current drugs, such as donepezil, galanthamine, rivastigmine as well as the restricted tacrine, have been used as the starting point for molecular inspiration in the drawing of innovative hybrid compounds with multifunctional mechanism of action. Molecular hybridization is the major tool in the design of new scaffolds, leading to the conception of new molecules with diverse and singular structural characteristics that could make possible to access simultaneously different molecular targets related to the pathophysiology of AD. One great challenge for this new concept of MTDLs, differently to the old paradigm of one potent and target-selective drug, is to achieve balanced properties of one single molecule to act in a multi-target environment or in interconnected diverse sites or biochemical ways related to the installation and progression of the disease. As discussed in this chapter and highlighted in Table 1, many bioactive and promising molecules have been discovered and proved to be interesting for further investments in the evaluation and development, stage where the interest and cooperation by the Pharma Industry is of capital importance. Despite none of these MTDL drug candidates have reached clinical phase of development, we hope that in the next few years the system biology approach could effectively contribute to the Medicinal Chemists in the discovery of innovative chemical entities able to understand and effectively modify the course of this devastating neurodegenerative disorder.

Acknowledgements

The authors are grateful to the Brazilian Agencies CNPq (#454088/2014-0, #400271/2014-1, #310082/2016-1), FAPEMIG (#CEX-APQ-00241-15), FINEP, INCT-INOVAR (#465.249/2014-0), PRPPG-UNIFAL, and CAPES for financial support and fellowships.

References

1. Bolognesi ML, Matera R, Minarini A, Rosini M, Melchiorre C (2009) Alzheimer's disease: new approaches to drug discovery. *Curr Opin Chem Biol* 13:303–308. <https://doi.org/10.1016/j.cbpa.2009.04.619>
2. Youdim MBH, Buccafusco JJ (2005) Multi-functional drugs for various CNS targets in the treatment of neurodegenerative disorders. *Trends Pharmacol Sci* 26:27–35. <https://doi.org/10.1016/j.tips.2004.11.007>
3. Thies W, Bleiler L (2011) 2011 Alzheimer's disease facts and figures. *Alzheimers Dement* 7:208–244. <https://doi.org/10.1016/j.jalz.2011.02.004>
4. Fraga CAM, Barreiro EJ (2008) New insights for multifactorial disease therapy: the challenge of the symbiotic drugs. *Curr Drug Ther* 3:1–13. <https://doi.org/10.2174/157488508783331225>
5. Zhang H-Y (2005) One-compound-multiple-targets strategy to combat Alzheimer's disease. *FEBS Lett* 579:5260–5264. <https://doi.org/10.1016/j.febslet.2005.09.006>
6. Rosini M, Simoni E, Bartolini M, Cavalli A, Ceccarini L, Pascu N, McClymont DW, Tarozzi A, Bolognesi ML, Minarini A, Tumiatti V, Andrisano V, Mellor IR, Melchiorre C (2008) Inhibition of acetylcholinesterase, β -amyloid aggregation, and NMDA receptors in Alzheimer's disease: a promising direction for the multi-target-directed ligands gold rush. *J Med Chem* 51:4381–4384. <https://doi.org/10.1021/jm800577j>
7. Möller HJ, Graeber MB (1998) The case described by Alois Alzheimer in 1911. *Eur Arch Psychiatry Clin Neurosci* 248:111–122. <https://doi.org/10.1007/s004060050027>
8. Sayeg N (2013) Aspectos socioeconômicos. <http://www.alzheimermed.com.br/conceitos/aspectos>
9. Instituto Brasileiro de Geografia e Estatística (IBGE) (2010) Censo Demográfico 2010
10. Alzheimer's Association (2012) 2012 Alzheimer's disease facts and figures
11. Alzheimer's Disease International (2012) Dementia: a public health priority. World Health Organization, Geneva, pp 1–102
12. Kalache A (1991) Ageing in developing countries. *Crit Public Health* 2:38–43. <https://doi.org/10.1080/09581599108406812>
13. IBGE (2013) Censo 2000. http://www.ibge.gov.br/home/estatistica/populacao/%0Acenso2000/populacao/censo2000_populacao.pdf
14. Francis PT, Palmer AM, Snape M, Wilcock GK (1999) The cholinergic hypothesis of Alzheimer's disease: a review of progress. *J Neurol Neurosurg Psychiatry* 66:137–147. <https://doi.org/10.1136/jnnp.66.2.137>
15. Chen S-Y, Chen Y, Li Y-P, Chen S-H, Tan J-H, Ou T-M, Gu L-Q, Huang Z-S (2011) Design, synthesis, and biological evaluation of curcumin analogues as multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem* 19:5596–5604. <https://doi.org/10.1016/j.bmc.2011.07.033>
16. Ray B, Lahiri DK (2009) Neuroinflammation in Alzheimer's disease: different molecular targets and potential therapeutic agents including curcumin. *Curr Opin Pharmacol* 9:434–444. <https://doi.org/10.1016/j.coph.2009.06.012>
17. Viegas FPD, Simões MCR, da Rocha MD, Castelli MR, Moreira MS, Viegas C Jr (2011) Alzheimer's disease: characterization, evolution and implications of the neuroinflammatory process. *Rev Virtual Química* 3:286–306. <https://doi.org/10.5935/1984-6835.20110034>
18. Schmitt B, Bernhardt T, Moeller HJ, Heuser I, Frolich L (2004) Combination

- therapy in Alzheimer's disease: a review of current evidence. *CNS Drugs* 18:827–844
19. Starkov AA, Beal FM (2008) Portal to Alzheimer's disease. *Nat Med* 14:1020–1021
 20. Reddy PH, Beal MF (2008) Amyloid beta, mitochondrial dysfunction and synaptic damage: implications for cognitive decline in aging and Alzheimer's disease. *Trends Mol Med* 14:45–53. <https://doi.org/10.1016/j.molmed.2007.12.002>. Amyloid
 21. Wang J-Z, Grundke-Iqbal I, Iqbal K (2007) Kinases and phosphatases and tau sites involved in Alzheimer neurofibrillary degeneration. *Eur J Neurosci* 25:59–68. <https://doi.org/10.1111/j.1460-9568.2006.05226.x>
 22. Zetterberg H, Blennow K (2006) Plasma A β in Alzheimer's disease – up or down? *Neurology* 5:638–639
 23. Campos HC, Divino M, Pereira F, Viegas D, Nicastró PC, Fossaluzza PC, Alberto C, Fraga M, Barreiro EJ, Viegas C Jr (2011) The role of natural products in the discovery of new drug candidates for the treatment of neurodegenerative disorders I: Parkinson's disease. *CNS Neurol Disord Drug Targets* 10:239–250
 24. Liu Q, Xie F, Rolston R, Moreira P, Nunomura A, Zhu X, Smith M, Perry G (2007) Prevention and treatment of Alzheimer disease and aging: antioxidants. *Mini Rev Med Chem* 7:171–180. <https://doi.org/10.2174/138955707779802552>
 25. Legg K (2011) Neurodegenerative diseases: an alternative path to reduce neuroinflammation. *Nat Rev Drug Discov* 10:901. <https://doi.org/10.1038/nrd3607>
 26. Gaggelli E, Kozłowski H, Valensin D, Valensin G (2006) Copper homeostasis and neurodegenerative disorders (Alzheimer's, prion, and Parkinson's diseases and amyotrophic lateral sclerosis). *Chem Rev* 106:1995–2044. <https://doi.org/10.1021/cr040410w>
 27. Akiyama H, Barger S, Barnum S, Bradt B, Bauer J, Cole GM, Cooper NR, Eikelenboom P, Emmerling M, Fiebich BL, Finch CE, Frautschy S, Griffin WST, Hampel H, Hull M, Landreth G, Lue LF, Mrazek R, Mackenzie IR, McGeer PL, Banion MKO, Pachter J, Pasinetti G, Salaman CP, Rogers J, Rydel R, Shen Y, Streit W, Strohmeyer R, Tooyoma I, Van Muiswinkel FL, Veerhuis R, Walker D, Webster S, Wegrzyniak B, Wenk G, Coray TW (2000) Inflammation and Alzheimer's disease. *Neurobiol Aging* 21:383–421
 28. Heneka MT, O'Banion MK (2007) Inflammatory processes in Alzheimer's disease. *J Neuroimmunol* 184:69–91. <https://doi.org/10.1016/j.jneuroim.2006.11.017>
 29. Giunta B, Fernandez F, Nikolic WV, Obregon D, Rrapo E, Town T, Tan J (2008) Inflammaging as a prodrome to Alzheimer's disease. *J Neuroinflammation* 5:51. <https://doi.org/10.1186/1742-2094-5-51>
 30. Kamal MA, Greig NH, Reale M (2009) Anti-inflammatory properties of acetylcholinesterase inhibitors administered in Alzheimer's disease. *Anti-Inflammatory Anti-Allergy Agents Med Chem* 8:85–100
 31. Mangialasche F, Solomon A, Winblad B, Mecocci P, Kivipelto M (2010) Alzheimer's disease: clinical trials and drug development. *Lancet Neurol* 9:702–716. [https://doi.org/10.1016/S1474-4422\(10\)70119-8](https://doi.org/10.1016/S1474-4422(10)70119-8)
 32. Piau A, Nourhashemi F, Hein C, Caillaud C, Vellas B (2011) Progress in the development of new drugs in Alzheimer's disease. *J Nutr Health Aging* 15:45–57. <https://doi.org/10.1007/s12603-011-0012-x>
 33. Cavalli A, Bolognesi ML, Minarini A, Rosini M, Tumiatti V, Recanatini M, Melchiorre C (2008) Multi-target-directed ligands to combat neurodegenerative diseases. *J Med Chem* 51:347–372. <https://doi.org/10.1021/jm7009364>
 34. Samadi A, Valderas C, Ríos CDL, Bastida A, Chioua M, González-Lafuente L, Colmena I, Gandía L, Romero A, Del Barrio L, Martín-De-Saavedra MD, López MG, Villarroya M, Marco-Contelles J (2011) Cholinergic and neuroprotective drugs for the treatment of Alzheimer and neuronal vascular diseases. II. Synthesis, biological assessment, and molecular modelling of new tacrine analogues from highly substituted 2-aminopyridine-3-carbonitriles. *Bioorg Med Chem* 19:122–133. <https://doi.org/10.1016/j.bmc.2010.11.040>
 35. Bolognesi ML, Cavalli A, Valgimigli L, Bartolini M, Rosini M, Andrisano V, Recanatini M, Melchiorre C (2007) Multi-target-directed drug design strategy: from a dual binding site acetylcholinesterase inhibitor to a trifunctional compound against Alzheimer's disease. *J Med Chem*

- 50:6446–6449. <https://doi.org/10.1021/jm701225u>
36. Cavalli A, Bolognesi ML, Capsoni S, Andrisano V, Bartolini M, Margotti E, Cattaneo A, Recanatini M, Melchiorre C (2007) A small molecule targeting the multifactorial nature of Alzheimer's disease. *Angew Chem Int Ed* 46:3689–3692. <https://doi.org/10.1002/anie.200700256>
 37. Shan WJ, Huang L, Zhou Q, Meng FC, Li XS (2011) Synthesis, biological evaluation of 9-N-substituted berberine derivatives as multi-functional agents of antioxidant, inhibitors of acetylcholinesterase, butyrylcholinesterase and amyloid- β aggregation. *Eur J Med Chem* 46:5885–5893. <https://doi.org/10.1016/j.ejmech.2011.09.051>
 38. Jiang H, Wang X, Huang L, Luo Z, Su T, Ding K, Li X (2011) Benzenediol-berberine hybrids: multifunctional agents for Alzheimer's disease. *Bioorg Med Chem* 19:7228–7235. <https://doi.org/10.1016/j.bmc.2011.09.040>
 39. Fernández-Bachiller MI, Pérez C, González-Muñoz GC, Conde S, López MG, Villarroya M, García AG, Rodríguez-Franco MI (2010) Novel tacrine-8-hydroxyquinoline hybrids as multifunctional agents for the treatment of Alzheimer's disease, with neuroprotective, cholinergic, antioxidant, and copper-complexing properties. *J Med Chem* 53:4927–4937. <https://doi.org/10.1021/jm100329q>
 40. Bolognesi ML, Cavalli A, Melchiorre C (2009) Memoquin: a multi-target – directed ligand as an innovative therapeutic opportunity for Alzheimer's disease. *Neurotherapeutics* 6:152–162
 41. Bolognesi ML, Cavalli A, Bergamini C, Fato R, Lenaz G, Rosini M, Bartolini M, Andrisano V, Melchiorre C (2009) Toward a rational design of multitarget-directed antioxidants: merging memoquin and lipoic acid molecular frameworks. *J Med Chem* 52:7883–7886. <https://doi.org/10.1021/jm901123n>
 42. Bolognesi ML, Bartolini M, Tarozzi A, Morroni F, Lizzi F, Milelli A, Minarini A, Rosini M, Hrelia P, Andrisano V, Melchiorre C (2011) Multitargeted drugs discovery: balancing anti-amyloid and anticholinesterase capacity in a single chemical entity. *Bioorg Med Chem Lett* 21:2655–2658. <https://doi.org/10.1016/j.bmcl.2010.12.093>
 43. Simoni E, Daniele S, Bottegoni G, Pizzirani D, Trincavelli ML, Goldoni L, Tarozzi G, Reggiani A, Martini C, Piomelli D, Melchiorre C, Rosini M, Cavalli A (2012) Combining galanthamine and memantine in multitargeted, new chemical entities potentially useful in Alzheimer's disease. *J Med Chem* 55:9708–9721. <https://doi.org/10.1021/jm3009458>
 44. Rook Y, Schmidtke KU, Gaube F, Schepmann D, Wunsch B, Heilmann J, Lehmann J, Winckler T (2010) Bivalent β -carbolines as potential multitarget anti-Alzheimer agents. *J Med Chem* 53:3611–3617. <https://doi.org/10.1021/jm1000024>
 45. Rizzo S, Tarozzi A, Bartolini M, Da Costa G, Bisi A, Gobbi S, Belluti F, Ligresti A, Allarà M, Monti J, Andrisano V, Di Marzo V, Hrelia P, Rampa A (2012) 2-Arylbenzofuran-based molecules as multipotent Alzheimer's disease modifying agents. *Eur J Med Chem* 58:519–532. <https://doi.org/10.1016/j.ejmech.2012.10.045>
 46. Piazzi L, Cavalli A, Colizzi F, Belluti F, Bartolini M, Mancini F, Recanatini M, Andrisano V, Rampa A (2008) Multi-target-directed coumarin derivatives: hAChE and BACE1 inhibitors as potential anti-Alzheimer compounds. *Bioorg Med Chem Lett* 18:423–426. <https://doi.org/10.1016/j.bmcl.2007.09.100>
 47. Jackson S, Ham RJ, Wilkinson D (2004) The safety and tolerability of donepezil in patients with Alzheimer's disease. *Br J Clin Pharmacol* 58:1–8. <https://doi.org/10.1111/j.1365-2125.2004.01848.x>
 48. Sugimoto H, Ogura H, Arai Y, Iimura Y, Yamanishi Y (2002) Research and development of donepezil hydrochloride, a new type of acetylcholinesterase inhibitor. *Jpn J Pharmacol* 89:7–20. <https://doi.org/10.1254/jjp.89.7>
 49. Wu M-Y, Esteban G, Brogi S, Shionoya M, Wang L, Campiani G, Unzeta M, Inokuchi T, Butini S, Marco-Contelles J (2015) Donepezil-like multifunctional agents: design, synthesis, molecular modeling and biological evaluation. *Eur J Med Chem* 121:1–16. <https://doi.org/10.1016/j.ejmech.2015.10.001>
 50. Pudlo M, Luzet V, Ismaïli L, Tomassoli I, Iutzeler A, Refouvet B (2014) Quinolone-benzylpiperidine derivatives as novel acetylcholinesterase inhibitor and antioxidant hybrids for Alzheimer disease. *Bioorg Med Chem* 22:2496–2507. <https://doi.org/10.1016/j.bmc.2014.02.046>

51. Samadi A, Chioua M, Bolea I, De Los Ríos C, Iriepa I, Moraleda I, Bastida A, Esteban G, Unzeta M, Gálvez E, Marco-Contelles J (2011) Synthesis, biological assessment and molecular modeling of new multipotent MAO and cholinesterase inhibitors as potential drugs for the treatment of Alzheimer's disease. *Eur J Med Chem* 46:4665–4668. <https://doi.org/10.1016/j.ejmech.2011.05.048>
52. Bolea I, Juárez-Jiménez J, de los Ríos C, Chioua M, Pouplana R, Luque FJ, Unzeta M, Marco-Contelles J, Samadi A (2011) Synthesis, biological evaluation, and molecular modeling of donepezil and N-[(5-(benzyloxy)-1-methyl-1H-indol-2-yl)methyl]-N-methylprop-2-yn-1-amine hybrids as new multipotent cholinesterase/monoamine oxidase inhibitors for the treatment of Alzheimer. *J Med Chem* 54:8251–8270. <https://doi.org/10.1021/jm200853t>
53. Huang L, Lu C, Sun Y, Mao F, Luo Z, Su T, Jiang H, Shan W, Li X (2012) Multitarget-directed benzylideneindanone derivatives: anti- β -amyloid ($A\beta$) aggregation, antioxidant, metal chelation, and monoamine oxidase B (MAO-B) inhibition properties against Alzheimer's disease. *J Med Chem* 55:8483–8492. <https://doi.org/10.1021/jm300978h>
54. De Ferrari GV, Canales M, Shin I, Weiner LM, Silman I, Inestrosa NC (2001) A structural motif of acetylcholinesterase that promotes amyloid beta-peptide fibril formation. *Biochemistry* 40:10447–10457
55. Guzior N, Bajda M, Rakoczy J, Brus B, Gobec S, Malawska B (2015) Isoindoline-1,3-dione derivatives targeting cholinesterases: design, synthesis and biological evaluation of potential anti-Alzheimer's agents. *Bioorg Med Chem* 23:1629–1637. <https://doi.org/10.1016/j.bmc.2015.01.045>
56. Więckowska A, Więckowski K, Bajda M, Brus B, Sałat K, Czerwińska P, Gobec S, Filipek B, Malawska B (2015) Synthesis of new N-benzylpiperidine derivatives as cholinesterase inhibitors with β -amyloid anti-aggregation properties and beneficial effects on memory in vivo. *Bioorg Med Chem* 23:2445–2457. <https://doi.org/10.1016/j.bmc.2015.03.051>
57. Wang L, Esteban G, Ojima M, Bautista-Aguilera OM, Inokuchi T, Moraleda I, Iriepa I, Samadi A, Youdim MBH, Romero A, Soriano E, Herrero R, Fernández Fernández AP, Ricardo-Martínez-Murillo, Marco-Contelles J, Unzeta M (2014) Donepezil + propargylamine + 8-hydroxyquinoline hybrids as new multifunctional metal-chelators, ChE and MAO inhibitors for the potential treatment of Alzheimer's disease. *Eur J Med Chem* 80:543–561. <https://doi.org/10.1016/j.ejmech.2014.04.078>
58. Bautista-Aguilera OM, Esteban G, Bolea I, Nikolic K, Agbaba D, Moraleda I, Iriepa I, Samadi A, Soriano E, Unzeta M, Marco-Contelles J (2014) Design, synthesis, pharmacological evaluation, QSAR analysis, molecular modeling and ADMET of novel donepezil–indolyl hybrids as multipotent cholinesterase/monoamine oxidase inhibitors for the potential treatment of Alzheimer's disease. *Eur J Med Chem* 75:82–95. <https://doi.org/10.1016/j.ejmech.2013.12.028>
59. Bautista-Aguilera OM, Samadi A, Chioua M, Nikolic K, Filipic S, Agbaba D, Soriano E, de Andrés L, Rodríguez-Franco MI, Alcaro S, Ramsay RR, Ortuso F, Yañez M, Marco-Contelles J (2014) N-methyl-N-((1-methyl-5-(3-(1-(2-methylbenzyl)piperidin-4-yl)propoxy)-1H-indol-2-yl)methyl)prop-2-yn-1-amine, a new cholinesterase and monoamine oxidase dual inhibitor. *J Med Chem* 57:10455–10463. <https://doi.org/10.1021/jm501501a>
60. Yerdelen KO, Koca M, Anil B, Sevindik H, Kasap Z, Halici Z, Turkaydin K, Gunesacar G (2015) Synthesis of donepezil-based multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem Lett* 25:5576–5582. <https://doi.org/10.1016/j.bmcl.2015.10.051>
61. Więckowska A, Kołaczowski M, Bucki A, Godyń J, Marcinkowska M, Więckowski K, Zareba P, Siwek A, Kazek G, Gluch-Lutwin M, Mierzejewski P, Bienkowski P, Sienkiewicz-Jarosz H, Knez D, Wichur T, Gobec S, Malawska B (2016) Novel multitarget-directed ligands for Alzheimer's disease: combining cholinesterase inhibitors and 5-HT₆ receptor antagonists. Design, synthesis and biological evaluation. *Eur J Med Chem* 124:63–81. <https://doi.org/10.1016/j.ejmech.2016.08.016>
62. Castañeda-Arriaga R, Alvarez-Idaboy JR (2014) Lipoic acid and dihydrolipoic acid. A comprehensive theoretical study of their antioxidant activity supported by available experimental kinetic data. *J Chem Inf Model* 54:1642–1652. <https://doi.org/10.1021/ci500213p>
63. Rosini M, Simoni E, Bartolini M, Tarozzi A, Matera R, Milelli A, Hrelia P, Andrisano V, Bolognesi ML, Melchiorre C (2011)

- Exploiting the lipoic acid structure in the search for novel multitarget ligands against Alzheimer's disease. *Eur J Med Chem* 46:5435–5442. <https://doi.org/10.1016/j.ejmech.2011.09.001>
64. Prezzavento O, Arena E, Parenti C, Pasquinucci L, Aricò G, Scoto GM, Grancara S, Toninello A, Ronsisvalle S (2013) Design and synthesis of new bifunctional sigma-1 selective ligands with antioxidant activity. *J Med Chem* 56:2447–2455. <https://doi.org/10.1021/jm3017893>
 65. Fava A, Pirritano D, Plastino M, Cristiano D, Puccio G, Colica C, Ermio C, De Bartolo M, Mauro G, Bosco D (2013) The effect of lipoic acid therapy on cognitive functioning in patients with Alzheimer's disease. *J Neurodegener Dis* 2013:7. <https://doi.org/10.1155/2013/454253>
 66. Estrada M, Pérez C, Soriano E, Laurini E, Romano M, Pricl S, Morales-García JA, Pérez-Castillo A, Rodríguez-Franco MI (2016) New neurogenic lipoic-based hybrids as innovative Alzheimer's drugs with σ -1 agonism and β -secretase inhibition. *Future Med Chem* 8:1191–1207. <https://doi.org/10.4155/fmc-2016-0036>
 67. Dias KST, de Paula CT, dos Santos T, Souza INO, Boni MS, Guimarães MJR, da Silva FMR, Castro NG, Neves GA, Veloso CC, Coelho MM, de Melo ISF, Giusti FCV, Giusti-Paiva A, da Silva ML, Dardenne LE, Guedes IA, Pruccoli L, Morrioni F, Tarozzi A, Viegas C (2017) Design, synthesis and evaluation of novel feruloyl-donepezil hybrids as potential multitarget drugs for the treatment of Alzheimer's disease. *Eur J Med Chem* 130:440–457. <https://doi.org/10.1016/j.ejmech.2017.02.043>
 68. Xu W, Wang X-B, Wang Z-M, Wu J-J, Li F, Wang J, Kong L-Y (2016) Synthesis and evaluation of donepezil–ferulic acid hybrids as multi-target-directed ligands against Alzheimer's disease. *Med Chem Commun* 7:990–998. <https://doi.org/10.1039/C6MD00053C>
 69. Wang J, Wang Z-M, Li X-M, Li F, Wu J-J, Kong L-Y, Wang X-B (2016) Synthesis and evaluation of multi-target-directed ligands for the treatment of Alzheimer's disease based on the fusion of donepezil and melatonin. *Bioorg Med Chem* 24:4324–4338. <https://doi.org/10.1016/j.bmc.2016.07.025>
 70. Qiang X, Sang Z, Yuan W, Li Y, Liu Q, Bai P, Shi Y, Ang W, Tan Z, Deng Y (2014) Design, synthesis and evaluation of genistein-O-alkylbenzylamines as potential multifunctional agents for the treatment of Alzheimer's disease. *Eur J Med Chem* 76:314–331. <https://doi.org/10.1016/j.ejmech.2014.02.045>
 71. Farina R, Pisani L, Catto M, Nicolotti O, Gadaleta D, Denora N, Soto-Otero R, Mendez-Alvarez E, Passos CS, Muncipinto G, Altomare CD, Nurisso A, Carrupt P-A, Carotti A (2015) Structure-based design and optimization of multitarget-directed 2 H-Chromen-2-one derivatives as potent inhibitors of monoamine oxidase B and cholinesterases. *J Med Chem* 58:5561–5578. <https://doi.org/10.1021/acs.jmedchem.5b00599>
 72. Claeysen S, Bockaert J, Giannoni P (2015) Serotonin: a new hope in Alzheimer's disease? *ACS Chem Neurosci* 6:940–943. <https://doi.org/10.1021/acschemneuro.5b00135>
 73. Van der Schyf CJ (2016) Psychotropic drug development strategies that target neuropsychiatric etiologies in Alzheimer's and Parkinson's diseases. *Drug Dev Res* 77:458–468. <https://doi.org/10.1002/ddr.21368>
 74. Lecoutey C, Hedou D, Freret T, Giannoni P, Gaven F, Since M, Bouet V, Ballandonne C, Corvaisier S, Malzert Fréon A, Mignani S, Cresteil T, Boulouard M, Claeysen S, Rochais C, Dallemagne P (2014) Design of donecopride, a dual serotonin subtype 4 receptor agonist/acetylcholinesterase inhibitor with potential interest for Alzheimer's disease treatment. *Proc Natl Acad Sci U S A* 111:E3825–E3830. <https://doi.org/10.1073/pnas.1410315111>
 75. Rochais C, Lecoutey C, Gaven F, Giannoni P, Hamidouche K, Hedou D, Dubost E, Genest D, Yahiaoui S, Freret T, Bouet V, Dauphin F, Sopkova de Oliveira Santos J, Ballandonne C, Corvaisier S, Malzert-Fréon A, Legay R, Boulouard M, Claeysen S, Dallemagne P (2015) Novel multitarget-directed ligands (MTDLs) with acetylcholinesterase (AChE) inhibitory and serotonergic subtype 4 receptor (5-HT₄R) agonist activities as potential agents against Alzheimer's disease: the design of donecopride. *J Med Chem* 58:3172–3187. <https://doi.org/10.1021/acs.jmedchem.5b00115>
 76. Mishra CB, Kumari S, Manral A, Prakash A, Saini V, Lynn AM, Tiwari M (2017) Design, synthesis, in-silico and biological evaluation of novel donepezil derivatives as multi-target-directed ligands for the treatment of

- Alzheimer's disease. *Eur J Med Chem* 125:736–750. <https://doi.org/10.1016/j.ejmech.2016.09.057>
77. Pereira JD, Caricati-Neto A, Miranda-Ferreira R, Smaili SS, Godinho RO, Rios CDL, León R, Villarroya M, Samadi A, Marco-Contelles J, Jurkiewicz NH, Garcia AG, Jurkiewicz A (2011) Effects of novel tacripyrines ITH12117 and ITH12118 on rat vas deferens contractions, calcium transients and cholinesterase activity. *Eur J Pharmacol* 660:411–419. <https://doi.org/10.1016/j.ejphar.2011.03.042>
 78. Tumiatti V, Minarini A, Bolognesi ML, Milelli A, Rosini M, Melchiorre C (2010) Tacrine derivatives and Alzheimer's disease. *Curr Med Chem* 17:1825–1838
 79. Marco-Contelles J, León R, López MG, García AG, Villarroya M (2006) Synthesis and biological evaluation of new 4H-pyrano [2,3-b]quinoline derivatives that block acetylcholinesterase and cell calcium signals, and cause neuroprotection against calcium overload and free radicals. *Eur J Med Chem* 41:1464–1469. <https://doi.org/10.1016/j.ejmech.2006.06.016>
 80. Minarini A, Milelli A, Tumiatti V, Rosini M, Simoni E, Bolognesi ML, Andrisano V, Bartolini M, Motori E, Angeloni C, Hrelia S (2012) Cystamine-tacrine dimer: a new multi-target-directed ligand as potential therapeutic agent for Alzheimer's disease treatment. *Neuropharmacology* 62:997–1003. <https://doi.org/10.1016/j.neuropharm.2011.10.007>
 81. Fang L, Kraus B, Lehmann J, Heilmann J, Zhang Y, Decker M (2008) Design and synthesis of tacrine-ferulic acid hybrids as multi-potent anti-Alzheimer drug candidates. *Bioorg Med Chem Lett* 18:2905–2909. <https://doi.org/10.1016/j.bmcl.2008.03.073>
 82. Marco-contelles J, Leo R, Ri DL, Samadi A, Bartolini M, Andrisano V, Huertas O, Barril X, Luque FJ, Rodri MI, Lo MG, Garci AG, Villarroya M (2009) Tacripyrines, the first tacrine-dihydropyridine hybrids, as multitarget-directed ligands for the treatment of Alzheimer's disease. *J Med Chem* 52:2724–2732. <https://doi.org/10.1021/jm801292b>
 83. Chao X, He X, Yang Y, Zhou X, Jin M, Liu S, Cheng Z, Liu P, Wang Y, Yu J, Tan Y, Huang Y, Qin J, Rapposelli S, Pi R (2012) Design, synthesis and pharmacological evaluation of novel tacrine-caffeic acid hybrids as multi-targeted compounds against Alzheimer's disease. *Bioorg Med Chem Lett* 22:6498–6502. <https://doi.org/10.1016/j.bmcl.2012.08.036>
 84. Wang Y, Wang F, Yu JP, Jiang FC, Guan XL, Wang CM, Li L, Cao H, Li MX, Chen JG (2012) Novel multipotent phenylthiazole-tacrine hybrids for the inhibition of cholinesterase activity, β -amyloid aggregation and Ca^{2+} overload. *Bioorg Med Chem* 20:6513–6522. <https://doi.org/10.1016/j.bmc.2012.08.040>
 85. Lan J-S, Xie S-S, Li S-Y, Pan L-F, Wang X-B, Kong L-Y (2014) Design, synthesis and evaluation of novel tacrine-(β -carboline) hybrids as multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem* 22:6089–6104. <https://doi.org/10.1016/j.bmc.2014.08.035>
 86. Mourad Chioua J, Pérez-Peña N, García-Font I, Moraleda II, Elena Soriano J (2015) Pyranopyrazolotacrine as nonneurotoxic, $\text{A}\beta$ -anti-aggregating and neuroprotective agents for Alzheimer's disease. *Future Med Chem* 7:845–855. <https://doi.org/10.4155/fmc.15.35>
 87. Fu Y, Mu Y, Lei H, Wang P, Li X, Leng Q, Han L, Qu X, Wang Z, Huang X (2016) Design, synthesis and evaluation of novel tacrine-ferulic acid hybrids as multifunctional drug candidates against Alzheimer's disease. *Molecules* 21:1338. <https://doi.org/10.3390/molecules21101338>
 88. Benchekroun M, Romero A, Egea J, León R, Michalska P, Buendía I, Jimeno ML, Jun D, Janockova J, Sepsóva V, Soukup O, Bautista-Aguilera OM, Refouvet B, Ouari O, Marco-Contelles J, Ismaili L (2016) The antioxidant additive approach for Alzheimer's disease therapy: new ferulic (lipoic) acid plus melatonin modified tacrine as cholinesterases inhibitors, direct antioxidants, and nuclear factor (erythroid-derived 2)-like 2 activators. *J Med Chem* 59:9967–9973. <https://doi.org/10.1021/acs.jmedchem.6b01178>
 89. García-Font N, Hayour H, Belfaitah A, Pedraz J, Moraleda I, Iriepa I, Bouraiou A, Chioua M, Marco-Contelles J, Oset-Gasque MJ (2016) Potent anticholinesterase and neuroprotective pyranotacrine as inhibitors of beta-amyloid aggregation, oxidative stress and tau-phosphorylation for Alzheimer's disease. *Eur J Med Chem* 118:178–192. <https://doi.org/10.1016/j.ejmech.2016.04.023>
 90. Wang X-Q, Xia C-L, Chen S-B, Tan J-H, Ou T-M, Huang S-L, Li D, Gu L-Q, Huang Z-S (2015) Design, synthesis, and biological

- evaluation of 2-arylethenylquinoline derivatives as multifunctional agents for the treatment of Alzheimer's disease. *Eur J Med Chem* 89:349–361. <https://doi.org/10.1016/j.ejmech.2014.10.018>
91. Anand P, Singh B, Singh N (2012) A review on coumarins as acetylcholinesterase inhibitors for Alzheimer's disease. *Bioorg Med Chem* 20:1175–1180. <https://doi.org/10.1016/j.bmc.2011.12.042>
92. Patil PO, Bari SB, Firke SD, Deshmukh PK, Donda ST, Patil DA (2013) A comprehensive review on synthesis and designing aspects of coumarin derivatives as monoamine oxidase inhibitors for depression and Alzheimer's disease. *Bioorg Med Chem* 21:2434–2450. <https://doi.org/10.1016/j.bmc.2013.02.017>
93. Chimenti F, Secci D, Bolasco A, Chimenti P, Bizzarri B, Granese A, Carradori S, Yáñez M, Orallo F, Ortuso F, Alcaro S (2009) Synthesis, molecular modeling, and selective inhibitory activity against human monoamine oxidases of 3-carboxamido-7-substituted coumarins. *J Med Chem* 52:1935–1942. <https://doi.org/10.1021/jm801496u>
94. Catto M, Nicolotti O, Leonetti F, Carotti A, Favia AD, Soto-Otero R, Méndez-Álvarez E, Carotti A (2006) Structural insights into monoamine oxidase inhibitory potency and selectivity of 7-substituted coumarins from ligand- and target-based approaches. *J Med Chem* 49:4912–4925. <https://doi.org/10.1021/jm060183l>
95. Xie SS, Wang X, Jiang N, Yu W, Wang KDG, Lan JS, Li ZR, Kong LY (2015) Multi-target tacrine-coumarin hybrids: cholinesterase and monoamine oxidase B inhibition properties against Alzheimer's disease. *Eur J Med Chem* 95:153–165. <https://doi.org/10.1016/j.ejmech.2015.03.040>
96. Boulebd H, Ismaili L, Bartolini M, Bouraiou A, Andrisano V, Martin H, Bonet A, Moraleda I, Iriepa I, Chioua M, Belfaitah A, Marco-Contelles J (2016) Imidazopyranotacrines as non-hepatotoxic, selective acetylcholinesterase inhibitors, and antioxidant agents for Alzheimer's disease therapy. *Molecules* 21:400. <https://doi.org/10.3390/molecules21040400>
97. Spilovska K, Korabecny J, Horova A, Musilek K, Nepovimova E, Drtinova L, Gazova Z, Siposova K, Dolezal R, Jun D, Kuca K (2015) Design, synthesis and in vitro testing of 7-methoxytacrine-amantadine analogues: a novel cholinesterase inhibitors for the treatment of Alzheimer's disease. *Med Chem Res* 24:2645–2655. <https://doi.org/10.1007/s00044-015-1316-x>
98. Martins C, Carreiras MC, León R, De Los Ríos C, Bartolini M, Andrisano V, Iriepa I, Moraleda I, Gálvez E, García M, Egea J, Samadi A, Chioua M, Marco-Contelles J (2011) Synthesis and biological assessment of diversely substituted furo[2,3-b]quinolin-4-amine and pyrrolo[2,3-b]quinolin-4-amine derivatives, as novel tacrine analogues. *Eur J Med Chem* 46:6119–6130. <https://doi.org/10.1016/j.ejmech.2011.09.038>
99. Chen Y, Lin H, Zhu J, Gu K, Li Q, He S, Lu X, Tan R, Pei Y, Wu L, Bian Y, Sun H (2017) Design, synthesis, in vitro and in vivo evaluation of tacrine-cinnamic acid hybrids as multi-target acetyl- and butyrylcholinesterase inhibitors against Alzheimer's disease. *RSC Adv* 7:33851–33867. <https://doi.org/10.1039/C7RA04385F>
100. Jeřábek J, Uliassi E, Guidotti L, Korábečný J, Soukup O, Sepsova V, Hrabínova M, Kuča K, Bartolini M, Peña-Altamira LE, Petralla S, Monti B, Roberti M, Bolognesi ML (2017) Tacrine-resveratrol fused hybrids as multi-target-directed ligands against Alzheimer's disease. *Eur J Med Chem* 127:250–262. <https://doi.org/10.1016/j.ejmech.2016.12.048>
101. Teponnou GAK, Joubert J, Malan SF (2017) Tacrine, trolox and tryptoline as lead compounds for the design and synthesis of multi-target agents for Alzheimer's disease therapy. *Open Med Chem J* 11:24–37. <https://doi.org/10.2174/1874104501711010024>
102. Viegas C, Bolzani VDS, Barreiro EJ, Fraga CAM (2005) New anti-Alzheimer drugs from biodiversity: the role of the natural acetylcholinesterase inhibitors. *Mini Rev Med Chem* 5:915–926. <https://doi.org/10.2174/138955705774329546>
103. Piazzi L, Rampa A, Bisi A, Gobbi S, Belluti F, Cavalli A, Bartolini M, Andrisano V, Valenti P, Recanatini M (2003) 3-(4-[[Benzyl(methyl)amino]methyl]phenyl)-6,7-dimethoxy-2 H-2-chromenone (AP2238) inhibits both acetylcholinesterase and acetylcholinesterase-induced β -amyloid aggregation: a dual function lead for Alzheimer's disease therapy. *J Med Chem* 46:2279–2282. <https://doi.org/10.1021/jm0340602>
104. Piazzi L, Cavalli A, Belluti F, Bisi A, Gobbi S, Rizzo S, Bartolini M, Andrisano V,

- Recanatini M, Rampa A (2007) Extensive SAR and computational studies of 3-[4-[(benzylmethylamino)methyl]phenyl]-6,7-dimethoxy-2 H-2-chromenone (AP2238) derivatives. *J Med Chem* 50:4250–4254. <https://doi.org/10.1021/jm070100g>
105. Rizzo S, Bartolini M, Ceccarini L, Piazzi L, Gobbi S, Cavalli A, Recanatini M, Andrisano V, Rampa A (2010) Targeting Alzheimer's disease: novel indanone hybrids bearing a pharmacophoric fragment of AP2238. *Bioorg Med Chem* 18:1749–1760. <https://doi.org/10.1016/j.bmc.2010.01.071>
 106. López-Iglesias B, Pérez C, Morales-García JA, Alonso-Gil S, Pérez-Castillo A, Romero A, López MG, Villarroya M, Conde S, Rodríguez-Franco MI (2014) New melatonin-N, N-dibenzyl(N-methyl)amine hybrids: potent neurogenic agents with antioxidant, cholinergic, and neuroprotective properties as innovative drugs for Alzheimer's disease. *J Med Chem* 57:3773–3785. <https://doi.org/10.1021/jm5000613>
 107. Luo XT, Wang CM, Liu Y, Huang ZG (2015) New multifunctional melatonin-derived benzylpyridinium bromides with potent cholinergic, antioxidant, and neuroprotective properties as innovative drugs for Alzheimer's disease. *Eur J Med Chem* 103:302–311. <https://doi.org/10.1016/j.ejmech.2015.08.052>
 108. Lu C, Guo Y, Yan J, Luo Z, Luo H, Yan M, Huang L, Li X (2013) Design, synthesis, and evaluation of multitarget-directed resveratrol derivatives for the treatment of Alzheimer's disease. *J Med Chem* 56:5843–5859. <https://doi.org/10.1021/jm400567s>
 109. Pan L-F, Wang X-B, Xie S-S, Li S-Y, Kong L-Y (2014) Multitarget-directed resveratrol derivatives: anti-cholinesterases, anti- β -amyloid aggregation and monoamine oxidase inhibition properties against Alzheimer's disease. *Med Chem Commun* 5:609. <https://doi.org/10.1039/c3md00376k>
 110. Pan W, Hu K, Bai P, Yu L, Ma Q, Li T, Zhang X, Chen C, Peng K, Liu W, Sang Z (2016) Design, synthesis and evaluation of novel ferulic acid-memoquin hybrids as potential multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem Lett* 26:2539–2543. <https://doi.org/10.1016/j.bmcl.2016.03.086>
 111. Pérez-Areales FJ, Di Pietro O, Espargaró A, Vallverdú-Queralt A, Galdeano C, Ragusa IM, Viayna E, Guillou C, Clos MV, Pérez B, Sabaté R, Lamuela-Raventós RM, Luque FJ, Muñoz-Torrero D (2014) Shogaol-huprine hybrids: dual antioxidant and anticholinesterase agents with β -amyloid and tau anti-aggregating properties. *Bioorg Med Chem* 22:5298–5307. <https://doi.org/10.1016/j.bmc.2014.07.053>
 112. Viayna E, Sola I, Bartolini M, De Simone A, Tapia-Rojas C, Serrano FG, Sabaté R, Juárez-Jiménez J, Pérez B, Luque FJ, Andrisano V, Clos MV, Inestrosa NC, Muñoz-Torrero D (2014) Synthesis and multitarget biological profiling of a novel family of rhein derivatives as disease-modifying anti-Alzheimer agents. *J Med Chem* 57:2549–2567. <https://doi.org/10.1021/jm401824w>
 113. Saura J, Luque JM, Cesura AM, Huber G, Lgffler J (1994) Increased monoamine oxidase B activity in plaque-associated astrocytes of Alzheimer brains revealed by quantitative enzyme radioautography. *Neuroscience* 62:15–30
 114. Opazo C, Huang X, Cherny RA, Moir RD, Roher AE, White AR, Cappai R, Masters CL, Tanzi RE, Inestrosa NC, Bush AI (2002) Metalloenzyme-like activity of Alzheimer's disease β -amyloid: Cu-dependent catalytic conversion of dopamine, cholesterol, and biological reducing agents to neurotoxic H₂O₂. *J Biol Chem* 277:40302–40308. <https://doi.org/10.1074/jbc.M206428200>
 115. Schieber M, Chandel NS (2014) ROS function in redox signaling and oxidative stress. *Curr Biol* 24:R453–R462. <https://doi.org/10.1016/j.cub.2014.03.034>
 116. Huang M, Xie S-S, Jiang N, Lan J-S, Kong L-Y, Wang X-B (2015) Multifunctional coumarin derivatives: monoamine oxidase B (MAO-B) inhibition, anti- β -amyloid (A β) aggregation and metal chelation properties against Alzheimer's disease. *Bioorg Med Chem Lett* 25:508–513. <https://doi.org/10.1016/j.bmcl.2014.12.034>
 117. Li Y, Qiang X, Luo L, Li Y, Xiao G, Tan Z, Deng Y (2016) Synthesis and evaluation of 4-hydroxyl aurone derivatives as multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem* 24:2342–2351. <https://doi.org/10.1016/j.bmc.2016.04.012>
 118. Li Y, Qiang X, Li Y, Yang X, Luo L, Xiao G, Cao Z, Tan Z, Deng Y (2016) Pterostilbene-O-acetamidoalkylbenzylamines derivatives as novel dual inhibitors of cholinesterase with anti- β -amyloid aggregation and antioxidant properties for the treatment of Alzheimer's disease. *Bioorg Med Chem Lett*

- 26:2035–2039. <https://doi.org/10.1016/j.bmc.2016.02.079>
119. Liu Q, Qiang X, Li Y, Sang Z, Li Y, Tan Z, Deng Y (2015) Design, synthesis and evaluation of chromone-2-carboxamido-alkylbenzylamines as multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem* 23:911–923. <https://doi.org/10.1016/j.bmc.2015.01.042>
 120. Shaik JB, Palaka BK, Penumala M, Kotapati KV, Devineni SR, Eadlapalli S, Darla MM, Ampasala DR, Vadde R, Amooru GD (2016) Synthesis, pharmacological assessment, molecular modeling and in silico studies of fused tricyclic coumarin derivatives as a new family of multifunctional anti-Alzheimer agents. *Eur J Med Chem* 107:219–232. <https://doi.org/10.1016/j.ejmech.2015.10.046>
 121. Lan J-S, Ding Y, Liu Y, Kang P, Hou J-W, Zhang X-Y, Xie S-S, Zhang T (2017) Design, synthesis and biological evaluation of novel donepezil-coumarin hybrids as multi-target agents for the treatment of Alzheimer's disease. *Eur J Med Chem* 139:48–59. <https://doi.org/10.1016/j.ejmech.2017.07.055>
 122. Puksasook T, Kimura S, Tadtong S, Jiaranaikulwanitch J, Pratuangdejkul J, Kitphati W, Suwanborirux K, Saito N, Nukoolkarn V (2017) Semisynthesis and biological evaluation of prenylated resveratrol derivatives as multi-targeted agents for Alzheimer's disease. *J Nat Med* 71:1–18. <https://doi.org/10.1007/s11418-017-1097-2>
 123. Yang H-L, Cai P, Liu Q-H, Yang X-L, Fang S-Q, Tang Y-W, Wang C, Wang X-B, Kong L-Y (2017) Design, synthesis, and evaluation of salicyladimine derivatives as multitarget-directed ligands against Alzheimer's disease. *Bioorg Med Chem* 25 (21):5917–5928. <https://doi.org/10.1016/j.bmc.2017.08.048>
 124. Prati F, De Simone A, Bisignano P, Armirotti A, Summa M, Pizzirani D, Scarpelli R, Perez DI, Andrisano V, Perez-Castillo A, Monti B, Massenzio F, Polito L, Racchi M, Favia AD, Bottegoni G, Martinez A, Bolognesi ML, Cavalli A (2015) Multitarget drug discovery for Alzheimer's disease: triazinones as BACE-1 and GSK-3 β inhibitors. *Angew Chem Int Ed* 54:1578–1582. <https://doi.org/10.1002/anie.201410456>
 125. Kumar J, Meena P, Singh A, Jameel E, Maqbool M, Mobashir M, Shandilya A, Tiwari M, Hoda N, Jayaram B (2016) Synthesis and screening of triazolopyrimidine scaffold as multi-functional agents for Alzheimer's disease therapies. *Eur J Med Chem* 119:260–277. <https://doi.org/10.1016/j.ejmech.2016.04.053>
 126. Liao S, Deng H, Huang S, Yang J, Wang S, Yin B, Zheng T, Zhang D, Liu J, Gao G, Ma J, Deng Z (2015) Design, synthesis and evaluation of novel 5,6,7-trimethoxyflavone–6-chlorotacrine hybrids as potential multifunctional agents for the treatment of Alzheimer's disease. *Bioorg Med Chem Lett* 25:1541–1545. <https://doi.org/10.1016/j.bmc.2015.02.015>
 127. Sang Z, Li Y, Qiang X, Xiao G, Liu Q, Tan Z, Deng Y (2015) Multifunctional scutellarin–rivastigmine hybrids with cholinergic, antioxidant, biometal chelating and neuroprotective properties for the treatment of Alzheimer's disease. *Bioorg Med Chem* 23:668–680. <https://doi.org/10.1016/j.bmc.2015.01.005>
 128. González-Naranjo P, Pérez-Macias N, Campillo NE, Pérez C, Arán VJ, Girón R, Sánchez-Robles E, Martín MI, Gómez-Cañas M, García-Arencibia M, Fernández-Ruiz J, Páez JA (2014) Cannabinoid agonists showing BuChE inhibition as potential therapeutic agents for Alzheimer's disease. *Eur J Med Chem* 73:56–72. <https://doi.org/10.1016/j.ejmech.2013.11.026>
 129. Greig NH, Lahiri DK, Sambamurti K (2002) Butyrylcholinesterase: an important new target in Alzheimer's disease therapy. *Int Psychogeriatr* 14:77–91. <https://doi.org/10.1017/S1041610203008676>
 130. Mesulam M, Guillozet A, Shaw P, Quinn B (2002) Widely spread butyrylcholinesterase can hydrolyze acetylcholine in the normal and Alzheimer brain. *Neurobiol Dis* 9:88–93. <https://doi.org/10.1006/nbdi.2001.0462>
 131. Digiacomio M, Chen Z, Wang S, Lapucci A, Macchia M, Yang X, Chu J, Han Y, Pi R, Rapposelli S (2015) Synthesis and pharmacological evaluation of multifunctional tacrine derivatives against several disease pathways of AD. *Bioorg Med Chem Lett* 25:807–810. <https://doi.org/10.1016/j.bmc.2014.12.084>
 132. Estrada M, Herrera-Arozamena C, Pérez C, Viña D, Romero A, Morales-García JA, Pérez-Castillo A, Rodríguez-Franco MI (2016) New cinnamic – N-benzylpiperidine and cinnamic – N,N-dibenzyl(N-methyl) amine hybrids as Alzheimer-directed multitarget drugs with antioxidant, cholinergic,

- neuroprotective and neurogenic properties. *Eur J Med Chem* 121:376–386. <https://doi.org/10.1016/j.ejmech.2016.05.055>
133. Chen Z, Digiacocono M, Tu Y, Gu Q, Wang S, Yang X, Chu J, Chen Q, Han Y, Chen J, Nesi G, Sestito S, Macchia M, Rapposelli S, Pi R (2017) Discovery of novel rivastigmine-hydroxycinnamic acid hybrids as multi-targeted agents for Alzheimer's disease. *Eur J Med Chem* 125:784–792. <https://doi.org/10.1016/j.ejmech.2016.09.052>
 134. Wang Z, Li W, Wang Y, Li XX, Huang L, Li XX (2016) Design, synthesis and evaluation of clioquinol–ebselen hybrids as multi-target-directed ligands against Alzheimer's disease. *RSC Adv* 6:7139–7158. <https://doi.org/10.1039/C5RA26797H>
 135. Wang Z-M, Xie S-S, Li X-M, Wu J-J, Wang X-B, Kong L-Y (2015) Multifunctional 3-Schiff base-4-hydroxycoumarin derivatives with monoamine oxidase inhibition, anti- β -amyloid aggregation, metal chelation, anti-oxidant and neuroprotection properties against Alzheimer's disease. *RSC Adv* 5:70395–70409. <https://doi.org/10.1039/C5RA13594J>
 136. Wang Z, Wang Y, Wang B, Li W, Huang L, Li X (2015) Design, synthesis, and evaluation of orally available clioquinol–moracin M hybrids as multitarget-directed ligands for cognitive improvement in a rat model of neurodegeneration in Alzheimer's disease. *J Med Chem* 58:8616–8637. <https://doi.org/10.1021/acs.jmedchem.5b01222>
 137. Maqbool M, Manral A, Jameel E, Kumar J, Saini V, Shandilya A, Tiwari M, Hoda N, Jayaram B (2016) Development of cyanopyridine–triazine hybrids as lead multitarget anti-Alzheimer agents. *Bioorg Med Chem* 24:2777–2788. <https://doi.org/10.1016/j.bmc.2016.04.041>
 138. Sheng R, Tang L, Jiang L, Hong L, Shi Y, Zhou N, Hu Y (2016) Novel 1-phenyl-3-hydroxy-4-pyridinone derivatives as multifunctional agents for the therapy of Alzheimer's disease. *ACS Chem Neurosci* 7:69–81. <https://doi.org/10.1021/acschemneuro.5b00224>
 139. Tang H, Wei Y, Zhang C, Ning F, Qiao W, Huang S, Ma L (2009) Synthesis, biological evaluation and molecular modeling of oxoisoaporphine and oxoaporphine derivatives as new dual inhibitors of acetylcholinesterase/butrylcholinesterase. *Eur J Med Chem* 44:2523–2532. <https://doi.org/10.1016/j.ejmech.2009.01.021>
 140. Wei S, Chen W, Qin J, Huangli Y, Wang L, Shen Y, Tang H (2016) Multitarget-directed oxoisoaporphine derivatives: anti-acetylcholinesterase, anti- β -amyloid aggregation and enhanced autophagy activity against Alzheimer's disease. *Bioorg Med Chem* 24:6031–6039. <https://doi.org/10.1016/j.bmc.2016.09.061>
 141. Ignasik M, Bajda M, Guzior N, Prinz M, Holzgrabe U, Malawska B (2012) Design, synthesis and evaluation of novel 2-(aminoalkyl)-isoindoline-1,3-dione derivatives as dual-binding site acetylcholinesterase inhibitors. *Arch Pharm (Weinheim)* 345:509–516. <https://doi.org/10.1002/ardp.201100423>
 142. Hebda M, Bajda M, Więckowska A, Szałaj N, Pasięka A, Panek D, Godyn J, Wichur T, Knez D, Gobec S, Malawska B (2016) Synthesis, molecular modelling and biological evaluation of novel heterodimeric, multiple ligands targeting cholinesterases and amyloid beta. *Molecules* 21:410. <https://doi.org/10.3390/molecules21040410>
 143. Li X, Wang H, Lu Z, Zheng X, Ni W, Zhu J, Fu Y, Lian F, Zhang N, Li J, Zhang H, Mao F (2016) Development of multifunctional pyrimidinylthiourea derivatives as potential anti-Alzheimer agents. *J Med Chem* 59:8326–8344. <https://doi.org/10.1021/acs.jmedchem.6b00636>
 144. De Simone A, Bartolini M, Baschieri A, Apperley KYP, Chen HH, Guardigni M, Montanari S, Kobrlova T, Soukup O, Valgimigli L, Andrisano V, Keillor JW, Basso M, Milelli A (2017) Hydroxy-substituted trans-cinnamoyl derivatives as multifunctional tools in the context of Alzheimer's disease. *Eur J Med Chem* 139:378–389. <https://doi.org/10.1016/j.ejmech.2017.07.058>
 145. Ozadali-Sari K, Tüylü Küçükılınç T, Ayazgok B, Balkan A, Unsal-Tan O (2017) Novel multi-targeted agents for Alzheimer's disease: synthesis, biological evaluation, and molecular modeling of novel 2-[4-(4-substitutedpiperazin-1-yl)phenyl]benzimidazoles. *Bioorg Chem* 72:208–214. <https://doi.org/10.1016/j.bioorg.2017.04.018>
 146. Cornec AS, Monti L, Kovalevich J, Makani V, James MJ, Vijayendran KG, Oukoloff K, Yao Y, Lee VMY, Trojanowski JQ, Smith AB, Brunden KR, Ballatore C (2017) Multitargeted imidazoles: potential therapeutic leads for Alzheimer's and other neurodegenerative diseases. *J Med Chem* 60:5120–5145. <https://doi.org/10.1021/acs.jmedchem.7b00475>
 147. Kořak U, Knez D, Brus B, Piřlar A, Kos J, Gobec S, Coquelle N, Colletier JP, Nachon F, Brazzolotto X (2017)

- N-Propargylpiperidines with naphthalene-2-carboxamide or naphthalene-2-sulfonamide moieties: potential multifunctional anti-Alzheimer's agents. *Bioorg Med Chem* 25:633–645. <https://doi.org/10.1016/j.bmc.2016.11.032>
148. Sang Z, Pan W, Wang K, Ma Q, Yu L, Yang Y, Bai P, Leng C, Xu Q, Li X, Tan Z, Liu W (2017) Design, synthesis and evaluation of novel ferulic acid-O-alkylamine derivatives as potential multifunctional agents for the treatment of Alzheimer's disease. *Eur J Med Chem* 130:379–392. <https://doi.org/10.1016/j.ejmech.2017.02.039>
149. Mohamed T, Rao PPN (2017) 2,4-Disubstituted quinazolines as amyloid- β aggregation inhibitors with dual cholinesterase inhibition and antioxidant properties: development and structure-activity relationship (SAR) studies. *Eur J Med Chem* 126:823–843. <https://doi.org/10.1016/j.ejmech.2016.12.005>
150. Sang Z, Pan W, Wang K, Ma Q, Yu L, Liu W (2017) Design, synthesis and biological evaluation of 3,4-dihydro-2(1H)-quinoline-O-alkylamine derivatives as new multipotent cholinesterase/monoamine oxidase inhibitors for the treatment of Alzheimer's disease. *Bioorg Med Chem* 25:3006–3017. <https://doi.org/10.1016/j.bmc.2017.03.070>

Part III

Case Studies



Virtual Screening for Dual Hsp90/B-Raf Inhibitors

Andrew Anighoro, Luca Pinzi, Giulio Rastelli, and Jürgen Bajorath

Abstract

In this chapter, we describe a computational strategy leading to the identification of the first dual inhibitors of Heat Shock Protein 90 (Hsp90) and protein kinase B-Raf. Both proteins are validated targets for anti-cancer drug discovery. There is strong evidence that the simultaneous inhibition of Hsp90 and B-Raf provides therapeutic benefits compared to exclusive engagement of one or the other target. Hence, we have been interested in searching for dual Hsp90/B-Raf inhibitors. Virtual compound screening led to the identification of two compounds with micromolar activity against both targets. The computational approach faced a number of challenges that needed to be overcome, as described herein.

Keywords B-Raf, Hsp90, Molecular docking, Multi-target inhibitors, Pharmacophores, Polypharmacology, Virtual screening

1 Introduction

A subset of drugs is known to elicit their therapeutic efficacy by binding to multiple biological targets, an effect commonly referred to as polypharmacology [1, 2]. Multi-target activities of compounds were mostly not achieved by design but uncovered retrospectively by experimental or computer-aided target deconvolution [3, 4]. Polypharmacology is also responsible for unwanted side effects of drugs. Whether or not multi-target activities are desirable depends on the specifics of therapeutic applications. Thus, it would be best to achieve a high level of control over drug specificity on one hand and polypharmacology on the other. This also explains why methods for the prospective design and discovery of multi-target ligands are of high interest [5, 6]. We discuss recent work leading to the identification of the first dual inhibitors of Hsp90 and B-Raf [7], which are popular targets in anti-cancer drug discovery. Hsp90 is a molecular chaperone that aids in the folding process of numerous client proteins. A schematic representation of the function of Hsp90 is presented in Fig. 1. Several members of the Hsp90 interactome are involved in signal transduction and regulatory mechanisms [8]. Therefore, Hsp90 plays a pivotal role in cell physiology. Among Hsp90 clients are many mutant or abundantly expressed oncoproteins that sustain the uncontrolled proliferation and apoptotic resistance of cancer cells [8, 9]. Although many Hsp90 inhibitors are known, none of

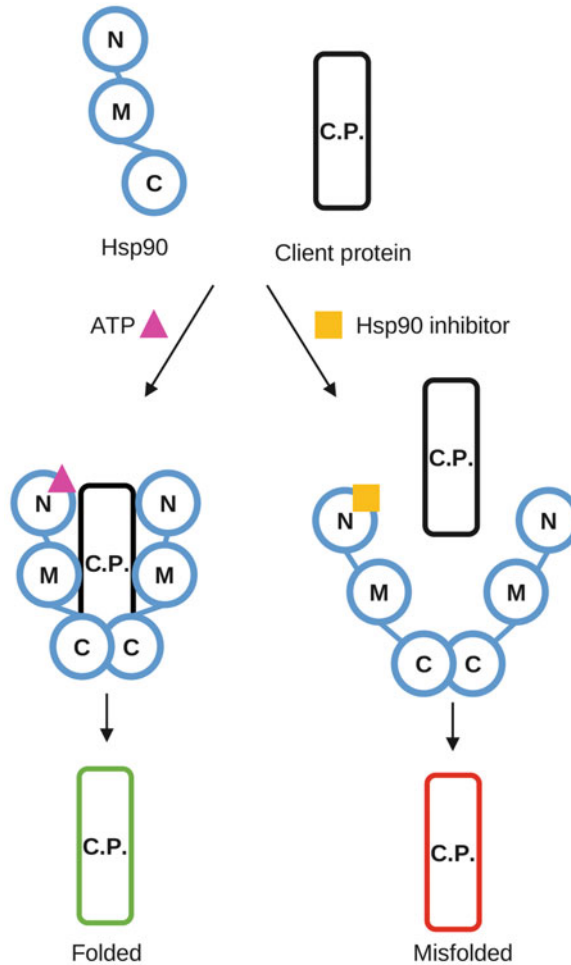


Fig. 1 A schematic representation of the Hsp90 function is shown, i.e., assisting the folding of a client protein (C. P.). Hsp90 contains an N-terminal domain (N) including the ATP binding site, a central or middle domain (M), and a C-terminal dimerization domain (C). In the presence of ATP, an Hsp90 dimer interacts with the client protein to facilitate correct folding (*green*). Hsp90 inhibitors prevent productive interaction of the chaperone with the client protein resulting in misfolding (*red*) followed by proteasome-dependent degradation

these compounds has so far been approved as a drug, which is mostly a consequence of non-optimal safety profiles or lack of efficacy [10, 11]. In cancer treatment, drug polypharmacology is typically desirable since multiple signalling pathways need to be controlled and side effects are more tolerated than in other therapeutic areas. Hence, Hsp90 engagement might be complemented by co-inhibition of other anti-cancer target(s) to further increase therapeutic relevance. A prime candidate is the Ser/Thr protein kinase B-Raf that is involved in the MAPK signaling cascade (Fig. 2). This signal transduction pathway contributes to the regulation of many essential cell processes such as cell proliferation and survival [12]. The naturally occurring B-Raf mutant V600E is critical for cancer development. This mutant is in a constitutively activated state that

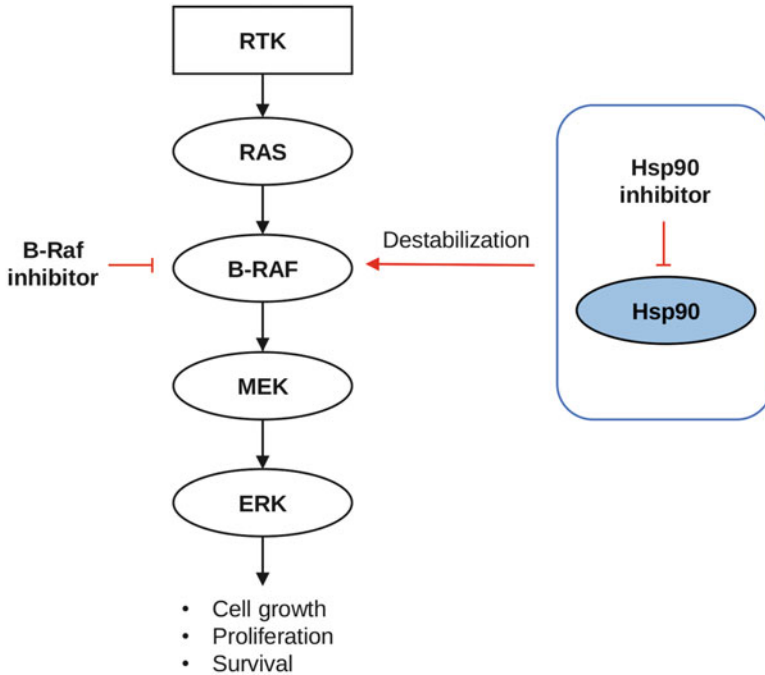


Fig. 2 A schematic representation of the MAPK signaling pathway is shown that originates from a receptor tyrosine kinase (RTK) and involves subsequent activation of the RAS, B-Raf, MEK, and ERK kinases. Functional relationships between B-Raf and Hsp90 inhibitors are also indicated. B-Raf inhibitors directly target the signaling pathway whereas Hsp90 inhibitors act indirectly by preventing correct folding of B-Raf

stimulates uncontrolled cell proliferation [13]. Such mutations may also cause resistance to drug treatment directed against the wild type [14]. Few B-Raf inhibitors are currently used in anti-cancer therapy, however, responses tend to be temporary, incomplete, and subject to drug resistance [15]. Hence, targeting B-Raf might also benefit from complementary target engagement. Potential benefits of simultaneously inhibiting Hsp90 and B-Raf are indicated by current clinical trials ([ClinicalTrials.gov](https://clinicaltrials.gov) identifiers: NCT02721459, NCT02097225, NCT01657591) that investigate combination therapies using Hsp90 and B-Raf inhibitors. Since B-Raf is a client of Hsp90, inhibition of the chaperone would contribute to the kinase degradation [16]. This relationship, schematically represented in Fig. 2, further increases the potential of Hsp90/B-Raf combination therapies. Representative Hsp90 and B-Raf inhibitors investigated for their potential application in combination therapies are shown in Fig. 3.

Given inherent complications of combination therapies such as coordinated dosing of different compounds or possible drug–drug interactions, it would be desirable to identify and develop dual Hsp90/B-Raf inhibitors. Therefore, we set out to search for such inhibitors focusing on a computational strategy [7].

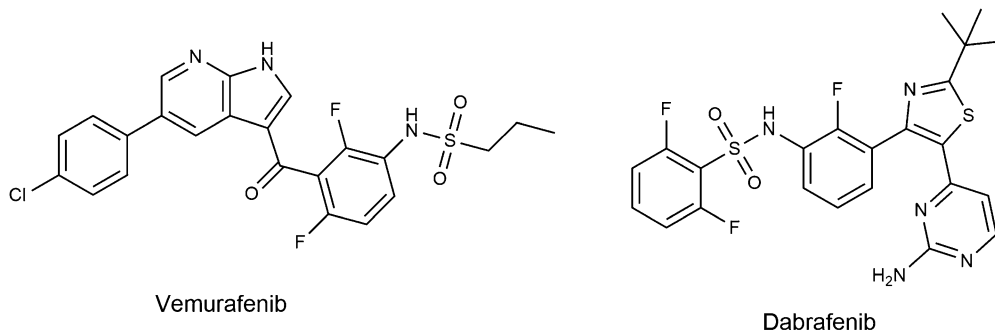
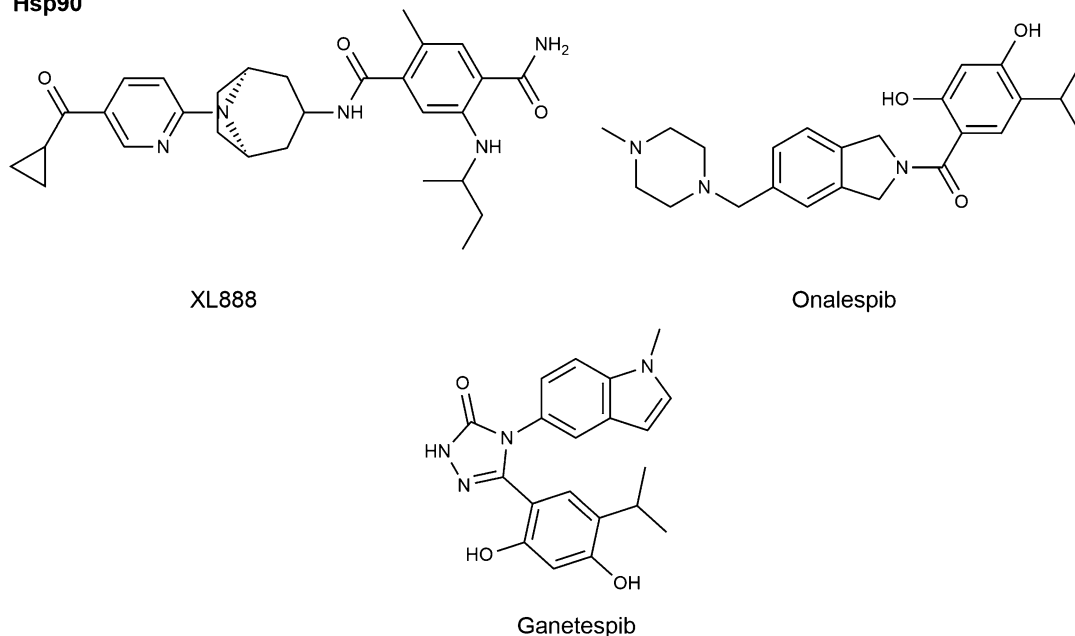
B-Raf**Hsp90**

Fig. 3 Structures of representative B-Raf and Hsp90 inhibitors are displayed that are investigated for their potential in combination therapy

Hsp90 and B-Raf have distinct folds and binding sites of different architecture, which represented an obstacle for identifying dual inhibitors [6]. However, there was at least remote binding site resemblance detectable between Hsp90 and few protein kinases [17] other than B-Raf. As a starting point of our analysis, we analyzed and compared available X-ray structures of Hsp90- and B-Raf-inhibitor complexes to explore potential similarities between target–ligand interactions. On the basis of our findings, a combined ligand- (LB) and structure-based (SB) virtual screening (VS) campaign was carried out in order to identify suitable candidate compounds.

Our computational strategy is schematically summarized in Fig. 4. Its application resulted in the identification of two Hsp90/B-Raf dual inhibitors sharing a thieno[2,3-*d*]pyrimidine scaffold. These inhibitors had micromolar activity against Hsp90, wild-type B-Raf, and the critically important B-RafV600E mutant.

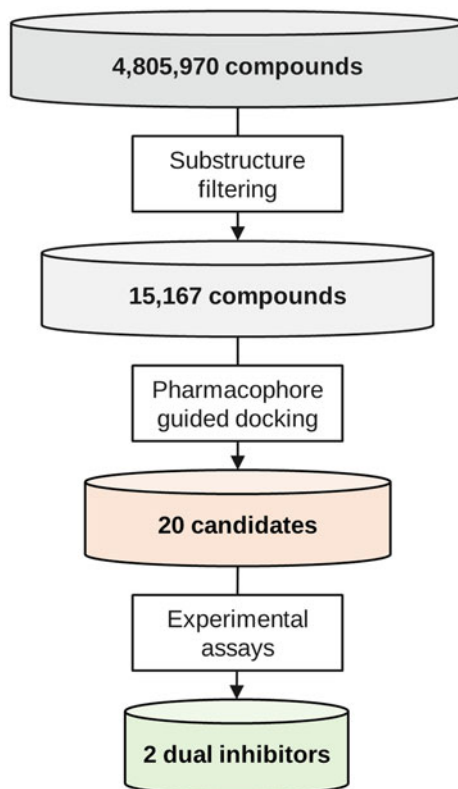


Fig. 4 A schematic representation of the applied virtual screening protocol is shown

2 Materials

2.1 Protein Structures

For B-Raf and Hsp90, eight and nineteen X-ray structures of complexes with different inhibitors were selected, respectively (Table 1). For B-Raf, structures with inhibitors for which high-confidence activity data were available were selected on the basis of a previous study [18]. For Hsp90, structures with a co-crystallized inhibitor having a K_i value of less than 10 μM were collected (see Note 1). All complex structures were taken from the Protein DataBank (PDB) [19].

2.2 Compound Database

A database of 4,805,970 unique compounds was assembled by merging the catalogs of 10 vendors including AMRI, Aronis, Asinex, Chem-Bridge, Enamine, InterBioScreen, Life Chemicals, Maybridge, Otava, and Vitas-MLab.

Table 1
X-ray structures of B-Raf and Hsp90

	Analyzed complexes	Selected complexes (PDB code)	Resolution (Å)	Ligand ID	K_i (nM)
B-Raf	8	3IDP	2.70	L1E	1
Hsp90	19	3RLR	1.70	3RR	30

Reported are the number of X-ray complex structures considered for B-Raf and Hsp90 and the two selected complexes. For inhibitors L1E and 3RR, K_i values are provided

3 Methods

3.1 Selection of Reference Ligands

Selected complex X-ray structures were analyzed to identify inhibitors potentially representing promising templates for searching for dual inhibitors. We especially focused on known inhibitors that were structurally similar and/or displayed similar interaction patterns in the binding sites of Hsp90 and B-Raf. Importantly, crystallographic ligands revealed bioactive conformations and detailed interactions with the binding site residues. This interactive analysis prioritized two known inhibitors as templates. These compounds were designated L1E and 3RR and contained in PDB entries 3IDP [20] and 3RLR [21], respectively. L1E and 3RR had a common scaffold (Fig. 5) and were thus structurally related (see Note 2).

3.2 Substructure Search

A SMARTS string representing the common scaffold of L1E and 3RR was generated and used to conduct a substructure search with OpenBabel 2.3.2 [22] in our database of commercial compounds. All ring substitutions preserving aromaticity of the core structure were permitted. A focused library of 15,167 compounds matching the SMARTS representation was obtained.

3.3 Pharmacophore Model

To further analyze the similarity of L1E and 3RR from a 3D perspective, their bioactive conformations were superposed on the basis of the shared scaffold (Fig. 5). The superposition highlighted additional common structural features including a hydrogen bond donor, a hydrogen bond acceptor, and an aromatic center, known to be important for activity against both Hsp90 and B-Raf. These features were combined into a pharmacophore model (Fig. 5) generated with the Molecular Operating Environment 2014.09 (MOE) [23].

3.4 Protein Preparation

Protein structures 3IDP (B-Raf) and 3RLR (Hsp90) were selected as templates for SBVS and prepared for docking calculations. 3IDP missed 15 residues (from 598 to 613) in the kinases activation loop [20]. Therefore, MOE was used to fill the gap by building the crystallographically not resolved residues and generating an initial 3D conformation by structural relaxation. Refinement of the loop conformation was carried

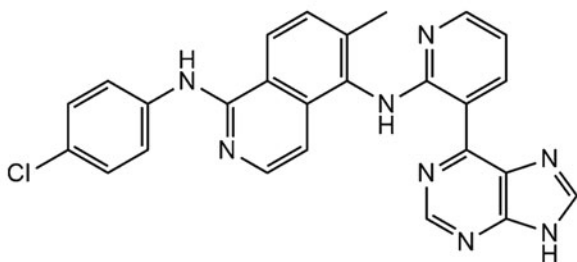
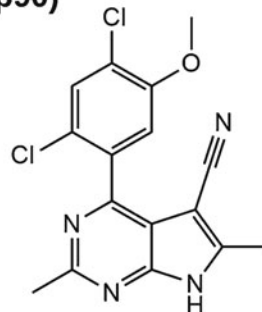
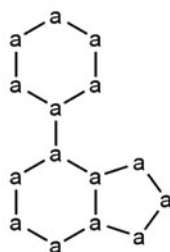
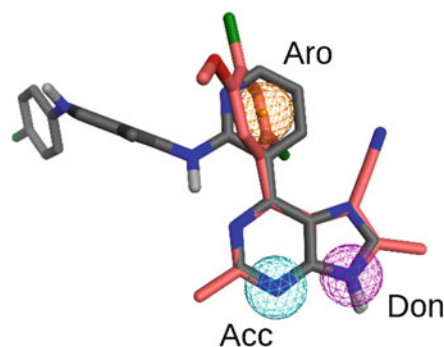
L1E (B-Raf)**3RR (Hsp90)****Common scaffold****Pharmacophore model**

Fig. 5 The structures of B-Raf and Hsp90 inhibitors selected as templates for computational compound screening are shown in the upper half of the figure. In the shared scaffold, shown in the *lower left panel*, “a” stands for any aromatic atom. The modeled pharmacophore is depicted in the *lower right panel*

out using ModLoop [24]. A total of 300 loop conformations were generated and the one yielding the most favorable force field energy score was retained. Both 3IDP and 3RLR structures were prepared with MOE. Preparation steps included the selection of L1E and 3RR as ligands that were automatically removed during docking calculations, addition of hydrogen atoms, and calculation of atomic partial charges according to the MMFF94x force field. All water molecules were removed except for one water molecule in Hsp90 (number 2 according to 3RLR residues numbering) that was known to be intimately involved in a conserved network of hydrogen bonds formed between Hsp90 and several other inhibitors [21].

3.5 Docking

Docking of the focused compound library generated by substructure search (see above) was carried out with the Dock module of MOE [23]. The library was also computationally screened to remove known pan assay interference compounds (PAINS) [25] using publicly available filters [26] and aggregating compounds using the ZINC/UCSF aggregator advisor [27]. Target sites for docking in 3IDP and 3RLR were defined using the experimental coordinates of L1E and 3RR,

respectively. A docking protocol was set up to exploit the feature information encoded in the pharmacophore model (see **Note 3**). The pharmacophore placement algorithm implemented in MOE was used to generate up to 1000 poses per test compound. The poses were scored and ranked based on the London dG function. The top 10 poses were selected and subjected to energy refinement and re-scoring with the more complex GBV/WSA dG function. Among poses matching the pharmacophore model, if any, the best scoring pose was ultimately obtained for each test compound. For both Hsp90 and B-Raf, highly ranked compounds were visually inspected for conformational and interaction commonalities to aid in a final selection of 20 candidates. The selected candidates were tested in vitro for activity against both targets. Experimental procedures were extensively described in the original publication [7].

4 Results Summary

Two of 20 candidate compounds were found to have micromolar activity against both Hsp90 and B-Raf in enzyme assays (compound **1** and **2** in Table 2 and Fig. 6). Interestingly, **1** and **2** also displayed comparable or higher activity against B-Raf mutant V600E (Table 2). As a consequence of the applied computational strategy, the newly identified compounds were structurally similar to L1E and 3RR. However, among known inhibitors, they were the first compounds with dual activity against Hsp90 and B-Raf.

Binding modes of **1** and **2** predicted by docking (Fig. 6) were characterized by the formation of known key interactions with target binding sites previously observed for other inhibitors such as hydrogen bonding with Asp93 in Hsp90 and with Cys532 of the hinge region in B-Raf. The modeled binding modes were further supported by superposition with known crystallographic inhibitors. Identification of **1** and **2** provides a basis for compound optimization and polypharmacological targeting of Hsp90 and B-Raf. The computational approach described herein (Fig. 4) provides an example for combining LBVS (using substructures and pharmacophores) and SBVS (docking) in the search for dual-target inhibitors. While we do not claim that this protocol might be generalizable, it nicely illustrates the potential of knowledge-based computational approaches for specialized drug design applications.

Table 2
Experimental IC₅₀ values of the virtual screening hits

Compound	IC ₅₀ B-Raf (μM)	IC ₅₀ B-Raf V600E (μM)	IC ₅₀ Hsp90 (μM)
1	28.9 ± 1.7	9.1 ± 0.4	1.2 ± 0.1
2	1.5 ± 0.1	2.5 ± 0.2	7.6 ± 0.6

IC₅₀ values of compounds **1** and **2** are reported for Hsp90, B-Raf, and B-Raf V600E

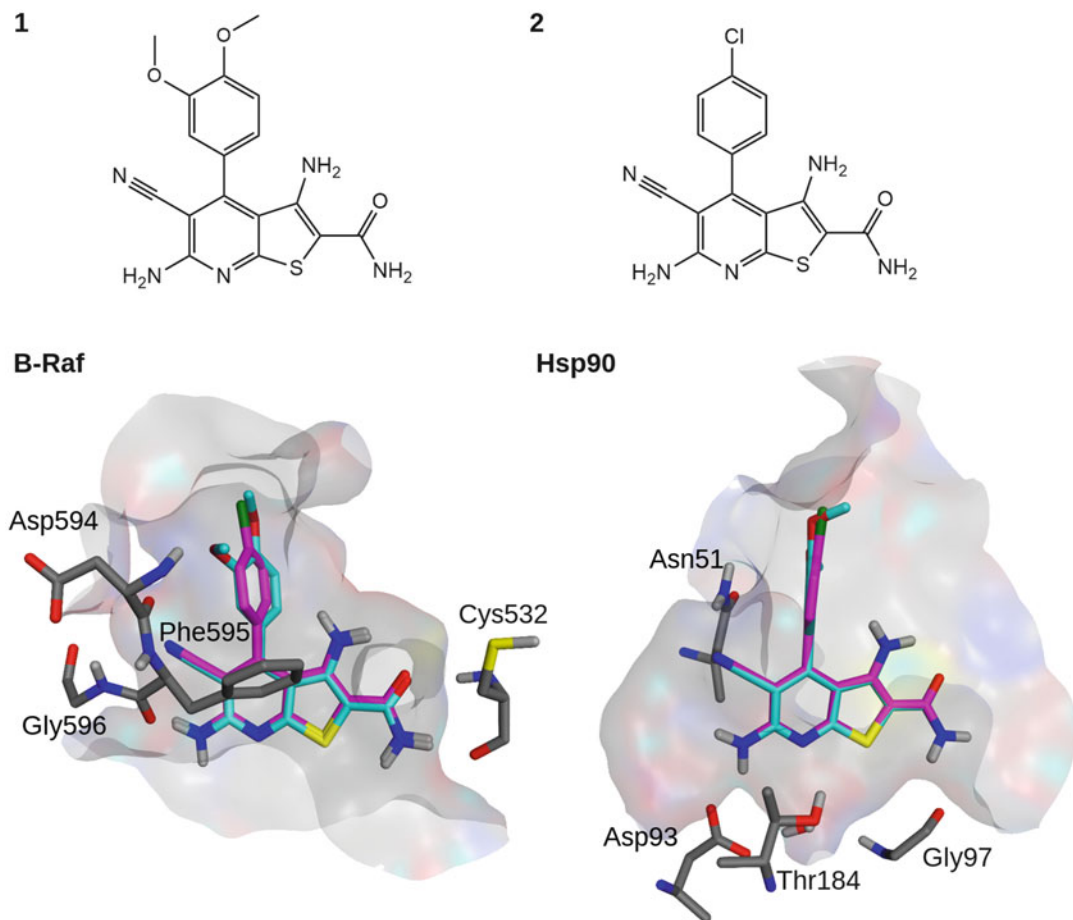


Fig. 6 The structures of the identified dual inhibitors are shown in the upper half of the figure. Binding modes predicted by docking into B-Raf and Hsp90 are shown in the lower half. In the docked complexes, compound **1** is colored in cyan and compound **2** in magenta

5 Notes

1. Our focus was on known inhibitors that were relatively potent in order to improve the odds of identifying comparably potent dual inhibitors. However, the choice of a potency threshold should also take into account the degree of chemical diversity among co-crystallized ligands. In the case of Hsp90 and B-Raf, a variety of crystallographic inhibitors existed, hence making it possible to raise the bar. In other cases, weakly potent compounds may require more extensive consideration.
2. The initial analysis of complex X-ray structures was interactive and largely relied on graphical assessment. As such, the approach was at least partly subjective in nature and involved chemical experience and intuition. Automated analysis would be indispensable if large

sets of complexes would need to be analyzed. For example, scaffolds could be extracted from candidate compounds and systematically compared using numerical similarity measures. Similarly, the geometry and chemical features of active sites might be automatically compared. However, while automated analysis will likely assign priorities to complexes for further study, it will typically not be able to entirely replace an in-depth analysis by experts.

3. Docking protocols including constraints in binding mode generation and selection such as those provided by a pharmacophore model or a set of pre-defined key interactions represent an intrinsically knowledge-based component of VS campaigns. Although such approaches are not without caveats and potential errors, e.g., the use of ill-defined constraints cannot be compensated for in subsequent computations, they typically help to narrow down candidate compounds and reduce noise of computational screening. In the case of our study, pharmacophore-constrained docking was designed to eliminate many compounds from further consideration that could not possibly satisfy interactions known to be critical for inhibition, although they might be able to fit the geometry and shape of the Hsp90 and/or B-Raf binding site. This type of negative selection helped to quickly focus docking calculations on more promising candidates.

Acknowledgement

We thank OpenEye Scientific Software, Inc., for a free academic license of the OpenEye Toolkit and Chemical Computing Group, Inc., for academic teaching licences of the Molecular Operating Environment.

References

1. Peters J-U (2013) Polypharmacology—foe or friend? *J Med Chem* 56:8955–8971
2. Anighoro A, Bajorath J, Rastelli G (2014) Polypharmacology: challenges and opportunities in drug discovery. *J Med Chem* 57:7874–7887
3. Terstappen GC, Schlüpen C, Raggiaschi R, Gaviraghi G (2007) Target deconvolution strategies in drug discovery. *Nat Rev Drug Discov* 6:891–903
4. Schenone M, Dančík V, Wagner BK, Clemons PA (2013) Target identification and mechanism of action in chemical biology and drug discovery. *Nat Chem Biol* 9:232–240
5. Morphy R, Rankovic Z (2005) Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem* 48:6523–6543
6. Rastelli G, Pinzi L (2015) Computational polypharmacology comes of age. *Front Pharmacol* 6:157
7. Anighoro A, Pinzi L, Marverti G, Bajorath J, Rastelli G (2017) Heat shock protein 90 and serine/threonine kinase B-Raf inhibitors have overlapping chemical space. *RSC Adv* 7:31069–31074
8. Li J, Buchner J (2013) Structure, function and regulation of the Hsp90 machinery. *Biomed J* 36:106–117

9. Echeverría PC, Bernthaler A, Dupuis P, Mayer B, Picard D (2011) An interaction network predicted from public data as a discovery tool: application to the Hsp90 molecular chaperone machine. *PLoS One* 6:e26044
10. Neckers L, Workman P (2012) Hsp90 molecular chaperone inhibitors: are we there yet? *Clin Cancer Res* 18:64–76
11. Bhat R, Tummalapalli SR, Rotella DP (2014) Progress in the discovery and development of heat shock protein 90 (Hsp90) inhibitors. *J Med Chem* 57:8718–8728
12. Zhang W, Liu HT (2002) MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Res* 12:9–18
13. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, Davis N, Dicks E, Ewing R, Floyd Y, Gray K, Hall S, Hawes R, Hughes J, Kosmidou V, Menzies A, Mould C, Parker A, Stevens C, Watt S, Hooper S, Wilson R, Jayatilake H, Gusterson BA, Cooper C, Shipley J, Hargrave D, Pritchard-Jones K, Maitland N, Chenevix-Trench G, Riggins GJ, Bigner DD, Palmieri G, Cossu A, Flanagan A, Nicholson A, Ho JWC, Leung SY, Yuen ST, Weber BL, Seigler HF, Darrow TL, Paterson H, Marais R, Marshall CJ, Wooster R, Stratton MR, Futreal PA (2002) Mutations of the BRAF gene in human cancer. *Nature* 417:949–954
14. Whittaker S, Kirk R, Hayward R, Zambon A, Viros A, Cantarino N, Affolter A, Nourry A, Niculescu-Duvaz D, Springer C, Marais R (2010) Gatekeeper mutations mediate resistance to BRAF-targeted therapies. *Sci Transl Med* 2:35ra41
15. Haarberg HE, Smalley KSM (2014) Resistance to Raf inhibition in cancer. *Drug Discov Today Technol* 11:27–32
16. Grbovic OM, Basso AD, Sawai A, Ye Q, Friedlander P, Solit D, Rosen N (2006) V600E B-Raf requires the Hsp90 chaperone for stability and is degraded in response to Hsp90 inhibitors. *Proc Natl Acad Sci U S A* 103:57–62
17. Anighoro A, Stumpf D, Heikamp K, Beebe K, Neckers LM, Bajorath J, Rastelli G (2015) Computational polypharmacology analysis of the heat shock protein 90 interactome. *J Chem Inf Model* 55:676–686
18. Furtmann N, Hu Y, Bajorath J (2015) Comprehensive analysis of three-dimensional activity cliffs formed by kinase inhibitors with different binding modes and cliff mapping of structural analogues. *J Med Chem* 58:252–264
19. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
20. Smith AL, DeMorin FF, Paras NA, Huang Q, Petkus JK, Doherty EM, Nixey T, Kim JL, Whittington DA, Epstein LF, Lee MR, Rose MJ, Babij C, Fernando M, Hess K, Le Q, Beltran P, Carnahan J (2009) Selective inhibitors of the mutant B-Raf pathway: discovery of a potent and orally bioavailable aminoisoquinoline. *J Med Chem* 52:6189–6192
21. Kung P-P, Sinnema P-J, Richardson P, Hickey MJ, Gajiwala KS, Wang F, Huang B, McClellan G, Wang J, Maegley K, Bergqvist S, Mehta PP, Kania R (2011) Design strategies to target crystallographic waters applied to the Hsp90 molecular chaperone. *Bioorg Med Chem Lett* 21:3557–3562
22. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) OpenBabel: an open chemical toolbox. *J Cheminform* 3:33
23. Molecular Operating Environment (MOE) (2014) Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7
24. Fiser A, Sali A (2003) ModLoop: automated modeling of loops in protein structures. *Bioinformatics* 19:2500–2501
25. Baell J, Walters MA (2014) Chemistry: chemical con artists foil drug discovery. *Nature* 513:481–483
26. ZINC PAINS Patterns Search. <http://zinc15.docking.org/patterns/home/>. Accessed 17 June 2016
27. Irwin JJ, Duan D, Torosyan H, Doak AK, Ziebart KT, Sterling T, Tumanian G, Shoichet BK (2015) An aggregation advisor for ligand discovery. *J Med Chem* 58:7076–7087



Strategies for Multi-Target Directed Ligands: Application in Alzheimer's Disease (AD) Therapeutics

Sucharita Das and Soumalee Basu

Abstract

Design of multi-target directed ligand (MTDL) is believed to be a novel and improved approach for diseases that elucidate a multifactorial nature. Alzheimer's disease (AD) is related to increased levels of the amyloid β peptide ($A\beta$), the hyperphosphorylated tau protein, free radicals, oxidized proteins and lipids, metal ion dysregulation and many more. For the multifactorial aetiology of AD and the fact that till date there is no effective treatment besides drugs alleviating associated symptoms, molecules designed to hit simultaneously different key targets of the complex pathological network emerges as a more realistic alternative. In this context, the present chapter puts forward a note of several strategies adopted for the development of MTDLs for the disease followed by a case study leading to in vitro validation.

Keywords Anti-amyloidogenic, Antioxidant, $A\beta$, BACE1, Inhibitor, MTDL, QSAR, Virtual screening

1 Introduction

Discovery and development of drugs is routinely becoming complex and challenging with the aggressive accumulation of molecular data and with rapid advances in technology. Drug research aiming at the discovery of a single drug, modulating the function of a single protein target, though prevalent was unsuccessful for multifactorial diseases. Of late, the philosophy of drug designing has remodelled from single target therapy to polypharmacology wherein either a single drug acts on multiple targets of an exclusive disease pathway or it acts on multiple targets related to multiple disease pathways. These multi-target ligands are more likely to be effective for multifactorial diseases due to reduced side effects of unintended drug–target interaction and lesser vulnerability to drug resistance. Moreover, highly selective ligand for a particular target does not always result in a clinically efficacious drug. Besides, drugs targeting only one protein are more vulnerable to resistance. Even a single mutation in the active site of the protein might affect ligand binding affinity reducing the efficacy. On the contrary, resistance to drugs targeting multiple proteins requires the unlikely event of concurrent mutations in multiple targets. In view of the prevalence

and expected increase in the incidence of diseases involving multiple pathways, the design and development of efficacious and safe agents has become a hotspot in the field of pharmaceutical research.

Diseases like neurodegenerative syndrome, diabetes, cardiovascular diseases, cancer and many more that involve multiple pathogenic factors are inadequately treated by drugs hitting a single target. Different pharmacological approaches are adopted to circumvent this issue. Multiple-medication therapy (MMT) also referred to as cocktail or combination of drugs is a blend of two or three drugs that combine several therapeutic mechanisms. This approach might not work well for patients with compliance problem. A second approach that handles this problem of compliance includes the use of multiple-compound medication (MCM) also referred to as “a single pill drug combination” wherein different drugs are incorporated into a single formulation thus simplifying the dosing regimen. The third emerging approach that forms the eventual choice for multifactorial pathoetiological diseases is the multi-target directed ligand (MTDL) in which a single compound simultaneously targets multiple proteins. Definitely, therapy with a single drug with multiple properties would be inherently advantageous over MMT and MCM whereby the administration of multiple single drug-entities with differential bioavailability, pharmacokinetics, metabolism and so forth is a challenge. However the reductionist approach centred around single drug for single molecular target would remain as a milestone for years to come. Needless to say that it has already yielded many successful drugs that are still in use.

It is worth mentioning that use of combination of drug entities for multiple targets has been in vogue for quite a few years through a polypharmaceutic approach. Polypharmacy is the approach whereby several drugs that independently act on different aetiological targets of a disease are combined. For instance, an HIV reverse transcriptase inhibitor and an HIV protease inhibitor are co-administered as a cocktail [1] in AIDS therapy. Another approved combination of drugs for the treatment of asthma is the corticosteroid—fluticasone and salmeterol, a bronchodilator. The drugs simultaneously target the inflammation and bronchoconstriction associated with the disease [2]. The combination of the calcium channel blocker amlodipine and the cholesterol-reducing agent atorvastatin in the treatment of cardiovascular disease [3] is yet another example. It may be borne in mind that there are several limitations behind the use of this kind of pharmaceutical combination. Apart from the demerit of different degrees of bioavailability, pharmacokinetics and metabolism [4, 5], polypharmacy might bring in or even multiply toxicity and side effects. Even worse is the possibility of an unforeseen drug–drug interaction [6] that can be fatal [7]. Research has therefore focused on the advantages associated with the design of single drug molecules acting on two

or more specific aetiological targets of a particular disease. In this case the possibility of unwanted side effects is much less and even if there are such effects it would be a much easier task to 'design out' the effect when only one ligand in contrast to many is used. Poly-pharmacology where compounds are designed as multiple-ligand drugs is the development of disease-modifying therapeutics addressing the principal causes of the disease. Elementary development of such compounds under numerous names such as 'multi-target directed ligands', 'designed multiple ligand', 'dual ligand', 'dual-mechanism', 'bifunctional', 'multi-functional', 'multimechanistic', 'multimodal', 'pan-agonist' or 'hybrid' drugs [8–14], started way back in 1990.

MTDLs are single chemical entities that simultaneously modulate multiple targets. The molecules are conceived to interact directly but weakly with multiple targets thus obviating the challenge of administering multiple drugs with different levels of bio-availability, pharmacokinetics and metabolism. Besides ensuring simplification of the therapeutic regimen to the patients it provides drugs that would prevent and cure diseases effectively and safely. MTDLs have found its use in a wide range of diseases [15]. Diseases associated with the central nervous system (CNS) that have been suggested to be catered by this strategy include movement disorders, cognitive deficit disorders, negative symptoms in schizophrenia, Lewy body disease and depressive illness [16–19].

Approved drugs currently prescribed for Alzheimer's disease are incapable of curing the disease and provide symptomatic relief. MTDLs thus emerge as a very significant strategy for designing of safer drugs in such complex diseases. It thus becomes absolutely essential to articulate the different methods adopted in designing such ligands for the scope of developing mechanism-modulating drugs for AD.

2 Materials

BACE1 and β -secretase activity assay kit, 2,2'-Azinobis-3-ethylbenzothiazoline-6-sulfonic acid (ABTS) and Thioflavin-T were purchased from Sigma-Aldrich. Hesperidin was purchased from EXTRASYNTHÈSE and A β (25–35) was purchased from ANA-SPEC. ANS was gifted by Dr. Rajat Banerjee (Department of Biotechnology, University of Calcutta). All solvents used in the study were of analytical grade (E. Merck). Deionized water from a Milli-Q system apparatus (Millipore Corp., Billerica, MA) was used throughout the experiments.

3 Methods for Development of MTDLs

The adopted methods that are commonly practised for this purpose may be very crudely classified to one in which a random screening approach is utilized and another in which a knowledge-based approach is used to combine scaffolds from different active molecules with known activity against a particular target. The latter is termed as framework combination approach. The scheme in Fig. 1 portrays the subgroups in each approach.

3.1 *Random Screening Approach*

Random screening is either carried out using virtual screening protocols or through multi-target quantitative structure activity relationship (mt-QSAR). The search by Huang et al. [20] may be taken up as an example of virtual screening. It is a bi-layered screening approach (Fig. 2) wherein core structural elements for H₃R antagonist and BACE1 along with AChE inhibitor were hybridized to generate the MTDL. Initially, a virtual database containing derivatives of quinoxaline was screened using a pharmacophore model of BACE1 inhibitor to arrive at 2073 compounds and next, docking studies with AChE were used for filtering the number to 17. The scaffolds were prepared such that they consisted of 2-amino-3,4-dihydroquinazoline of BACE 1 inhibitor joined to a benzyl pyrrolidine fragment of BYYT-25, an AChE inhibitor. The ring core moiety of the BACE1 inhibitor combined to the basic centre (contributed by BYYT-25) resulted in the scaffold that

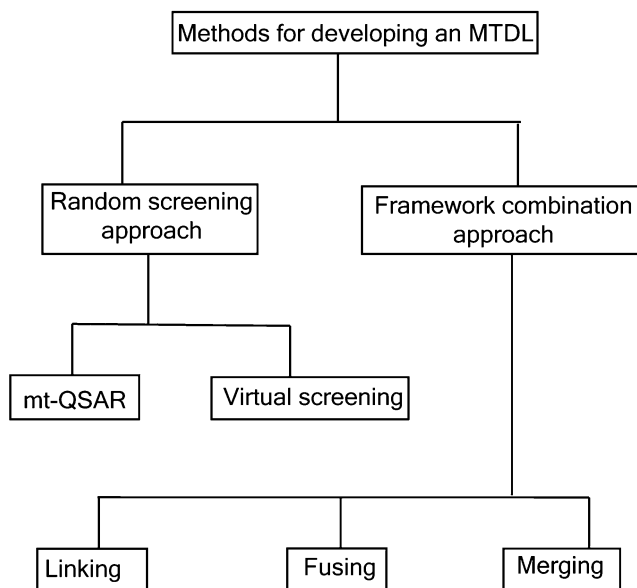


Fig. 1 Methods used for development of MTDL

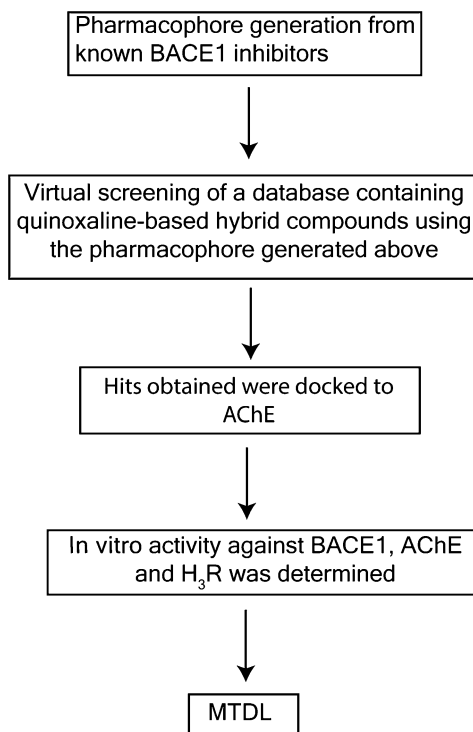


Fig. 2 A schematic representation of the bi-layered screening approach

consisted of all the required pharmacophores of an H₃R antagonist. These 17 hits were synthesized and experimentally validated.

mt-QSAR has certain merits over structural similarity and docking methods such as for accuracy of prediction and scaffold hopping. A study by Fang et al. [21] is the first report where the mt-QSAR approach has been used, validated and successfully proven in favour of polypharmacology for anti-AD drugs and eventually for discovering MTDLs. Here, naïve Bayesian (NB) and recursive partitioning (RP) algorithm have been employed for constructing classifiers using two kinds of fingerprints, namely ECFP₆ and MACCS and predicting active molecules against 25 key targets related to AD using the multitarget-quantitative structure–activity relationships (mt-QSAR) method. One hundred classifiers were constructed to predict the chemical protein interaction using the mt-QSAR method and the predictability of the model has been validated by cross-validation and test set validation.

3.2 Framework Combination Approach

In this knowledge-based approach, the strategy is to use the knowledge of activity of potent compounds for each target from individual SAR (structure–activity relationship) for the combination of different activities in a single molecule by integrating the pharmacophores of the selected molecules. The integration may occur in a

few different ways as shown in Fig. 3. In linked ligand they are separated by a distinct linker group that is absent in the structure of either of the ligands. The pharmacophore of this resulting molecule is actually responsible for the activity with different targets [22]. In some cases the linker may be designed to be cleavable so that it may be metabolized to release the molecules to interact with their targets independently. In yet another case the molecules might be in contact to each other without requiring a linker to link them and thus be fused [23]. The framework in merged ligands is overlapped on the basis of structural similarity with the starting structure. This combination strategy almost always produces dual ligands, discovering ligands that bind to more than two targets usually demanding a screening approach.

The difference between screening approach and framework combination is that in the former a novel chemotype may be selected whereas in framework combination selective ligands for known targets are considered from the beginning. However in framework combination, incorporation of subsequent activities retaining the former may be difficult to be employed.

3.3 A Hybrid Approach

A more recent approach is the one in which combination of docking studies and 2-D QSAR models for different targets have been successfully applied for a multi-tier screening [24]. In this study, the authors choose the targets supporting the A β hypothesis which emerges as the prominent pathological mechanism for the complex neurological disorder, AD. Aberrant processing of the amyloid precursor protein (APP) by BACE1 results in toxicity and can cause oxidative stress mediated neuronal death. The study by

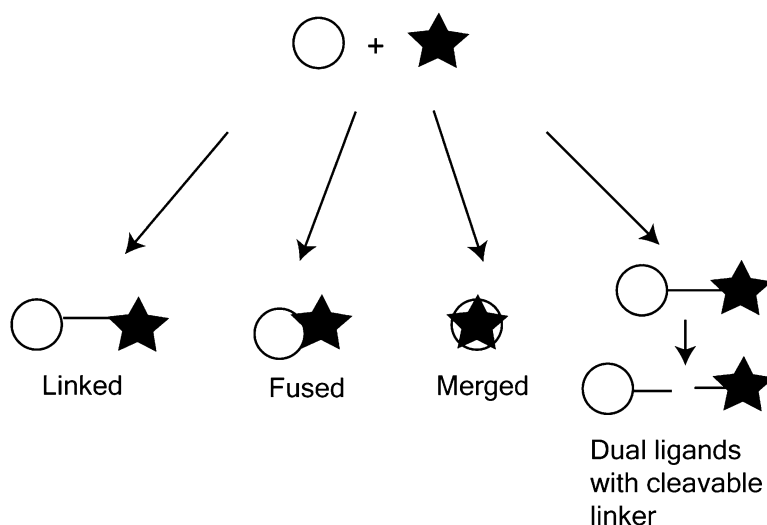


Fig. 3 A schematic representation of the different methods of framework combination approach

Chakraborty et al. has exploited these facts for the design of effective multi-potent agents from natural origin. Therefore an MTDL was developed (Fig. 4) that could inhibit the protease BACE1 as well as exhibit anti-amyloidogenic and antioxidant activities. The steps undertaken for the method are described below.

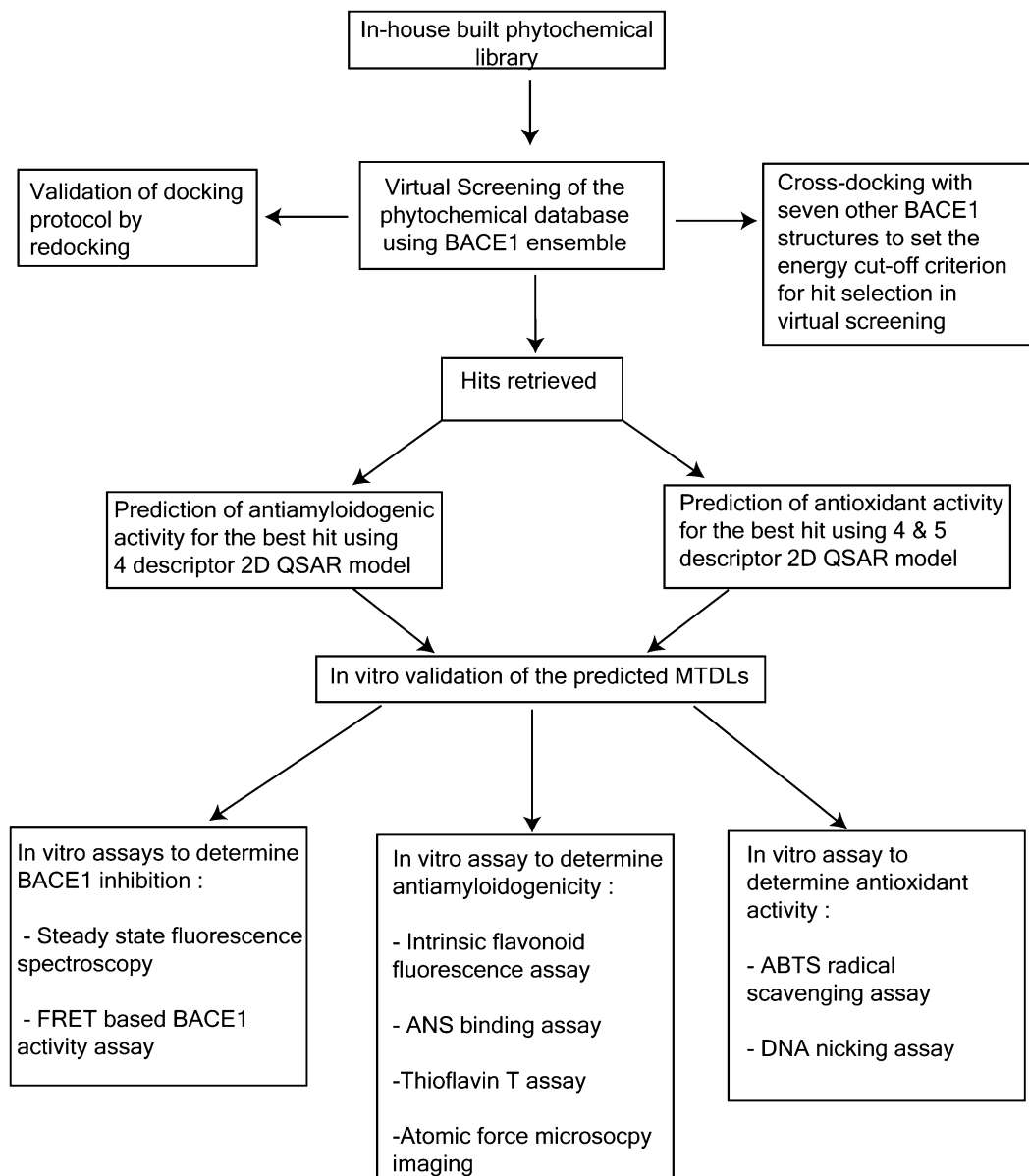


Fig. 4 A flowchart of the steps of hybrid approach

3.3.1 Development of an In-House Phytochemical Library

An in-house phytochemical library with 200 entries was built as the first step. The 3-D structure of each phytochemical was built from the respective 2D structures using HYPERCHEM 8.0. [25]. Minimization using MM+ molecular mechanical force field with Polak-Ribiere conjugate gradient algorithm was carried out to RMS gradient of 0.001 kcal/Å mol. 3D structures of all the phytochemicals were prepared in the SYBYL mol2 format. This database of phytochemicals was screened simultaneously by molecular docking algorithm with receptor ensemble of BACE1.

3.3.2 Ensemble Docking

A structure-based docking was next carried out with BACE1 to mine inhibitors for the highly flexible enzyme. It is due to its flexibility that BACE1 exhibits structural reorientation [26] and is capable of binding ligands that range from smaller molecules to larger peptide fragments. Choosing the appropriate receptor conformation for docking thus forms a crucial step here. Several crystal structures for BACE1-inhibitor complexes available in the Protein Data Bank (PDB) were systematically searched and aligned for comparison of the active site. Two different BACE1-inhibitor crystal structures (PDBID: 3TPP and 3IND), one with a squeezed active site cavity and another with a wider one were chosen for docking (Fig. 5). Before using the method of molecular docking for screening, the docking method was validated by re-docking. In this approach, the bound inhibitors of the 3TPP and 3IND are docked back to their respective receptors using the docking method of FlexX. Re-docking with 3IND yielded two solutions in terms of

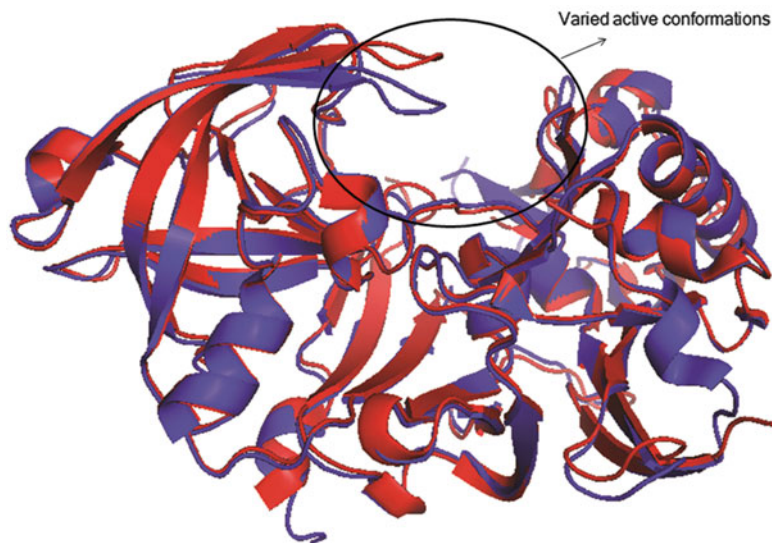


Fig. 5 Superimposed images of the two BACE1 structures—3IND (red) and 3TPP (blue)

the most favourable docking energy and best root mean square deviation (RMSD) with respect to the crystallographic orientation of the bound inhibitor in 3IND. For 3TPP, the solution with the most favourable docking energy exhibited the lowest RMSD with the corresponding crystallographic orientation. These results suggest that the lowest energy docking solution exhibited by FlexX reproduces the pharmacophoric features for the inhibition of the enzyme. Further, cross-docking was also applied such that the inhibitors of 3IND and 3TPP were interchangeably docked and as expected the pharmacophoric features could not be unfolded from cross-docking. This is so as the two inhibitors largely differ in size as do the binding cavities of their respective receptors. Results of re-docking and cross-docking advocates us to use the small as well as the large cavity receptors for virtual screening of unknown ligands which in turn calls for the application and hence validation of ensemble docking. That ensemble docking can be successfully applied to BACE1 was already tested by our group [27]. In this study, multiple receptor conformations (MRC) of BACE1 was generated by normal mode analysis and the structures were subsequently used to dock myricetin, a polyphenol for the elucidation of pharmacophoric features necessary for it to dose dependently inhibit the enzyme with an IC_{50} of 2.5 μ M. A similar validation with seven BACE1-inhibitor crystal complexes obtained from PDB was carried out. A dataset with diverse structural scaffolds and the binding potency ranging from μ M to nM was considered. The seven inhibitors were re-docked into the respective receptors reproducing the inhibitor orientation as observed in the crystal structure judged by the RMSD (<1 Å) between the docked pose and the crystallographic orientation. These seven inhibitors were further docked to the receptor ensemble of BACE1 (3IND and 3TPP) and the results implied that inhibitors with higher binding affinity reveal exact crystallographic contacts as observed in the crystal structure. Highly active inhibitors were found to dock with much higher affinity which led us to set an energy cut-off criterion of -20 kJ/mol for a compound to be judged active during virtual screening.

3.3.3 Virtual Screening

The in-house developed library was subsequently screened adopting the high-throughput ensemble docking approach using both 3IND and 3TPP as the structure of BACE1. Apart from selecting the phytochemicals with a docking energy cut-off criterion of -20 kJ/mol, the docked conformation corresponding to the higher binding affinity was selected. Sixty-four phytochemicals were mined, out of which the highest binding affinity was calculated for hesperidin.

3.3.4 Development of a QSAR Model for Anti-Amyloidogenic Activity

Eighteen polyphenols from different structural classes of flavonoids were selected from literature survey with known anti-amyloidogenic activity expressed as their IC_{50} value and converted to the logarithmic scale. The structures were drawn using the molecular builder interface of HYPERCHEM 8.0 and minimized following the steps discussed in Subheading 3.3.1. The CODESSA package [28] was used for the calculation of several descriptors and also for QSAR modelling. Descriptors are classified into six sub-groups (constitutional, topological, geometrical, electrostatics, quantum-chemical and thermodynamic) in CODESSA. Out of these, for the calculation of quantum-chemical and thermodynamic descriptors, wave-function optimization was carried out on all the pre-optimized structures using the AM1 semi empirical method from the AMPAC 9.2 packages [29]. Other descriptors of each structure were calculated using the descriptor calculation tool available in CODESSA. 1-D, 2-D and multivariate analysis tools of the CODESSA package were utilized for preliminary analysis of the physico-chemical properties and descriptors. Heuristic methodology of the package was used to generate the QSAR models. Initially, 458 descriptors were calculated for each of the dataset compounds. Descriptors with missing and unacceptable values were eliminated along with the highly correlated ones obtained from the one-parameter regression models. Next, two-parameter regression models computed on the remaining descriptors were arranged on the basis of squared correlation coefficient. Then, multi-parameter regression models were generated by step-wise addition of descriptors. Highest values of R^2 , R_{cv}^2 and the F value judge the best model. The predictive potential of these models is assessed using the squared correlation coefficient R^2 , root mean square error (RMSE) and F -test values. Finally, validation of the designed QSAR model was performed by cross-validation techniques using the leave-one-out (LOO) method. R_{cv}^2 is calculated by removing a data point from the data set and the regression is recalculated. This process is repeated until each datum has been omitted at least once. One more thing that becomes very important is to determine the number of descriptors required to explain the observed anti-amyloidogenic activity. A simple breaking point rule was applied to determine the optimum number of descriptors. Multi-parameter QSAR models were generated using 2–6 descriptors. For a given number of descriptors, heuristic method generates many possible QSAR models. The QSAR model selected based on the highest predicted R^2 value among all generated models for a given number of descriptor is supposed to be the best model. To judge the applicability of the model, cross-validation techniques have been used. Variations of the predicted R^2 and R_{CV}^2 of the QSAR models with the number of descriptors increase significantly for QSAR model with four descriptors. Both the squared correlation terms thereafter increase

marginally with the increasing number of descriptors indicating the “breaking point” to have occurred at 4, i.e., at least four descriptors are required to model the anti-amyloidogenic properties of the polyphenols with reasonable accuracy. The descriptors include HBSA (hydrogen bonding surface area), average information content of zero order (topological descriptor), relative negative charged surface area and the vibrational entropy per atom (thermodynamic descriptor).

Anti-amyloidogenic activity of the hits obtained using docking energy cut-off was assessed using this 5-parameter QSAR model and hesperidin was predicted to show high anti-amyloidogenic activity. Thereafter another QSAR model for antioxidant activity was developed for the progress of the process of MTDL development.

3.3.5 QSAR for Antioxidant Activity

Literature study yielded antioxidant activity of 33 phytochemicals expressed as TEAC values that were converted to the logarithmic scale. These compounds were chosen such that the activities spanned over tenfold in magnitude and their structures were drawn using the molecular builder interface of HYPERCHEM 8.0 and minimized following the steps discussed in Subheading 3.3.1. The CODESSA package was used for the calculation of several descriptors and also for QSAR modelling following the same method as described in the section for anti-amyloidogenic activity.

A total of 458 descriptors belonging to different classes were calculated which was reduced by eliminating descriptors with missing values and also those that are highly correlated. Bayesian multiple linear regression (BMLR) was finally carried out on 426 descriptors. The dependent variable in the QSAR model was the observed free radical scavenging activity and the independent variable constituted of the various combination of the chosen descriptors. Determination coefficient (R^2), standard deviation (σ), F -statistics (F) and t -statistics (t) were considered to evaluate the quality of the model statistically.

Another challenge was to determine the optimum number of descriptors required to explain the modelled activity. An addition of descriptors no doubt increases the R^2 value but simultaneously, the model suffers from over-parameterization. Series of models comprising of 2–7 descriptors were obtained using BMLR method. Variation of the predicted R^2 with the number of descriptors associated with the particular model increases up to four, after which the value saturates. Thus the slope of the line changes at 4 giving rise to a break point at 4, implying that a model with four descriptors is necessary and sufficient to predict the antioxidant activity with reasonable accuracy. It may be noted that the R^2 marginally increases from 0.8903 for four-descriptor model to 0.9094 and

0.9094 for five- and six-descriptor model, respectively. These three models were further validated using cross-validation techniques. Cross-validated squared correlation coefficient (R_{cv}^2) was calculated using the leave-one-out (LOO) method. R_{cv}^2 is supposed to be a better measure of the predictive power of a regression equation. It is highly sensitive to the number of descriptors and decreases when a model is over-parameterized. $R_{cv}^2 > 0.5$ is generally supposed to be an acceptable prediction ability. For this study, R_{cv}^2 with four-descriptor model is 0.8318, but for five-descriptor model it significantly increases to 0.8706 although for six-descriptor model it marginally increases to 0.8723. Finally carrying out principal component analysis for all the three models and subsequently plotting the first two score vectors in the form of a bi-dimensional projection, it was concluded that the descriptors in the four- and the five-descriptor models being widely distributed unlike the six-descriptor QSAR model, the former two descriptor models were considered for further studies. The following are the descriptors of the four-descriptor QSAR model—number of benzene ring, zero point vibrational energy (a thermodynamic descriptor that describes the vibrational property of the molecule due to atomic fluctuations at the lowest possible energy state), maximum partial charge for an Oxygen atom and minimum valency of a Carbon atom (a quantum-chemical descriptor negatively correlated with the antioxidant activity). For the five-descriptor QSAR model, FNSA-2 (fractional partial negative surface area – total charge weighted partial negative surface area/total molecular surface area) is an important descriptor and so is HACA-2/TMSA which is also an electrostatic descriptor calculated from the hydrogen acceptor charged surface area and the total molecular accessible surface area.

Hits obtained using docking energy cut-off after being assessed for anti-amyloidogenic activity using a four-parameter 2-D QSAR model [30] were further assessed for predicting antioxidant activity using two (four and five parameter) QSAR models [31]. Moderate antioxidant activity of the most potent compound hesperidin was thus predicted. These predicted activities were further verified by *in vitro* assays.

3.3.6 Experimental Validation

Steady-state fluorescence spectroscopy confirms ligand binding and finds the binding constant. ANS binding assay confirms the binding of the compound at the hydrophobic core of the A β aggregate, Thioflavin T assay shows the ability of the compound to destabilize the fibrils with final confirmation through atomic force microscopy. Finally, ABTS⁺ assay and DNA nicking assay estimate the antioxidant activity of the screened compound.

3.3.7 Steady-State Fluorescence Studies for BACE1–Hesperidin Interaction

BACE1 was prepared in 0.2 M acetate buffer, pH 4.8 and concentration was 100 nM for fluorescence experiments. Concentrated stock solution of hesperidin was prepared in DMSO. To perform the binding studies, fluorescence emission experiments were carried out using the titration method. Aliquot of the inhibitor from the stock solution was gradually added to the protein solution to obtain the desired final concentration of the inhibitor. In all samples, the final concentration of DMSO was kept <1% (by volume). Emission spectra were obtained in the wavelength range from 300 to 450 nm using $\lambda_{\text{ex}} = 295$ nm, whereas excitation spectra were obtained in the wavelength range from 260 to 310 nm using $\lambda_{\text{em}} = 340$ nm.

The equilibrium constant (K) was determined by the above-mentioned fluorescence titration experiments using the Benesi-Hildebrand equation. In the titration experiments, as the ligand is present in excess when compared to the protein and there is an increase in fluorescence intensity with increasing ligand concentration hence it is possible to apply this equation.

The equation used is as follows:

$$1/\Delta F = 1/\Delta F_{\text{max}} + 1/\Delta F_{\text{max}}K [\text{Ligand}] \quad (1)$$

where $\Delta F = F_x - F_0$, F_x and F_0 represent the fluorescence intensities of BACE1 in the presence and absence of added compound, respectively. F_{max} is the maximum change in fluorescence intensity and K is the binding constant for the 1:1 complex. A plot of $1/\Delta F$ vs. $1/[\text{ligand}]$ exhibits strong linearity indicating 1:1 complex formation.

The equilibrium constant (K) for the BACE1-hesperidin complex obtained by fitting the fluorescence data to Eq. 1 is 3.02×10^5 , which implies that hesperidin binds strongly at the active site of the protein.

Therefore, inhibitory effect of hesperidin on the activity of the enzyme was further evaluated using the BACE1 assay kit and interestingly, it has been observed that only 500 nM of hesperidin causes 100% inhibition of BACE1 activity.

Steady-state fluorescence spectroscopy reveals that hesperidin binds at the active site of BACE1. Docking studies with hesperidin and BACE1 illustrate that hesperidin docks close to the catalytic residue and orients itself in a manner such that it blocks the cavity opening, precluding substrate binding.

3.3.8 Preparation of A β (25–35) and Compound Stock Solution

Hesperidin stock solution was prepared in DMSO at 10 mM concentration and was diluted in appropriate concentration using PBS (pH 7.4). A β (25–35) was dissolved in 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) at concentration of 1 mM to produce uniform, non-aggregated A β , immediately stored at -80 °C. The amyloid samples were cryo-lyophilized and dissolved in PBS (pH 7.4) for aggregation study during the experiment. A β (25–35) peptide in

absence and presence of hesperidin in different concentrations were immediately suspended in PBS buffer (pH 7.4) at final concentration of 100 μM and incubated at 37 °C for 3 days.

3.3.9 *Intrinsic Flavonoid Fluorescence Assay*

Incubated A β (25–35) peptide in absence and presence of hesperidin at 1:1 peptide: hesperidin concentration ratio was immediately diluted five times with PBS buffer (pH 7.4) such that the final peptide and hesperidin concentration became 20 μM . In another experiment, A β (25–35) solution incubated for 3 days at 37 °C was diluted with PBS such that the final peptide concentration became 20 μM and to this solution, hesperidin (20 μM) was added. Intrinsic flavonoid fluorescence measurement was done using excitation at $\lambda_{\text{ex}} = 340$ nm and emission measurements were monitored in the wavelength region of 410–520 nm.

A study of the intrinsic fluorescence of hesperidin reveals that in bulk solvent it is non-fluorescent. Upon addition of hesperidin in 3 days aged A β (25–35) solution, a significant rise in fluorescence intensity has been observed. When co-incubated with A β , fluorescence intensity decreases. Since the core of the fibril is highly ordered and essentially hydrophobic in nature therefore in the former the fluorescence increases but in the latter, decrease in fluorescence indicates reduction of fibrillar species in solution.

3.3.10 *ANS Binding Assay*

Reaction mixtures of peptide in presence and absence of hesperidin after 3 days incubation at 37 °C were diluted five times with PBS buffer such that the final peptide concentration became 20 μM . Aliquots of ANS were added from the concentrated stock solution of 1 mM in Tris buffer (pH 8.0) such that the final concentration of ANS was 4 μM . Fluorescence measurements using excitation at $\lambda_{\text{ex}} = 350$ nm and emission measurements in the wavelength region of 410–520 nm were monitored. Fluorescence was corrected by subtracting the appropriate hesperidin fluorescence from the ANS fluorescence of peptide in presence of hesperidin.

3.3.11 *Thioflavin T Assay*

Both reaction mixtures (with and without hesperidin), after 3 days incubation at 37 °C, for fibrillation, were diluted five times with PBS buffer such that the final peptide concentration became 20 μM . Aliquots of Thioflavin T (ThT) were added from the concentrated stock solution of 5 mM ThT in DMSO such that the final concentration of ThT was 15 μM . Fluorescence measurements were carried out using excitation at $\lambda_{\text{ex}} = 440$ nm and emission measurements were monitored in the wavelength region of 460–550 nm.

3.3.12 *Atomic Force Microscope (AFM) Imaging*

A β (25–35) peptide in absence and presence of hesperidin was immediately suspended in PBS buffer (pH 7.4) such that both the peptide and ligand concentrations were 100 μM and incubated at

37 °C for 7 days. Five microlitres of sample was placed on freshly cleaved mica surface for 15 min, rinsed gently with deionized water, and dried in a desiccator. Images were acquired in air using a VEECO VI INNOVA AFM (Bruker AXS Pte. Ltd) operating in tapping mode using antimony doped silicon probes (Bruker).

Hesperidin shows complete inhibition of the amyloid fibril supported by ANS, Thioflavin T binding assay and AFM study. Amyloid fibril formation has not been observed in images obtained from AFM when A β (25–35) peptide was co-incubated with hesperidin for 7 days at 37 °C.

3.3.13 Antioxidant Activity of Hesperidin

ABTS⁺ Radical Scavenging Assay

The method assesses the ABTS⁺ scavenging ability of antioxidant molecules by the change in the characteristic absorption peak of ABTS⁺ at 405 nm [32]. A stable stock solution of ABTS⁺ was prepared by reacting a 7 mM aqueous solution of ABTS with 2.45 mM potassium persulfate (final concentration). The mixture was allowed to stand in the dark at room temperature for 12 h before use. Thereafter, the ABTS⁺ solution mixture was diluted with PBS buffer at pH 7.4 to an absorbance of 0.60 at 405 nm. Varying final concentrations of hesperidin was obtained by adding aliquots from the stock solutions. Percentage inhibition of the test compounds has been calculated using the equation mentioned below.

$$I\% = [(A_{\text{Control}} - A_{\text{Sample}}) / A_{\text{Control}}] \times 100 \quad (2)$$

where A_{Control} is the absorbance of ABTS⁺ solution and A_{Sample} is the absorbance of ABTS⁺ solution in presence of increasing concentration of hesperidin.

DNA Nicking Assay

As described in Satyamitra et al. [32], the experiments were conducted in potassium phosphate buffer (pH 7.4, 50 mM). Briefly, 2 μ l each of ethylenediaminetetraacetic acid (EDTA), phosphate buffer, H₂O₂ (25 mM), FeSO₄ (16 mM), hesperidin (250 μ M) and plasmid pcDNA Hismyc B4 (0.3 μ g/ μ l) were mixed. The final volume of the reaction mixture was brought to 12 μ l with deionized DDW incubated for 1 h at 37 °C in all the cases. After incubation, 2 μ l loading dye (6XEESB-20XTEA buffer, glycerol, bromophenol blue, xylene cyanole and water) was added and 14 μ l of this mixture was loaded onto a 0.7% agarose gel. Electrophoresis was conducted at 60 V in Tris, boric acid, EDTA (TBE) buffer for 2 h. The gel was stained with ethidium bromide (0.5 μ g/ml in DDW) for 30 min and DNA bands were visualized under UV light and photographed using a gel documentation system (Amersham Biosciences, GE Healthcare, UK).

Finally, ABTS⁺ radical scavenging assay demonstrates that hesperidin shows moderate antioxidant potential but DNA nicking assay shows that it can strongly scavenge hydroxyl radicals.

4 Pitfalls

MTDLs that are designed to address two or more targets should potentially lead to additive response if not acting in synergism. Yet there are certain drawbacks and disadvantages which limit its application. It is well understood that high quality negative data chosen from experimentally validated inactive compounds are hard to find. DUD, the online database, or compounds randomly extracted from the commercial database serves the purpose for decoy generation. But this brings in the involvement of some noisy compounds. A second concern is a good classification model. This implies a large and diverse chemical space in the training set and test set. Thus in case of targets with only few ligands or no ligand (such as orphan receptors), structural similarity or inverse docking may be preferred for target fishing.

Finally, compared with docking method, machine learning models and structural similarity cannot directly interpret receptor–ligand interaction, which aids in understanding the mechanism of action and rational structural modification. Thus, an appropriate application for each method, or a combination of different methods, provides a new perspective to overcome their own shortcomings.

References

1. Rutherford GW, Sangani PR, Kennedy GE (2003) Three- or four- versus two-drug antiretroviral maintenance regimens for HIV infection. *Cochrane Database Syst Rev* (4): CD002037
2. Chung KF, Adcock IM (2004) Combination therapy of long-acting β_2 -adrenoceptor agonists and corticosteroids for asthma. *Treat Respir Med* 3(5):279–289
3. Frishman WH, Zuckerman AL (2004) Amlodipine/atorvastatin: the first cross risk factor polypill for the prevention and treatment of cardiovascular disease. *Expert Rev Cardiovasc Ther* 2(5):675–681
4. Zerkak D, Dougados M (2004) Benefit/risk of combination therapies. *Clin Exp Rheumatol* 22(5 Suppl 35):S71–S76
5. Keith CT, Borisy AA, Stockwell BR (2005) Multicomponent therapeutics for networked systems. *Nat Rev Drug Discov* 4(1):71–78
6. Smid P, Coolen HK, Keizer HG et al (2005) Synthesis, structure-activity relationships, and biological properties of 1-heteroaryl-4-[w-(1H-indol-3-yl)alkyl]piperazines, novel potential antipsychotics combining potent dopamine D2 receptor antagonism with potent serotonin reuptake inhibition. *J Med Chem* 48(22):6855–6869
7. Holmes HM, Sachs GA, Shega JW et al (2008) Integrating palliative medicine into the care of persons with advanced dementia: identifying appropriate medication use. *J Am Geriatr Soc* 56(7):1306–1311
8. Van der Schyf CJ, Youdim MB (2009) Multifunctional drugs as neurotherapeutics. *Neurotherapeutics* 6(1):1–3
9. Morphy R, Rankovic Z (2005) Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem* 48(21):6523–6543
10. Morphy R, Rankovic Z (2009) Designing multiple ligands – medicinal chemistry strategies and challenges. *Curr Pharm Des* 15(6):587–600
11. Stahl SM (2009) Multifunctional drugs: a novel concept for psychopharmacology. *CNS Spectr* 14(2):71–73
12. Youdim MB, Van der Schyf CJ (2007) Magic bullets or novel multimodal drugs with various targets for Parkinson's disease? *Nat Rev Drug Discov* 6(6):iii–ivi
13. Bolognesi ML, Matera R, Minarini A et al (2009) Alzheimer's disease: new approaches

- to drug discovery. *Curr Opin Chem Biol* 13 (3):303–308
14. Bolognesi ML, Rosini M, Andrisano V (2009) MTDL design strategy in the context of Alzheimer's disease: from lipocrine to memoguin and beyond. *Curr Pharm Des* 15(6):601–613
 15. Pruss RM (2010) Phenotypic screening strategies for neurodegenerative diseases: a pathway to discover novel drug candidates and potential disease targets or mechanisms. *CNS Neurol Disord Drug Targets* 9(6):693–700
 16. Levy OA, Malagelada C, Greene LA (2009) Cell death pathways in Parkinson's disease: proximal triggers, distal effectors, and final steps. *Apoptosis* 14(4):478–500
 17. Bar-Am O, Amit T, Weinreb O et al (2010) Propargylamine containing compounds as modulators of proteolytic cleavage of amyloid- β protein precursor: involvement of MAPK and PKC activation. *J Alzheimers Dis* 21(2):361–371
 18. Van der Schyf CJ, Geldenhuys WJ, Youdim MB (2006) Multifunctional drugs with different CNS targets for neuropsychiatric disorders. *J Neurochem* 99(4):1033–1048
 19. Chertkow Y, Weinreb O, Youdim MB et al (2009) Molecular mechanisms underlying synergistic effects of SSRI-antipsychotic augmentation in treatment of negative symptoms in schizophrenia. *J Neural Transm* 116 (11):1529–1541
 20. Huang W, Tang L, Shi Y, Huang S, Xu L, Sheng R, Wu P, Li J, Zhou N, Hu Y (2011) Searching for the multi-target-directed ligands against Alzheimer's disease: discovery of quinoxaline-based hybrid compounds with AChE, H₃R and BACE 1 inhibitory activities. *Bioorg Med Chem* 19(23):7158–7167
 21. Fang J, Li Y, Liu R, Pang X, Li C, Yang R, He Y, Lian W, Liu AL, Du GH (2015) Discovery of multi target-directed ligands against Alzheimer's disease through systematic prediction of chemical-protein interactions. *J Chem Inf Model* 55(1):149–164
 22. Rosini M, Simoni E, Minarini A, Melchiorre C (2014) Multi-target design strategies in the context of Alzheimer's disease: acetylcholinesterase inhibition and NMDA receptor antagonism as the driving forces. *Neurochem Res* 39(10):1914–1923
 23. Luo Z, Sheng J, Sun Y, Lu C, Yan J, Liu A, Luo HB, Huang L (2013) Synthesis and evaluation of multi-target-directed ligands against Alzheimer's disease based on the fusion of donepezil and ebselen. *J Med Chem* 56(22):9089–9099
 24. Chakraborty S, Bandyopadhyay J, Chakraborty S, Basu S (2016) Multi-target screening mines hesperidin as a multi-potent inhibitor: implications in Alzheimer's disease therapeutics. *Eur J Med Chem* 121:810–822
 25. Hyperchem (2002) Hypercube, Inc., USA
 26. Chakraborty S, Basu S (2015) Structural insight into the mechanism of amyloid precursor protein recognition by β -secretase 1: a molecular dynamics study. *Biophys Chem* 202:1–12
 27. Chakraborty S, Kumar S, Basu S (2011) Conformational transition in the substrate binding domain of β -secretase exploited by NMA and its implication in inhibitor recognition: BACE1-myricetin a case study. *Neurochem Int* 58(8):914
 28. CODESSA [computer program] (1995) Version 2.7.16. Comprehensive descriptor for structure and statistical analysis. Copyright center of heterocyclic chemistry, University of Florida and Institute of Chemical Physics, University of Tartu, Estonia and SemiChem Inc
 29. AMPAC 9.0, © (1994) Semichem, 7128 Summit, Shawnee, KS 66216
 30. Chakraborty S, Basu S (2014) Insight into the anti-amyloidogenic activity of polyphenols and its application in virtual screening of phytochemical database. *Med Chem Res*. <https://doi.org/10.1007/s00044-014-1081-2>
 31. Chakraborty S, Basu S (2014) Mechanistic insight into the radical scavenging activity of polyphenols and its application in virtual screening of phytochemical library: an in silico approach. *Eur Food Res Technol*. <https://doi.org/10.1007/s00217-014-2285-x>
 32. Satyamitra M, Mantena S, Nair CKK, Chandna S, Dwarakanath BS, Devi PU (2014) The antioxidant flavonoids, orientin and vicenin enhance repair of radiation-induced damage. *SAJ Pharm Pharmacol* 1:1–9



Computational Design of Multi-target Kinase Inhibitors

Sinoy Sugunan and Rajanikant G. K.

Abstract

As key regulators of every aspect of cell function, protein kinases are frequently associated with various human diseases. Therefore, protein kinase inhibition has become the second most important group of drug targets, after G-protein-coupled receptors. Owing to the complex and polygenic nature of diseases, designing multi-kinase small molecule inhibitors as potential therapies is gaining major consideration. Effective *in silico* drug design strategies are desired to identify multi-target kinase inhibitors. In this chapter, we summarize the two such effective computational strategies reported by our group to identify multi-target kinase inhibitors.

Keywords Ensemble pharmacophore, Molecular docking, Molecular dynamics simulation, Multi-target inhibitor, Protein kinase, Virtual screening

1 Introduction

Protein kinases play a critical role in signal transduction and regulation of a range of cellular activities and hence have established themselves as promising drug targets for the treatment of wide variety of human diseases [1–3]. The notion of kinase inhibition was originated during the 1950s and 1960s, when multiple studies on protein kinases characterization and elucidation of their signaling cascade were initiated [1, 3]. The strategy of kinase inhibition was introduced during late 1980s, when inhibitors against the epidermal growth factor receptor (EGFR) were reported [4]. Subsequently, a large number of kinase inhibitors based on diverse structural frameworks and pharmacological profiles have been reported [1, 5–8]. Currently, there are around 46 approved small molecule kinase inhibitors in the market, along with these inhibitors; a large number of molecules are undergoing clinical trials at different stages [9–11].

Amid 46 FDA approved drugs affecting the human kinome, 15 are multi-target drugs, and 14 are mainly used in combination with other drugs [9]. The approved multi-target kinase inhibitors manage the multifaceted nature of complex diseases by simultaneously modulating multiple targets. The multi-target inhibitors provide an effective approach to avoid the untoward effects caused by drug–drug interactions [9, 12, 13], but it is very challenging to design an effective multi-target molecule. Even though the combination of individual

single-target drugs represents the most clinically practiced therapeutic tactics, the modulation of multiple targets by a single molecule has an improved chance of affecting the multifarious equilibrium of complete cellular networks than drugs acting on a single target [9, 14].

Computer-aided drug design methodologies are being used extensively in drug discovery and development arena [15]. To design and identify novel small molecules as effective therapeutics, various promising computational strategies have been formulated, depending on the purpose and target system. Two important *in silico* approaches in the drug discovery process are ligand- and structure-based design. These potent approaches are widely applied in virtual screening for lead identification and optimization. Considering the potential of *in silico* approaches in drug discovery process, various studies have reported the use of computational methodology to derive multi-target molecules with appreciable success [16–18]. This chapter discusses the two such effective *in silico* strategies reported by our group to identify multi-target kinase inhibitors.

The first approach involved a combination of merged pharmacophore matching, database screening, and molecular docking to reliably identify the potential multi-target kinase inhibitors [19] (Fig. 1). The second strategy employed ensemble pharmacophore-based screening (EPS) of a compound database, post-EPS filtration (PEPSF) of the ligand hits, and multiple dockings. Consequently, the binding affinity of small molecules with the desired protein structures was validated by molecular dynamics (MD) simulation [20]

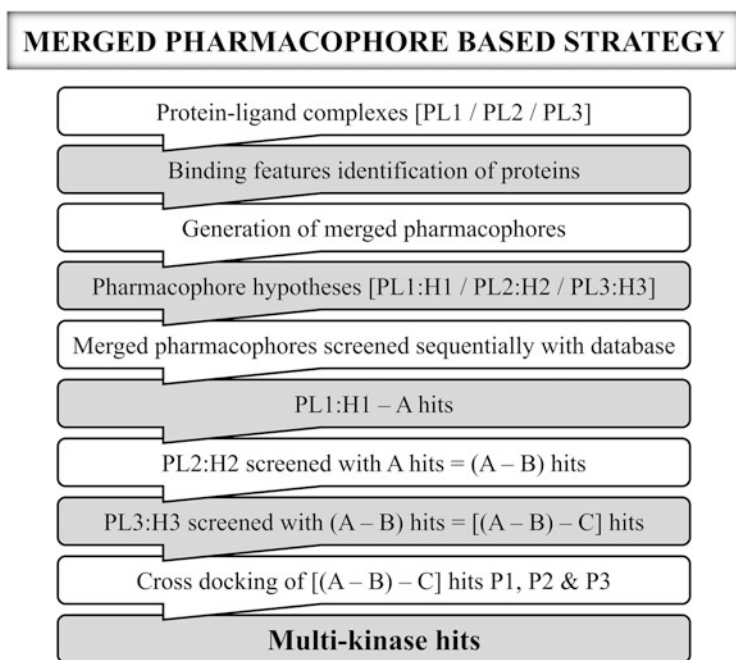


Fig. 1 Merged pharmacophore-based strategy to identify multi-target kinase inhibitors

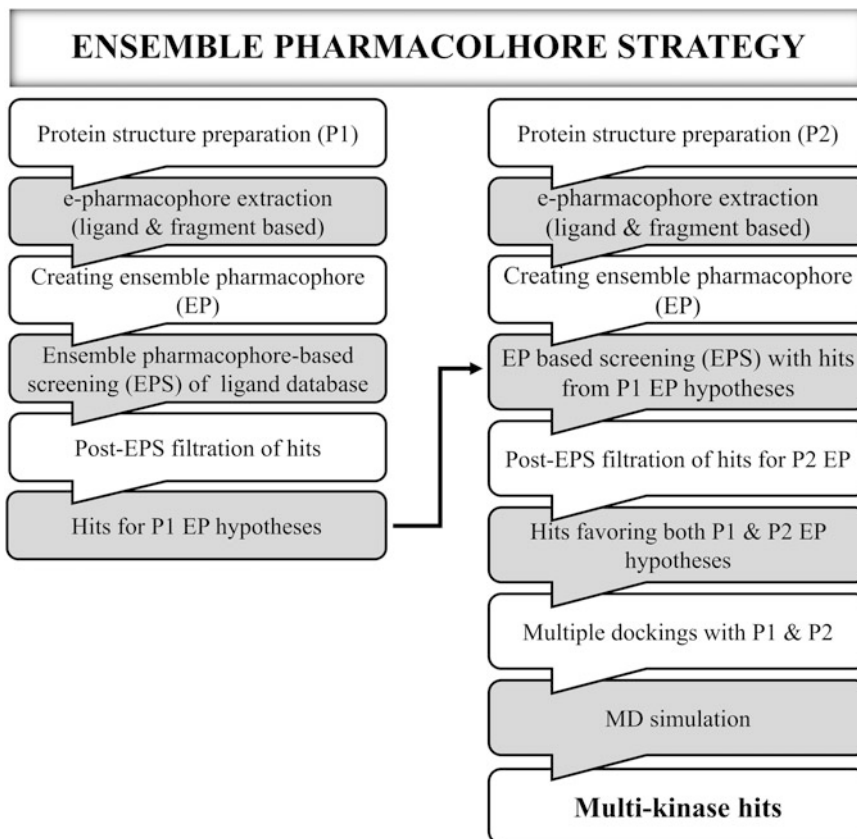


Fig. 2 Ensemble pharmacophore-based strategy to identify multi-target kinase inhibitors

(Fig. 2). These strategies were effective in mining structurally diverse inhibitors through combined pharmacophoric features, which in turn increased the speed of the virtual screening. In general, the collection of multiple protein structures for pharmacophore extraction is considered as an improvement over the standard technique of using a single structure. The study by Zou et al. [21] had shown that multicomplex-based pharmacophore model derived from many holoprotein (protein–ligand complex) structures would determine all of the key protein–ligand interactions. It was also demonstrated that an ensemble pharmacophore derived from multiple apoprotein (protein only) and holoprotein structures could help in extracting all of the features available at the binding site of the protein [21]. The usage of merged and ensemble pharmacophore for screening of multi-target ligands is preferred when the conformations of protein structures vary considerably and ligands dock to them with different pharmacophoric features. This in turn would effectively screen ligands with different scaffolds.

2 Materials

1. High-end computer work station
2. Small molecule databases such as ZINC
3. Schrödinger Suite (module: Prime, Glide, E-pharmacophore, Ligprep, and Phase)
4. SMAP webservice
5. Molecular dynamics (MD) simulation software such as Gromacs

3 Methods

3.1 Merged Pharmacophore Strategy

3.1.1 Protein Structure Preparation

The protein–ligand complexes of proteins are retrieved from the PDB and are prepared using the protein preparation wizard of Schrödinger (Schrödinger, LLC, New York, NY, USA). Missing hydrogens are added, proper bond orders are assigned, and water molecules further than 5 Å from the heterogeneous groups are deleted. The H-bonds are then optimized and the protein structures should be minimized to the default root mean square deviation (RMSD) value of 0.30 Å. The minimized protein structures will be further considered for pharmacophore analysis and docking.

3.1.2 Analysis of Active Site

The active sites of proteins could be analyzed by selecting neighboring region within 5 Å of their respective ligands. This analysis could give information about amino acids involved in H-bond and hydrophobic interactions with ligands. It could reveal information of shared amino acid features present in proteins of the same family. Potential of this analysis is reported in our study related to screening of multi-target inhibitors against DAP kinases [19].

3.1.3 Superimposition of Binding Sites

The binding sites are compared by superimposing the proteins. This could be achieved by using the SMAP software package [19, 22]. The proteins should be superimposed in such a way that the binding sites occupy approximately the same region of space. It could also provide information about position of active and conservation among all the members of a protein family. Significantly, SMAP server provides vital data regarding ligand-binding sites and feasibility of identifying common ligands for different proteins, by comparing the binding site similarity [19]. The server computes p -value, raw alignment score, RMSD, and Tanimoto coefficient of overlap for each of the hits. The hits producing SMAP p value $< 1.0e-3$ and high Tanimoto coefficient (near to one) could be considered significant. We have calculated these values for DAP kinases and the results are mentioned in a Table of our previously reported study [19].

3.1.4 *Generation and Evaluation of Pharmacophore*

Pharmacophores could be generated for each protein–ligand complex using an e-pharmacophore script of Schrödinger Suite [19]. The script mines the energetic descriptors of the Glide XP score and allocates them to pharmacophore features. The resulting energy e-pharmacophore hypotheses can be opened in a Phase search of Schrödinger or could be modified manually based on energetics and visualization. In order to generate the hypotheses, Phase uses a default set of six chemical features: namely, H-bond acceptor (A), H-bond donor (D), hydrophobic (H), negative ionizable (N), positive ionizable (P), and aromatic ring (R) site. The Phase definitions could be edited so that the positive and negative groups are also labeled as donors or acceptors, respectively. The successful generation of merged pharmacophore and alignment of ligands to pharmacophore for DAP kinases was presented in Fig. 4 of previously reported study [19].

3.1.5 *Preparation of Ligand Dataset*

The strategy of merged pharmacophore is mainly useful if the ligand dataset is small. These ligands are prepared using the LigPrep module of Schrödinger. Ligands are protonated and subjected to removal of salt, ionization, and generation of low-energy ring conformations. The chiralities of the original compounds should be preserved. Lastly, low-energy 3D structures of all the ligands will be produced.

3.1.6 *Multiple Merged Pharmacophore-Based Ligand Screening*

To identify the multi-targeting kinase inhibitors, the dataset of ligands should be screened in a sequential way with all merged pharmacophore hypotheses. Constraints given during this process should be based on shared features among the pharmacophores. We have effectively screened dataset with merged pharmacophore in our previously described study [19].

3.1.7 *Cross-Docking of Ligand Hits to Proteins*

The ligand hits retrieved through screening of the ligand dataset, using merged pharmacophores of proteins, are docked into the binding sites of each protein. Glide Extra Precision (XP) present in the Glide module of Schrödinger is used for these docking studies. All Glide protocols should run on default parameters. Throughout the docking process, glide scoring function (*G*-score) is used to select the best conformation for each ligand. The similar docking study of DAP kinases was reported earlier and the results were presented in a Table of our previous study [19].

3.2 **Ensemble Pharmacophore-Based Strategy**

All protein–ligand complexes should be prepared as mentioned earlier in Sect. 3.1.1 of this chapter.

3.2.1 *Protein Structure Preparation*

3.2.2 Receptor Grid Generation

Glide energy grids are produced for all of the prepared protein structures through Receptor Grid Generation panel. The co-crystallized ligand has to be differentiated from the active site of the receptor. The grid is defined by a rectangular box surrounding the co-crystallized ligand. The atoms should be scaled by van der Waals radii of 1.0 Å with the partial atomic charge less than 0.25 defaults. No constraints have to be defined. The grid should cover the active site of the protein. These grids are to be engaged in all of the docking studies.

3.2.3 E-Pharmacophore Extraction

Structure-based pharmacophores could be created from protein structures using the energy-optimized pharmacophore (e-pharmacophore) script of Schrödinger [20]. Both ligand- and fragment-based e-pharmacophores have to be extracted. During ligand-based method, various protein structures are treated as holoproteins and during fragment-based method, they are treated as apoproteins. Pharmacophore features from holoproteins are extracted through ligand-based e-pharmacophore approach. In this step, ligands are separated from their respective receptors and then re-docked using the Score in Place option in Glide XP.

Extraction of pharmacophoric features from apoproteins is achieved through fragment-based e-pharmacophores. To extract these pharmacophores, the glide fragment library is docked to each of the proteins using Glide XP. In our study, glide fragment library consisted of a set of 441 unique small fragments (1–7 ionization/tautomer variants; 6–37 atoms; MW range 32–226), derived from molecules in the medicinal chemistry literature [20].

The script extracts the energetic descriptors, and the later steps are carried as mentioned earlier in this chapter. The generated pharmacophore sites and the Glide XP energies from the atoms that comprised each pharmacophore sites are compiled. Then, these sites are ranked based on the individual energies. The successful e-pharmacophore extraction was carried out on RIPK1 protein and was mentioned in Figs. 6a–d and 7 of our previously described study [20].

3.2.4 Ensemble Pharmacophore Construction

The ligand- and the fragment-based e-pharmacophores obtained from the proteins should be carefully evaluated for their diverse features. In order to retrieve all these features in the final pharmacophore, all the e-pharmacophores are superposed and their coordinates are manually compiled into a single file to generate an ensemble pharmacophore. If common features corresponding to the same amino acid feature in the binding site are present, then they are considered as a single feature and only one of them was represented in the ensemble pharmacophore. The results of ensemble pharmacophore construction for RIPK1 and RIPK3 are presented in Tables of our previously reported study [20].

3.2.5 Ensemble Pharmacophore-Based Screening for Proteins

Ensemble pharmacophore obtained from the previous step is screened against ligand databases such as ZINC (~2 million unique structure records) to retrieve lead-like compounds [31], using the Phase module. The database molecules are filtered explicitly with a distance-matching tolerance of 2.0 Å, and matching of a minimum of four sites is required. The database hits are ranked based on their fitness scores, which measures how well the aligned ligand conformers matched the hypothesis.

3.2.6 Post-Pharmacophore-Based Screening Filtration

To filter out the false positives from EPS hits, PEPSF has to be carried out, which separates the compounds into groups with features from individual pharmacophores. The output of protein EPS gives information on the pharmacophore features present in each ligand hit that matched with the protein ensemble pharmacophore features. The EPS result retrieved in the form of an Excel worksheet, and the ligand hits are arranged based on their matching sites. These sites are compared with each of the protein fragment-based e-pharmacophores, and the compounds with sites that matched with any of these pharmacophores are considered. Compounds with sites that matched with more than one pharmacophore are identified as false positives since they contained cross-features, and removed from the study. PEPSF could eliminate a large number of compounds, greatly reducing the computational time required for the docking step. The compounds retrieved through PEPSF of a protein are subjected to EPS and PEPSF with the second protein. These compounds will contain pharmacophore features of both proteins (like RIPK1 and RIPK3 in our study) and will not harbor any cross-features [20].

3.2.7 Multiple Dockings

The multiple groups of compounds generated from PEPSF steps are subjected to multiple dockings with their respective protein structures. The resultant top-ranked compounds are expected to show good fitness and Glide scores and also exhibit good interactions with the desired amino acids of both the proteins.

Docking is performed with the Glide module using the OPLS 2005 force field. Default settings are selected for all of the docking calculations, and the Glide XP descriptor information is used to construe energy terms such as H-bond interactions, electrostatic interactions, hydrophobic enclosure, and π - π stacking interactions (Glide v.5.7, Schrodinger, LLC). At final stage, post-docking minimization is carried out to augment the ligand geometries.

3.2.8 Molecular Dynamic Simulation of Lead Compounds

Based on docking results, top hit compounds could be subjected to molecular dynamics (MD) simulation studies with both the protein structures to validate binding affinities of compounds. MD simulations may be carried out using the Gromacs v4.5.5 software with Gromos53a6 force field [20, 23].

Table 1

It describes the few notable computational strategies reported by various researchers to identify multi-target kinase inhibitors

Strategy adopted	References
Docked with targets separately	[24–26]
Docked with targets separately and comparison of binding sites	[27]
Neural network-QSAR based strategy	[28]
QSAR based strategy	[29]
mtSAR approach	[30]
Machine learning–datasets–structure-based drug designing approach	[31]
CoreSiMMap based approach	[32]

4 Conclusion

In this chapter, we have described two computational strategies to screen and identify multi-target kinase inhibitors, which were successful in identifying multi-kinase inhibitors [19, 20]. According to us, both strategies are effective in identifying multi-target inhibitors in a swift way, the choice of strategy depends largely on size of ligand database and on discretion regarding the curation of pharmacophoric features. Furthermore, other computational approaches that have been utilized successfully for screening small molecule multi-target kinase inhibitors are listed in Table 1.

5 Notes

1. A molecular dynamic simulation step could also be included in merged pharmacophore-based strategy to verify the ligand–target binding efficiency.
2. Comparatively, ensemble pharmacophore-based strategy is potent in screening big ligand databases and PEPSF step shortens the time consumed in screening. This strategy compiles the pharmacophoric features in more accommodative manner, as it considers both ligand- and fragment-based pharmacophores.
3. The speed of whole screening process also depends upon the computer-hardware configuration.
4. Modeled protein structures could also be used for ensemble pharmacophore generation as described in our previous study [20].

References

1. Wu P, Nielsen TE, Clausen MH (2016) Small-molecule kinase inhibitors: an analysis of FDA-approved drugs. *Drug Discov Today* 21:5–10
2. Zhang J, Yang PL, Gray NS (2009) Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* 9:28–39
3. Wu P, Nielsen TE, Clausen MH (2015) FDA-approved small-molecule kinase inhibitors. *Trends Pharmacol Sci* 36:422–439
4. Yaish P, Gazit A, Gilon C et al (1988) Blocking of EGF-dependent cell proliferation by EGF receptor kinase inhibitors. *Science* 242:933–935
5. Divito SJ, Kupper TS (2014) Inhibiting Janus kinases to treat alopecia areata. *Nat Med* 20:989–990
6. Hendriks RW, Yuvaraj S, Kil LP (2014) Targeting Bruton's tyrosine kinase in B cell malignancies. *Nat Rev Cancer* 14:219–232
7. Wu P, Hu Y (2012) Small molecules targeting phosphoinositide 3-kinases. *MedChemComm* 3:1337–1355
8. Wu P, Hu YZ (2010) PI3K/Akt/mTOR pathway inhibitors in cancer: a perspective on clinical progress. *Curr Med Chem* 17:4326–4341
9. Li YH, Wang PP, Li XX et al (2016) The human kinome targeted by FDA approved multi-target drugs and combination products: a comparative study from the drug-target interaction network perspective. *PLoS One* 11: e0165737
10. Rask-Andersen M, Zhang J, Fabbro D et al (2014) Advances in kinase targeting: current clinical use and clinical trials. *Trends Pharmacol Sci* 35:604–620
11. Yang H, Qin C, Li YH et al (2016) Therapeutic target database update 2016: enriched resource for bench to clinical drug target and targeted pathway information. *Nucleic Acids Res* 44: D1069–D1074
12. Boran ADW, Iyengar R (2010) Systems approaches to polypharmacology and drug discovery. *Curr Opin Drug Discov Devel* 13:297–309
13. Dar AC, Das TK, Shokat KM et al (2012) Chemical genetic discovery of targets and anti-targets for cancer polypharmacology. *Nature* 486:80–84
14. Korcsmáros T, Szalay MS, Böde C et al (2007) How to design multi-target drugs: target search options in cellular networks. *Expert Opin Drug Discov* 2:799–808
15. Zhang S (2011) Computer-aided drug discovery and development. In: Satyanarayanan SD (ed) *Drug design and discovery: methods and protocols*. Humana Press, Totowa, NJ, pp 23–38
16. Ma XH, Shi Z, Tan C et al (2010) In-silico approaches to multi-target drug discovery. *Pharm Res* 27:739–749
17. Clemente JC, Govindasamy L, Madabushi A et al (2006) Structure of the aspartic protease plasmeprin 4 from the malarial parasite *Plasmodium malariae* bound to an allophenylnorstatine-based inhibitor. *Acta Crystallogr Sect D* 62:246–252
18. Wei D, Jiang X, Zhou L et al (2008) Discovery of multitarget inhibitors by combining molecular docking with common pharmacophore matching. *J Med Chem* 51:7882–7888
19. Nair SB, Fayaz SM, Rajanikant GK (2013) A novel multi-target drug screening strategy directed against key proteins of DAPK family. *Comb Chem High Throughput Screen* 16:449–457
20. Fayaz SM, Rajanikant GK (2015) Ensembling and filtering: an effective and rapid in silico multitarget drug-design strategy to identify RIPK1 and RIPK3 inhibitors. *J Mol Model* 21:314
21. Zou J, Xie HZ, Yang SY et al (2008) Towards more accurate pharmacophore modeling: multicomplex-based comprehensive pharmacophore map and most-frequent-feature pharmacophore model of CDK2. *J Mol Graph Model* 27:430–438
22. Ren J, Xie L, Li WW et al (2010) SMAP-WS: a parallel web service for structural proteome-wide ligand-binding site comparison. *Nucleic Acids Res* 38:441–444
23. Fayaz SM, Rajanikant GK (2014) Ensemble pharmacophore meets ensemble docking: a novel screening strategy for the identification of RIPK1 inhibitors. *J Comput Aided Mol Des* 28:779–794
24. Cui Z, Chen S, Wang Y et al (2017) Design, synthesis and evaluation of azaacridine derivatives as dual-target EGFR and Src kinase inhibitors for antitumor treatment. *Eur J Med Chem* 136:372–381
25. Cui Z, Li X, Li L et al (2016) Design, synthesis and evaluation of acridine derivatives as multi-target Src and MEK kinase inhibitors for anti-tumor treatment. *Bioorg Med Chem* 24:261–269
26. Reker D, Seet M, Pillong M et al (2014) Deorphaning pyrrolopyrazines as potent multi-target antimalarial agents. *Angew Chem Int Ed* 53:7079–7084

27. Rehan M (2015) A structural insight into the inhibitory mechanism of an orally active PI3K/mTOR dual inhibitor, PKI-179 using computational approaches. *J Mol Graph Model* 62:226–234
28. Ajmani S, Viswanadhan VN (2013) A neural network-based QSAR approach for exploration of diverse multi-tyrosine kinase inhibitors and its comparison with a fragment-based approach. *Curr Comput Aided Drug Des* 9:482–490
29. Marzaro G, Chilin A, Guiotto A et al (2011) Using the TOPS-MODE approach to fit multi-target QSAR models for tyrosine kinases inhibitors. *Eur J Med Chem* 46:2185–2192
30. Wassermann AM, Peltason L, Bajorath J (2010) Computational analysis of multi-target structure-activity relationships to derive preference orders for chemical modifications toward target selectivity. *ChemMedChem* 5:847–858
31. Allen BK, Mehta S, Ember SWJ et al (2015) Large-scale computational screening identifies first in class multitarget inhibitor of EGFR kinase and BRD4. *Sci Rep* 5:16924
32. Hsu K-C, Cheng W-C, Chen Y-F et al (2012) Core site-moiety maps reveal inhibitors and binding mechanisms of orthologous proteins by screening compound libraries. *PLoS One* 7:e32142



Proteochemometrics for the Prediction of Peptide Binding to Multiple HLA Class II Proteins

Ivan Dimitrov, Ventsislav Yordanov, Darren R. Flower, and Irini Doytchinova

Abstract

Proteochemometrics (PCM) is a method for deriving quantitative structure–activity relationships (QSAR). It models the bioactivity of multiple ligands against multiple target proteins. Thus, PCM is a key method used to develop multi-target drug molecules. In the present protocol, the PCM method is applied to a set of peptides binding to seven polymorphic HLA class II proteins from locus DP. The peptides bind to an open-ended binding site on DP proteins which accepts a nonameric binding core. As the peptides in the studied set are 15-mers, the initial set is processed by an iterative self-consistent algorithm to select the most probable binding nonamer for each 15-mer peptide. The final set containing the most probable binding nonamers is used to derive the PCM model. The model is validated by external set of proteins and enters the server EpiTOP which is freely accessible at <http://www.ddg-pharmfac.net/EpiTOP3/>.

Keywords Proteochemometrics, HLA class II binding prediction, Partial least squares, EpiTOP

1 Introduction

The main steps of the immune response mounted by the host after microbiological invasion are: antigen processing in antigen-presenting cells (APCs), presentation on the cell membrane of foreign peptide fragments, and recognition by the cells of the host immune system. A major role in this process is played by major histocompatibility complex (MHC) proteins, a family of proteins that bind antigens and present them on the cell surface. In humans, MHC proteins are often referred to as HLA (human leukocyte antigen). There are several classes of HLA proteins, the most important of which are class I and class II [1]. HLAs are the most polymorphic proteins in higher vertebrates, with more than 17,000 class I and class II molecules listed in IMGT/HLA [2]. The HLA class II proteins include three main loci: DR, DQ, and DP and consist of two chains: α and β . Each chain contains one variable and one constant domain. The variable domain forms the peptide binding site where the foreign peptide fragments bind to the HLA molecule. The binding site itself accepts only nine residues but it

is open at both ends and peptides longer than nine amino acids are able to bind (usually between 13 and 25 residues).

Quantitative structure–activity relationship (QSAR) is a drug discovery method which uses statistical approaches to model the relationship between the chemical structure of a ligand and its biological activity, typically its affinity for a given target biomacromolecule. This method is based on the principle of similarity of compounds—or active analogue approach (similar structures—similar activity); it is often used for lead optimization. A drawback of this method is the consideration of the ligand structure only. So, QSAR is not able to model interactions with multiple biological targets adequately [3–5].

Proteochemometrics (PCM) is a QSAR method developed by Lapinsh et al. [6] to model simultaneously the bioactivity of multiple ligands against multiple target proteins [6, 7]. PCM can be seen as a generalization of QSAR, combining information from ligands and target molecules into a single matrix, allowing extrapolation of results and predicting biological activity of new compounds to new target molecules [5]. The ability of PCM to link such data mimics inductive learning, a process from artificial intelligence, which acquires knowledge from one system and uses it to solve a related problem in a similar system. PCM works with multiple target molecules, which is the key in the development of multi-target drug molecules [4]. A drawback of PCM is its dependence on the assumed variability of ligand and target molecules; if an amino acid plays an important role in a given interaction but remains constant across the dataset, PCM will struggle to evaluate its importance [8].

In the present protocol, we apply the PCM to derive a model describing the interactions between peptides of foreign origin and seven polymorphic HLA-DP proteins. The model is implemented in a server for HLA binding prediction EpiTOP, which is freely accessible via the web.

2 Datasets

2.1 HLA-DP Proteins

The HLA class II proteins considered in the protocol presented here are from locus DP:

- DP1 (DPA*02:01/DPB1*01:01),
- DP2 (DPA*01:03/DPB1*02:01),
- DP3 (DPA*01:03/DPB1*03:01),
- DP41 (DPA*01:03/DPB1*04:01),
- DP42 (DPA*01:03/DPB1*04:02, assigned as DP42a and DPA*03:01/DPB1*04:02 assigned as DP42b), and
- DP5 (DPA*02:01/DPB1*05:01).

Taken together, these DP proteins are found in >90 % of the human population [9]. The peptide binding site is formed by the first 80 residues of α chain and the first 90 residues of β chain (Fig. 1) [10–12]. Among them are four polymorphic residues from α chain (Ala/Met^{11 α} , Met/Gln^{31 α} , Gln/Arg^{50 α} , and Leu/Ser^{66 α}) and 15 polymorphic ones from β chain (Val/Leu^{8 β} , Tyr/Phe^{9 β} , Gly/Leu^{11 β} , Tyr/Phe/Leu^{35 β} , Ala/Val^{36 β} , Ala/Asp/Glu^{55 β} , Ala/Glu^{56 β} , Glu/Asp^{57 β} , Ile/Leu^{65 β} , Lys/Glu^{69 β} , Val/Met^{76 β} , Asp/Gly^{84 β} , Glu/Gly^{85 β} , Ala/Pro^{86 β} , and Val/Met^{87 β}).

The peptide binding site on the DP proteins consists of 54 residues: 25 residues belong to α chain (9, 10, 11, 22, 24, 31, 32, 43, 52, 53, 54, 55, 58, 59, 61, 62, 63, 65, 66, 68, 69, 70, 72, 73, and 76), and 29 residues are from β chain (9, 11, 12, 13, 24, 26, 27, 28, 35, 36, 45, 54, 55, 58, 59, 65, 68, 69, 72, 75, 76, 77, 79, 80, 81, 82, 83, 84, and 88) (Fig. 2). The contacts between the bound peptide and the protein are found by the software Chimera (UCSF) and a combinatorial library of 172 peptides (9 positions \times 19 amino acids + 1 parent peptide from the X-ray structure) [12]. Among the contact residues, only three residues from α chain (Ala/Met^{11 α} ,

AA Pos.	10	20	30	40	50	60	70	80
DPA1*01:03	IKADHVSTYA	AFVQTHRPTG	EFMFEPDEDE	MFYVDLOKKE	TVVHLEEPGQ	AFSFEAQGGL	ANIAILNNNL	NTLIQRSNHT
DPA1*02:01	-----	-----	-----	Q-----	-----R	-----	-----	-----
DPA1*03:01	**-----	M-----	-----	-----	-----	-----	-----S	-----

AA Pos.	10	20	30	40	50	60	70	80	90
DPB1*01:01	RATPENYYVQ	GRQECYAFNG	TQRFLERYIY	NREEYARFDS	DVGEFRAVTE	LGRPAASYNN	SQKDILEEKR	AVPDRVCRNN	YELDEAVTLQ
DPB1*02:01	-----LF-	-----	-----	---FV---	-----	---DE---	---E---	---M---	---GGFM---
DPB1*03:01	-----	L-----	-----	---FV---	-----	---DED---	---L---	-----	-----
DPB1*04:01	-----LF-	-----	-----	---F---	-----	---DE---	-----	---M---	---GGFM---
DPB1*04:02	-----LF-	-----	-----	---FV---	-----	---DE---	-----	---M---	---GGFM---
DPB1*05:01	-----LF-	-----	-----	---LV---	-----	---E---	-----	---M---	-----

Fig. 1 Protein sequences of α and β chains forming the peptide binding site on HLA-DP. Data are taken from IMGT/HLA database (<http://www.ebi.ac.uk/imgt/hla>)

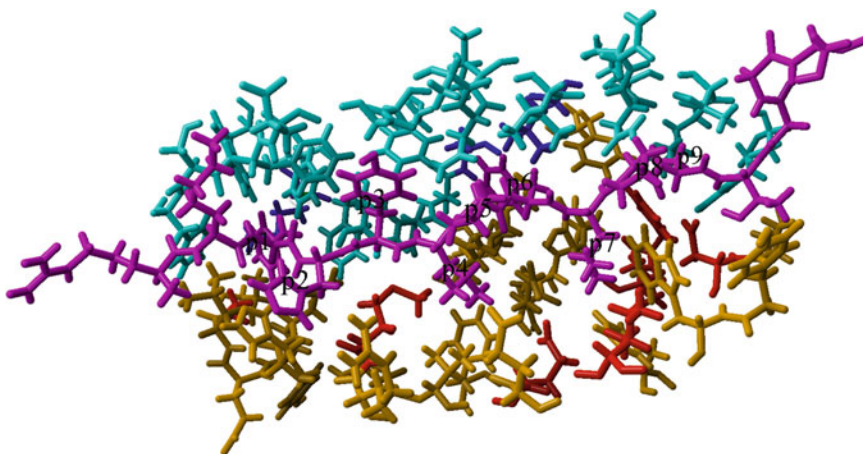


Fig. 2 A peptide (purple) bound in the peptide binding site on HLA-DP2 protein (pdb code: 3lqz). The α chain is given in cyan, the β chain in yellow. The polymorphic residues are shown in blue and red, respectively

Table 1
Contacts between peptide and HLA-DP protein

Peptide position	Protein position
p1	31 α , 76 β , 84 β
p2	76 β
p3	76 β
p4	69 β , 76 β
p5	–
p6	11 α , 66 α , 11 β
p7	66 α , 55 β , 65 β , 69 β
p8	55 β
p9	9 β , 35 β , 36 β , 55 β

Met/Gln^{31 α} , and Leu/Ser^{66 α}) and nine residues from β chain (Tyr/Phe^{9 β} , Gly/Leu^{11 β} , Tyr/Phe/Leu^{35 β} , Ala/Val^{36 β} , Ala/Asp/Glu^{55 β} , Ile/Leu^{65 β} , Lys/Glu^{69 β} , Val/Met^{76 β} , and Asp/Gly^{84 β}) are polymorphic (Table 1). Asp^{55 β} forms a salt bridge with the peptide Ser9.

2.2 Peptides Binding to HLA-DP Proteins

A set of 4304 15-mer peptides binding to the seven HLA-DP proteins was collected from Immune Epitope Database (IEDB) (<http://www.iedb.org>) in February 2016 at the following settings: alleles—HLA-DP, assay—purified/competitive/radioactivity, and units—IC₅₀ nM. In case of overlapping peptides of different lengths, the longest peptide was considered. The affinity was presented as $\text{pIC}_{50} = \log(1/\text{IC}_{50})$. Duplicates were removed. The final set consisted of 3864 15-mer peptides. They were grouped by allele. Each 15-mer peptide was presented as a set of overlapping 9-mers as is shown in Fig. 3.

2.3 Proteins Used for Model Validation

Some of the peptides considered in the present study have natural origin, others are synthetically derived. The 72 parent proteins of the naturally occurring peptides were collected and used to validate the PCM model.

3 Methods

3.1 Proteochemometrics

1. The chemical structure of ligands and proteins in PCM is described by three descriptor blocks: ligand (L), protein (P), and ligand–protein (LP) (Fig. 4) [7]. The chemical structure of the peptides and proteins used in the present protocol is described by three z -scales obtained by Hellberg et al.

15-mer peptide	pIC ₅₀	Overlapping 9-mer peptides
GGSILKISNKYHTKG	7.369	G G S I L K I S N
	7.369	G S I L K I S N K
	7.369	S I L K I S N K Y
	7.369	I L K I S N K Y H
	7.369	L K I S N K Y H T
	7.369	K I S N K Y H T K
	7.369	I S N K Y H T K G

Fig. 3 Presentation of a 15-mer peptide as a set of overlapping 9-mers. Each 9-mer takes the pIC₅₀ value of the parent 15-mer

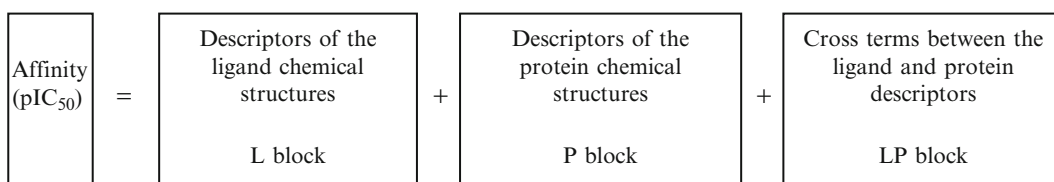


Fig. 4 Blocks of descriptors used in proteochemometrics (PCM)

[13, 14]. They describe three main properties of amino acids: the hydrophobicity (described by z_1), size (described by z_2), and electronic properties (described by z_3) (Table 2). Amino acids with similar properties show similar z values: aromatic amino acids or long chain aliphatic chains show negative values for z_1 (Ile, Leu, Phe, Trp, etc.); glycine is the smallest amino acid and therefore has the lowest value for z_2 ; for descriptor z_3 has a positive correlation with the negative charge in molecule (Arg and Lys are positively charged and have the lowest values).

- Each 15-mer is presented as a set of overlapping nonamers. Thus, from 3094 15-mers are generated 21,658 nonamer peptides. Each nonamer is described by 27 descriptors (9 positions \times 3 z -scales). They form the L block.
- The 19 polymorphic residues from α and β chain also are described by the z -scales (19 positions \times 3 z -scales = 57 descriptors) and form the P block.
- The peptide-protein contact residues given in Table 1 are described by 57 descriptors (19 contacts \times 3 z -scales) and form the LP block.

The final X matrix consists of 141 descriptors grouped into three blocks: L, P, and LP. This matrix is used by the iterative self-consistent algorithm to identify the most probable binding core of each 15-mer.

Table 2
z-scales [13]

ακ	z_1	z_2	z_3
Ala	0.07	-1.73	0.09
Arg	2.88	2.52	-3.44
Asn	3.22	1.45	0.84
Asp	3.64	1.13	2.36
Cys	0.71	-0.97	4.13
Gln	2.18	0.53	-1.14
Glu	3.08	0.39	-0.07
Gly	2.23	-5.36	0.30
His	2.41	1.74	1.11
Ile	-4.44	-1.68	-1.03
Leu	-4.19	-1.03	-0.98
Lys	2.84	1.41	-3.14
Met	-2.49	-0.27	-0.41
Phe	-4.92	1.30	0.45
Pro	-1.22	0.88	2.23
Ser	1.96	-1.63	0.57
Thr	0.92	-2.09	-1.40
Trp	-4.75	3.65	0.85
Tyr	-1.39	2.32	0.01
Val	-2.69	-2.53	-1.29

3.2 Iterative Self-consistent Algorithm Based on Partial Least Squares (ISC-PLS)

1. The initial X matrix (Set 0) is processed by multiple linear regression based on partial least squares (PLS) [15] and the initial model M0 is derived (Fig. 5). The M0 calculates pIC_{50} value (pIC_{50calc}) for each nonamer. The nonamer whose pIC_{50calc} is closest to pIC_{50exp} of the parent 15-mer is selected as the most probable binding core. Thus, 3094 most probable binding nonamers are collected from each 15-mer and enter a new set (Set 1).
2. From Set 1 is derived PLS model M1. M1 is used to predict the pIC_{50} s of Set 0 and again the most probable binding nonamer cores are selected and enter a new set (Set 2).
3. The Set 2 is used to derive model M2. M2 predicts the pIC_{50} s of Set 0 and the most probable binding nonamer cores are selected and enter a new set (Set 3).

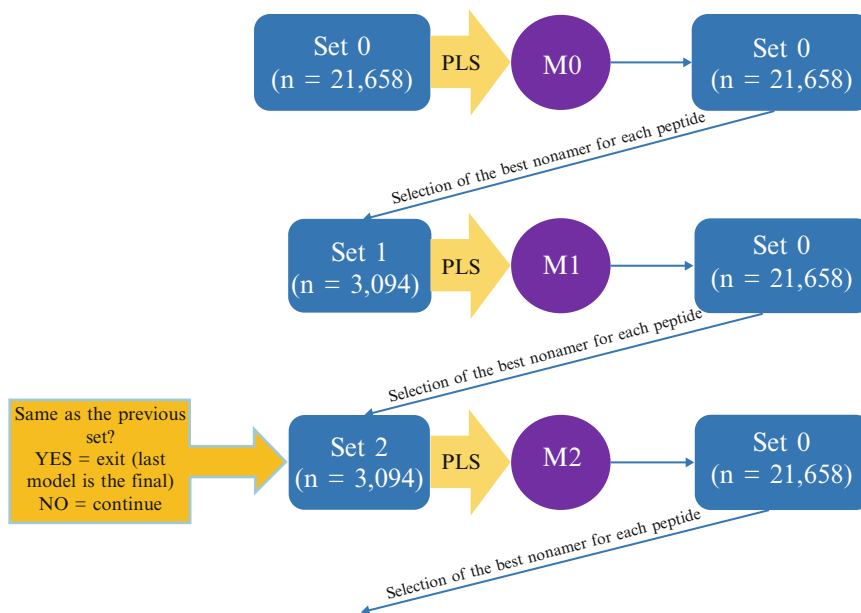


Fig. 5 Iterative self-consistent algorithm

4. After each iteration, the identity of two subsequent sets is registered.
5. Steps 2 and 3 are repeated until the identity of two subsequent set reaches 100 % or close to it.
6. The final set is used to derive the PCM model.

The calculations are performed by the software SIMCA 13.0.3 (Umetrics, Sartorius Stedim Biotech).

3.3 Model Validation

1. The test set of 72 parent proteins is presented as a set of 4239 overlapping peptides.
2. The derived PCM model is used to predict the pIC_{50} values of the test peptides for each of the seven HLA-DP proteins.
3. The set of peptides binding to a specific HLA-DP is divided into binders (predicted $pIC_{50} > 5.3$, $IC_{50} < 5000$ nM) and non-binders (predicted $pIC_{50} < 5.3$, $IC_{50} > 5000$ nM). Common nonamers between the subsets are classified as non-binders.
4. The predicted binders for each HLA-DP protein are ranked by descending pIC_{50} .
5. The top 5 % of the predicted binders are compared to the known binders from the training set. The common nonamers are considered as true binders and expressed as % of the total known binders originating from one parent protein. The PCM model recognizes 75 % of known binders.
6. Step 5 is repeated for the top 10 % of the predicted binders. The PCM model recognizes 95 % of the known binders.

3.4 EpiTOP

The model derived here is implemented by the EpiTOP server, which is freely accessible at <http://www.ddg-pharmfac.net/EpiTOP3/>.

EpiTOP is a proteochemometric tool for HLA class II binding prediction [16–19]. It predicts peptide binding to 24 most frequent HLA class II proteins: DRB1*01:01; DRB1*03:01; DRB1*04:01; DRB1*04:04; DRB1*04:05; DRB1*07:01; DRB1*08:02; DRB1*09:01; DRB1*11:01; DRB1*12:01; DRB1*13:01; DRB1*15:01; DQA1*01:01/DQB1*0501; DQA1*01:02/DQB1*0601; DQA1*03:01/DQB1*0302; DQA1*04:01/DQB1*0402; DQA1*05:01/DQB1*0301; DPA*02:01/DPB1*01:01; DPA*01:03/DPB1*02:01; DPA*01:03/DPB1*03:01; DPA*01:03/DPB1*04:01; DPA*01:03/DPB1*04:02; DPA*03:01/DPB1*04:02; and DPA*02:01/DPB1*05:01.

The EpiTOP home page is shown in Fig. 6. The amino acid sequence of the tested protein is given using single letter format. The HLA class II locus (HLA DRB1, HLA DQ, or HLA DP) of interest is selected and the binding threshold (pIC_{50}) is set. Two thresholds are optional: 5.3 (up to $IC_{50} = 5000$ nM, includes weak, median, and strong binders) and 6.3 (up to $IC_{50} = 500$ nM, includes median and strong binders). On the results page (Fig. 7), the protein is presented as a set of overlapping nonamers with predicted pIC_{50} values. The first column indicates the starting position of the nonamer in the protein. In the second column, the nonamer

Fig. 6 Home page of EpiTOP

Position	Peptide	pIC50 values for HLA-DP							Hit counts
		DPA1*0103/DPB1*0201	DPA1*0103/DPB1*0301	DPA1*0103/DPB1*0401	DPA1*0103/DPB1*0402	DPA1*0201/DPB1*0101	DPA1*0201/DPB1*0501	DPA1*0301/DPB1*0402	
1	IQGIPWEFL	9.072	9.155	8.277	9.042	7.942	8.625	8.754	7
2	QGGIPWEFL	5.484	6.904	4.967	5.434	5.518	5.488	5.39	1
3	SPWEFLPV	7.573	6.925	7.556	7.515	7.04	7.627	7.186	7
4	IPWEFLPV	6.449	6.556	5.816	6.392	5.273	5.819	6.355	4
5	IPWEFLPVV	6.516	7.153	6.623	6.421	6.764	6.199	5.983	5
6	WEFLPVV	6.738	7.127	6.592	6.685	6.269	6.44	6.707	6
7	EFLPVVFFQ	7.686	8.199	7.164	7.595	7.809	7.905	7.339	7
8	FLPVVFFQR	7.108	7.165	7.209	7.096	6.695	6.865	6.775	7
9	LPVFFQRL	8.023	8.268	8.324	7.979	8.201	7.664	7.69	7
10	IPVFFQRLA	6.751	7.363	5.945	6.795	5.339	6.241	6.735	4
11	VVFFQRLAI	6.252	7.314	5.86	6.191	5.215	5.89	6.336	2
12	FFVFFQRLAI	7.976	7.903	7.018	7.903	6.75	7.489	7.753	7
13	FFVFFQRLAI	6.383	7.138	6.149	6.372	6.015	6.004	6.163	3
14	FFVFFQRLAI	6.268	6.888	5.639	6.199	5.82	5.5	5.996	1
15	FFVFFQRLAI	4.335	5.596	4.83	4.238	5.411	4.675	4.154	0
16	RLAIIVD	8.637	9.364	7.457	8.585	8.306	9.04	8.631	7
17	LAIVDIN	6.382	7.866	5.668	6.324	5.798	5.795	5.985	3
18	LAIVDINR	6.661	6.852	6.998	6.597	7.107	6.746	6.397	7

Fig. 7 Results page of EpiTOP

sequence is shown and in the following columns the affinity for each DP protein, expressed as pIC_{50} , is given. The last column shows the number of DP proteins to which the peptide binds with pIC_{50} above the selected threshold.

4 Notes

The search in IEDB is case-sensitive. For full search, one has to enter all possible combinations for specific terms, e.g., the molecular weight is given in kDa, KDA, or kda units.

References

1. Janeway CA (2001) Immunobiology: the immune system in health and disease. Churchill Livingstone, New York, NY
2. Robinson J, Halliwell JA, Hayhurst JH, Flicek P, Parham P, Marsh SGE (2015) The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res* 43:D423–D431
3. Lapins M, Wikberg J (2010) Kinome-wide interaction modelling using alignment-based and alignment-independent approaches for kinase description and linear and non-linear data analysis techniques. *BMC Bioinformatics* 11:339
4. van Westen GJP, Wegner JK, IJzerman AP, van Vlijmen HWT, Bender A (2011) Proteochemometric modeling as a tool to design selective compounds and for extrapolating to novel targets. *Med Chem Commun* 2:16–30
5. Cortés-Ciriano I, Ain QU, Subramanian V, Lenselink EB, Méndez-Lucio O, IJzerman AP, Wohlfahrt G, Prusis P, Malliavin TE, van Westen GJP, Bender A (2015) Polypharmacology modelling using proteochemometrics (PCM): recent methodological developments, applications to target families, and future prospects. *Med Chem Commun* 6:24–50
6. Lapinsh M, Prusis P, Gutcaits A, Lundstedt T, Wikberg JE (2001) Development of proteochemometrics: a novel technology for the analysis of drug-receptor interactions. *Biochim Biophys Acta* 1525:180–190
7. Prusis P, Uhlén S, Petrovska R, Lapinsh M, Wikberg JE (2006) Prediction of indirect interactions in proteins. *BMC Bioinformatics* 7:167
8. Lapinsh M, Prusis P, Lundstedt T, Wikberg JE (2002) Proteochemometrics modeling of the interaction of amine G-protein coupled receptors with a diverse set of ligands. *Mol Pharmacol* 61:1465–1475
9. Sidney J, Steen A, Moore C, Ngo S, Chung J, Peters B, Sette A (2011) Five HLA-DP molecules frequently expressed in the worldwide human population share a common HLA supertypic binding specificity. *J Immunol* 184:2492–2503

10. Dai S, Murphy GA, Crawford F, Mack DG, Falta MT, Marrack P, Kappler JW, Fontenot AP (2010) Crystal structure of HLA-DP2 and implications for chronic beryllium disease. *Proc Natl Acad Sci U S A* 107:7425–7430
11. Patronov A, Dimitrov I, Flower DR, Doytchinova I (2011) Peptide binding prediction for the human class II MHC allele HLADP2: a molecular docking approach. *BMC Struct Biol* 11:32
12. Patronov A, Dimitrov I, Flower DR, Doytchinova I (2012) Peptide binding to HLA-DP proteins at pH 5.0 and pH 7.0: a quantitative molecular docking study. *BMC Struct Biol* 12:20
13. Hellberg S, Sjostrom M, Skagerberg B, Wold S (1987) Peptide quantitative structure-activity relationships, a multivariate approach. *J Med Chem* 30:1126–1135
14. Eriksson L, Jonsson J, Sjostrom M, Wold S (1989) Multivariate parametrization of coded and non-coded amino acids by thin layer chromatography. *Prog Clin Biol Res* 291:131–134
15. Wold S, Ruhe A, Wold H, Dunn WJ (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J Sci Stat Comput* 5:735–743
16. Dimitrov I, Garnev P, Flower DR, Doytchinova I (2010) Peptide binding to the HLA-DRB1 supertype: a proteochemometric analysis. *Eur J Med Chem* 45:236–243
17. Dimitrov I, Garnev P, Flower DR, Doytchinova I (2010) Epi TOP – a proteochemometric tool for MHC class II binding prediction. *Bioinformatics* 26:2066–2068
18. Dimitrov I, Doytchinova I (2015) Peptide binding prediction to five most frequent HLA-DQ proteins – a proteochemometric approach. *Mol Inf* 34:467–476
19. Yordanov V, Dimitrov I, Doytchinova I (2018) Proteochemometrics-based prediction of peptide binding to HLA-DP proteins. *J Chem Inf Model* 58:297–304. <https://doi.org/10.1021/acs.jcim.7b00026>



Linked Open Data: Ligand-Transporter Interaction Profiling and Beyond

Stefanie Kickinger, Eva Hellsberg, Sankalp Jain, and Gerhard F. Ecker

Abstract

Multi-target drug design is an innovative new paradigm in the drug development process. With the help of growing open data sources, *in silico* modeling approaches have become successful tools to discover and investigate multi-target drugs. In this chapter, we describe a workflow for retrieving and curating information for multiple drug targets from the open domain, provide insights into how the retrieved data can be employed in ligand and structure-based approaches, and discuss the hurdles to consider with respect to data analysis.

Keywords KNIME workflow, Ligand-based design, Molecular docking, Molecular dynamics simulation, Multi-target drug design, Open data, Protein homology modeling, Structure-based design

1 Introduction

Multi-target drug design is an emerging new paradigm to treat complex diseases by regulating multiple targets at the same time to achieve the desired physiological responses [1–4]. Traditionally, drugs have been designed to selectively modulate a so-called on-target in order to avoid side effects by modulating “off-targets.” However, several approved drugs retrospectively have been shown to hit more than one target, which turned out to contribute to the therapeutic efficacy [5, 6]. Furthermore, in recent years many drugs failed in phase II clinical trials because of a lack of therapeutic efficacy [7]. Therefore, multi-target drug design represents an innovative principle to overcome lack of efficacy. Different approaches to discover and investigate multi-target drugs have been reviewed by Zhang et al. [8] addressing data-driven, ligand-based, or structure-based methods [4, 9–14]. Most of these methods focus on drug repurposing (i.e., to find new targets for known drugs) such as the ligand-based methods SPiDER [15] and SEA [16], which are based on 2D fingerprint or 3D shape similarity. Furthermore, structure-based methods such as TarFisDOCK [17], INVDOCK [14, 18], or

VinaMPI [18] could be used to dock potential ligands into many target structures at the same time [19, 20]. With the help of growing open data sources such as Open PHACTS [21], ChEMBL [22], and freely available medicinal chemistry literature, data-driven *in silico* modeling approaches have also proven to be capable of effectively identifying protein-ligand interactions at an early stage in the drug discovery pipeline [23]. However, increase in complexity and size and diversity of public data sources necessitate judicious curation of the data. With the availability of workflow tools like KNIME [24] or pipeline pilot [25], complex querying for multiple drug targets became a feasible task without the need of comprehensive programming skills [26]. In this chapter, we present a protocol which starts with mining the Open PHACTS Discovery Platform to collect a data-set of suitable size and quality for subsequent structure-based selectivity profiling studies. As concrete case study, we chose the human serotonin (hSERT) and dopamine transporter (hDAT). Both proteins belong to the neurotransmitter sodium symporter family which represents the largest group of transporters in the human genome. hSERT and hDAT are responsible for the reuptake of serotonin and dopamine, respectively, from the presynaptic cleft after signaling [27, 28]. Numerous drugs have been developed which interact with these transporters and are used as therapeutic agents to treat neurological disorders such as depression. In addition, there is a wealth of compounds which are abused as illicit drugs [28–30]. Even though hSERT and hDAT share high sequence and structural similarity, they fulfill different physiological roles. Substances increasing dopamine levels in the mesolimbic pathway of the brain can influence the reward system, whereas increased levels of serotonin are involved in several other neurotransmitter systems, most importantly influencing mood [31]. A profound understanding of the structural basis for hSERT and hDAT ligand selectivity is therefore of major interest for designing ligands that either hit one of these transporters or both. This chapter will tackle this research question by reviewing the data mining and curating process for hSERT and hDAT bioactivities present in the linked open data domain. This is followed by a comprehensive scaffold analysis in order to analyze the chemical space, which allowed to identify a congeneric series of compounds suitable for structure-activity relationship studies and experimental data guided ligand docking. The power of this protocol is based on the combination of mining the available knowledge in the open data domain and its breakdown to concrete molecular interactions. This chapter thus gives an overview of the overall workflow, points out the potential of retrieving data for multiple drug targets from the open domain, provides insights into structure-based approaches, and discusses the hurdles to be considered in data analysis.

2 Materials

Data retrieval and scaffold analysis

- Knime [24]: Knime is an open-source platform that provides an integrated solution for the data mining process across the drug discovery pipeline. It can be downloaded from <https://www.knime.com/software>. It also provides a visual assembly of data workflows drawn from an extensive repository of tools. Additionally, it also offers nodes for machine learning (classification and regression analysis).

Homology modeling

- MODELLER [32]: Modeller is a widely used open-source software for comparative modeling of protein three-dimensional structures. The program also incorporates limited functions for ab initio structure prediction of loop regions of proteins, which are often highly variable even among homologous proteins and thus difficult to predict by homology modeling. It can be downloaded from <https://salilab.org/modeller>.

Molecular docking and visualization

- Schrödinger [33]: Schrödinger is one of the leading commercial software packages in the field of drug design. It includes small molecule modeling and simulations, macromolecular modeling and simulations, lead discovery, and lead optimization, visualization, and automation (<https://www.schrodinger.com/maestro>). Glide [34] is the molecular docking module in Schrödinger that places the ligand in the protein binding pocket and ranks the generated poses with an empirical scoring function.
- Molecular Operating Environment (MOE) [35]: MOE is a commercial drug discovery software platform that integrates visualization, modeling, and simulations, as well as methodology development, in one package (<http://www.chemcomp.com/>).

Molecular dynamics

- Desmond [36]: Desmond is a freely available software package developed at D. E. Shaw Research to perform high-speed molecular dynamics simulations of biological systems (<http://www.deshawresearch.com/index.html>). Schrödinger provides an easy-to-use graphical user interface for performing molecular dynamics simulations with Desmond [37].

3 Methods

Sophisticated approaches are necessary to tackle multi-target drug design. The great variety of methodological possibilities demands well-informed decisions on which individual path to embark. In this section, we describe the methods in detail which we used to retrieve and curate information on two drug targets. Note that this example was driven by the solid basis of available experimental data and previous findings on these drug targets. All technical parameters described in the methods section are either the default options recommended by the software developers or adapted due to specific biological evidence relevant for the focus of the study.

3.1 Data Collection and Data Mining

Open data sources such as ChEMBL [22], DrugBank [38], KEGG [39], or Open PHACTS [21] provide a large collection of linked information on compounds including their structures, biological targets, pathways, bioactivities, and experimental details on biological assays. ChEMBL and other resources extract their information from the literature in an automated or semiautomated fashion. The collected data therefore originate from a variety of different resources resulting in a collection of bioactivity data of different activity endpoints (K_i , IC_{50} , % inhibition, etc.) that was measured in different assay types and under varying assay conditions (*see Note 1*). However, using such diverse data for modeling or virtual screening was reported to show inconsistent performance, and hence recommendations were proposed to deal with the experimental uncertainty associated with such data [40, 41]. For our case study, bioactivity data for hSERT and hDAT were extracted from the Open PHACTS Discovery Platform via a KNIME workflow. The application programming interface (API) call was used to retrieve pharmacology data from ChEMBL20 for both proteins. In the present case study, we decided to include the bioactivity endpoints IC_{50} and K_i , because these bioactivities have been demonstrated to be most reliably in large data analysis [41, 42] and because they can be correlated with each other. In order to investigate the uncertainty of the data that was introduced by combining these different activity endpoints from different assays, the correlation between pIC_{50} and pK_i ($p = \text{negative log}$) values from duplicate measurements for hSERT and hDAT was calculated. This showed that the observed correlations are within the same range as the calculated correlations for duplicate measurements within only one of the activity endpoints [43]. As a next step, classification of the data into active and inactive compounds has to be performed in order to extract the actives. Setting reasonable activity thresholds is a challenging task, and it requires considering the focus of the study. In the present case, the thresholds were tailored according to the lowest known activity endpoints (IC_{50} and K_i) that still showed pharmacological activity.

If a dataset is used for calculating structure-activity relationships (SAR), the compounds must be measured for the same activity endpoint (i.e., either IC_{50} alone or K_i alone). However, if a dataset is designed for the construction of machine learning models, also the use of activity annotations is possible (i.e., active, 1, inactive, 0). In this scenario, the data from different endpoints can be merged (as described above). To increase the accuracy of the classification of the dataset, data points close to the activity thresholds might be omitted. Inconsistent data points with conflicting activity data should in general be omitted from the dataset. In order to visualize the diversity of the dataset and to see if there are scaffolds showing pronounced selectivity for one or both targets, Bemis-Murcko scaffold analysis [44] was performed. Out of the 53 most populated scaffolds, four scaffolds were identified as hDAT selective, 10 as hSERT selective, and 24 as promiscuous. In order to perform quantitative structure-activity relationship (SAR) calculations, scaffolds that contained congeneric series of compounds, which showed selectivity for one of the targets and were measured in the same assay, were prioritized. A congeneric series of 56 compounds sharing a cathinone substructure was identified that showed pronounced selectivity for hDAT over hSERT. A detailed description of the KNIME workflow for data retrieval, filtering, preprocessing, and analyses can be found in [43]. The whole workflow can be downloaded from myExperiment [45]. Out of the whole set of derivatives, six compounds were further selected for subsequent structure-based studies in order to link the observed selectivity profile to specific molecular interactions.

3.2 Ligand-Based Methods

In general, ligand-based methods can be used to find trends in the data (as discussed above) or to classify compounds with machine learning methods. However, their application depends strongly on the data quality. In our case study, we analyzed the SAR of the 56 cathinones to get first insights which molecular features trigger their selectivity profiles. Since the compounds show selectivity for hDAT (over hSERT), we performed multiple linear regression (classical Hansch analysis) with hDAT pIC_{50} values and selectivity ($= \log(hSERT IC_{50}/hDAT IC_{50})$) as dependent variables using a limited set of descriptors characterizing the molecules (Van der Waals volume (overall, $C\alpha$ - and N-substituents), partition coefficient ($\log P(o/w)$), molar refractivity, constants for the substituents on the aromatic ring, and indicator variables for meta- and para-substitutions). Briefly, both calculated equations showed a first trend that the substituent on the $C\alpha$ -atom to the carbonyl group of the compounds influences hDAT activity and selectivity. Details on the approach can be found in [43]. This information is subsequently used to guide the prioritization of docking poses.

3.3 Structure-Based Methods

Structure-based methods require 3D coordinates from available high-resolution crystal structures, NMR experiments, or homologous template structures. A plethora of crystal structures is deposited in the Protein Data Bank [46] (PDB, www.rcsb.org) and can be downloaded free of charge. All selected crystal structures should be checked thoroughly whether the resolution and B-factors are appropriate, if certain amino acids are annotated with multiple possible rotamers, and if there are relevant amino acids missing (*see Note 2*). This procedure can be performed with commercial protein visualization software (MOE [35] or Schrödinger Suite [33]) or free software (VMD [47] or pymol [48]). A lot of information can be already taken from the downloaded pdb files themselves, as they are written in text format and include the experimental data and setup. A visual inspection of PDB structures is also possible in a web browser using the LiteMol viewer [49, 50] in PDBe (<https://www.ebi.ac.uk/pdbe/>) [51]. Since many crystal structures are models retrieved by X-ray crystallography based on experimentally measured diffraction patterns, it is furthermore advisable to check the placement of the protein and its ligands in the experimentally measured electron density map [31, 52]. Electron density maps can be visualized with commercial software (Schrödinger [33], MOE [35]) and free software (Coot [53]) or in the web browser (LiteMol [49, 50], PDBe [51]). By considering the abovementioned procedures, one can identify the areas of the crystal structure where the structure can be trusted or should be taken with caution. In the case of our study, no crystal structures of hSERT and hDAT were resolved back then. Consequently, homology modeling needs to be performed to obtain decent models based on suitable template crystal structures.

3.3.1 Homology Modeling

Homology modeling or comparative modeling refers to the technique of using a resolved crystal structure to model an unknown homologous protein structure. It is believed that overall fold is far more conserved among different proteins than sequence identity [54]. There are four crucial steps in homology modeling. First, a suitable crystal structure is chosen as a template. At the time this analysis was performed, the PDB provided two different types of homologous template structures for modeling hSERT and hDAT: crystal structures of the bacterial leucine transporter (LeuT, sequence 20%) [55] and the drosophila dopamine transporter (dDAT, sequence identity 70%) [56]. In the present case study, the dDAT PDB structure 4M48 was chosen as the most suitable template due to higher sequence identity and the fact that it shows the desired outward-open conformation (*see Note 3*). Second, the desired protein sequence needs to be aligned with the template structure. This task was performed with the tool ClustalX [57]. All 12 transmembrane helices (TMs) of hSERT and hDAT are highly

conserved and can be easily aligned with the template structure. Third, models are generated and refined, e.g., with the program Modeller, which was also the program of choice in this study [32]. Within Modeller it is possible to also implement experimental data in the model generation process by setting restraints for secondary structure elements, disulfides or salt bridges. Fourth, the models' quality needs to be assessed with help of, e.g., the DOPE score (*see Note 4*) [58]. Additional quality assessment can be performed with ProCHECK (<https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/index.html>) [58, 59] and ProQM (<http://bioinfo.ifm.liu.se/ProQM/index.php>) [60]. Procheck additionally provides Ramachandran plots and information on residue properties. ProQM was specifically optimized for membrane proteins. Nevertheless, the quality of the homology model depends highly on the quality of the available crystal structures and the amount of available structural information. A more detailed description of homology modeling was recently provided by Lushington [61]. The generated hSERT and hDAT homology models were then further used for molecular docking experiments.

3.3.2 Docking

Molecular Docking is a common method in structure-based drug design to calculate the possible positions of a ligand in the binding site of its target protein. A great variety of software packages is available that provide different algorithms and all kinds of settings [62]. In the present example, six selected compounds of a congeneric series sharing a cathinone scaffold were docked into the central binding site of both the homology models of hDAT and hSERT with Glide 6.8 [34] from the Schrödinger release 2015-2 [33]. In Glide, the protein is kept rigid during the docking process, and the ligands are placed into the space between defined binding site residues. This setting was sufficient for our task, as we were docking small compounds with respect to the outward-open binding site of the transporters and we wanted to keep the side chain rotamers of the homology models as close as possible to the dDAT template crystal structure 4M48 at this stage. Furthermore, we restrained the cationic amine function of the compounds to be placed within 2–4 Å to the carbonyl oxygen of F76 in hDAT and Y95 of hSERT, because several X-ray structures of related proteins with co-crystallized ligands are available in the PDB showing a similar distance (for further details, *see* [43]). The decision on how much flexibility should be allowed during the docking process is strongly depending on the availability of experimental data—which is very rich in this case. Consequently, the introduced bias caused by applying docking constraints was justified by the available experimental data. The models and ligands were prepared in the Schrödinger suite using default options (*see Note 5*). Once the docking output is generated, which usually results in about 100 poses per ligand, a reasonable pose

analysis and interpretation approach are needed. The poses are ranked by a specific docking score, which gives an orientation how well the program was able to place the ligand into the defined binding site. The docking score includes relevant energetic and steric terms to achieve a most accurate placement and ranking. The Glide-Score (used in this study) consists of such components (van der Waals energy, Coulomb energy, lipophilic term, hydrogen-bonding term, metal-binding term, as well as several rewards and penalties for relevant features) [33] to predict the binding mode of the ligand most accurately. However, these algorithms cannot include individual information such as the details known from biological experiments about proposed binding modes for a certain target. In this case a common scaffold clustering approach of all gained poses is recommended [63]. In this approach, the common scaffold shared by all docked ligands is extracted, and an RMSD matrix of all poses is generated from these atoms. Subsequently, the clusters are calculated at a defined similarity level which corresponds to the maximal distance within a cluster in Ångström. This helps to bundle the large amount of poses into assessable bins which can be analyzed for common characteristics and compared with the knowledge from biological experiments in a more quantitative way. The analysis of the docking study revealed certain trends explaining the observed ligand selectivity of hSERT over hDAT showing slightly more negative overall glide scores, less steric clashes, and hydrogen bonding exclusively in hDAT.

3.3.3 Molecular Dynamics Simulations

In general, molecular dynamics (MD) simulations are used to study the motions of molecules over time and are therefore the method of choice to characterize dynamic interactions within and between biomolecules. Using such methods requires a lot of considerations regarding the force field, ligand parameters, membrane and solvent type, ion concentration, system size, and many more. Experimental data about the respective systems and facts from profound literature ideally guide these decisions. The book *Molecular Modeling of Proteins* [64] provides an excellent review on various aspects of these issues. This case study focuses on the protein-ligand interactions between cathinone compounds and hSERT and hDAT. Investigating the structure-activity relationships of these compounds and a subsequent docking study showed trends in the ligand selectivity and provided possible binding modes. To further evaluate these hypotheses, MD simulations of one compound representing the previous findings (*see* [43]) were conducted. In this context, the primary aim is to verify the stability of the complexes gained from docking and to review the motions of the ligand inside the binding site over time. MD simulations are computationally expensive and need comprehensive analysis, so it is crucial to take the actual research question into consideration before choosing the simulation

settings. For example, the simulation time to check the ligand stability can be short (20 ns) if the binding mode is well defined, whereas free simulation of unbinding might take up to micro or even milliseconds [65–67]. For this study, a system instability or an unfavorable starting pose of the ligand would already be observed within the first nanoseconds of the simulation, because the biological data provide a solid basis for our current understanding. The major criteria to prove stability is a convergence of the root-mean-square deviation (RMSD) of the protein and the ligand in unrestrained simulations. For the protein, it is important to solely consider the RMSD of the backbone atoms as the higher side chain movement could hide major conformational changes in the backbone. The stability of the protein-ligand interactions can be observed by investigating all interactions of the ligand with the protein residues over the whole simulation time. This identifies the involved residues and shows the duration of each interaction. Key interactions should be present over the whole simulation time. The structure-based part of this work was all done in the Schrödinger software suite [33]. The MD simulations were prepared in Maestro 10.2 [68] and conducted in 20 ns simulations with Desmond 4.2 [69]. The MD studies showed that the selected poses were stable and could also confirm the observed trends in the ligand selectivity profiles for the two target proteins.

3.4 Summary

Designing ligands which target multiple targets with a defined affinity pattern represents a powerful approach to overcome lack of efficacy. With this case study, we present a holistic workflow starting from data mining across public data sources and ending with molecular dynamics simulations of a concrete ligand-transporter complex, which revealed the stability of the ligand-binding mode suggested by experimental data guided docking. As parts of the protocols described are implemented in KNIME workflows, they can be easily adapted to other targets of interest.

4 Notes

1. In ChEMBL, more than 5000 measurement types are considered including, e.g., “%max,” “Activity,” “Efficacy,” “EC50,” “Kd,” and “Residual Activity” [41]. Depending on the focus of the study, these filters can be modified.
2. If there are several rotamers possible fitting in the observed electron density, the “right” rotamer is not necessarily the one selected by the crystallographer! High *B*-values are also an indicator for high flexibility. Make sure to check which rotamer is relevant for the specific research question.

3. When dealing with flexible proteins such as transporters, choosing the right conformation of your template structure is essential. We believe that classical inhibitors most probably bind and stabilize the outward-open conformation of the transporter and therefore hinder the transporter from adopting other conformations in the transport cycle [55]. Substrates most likely bind to the occluded transporter state as the translocation process requires among others the adaptation of an outward-occluded transporter conformation [70].
4. The DOPE score is the most widely used quality assessment parameter even though it is only optimized for soluble proteins [58]. It has been successfully used for scoring homology models of different membrane proteins [71, 72], nevertheless, it is advisable to not only rely on this parameter when modeling membrane proteins. Scores specifically optimized for membrane proteins such as the ProQM score should be taken into consideration as well for selecting the best model.
5. The Schrödinger Suite [33] offers preparation modules for both proteins and ligands. It is strongly recommended to conduct both preparation and docking procedure in the same software package as the used algorithms are compatible.

Acknowledgment

We gratefully acknowledge financial support provided by the Austrian Science Fund, grants #F03502 (SFB35) and W1232 (MolTag). Stefanie Kickinger, Eva Hellsberg, and Sankalp Jain contributed equally to this chapter.

References

1. Morphy R, Rankovic Z (2005) Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem* 48:6523–6543
2. Hopkins AL, Mason JS, Overington JP (2006) Can we rationally design promiscuous drugs? *Curr Opin Struct Biol* 16:127–136
3. Peters J-U (2013) Polypharmacology – foe or friend? *J Med Chem* 56:8955–8971
4. Anighoro A, Bajorath J, Rastelli G (2014) Polypharmacology: challenges and opportunities in drug discovery. *J Med Chem* 57:7874–7887
5. Bolognesi ML, Cavalli A (2016) Multitarget drug discovery and polypharmacology. *ChemMedChem* 11:1190–1192
6. Lu J-J, Pan W, Hu Y-J, Wang Y-T (2012) Multi-target drugs: the trend of drug research and development. *PLoS One* 7:e40262
7. Harrison RK (2016) Phase II and phase III failures: 2013–2015. *Nat Rev Drug Discov* 15:817–818
8. Zhang W, Pei J, Lai L (2017) Computational multitarget drug design. *J Chem Inf Model* 57:403–412
9. Ma XH, Shi Z, Tan C, Jiang Y, Go ML, Low BC, Chen YZ (2010) In-silico approaches to multi-target drug discovery. *Pharm Res* 27:739–749
10. Koutsoukas A, Simms B, Kirchmair J et al (2011) From in silico target prediction to multi-target drug design: current databases, methods and applications. *J Proteome* 74:2554–2574
11. Lavecchia A, Cerchia C (2016) In silico methods to address polypharmacology: current

- status, applications and future perspectives. *Drug Discov Today* 21:288–298
12. Taboureau O, Baell JB, Fernández-Recio J, Villoutreix BO (2012) Established and emerging trends in computational drug discovery in the structural genomics era. *Chem Biol* 19:29–41
 13. Kuyoc-Carrillo VF, Medina-Franco JL (2014) Progress in the analysis of multiple activity profile of screening data using computational approaches. *Drug Dev Res* 75:313–323
 14. Ellingson SR, Smith JC, Baudry J (2014) Polypharmacology and supercomputer-based docking: opportunities and challenges. *Mol Simul* 40:848–854
 15. Reker D, Rodrigues T, Schneider P, Schneider G (2014) Identifying the macromolecular targets of de novo-designed chemical entities through self-organizing map consensus. *Proc Natl Acad Sci U S A* 111:4067–4072
 16. Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK (2007) Relating protein pharmacology by ligand chemistry. *Nat Biotechnol* 25:197–206
 17. Li H, Gao Z, Kang L et al (2006) TarFisDock: a web server for identifying drug targets with docking approach. *Nucleic Acids Res* 34:W219–W224
 18. Chen YZ, Zhi DG (2001) Ligand-protein inverse docking and its potential use in the computer search of protein targets of a small molecule. *Proteins* 43:217–226
 19. Meslamani J, Rognan D, Kellenberger E (2011) sc-PDB: a database for identifying variations and multiplicity of “druggable” binding sites in proteins. *Bioinformatics* 27:1324–1326
 20. Meslamani J, Li J, Sutter J, Stevens A, Bertrand H-O, Rognan D (2012) Protein-ligand-based pharmacophores: generation and utility assessment in computational ligand profiling. *J Chem Inf Model* 52:943–955
 21. Williams AJ, Harland L, Groth P et al (2012) Open PHACTS: semantic interoperability for drug discovery. *Drug Discov Today* 17:1188–1198
 22. Bento AP, Gaulton A, Hersey A et al (2014) The ChEMBL bioactivity database: an update. *Nucleic Acids Res* 42:D1083–D1090
 23. Montanari F, Zdrazil B, Digles D, Ecker GF (2016) Selectivity profiling of BCRP versus P-gp inhibition: from automated collection of polypharmacology data to multi-label learning. *J Cheminform.* <https://doi.org/10.1186/s13321-016-0121-y>
 24. Berthold MR et al (2008) KNIME: the Konstanz information miner. In: Preisach C, Burkhardt H, Schmidt-Thieme L, Decker R (eds) Data analysis, machine learning and applications, Studies in classification, data analysis, and knowledge organization. Springer, Berlin, Heidelberg
 25. Pipeline pilot. <http://accelrys.com/products/collaborative-science/biovia-pipeline-pilot/>
 26. Montanari F, Zdrazil B (2017) How open data shapes in silico transporter modeling. *Molecules.* <https://doi.org/10.3390/molecules22030422>
 27. César-Razquin A, Snijder B, Frappier-Brinton T et al (2015) A call for systematic research on solute carriers. *Cell* 162:478–487
 28. Kristensen AS, Andersen J, Jørgensen TN, Sørensen L, Eriksen J, Loland CJ, Strømgaard K, Gether U (2011) SLC6 neurotransmitter transporters: structure, function, and regulation. *Pharmacol Rev* 63:585–640
 29. Koldsø H, Christiansen AB, Sinning S, Schiøtt B (2013) Comparative modeling of the human monoamine transporters: similarities in substrate binding. *ACS Chem Neurosci* 4:295–309
 30. Sitte HH, Freissmuth M (2015) Amphetamines, new psychoactive drugs and the monoamine transporter cycle. *Trends Pharmacol Sci* 36:41–50
 31. Schultz W (2010) Dopamine signals for reward value and risk: basic and recent data. *Behav Brain Funct* 6:24
 32. Webb B, Sali A (2016) Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci* 86:2.9.1–2.9.37
 33. (2015) Schrödinger Release 2015-2. Schrödinger, LLC, New York, NY
 34. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* 47:1750–1759
 35. Molecular Operating Environment (MOE), 2013.08. Chemical Computing Group Inc., Montreal, Canada
 36. Bowers K, Chow E, Xu H, et al (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters. In: ACM/IEEE SC 2006 conference (SC’06). <https://doi.org/10.1109/sc.2006.54>
 37. Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2017. Maestro-Desmond interoperability tools, Schrödinger, New York, NY, 2017
 38. Wishart DS, Feunang YD, Guo AC et al (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 46: D1074–D1082

39. Aoki-Kinoshita KF, Kanehisa M (2007) KEGG primer: an introduction to pathway analysis using KEGG. NCI Nature Pathway Interaction Database. <https://doi.org/10.1038/pid.2007.2>
40. Kramer C, Fuchs JE, Whitebread S, Geddeck P, Liedl KR (2014) Matched molecular pair analysis: significance and the impact of experimental uncertainty. *J Med Chem* 57:3786–3802
41. Hu Y, Bajorath J (2014) Influence of search parameters and criteria on compound selection, promiscuity, and pan assay interference characteristics. *J Chem Inf Model* 54:3056–3066
42. Hu Y, Bajorath J (2015) Structural and activity profile relationships between drug scaffolds. *AAPS J* 17:609–619
43. Zdrzil B, Hellsberg E, Viereck M, Ecker GF (2016) From linked open data to molecular interaction: studying selectivity trends for ligands of the human serotonin and dopamine transporter. *Medchemcomm* 7:1819–1831
44. Bemis GW, Murcko MA (1996) The properties of known drugs. 1. Molecular frameworks. *J Med Chem* 39:2887–2893
45. myExperiment – Workflows – KNIME workflow without hERG labels included from Zdrzil et al., *MedChemComm*, 2016: “From linked open data to molecular interaction: studying selectivity trends for ligands of the human serotonin and dopamine transporter” (Barbara Zdrzil) [KNIME Workflow]. <https://www.myexperiment.org/workflows/4911.html>. Accessed 5 Feb 2018
46. Berman HM (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242
47. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:33–38
48. PyMOL, The PyMOL Molecular Graphics System, version 2.0. Schrödinger, LLC
49. Conroy MJ, Sehnal D, Deshpande M, Svobodova R, Mir S, Berka K, Midlik A, Velankar S, Koca J (2017) LiteMol: web-based three-dimensional visualization of macromolecular structure data. *Acta Crystallogr A* 73:C669
50. Sehnal D, Deshpande M, Vařeková RS, Mir S, Berka K, Midlik A, Pravda L, Velankar S, Koča J (2017) LiteMol suite: interactive web-based visualization of large-scale macromolecular structure data. *Nat Methods* 14:1121–1122
51. Mir S, Alhroub Y, Anyango S et al (2018) PDBe: towards reusable data delivery infrastructure at protein data bank in Europe. *Nucleic Acids Res* 46:D486–D492
52. Pozharski E, Weichenberger CX, Rupp B (2013) Techniques, tools and best practices for ligand electron-density analysis and results from their application to deposited crystal structures. *Acta Crystallogr D Biol Crystallogr* 69:150–167
53. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501
54. Qu X, Swanson R, Day R, Tsai J (2009) A guide to template based structure prediction. *Curr Protein Pept Sci* 10:270–285
55. Singh SK, Piscitelli CL, Yamashita A, Gouaux E (2008) A competitive inhibitor traps LeuT in an open-to-out conformation. *Science* 322:1655–1661
56. Penmatsa A, Wang KH, Gouaux E (2013) X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature* 503:85–90
57. Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics*. Chapter 2:Unit 2.3
58. Shen M-Y, Sali A (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci* 15:2507–2524
59. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291
60. Ray A, Lindahl E, Wallner B (2010) Model quality assessment for membrane proteins. *Bioinformatics* 26:3067–3074
61. Lushington GH (2015) Comparative modeling of proteins. *Methods Mol Biol* 1215:309–330
62. Chen Y-C (2015) Beware of docking! *Trends Pharmacol Sci* 36:78–95
63. Richter L, de Graaf C, Sieghart W, Varagic Z, Mörzinger M, de Esch IJP, Ecker GF, Ernst M (2012) Diazepam-bound GABAA receptor models identify new benzodiazepine binding-site ligands. *Nat Chem Biol* 8:455–464
64. Kukol A (2017) Molecular modeling of proteins. Humana Press, New York
65. Schuetz DA, de Witte WEA, Wong YC et al (2017) Kinetics for drug discovery: an industry-driven effort to target drug residence time. *Drug Discov Today* 22:896–911
66. De Vivo M, Masetti M, Bottegoni G, Cavalli A (2016) Role of molecular dynamics and related methods in drug discovery. *J Med Chem* 59:4035–4061

67. Huang D, Caffisch A (2011) The free energy landscape of small molecule unbinding. *PLoS Comput Biol* 7:e1002002
68. Schrödinger Release 2015-2: Maestro, version 10.2. Schrödinger, LLC, New York, NY
69. Schrödinger Release 2015-2: Desmond Molecular Dynamics System, version 4.2. Schrödinger, LLC, New York, NY
70. Wang H, Gouaux E (2012) Substrate binds in the S1 site of the F253A mutant of LeuT, a neurotransmitter sodium symporter homologue. *EMBO Rep* 13:861–866
71. Jurik A, Zdrazil B, Holy M, Stockner T, Sitte HH, Ecker GF (2015) A binding mode hypothesis of tiagabine confirms liothyronine effect on γ -aminobutyric acid transporter 1 (GAT1). *J Med Chem* 58:2149–2158
72. Saha K (2015) “Second generation” mephedrone analogs, 4-MEC and 4-MePPP, differentially affect monoamine transporter function. *Intrinsic Activity* 3:A2.18



Design of Novel Dual-Target Hits Against Malaria and Tuberculosis Using Computational Docking

Manoj Kumar and Anuj Sharma

Abstract

Drugs which are purposefully designed to hit more than one target (multi-target drugs) promise a better safety profile and low resistance probability. Multi-target therapy also offers a cost-effective model for pharmaceutical R&D, making it quite an appealing strategy in the domain of neglected tropical diseases (NTDs) and other infections/coinfections of the global impact such as malaria, tuberculosis, and AIDS. We reviewed herein different approaches (knowledge base and screening base) for designing multi-target inhibitors with the special emphasis on the research work of the authors. Additionally, a step-by-step guide (protocol) and different computational resources are also discussed in detail to design multi-target hits for malaria and tuberculosis.

Keywords AIDS, AutoDock, Computational docking, Druglikeness, Infectious diseases, Ligand efficiency, Malaria, Multi-target drugs, Multi-target screening, Neglected tropical diseases, Tuberculosis

Abbreviations

2D	Two dimensional
3D	Three dimensional
ACD	Advanced Chemistry Development, Inc.
ADMET	Absorption, distribution, metabolism, excretion, and toxicity
ADT	AutoDock Tool
AIDS	Acquired immune deficiency syndrome
AMI	Austin model 1
AMBER	Assisted model building with energy refinement
AMR	Antimicrobial resistance
BE	Binding energy
BMRB	Biological magnetic resonance data bank
CCDC	The Cambridge Crystallographic Data Centre
CHARMM	Chemistry at Harvard Macromolecular Mechanics
ChEMBL or ChEMBLdb	Chemical database of bioactive molecules with drug-like properties
COX	Cyclooxygenase
DHFR	Dihydrofolate reductase
DMEs	Disease-modifying agents
DMLs	Designed multiple ligands
DSV	Discovery Studio Visualizer

DUD	Directory of useful decoys
EGFR	Epidermal growth factor receptor
EMBL-EBI	European Bioinformatics Institute
FDA	Food and Drug Administration, USA
GA	Genetic algorithm
HIV	Human immunodeficiency virus
In silico (syn in computo)	Performed on computer
LE	Ligand efficiency
LGA	Lamarckian genetic algorithm
LS	Local search
MAPK	Mitogen-activated protein kinase
MD	Molecular dynamics
MDR	Multidrug resistance
MGLTools	Molecular Graphics Laboratory tools
MM2	Molecular mechanics 2
MMV	Molegro Molecular Viewer
MTDs	Multi-target drugs
MW	Molecular weight
NSAIDs	Nonsteroidal anti-inflammatory drugs
NTDs	Neglected tropical diseases
OMICS	Genomics, proteomics, or metabolomics
PAINS	Pan-assay interference compounds
PDB	Protein Data Bank
PM3	Parameterized model number 3
QSAR	Quantitative structure activity relationship
R&D	Research and development
RCSB	Research Collaboratory for Structural Bioinformatics
RMSD/rmsd	Root-mean-square deviation
SA	Simulated annealing
SHAFTS	Shape-Feature Similarity
SITITCH	Search tool for interacting chemicals
STRING	Search tool for the retrieval of interacting genes/proteins
TB	Tuberculosis
TCM	Traditional Chinese medicines
TDR	Total drug resistance
TS	Thymidylate synthase
TTD	Therapeutic target database
WT	Wild type
XRD	Extreme drug resistance
ZINC	Zinc Is Not Commercial (ZINC database)

1 Introduction

1.1 *Multi-Target Therapies*

Traditional pharmacology-based approaches are focused on the identification of compounds/inhibitors which can selectively and effectively bind to a single therapeutically important target (mostly protein). This one target-one drug strategy is historically the main

theme of most of the drug discovery endeavors. With the emergence of new areas such as system biology, network pharmacology, polypharmacology, and drug repurposing, the concept of inhibition of single selective target by a drug is becoming increasingly questionable and outdated. These new disciplines believe that the effect of a drug on an organism is the result of its several complex and coordinated interactions with several biomolecules/targets [1–5]. In reality, most of the conventional drugs are essentially multi-targets (i.e., they interact with more than one target), and their major interaction/s depend on the site of action. For example, HIV-1 protease inhibitors mainly interact with protease molecule in the main event, but it has to be catabolized by some oxidases such as cytochrome P450 (CYP450) enzymes after its action [6]. Side effects by many known therapies/drugs are also indicative of low *off-target* selectivity of a drug and support the concept of multi-target drug models.

At present, numerous terms have been used to define the compounds which are intentionally designed to inhibit more than one target such as dual inhibitors, triple inhibitors, multi-target drugs, dual blocker, dual antagonist, multipotent agents, multi-functional compound, and heterodimer just to name a few [5]. Morphy, who has done pioneering work in this field, suggested to use the term designed multiple ligands (DMLs), to describe “compounds whose multiple biological profile is rationally designed to address a particular disease, with the overall goal of enhancing efficacy and/or improving safety” [5]. The term multi-target drug/inhibitor (hereafter abbreviated as MTDs) is still very common in literature and has been used throughout this chapter.

It is also important to realize that MTDs are significantly different from highly promiscuous compounds such as pan-assay interference compounds (PAINS), which have very high activity against a range of biological targets [7]. No doubt, MTDs are active against more than one target, but these targets are purposefully selected for a particular disease/s without negotiating its safety profile. The study of activity cliff, PAINS, and bright and dark spots in biochemical space can undeniably provide valuable information/s for designing newer and more effective MTDs [7].

Some examples of successful MTDs include anticancer kinase inhibitors such as sunitinib, dasatinib, and lapatinib (Fig. 1) [8, 9]. These drugs target the main therapeutic target along with the alternative signaling pathway (escape mechanism) hence resulting in effectively breaking the cross talk between the two paths. Similarly, the nonsteroidal anti-inflammatory drug (NSAID) aspirin was initially known as a selective COX inhibitor. Later on, it was found that it renders its anti-inflammatory properties by multiple pathways along with COX, such as interference with transcription

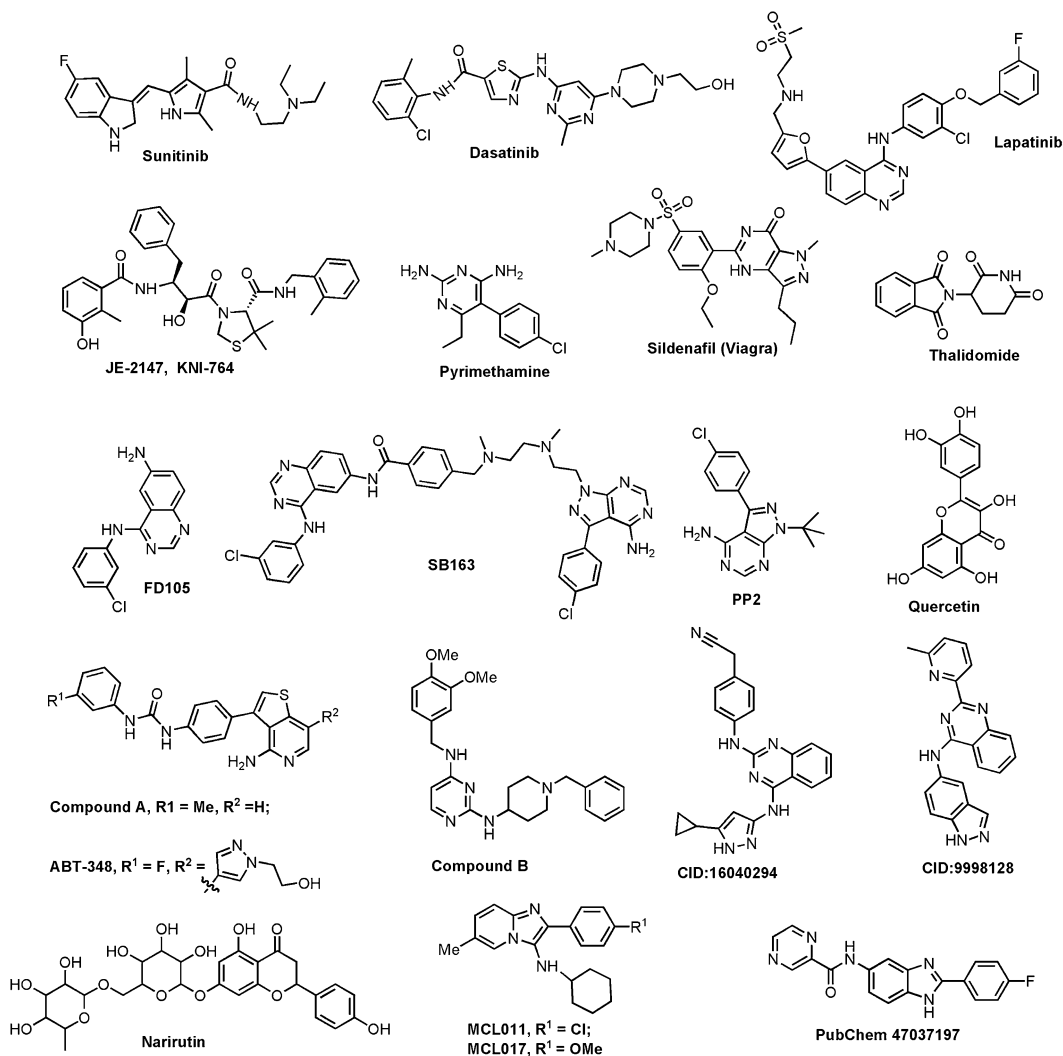


Fig. 1 Structures of selected multi-target inhibitors and other bioactive compounds mentioned in this chapter

factors [e.g., nuclear factor (NF)-κB and mitogen-activated protein kinases (MAPKs)] etc. which makes it even more effective. In fact, most of the known MTDs were first discovered as single target drugs for a protein, and later on, some additional target/s were identified [10, 11].

1.2 Multi-Target Therapies and Microbial Infectious Diseases: A Unique Connection

In the particular area of infectious diseases such as neglected tropical diseases (NTDs) including malaria and tuberculosis, application of MTDs can be very crucial. Genesis and spread of resistant pathogenic forms has seriously compromised the efficiency of current drug arsenal. For example, quinolones, antifolates, and related

compounds, which were initially the first-line medications for malaria, have lost their effectiveness due to the emergence of new mutant forms of *Plasmodium* species [12, 13]. There are also some initial reports of artemisinin resistance in western Cambodia and other regions of East Asia [14, 15]. Similarly, multidrug resistant (MDR), extreme drug resistant (XDR), and total drug resistant (TDR) forms of *Mycobacterium tuberculosis* are also difficult to treat by drugs traditionally used to treat TB [16–18]. Most of these drugs are the outcomes of single target-based strategies, and even a single mutation in amino acid sequence of target protein is enough to render them ineffective. As what has happened with antibiotic resistance (or antimicrobial resistance, AMR), it is quite likely that the single target drug/s of future may also meet similar fate. Considering the present situation and unavoidable resistance of future single target therapies, the design and screening of MTDs is the need of the hour. Drugs, which are deliberately aimed for more than one target of therapeutic interest, offer better safety profile and low resistance against pathogens. Since MTDs simultaneously interact with more than one target, their low dosage can bring the desired effect. Furthermore, possibility of concurrent mutation in all the involved targets is highly unlikely. These two factors contribute to overall low resistance probability of drug by pathogen (Fig. 2) [5, 6, 19].

At this point, it would also be interesting to compare MTDs with combination therapies, which are now very common in some of the infectious diseases such as malaria (e.g., quinine + tetracycline/doxycycline, artesunate + mefloquine etc.). MTDs are expected to have a better safety profile as compared to the cocktail of drugs individually targeting a specific protein (combination therapy), since in the latter case, each component drug has its own selectivity profile [3, 4]. As compared to the combination therapy, MTDs are also theoretically easy to design as well. Consider an imaginary scenario, where each




 Multidrug regimen and combination therapy	 Single target Therapy	 Multi-target therapy
Many drugs- Many targets	One drug-One target	One drug-Many targets
More side effects & Drug-Drug interactions	Low side effects	Low side effects
High cost in R&D \$\$\$	Low cost in R&D \$\$	Low cost in R&D \$\$
Low resistance probability	High resistance probability	Low resistance probability
Sulfadoxine-pyrimethamine (SP) (Fansidar)	Vancomycin	Kinase inhibitors sunitinib, dasatinib, lapatinib

Fig. 2 A comparison between different therapies of drug designs

candidate has to pass ten requirements such as potency, selectivity, half-life, ADMET, etc. to be a successful drug. In this situation, the total numbers of factors for a dual inhibitor (MTD targeting two targets of therapeutic relevance) will be $10 + 1 = 11$ (ten initial factors + one additional factor, potency). While at the same time, for a new combinational therapy that targets the same two targets, the total number of criteria will be at least $10 + 10 = 20$ (ten for each component drug and two drugs). Hence theoretically, it will be easier to design a MTD than to design two or more component drug for combination therapy (Fig. 2) [19].

In the specific context of NTDs, where economics of drug research and development (R&D) is largely a limiting factor, MTDs offer a cost-effective alternative. In principle, the cost of development of a new MTD should be comparable to that of a single target drug, which is supposedly lower than the development of two or more component drugs for combination therapy/multi-drug regimen (Fig. 2) [6].

Additionally, many of the infectious diseases such as malaria and tuberculosis are opportunistic in nature. For example, infection of HIV is largely encountered with secondary infections such as tuberculosis and/or malaria. As the immune system of AIDS patient ceases, he/she becomes more and more prone to microbial attack. A combination of several drugs (multidrug regimens) is used to relieve the patient with many possible side effects and drug-drug interaction/s [20, 21]. Design and development of a MTD, which can suppress both the coinfection (HIV and opportunistic infections) at once, would be a highly desirable situation. In fact, some initial attempts have already been made in this direction. For example, KNI-764 (or JE-2147) (Fig. 1) shows significant HIV-1 protease and plasmepsin II (*Plasmodium malariae*) inhibitory activity [22]. If successful, current multidrug regimen for HIV treatment may be replaced by a single and safer multi-target regimen in the foreseeable future.

In the next sections, we will discuss different strategies for designing MTDs, available resources, and a case study for designing dual inhibitor against tuberculosis and malaria in silico.

2 Methods

2.1 Overview of Multi-Target Screening Methods

There are broadly two different methods for designing MTDs, relying on two different concepts. Knowledge-based approaches are based on the existing knowledge (such as scientific databases, literature including patents, etc.) to commence a MTD project, while in screening-based approach, different filters are applied to find out some novel multi-target hits.

2.1.1 Knowledge-Based Approaches

System Biology,
Networking, and
Polypharmacology-Based
Approaches

With the advancements made in the field of structural biology and instrumental techniques, we have now more access to biological information which is stored in different structural data banks (such as 3D structures of targets, activity and toxicity data, *off-target* selectivity, etc.). Capturing this existing knowledge is very important for the successful completion of a MTD project [3, 4].

Network-based approaches only navigate the available biochemical space for the search of new MTDs. Bigger size of this accessible biochemical space means greater chances of success in MTD efforts. Unfortunately, there is a striking contrast in the speed of chemical and biological data with which it flows toward public databases. While most of the biological data such as protein structure, sequence, and function had been historically available in the open/publically accessible repositories such as PDB, UniProt, etc., most of the chemical databases have remain commercial/private and not freely available. Only in the past decades, several public collections of chemical databases have emerged such as PubChem, ChemDB, ZINC, etc. Finally, it is also important to integrate these two sets of information through appropriate connections/projections, so that important inferences can be mined [23]. Programs such as STITCH [24] can be very useful to study these interconnections. An example of this approach is explained below.

Pyrimethamine (Daraprim, Fig. 1) is primarily used as a medication for protozoal infections including *Plasmodium* and *Toxoplasma*. It mainly inhibits bifunctional dihydrofolate reductase/thymidylate synthase (DHFR-TS) of the pathogen and inhibits its effect. It also binds with some important human enzymes such as cytochrome P450 (CYP2), mitochondrial peptide deformylase (PDF), signal transducer and activator of transcription 3 (or acute phase response factor, STAT3), hexosaminidase B (HEXB), etc. (Fig. 3). Side effects by this drug such as nausea, vomiting, glossitis, anorexia, and diarrhea, etc. can be the result of these interactions [25]. Its interaction with HEXB (Fig. 3) is clinically important as it slows down the progression of late-onset Tay-Sachs and Sandhoff disease [26]. Pyrimethamine is under clinical trial for this application. This information on the binding of pyrimethamine with important human *off targets* has been collected from different data banks using STITCH program [24]. Color code has been used to display protein-protein and chemical-protein interactions (Fig. 3).

Approaches Based on
Protein-Protein
Relationship and
Evolutionary History of
Proteins

Along with this network information, the evolutionary history of different protein families can also be a valuable resource for designing newer MTDs. Proteins which are close in the evolutionary tree may be structurally and functionally similar, and it would be easier to design a common inhibitor for these proteins as compared to the proteins with different origins. Evolutionary dendrograms (tree

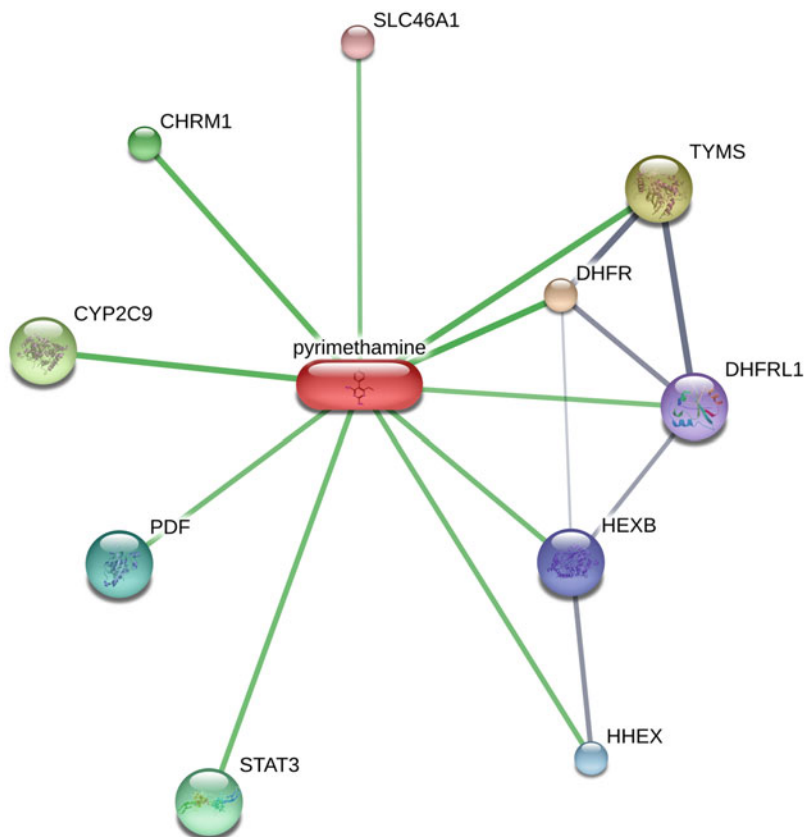


Fig. 3 Drug-target relationship network for pyrimethamine extracted from STITCH program [24] (stronger associations are represented by thicker lines. Protein-protein interactions are shown in gray, chemical-protein interactions in green and interactions between chemicals in red)

diagram) should always be examined to find out protein-protein relationship [6, 23].

Approaches Based on Drug Repurposing/Drug Repositioning

One other knowledge-based route to find new multi-target leads can be drug repurposing (also known as drug repositioning). The main benefit of this strategy is that researcher needs not to think much about other liabilities such as selectivity and safety. Pfizer’s Viagra and Celgene’s thalidomide (Fig. 1) are the two well-known examples of drug repurposing [27]. Many FDA-approved drugs of the developed world had already been tested in target-specific assays to examine their interactions with important protein families; unfortunately, this is not the case for many of the drugs related to neglected tropical/infectious diseases, and drug repurposing can be an important starting point for the discovery of some important multi-target leads [19].

Multi-Target Inhibitors by Design

Conceptually, hybrid pharmacophores (compounds having more than one pharmacophore, sometimes called combi-molecules) can retain their activity against each of the specific targets. These structurally designed MTDs fall into categories such as conjugates, fused, and merged [5, 28]. In conjugates, all the distinct pharmacophore domains are connected by a cleavable or non-cleavable linker. Linker of cleavable conjugates is metabolically unstable and released both (or all) the connected units after enzymatic action and all the free ligands then act separately. As the size of linker decreases, a situation comes where all the pharmacophoric frameworks just touch each other and this results into a fused hybrid pharmacophore. In another scenario, the common structural features of different functional domains can be coalesced, giving rise to a smaller “merged pharmacophores.” Most of the hybrid pharmacophore projects either commence with the incorporation of two already known and distinct functional pharmacophores or by adding a structural component in a known drug. Fu’s group has extensively reviewed many structurally designed multi-target inhibitors in anticancer drug discovery paradigm [28]. A classic example of multi-target agent targeting both Src family kinase and EGFR tyrosine kinase is presented by Barchéath S and coworkers [29]. Combi-molecule SB163 is a conjugate hybrid pharmacophore which is designed by connecting PP2 (Src kinase inhibitor) and FD105 (EGFR tyrosine kinase inhibitor) using a suitable linker (Fig. 1). The hybrid SB163 displays an IC_{50} of 2.9 μ M (EGFR) and 0.32 μ M (Src) which is superior to that of an equimolar combination of Iressa and PP2 in the Boyden chamber assay [29].

2.1.2 Screening-Based Approaches

Screening-based approaches are based on the actual or computational filtering of large databases. As compared to single target screening, these multi-target screening approaches are less common and only recently reviewed [1–6]. Since more than one target or activity series (QSARs) are involved in designing a multi-target model, extra precision should be taken to ensure the robustness of the model [4, 5].

Multiple Phenotypical/Enzyme-Based Assays in Parallel

A typical example of screening-based approach was the discovery of a multi-target kinase inhibitor involved in tumor progression [30]. Abbott kinase proprietary was screened against a broad panel of kinases (pan-kinase assays), and it was found out that compound A (a thienopyridine urea analogue, Fig. 1) displays inhibitory activity of $IC_{50} = 9.0$ nM (KDR, cellular) and $IC_{50} = 32.0$ nM (enzymatic) along with moderate aurora B activity [$IC_{50} = 487.0$ nM (enzymatic), 42,000 nM (cellular)]. Further structural modification at two different sites (thienopyridine C7 position and urea terminal phenyl position) led to the discovery of ABT-348 with acceptable

aurora B VEGF and PDGF activities (Fig. 1). This molecule is currently in phase I/II clinical trial against solid and hematological cancers [30]. Similarly, the screening of a collection of 2-benzylpiperidin-4-amines against anti-cholinesterase (ACHE and BuCHE), anti-A β aggregation (ACHE and self-induced), and anti- β secretases (BACE-1) led to the identification of multi-target inhibitor compound B, [N²-(1-benzylpiperidin-4-yl)-N⁴-(3,4-dimethoxybenzyl)pyrimidine-2,4-diamine] (Fig. 1) [31]. This compound has a potential to develop as a disease-modifying agent (DME) for Alzheimer's diseases.

In Silico Screening of Multi-Target Inhibitors

While experimental assays are time, labor, and cost intensive with high attrition rate, *in silico* screening approaches can find a solution. Now, we will discuss some computational (*in silico* or *in computo*) models for MTD screening which is also the theme of this chapter.

In Silico Target Identification or Reverse Docking

The simplest screening method is based on *in silico* target identification (or reverse docking). By structural similarity (2D or 3-D), a compound can be screened against all the cognate ligands/innate inhibitors available in Protein Data Bank (PDB) or similar structural databases. In theory, if a compound is structurally similar to a known inhibitor, chances are there that it may also act as an inhibitor for that specific target. Computational target identification methods such as SHAFTS [23] use this concept for identification of targets. Multi-target activity of quercetin (a polyphenolic phytochemical, Fig. 1) was determined by screening it against cognate inhibitors in PDB by SHAFT and binding site similarity (idTarget, LIBRA). The identified targets fall into two broad categories protein kinases and poly [ADP-ribose] polymerases which are important in conditions like inflammation, neurodegeneration, and cancer [32].

Parallel Docking, Multiple QSAR, and Other Computational Screening Methods

Docking- and multi-QSAR-based screening approaches have also been frequently employed for identification of multi-target hits. In the simplest form, if a ligand collection is screened using two different proteins (in docking, structure based) or using two distinct QSAR models (ligand based), then the common hits in parallel runs represent the starting point for multi-target leads. An example of this parallel screening approach was presented by Li group [33]. In a quest of novel multi-target inhibitors against Alzheimer's disease, this group screened 1.4 million compounds from PubChem data bank using their druglikeness and molecular docking. Five targets including amyloid-beta fibril, peroxisome proliferator-activated receptor γ , retinoic X receptor α , and β - and

γ -secretases were used in the study. Only common hits to all the five receptors with AutoDock Vina score $\Delta E \geq -8.0$ kcal/mol were selected for further MD simulation. Finally, two compounds CID 16040294 and CID 9998128 were identified as multi-target hits (Fig. 1). Similar structure-based screening of ZINC + PubChem + Molecule database (about three million molecules) led to the identification of multi-targeted (EGFR + HER2 + HSP90) inhibitors against breast cancer (Fig. 1) [34].

In most of the cases, computational docking- and QSAR-based approaches are used in stepwise manners along with several other filters such as Lipinski's rule of fives, druglikeness, ADMET, etc. For example, Basu and coworkers successfully applied structure-based ensemble docking (using different closed conformations of BACE1 such as 3IND and 3TPP) and a four parameter-based 2D QSAR model for the screening of an *in-house* library of 200 phytochemicals [35]. The top 65 compounds (based on energy criteria) were selected for the second round of screening where these were again filtered using 2D QSAR model for anti-amyloidogenic and antioxidant activity. A polyphenol narirutin (Fig. 1) was the common hit in both the 2D QSAR exercises and eventually tested for its BACE inhibitory activity (enzyme-based fluorescence titration assay), anti-amyloidogenic (ANS and ThT assay), and antioxidant activity (ABTS scavenging assay). Narirutin displayed significant BACE and anti-amyloidogenic activity with moderate antioxidant activity, making it a suitable multi-target hit for Alzheimer's disease. This group also explored some other multi-target hits for the same disease using similar computational strategy [36, 37]. Jiang group has also explored PubChem database by a three-tier screening for the discovery of a multi-target (EGFR, VEGFR-2, and PDGFR) compound CID:47037197 (Fig. 1) [38]. Their method involved the use of support vector model (SVM), druglikeness criteria, and computational docking by Discovery Studio 2.5/CDOCKER.

Despite its relevance, there are only a few scholarly articles dedicated to the design of MTDs especially in the domain of neglected tropical and infectious diseases [6, 39–42]. In the next section, a molecular docking-based strategy has been presented for the discovery of dual inhibitors against malaria and tuberculosis. This case study is based on authors' published research work [39].

2.2 Screening Dual Inhibitors Against Malaria and Tuberculosis by a Three-Tier Computational Approach: A Case Study

Presented three-tier screening is based on parallel computational docking, and several variations are possible depending upon the preferences and resources available to medicinal/computational chemists (*see Note 1*) [39]. Conceptually, the similar methods can be extended to other research domains of MTDs. Computational docking by AutoDock is the key event of this strategy, so a step-by-step protocol is given in detail (Fig. 4).

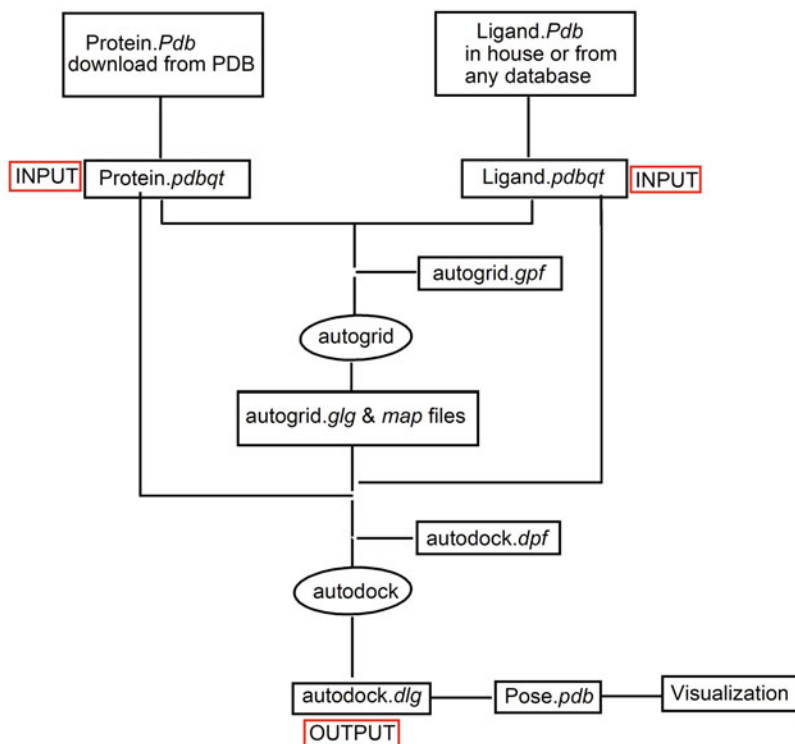


Fig. 4 A step-by-step flowchart of a docking operation using AutoDock

Selection of Protein: Selection of suitable targets (receptors) is a key step for any multi-target screening exercise.

1. Two essential criteria which must be satisfied by a target are as follows: (1) it should have therapeutic weight (in case of infectious diseases, it should be an enzyme/protein of vital importance for the survival of microbe/pathogen), and (2) it must be druggable (it must be capable of binding to a small molecule, and its activity can be modulated by that small molecule).
2. It is highly recommended to survey the literature to examine the validity of a target (linkage and site mutation studies), its physiological role, *off-target* relationship, etc. before finalizing its selection. Previous docking studies can be very helpful in this regard.

In the presented case study, ten well-validated targets of *Plasmodium*, *Mycobacterium*, and *Trypanosoma* were selected [39]. These include wild-type (WT) forms of *Pf*-DHFR-TS (1J3I.pdb, [43]), *Pf* enoyl-ACP-reductase (1NHG.pdb, [44]), *Pf*-PK7 (2PMN.pdb, [45]), *Mt*-SK (2DFN.pdb, [46]), *Mt*-PS (1N2H.pdb, [47]), *Mt*-TMPK (1G3U.pdb, [48]), *Mt*-MurE ligase (2WTZ.pdb, [49]), *Tc*-TR (1BZL.pdb, [50]), *Tc*-G3PD (1QXS.pdb, [51]), and *Tc*-TS (1S0I.pdb, [52]).

2.2.1 Protein Preparation

1. Protein files can be retrieved from Protein Data Bank (RCSB PDB, www.rcsb.org, [53]) in *pdb* format with maximum possible resolution.
2. File should be examined for any structural error/s, missing residue/s, inappropriate charge/s, wrong tautomeric state/s, etc. and should be corrected.
3. If protein molecule contains any cognate ligand/inhibitor/substrate, it should be removed first. [For example, *Pf* DHFR-TS (1J3I) contains a third-generation inhibitor WR99210 (or WR 609A).]
4. If the protein is homodimer (homomultimer) and the binding cavity is not present at the area of intersection, a single monomer can be dissected.
5. Polar hydrogen atoms are added, and Gasteiger charges are computed using AutoDock tool (ADT 4.2, <http://autodock.scripps.edu/>) [54].
6. The target structures are cleaned by removing crystallographic/bound water molecules. Finally, protein files are saved as *pdabt* files.

2.2.2 Ligand Preparation

1. A ligand is the small molecule which will be docked inside the binding cavity of the macromolecule. There are many ready to use small molecule repositories available for this purpose (can be directly downloaded in various file formats such as *sdf*) such as natural products, phytochemicals, commercially available compounds from different suppliers, known drugs, etc. *In-house* collection of focused or diverse set of compounds can also be built.

A small collection of 30 imidazo-azines (with different substituent patterns) was used by the authors [39]. The selection of this library was based on the therapeutic profile of this class of *N*-fused bicyclic heterocycles and their easy accessibility through Groebke-Blackburn-Bienymé multicomponent reaction [55].

2. 2D structure of a ligand is drawn using a suitable drawing tool such as ChemDraw (<http://www.cambridgesoft.com>). Chiral and stereochemical positioning of a group in space should be drawn with extreme care using suitable bond representation such as wedge bond (up to plane), hatch bond (down to plane), or wiggly bond (stereo configuration is irrelevant or unavailable).
3. The 2D structures are converted to 3D structures. Simplest way to do this is to copy 2D file and paste it to ChemDraw 3D or by using a file format converter tool such as Open Babel (<http://openbabel.org>).

4. 3D structures are energy minimized/geometry optimized using a suitable force field such as MM2, AM1, PM3, CHARMM, etc. depending on the size and flexibility of the ligand.
5. Finally, all the 3D structures are saved as *pdb* format. The *pdb* format contains information regarding the coordinate and connectivity of all the constituent atoms along with the annotation of side-chain rotamers and secondary configuration.
6. Finally, ligand *pdb* file is loaded to ADT, and information such as AutoDock 4 atom types, Gasteiger charges (if necessary), nonpolar hydrogens, aromatic carbons (if any), and the torsion tree (for information of bond flexibility) are added to it and again saved back to *pdabt* [*pdb* (atom coordinate and connectivity) + *g* (partial charge) + *t* (AutoDock 4 atom types)] format. This *pdabt* file is the main input file for grid setting and docking by AutoDock.

2.2.3 Grid Setting and Calculation of Pre-energy Maps

1. Setting a grid (a square box) around binding cavity of protein is an important step of the AutoDock protocol. Each atom type of the ligand (aliphatic carbon, aromatic carbon, hydrogen-bonded oxygen, etc.) is used as an imaginary probe, and their interaction energy at each regularly spaced point (grid points) is calculated. In this way, energy files (*map* files) for each atom type of ligand are calculated. In addition, AutoDock version 4 and electrostatic and desolvation energy files are also calculated.
2. Size of the grid box should be larger than the binding cavity of the target protein and that of the fully extended form of the ligand/s under study. The default spacing is 0.375 Å between grid points which can be changed especially when a large volume is to explore.
3. Binding site residues can be identified by opening the protein *pdb* file into some visualizer such as Discovery Studio (DSV, <http://accelrys.com>), Molegro Molecular Viewer (MMV, www.qiagenbioinformatics.com), PyMOL (<https://pymol.org>), etc. or by browsing the motif and site option available in Protein Data Bank in Europe (EMBL-EBI, <http://pdbe.org/motif/>) for each cognate ligand present in the protein.
4. All the information about position and size of the grid, the types of map to be calculated etc. are stored in the form of *gpf* (grid parameter file). Once *gpf* is generated, AutoGrid is run to calculate all *maps* and *glg* (grid log file) files which are the precondition for running AutoDock (Fig. 4).

2.2.4 Execution of Docking Program

1. A specific file called docking parameter file (*dppf*) has to be built before performing docking experiments. This file contains all the information about the size of docking calculation (number of run, number of final pose, etc.), which algorithm is to use and which grid maps are to use etc. for docking.
2. There are four different types of algorithm presently available in AutoDock 4.2 [49]: Monte Carlo simulated annealing (SA), traditional Darwinian genetic algorithm (GA), local search (LS), and a combination of genetic algorithms and local search [GA-LS, also called LGA (Lamarckian genetic algorithm, [56]). LGA was shown to be the better than other three and mostly used to run the docking experiments. In LGA local environmental adaptations of a “parent” are allowed to inherit back to “offspring” [56].
3. Some important parameters of the search algorithm are *ga_pop_size* (number of individuals in population, default value 150), *ga_num_evals* (maximum number of energy evaluations, default value 2.5×10^7), *ga_num_generation* (maximum number of generations, default value 27,000), *ga_elitism* (number of top individuals to survive to next generation, default value 1), *ga_mutation_rate* (rate of gene mutation, default value 0.02), *ga_crossover_rate* (rate of crossover, default value 0.8), *ga_run* (number of hybrids GA-LS run, default value 10), etc.

ga_num_evals and *ga_num_generations* determine the size of calculations, while *ga_run* determines the numbers of final poses in the output file. If ligand has more flexible and rotatable bonds, then size of the computational calculations can be increased.

4. Once all the files (*pdbqt* of receptor and ligand + *map* files for all atom types + electrostatic and desolvation map files + *dppf*) are generated, AutoDock can be launched (Fig. 4).

2.2.5 Visualization of Docking Output

1. The docking log file (*dlog*) is the output file of the docking process; it contains important information such as spatial coordination of all the poses, binding energies ($-\Delta G$), clustering, histogram, root-mean-square deviation (RMSD), etc. By opening this file as a notepad and then typing a keyword “histogram,” the entire docking summary can be viewed in the table format.
2. In ADT all the conformations can be visualized separately or at once. Mostly *dlog* files are converted back to *pdb* files via *pdbqt* using ADT itself or with the help of a file converter. These files then can be opened in any 3D visualize tool such as DSV, MMV, PyMOL, etc.
3. Finally, these conformations are studied for their interactions with the neighboring residues of the proteins. Van der Waals

(VdW) electrostatic hydrogen interactions are especially looked for. The location of hydrophobic and hydrophilic residues around the ligand is also important. These interactions provide a rational base for the calculated binding energy.

2.2.6 Analysis of Results

1. Analysis of docking result is a tricky and subjective exercise (*see Note 2*). While mostly the lowest energy poses are given preference in the ranking, pose belonging to the most populated cluster is also important (because of their high reproducibility in the repeated runs) [57].
2. Binding energy (BE) is definitely an important criterion. Generally it is advisable to compare the binding energy of the docking pose with that of the already known inhibitors using it as a benchmark for the study.
3. In docking-based multi-target screening, all the docking exercises are set up in parallel, and all the common hits are considered as a starting point, though more optimization may require further. For example in multi-tiers strategy proposed by the authors, the following criteria were selected (Fig. 5):
 - (a) Re-docking/control docking including flexibility and clustering pattern of the bench marks for those protein ligand sets where docking engine was failed to meet rmsd criteria of 2 Å in validation experiment was not simply considered for the next round. Similarly the cases where too many clusters were generated due to high flexibility of benchmark (more than 15 flexible bonds) were also rejected (*see also Notes 3, 4 and 5*). Although implementation of this criterion led to the elimination of four proteins from the multi-tier screening, yet it significantly improved the reliability of the model (Fig. 5).
 - (b) Filter based on potency/ligand efficiency (LE): it has been shown that as compared to binding energy, binding energy per heavy atom (or ligand efficiency, LE) is a better representative of potency [58–60]. So a common baseline was set up by averaging the LEs of all the cognate ligand (also served as benchmark). This LE base line (≈ -0.35 kcal/mol per heavy atom) was quite close to consensus LE (-0.30 kcal/mol per heavy atom) proposed for fragment-based lead discovery. The hits ($LE \geq -0.35$ kcal/mol per heavy atom) which were common in all the proteins were selected for the next phase of multi-target screening (Fig. 5) [39].
 - (c) Filter based on druglikeness/leadlikeness: nowadays the potency of a molecule is not a sole criterion for making it hit; other factors such as druglikeness are also very decisive (Fig. 5) [59, 60]. The main benefit of this approach is to

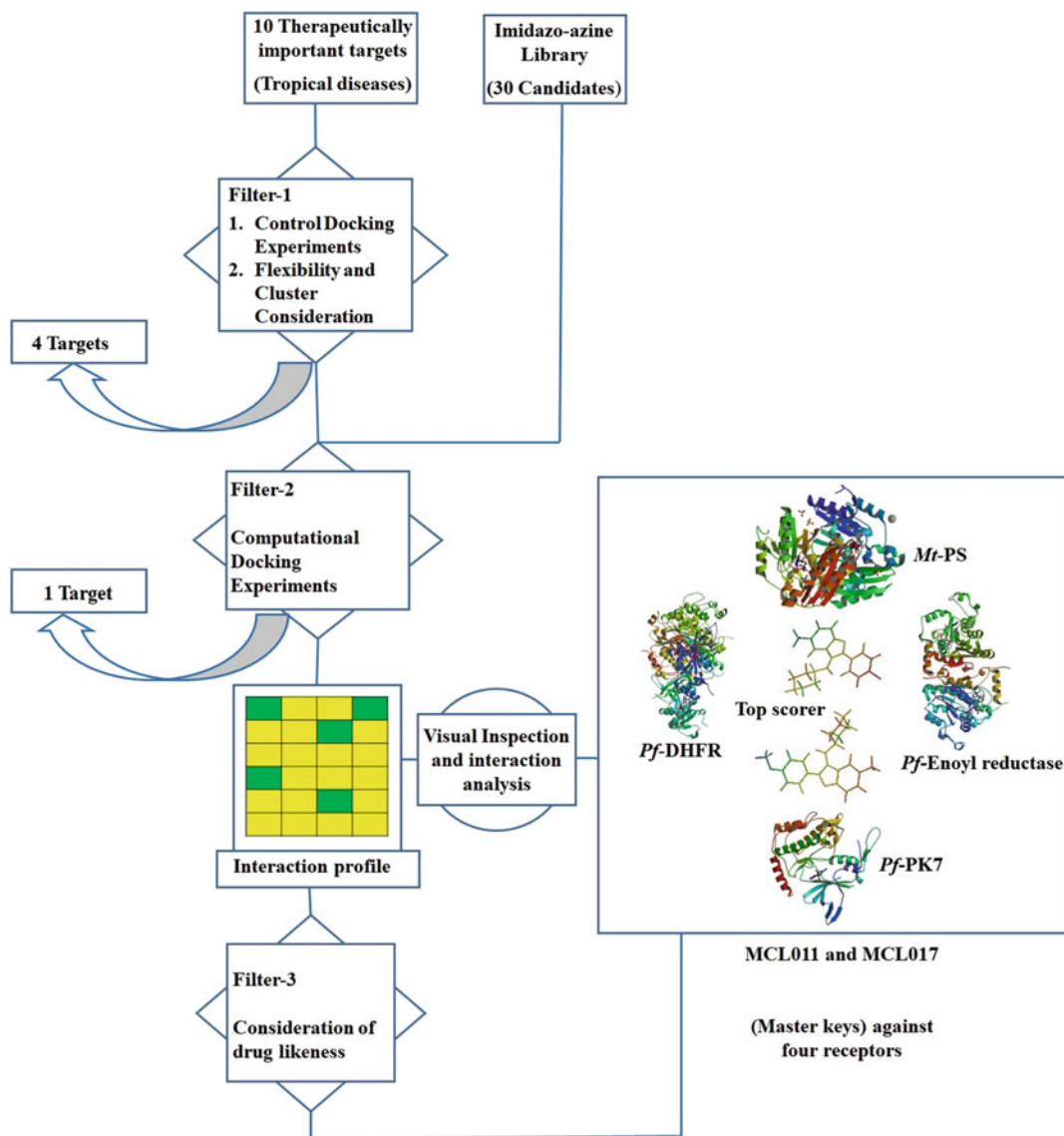


Fig. 5 Details of the strategy used in multi-target screening of imidazo-azine library (adapted from ref. 37)

identify drug-like hits in the initial stages of drug discovery with minimum numbers of liabilities in the advance stages. The following criteria were selected based on various literature reports [61, 62]:

- (d) Molecular weight (MW) < 350 Da, $\log P(\text{octanol/water}) = 3-5$, maximum no. of rings (N_{ring}) < 4, maximum no. of nonterminal single bonds (N_{ter}) < 10, maximum no. of hydrogen bond donors (HBD) < 5, maximum no. of hydrogen bond acceptors (HBA) < 8, and ligand efficiency (LE) ≥ 0.30 kcal/mol/non-hydrogen atom.

After applying all these filters, two compounds 2-(4-chlorophenyl)-*N*-cyclohexyl-6-methyl*H*-imidazo[1,2-*a*]pyridine-3-amine (MCL011) and *N*-cyclohexyl-2-(4-methoxyphenyl)-6-methyl*H*-imidazo[1,2-*a*]pyridine-3-amine (MCL017) (Fig. 1) were selected as a multi-target hits. Like a master key (which can be used to open more than one lock), these hits can be used to design multi-target leads against multiple targets [39].

3 Materials and Tools

Many computational tools are available for multi-target screening. Some of the tools/software/online servers/repositories used in the three-tier protocol are given in detail:

3.1 Macromolecular Database

The Protein Data Bank (RCSB PDB, www.rcsb.org, [51]) is a freely available database for 3D structures of macromolecules such as protein and nucleic acid. The structural data are obtained by techniques such as X-ray crystallography, NMR spectroscopy, cryo-electron microscopy, etc. It was established in 1971 at the Brookhaven National Laboratory (BNL), New York. It also has some sister organizations such as PDBe (Protein Data Bank in Europe), PDBj (Protein Data Bank in Japan), etc. Some other databases also rely on PDB for structural information such as SCOP (<http://scop.mrc-lmb.cam.ac.uk/scop/>, for protein structural classification), PDBREPORT (<http://swift.cmbi.ru.nl/gv/pdbreport/>, database of structure error in macromolecule structure), PDBsum (<http://www.ebi.ac.uk/pdbsum/>, summarize data from other databases about PDB structures), etc. From 2003, worldwide Protein Data Bank (wwPDB, <https://www.wwpdb.org/>) came to existence as a single macromolecule data bank founded by RCSB PDB, PDBe, PDBj, and Biological Magnetic Resonance Data bank (BMRB, <http://www.bmrwisc.edu/>). Some other data banks and resources are UniProtKB/Swiss-Prot (<http://www.uniprot.org/>, protein sequence database), STRING (<https://string-db.org/>, protein-protein interactions), SWISS-MODEL Repository (protein homology models, <https://swissmodel.expasy.org/repository>), TTD (Therapeutic Target Database, <http://bidd.nus.edu.sg/group/cjttd/>), etc.

3.2 Small Molecule Databases

There are several private and publically available databases for small molecules which can be directly used for a screening campaign. Some notable examples are:

ZINC (Zinc Is Not Commercial, <http://zinc.docking.org/>) is a freely available collection of commercially available compounds (over 35 million compounds, in ready to dock format) including information about their suppliers.

ChEMBL (<https://www.ebi.ac.uk/chembl/>) is an open data bank of small drug-like compounds along with information about their binding and ADMET properties.

DrugBank (<https://www.drugbank.ca/>) is a freely available database for comprehensive information about drugs and drugs targets (mainly important for drug repurposing). It is both bioinformatics and a cheminformatics resource.

TCM-Taiwan (traditional Chinese medicines, <http://tcm.cmu.edu.tw/>) is an natural product database which contain information about different TCM constituents (1,70,000 entries in 2D and 3D format).

PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) is an open chemistry database maintained by National Institute of Health (NIH), USA. It contains information about chemical structures, identifiers, chemical and physical properties, biological activities, patents, health, safety, toxicity data, etc.

Other important database includes Mcule (<https://mcule.com/>), ChemDB (<http://cdb.ics.uci.edu/>), ChemBank (<http://chembank.broadinstitute.org/>), DUD (<http://dud.docking.org/>), KEGG DRUG (<http://www.genome.jp/kegg/drug/>), SuperDRUG2 (<http://cheminfo.charite.de/superdrug2/>), eMolecules (<https://www.emolecules.com/>), etc.

3.3 Tools for Computational Docking

AutoDock (<http://autodock.scripps.edu/>, [54]) is a suite of automated docking tools. It is freely available and maintained by the Molecular Graphics Laboratory (MGL, <http://mgl.scripps.edu/>) and the Scripps Research Institute, La Jolla, USA. Its current version AutoDock 4 consists of two main programs: *AutoGrid* for calculating pre-energy maps which were then used by another program *AutoDock* for molecular docking. It has option for four different types of search algorithm, but genetic algorithms and local search [GA-LS, also called LGA (Lamarckian genetic algorithm)] are the mostly implemented. AutoDock 4.2 has an empirical free energy-based scoring function which uses linear regression analysis in AMBER force field.

There are several other docking software tools are UCSF DOCK [<http://dock.compbio.ucsf.edu/>, freeware for academic user, Kuntz's Group in University of California, San Francisco (UCSF)], GOLD [<https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/>, commercial, from the Cambridge Crystallographic Data Centre (CCDC)], FlexX (<https://www.biosolveit.de/FlexX/>, commercial, from BioSolveIT), Glide (<https://www.schrodinger.com/glide>, Commercial, from Schrödinger), ICM-Dock (<http://www.molsoft.com/docking.html>, commercial, MolSoft Inc.), MOE (<https://www.chemcomp.com/>, commercial, from Chemical Computing Group), Surflex-Dock (<http://www.jainlab.org/downloads.html>, commercial, from Tripos), SwissDock [<http://www.swissdock.ch/>, free web

service for academic user, from Swiss Institute of Bioinformatics (SIB)], GEMDOCK (<http://gemdock.life.nctu.edu.tw/dock/>), freeware, from National Chiao Tung University, Taiwan), etc.

3.4 Software Tools for Visualization

These software tools are used for the visualization, editing, and studying protein-ligand interaction. Some examples are Discovery Studio Visualizer (DSV, <http://accelrys.com>) (freeware for academics, now distributed by BIOVIA Inc.), UCSF Chimera [<https://www.cgl.ucsf.edu/chimera/>], freeware, from Resource for Biocomputing, Visualization, and Informatics (RBVI)], ICM Browser (http://www.molsoft.com/icm_browser.html), freeware, from Molsoft LLC), Jmol (<http://jmol.sourceforge.net/>), freeware, widely used, developed by Jmol developer team), PyMOL (<https://pymol.org>), commercial, from Schrödinger), RasMol (<http://www.openrasmol.org/>), freeware, developed by Roger Sayle), Swiss-PdbViewer aka DeepView [<https://spdbv.vital-it.ch/>], freeware, by Nicolas Guex, Swiss Institute of Bioinformatics (SIB)], etc.

3.5 Drawing Tools (Molecular Editors) and File Format Converters

Drawing tools are used for the drawing, processing, storing, rendering, and editing the 2D structural information. Some example includes ChemDraw (<http://www.cambridgesoft.com>), developed by CambridgeSoft), BIOVIA Draw (<http://accelrys.com>), formerly Accelrys Draw and before that Symyx Draw, by BIOVIA), ChemDoodle (www.chemdoodle.com), by iChemLabs), MarvinSketch/View (<https://chemaxon.com>), by ChemAxon), etc.

During different steps of computational docking (and other operations), formats of different files need to be converted to other suitable formats. Open Babel (<http://openbabel.org>) is a toolbox which can read, write, and convert over 110 chemical file formats.

4 Notes

1. In theory, tools for computational screening of multi-target inhibitors are similar to that of single target inhibitors. It's just the integration of these tools which is differing.
2. Multi-target screening exercises are more prone to errors and false input etc. Therefore, special care should be taken care to minimize the errors.
3. Validity of a docking protocol must be confirmed by a re-docking or control docking exercise. Validation experiment confirms how accurately docking engine mimics the conformational orientation of the substrate or inhibitor within the active site of the protein. For this purpose, the cognate (in bound) ligand from protein structure is removed and again docked back into the active site of the protein. For successful docking

model, RMSD of spatial coordinates of cognate and docked ligand should be ≤ 2 Å.

4. Traditional rigid docking-based approach considers the protein/macromolecule as a rigid body. This consideration no doubt saves some computational time but affects the predicting power of the model. The use of fast molecular dynamic (MD) simulation can be useful and can provide more realistic picture. Alternatively, structure-based ensemble docking can also be useful.
5. Flexibility of the ligand is also an important factor for computational docking. Compounds with too many rotatable bonds may lead to too many clusters and can produce inconclusive outcomes.

5 Future Prospects and Concluding Remarks

Devastating nature of many infectious diseases and continuous reports of drug resistance demand new strategic advancements in the field of NTDs and tropical infectious diseases. Multi-target screening is one such promising concept which offers better safety profile with low resistance possibility than the current therapeutical models. Multi-target screening and design also present an economical and more cost-effective drug discovery model than the traditional single target and combination therapies. Moreover, multi-target drugs can also replace multidrug regimen used at the time of coinfection by HIV and opportunistic pathogens.

Following developments will definitely help in R&D in the domain of MTDs:

1. Theoretical and computational progress in the field of structural biology and OMICS science.
2. Ever-expanding biochemical space available for drug discovery endeavors.
3. New efforts focused on the drug repurposing/drug repositioning of current medications of infectious disease.

It is highly likely that computational multi-target screening methods may lead to safer and superior drugs in the area of infectious disease in the foreseeable future.

Acknowledgments

The authors gratefully acknowledge Science and Engineering Research Board (SERB), Govt. of India (Grant No. SER-892-CMD), to financially assist this work.

References

- Bolognesi ML, Cavalli A (2016) Multitarget drug discovery and polypharmacology. *Chem-MedChem* 11:1190–1192
- Koutsoukas A, Simms B, Kirchmair J, Bond PJ, Whitmore AV, Zimmer S, Young MP, Jenkins JL, Glick M, Glen RC, Bender A (2011) From *in silico* target prediction to multi-target drug design: current databases, methods and applications. *J Proteome* 74:2554–2574
- Korcsmáros T, Szalay MS, Böde C, Kovács IA, Csérmelyi P (2007) How to design multi-target drugs: target search options in cellular networks. *Expert Opin Drug Discov* 2:1–10
- Ma XH, Shi Z, Tan C, Jiang Y, Go ML, Low BC, Chen YZ (2010) *In-silico* approaches to multi-target drug discovery computer aided multi-target drug design, multi-target virtual screening. *Pharm Res* 27:739–749
- Morphy R, Rankovic Z (2005) Designed multiple ligands. An emerging drug discovery paradigm. *J Med Chem* 48:6523–6543
- Jenwitheesuk E, Horst JA, Rivas KL, Voorhis WCV, Samudrala R (2008) Novel paradigms for drug discovery: computational multitarget screening. *Trends Pharmacol Sci* 29:62–71
- Senger MR, Fraga CA, Dantas RF, Silva FP Jr (2016) Filtering promiscuous compounds in early drug discovery: is it a good idea? *Drug Discov Today* 21:868–872
- Krug M, Hilgeroth A (2008) Recent advances in the development of multi-kinase inhibitors. *Mini Rev Med Chem* 8:1312–1327
- Gill AL, Verdonk M, Boyle RG, Taylor R (2007) A comparison of physicochemical property profiles of marketed oral drugs and orally bioavailable anti-cancer protein kinase inhibitors in clinical development. *Curr Top Med Chem* 7:1408–1422
- Koeberle A, Werz O (2014) Multi-target approach for natural products in inflammation. *Drug Discov Today* 19:1871–1882
- Amann R, Peskar BA (2002) Anti-inflammatory effects of aspirin and sodium salicylate. *Eur J Pharmacol* 447:1–9
- Rockett KA, Clarke GM et al (2014) Reappraisal of known malaria resistance loci in a large multicenter study. *Nat Genet* 46:1197–1204
- Sidhu AB, Verdier-Pinard D, Fidock DA (2002) Chloroquine resistance in *Plasmodium falciparum* malaria parasites conferred by *pfcr* mutations. *Science* 298:210–213
- Leang R, Taylor WR, Bouth DM, Song L, Tarning J, Char MC, Kim S, Witkowski B, Duru V, Domergue A, Khim N, Ringwald P, Menard D (2015) Evidence of *Plasmodium falciparum* malaria multidrug resistance to artemisinin and piperazine in western Cambodia: dihydroartemisinin-piperazine open-label multicenter clinical assessment. *Antimicrob Agents Chemother* 59:4719–4726
- Ashley EA, Dhorda M, Fairhurst RM et al (2014) Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *N Engl J Med* 371:411–423
- Wright A, Bai G, Barrera L et al (2006) Emergence of *Mycobacterium tuberculosis* with extensive resistance to second-line drugs – worldwide, 2000–2004. *MMWR Morb Mortal Wkly Rep* 55:301–305
- Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB (2009) Tuberculosis drug resistance mutation database. *PLoS Med* 6(2):e1000002
- Johnson R, Streicher EM, Louw GE, Warren RM, Helden PD, Victor TC (2006) Drug resistance in *Mycobacterium tuberculosis*. *Curr Issues Mol Biol* 8:97–112
- Oldfield E, Feng X (2014) Resistance-resistant antibiotics. *Trends Pharmacol Sci* 35:664–674
- Bisson GP, Frank I, Gross R, Lo Re V 3rd, Strom JB, Wang X, Mogorosi M, Gaolathe T, Ndwapi N, Friedman H, Strom BL, Dickinson D (2006) Out-of-pocket costs of HAART limit HIV treatment responses in Botswana's private sector. *AIDS* 20:1333–1336
- Moya J, Casado JL, Aranzabal L, Moreno A, Antela A, Dronda F, Perez-Elías MJ, Marin A, Moreno S (2006) Limitations of a simplification antiretroviral strategy for HIV-infected patients with decreasing adherence to a protease inhibitor regimen. *HIV Clin Trials* 7:210–214
- Gutiérrez-de-Terán H, Nervall M, Dunn BM, Clemente JC, Aqvist J (2006) Computational analysis of plasmepsin IV bound to anello phenyl norstatine inhibitor. *FEBS Lett* 580:5910–5916
- Kuhn M, Campillos M, González P, Jensen LJ, Bork P (2008) Large-scale prediction of drug–target relationships. *FEBS Lett* 582:1283–1290
- Kuhn M, Mering C, Campillos M, Jensen LJ, Bork P (2008) STITCH: interaction networks of chemicals and proteins. *Nucleic Acids Res* 36:D684–D688
- Drug.com database (Daraprim Side Effects) <https://www.drugs.com/sfx/daraprim-side-effects.html>. Accessed 11 Apr 2018
- Osher E, Fattal-Valevski A, Sagie L, Urshanski N, Amir-Levi Y, Katzburg S, Peleg L, Lerman-Sagie T, Zimran A, Elstein D,

- Navon R, Stern N, Valevski A (2011) Pyrimethamine increases β -hexosaminidase A activity in patients with late onset Tay Sachs. *Mol Genet Metab Rep* 102:356–363
27. Ashburn TT, Thor KB (2004) Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov* 3:673–683
 28. Fu RG, Sun Y, Sheng WB, Liao DF (2017) Designing multi-targeted agents: an emerging anticancer drug discovery paradigm. *Eur J Med Chem* 136:195–211
 29. Barchéchath S, Williams C, Saade K, Lauwagie S, Jean-Claude B (2009) Rational design of multitargeted tyrosine kinase inhibitors: a novel approach. *Chem Biol Drug Des* 73:380–387
 30. Curtin ML, Frey RR, Heyman HR, Soni NB, Marcotte PA, Pease LJ, Glaser KB, Magoc TJ, Tapang P, Albert DH, Osterling DJ, Olson AM, Bouska JJ, Guan Z, Preusser LC, Polakowski JS, Stewart KD, Tse C, Davidsen SK, Michaelides MR (2012) Thienopyridine ureas as dual inhibitors of the VEGF and Aurora kinase families. *Bioorg Med Chem Lett* 22:3208–3212
 31. Mohamed T, Yeung JC, Vasefi MS, Beazely MA, Rao PP (2012) Development and evaluation of multifunctional agents for potential treatment of Alzheimer's disease: application to a pyrimidine-2,4-diamine template. *Bioorg Med Chem Lett* 22:4707–4712
 32. Carvalho D, Paulino M, Polticelli F, Arredondo F, Williams RJ, Abin-Carriquiry JA (2017) Structural evidence of quercetin multi-target bioactivity: a reverse virtual screening strategy. *Eur J Pharm Sci* 106:393–403
 33. Thai NQ, Nguyen HL, Linh HQ, Li MS (2017) Protocol for fast screening of multi-target drug candidates: application to Alzheimer's disease. *J Mol Graph Model* 77:121–129
 34. Yousuf Z, Iman K, Iftikhar N, Mirza MU (2017) Structure-based virtual screening and molecular docking for the identification of potential multi-targeted inhibitors against breast cancer. *Breast Cancer* 9:447–459
 35. Chakraborty S, Basu S (2017) Multi-functional activities of citrus flavonoid narirutin in Alzheimer's disease therapeutics: an integrated screening approach and in vitro validation. *Int J Biol Macromol* 103:733–743
 36. Chakraborty S, Basu S (2017) Dual inhibition of BACE1 and A β aggregation by β -ecdysone: application of a phytoecdysteroid scaffold in Alzheimer's disease therapeutics. *Int J Biol Macromol* 95:281–287
 37. Chakraborty S, Bandyopadhyay J, Chakraborty S, Basu S (2016) Multi-target screening mines hesperidin as a multi-potent inhibitor: implication in Alzheimer's disease therapeutics. *Eur J Med Chem* 121:810–822
 38. Li Y, Tan C, Gao C, Zhang C, Luan X, Chen X, Liu H, Chen Y, Jiang Y (2011) Discovery of benzimidazole derivatives as novel multi-target EGFR, VEGFR-2 and PDGFR kinase inhibitors. *Bioorg Med Chem* 19:4529–4535
 39. Kumar M, Makhil B, Gupta VK, Sharma A (2014) *In silico* investigation of medicinal spectrum of imidazo-azines from the perspective of multitarget screening against malaria, tuberculosis and Chagas disease. *J Mol Graph Model* 50:1–9
 40. Sidorov P, Desta I, Chessé M, Horvath D, Marcou G, Varnek A, Davioud-Charvet E, Elhabiri M (2016) Redox polypharmacology as an emerging strategy to combat malarial parasites. *ChemMedChem* 11:1339–1351
 41. Aguilera E, Varela J, Birriel E, Serna E, Torres S, Yaluff G, Bilbao NV, Aguirre-López B, Cabrera N, Mazariegos SD, Gómez-Puyou MT, Gómez-Puyou A, Pérez-Montfort R, Minini L, Merlino A, Cerecetto H, González M, Alvarez G (2016) Potent and selective inhibitors of *Trypanosoma cruzi* triosephosphate isomerase with concomitant inhibition of cruzipain: inhibition of parasite growth through multitarget activity. *ChemMedChem* 11:1328–1338
 42. Álvarez G, Perdomo C, Coronel C, Aguilera E, Varela J, Aparicio G, Zolessi FR, Cabrera N, Vega C, Rolón M, Arias AR, Pérez-Montfort R, Cerecetto H, González M (2017) Multi-anti-parasitic activity of arylidene ketones and thiazolidene hydrazines against *Trypanosoma cruzi* and *Leishmania* spp. *Molecules* 22:709–725
 43. Yuvaniyama J, Chitnumsub P, Kamchonwongpaisan S, Vanichtanankul J, Sirawaraporn W, Taylor P, Walkinshaw MD, Yuthavong Y (2003) Insights into antifolate resistance from malarial DHFR-TS structures. *Nat Struct Mol Biol* 10:357–365
 44. Perozzo R, Kuo M, Sidhu AS, Valiyaveetil JT, Bittman R, Jacobs WR Jr, Fidock DA, Sacchetti JC (2002) Structural elucidation of the specificity of the antibacterial agent triclosan for malarial enoyl acyl carrier protein reductase. *J Biol Chem* 277:13106–13114
 45. Merckx A, Echalié A, Langford K, Sicard A, Langsley G, Joore J, Doerig C, Noble M, Endicott J (2008) Structures of *P. falciparum* protein kinase 7 identify an activation motif and leads for inhibitor design. *Structure* 16:228–238

46. Dias MV, Faim LM, Vasconcelos IB, de Oliveira JS, Basso LA, Santos DS, de Azevedo WF (2007) Effects of the magnesium and chloride ions and shikimate on the structure of shikimate kinase from *Mycobacterium tuberculosis*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 63:1–6
47. Wang S, Eisenberg D (2003) Crystal structures of a pantothenate synthetase from *M. tuberculosis* and its complexes with substrates and a reaction intermediate. *Protein Sci* 12:1097–1108
48. Li de la Sierra I, Munier-Lehmann H, Gilles AM, Bârzu O, Delarue M (2001) X-ray structure of TMP kinase from *Mycobacterium tuberculosis* complexed with TMP at 1.95 Å resolutions. *J Mol Biol* 311:87–100
49. Basavannacharya C, Robertson G, Munshi T, Keep NH, Bhakta S (2010) ATP-dependent mure ligase in *Mycobacterium tuberculosis*: biochemical and structural characterisation. *Tuberculosis (Edinb)* 90:16–24
50. Bond CS, Zhang Y, Berriman M, Cunningham ML, Fairlamb AH, Hunter WN (1999) Crystal structure of *Trypanosoma cruzi* trypanothione reductase in complex with trypanothione, and the structure-based discovery of new natural product inhibitors. *Structure* 7:81–89
51. Ladame S, Castilho MS, Silva CH, Denier C, Hannaert V, Périé J, Oliva G, Willson M (2003) Crystal structure of *Trypanosoma cruzi* glyceraldehyde-3-phosphate dehydrogenase complexed with an analogue of 1,3-bisphospho-D-glyceric acid. *Eur J Biochem* 270:4574–4586
52. Amaya MF, Watts AG, Damager I, Wehenkel A, Nguyen T, Buschiazzo A, Paris G, Frasch AC, Withers SG, Alzari PM (2004) Structural insights into the catalytic mechanism of *Trypanosoma cruzi* trans-sialidase. *Structure* 12:775–784
53. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
54. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 16:2785–2791
55. Shaaban S, Abdel-Wahab BF (2016) Groebke–Blackburn–Bienaymé multicomponent reaction: emerging chemistry for drug discovery. *Mol Divers* 20:233–254
56. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19:1639–1662
57. Kumar M, Dagar A, Gupta VK, Sharma A (2014) *In silico* docking studies of bioactive natural plant products as putative DHFR antagonists. *Med Chem Res* 23:810–817
58. Hopkins AL, Groom CR, Alex A (2004) Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* 9:430–431
59. Abad-zapatero C, Metz JT (2005) Ligand efficiency indices as guideposts for drug discovery. *Drug Discov Today* 10:464–469
60. Abad-zapatero C (2007) Ligand efficiency indices for effective drug discovery. *Expert Opin Drug Discov* 4:469–488
61. Lajiness MS, Vieth M, Erickson J (2004) Molecular properties that influence oral drug-like behavior. *Curr Opin Drug Discov Devel* 7:470–477
62. Wunberg T, Hendrix M, Hillisch A, Lobell M, Meier H, Schmeck C, Wild H, Hinzen B (2006) Improving the hit-to-lead process: data-driven assessment of drug-like and lead-like screening hits. *Drug Discov Today* 11:175–180



Computational Design of Multi-Target Drugs Against Breast Cancer

Shubhandra Tripathi, Gaurava Srivastava, and Ashok Sharma

Abstract

One drug one target approach has nowadays evolved in the form of one drug multiple targets strategy for the drug designing. In recent years, polypharmacology has gained much attention for the identification of multi-targeting drugs, i.e., a single drug molecule interacting to more than one target. For the treatment of HER-2-positive breast cancer, lapatinib is used as a dual inhibitor against the EGFR and HER-2 receptors. Several other multi-target inhibitors are also reported for the treatment of breast cancer, which are basically dual inhibitors of PI3K and mTOR, mTORC1 and mTORC2, and multikinase anti-angiogenesis inhibitors. Strategies for identifying multi-target drugs for the breast cancer treatment are still needed. Computational methods for identifying multi-target drugs are mainly of two types, i.e., structure-based and ligand-based. Structure-based drug designing methods are based on physical interaction studies and include molecular docking and molecular dynamics simulations for screening of multi-target drug molecules. On the other hand, ligand-based methods are knowledge-based methods which analyze similar properties of ligand molecules for identifying multi-targeting molecules.

Keywords Breast cancer, Computer-aided drug design, Molecular descriptors, Molecular dynamics simulation, Molecular interaction, Pharmacophore

1 Introduction

1.1 Breast Cancer and Therapeutic Targets

Breast cancer is one of the most prevalent types of cancer among women worldwide raising serious concerns. It is mainly of three types. The most common type of breast cancer with nearly 60% of cases is due to the overexpression of estrogen receptor (ER) and/or progesterone receptor (PR) known as hormone receptor-positive breast cancer [1]. The second type comprising 15–20% of all cases is due to overexpression of human epidermal growth factor receptor-2 (HER-2) [2], while in the rest of the breast cancer cases, none of these three receptors, i.e., ER, PR, and HER-2, are overexpressed, and this type is known as triple-negative breast cancer (TNBC) [1]. For the treatment of breast cancer, overexpressed receptors and various pathways assisting tumor development are targeted using drugs (Table 1).

For the treatment of hormone receptor-positive breast cancer (ER or PR positive), endocrine therapy is used, i.e., selective

Table 1
Various mono-target drugs and their respective target, used for the treatment of various types of breast cancers

Breast cancer type	Drug type	Drug molecule	Reference
Hormone receptor positive (ER+ or PR+)	SERMs	Tamoxifen	[3]
	SERMs	Raloxifen	[4]
	SERMs	Bazedoxifene	[5]
	SERDs	Fulvestrant	[6]
	SERDs	ARN-810	[7]
	Aromatase inhibitors	Letrozole	[8]
	Aromatase inhibitors	Anastrozole	[9]
	Aromatase inhibitors	Exemestane	[10]
HER-2 receptor positive	EGFR inhibitor	Gefitinib	[11]
	EGFR inhibitor	Erlotinib	[12]
	HER-2 inhibitor	Trastuzumab	[13]
	HER-2 inhibitor	Pertuzumab	[14]
TNBC (in combination therapy in all types of breast cancer)	PI3K inhibitor	PX-866	[15]
	PI3K inhibitor	GDC-0941	[16]
	PI3K inhibitor	XL-147	[17]
	PI3K inhibitor	BKM-120	[18]
	AKT inhibitor	GSK690693	[19]
	AKT inhibitor	A-443654	[20]
	AKT inhibitor	MK-2206	[21]
	mTOR inhibitor	Everolimus	[22]
	VEGFR inhibitor	Bevacizumab	[23, 24]
VEGFR inhibitor	Aflibercept	[24, 25]	

estrogen receptor modulators (SERMs) are administered. These SERM molecules, viz., tamoxifen [3], raloxifen [4], and bazedoxifene [5], are antagonists to ER [1]. Similarly, molecules that downregulate ER, i.e., selective estrogen receptor downregulators (SERDs), are also utilized for the treatment of hormone receptor-positive breast cancer. SERDs like fulvestrant and ARN-810 disrupt the dimerization and nucleation of ERs [6, 7, 26]. In addition, aromatase receptor responsible for the conversion of androgens into estrogens is also targeted. Aromatase inhibitors (AIs) (targeting aromatase receptor responsible for estrogen production), viz., letrozole, anastrozole, and exemestane, are used for treatment of hormone receptor-positive breast cancer (Table 1) [1, 8–10, 27]. In hormone receptor-positive breast cancer, hyperactivation of phosphatidylinositol 3-kinase (PI3K) pathway promotes antiestrogen resistance. In such cases PI3K signaling pathway kinases, viz. PI3K, protein kinase B (Akt), and mammalian target of rapamycin (mTOR) inhibitors, are targeted in a combination therapy [22, 27].

For the treatment of HER-2-positive breast cancer, drugs targeting HER-2 receptor are provided along with chemotherapeutic drugs [1, 2]. HER-2 receptor is a tyrosine kinase class of receptor.

It belongs to human epidermal growth factor receptor family, having four members, viz., epidermal growth factor receptor (EGFR/HER-1), HER-2, HER-3, and HER-4 [28]. All these receptors have a role in controlling cell growth, proliferation, and cell survival [28]. HER-2-positive breast cancer is characterized by overexpression of HER-2 receptor. HER-2 targeting drugs are mainly used for the treatment. Trastuzumab is the first monoclonal antibody targeting HER-2 receptor, for HER-2-positive breast cancer treatment [13, 29, 30]. Pertuzumab targeting HER-2 and inhibiting HER-2 dimerization is also used in combination therapy [14]. EGFR-targeting drugs like erlotinib [12, 31] and gefitinib [11] show activity in combination therapy against HER-2-positive breast cancer [29]. Moreover dual inhibitors targeting both EGFR and HER-2 are successful in HER-2-positive breast cancer treatment.

In the absence of ER/PR or HER-2-targeted therapy for TNBC type, other conventional chemotherapeutic agents like alkylating agents, drugs targeting cytoskeleton proteins (taxanes), focal adhesion kinase (FAK) inhibitors, and vascular endothelial growth factor receptor (VEGFR)-targeting drugs are mainly used in TNBC treatment. For TNBC treatment, target-specific drug is not utilized, and it is mainly treated using cytotoxic chemotherapy drugs [1]. Due to overexpression of EGFR in nearly half of the TNBC cases, EGFR inhibitors are also used for TNBC treatment [29, 32]. PARP inhibitors in combination therapy are also used for TNBC treatment [33]. In addition, drugs targeting PI3K/Akt/mTOR pathway are considered promising therapeutics agents for TNBC treatment [34]. FAK is an intracellular tyrosine kinase having a role in signal transduction and other cellular responses like cell adhesion, migration, survival, proliferation, and angiogenesis [32]. FAK is overexpressed in most breast cancer cases [35]. Therefore FAK inhibitors like VS-6063 and VS-4718 are important candidates for breast cancer treatment in particular for TNBC [36].

In addition, angiogenesis, i.e., formation of new blood vessels, is essential for tumor growth and metastasis. Various kinases are involved in angiogenesis, and VEGFR receptor tyrosine kinase plays a key role in tumor angiogenesis and breast cancer development [25]. Therefore VEGFR-targeting drugs like bevacizumab, aflibercept, etc. are used for breast cancer treatment [23, 29, 37]. In monotherapy, these drugs demonstrate limited activity; therefore they are generally used in combination with chemotherapies (with taxanes like paclitaxel or docetaxel) [24, 37]. Multikinase anti-angiogenesis inhibitors like sunitinib, sorafenib, vandetanib, vatalanib, and axitinib targeting VEGFRs (different VEGFR variants like VEGFR-1, VEGFR-2, and VEGFR-3) show significant activity for breast cancer treatment (Table 1) [25, 29, 37]. Apart from VEGFRs, these multikinase inhibitors also target other kinases like platelet-derived growth factor receptor (PDGFR),

stem cell factor receptor (KIT), feline McDonough sarcoma (FMS)-related tyrosine kinase 3 (FLT-3), colony-stimulating factor 1 receptor, rearranged during transfection (RET), C-rapidly accelerated fibrosarcoma (CRAF), and B-type RAF kinase (BRAF) [37]. Since they inhibit multiple kinases of the angiogenesis pathway, these inhibitors show high anti-angiogenesis activity.

PI3K/Akt/mTOR pathways regulate signaling of cell proliferation, adhesion, survival, and motility, and aberrant activation of this pathway in tumor cells leads to cancer growth, survival, motility, and development of resistance against targeted therapy [34, 38]. PI3K pathway kinases are the key target in treatment of all breast cancer types for use in combination therapy [34]. Analysis of cancer samples showed that nearly 38% of solid tumors have alteration in PI3K pathway receptors [39]. PIK3A, phosphatase and tensin homologue deleted on chromosome 10 (PTEN), and AKT1 mutations are frequent in hormone receptor-positive breast cancer (ER+/PR+) [22]. Likewise, PIK3CA mutations are common in HER-2-positive breast cancer [39]. This reveals the importance of targeting PI3K/Akt/mTOR pathway, in particular for TNBC and in combination therapy for hormone receptor-positive and HER-2-positive breast cancer.

In the PI3K/Akt/mTOR pathways, PI3K is the initial signaling kinase required for cell growth, cell proliferation, and important metabolic activities. PI3K consists of three classes of heterodimeric lipid kinases with catalytic and regulatory subunit catalyzing phosphorylation of 3'-hydroxyl group of phosphatidylinositols [34]. PI3K activation occurs through growth factor stimulation via receptor tyrosine kinases like EGFR. In contrast, PI3K is negatively regulated by PTEN, promoting dephosphorylation of phosphatidylinositol phosphates [40]. The second kinase is serine/threonine kinase AKT, having three isoforms AKT1, AKT2, and AKT3. The downstream effector of AKT is mTOR, the third important serine/threonine kinase of the pathway [34]. mTOR comprises two cell signaling complexes mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2) [41]. This pathway also suppresses expression of apoptosis receptors like caspases. Therefore, drugs targeting PI3K pathway activate caspases, thus leading to apoptosis of cancer cells.

PX-866 is a pan-PI3K inhibitor with significant inhibition of PI3K signaling [15]. Various Class I PI3K selective compounds, viz., GDC-0941 [16], XL-147 [17], and BKM120 [18], are also reported for inhibiting PI3K signaling in breast tumor cells (Table 1). AKT inhibitors developed for breast cancer treatment include GSK690693, A-443654, and MK-2206 [19–21]. MK-2206 is a highly selective allosteric pan-inhibitor of AKT [21]. These AKT inhibitors show significant antitumor activity in breast cancer. Everolimus is a mTOR inhibitor, showing allosteric binding to mTORC1 [22]. Everolimus combined with aromatase

inhibitors is used for hormone receptor-positive breast cancer patient treatment [22]. Dual inhibitors targeting both PI3K and mTOR receptors are reported for treatment of breast cancer.

In a better approach to combat breast cancer, multiple targets are inhibited using drug combination. In case of dual inhibition, two drug molecules targeting two different receptors for treatment of breast cancer are utilized for synergistic activity. For instance, the dual inhibition of FAK (by FAK-CD) and EGFR (by AG1478) protein tyrosine kinases was reported to have synergistic effect and trigger apoptosis in breast cancer cells [42]. Likewise, dual inhibition of mTOR (by everolimus) and aromatase receptor (by letrozole) in combination induces apoptosis in hormone-dependent breast cancer [43]. Dual inhibition of EGFR (by erlotinib) and VEGFR (by bevacizumab) in combination was reported for improved antitumor activity in patients with metastatic breast cancer [31].

Contrary to dual inhibition (using two different molecules targeting two receptors), in multi-targeted approach, a single molecule targets multiple inhibitors (e.g., one molecule targeting two receptors). This ability of drug molecules to bind multiple targets is the characteristic feature of multi-targeting drugs. Multi-targeted approach is very useful for the treatment of breast cancer. In this approach single drug molecule is able to inhibit multiple receptors, thus providing better opportunity for breast cancer therapy.

1.2 Multi-Target Approach in Breast Cancer Treatment

Multi-target therapeutics overcomes the limitations of the mono-target therapies. Multi-target drugs are more efficacious and less prone to resistance [44]. The process of oncogenesis is multigenic; therefore, multi-target drugs are more useful for cancer treatment [44]. Multi-target approach is utilized for breast cancer treatment. Lapatinib, one of the pioneer dual inhibitors of HER-2 and EGFR receptor, was reported for breast cancer treatment [45]. For breast cancer treatment, various multi-targeted drugs are used, viz., dual inhibitors of HER-2 and EGFR receptors, dual inhibitors of PI3K and mTOR kinases, dual inhibitors of mTORC1 and mTORC2 kinase, and multikinase inhibitors of angiogenesis (Fig. 1).

1.2.1 Dual Inhibition of EGFR and HER-2 Receptors

EGFR and HER-2 receptor targeting dual inhibitor lapatinib was the major breakthrough for the treatment of breast cancer (Table 2) [45]. Lapatinib is a tyrosine kinase inhibitor that binds at the intracellular domain of both the EGFR and HER-2 receptor and causes complete blockage of phosphorylation and restricts the downstream cascade [46, 47, 54]. Another dual inhibitor molecule neratinib was reported for HER-2-positive breast cancer [48]. BIBW-2992 was reported for targeting both EGFR and HER-2 receptors in HER-2-positive breast cancer [49]. All these dual inhibitors of EGFR and HER-2 receptor show activity against HER-2-positive metastatic breast cancer.

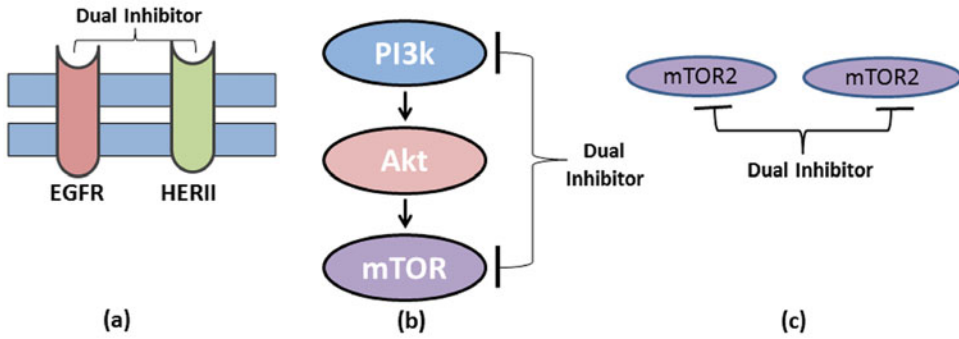


Fig. 1 This figure shows the way dual inhibitors concurrently inhibit multiple inhibitors. (a) EGFR and HER-2 dual inhibitor, (b) PI3K and mTOR dual inhibitor, and (c) mTOR1 and mTOR2 dual inhibitor

Table 2
Various multi-target inhibitors used for breast cancer treatment

Drug targets	Drug molecule	References
EGFR + HER-2 receptor	Lapatinib Neratinib BIBW 2992	[45–47] [48] [49]
PI3K + mTOR	NVP-BEZ235 GSK2126458 VS-5584	[18] [50] [51]
mTORC1 + mTORC2	AZD8055 OSI-027	[52] [53]
Multikinase anti-angiogenesis inhibitors		
VEGFR1-3, PDGFR KIT, FLT-3, MET	Sunitinib	[25, 37]
VEGFR1-3, PDGFR KIT, FLT-3, CRAF BRAF	Sorafenib	[25, 37]
VEGFR1-3, PDGFR-b, KIT	Axitinib	[37]
VEGFR1-3, PDGFR, KIT, RET	Motesanib	[37]
VEGFR1-3 PDGFR, KIT	Pazopanib	[25, 37]
VEGFR1-3, EGFR KIT, RET	Vandetanib	[37]
VEGFR1-3, PDGFR-B KIT	Cediranib	[37]
FGFR1-3, VEGFR1-3 PDGFR-b, C-KIT FLT-3	Dovitinib	[37]
MET, VEGFR2	Cabozantinib	[37]

1.2.2 Dual Inhibition of PI3K and mTOR

A dual inhibitor of PI3K and mTOR, NVP-BEZ235, was reported for breast cancer activity (Table 2) [18]. The use of NVP-BEZ235 for breast cancer in HER-2-positive and PIK3CA mutations was later reported [55]. Another dual inhibitor reported for PI3K and mTOR is GSK2126458 [50]. VS-5584 is another dual inhibitor of PI3k and mTOR [51]. PI3K/mTOR dual inhibitors are effective for hormone-resistant ER+ breast cancer and TNBC. These dual inhibitor molecules effectively and specifically block PI3K/Akt/mTOR signaling pathway.

1.2.3 Dual Inhibition of mTOR1 and mTOR2

The traditional targets of mTOR are rapamycin analogs known as rapalogs, everolimus, temsirolimus, and deferolimus that show allosteric inhibition of mTORC1 [34]. But the feedback activation continues via mTORC2. Therefore the ATP-competitive inhibitors binding to both mTORC1 and mTORC2 can completely inhibit the mTOR signaling. AZD-8055 and OSI-027 are the selective competitive dual inhibitors of mTOR (binding to mTORC1 and mTORC2); these inhibitors interact with both mTOR receptors and have promising results against breast cancer (Table 2) [52, 53].

1.2.4 Multikinase Inhibition of Angiogenesis

These multikinase anti-angiogenesis inhibitors target multiple proangiogenic pathways [25]. Sunitinib and sorafenib were initially explored for multiple kinase activity on angiogenesis signaling pathways [25]. Sunitinib prevents angiogenesis by targeting VEGFR1-3, PDGFR, KIT, FLT-3, colony-stimulating factor 1 receptor, and RET kinases [25, 37]. Similarly sorafenib interacts with VEGFR1-3, PDGFR, KIT, and FLT-3 cell surface kinase and CRAF and BRAF intracellular kinases [25, 37]. Other multikinase inhibitors include axitinib, motesanib, pazopanib, vandetanib, and cediranib targeting various angiogenesis signaling pathway receptors (Table 2) [37]. These multikinase inhibitors due to the property of concomitantly inhibiting various kinases prevent angiogenesis and, thus, tumor progression and metastasis in breast cancer.

2 Methodology

With the utility of multi-targeted inhibitors in breast cancer treatment, efforts are required to search for novel molecules for breast cancer treatment. Computational methods aid in the development and identification of multi-target molecules. The computational methods for development and identification of multi-target lead molecules for breast cancer treatment are discussed in detail.

2.1 Virtual Screening and Molecular Docking

Virtual screening is a computational method to discover new leads against biological drug target via screening ligand libraries [56–58]. It is a computational method in drug discovery for new lead molecule identification, compared to experimental method of

physical screening of large compound libraries against the disease target, i.e., high-throughput screening [59]. Virtual screening is an important tool in computer-aided drug designing based on molecular docking. It requires structure preparation of ligand libraries and protein target. For multi-target approach, two or more breast cancer target proteins (e.g., HER-2 and EGFR), blocking a pathway, should be identified. The experimental structures (X-ray, NMR, Cryo-EM) of breast cancer drug targets can be obtained from PDB (www.PDB.org) database. On the basis of active site of multiple targets, ligand libraries should be prepared. The protonation state of the target proteins can be determined using servers like PROPKA [60], H++ [61], etc.

The ligand libraries can be prepared using the public chemical compound databases like PubChem, Zinc, DrugBank, etc. Ligand libraries can be prepared using the knowledge from the pharmacophore of the receptor protein molecules. Ligand libraries with pharmacophore of both the receptor molecules (e.g., HER-2 and EGFR) can be utilized for virtual screening, and they would provide lead molecules for further evaluation [57]. The ligand libraries can be screening for absorption, desorption, metabolism, excretion, and toxicity (ADMET) properties. The significance of performing ADMET profiling is that it rejects those molecules which can be later rejected after performing whole study due to poor ADMET parameters [62]. The best set of identified lead molecules showing good binding affinity in both the targets (e.g., HER-2 and EGFR) are utilized for further assessment.

After virtual screening, the set of molecules showing good binding affinity for both the targets are further evaluated by performing molecular docking. In molecular docking as compared with virtual screening, priority is given for accuracy as compared to time cost. For molecular docking, freely available docking software AutoDock could be utilized which also provides the option for the flexible docking [63, 64]. AutoDock uses Lamarckian genetic algorithm to perform molecular docking [64]. Other docking software includes DOCK [65], GLIDE [66], Surflex [67], and AutoDock Vina [68]. The lead molecules showing good docking scores in both the receptor molecules (e.g., HER-2 and EGFR) are further evaluated using molecular dynamics simulation studies. A complete protocol for identification of new lead molecules using computational methods is called as relaxed complex scheme which could be used for identification of common lead molecules among multiple targets [69, 70].

2.2 Molecular Dynamics Simulation

After molecular docking, protein-ligand docked complexes are further evaluated using molecular dynamics simulation [71, 72]. As compared to molecular docking, in molecular dynamics whole system is in fully flexible explicit solvent [72]. In molecular dynamics simulation, the ligand-bound docked complex of the protein

molecule is simulated using classical force fields. The simulation time period is normally in the range of 10–50 ns with time step of 1–2 fs. Various packages to perform molecular dynamics simulation include GROMACS [73], AMBER [74], CHARMM [75], NAMD [76], etc. Various force fields used to perform molecular dynamics simulation are GROMOS [77], AMBER [78], OPLS [79], and CHARMM [80]. In brief, the steps involved in molecular dynamics include preparing the system with periodic boundary condition; solvating with water molecules, viz., SPCE [81] and TIP3P [81]; and neutralizing the system with counterions Na^+ or Cl^- . After system preparation, energy minimization is performed using steepest descent and/or conjugate gradient algorithms. Energy minimization is followed by equilibration of the system in NPT ensembles. Equilibration should be performed in a stepwise manner initially relaxing only the solvent molecules that the protein of the protein molecule followed by relaxing the amino acid side chain, and in last the entire system should be relaxed. Then the final production run should be performed in NVT ensemble to generate the molecular dynamics trajectory. To maintain the temperature of the system, thermostat like velocity-rescale thermostat [82] is utilized. Likewise, to maintain the pressure, barostat like Parrinello-Rahman barostat [83] is utilized.

From the obtained molecular dynamics trajectory, the binding of ligand molecules is analyzed monitoring various parameters like root-mean-square deviation (RMSD), root-mean-square fluctuations (RMSF), solvent-accessible surface area (SASA), and radius of gyration (R_g) [84, 85]. The evaluation of hydrogen bonds formed between the ligand molecule and protein molecule is also another parameter monitored for protein-ligand binding analysis. Similarly, secondary structure analysis of protein is also performed to monitor any changes induced by ligand binding in the protein. If the ligand binding creates huge fluctuations in the protein RMSD, or if the molecular dynamics trajectory ligand is moving far apart from the active site, then only those ligand molecules having unstable interactions with the protein molecules are considered. Other ligand molecules are discarded and are not further included in binding free energy estimation.

2.3 Binding Free Energy Estimation

To monitor the binding of ligand molecules with receptor, the affinity of a ligand with protein molecules can be evaluated by two approaches. In the first approach, from the MD trajectory frames, ligand binding affinity is evaluated using the implicit solvent method. For the implicit solvent binding free estimation, molecular mechanics Poisson-Boltzmann/generalized Born surface area (MMPB/GBSA) methods are utilized [86]. The benefit with MMPBSA and/or MMGBSA method is that they are not computationally expensive and use the MD simulation frame. However,

the main drawback with these methods is that they do not include entropic changes and are thus less reliable.

In the second approach from the equilibrated structure, binding affinity is evaluated using more rigorous explicit solvent methods. The rigorous explicit solvent methods include linear interaction energy (LIE) [87], thermodynamic integration (TI) [88], and free energy perturbation (FEP) [89]. LIE, TI, and FEP are computationally expensive methods but provide reliable estimates of binding free energy. Free energy workflow facilitates easy calculation of binding free energy with AMBER simulation package [90]. Using binding free energy methods, lead molecules showing a significant activity in the multiple targets can be evaluated further.

2.4 Advanced Molecular Dynamics Simulation Methods

For further evaluation of identified lead molecules, they can be sampled using advanced MD simulation method like accelerated molecular dynamics (AMD) [91], umbrella sampling (US) [92], and replica exchange molecular dynamics (REMD) [93, 94]. PLUMED is a very useful plugin which can be patched with any of molecular dynamics simulation package like GROMACS and AMBER to perform these advanced methods [95]. Although currently these methods are very rigorous and computationally expensive, but in the future with increasing computational power and with better protocols, time of residence will prove an important tool in lead identification.

After evaluating the lead molecules on the binding free energy criterion, molecules showing good binding free energies are now selected. These selected lead molecules can be evaluated using experimental methods. Molecules showing good binding energy in both the receptor molecules are further evaluated using in vitro and in vivo testing.

In summary, an example of dual inhibitor identification using computational methods for HER-2 and EGFR receptor is discussed here. Initially, we obtain the X-ray structure of the HER-2 and EGFR receptor from the PDB database. Then, using PROPKA server the protonation state of the amino acid residues is decided. The bound heteroatoms in the protein, i.e., water and ligand molecules, are removed. Now our protein structure for the HER-2 and EGFR receptors is ready. To prepare the ligand library, chemical compounds will be retrieved from database like ZINC. The downloaded compound library will be further screened using ADMET criterion. This filtered library of chemical compounds will be screened against both HER-2 and EGFR receptors via virtual screening. On the basis of binding affinity molecules showing appropriate binding with both the HER-2 and EGFR receptors will be separated. Now these lead molecules will be evaluated with rigorous docking at both the receptors. The lead molecules showing significant binding affinity to HER-2 and EGFR receptors will be selected. These lead dual inhibitors will be further tested for

their binding affinity to both HER-2 and EGFR receptors via molecular dynamics simulation. For the filtered molecules showing stability in molecular dynamics simulations, binding free energies will be estimated. The dual inhibitor lead molecules having good binding free energies with both HER-2 and EGFR receptors will be our lead dual inhibitors. These dual inhibitor lead molecules can be further evaluated for in vitro and in vivo testing.

3 Notes

The protocols discussed in the above book chapter are the computational practices for screening of multi-target drug molecules for breast cancer receptors. Further validation of the screened molecules is essential in the in vitro and in vivo conditions. Moreover, the screened molecules should be tested for the enzyme kinetics against the targeting receptors. The toxicity and other ADMET profiles of the screened molecule should also be evaluated.

Acknowledgments

S.T. is thankful to CSIR, New Delhi, for providing the CSIR-SRF fellowship, and G.S. is thankful to ICMR, New Delhi, for SRF fellowship.

References

1. Tinoco G, Warsch S, Glück S et al (2013) Treating breast cancer in the 21st century: emerging biological therapies. *J Cancer* 4:117–132. <https://doi.org/10.7150/jca.4925>
2. Hernández-Blanquisset A, Touya D, Strasser-Weippl K et al (2016) Current and emerging therapies of HER2-positive metastatic breast cancer. *Breast* 29:170–177. <https://doi.org/10.1016/j.breast.2016.07.026>
3. Jordan VC, Koerner S (1975) Tamoxifen (ICI 46,474) and the human carcinoma 8S oestrogen receptor. *Eur J Cancer* 11:205–206. [https://doi.org/10.1016/0014-2964\(75\)90119-X](https://doi.org/10.1016/0014-2964(75)90119-X)
4. Black LJ, Sato M, Rowley ER et al (1994) Raloxifene (LY139481 HCl) prevents bone loss and reduces serum cholesterol without causing uterine hypertrophy in ovariectomized rats. *J Clin Invest* 93:63–69. <https://doi.org/10.1172/JCI116985>
5. Komm BS, Chines AA (2012) Bazedoxifene: the evolving role of third-generation selective estrogen-receptor modulators in the management of postmenopausal osteoporosis. *Ther Adv Musculoskelet Dis* 4:21–34. <https://doi.org/10.1177/1759720X11422602>
6. Robertson JF (2002) Estrogen receptor down-regulators: new antihormonal therapy for advanced breast cancer. *Clin Ther* 24(Suppl A):A17–A30
7. Lai A, Kahraman M, Govek S et al (2015) Identification of GDC-0810 (ARN-810), an orally bioavailable selective estrogen receptor degrader (SERD) that demonstrates robust activity in tamoxifen-resistant breast cancer xenografts. *J Med Chem* 58:4888–4904. <https://doi.org/10.1021/acs.jmedchem.5b00054>
8. Mouridsen H, Gershanovich M, Sun Y et al (2001) Superior efficacy of letrozole versus tamoxifen as first-line therapy for postmenopausal women with advanced breast cancer: results of a phase III study of the International Letrozole Breast Cancer Group. *J Clin Oncol* 19:2596–2606. <https://doi.org/10.1200/JCO.2001.19.10.2596>

9. Baum M, Budzar AU, Cuzick J et al (2002) Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. *Lancet* 359:2131–2139
10. Paridaens RJ, Dirix LY, Beex LV et al (2008) Phase III study comparing exemestane with tamoxifen as first-line hormonal treatment of metastatic breast cancer in postmenopausal women: the European Organisation for Research and Treatment of Cancer Breast Cancer Cooperative Group. *J Clin Oncol* 26:4883–4890. <https://doi.org/10.1200/JCO.2007.14.4659>
11. Anido J, Matar P, Albanell J et al (2003) ZD1839, a specific epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor, induces the formation of inactive EGFR/HER2 and EGFR/HER3 heterodimers and prevents heregulin signaling in HER2-overexpressing breast cancer cells. *Clin Cancer Res* 9:1274–1283
12. Pollack VA, Savage DM, Baker DA et al (1999) Inhibition of epidermal growth factor receptor-associated tyrosine phosphorylation in human carcinomas with CP-358,774: dynamics of receptor inhibition in situ and antitumor effects in athymic mice. *J Pharmacol Exp Ther* 291:739–748
13. Leyland-Jones B (2002) Trastuzumab: hopes and realities. *Lancet Oncol* 3:137–144
14. Baselga J, Gelmon KA, Verma S et al (2010) Phase II trial of pertuzumab and trastuzumab in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer that progressed during prior trastuzumab therapy. *J Clin Oncol* 28:1138–1144. <https://doi.org/10.1200/JCO.2009.24.2024>
15. Ihle NT (2005) The phosphatidylinositol-3-kinase inhibitor PX-866 overcomes resistance to the epidermal growth factor receptor inhibitor gefitinib in A-549 human non-small cell lung cancer xenografts. *Mol Cancer Ther* 4:1349–1357. <https://doi.org/10.1158/1535-7163.MCT-05-0149>
16. O'Brien C, Wallin JJ, Sampath D et al (2010) Predictive biomarkers of sensitivity to the phosphatidylinositol 3' kinase inhibitor GDC-0941 in breast cancer preclinical models. *Clin Cancer Res* 16:3670–3683. <https://doi.org/10.1158/1078-0432.CCR-09-2828>
17. Chakrabarty A, Sánchez V, Kuba MG et al (2012) Feedback upregulation of HER3 (ErbB3) expression and activity attenuates antitumor effect of PI3K inhibitors. *Proc Natl Acad Sci U S A* 109:2718–2723. <https://doi.org/10.1073/pnas.1018001108>
18. Maira S-M, Stauffer F, Brueggen J et al (2008) Identification and characterization of NVP-BEZ235, a new orally available dual phosphatidylinositol 3-kinase/mammalian target of rapamycin inhibitor with potent in vivo antitumor activity. *Mol Cancer Ther* 7:1851–1863. <https://doi.org/10.1158/1535-7163.MCT-08-0017>
19. Carol H, Morton CL, Gorlick R et al (2010) Initial testing (stage 1) of the Akt inhibitor GSK690693 by the pediatric preclinical testing program. *Pediatr Blood Cancer* 55:1329–1337. <https://doi.org/10.1002/pbc.22710>
20. Chan TO, Zhang J, Rodeck U et al (2011) Resistance of Akt kinases to dephosphorylation through ATP-dependent conformational plasticity. *Proc Natl Acad Sci U S A* 108: E1120–E1127. <https://doi.org/10.1073/pnas.1109879108>
21. Yap TA, Yan L, Patnaik A et al (2011) First-in-man clinical trial of the oral pan-AKT inhibitor MK-2206 in patients with advanced solid tumors. *J Clin Oncol* 29:4688–4695. <https://doi.org/10.1200/JCO.2011.35.5263>
22. Baselga J, Campone M, Piccart M et al (2012) Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer. *N Engl J Med* 366:520–529. <https://doi.org/10.1056/NEJMoa1109653>
23. Miller K, Wang M, Gralow J et al (2007) Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer. *N Engl J Med* 357:2666–2676. <https://doi.org/10.1056/NEJMoa072113>
24. Zelnak AB, O'Regan RM (2007) Targeting angiogenesis in advanced breast cancer. *BioDrugs* 21:209–214. <https://doi.org/10.2165/00063030-200721040-00001>
25. Gotink KJ, Verheul HMW (2010) Anti-angiogenic tyrosine kinase inhibitors: what is their mechanism of action? *Angiogenesis* 13:1–14. <https://doi.org/10.1007/s10456-009-9160-6>
26. McDonnell DP, Wardell SE, Norris JD (2015) Oral selective estrogen receptor downregulators (SERDs), a breakthrough endocrine therapy for breast cancer. *J Med Chem* 58:4883–4887. <https://doi.org/10.1021/acs.jmedchem.5b00760>
27. Fox EM, Arteaga CL, Miller TW (2012) Abrogating endocrine resistance by targeting ER α and PI3K in breast cancer. *Front Oncol* 2. <https://doi.org/10.3389/fonc.2012.00145>
28. O'Sullivan CC, Smith KL (2014) Therapeutic considerations when treating HER2-positive

- metastatic breast cancer. *Curr Breast Cancer Rep* 6:169–182. <https://doi.org/10.1007/s12609-014-0155-y>
29. Alvarez RH, Valero V, Hortobagyi GN (2010) Emerging targeted therapies for breast cancer. *J Clin Oncol* 28:3366–3379. <https://doi.org/10.1200/JCO.2009.25.4011>
30. Cobleigh MA, Vogel CL, Tripathy D et al (1999) Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J Clin Oncol* 17:2639–2648. <https://doi.org/10.1200/JCO.1999.17.9.2639>
31. Dickler M, Rugo H, Caravelli J et al (2004) Phase II trial of erlotinib (OSI-774), an epidermal growth factor receptor (EGFR)-tyrosine kinase inhibitor, and bevacizumab, a recombinant humanized monoclonal antibody to vascular endothelial growth factor (VEGF), in patients (pts) with metastatic breast cancer (MBC). *J Clin Oncol* 22:2001–2001. <https://doi.org/10.1200/jco.2004.22.90140.2001>
32. Nielsen TO, Hsu FD, Jensen K et al (2004) Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 10:5367–5374. <https://doi.org/10.1158/1078-0432.CCR-04-0220>
33. Dréan A, Lord CJ, Ashworth A (2016) PARP inhibitor combination therapy. *Crit Rev Oncol Hematol* 108:73–85. <https://doi.org/10.1016/j.critrevonc.2016.10.010>
34. Cidado J, Park BH (2012) Targeting the PI3K/Akt/mTOR pathway for breast cancer therapy. *J Mammary Gland Biol Neoplasia* 17:205–216. <https://doi.org/10.1007/s10911-012-9264-2>
35. Luo M, Guan J-L (2010) Focal adhesion kinase: a prominent determinant in breast cancer initiation, progression and metastasis. *Cancer Lett* 289:127–139. <https://doi.org/10.1016/j.canlet.2009.07.005>
36. Kolev VN, Tam WF, Wright QG et al (2017) Inhibition of FAK kinase activity preferentially targets cancer cells. *Oncotarget* 8. <https://doi.org/10.18632/oncotarget.18517>
37. Reddy S, Raffin M, Kaklamani V (2012) Targeting angiogenesis in metastatic breast cancer. *Oncologist* 17:1014–1026. <https://doi.org/10.1634/theoncologist.2012-0043>
38. Samuels Y (2004) High frequency of mutations of the PIK3CA gene in human cancers. *Science* 304:554–554. <https://doi.org/10.1126/science.1096502>
39. Millis SZ, Ikeda S, Reddy S et al (2016) Landscape of phosphatidylinositol-3-kinase pathway alterations across 19 784 diverse solid Tumors. *JAMA Oncol* 2:1565. <https://doi.org/10.1001/jamaoncol.2016.0891>
40. Maehama T, Dixon JE (1998) The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *J Biol Chem* 273:13375–13378
41. Leung EY, Askarian-Amiri M, Finlay GJ et al (2015) Potentiation of growth inhibitory responses of the mTOR inhibitor everolimus by dual mTORC1/2 inhibitors in cultured breast cancer cell lines. *PLoS One* 10: e0131400. <https://doi.org/10.1371/journal.pone.0131400>
42. Golubovskaya V, Beviglia L, Xu L-H et al (2002) Dual inhibition of focal adhesion kinase and epidermal growth factor receptor pathways cooperatively induces death receptor-mediated apoptosis in human breast cancer cells. *J Biol Chem* 277:38978–38987. <https://doi.org/10.1074/jbc.M205002200>
43. Boulay A, Rudloff J, Ye J et al (2005) Dual inhibition of mTOR and estrogen receptor signaling in vitro induces cell death in models of breast cancer. *Clin Cancer Res* 11:5319–5328. <https://doi.org/10.1158/1078-0432.CCR-04-2402>
44. Zimmermann GR, Lehar J, Keith CT (2007) Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov Today* 12:34–42. <https://doi.org/10.1016/j.drudis.2006.11.008>
45. Burris HA (2004) Dual kinase inhibition in the treatment of breast cancer: initial experience with the EGFR/ErbB-2 inhibitor lapatinib. *Oncologist* 9(Suppl 3):10–15
46. Ahn ER, Vogel CL (2012) Dual HER2-targeted approaches in HER2-positive breast cancer. *Breast Cancer Res Treat* 131:371–383. <https://doi.org/10.1007/s10549-011-1781-y>
47. Konecny GE, Pegram MD, Venkatesan N et al (2006) Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells. *Cancer Res* 66:1630–1639. <https://doi.org/10.1158/0008-5472.CAN-05-1182>
48. Tsou H-R, Overbeek-Klumpers EG, Hallett WA et al (2005) Optimization of 6,7-disubstituted-4-(arylamino)quinoline-3-carbonitriles as orally active, irreversible inhibitors of human

- epidermal growth factor receptor-2 kinase activity. *J Med Chem* 48:1107–1131. <https://doi.org/10.1021/jm040159c>
49. Eskens FALM, Mom CH, Planting AST et al (2008) A phase I dose escalation study of BIBW 2992, an irreversible dual inhibitor of epidermal growth factor receptor 1 (EGFR) and 2 (HER2) tyrosine kinase in a 2-week on, 2-week off schedule in patients with advanced solid tumours. *Br J Cancer* 98:80–85. <https://doi.org/10.1038/sj.bjc.6604108>
 50. Knight SD, Adams ND, Burgess JL et al (2010) Discovery of GSK2126458, a highly potent inhibitor of PI3K and the mammalian target of rapamycin. *ACS Med Chem Lett* 1:39–43. <https://doi.org/10.1021/ml900028r>
 51. Kolev VN, Wright QG, Vidal CM et al (2015) PI3K/mTOR dual inhibitor VS-5584 preferentially targets cancer stem cells. *Cancer Res* 75:446–455. <https://doi.org/10.1158/0008-5472.CAN-14-1223>
 52. Chresta CM, Davies BR, Hickson I et al (2010) AZD8055 is a potent, selective, and orally bioavailable ATP-competitive mammalian target of rapamycin kinase inhibitor with in vitro and in vivo antitumor activity. *Cancer Res* 70:288–298. <https://doi.org/10.1158/0008-5472.CAN-09-1751>
 53. Bhagwat SV, Gokhale PC, Crew AP et al (2011) Preclinical characterization of OSI-027, a potent and selective inhibitor of mTORC1 and mTORC2: distinct from rapamycin. *Mol Cancer Ther* 10:1394–1406. <https://doi.org/10.1158/1535-7163.MCT-10-1099>
 54. Paul B, Trovato JA, Thompson J (2008) Lapatinib: a dual tyrosine kinase inhibitor for metastatic breast cancer. *Am J Health Syst Pharm* 65:1703–1710. <https://doi.org/10.2146/ajhp070646>
 55. Brachmann SM, Hofmann I, Schnell C et al (2009) Specific apoptosis induction by the dual PI3K/mTor inhibitor NVP-BEZ235 in HER2 amplified and PIK3CA mutant breast cancer cells. *Proc Natl Acad Sci U S A* 106:22299–22304. <https://doi.org/10.1073/pnas.0905152106>
 56. Shoichet BK (2004) Virtual screening of chemical libraries. *Nature* 432:862–865. <https://doi.org/10.1038/nature03197>
 57. Kumar V, Krishna S, Siddiqi MI (2015) Virtual screening strategies: recent advances in the identification and design of anti-cancer agents. *Methods* 71:64–70. <https://doi.org/10.1016/j.ymeth.2014.08.010>
 58. Okimoto N, Futatsugi N, Fuji H et al (2009) High-performance drug discovery: computational screening by combining docking and molecular dynamics simulations. *PLoS Comput Biol* 5:e1000528. <https://doi.org/10.1371/journal.pcbi.1000528>
 59. Liu B, Li S, Hu J (2004) Technological advances in high-throughput screening. *Am J Pharmacogenomics* 4:263–276
 60. Bas DC, Rogers DM, Jensen JH (2008) Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins* 73:765–783. <https://doi.org/10.1002/prot.22102>
 61. Anandakrishnan R, Aguilar B, Onufriev AV (2012) H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res* 40:W537–W541. <https://doi.org/10.1093/nar/gks375>
 62. Wang J, Skolnik S (2009) Recent advances in physicochemical and ADMET profiling in drug discovery. *Chem Biodivers* 6:1887–1899. <https://doi.org/10.1002/cbdv.200900117>
 63. Goodsell DS, Morris GM, Olson AJ (1996) Automated docking of flexible ligands: applications of AutoDock. *J Mol Recognit* 9:1–5. [https://doi.org/10.1002/\(SICI\)1099-1352\(199601\)9:1<1::AID-JMR241>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1099-1352(199601)9:1<1::AID-JMR241>3.0.CO;2-6)
 64. Morris GM, Goodsell DS, Halliday RS et al (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19:1639–1662. [https://doi.org/10.1002/\(SICI\)1096-987X\(199811\)19:14<1639::AID-JCC10>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1096-987X(199811)19:14<1639::AID-JCC10>3.0.CO;2-B)
 65. Allen WJ, Balias TE, Mukherjee S et al (2015) DOCK 6: impact of new features and current docking performance. *J Comput Chem* 36:1132–1156. <https://doi.org/10.1002/jcc.23905>
 66. Friesner RA, Murphy RB, Repasky MP et al (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. *J Med Chem* 49:6177–6196. <https://doi.org/10.1021/jm051256o>
 67. Jain AN (2003) Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem* 46:499–511. <https://doi.org/10.1021/jm020406h>
 68. Trott O, Olson AJ (2009) AutoDockVina: improving the speed and accuracy of docking with a new scoring function, efficient

- optimization, and multithreading. *J Comput Chem*. <https://doi.org/10.1002/jcc.21334>
69. Lin J-H, Perryman AL, Schames JR, McCammon JA (2002) Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *J Am Chem Soc* 124:5632–5633
70. Amaro RE, Baron R, McCammon JA (2008) An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J Comput Aided Mol Des* 22:693–705. <https://doi.org/10.1007/s10822-007-9159-2>
71. Durrant JD, McCammon JA (2011) Molecular dynamics simulations and drug discovery. *BMC Biol* 9:71. <https://doi.org/10.1186/1741-7007-9-71>
72. Zhao H, Caffisch A (2015) Molecular dynamics in drug design. *Eur J Med Chem* 91:4–14. <https://doi.org/10.1016/j.ejmech.2014.08.004>
73. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4:435–447. <https://doi.org/10.1021/ct700301q>
74. Case DA, Cheatham TE, Darden T et al (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688. <https://doi.org/10.1002/jcc.20290>
75. Brooks BR, Bruccoleri RE, Olafson BD et al (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217. <https://doi.org/10.1002/jcc.540040211>
76. Phillips JC, Braun R, Wang W et al (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802. <https://doi.org/10.1002/jcc.20289>
77. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) Abiomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 25:1656–1676. <https://doi.org/10.1002/jcc.20090>
78. Lindorff-Larsen K, Piana S, Palmo K et al (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78:1950–1958. <https://doi.org/10.1002/prot.22711>
79. Jorgensen WL, Tirado-Rives J (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* 110:1657–1666. <https://doi.org/10.1021/ja00214a001>
80. Vanommeslaeghe K, Hatcher E, Acharya C et al (2009) CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem*. <https://doi.org/10.1002/jcc.21367>
81. Mark P, Nilsson L (2001) Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J Phys Chem A* 105:9954–9960. <https://doi.org/10.1021/jp003020w>
82. Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. *J Chem Phys* 126:014101. <https://doi.org/10.1063/1.2408420>
83. Parrinello M (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys* 52:7182. <https://doi.org/10.1063/1.328693>
84. Srivastava G, Tripathi S, Kumar A, Sharma A (2017) Molecular investigation of active binding site of isoniazid (INH) and insight into resistance mechanism of S315T-MtKatG in Mycobacterium tuberculosis. *Tuberculosis* 105:18–27. <https://doi.org/10.1016/j.tube.2017.04.002>
85. Tripathi S, Srivastava G, Sharma A (2016) Molecular dynamics simulation and free energy landscape methods in probing L215H, L217R and L225M β -tubulin mutations causing paclitaxel resistance in cancer cells. *Biochem Biophys Res Commun* 476:273–279. <https://doi.org/10.1016/j.bbrc.2016.05.112>
86. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* 10:449–461. <https://doi.org/10.1517/17460441.2015.1032936>
87. Singh N, Warshel A (2010) Absolute binding free energy calculations: on the accuracy of computational scoring of protein-ligand interactions. *Proteins*. <https://doi.org/10.1002/prot.22687>
88. Jorge M, Garrido NM, Queimada AJ et al (2010) Effect of the integration method on the accuracy and computational efficiency of free energy calculations using thermodynamic integration. *J Chem Theory Comput* 6:1018–1027. <https://doi.org/10.1021/ct900661c>
89. Liu P, Dehez F, Cai W, Chipot C (2012) A toolkit for the analysis of free-energy perturbation calculations. *J Chem Theory Comput* 8:2606–2616. <https://doi.org/10.1021/ct300242f>
90. Homeyer N, Gohlke H (2013) FEW: a workflow tool for free energy calculations of ligand

- binding. *J Comput Chem* 34:965–973. <https://doi.org/10.1002/jcc.23218>
91. Hamelberg D, Mongan J, McCammon JA (2004) Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J Chem Phys* 120:11919–11929. <https://doi.org/10.1063/1.1755656>
92. Wojtas-Niziurski W, Meng Y, Roux B, Bernèche S (2013) Self-learning adaptive umbrella sampling method for the determination of free energy landscapes in multiple dimensions. *J Chem Theory Comput* 9:1885–1895. <https://doi.org/10.1021/ct300978b>
93. Zhang W, Wu C, Duan Y (2005) Convergence of replica exchange molecular dynamics. *J Chem Phys* 123:154105. <https://doi.org/10.1063/1.2056540>
94. Sinko W, Lindert S, McCammon JA (2013) Accounting for receptor flexibility and enhanced sampling methods in computer-aided drug design: accounting for receptor flexibility. *Chem Biol Drug Des* 81:41–49. <https://doi.org/10.1111/cbdd.12051>
95. Bonomi M, Branduardi D, Bussi G et al (2009) PLUMED: a portable plugin for free-energy calculations with molecular dynamics. *Comput Phys Commun* 180:1961–1972. <https://doi.org/10.1016/j.cpc.2009.05.011>



Computational Methods for Multi-Target Drug Designing Against *Mycobacterium tuberculosis*

Gaurava Srivastava, Ashish Tiwari, and Ashok Sharma

Abstract

Despite the availability of several drugs, *Mycobacterium tuberculosis* is still a big concern for public health. Such situation exists because of continuous emergence of TB-resistant strains. Possible reasons of developing resistance include long therapy and combination therapy. Therefore new potential leads are needed to be identified, and at the same time, the number of drugs in the combination therapy should also be reduced that will make administration of drug doses easier. In the present scenario, developing drug having the ability to interact with multiple targets, simultaneously, is a promising approach to treat the complicated diseases. These multi-target drug therapies have advantage of improved safety profile and high drug efficacy with easier administration over the single-target drug therapies. Many of in silico methods have been applied to reach different polypharmacologically directed drug designing, mainly for multi-target drug designing. In this chapter, we have discussed about the available strategies for computational multi-target drug designing with their advantages and disadvantages. We have also discussed an easy, fast, and equally accurate method for multi-target drug designing against the *Mycobacterium tuberculosis*.

Keywords De novo methods, Docking, FEL, MDR-TB, MM-PBSA, Molecular dynamics simulation, Multi-target drug designing, *Mycobacterium tuberculosis*, PCA, Pharmacophore

Abbreviations

BCG	Bacillus Calmette-Guerin
CS	Cycloserine
DDI	Drug-drug interactions
DOTS	Directly observed short-course chemotherapy
EMB	Ethambutol
ETA	Ethionamide
FEL	Free energy landscape
IFN	Interferon
IGRA	Interferon-gamma release assay
INH	Isoniazid
MDR	Multidrug resistant
Mtb	<i>Mycobacterium tuberculosis</i>
ODE	Ordinary differential equation
PAS	Para-amino salicylate
PCA	Principal component analysis
PZA	Pyrazinamide

Rg	Radius of gyration
RIF	Rifampin
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
SASA	Solvent-accessible surface area
SMD	Steered molecular dynamics
TB	Tuberculosis
TDR	Totally drug resistant
TST	Tuberculin skin test
XDR	Extensively drug resistant

1 Introduction

Tuberculosis (TB), which is one of the most infectious diseases, still persists as a major global health problem [1]. TB is an airborne disease caused by *Mycobacterium tuberculosis* (Mtb). Mtb, a species belonging to the family *Mycobacteriaceae*, was first discovered by Robert Koch in the year 1882 [2]. Mtb has a special composition of its cell wall which makes it unique to the other bacteria [3, 4]. Cell wall of Mtb is made up of high lipid constituents (about >60%), which include mycolic acid, cord factor, and wax-D. Out of them, mycolic acid is a hydrophobic alpha-branched lipid, whereas cord factor influences the arrangement of Mtb cells into long slender formation [5–7]. Such unique composition of Mtb cell wall makes it impervious against the gram (–) or gram (+) staining, that is why acid-fast techniques are used to identify the Mtb under the microscopes [8].

Mtb, an infectious disease, transmits through droplet infection by sneezing or cough. After reaching into the host, Mtb resides into the alveoli. Initially, Mtb appears to gain entry into the macrophages of lung cells through phagocytosis. But inside the macrophages, they do not get affected because of having unique cell wall. Inside the macrophages they multiply, survive, and sustain for a long time period. The success of *Mycobacterium* in producing disease relies entirely on its ability to utilize macrophages for its replication and, more importantly, the maintenance of viability of host macrophages that sustain *Mycobacterium*. *Mycobacterium* has several mechanisms to maintain the hostile environment of macrophages, its primary host cell [9–12]. The mechanism includes:

- Inhibition of fusion of phagosome with lysosomes
- Inhibition of acidification of phagosome by proton ATPase pump
- Protection from reactive oxidative radicals

- Recruitment and retention of tryptophan/aspartate-containing coat protein on phagosomes to prevent their delivery to the lysosomes
- Expression of virulence proteins of PE-PGRS family

TB may be classified into two categories: first is a latent infection, and another is an active disease. Latent TB refers to the condition when the patients have TB infection but the bacteria remains in inactive form and the patient does not exhibit the disease conditions. It has no symptoms and their chest X-ray may also be normal. The only manifestation of this encounter may be diagnosed through the tuberculin skin test (TST) or interferon-gamma release assay (IGRA). Latent TB is not contagious, but there is an ongoing risk that the latent infection may escalate the chances of getting active TB disease [13, 14]. Active TB is an illness in which the TB bacteria seem to rapidly multiply and invade toward the different organs of the body. The typical symptoms of active TB variably include cough, phlegm, chest pain, weakness, weight loss, fever, chills, and sweating at night [14, 15]. The most common form of active TB is pulmonary TB (lung disease), but depending upon the site of infection, other types of TB are loosely classified as “extrapulmonary TB.” Symptoms of extrapulmonary TB are listed in Table 1. Since most of the TB cases follow a general pattern of infection course, the whole pathophysiology has been classified in five stages, named as stage 1, stage 2, stage 3, stage 4, and stage 5 [16–19]. All stages are discussed below.

Stage 1: Droplet nuclei are inhaled through talking, coughing, and sneezing. Once nuclei are inhaled, the bacteria are taken up by

Table 1
Type of extrapulmonary TB with their symptoms

Type of extrapulmonary TB	Symptoms
Skeletal TB/Pott's disease	Spinal pain, back stiffness, paralysis is also possible
TB meningitis	Headaches, mood swings, coma
TB arthritis	Pain in single joint (most commonly in hips and knees)
Genitourinary TB	Dysuria, flank pain, masses or lumps (granulomas)
Gastrointestinal TB	Difficulty in swallowing, nonhealing ulcers, abdominal pains, malabsorption, diarrhea (may be bloody)
Miliary TB	Many small nodules widespread in organs that resemble millet seeds
Pleural TB	Empyema and pleural effusions
Caseous TB	Appearance of necrotic tissues with a soft, dry, and cheesy appearance

alveolar macrophages. The macrophages will not be activated, therefore unable to destroy the intracellular organism. The large droplet nuclei reach the upper respiratory tract, and the small droplet nuclei reach to the air sacs of the lung (alveoli) where infection begins. The disease starts when droplet nuclei reach to the alveoli.

Stage 2: It begins after 7–21 days post initial infection. TB bacilli multiply within the inactivated macrophages until macrophages burst. Other macrophages diffuse from the peripheral blood and phagocytose the TB bacterium, but because of the inactivated form, macrophages are not able to destroy TB.

Stage 3: Lymphocytes, specifically T cells, recognize TB antigen. This results in T-cell activation and the release of cytokines, including interferon (IFN). The release of IFN causes the activation of macrophages, which can release lytic enzymes and reactive intermediates that facilitate immune pathology. The Tubercle thus formed, which have a semi-solid or cheesy mass like structure. TB cannot multiply within tubercles due to low pH and anoxic environment, but TB can persist within these tubercles for extended periods.

Stage 4: Although many activated macrophages surround the tubercles, many other macrophages are inactivated or poorly activated. TB uses these macrophages to replicate causing the tubercle to grow. The growing tubercle may invade a bronchus, causing an infection which may spread to other parts of the lungs. Tubercle may also invade the artery or other blood supply. Spreading of TB may cause miliary tuberculosis, which can cause secondary lesions. Secondary lesions occur in the bones, joints, lymph nodes, genitourinary system, and peritoneum.

Stage 5: The caseous centers of the tubercles liquefy. This liquid is very crucial for the growth of TB, and therefore it multiplies rapidly (extracellularly). This later becomes a large antigen load, causing the walls of nearby bronchi to become necrotic and rupture. This results in cavity formation and allows TB to spread rapidly into other airways and to other parts of the lung.

In the year 2015, TB was found in the list of top ten infectious diseases, caused deaths throughout the world, and ranked above HIV/AIDS [1]. According to WHO report (2016), 1.4 million TB deaths were reported in the year 2015, and in addition 0.4 million also died because of TB along with HIV-2. Another estimation showed a total of 10.4 million new TB cases in the year 2015. Among them 1.2 million were HIV-positive. Out of 10.4 million, 5.9 million were men, 3.5 million were women, and 1.0 million were children [1].

Several drugs have been developed to combat the TB infection. These drugs can be classified into two categories: first-line drugs [isoniazid (INH), rifampin (RIF), pyrazinamide (PZA), and ethambutol (EMB)] and second-line drugs (para-amino salicylate

(PAS), kanamycin, cycloserine (CS), ethionamide (ETA), amikacin, capreomycin, thiacetazone, and fluoroquinolones) [20]. DOTS (directly observed short-course chemotherapy), current TB therapy, is a predefined therapy program which includes combinations of some of these drugs to treat the disease. It consists of an initial phase of treatment with four drugs, INH, RIF, PZA, and EMB, for 2 months daily, followed by treatment with INH and RIF for another 4 months, three times a week. But, this treatment has not been improved for over 30 years [21]. DOTS are singularly inefficient by the standards of today's pharmaceutical industry in terms of drug activity and toxicity, and its efficacy is threatened by increasingly widespread drug resistance [22, 23]. A brief of the available TB drugs is tabulated in Table 2.

Table 2
Illustration of anti-TB drugs and their target with encoding genes

Antibiotics	Targets proteins	Encoding genes
<i>First line</i>		
Isoniazid	Catalase/peroxidase Enoyl reductase Alky hydroperoxide reductase NADH dehydrogenase	katG inhA ahpC ndh
Pyrazinamide	Fatty acid synthetase I Pyrazinamidase/nicotinamidase	rpsA pncA
Rifampin	RNA polymerase B chain	rpoB
Streptomycin	S12 ribosomal protein 16S rRNA 7-Methylguanosine methyltransferase	rpsL rrs gidB
Ethambutol	Arabinosyl transferase	embB
<i>Second line</i>		
Levofloxacin/moxifloxacin	DNA gyrase	gyrA/gyrB
Capreomycin	rRNA methyltransferase	tlyA
Amikacin/kanamycin	16S rRNA Enhanced intracellular survival protein	rrs eis
Cycloserine	D-Alanine racemase	alrA
p-Aminosalicylic acid	Thymidylate synthase A	thyA
Ethionamide	Enoyl reductase Monooxygenase TetR family transcriptional repressor NADH dehydrogenase	inhA ethA ethR Ndh
Fluoroquinolones	DNA gyrase	gyrAB

Apart from these, active immunization techniques have also been applied to control TB. Although vaccination is not effective nowadays, bacillus Calmette-Guerin (BCG) is the most commonly used vaccine against tuberculosis. This is an attenuated strain of *Mycobacterium bovis* (MB), and mostly, it was found effective in children to protect them from the TB disease, while in adults it was not found effective, especially against pulmonary TB. This again shows a predominant necessity of developing more effective vaccines against the Mtb [24, 25].

Despite presence of a range of drugs against the mycobacterium, continuous emergence of various drug-resistant TB strains has made these treatments less effective and necessitated attention on this emerging problem which if avoided may become a serious health issue. Drug-resistant tuberculosis has been classified into three groups [26–28].

1.1 Multidrug-Resistant TB (MDR-TB)

MDR-TB strains are defined as those strains which have resistance against two major first-line anti-TB drugs (isoniazid and rifampicin) with or without having resistance to other first-line drugs. MDR-TB is a serious health condition because patients with MDR-TB respond very poorly to the standard anti-TB treatment with first-line drugs. In addition, MDR-TB needs relatively costly laboratory diagnosis as well as a long-term treatment of at least 2 years with drugs that are expensive, toxic, and not specifically potent. A case of MDR-TB is about 20–40 times more expensive to manage than a case of drug-sensitive TB.

1.2 Extensively Drug-Resistant TB (XDR-TB)

XDR-TB strains are MDR-TB strains having resistance against second-line drugs (i.e., fluoroquinolones) along with at least one of the injectable aminoglycosides (like capreomycin).

1.3 Totally Drug-Resistant TB (TDR-TB)

TDR-TB strains are those resistant TB strains which have resistance against all first-line and second-line drugs. TDR-TB has been reported in India also.

Above all, association of HIV with TB has made the DOTS therapy inefficient. Tuberculosis (TB) is the largest cause of death in human immunodeficiency virus type 1 (HIV-1) infection, having claimed an estimated one third to one half of the 30 million AIDS deaths that have occurred worldwide [1, 28]. In such circumstances, the second-line drugs are prescribed in combination with DOTS. However, this combination of drugs is very expensive and has to be administered for a longer duration with significant side effects.

One major drawback of current TB therapy is that the drugs are administered for at least 6 months. The length of therapy makes patient compliance difficult, and such patients become potent source of drug-resistant strains. The second major and serious

problem of current therapy is that most of the TB drugs available today are ineffective against persistent bacilli, except for RIF and PZA. RIF is active against both actively growing and slow metabolizing nongrowing bacilli, whereas PZA is active against semi-dormant nongrowing bacilli. However, there are still persistent bacterial populations that are not killed by any of the available TB drugs. Therefore, there is a need to design new drugs that are more active against slowly growing or nongrowing persistent bacilli to treat the population at risk of developing active disease through reactivation. Secondly, it is important to achieve a shortened therapy schedule to encourage patient's compliance and to slow down the development of drug resistance in *Mycobacterium*.

Multi-target drug therapy may help a lot to reduce the chances of resistance development, and since single molecule will act on multiple targets, so it will also be able to reduce the toxicity or side effect of the drug. Although most of the presently available drugs have multiple targets, but when we use the term "multi-target drug," it senses for the drug having the ability to target different known targets, and all targets must have some contribution to control that specific disease.

Recent advancement in computational modeling and molecular interaction study open the window of prelaboratory screening of molecules. This increases the success rate of bench experiments [29, 30]. These computational experiments are used to direct the drug discovery against a specific disease [31]. The drug may act against a single target, or it may subvene multiple targets against the same disease. This multi-targeting approach increases the rate of success of that specific drug with the high chance of successful inhibition of at least one target and accomplishes the purpose of multi-targeting lead identification. This approach also decreases the probability of resistance development and lowers the toxicity effect by decreasing effective dose [29, 32–34]. Advantages of *multi-target* drugs are as follows:

- Multi-target or multifunctional lead molecules may replace the combination therapy of drugs and help to decrease the side effects of combination therapies.
- Toxicity-related issues, arising because of multiple drug intakes, can be minimized with the treatment of single multi-target drug.
- From the patient's point of view, in multi-target drug therapy, patients only have to remember to take a single drug instead of remembering to take multiple drugs.
- In order to minimize the drug resistance, multi-target drugs can be directed against the key disease targets.
- In the multi-target drug therapy, chemical or drug metabolism-related issues, i.e., drug-drug interactions (DDIs), may also be avoided.

1.4 Multi-Targeting Drug Designing Against *Mycobacterium tuberculosis*

Several methods have been proposed in multi-target drug designing for diseases like cancer, Alzheimer, etc. [83–88]. However, in case of *Mycobacterium tuberculosis*, not much has been done instead of having very clear targets on those needed to be targeted simultaneously. Since combination therapy is recommended in TB, multi-target drug designing approach may be a very promising approach for targeting all the possible targets with a single molecule. Recently, Kai Li et al. have presented multi-target drug designing approach against *Mycobacterium tuberculosis* and proposed a series of new leads that have potential to restrict the pathogenicity of *Mycobacterium tuberculosis* as well as some other bacteria, fungi, and malaria parasite [89]. The proposed compounds were basically the analogues of a new anti-TB drug, which was reported for the inhibition of MmpL3. MmpL3 is an inner membrane transporter protein which exports mycolic acids in the form trehalose monomycolate (TMM) into the cell envelope. Different analogues were prepared by varying the nature of the ethylenediamine linker to provide cationic, protonatable, or neutral variants. In addition adamantyl head group was also altered. The proposed compounds were also targeting the enzymes involved in menaquinone biosynthesis and electron transport, inhibiting respiration and ATP biosynthesis. Study has shown potent inhibition of TB cell growth, as well as very low rates of spontaneous drug resistance.

In another study Alejandro Speck-Planche et al. have introduced the first chemo-bioinformatic approach for the in silico designing and virtual screening of anti-TB compounds against different MTB targets by developing a multi-target (mt) QSAR discriminant model [90]. The model was developed on the basis of datasets of 124 compounds, having inhibitory activity against six potent MTB targets. These protein targets were enoyl-[acyl-carrier-protein] reductase (InhA), DNA gyrase subunit A (GyrA), DNA gyrase subunit B (GyrB), pantothenate synthetase (PS), fibronectin-binding protein C (Ag85C), and peptide deformylase (PDF). Classical chemoinformatics- or bioinformatics-based approaches consider only small series of structurally related compounds and generally target only one protein to derive new leads, while in this study, Alejandro Speck-Planche et al. [90] have shown an effort to overcome this problem. During the training as well as prediction phase, more than 90% of active and inactive compounds were classified correctly through this model. Proposed mt-QSAR model has extracted significant fragments of small molecules and their contributions for anti-TB activity against the selected target proteins. Several fragments were identified. Further new molecular entities were also designed as possible anti-TB agents on the basis of extracted fragments.

2 Methodology

Application of computational biology has shown very promising results in the field of multi-target drug designing and also gathered a huge attention of researchers in this field. Several methods have been proposed to discover ligands with the ability of desired multi-target activities. To reach the goal, known ligands, having multiple pharmacological activities, are always a primary choice to renovate them by removing the unwanted activity and retaining only desired activity profile. In case of unavailability of known multi-target leads of the desired set of pharmacological activity, a new molecule may be designed using the target-centric methods. This includes classical computational drug design methods like searching of 3D similarity, pharmacophore studies, docking, and molecular dynamics simulation studies, in series or parallel [35, 36]. In this chapter, we have discussed some recent multi-target drug designing methods including docking-based, pharmacophore-based, and de novo methods.

2.1 Selection of Target Combinations for Multi-Target Drug Design

In a multi-target drug designing approach, the first question arises that how to select combination of targets, against which drug needs to be developed. Identification of a set of feasible targets against a specific disease is one of the key task [37, 38]. To resolve this issue, network analysis of disease is one of the promising methods. Topology and dynamics of a disease network provide a valuable insight to develop potential therapeutic interventions and manipulations [39–43]. Previous studies indicated that developing a set of novel synergistic drug, based on network topology, is a promising approach in multi-target drug designing. In a network with realistic dynamics, a phenotypic response could be related to the state of the network. A network model could be used to identify potential intervention through multiple drug-target interactions, which drive the network from a disease state to a healthy state [44].

A Monte Carlo simulation annealing-based algorithm, also called as multiple target optimal intervention, has also been proposed to identify combinations of drug targets, which work on ordinary differential equation (ODE)-based network models. This method has been applied on human arachidonic acid metabolic network and predicted combinations of potential drug target with high efficacy and lesser side effects. The resulting targets have also been validated experimentally [45–47].

After detection or recognition of combinations of target proteins, their interaction sites with small molecules should also be known [48]. This can be achieved through different detection methods like Q-SiteFinder, CAVITY, LIGSITE, SiteMap, etc. [49–52]. These programs identify the binding sites of target proteins on the basis of their structures. In this way this is highly possible

that proteins with similar binding sites bind with the same ligand, but in case of distinctly related or unrelated biological targets, comparison of binding sites through computational methods may provide significant indications. One of the binding site comparison programs is Apoc [53]. It gives promising results for binding site selection in multi-target drug designing. In the Apoc program, pocket similarity score (PS-score) is calculated on the basis of combination of side chain orientation, backbone geometry, and chemical similarity of binding site residue. This score can be used to judge the similarity or dissimilarities between the sites of distantly related or unrelated proteins. After target selection and their structure-based binding site prediction, the next step in the multi-target drug designing is lead identification against different targets. Existing methods for multi-target drug designing are discussed below.

2.2 Computational Lead Identification Methods

2.2.1 Pharmacophore- and Docking-Based Multi-Target Drug Design Methods

Docking- and pharmacophore-based screening of lead molecules are two highly popular computational methods in the field of single-target-based drug discovery. Many related tools have been extended for finding ligands with the required biological profile. A pharmacophore model can be built on the basis of three-dimensional structure of target binding site or according to the structure of known ligand of that specific target, and then different conformations of lead molecules are mapped against that pharmacophore model with some fitness score. In the docking-based approaches, each ligand molecule needs to be placed into the binding sites and evaluated with different scoring functions. These approaches can be used sequentially or parallel, to screen the molecules having the ability to target multiple proteins. In this process, top-ranked screened molecules, showing proper interaction with two or more than two targets, are subjected to the next step. One way is to check each model and then select top-ranked molecules from each screening. However, to approach these strategies and identify common hits, multiple computational screening steps are required which raise computational expense and have to face different challenges like fitness scores for different targets and comparison of binding scores. Alternatively to lower the computational power, a pharmacophore-guided multi-target drug design strategy was also proposed [54, 55]. Initially the pharmacophore models were developed on the basis of three-dimensional structure of target sites or available active ligands (Fig. 1) [56]. Further a common pharmacophore is built by aligning all possible pharmacophore combinations. Top-ranked compounds are then passed through molecular docking experiments to identify the well-bound molecules in the initial screening. The common pharmacophore may also be used for post-filtering after multiple docking to select compounds, binding to all targets. In addition, a shape-based comparison is also proposed in this method as an alternative way for docking [57]. Similar to single-target drug designing, implementation of multi-target drug

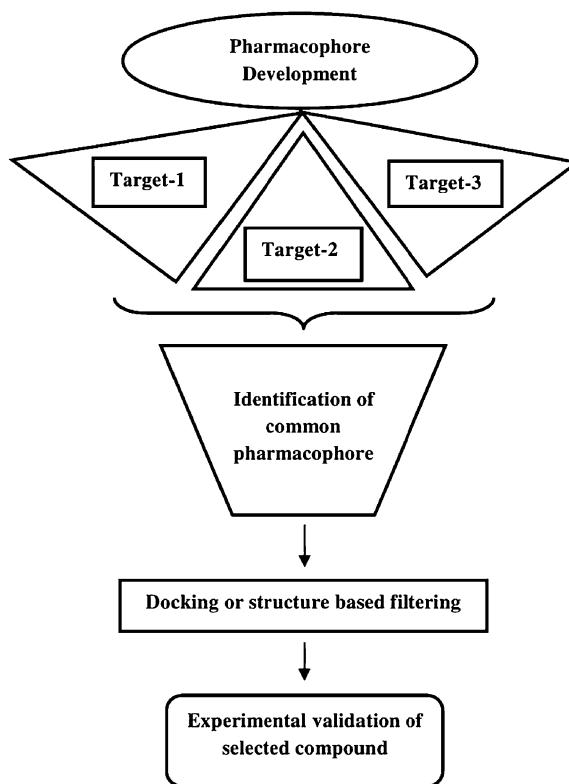


Fig. 1 Common pharmacophore-based multi-target drug design

designing methods should also address the same concerns. The first ones are false-positive hits, which should be avoided. False-positive molecules show very minor variation in comparison to active molecule; that is why they usually pass the pharmacophore screening and further show incompatibility with the target protein. In pharmacophore screening, application of volume- or shape-based filters can address these issues to some extent. In the docking experiment, false-positive results come because of scoring function and protein flexibility [58–60]. The overall docking scores and pharmacophore fitness scores are used to rank the ligands against a single target. Therefore, to avoid these errors, some statistical significance consideration should be applied to compare the scores against different targets [61–63]. In the docking-based methods, because of limited computational resources, target flexibility is not supposed to be fully taken into consideration; therefore the fit criteria of pharmacophore models may get changed to provide some sort of flexibility in the model [64]. These limitations can be further compensated by using molecular dynamics (MD) simulation-based validation, like MM-PB/GBSA free energy calculation, [65] or by examining different conformational representatives from the FEL analysis on the trajectories from MD simulations [66, 67]. The overall challenges in

developing a good pharmacophore model for multi-target drug discovery are quite similar to the single-target ligand designing which includes proper selection of ligands, handling of ligand conformational flexibility, selection of key chemical features from ligand functional groups or target residues, and proper pharmacophore alignments [64]. Although low pocket similarities do not allow an easy identification of sufficient number of multi-target candidates, a well-designed library, covering a broad range of chemical space with a sufficient number of candidates, can resolve this issue.

2.2.2 De Novo Design-Based Methods for Multi-Target Drug Designs

Apart from the pharmacophore-based or molecular docking-based approaches, de novo drug designing method is a way to develop new molecules to ensure that the resulting molecules meet multiple predefined objectives. This approach designs molecules on demand and could address the goal of multi-target drug designing, i.e., to design molecules of specific activity profiles. As compared to virtual screening-based methods, in de novo drug designing, the candidate pool is usually diverse and broad, which is a better way for the rational of designing and optimizing highly integrated multi-target drugs, especially for proteins with dissimilar binding pockets. Different de novo drug design programs are available which include DOGS [68], PhDD [69], SMOG [70, 71], SYNOPSIS [72], FLUX [73, 74], LUDI [75], LigBuilder [76, 77], etc. De novo drug design can be classified as structure-based or ligand-based. De novo designed compounds usually are not usually commercially available, and their synthesis is required afterward. In structure-based methods, embedded chemical reaction database and a retrosynthesis analyzer (first introduced by LigBuilder 2) methods are used to analyze the synthetic accessibility of designed compounds [77], while in ligand-based methods, the compounds get synthesized by transformations of most common chemical reaction, e.g., SYNOPSIS and DOGS [68, 72]. A typical scheme of de novo multi-target drug designing includes several steps which are shown in Fig. 2. In this way several of ligand-based [78] and structure-based de novo multi-target drug designing methods [79–82] have shown promising results, which are discussed below:

Ligand-Based De Novo Methods for Multi-Target Drug Design

In the ligand-based de novo multi-target drug design methods, the known active ligands of each of the desired targets can be taken as a reference or a starting structure to generate a virtual ligand library. These can also be used in the training dataset for prediction modeling of ligand target interactions. The initial reference structure of ligand which satisfies at least part of the objective functions is used as the starting point, and further different candidate structures are generated using a predefined transformation/growth scheme (Fig. 2). All the compounds are then examined for their fitness against several optimal objective functions. These objectives usually include the desired interaction pattern with the specific target which may get calculated using

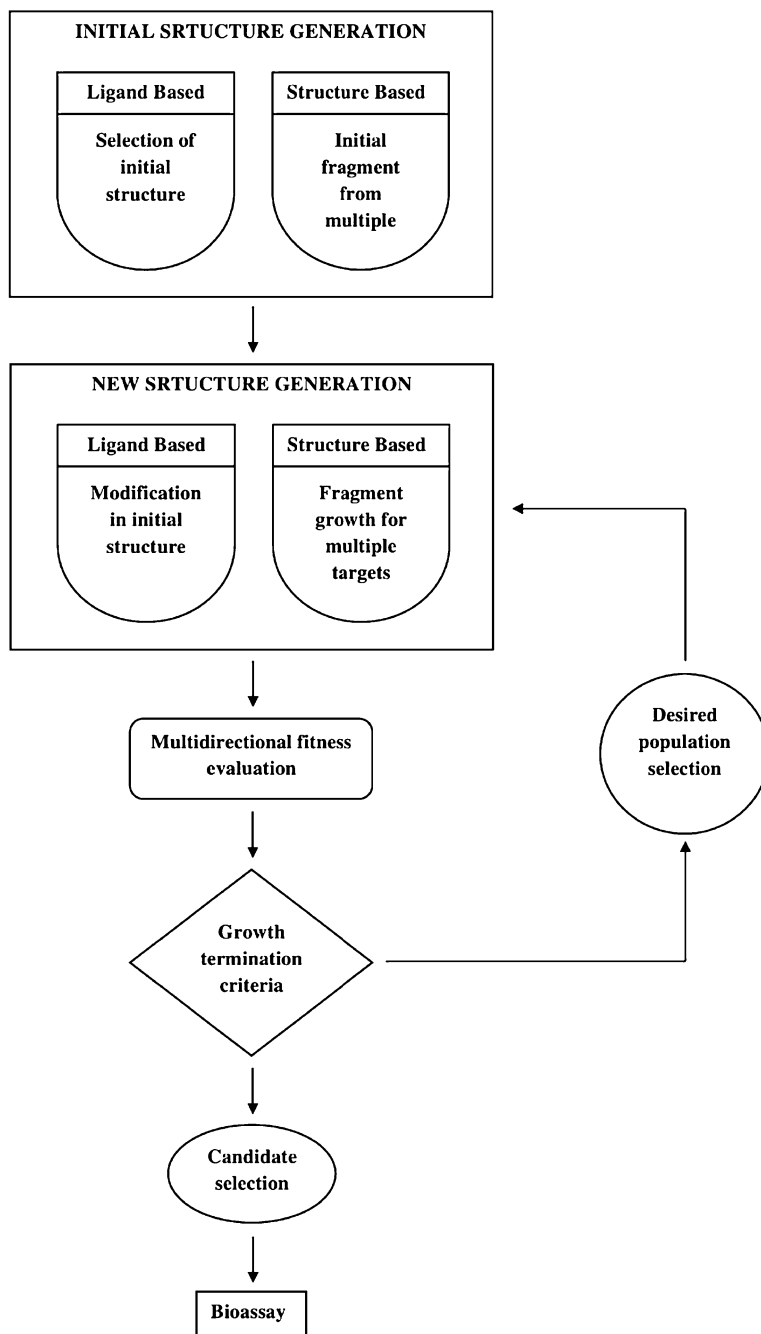


Fig. 2 De novo multi-target drug design scheme

similarity-comparison or machine learning methods. The top-ranking compounds are then moved to another round of transformation until the desired criteria are achieved. Top representative structures are then selected for the synthesis and experimental validation such as synthetic

accessibility, novelty, and ADMET properties. Various target activity profiles and blood-brain barrier penetration ability were also utilized to validate the fitness of the generated ligands. Several successful examples of ligand-based de novo methods for multi-target drug designing have been reported [79–82].

Structure-Based De Novo Methods for Multi-Target Drug Design

Another de novo method, known as structure-based multi-target de novo approach, is an intuitive and rational method because it directly peruses the interaction of a ligand with different targets. In this method the initial seeds are first placed into the desired binding pocket with the physical energy-based assistance method like docking experiments. Candidate pools are incubated on the basis of genetic algorithm-based growth algorithm. The explicit interactions with proposed targets can be used as fitness criteria. In this way all the generated compounds are tested against fitness criteria. The top-ranked candidates are then passed to the next round of screening. However, in the structure-based de novo designing, conformations of multiple targets can be used to address the target flexibility-related issues. LigBuilder 3, an updated version of LigBuilder 2, is a successful example of structure-based, multi-target, de novo designing [77]. LigBuilder 3 can judge the accessibility of synthetically designed compound. The earlier studies have demonstrated that in case of high similarity with the binding site, common pharmacophore approaches should be used; otherwise docking-based or de novo design may be recommended.

Moreover, these computational multi-target drug designing methods may also include MD simulation, binding free energy calculation through MM-PB/GBSA method, FEL analysis, and steered molecular dynamics (SMD) study to validate and strengthen the findings. In case of *Mycobacterium tuberculosis*, continuous emergence of resistant strains is still demanding to replace the available drugs with more effective and potential lead molecules which may be able to interact with two or more than two drug targets, especially with first-line drugs. As it was already discussed that in case of tuberculosis not much has been done by utilizing multi-target drug designing approaches, therefore, a method is discussed below which includes a generalized protocol of MD simulation, MM-PBSA calculation, FEL analysis, and SMD for the computational validation of the results.

2.2.3 An Illustration for Designing Multi-Targeting Drug Against Mtb

Target Selection

In multi-target drug designing, target selection is the first and very important step toward drug designing which is usually done with the network or pathway analysis of diseases. This approach may also be used to recognize some new and potential targets against the Mtb or targets of known TB drugs; potential to restrict the TB virulence will also be a good option. Several first-line drugs like isoniazid, pyrazinamide, rifampin, streptomycin, and ethambutol

are available against tuberculosis (Table 2). DOTS, the current TB therapy, consists of an initial phase of treatment with four drugs, isoniazid (INH), rifampin (RIF), pyrazinamide (PZA), and ethambutol (EMB), for a time period, followed by the treatment with INH and RIF for another fixed time period. So the targets of these four important drugs may be taken into consideration as the key targets to develop multi-target drugs. INH, RIF, PZA, and EMB mainly target enoyl-acyl reductase (InhA), RNA polymerase B subunit (RpoB), fatty acid synthetase I (RpsA), and arabinosyl transferase (EmbB), respectively. The three-dimensional crystal structures of these target proteins are available in PDB with the PDBID of 1BVR [91], 5UHB [92], and 4NNI [93] for InhA, RpoB, and RpsA, respectively. In this case, binding pocket of all the three targets has already been well studied, and there is no need of predicting active sites of these target proteins [91–93].

Retrieval of Small Molecules

Around 1.4 million anti-TB compounds are available in the PubChem [94], which can be retrieved. At the same time, reference compounds for each target should also be retrieved. Reference compounds are used to compare the *in silico* findings with the experimentally proven data. In this case INH, RIF, PZA, and EMB will be the reference compounds for InhA, RpoB, RpsA, and EmbB, respectively. All the retrieved compounds along with reference compounds should pass through the Lipinski filters. Lipinski rule of five helps to distinguish drug-like and non-drug-like molecules. It predicts high probability of success or failure due to drug likeness for molecules complying with two or more of the Lipinski parameters [95]. To pass the Lipinski filter, molecules should pass the criteria of having:

- Molecular mass should be less than 500 Da.
- Optimum lipophilicity (expressed as LogP should be less than 5).
- Less than five hydrogen bond donors.
- Less than ten hydrogen bond acceptors.
- Molecular refractivity should be within 40–130.

These filters help in early preclinical development and could help to avoid costly late-stage preclinical or clinical failures [96].

Molecular Docking

After retrieval and preparation of target proteins and screening of small molecule library with the Lipinski filter, the next step is docking of each target molecule with each drug-like small molecule. PyRx virtual screening tool [97], connected with Autodock Vina [98], is a widely used tool for fast library screening through molecular docking. In this way, to prepare the PDBQT file of both receptor protein and ligand, Autodock tool [99] may also be used

with merging of nonpolar hydrogens and addition of Gasteiger charges. Charge deficits should spread all over the atoms of related residues. These PDBQT files are further utilized in Autodock Vina for the molecular docking. In the docking experiment, some parameters need to be set according to the systems. For example, when the binding pocket of the target is well-known, the grid box should focus that specific pocket with settling the grid dimensions accordingly; otherwise grid box should cover entire target protein. Binding sites of all the target proteins (i.e., InhA, RpoB, RpsA, and EmbB) are well studied. For a reliable docking result in a global search, exhaustiveness should be near to 600. In Autodock Vina, binding energy is used as scoring function. After virtual screening, this need to be manually observed that how many common leads exist in the drug library showing good binding against all the target protein. This is done on the basis of binding energy of each small molecule in comparison to the docking score of respective reference compound against each target. Common molecules, found to have good binding energy against all the targets, should again be docked individually with each target through Autodock tool. This will validate the result as well as achieve better binding conformation. Docking parameters may be set as follows: number of GA runs 200, population size 200, maximum number of evaluations 25,000,000, maximum number of generations 2,700,000, maximum number of top individuals that automatically survive 1, rate of gene mutations 0.02, rate of crossover 0.8, GA crossover mode 2 points, mean of Cauchy distribution for gene mutation 0.0, variance of Cauchy distribution for gene mutation 1.0, and number of generations for picking worst individuals 10. Binding poses may be generated through Lamarckian genetic algorithm (LGA) or simulation annealing or through local search parameters [100]. Further, docked conformations may be clustered on the basis of root mean square deviation (RMSD) of ligand. The best conformations from each docked complexes should be obtained on the basis of binding energies of clusters. The selected conformations may be further analyzed through MD simulation to evaluate the interaction of each lead molecule with their respective target in comparison to interaction of their reference molecules.

Molecular Dynamics Simulation

Several tools like GROMACS [101], AMBER [102], and NAMD [103] are available for MD simulation study. Out of these, GROMACS is an open-source and highly accepted platform to perform MD simulation studies. In GROMACS, force field parameters of small molecules are required to be generated outside the GROMACS platform, e.g., prodrug server or other pipelines which utilize Antechamber [104] and Acypype [105]. GROMACS provides several options of force field to generate the topology of macromolecules. It includes AMBER, CHARMM, GROMOS, and

OPLS force fields with their different variants. Protein-ligand complexes or apoproteins must be placed in the center of the box having a distance of 1.0 nm between the protein and edge of the simulation box. In GROMACS, the simulation box may be of triclinic, cubic, dodecahedron, or octahedron nature. To solvate the system, GROMACS provides the option of different types of solvent models which include SPC, SPC/E, or TIP3P, TIP4P, and TIP5P. According to the selected force field, the solvent model is also recommended, e.g., GROMOS force field and SPC water model, and with AMBER force field, TIP3P water model is recommended. After solvation, the desired numbers of positive or negative ions are to be added to neutralize the charges of the system. Na^+ is used as positive ion and Cl^- is used as negative ion. After the system assembly, the system must pass through an energy minimization step by using steepest descent or conjugate gradient method [106], and then the system is equilibrated. Equilibration is often conducted in two phases. The first phase is conducted under an *NVT* ensemble (constant number of particles, volume, and temperature). This ensemble is also referred to as “isothermal-isochoric” or “canonical.” The timeframe for such a procedure is dependent upon the contents of the system, but in *NVT*, the temperature of the system should reach a plateau at the desired value. If the temperature has not yet stabilized, additional time will be required, typically, e.g., 50–100 ps should suffice. *NVT* equilibration stabilizes the temperature of the system. The next is to stabilize the pressure (and thus also the density) of the system. Equilibration of pressure is conducted under an *NPT* ensemble, wherein the number of particles, pressure, and temperature are all constant. The ensemble is also called the isothermal-isobaric ensemble, and most closely resembles experimental conditions. Finally, the production MD is performed for a time period depending upon the availability of the computational power. A stepwise protocol of GROMACS tutorial has been available by Bevan Lab (<http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/index.html>).

Trajectory Analysis

The generated trajectories may be analyzed further through different GROMACS analysis tools like *g_rms* for root mean square deviation (RMSD), *g_rmsf* for root mean square fluctuation (RMSF), *g_gyrate* for radius of gyration (Rg), *g_sas* for solvent-accessible surface area (SASA), *g_hbond* for hydrogen bonds, etc. H-bond occupancy can also be calculated by a python script. The resulting plots can be analyzed through Xmgrace [107], while the visual analysis and figure preparation can be done with VMD [108], CHIMERA [109], and Pymol [110, 111].

- Binding Energy Calculation** Binding free energy calculation through the MM-PB/GBSA method may be done on different platforms, depending upon the trajectory format. In AMBER and NAMD, MM-PB/GBSA can be calculated on the same platform, while against the GROMACS MD trajectory, the `g_mmpbsa` tool is used for the calculation of binding free energy between protein-ligand complexes [112, 113]. The `g_mmpbsa` uses the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) approach for the binding free energy estimation. It calculates the molecular mechanics potential energy (electrostatic and van der Waals interactions) and solvation free energy (polar and nonpolar solvation energies) but do not calculate the entropy function of the system. MM-PBSA calculations should be done on the last stable trajectories (by excluding initial unstable trajectories).
- PCA and FEL** PCA of any system can be done using GROMACS tools. PCA defines all the essential motions, governing conformational transitions throughout the simulation period [114, 115]. Collective motion of most dominant, initial eigenvectors determines the sub-conformational structural transitions of protein [115]. FELs [116] can also be generated on the basis of estimation of joint probability distribution of top two highly contributing eigenvectors.
- Steered Molecular Dynamics (SMD)** Parallel to MM-PBSA, the use of SMD studies strengthens the findings and reduces the chances of getting false-positive results. SMD is a promising tool to analyze single biomolecules using the external force as an additional variable [117, 118]. This can also be utilized to probe the binding affinity of ligand by pulling the ligand from the receptor binding site [119]. Previous studies have reported that the accuracy of the SMD method is compatible with that of the MM-PBSA method, but its computational speed is much higher [120–122]. Since SMD is about 133-fold faster than MM-PBSA, it can be used to refine the docking results in virtual screening. In SMD the ligand is attached to a dummy atom via a spring with spring constant k , and the dummy atom is moved with a constant velocity v along the direction allowing a smooth exit from the binding site. Thus, during pulling the force, exerted by the dummy atom on the ligand, is
- $$F = k(\Delta x - vt)$$
- where Δx is a displacement of pulled atom from the initial position [120–122]. To prevent the receptor from drifting together with the ligand during the pulling, C α -atoms should be restrained with maintaining flexibility of side chain.
- Possible pathways for ligand to escape from the binding pocket can be determined using CAVER 3.0 [123], a plug-in of Pymol. The easiest or optimum path with the lowest rupture force should be chosen [124]. The ligand changes direction during exit from the

binding site. Therefore the drawback of SMD with a single pulling direction is that it does not take into account multidirectional movement. Recently, Yang et al. [125] and Gu et al. [126] have proposed a SMD method with adaptive direction adjustments where the optimum path of ligand is navigated by minimizing the pulling force automatically during the simulation.

The self-adaptive SMD also yielded a good correlation between the rupture forces and experimentally measured binding free energies for two sets of protein-ligand complexes [126]. As the correlation level from standard SMD [122] is compatible with that of self-adaptive SMD, SMD with single pulling direction is also a good option which can be implemented in GROMACS. In GROMACS, receptor-ligand complexes should be solvated in a box filled with water. After equilibration, 500 ns of SMD simulations can be performed in the NPT mode. These runs are long enough to get the ligand out from the active site completely. To obtain good statistics, five independent runs, starting from the same initial conformation, may be carried out but with different velocity distributions. In SMD one can choose either rupture force F_{\max} or non-equilibrium work W_{pull} as a scoring function to rank binding affinities.

In this way the resulting lead molecules, showing good results in each step of analysis as compared to the control one, may be further evaluated using in vitro and in vivo testing.

3 Conclusion

TB, an infectious disease, is still sustaining as one of the major concerns for the scientific fraternity, dealing with the medication of TB disease. Although a range of drug regimens are available against TB, consistently emerging TB-resistant strains are making the available treatment inefficient. Previous studies have indicated that long treatment course and inefficacy of available drug against the persistent TB infections are some of the major reasons behind the emerging resistance. So, the current TB status is demanding for some new alternatives of drug regimens against its emerging resistant strains. To handle these issues to some extent, a combination therapy is very popular for the TB treatment, but it again raises the toxicity level as well as makes the drug administration more complex. To overcome these challenges, multi-target drug designing approach is one of the options through which multiple targets may be addressed simultaneously. Currently, multi-target drug designing is a highly acceptable and promising approach in the field of drug designing. It raises the success rate of any proposed compound with more chances of successful inhibition of at least one target. Multi-target drugs reduce the complexity of the treatment,

ease the combination therapy, and reduce the toxicity by reducing the number of drugs. In case of Alzheimer and cancer, multi-target drug designing is quite popular, while in case of TB, not much work has been done. Since combination therapy is highly required for the treatment of Mtb, therefore, multi-target drug designing approach may prove as a more fruitful approach. Studies have already shown that replacement of the first-line anti-TB drugs with a single or less number of drugs will be a great breakthrough in TB treatment, and presently, this seems possible with multi-target drug designing approach.

Acknowledgments

G.S. and A.T. are thankful to ICMR, New Delhi, for their ICMR-SRF and ICMR-RA fellowship. Authors are also thankful to BTIS-net program of DBT, New Delhi.

References

1. World Health Organization (2016) Global TB Rep 2016
2. Cambau E, Drancourt M (2014) Steps towards the discovery of Mycobacterium tuberculosis by Robert Koch, 1882. *Clin Microbiol Infect* 20(3):196–201
3. Koike M, Takeya K (1961) Fine structures of intracytoplasmic organelles of mycobacteria. *J Cell Biol* 9(3):597–608
4. Inaeda T, Ogura M (1963) Formation of intracytoplasmic membrane system of mycobacteria related to cell division. *J Bacteriol* 85 (1):150–163
5. Draper P (1971) The walls of Mycobacterium lepraemurium: chemistry and ultrastructure. *Microbiology* 69(3):313–324
6. Draper P, Kandler O, Darbre A (1987) Peptidoglycan and arabinogalactan of Mycobacterium leprae. *Microbiology* 133(5):1187–1194
7. Draper P (1998) The outer parts of the mycobacterial envelope as permeability barriers. *Front Biosci* 3:D1253–D1261
8. Moore DF, Curry JI (1998) Detection and identification of Mycobacterium tuberculosis directly from sputum sediments by ligase chain reaction. *J Clin Microbiol* 36 (4):1028–1031
9. Bloom BR (ed) (1994) Tuberculosis: pathogenesis, protection, and control. ASM Press, Washington, DC
10. Baltimore RS (2001) Tuberculosis: current concepts and treatment. *Yale J Biol Med* 74 (6):413
11. Grange JM (1988) Mycobacteria and human disease. Edward Arnold (Publishers) Ltd, London
12. Rom WN, Garay S (1996) Tuberculosis Boston. Little, Brown and Company, New York, Toronto, London
13. Barry CE, Boshoff HI, Dartois V, Dick T, Ehrt S, Flynn J, Schnappinger D, Wilkinson RJ, Young D (2009) The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat Rev Microbiol* 7 (12):845–855
14. Harries AD, Dye C (2006) Tuberculosis. *Ann Trop Med Parasitol* 100(5–6):415–431
15. Sutherland I, Švandová E, Radhakrishna S (1982) The development of clinical tuberculosis following infection with tubercle bacilli: 1. A theoretical model for the development of clinical tuberculosis following infection, linking from data on the risk of tuberculous infection and the incidence of clinical tuberculosis in the Netherlands. *Tubercle* 63(4):255–268
16. Prasanthi K, Murty DS (2014) A brief review on ecology and evolution of mycobacteria. *Mycobact Dis* 4(6). <https://doi.org/10.4172/2161-1068.1000172>
17. Ernst JD (2012) The immunological life cycle of tuberculosis. *Nat Rev Immunol* 12 (8):581–591

18. Zahrt TC (2003) Molecular mechanisms regulating persistent Mycobacterium tuberculosis infection. *Microbes Infect* 5(2):159–167
19. Bodnar KA, Serbina NV, Flynn JL (2001) Fate of Mycobacterium tuberculosis within murine dendritic cells. *Infect Immun* 69(2):800–809
20. Nguyen L (2016) Antibiotic resistance mechanisms in *M. tuberculosis*: an update. *Arch Toxicol* 90(7):1585
21. Davies PD (2003) The role of DOTS in tuberculosis treatment and control. *Am J Respir Med* 2(3):203–209
22. Zignol M, Gemert WV, Falzon D, Sismanidis C, Glaziou P, Floyd K, Raviglione M (2012) Surveillance of anti-tuberculosis drug resistance in the world: an updated analysis, 2007–2010. *Bull World Health Organ* 90(2):111–119
23. Kahana LM (1996) The problem of drug resistance in tuberculosis. *Chest* 110(1):8–10
24. Colditz GA, Brewer TF, Berkey CS, Wilson ME, Burdick E, Fineberg HV, Mosteller F (1994) Efficacy of BCG vaccine in the prevention of tuberculosis: meta-analysis of the published literature. *JAMA* 271(9):698–702
25. Horwitz MA, Harth G, Dillon BJ, Masleša-Galić S (2000) Recombinant bacillus Calmette–Guérin (BCG) vaccines expressing the Mycobacterium tuberculosis 30-kDa major secretory protein induce greater protective immunity against tuberculosis than conventional BCG vaccines in a highly susceptible animal model. *Proc Natl Acad Sci U S A* 97(25):13853–13858
26. Unissa AN, Selvakumar N, Narayanan S, Suganthi C, Hanna LE (2015) Investigation of Ser315 substitutions within katG gene in isoniazid-resistant clinical isolates of Mycobacterium tuberculosis from south India. *Biomed Res Int* 2015:257983
27. Centers for Disease Control and Prevention (2006) Revised definition of extensively drug-resistant tuberculosis. *MMWR Morb Mortal Wkly Rep* 55(1176):1
28. Revised National Tuberculosis Control Programme: National Strategic Plan for Tuberculosis Control 2012–2017
29. Jenwitheesuk E, Horst JA, Rivas KL, Van Voorhis WC, Samudrala R (2008) Novel paradigms for drug discovery: computational multitarget screening. *Trends Pharmacol Sci* 29(2):62–71
30. Costin JM, Jenwitheesuk E, Lok SM, Hunsperger E, Conrads KA, Fontaine KA, Rees CR, Rossmann MG, Isern S, Samudrala R, Michael SF (2010) Structural optimization and de novo design of dengue virus entry inhibitory peptides. *PLoS Negl Trop Dis* 4(6):e721
31. Schneider G, Fechner U (2005) Computer-based de novo design of drug-like molecules. *Nat Rev Drug Discov* 4(8):649–663
32. Rogawski MA (2000) Low affinity channel blocking (uncompetitive) NMDA receptor antagonists as therapeutic agents—toward an understanding of their favorable tolerability. *Amino Acids* 19(1):133–149
33. Nezami A, Kimura T, Hidaka K, Kiso A, Liu J, Kiso Y, Goldberg DE, Freire E (2003) High-affinity inhibition of a family of Plasmodium falciparum proteases by a designed adaptive inhibitor. *Biochemistry* 42(28):8459–8464
34. Csermely P, Agoston V, Pongor S (2005) The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol Sci* 26(4):178–182
35. Pei J, Yin N, Ma X, Lai L (2014) Systems biology brings new dimensions for structure-based drug design. *J Am Chem Soc* 136(33):11556–11565
36. Schneider G (2014) Future de novo drug design. *Mol Inform* 33(6–7):397–402
37. Ramaswamy S (2007) Rational design of cancer-drug combinations. *N Engl J Med* 357(3):299–300
38. Kitano H (2007) A robustness-based approach to systems-oriented drug design. *Nat Rev Drug Discov* 5(3):202–210
39. Albert R, Jeong H, Barabási AL (2000) Error and attack tolerance of complex networks. *Nature* 406(6794):378–382
40. Ma W, Trusina A, El-Samad H, Lim WA, Tang C (2009) Defining network topologies that can achieve biochemical adaptation. *Cell* 138(4):760–773
41. Boran AD, Iyengar R (2010) Systems approaches to polypharmacology and drug discovery. *Curr Opin Drug Discov Devel* 13(3):297
42. Radhakrishnan ML, Tidor B (2008) Optimal drug cocktail design: methods for targeting molecular ensembles and insights from theoretical model systems. *J Chem Inf Model* 48(5):1055–1073
43. Hu Y, Bajorath J (2013) Systematic identification of scaffolds representing compounds active against individual targets and single or multiple target families. *J Chem Inf Model* 53(2):312–326
44. Zhao S, Iyengar R (2012) Systems pharmacology: network analysis to identify multiscale mechanisms of drug action. *Annu Rev Pharmacol Toxicol* 52:505–521

45. Meng H, Liu Y, Lai L (2015) Diverse ways of perturbing the human arachidonic acid metabolic network to control inflammation. *Acc Chem Res* 48(8):2242–2250
46. Yang K, Bai H, Ouyang Q, Lai L, Tang C (2008) Finding multiple target optimal intervention in disease-related molecular network. *Mol Syst Biol* 4(1):228
47. Yang K, Ma W, Liang H, Ouyang Q, Tang C, Lai L (2007) Dynamic simulations on the arachidonic acid metabolic network. *PLoS Comput Biol* 3(3):e55
48. Overington JP, Al-Lazikani B, Hopkins AL (2006) How many drug targets are there? *Nat Rev Drug Discov* 5(12):993–996
49. Halgren T (2007) New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des* 69(2):146–148
50. Laurie AT, Jackson RM (2005) Q-SiteFinder: an energy-based method for the prediction of protein–ligand binding sites. *Bioinformatics* 21(9):1908–1916
51. Hendlich M, Rippmann F, Barnickel G (1997) LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *J Mol Graph Model* 15(6):359–363
52. Yuan Y, Pei J, Lai L (2013) Binding site detection and druggability prediction of protein targets for structure-based drug design. *Curr Pharm Des* 19(12):2326–2333
53. Gao M, Skolnick J (2013) APoc: large-scale identification of similar protein pockets. *Bioinformatics* 29(5):597–604
54. Haupt VJ, Daminelli S, Schroeder M (2013) Drug promiscuity in PDB: protein binding site similarity is key. *PLoS One* 8(6):e65894
55. Günther S, Senger C, Michalsky E, Goede A, Preissner R (2006) Representation of target-bound drugs by computed conformers: implications for conformational libraries. *BMC Bioinformatics* 7(1):293
56. Dixon SL, Smondryev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA (2006) PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J Comput Aided Mol Des* 20(10–11):647–671
57. Moser D, Wisniewska JM, Hahn S, Achenbach J, Buscató EL, Klingler FM, Hofmann B, Steinhilber D, Proschak E (2012) Dual-target virtual screening by pharmacophore elucidation and molecular shape filtering. *ACS Med Chem Lett* 3(2):155–158
58. Irwin JJ, Shoichet BK (2016) Docking screens for novel ligands conferring new biology. *J Med Chem* 59(9):4103–4120
59. Liu J, He X, Zhang JZ (2013) Improving the scoring of protein–ligand binding affinity by including the effects of structural water and electronic polarization. *J Chem Inf Model* 53(6):1306–1314
60. Verdonk ML, Giangreco I, Hall RJ, Korb O, Mortenson PN, Murray CW (2011) Docking performance of fragments and druglike compounds. *J Med Chem* 54(15):5422–5431
61. Schomburg KT, Bietz S, Briem H, Henzler AM, Urbaczek S, Rarey M (2014) Facing the challenges of structure-based target prediction by inverse virtual screening. *J Chem Inf Model* 54(6):1676–1686
62. Lauro G, Romano A, Riccio R, Bifulco G (2011) Inverse virtual screening of antitumor targets: pilot study on a small database of natural bioactive compounds. *J Nat Prod* 74(6):1401–1407
63. Wang X, Pan C, Gong J, Liu X, Li H (2016) Enhancing the enrichment of pharmacophore-based target prediction for the polypharmacological profiles of drugs. *J Chem Inf Model* 56(6):1175–1183
64. Yang SY (2010) Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov Today* 15(11):444–450
65. Xie L, Evangelidis T, Xie L, Bourne PE (2011) Drug discovery using chemical systems biology: weak inhibition of multiple kinases may contribute to the anti-cancer effect of nelfinavir. *PLoS Comput Biol* 7(4):e1002037
66. Wu Y, He C, Gao Y, He S, Liu Y, Lai L (2012) Dynamic modeling of human 5-lipoxygenase–inhibitor interactions helps to discover novel inhibitors. *J Med Chem* 55(6):2597–2605
67. Shang E, Wu Y, Liu P, Liu Y, Zhu W, Deng X, He C, He S, Li C, Lai L (2014) Benzo[d]isothiazole 1,1-dioxide derivatives as dual functional inhibitors of 5-lipoxygenase and microsomal prostaglandin E(2) synthase-1. *Bioorg Med Chem Lett* 24(12):2764–2767
68. Hartenfeller M, Zettl H, Walter M, Rupp M, Reisen F, Proschak E, Weggen S, Stark H, Schneider G (2012) DOGS: reaction-driven de novo design of bioactive compounds. *PLoS Comput Biol* 8(2):e1002380
69. Huang Q, Li LL, Yang SY (2010) PhDD: a new pharmacophore-based de novo design

- method of drug-like molecules combined with assessment of synthetic accessibility. *J Mol Graph Model* 28(8):775–787
70. DeWitte RS, Ishchenko AV, Shakhnovich EI (1997) SMOG: de novo design method based on simple, fast, and accurate free energy estimates. 2. Case studies in molecular design. *J Am Chem Soc* 119(20):4608–4617
 71. DeWitte RS, Shakhnovich EI (1996) SMOG: de novo design method based on simple, fast, and accurate free energy estimates. 1. Methodology and supporting evidence. *J Am Chem Soc* 118(47):11733–11744
 72. Vinkers HM, de Jonge MR, Daeyaert FF, Heeres J, Koymans LM, van Lenthe JH, Lewi PJ, Timmerman H, Van Aken K, Janssen PA (2003) Synopsis: synthesize and optimize system in silico. *J Med Chem* 46(13):2765–2773
 73. Fechner U, Schneider G (2006) Flux (1): a virtual synthesis scheme for fragment-based de novo design. *J Chem Inf Model* 46(2):699–707
 74. Fechner U, Schneider G (2007) Flux (2): comparison of molecular mutation and crossover operators for ligand-based de novo design. *J Chem Inf Model* 47(2):656–667
 75. Böhm HJ (1992) The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J Comput Aided Mol Des* 6(1):61–78
 76. Wang R, Gao Y, Lai L (2000) LigBuilder: a multi-purpose program for structure-based drug design. *J Mol Model* 6(7–8):498–516
 77. Yuan Y, Pei J, Lai L (2011) LigBuilder 2: a practical de novo drug design approach. *J Chem Inf Model* 51(5):1083–1091
 78. Shang E, Yuan Y, Chen X, Liu Y, Pei J, Lai L (2014) De novo design of multitarget ligands with an iterative fragment-growing strategy. *J Chem Inf Model* 54(4):1235–1241
 79. Besnard J, Ruda GF, Setola V, Abecassis K, Rodriguiz RM, Huang XP, Norval S, Sassano MF, Shin AI, Webster LA, Simeons FR (2012) Automated design of ligands to polypharmacological profiles. *Nature* 492(7428):215–220
 80. Reutlinger M, Rodrigues T, Schneider P, Schneider G (2014) Multi-objective molecular de novo design by adaptive fragment prioritization. *Angew Chem Int Ed* 53(16):4244–4248
 81. Reutlinger M, Rodrigues T, Schneider P, Schneider G (2014) Combining on-chip synthesis of a focused combinatorial library with computational target prediction reveals imidazopyridine GPCR ligands. *Angew Chem Int Ed* 53(2):582–585
 82. Rodrigues T, Hauser N, Reker D, Reutlinger M, Wunderlin T, Hamon J, Koch G, Schneider G (2015) Multidimensional De novo design reveals 5-HT_{2B} receptor-selective ligands. *Angew Chem* 127(5):1571–1575
 83. Giordano S, Petrelli A (2008) From single- to multi-target drugs in cancer therapy: when aspecificity becomes an advantage. *Curr Med Chem* 15(5):422–432
 84. Lu JJ, Pan W, Hu YJ, Wang YT (2012) Multi-target drugs: the trend of drug research and development. *PLoS One* 7(6):e40262
 85. Cavalli A, Bolognesi ML, Minarini A, Rosini M, Tumiatti V, Recanatini M, Melchiorre C (2008) Multi-target-directed ligands to combat neurodegenerative diseases. *J Med Chem* 51(3):347–372
 86. Kumar A, Sharma A (2018) Computational modeling of multi-target-directed inhibitors against Alzheimer's disease. In: Computational modeling of drugs against Alzheimer's disease. Humana Press, New York, NY, pp 533–571
 87. Bajda M, Guziar N, Ignasik M, Malawska B (2011) Multi-target-directed ligands in Alzheimer's disease treatment. *Curr Med Chem* 18(32):4949–4975
 88. Domínguez JL, Fernández-Nieto F, Castro M, Catto M, Paleo MR, Porto S, Sardina FJ, Brea JM, Carotti A, Villaverde MC, Sussman F (2014) Computer-aided structure-based design of multitarget leads for Alzheimer's disease. *J Chem Inf Model* 55(1):135–148
 89. Li K, Schurig-Briccio LA, Feng X, Upadhyay A, Pujari V, Lechartier B, Fontes FL, Yang H, Rao G, Zhu W, Gulati A (2014) Multitarget drug discovery for tuberculosis and other infectious diseases. *J Med Chem* 57(7):3126–3139
 90. Speck-Planche A, Kleandrova V, Luan F, Cordeiro ND (2012) In silico discovery and virtual screening of multi-target inhibitors for proteins in Mycobacterium tuberculosis. *Comb Chem High Throughput Screen* 15(8):666–673
 91. Rozwarski DA, Vilchèze C, Sugantino M, Bittman R, Sacchettini JC (1999) Crystal structure of the Mycobacterium tuberculosis enoyl-ACP reductase, InhA, in complex with NAD⁺ and a C16 fatty acyl substrate. *J Biol Chem* 274(22):15582–15589

92. Lin W, Mandal S, Degen D, Liu Y, Ebright YW, Li S, Feng Y, Zhang Y, Mandal S, Jiang Y, Liu S (2017) Structural basis of Mycobacterium tuberculosis transcription and transcription inhibition. *Mol Cell* 66(2):169–179
93. Fan Y, Dai Y, Hou M, Wang H, Yao H, Guo C, Lin D, Liao X (2017) Structural basis for ribosome protein S1 interaction with RNA in trans-translation of Mycobacterium tuberculosis. *Biochem Biophys Res Commun* 487(2):268–273
94. Bolton EE, Wang Y, Thiessen PA, Bryant SH (2008) PubChem: integrated platform of small molecules and biological activities. *Annu Rep Comput Chem* 4:217–241
95. Tiwari V, Patel V, Tiwari M (2018) In-silico screening and experimental validation reveal L-Adrenaline as anti-biofilm molecule against biofilm-associated protein (Bap) producing *Acinetobacter baumannii*. *Int J Biol Macromol* 107:1242–1252
96. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 23(1–3):3–25
97. Wolf LK (2009) Digital briefs. *Chem Eng News* 87:31
98. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31(2):455–461
99. Goodsell DS, Morris GM, Olson AJ (1996) Automated docking of flexible ligands: applications of AutoDock. *J Mol Recognit* 9(1):1–5
100. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19(14):1639–1662
101. Hess B, Kutzner C, Van Der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4(3):435–447
102. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26(16):1668–1688
103. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26(16):1781–1802
104. Wang J, Wang W, Kollman PA, Case DA (2001) Antechamber: an accessory software package for molecular mechanical calculations. *J Am Chem Soc* 222:U403
105. da Silva AW, Vranken WF (2012) ACPYPE-Antechamber python parser interface. *BMC Res Notes* 5(1):367
106. Mozolewska MA, Krupa P, Scheraga HA, Liwo A (2015) Molecular modeling of the binding modes of the iron-sulfur protein to the Jac1 co-chaperone from *Saccharomyces cerevisiae* by all-atom and coarse-grained approaches. *Proteins* 83(8):1414–1426
107. Turner PJ (2005) XMGRACE, Version 5.1.19. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR
108. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:33–38
109. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612
110. DeLano WL (2002) Pymol: an open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography* 40:82–92
111. DeLano WL (2009) PyMOL molecular viewer: updates and refinements. In: Abstracts of papers of the American Chemical Society, 2009 Aug 16, vol 238. American Chemical Society, Washington, DC
112. Kumari R, Lynn A (2011) Application of MM/PBSA in the prediction of relative binding free energy: re-scoring of docking hit-list. *J Nat Sci Biol Med* 2(3):92
113. Kumari R, Kumar R, Open Source Drug Discovery Consortium, Lynn A (2014) g_mmpbsa—a GROMACS tool for high-throughput MM-PBSA calculations. *J Chem Inf Model* 54(7):1951–1962
114. Amadei A, Linssen A, Berendsen HJ (1993) Essential dynamics of proteins. *Proteins* 17(4):412–425
115. Amadei A, Linssen AB, De Groot BL, Van Aalten DM, Berendsen HJ (1996) An efficient method for sampling the essential subspace of proteins. *J Biomol Struct Dyn* 13(4):615–625
116. Frauenfelder H, Sligar SG, Wolynes PG (1991) The energy landscapes and motions of proteins. *Urbana* 51(61801):61801

117. Isralewitz B, Gao M, Schulten K (2001) Steered molecular dynamics and mechanical functions of proteins. *Curr Opin Struct Biol* 11(2):224–230
118. Kumar S, Li MS (2010) Biomolecules under mechanical force. *Phys Rep* 486(1):1–74
119. Grubmüller H, Heymann B, Tavan P (1996) Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* 271(5251):997–999
120. Mai BK, Li MS (2011) Neuraminidase inhibitor R-125489—a promising drug for treating influenza virus: steered molecular dynamics approach. *Biochem Biophys Res Commun* 410(3):688–691
121. Suan Li M, Khanh Mai B (2012) Steered molecular dynamics—a promising tool for drug design. *Curr Bioinforma* 7(4):342–351
122. Van Vuong Q, Nguyen TT, Li MS (2015) A new method for navigating optimal direction for pulling ligand from binding pocket: application to ranking binding affinity by steered molecular dynamics. *J Chem Inf Model* 55(12):2731–2738
123. Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, Gora A, Sustr V, Klvana M, Medek P, Biedermannova L (2012) CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol* 8(10):e1002708
124. Mai BK, Viet MH, Li MS (2010) Top leads for swine influenza A/H1N1 virus revealed by steered molecular dynamics approach. *J Chem Inf Model* 50(12):2236–2247
125. Yang K, Liu X, Wang X, Jiang H (2009) A steered molecular dynamics method with adaptive direction adjustments. *Biochem Biophys Res Commun* 379(2):494–498
126. Gu J, Li H, Wang X (2015) A self-adaptive steered molecular dynamics method based on minimization of stretching force reveals the binding affinity of protein–ligand complexes. *Molecules* 20(10):19236–19251

Part IV

Databases and Web Servers



Development of a Web-Server for Identification of Common Lead Molecules for Multiple Protein Targets

Abhilash Jayaraj, Ruchika Bhat, Amita Pathak, Manpreet Singh, and B. Jayaram

Abstract

Due to increasing unresponsiveness of drugs to single targets in the form of resistance or presence of alternate mechanisms in case of complex diseases and disorders, etc., the focus is shifting towards polypharmacology. It is desirable that a drug works on multiple targets to elicit guaranteed/multiplier effect. Here, we provide a one stop solution to the quest of finding common leads for multiple protein targets. The computational protocol designed involves screening, docking, and scaffold-based optimization of hit molecules from a variety of compound libraries against any two specified protein targets. The protocol is validated with five case studies involving five pairs of proteins with varying active site similarities. The methodology is able to recover the known common FDA approved drugs against them. A web-server named “Multi-Target Ligand Design” is created and made freely accessible at <http://www.scfbio-iitd.res.in/multitarget/>.

Keywords Multi-target drug design, Polypharmacology, Scaffold-based optimization, Screening and docking, Structure based ligand design

1 Introduction

Drug discovery is a century old process, but the methods used to discover therapeutically relevant compounds have evolved over the years. Pre 1980 era, medicine discovery was based on phenotypic response of animals to compounds. The compounds were then tested on humans to test if it remedied a medical complication. Post 1980s saw the evolution of the concept of “magic bullet,” an idea first established by Paul Ehrlich [1]. This propelled the reductionist approach which saw the development of a new class of drugs which were specifically tailored to bind to a key bio-molecular target [2]. This method of drug discovery involves the selection, synthesis, and testing of compounds which are highly specific and selective to their targets. The molecules that affect one target only might not always affect the complex biosystems in the way

expected, even if they inhibit the immediate target because of the alternate mechanisms the complex organisms possess. In addition to this, many academicians, researchers, and industry experts of drug development have hinted at the slowing rate of new drug development and rise in the rate of failures and cost [3, 4]. The failures have mainly manifested themselves in clinical trials where the drug candidates fail to show efficacy or are rendered unviable due to adverse drug reaction/toxicity [5]. Based on the above facts and the evidences collected over the past two decades there is a transition from single target drug design to polypharmacology [6–14].

Polypharmacological processes include a single drug acting on multiple targets of a unique disease pathway or even multiple disease pathways [15]. Another approach towards polypharmacology is to discover the unknown off targets for the existing drugs which is also known as drug repurposing. The studies by Grant et al. suggest that multi-target drugs can help in bridging this weak link in the complex biosystems caused due to their alternative networks [16]. Certain examples of multi-targeting drugs are dicoumarol which inhibits vitamin K epoxide reductase (VKOR) as well NAD(P)H quinone oxidoreductase 1 (NQO1) [17], Gleevec (Imatinib) initially used for inhibiting Bcr-Abl fusion gene which later was shown to inhibit other kinases such as c-kit and PDGFR as well [18], and letermovir [19–23] which acts on three protein targets and cures prophylaxis of cytomegalovirus (CMV) infection. Some drugs like asenapine, dronedarone, iloperidone, pazopanib, and milnacipran have 20, 18, 11, 10, and 9 protein targets, respectively [24]. Other examples are of complex disorders and diseases such as cancer, depression, diabetes, and infectious diseases, etc. which showcase the success of multi-targeted agents in small molecule therapy [16, 25]. Polypharmacology of certain natural products like curcumin, berberine, baicalein, and flavonoids has also been studied [26–32].

For drug discovery approach, in general, efforts are crucially dependent on identifying compounds that are able to inhibit the target proteins [33]. The structure based drug discovery approaches mainly work on the ideology of active site inhibition [34–36]. Thus, the active site similarity and the pattern of residues found in multiple drug targets can greatly influence the probability of a ligand to be a multiple target binder [37–40]. Targeting multiple protein targets is beneficial in terms of saving economic costs and time involved in phase trials and marketability [14, 26, 41–46]. Our work focuses on development of a protocol with advanced and refined techniques which convert the uni-dimensional one-target one-drug approach to a multi-dimensional many-target one-drug approach.

2 Methods

Multi-Target Ligand Design (MTLD) is a web server that helps to identify common leads for any two protein targets. The basic principle of identifying common small molecules inhibiting multiple targets remains the active site similarity and common interactive residues in their binding pockets. The protocol is based on the utilization of already established and validated software tools which are harnessed here to provide a set of common ligands for multiple proteins. A schematic representation of the methodology is provided in Fig. 1.

To predict common ligand molecules for two protein targets, MTLD requires either the structure or sequence of two proteins. In the first step, if the structure is unavailable, then it is predicted for each protein using BhageerathH+ methodology [47–50]. BhageerathH+ is a leading state-of-the-art software for prediction of tertiary structures of proteins. It is ranked as one of the leading servers for tertiary structure prediction in the biennial Critical Assessment of Protein Structure prediction (CASP) experiment. This is followed by model evaluation using ProtSAV [51] meta-server. The end result is a combined evaluation of the models predicted by BhageerathH+. The best model is utilized for MTLD methodology. Alternatively, user can input the known structures of the two

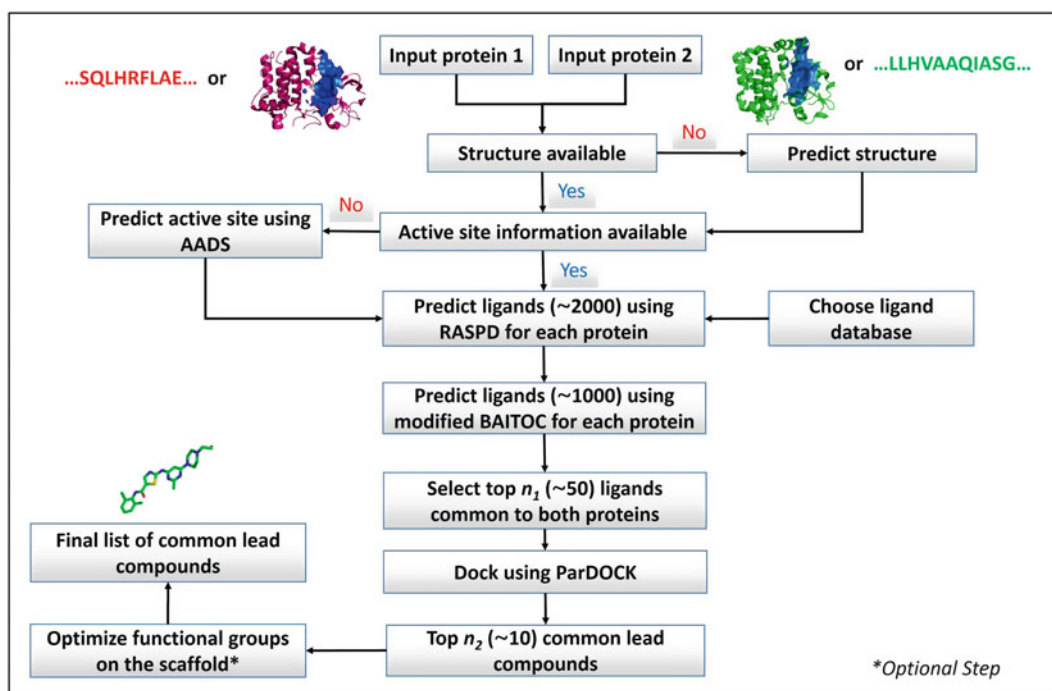


Fig. 1 Methodology for identifying lead molecules common to both input protein targets

protein targets. The active site can be either user defined in the protein structure or is predicted for each protein in step two using the in-house (Automated Active site detection, Docking and Scoring) AADS [52] methodology. AADS utilizes physico-chemical properties of the residues in cavities of the protein to predict biologically relevant binding pockets in a given protein. It has an accuracy of 67% in top 1 and 100% in top 10. In the third step, the protocol utilizes RASPD methodology on each protein to predict potentially viable binders [53]. It utilizes a quantitative structure–property relationship (QSPR) type equation to predict the binding affinity between a protein–ligand complex using 1D and 2D descriptors without actual docking. The prediction is based upon physico-chemical descriptors like hydrogen bond donors and acceptors, logP, molar refractive index, volume and wiener index of both the active site and the ligand. The methodology is capable of scanning million molecule library of small organic compounds under 10 min. User can choose from any of the following four databases: (1) Million molecule database, which is a collection of over one million compounds from ZINC database [54]; (2) Natural compound library comprising 0.2 million compounds; (3) National cancer database which is a collection of compounds from NCI [55] and finally (4) FDA approved drugs [46, 56, 57]. In the fourth step, the compounds selected through RASPD are re-evaluated using BAITOC methodology. This methodology selects ligands based on the complementarity of distances between pairs of hydrogen bond donors and acceptors, ratios of volumes of active sites with the ligands of interest and ratio of nonpolar surface areas of the ligands and the active sites of the proteins of interest (<http://www.scfbio-iitd.res.in/software/drugdesign/baitocnew.jsp>). This is followed by selection of common ligands between the proteins. The top 50 ligands based on calculated binding affinities with both proteins are selected. The selected ligands are then subjected to atomic-level docking and scoring using ParDOCK [58] methodology. The methodology uses Monte Carlo technique to predict the best pose of a ligand in the active site. The resulting structures are energy minimized using steepest descent and conjugate gradient minimization techniques. The poses are then evaluated for binding affinity using BAPPL [59] scoring methodology. Top 10 compounds showing best affinity to both chosen targets are then reported. Optionally the resulting ligands can be optimized for better binding using ligand optimization step. In this step, functional groups are added in place of hydrogen atoms while retaining the scaffold and re-evaluated for binding affinity. Ligands showing better binding affinity compared to original input ligands with both proteins are selected. In cases where optimization is unable to find better binding ligands, original unmodified ligands are reported.

3 Results and Discussion

Here we report the development of a multi-target ligand design protocol. The protocol works on predicting a list of common ligands against both the input protein targets. To assess the protocol, we validated it on five well-known cases of multi-target FDA approved drugs. The protein targets in the case studies vary in active site similarities and were scanned against available FDA approved drug library, thus, testing the methodology. The experimental data of these systems along with their predicted binding affinities are shown in Table 1.

In the first case study two proteins maltase-glucoamylase and lysosomal alpha-glucosidase were considered to identify the common FDA approved drugs against them. The protocol presented here identified MIG as one of the common hits among the top 10 molecules which is experimentally known to work against both these protein targets [64, 65]. Miglitol is known as an antihyperglycemic drug which results in reversible inhibition of membrane-bound intestinal α -glucoside hydrolase enzymes called alpha-glucosidases. In diabetic patients, primarily diabetes mellitus type 2,

Table 1
Experimental and computational binding energy values for various known multi-targets

Ligand name [57]	Disease	Protein target [60]	PDB ID [60]	Predicted binding energy [58]	Experimental binding energy [61]	Rank in the hit list	Active site similarity ^a [62, 63]
Miglitol (DB00491)	Diabetes mellitus type 2	Maltase-glucoamylase	3L4W	-4.22	-6.00	2	0.93
		Lysosomal alpha-glucosidase	5NN6	-4.16	N.A.		
Progesterone (DB00396)	Progesterone hormone deficiency	Mineralocorticoid receptor	2AA5	-8.16	N.A.	4	0.88
		Ancestral steroid receptor 2	4LTW	-8.87	N.A.		
Crizotinib (DB08865)	Non-small cell lung cancer	Hepatocyte growth factor receptor	2WGJ	-9.80	-8.70	7	0.87
		ALK tyrosine kinase receptor	4ANQ	-7.89	-6.22		
Imatinib (DB00619)	Chronic myelogenous leukemia and gastrointestinal stromal tumors	Mast/stem cell growth factor receptor	1T46	-12.19	-6.43	7	0.69
		vWF glycoprotein	3PYY	-6.81	N.A.		
Dasatinib (DB01254)	Chronic myelogenous leukemia	DDR1 and DDR2 receptor tyrosine kinases	5BVV	-9.35	-8.82	9	0.57
		Ephrin A2 (EphA2) receptor protein kinase	519Y	-8.85	-8.55		

N.A. not available

Predicted binding energy reported is the result after atomic-level docking (in kcal/mol). Experimental energy (kcal/mol) is adapted from PDBbind-CN Database [61]. Active site similarities are calculated using PocketMatch [62, 63]

^aActive site similarity scores near 0 imply no similarity and near 1 imply complete similarity. Pockets are considered very similar if they have scores of 0.8 or greater. Pockets are considered somewhat similar if they have scores of 0.6 or greater

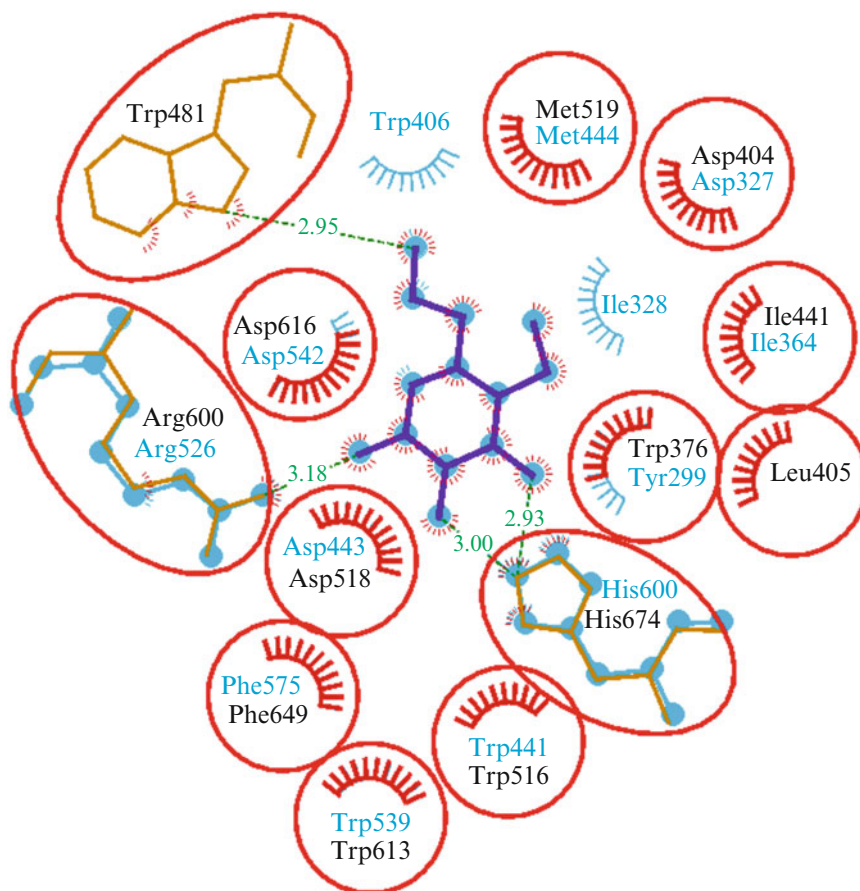


Fig. 2 The interaction patterns of active site residues of proteins maltase-glucoamylase and lysosomal alpha-glucosidase with ligand miglitol (ball and stick model representation shown in purple color) are illustrated in which nonpolar and polar residues are labelled black in the former case and in deep-skyblue in the latter case. Hydrogen bonds are shown in green color. The 2D interaction diagrams are made using LigPlot [66]

this enzyme inhibition results in delayed glucose absorption and lowering of postprandial hyperglycemia. The active sites of both the protein targets chosen had almost 93% similarity as shown in Fig. 2.

In the second case study, the common ligands against mineralocorticoid receptor and ancestral steroid receptor 2 (Table 1) were considered which showed progesterone as one of the best common identified leads using our protocol. Progesterone, a naturally occurring progestin, is reported to bind to the mineralocorticoid receptor and ancestral steroid receptor 2 [67, 68] and its levels dictate the release of eggs, thus, maintaining sexual fertility. The active sites of both the protein targets have almost 88% similarity as shown in Fig. 3.

In the third case study we considered hepatocyte growth factor receptor and ALK tyrosine kinase receptors. The protocol identified crizotinib as the common FDA approved drug against the two

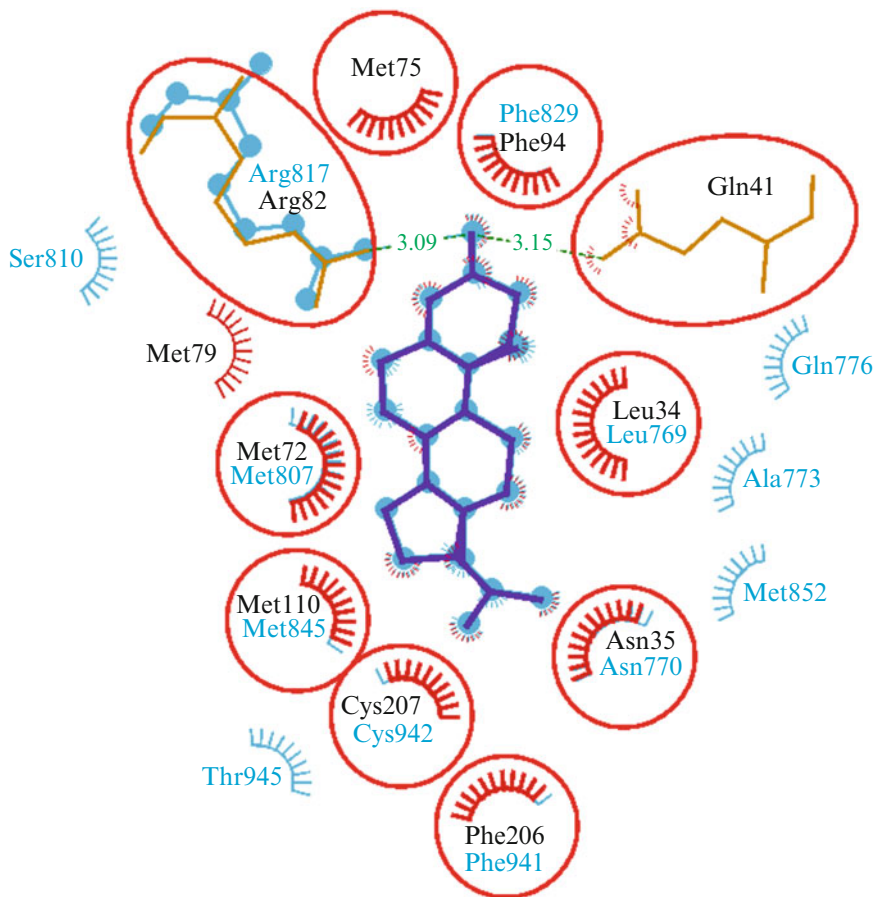


Fig. 3 The interaction patterns of active site residues of proteins mineralocorticoid receptor and ancestral steroid receptor 2 with ligand progesterone (ball and stick model representation shown in purple color) are illustrated in which nonpolar and polar residues are labelled black in the former case and in deep-skyblue in the latter case. Hydrogen bonds are shown in green color

proteins. It was ranked seventh among the top 10 common identified hits among all the FDA library of compounds as shown in Table 1. Crizotinib (DB08865) is a well-known tyrosine kinase receptor inhibitor. Experimentally it is known to inhibit both hepatocyte growth factor receptor (HGFR, c-MET) and ALK tyrosine kinase [69, 70]. The inhibition of tyrosine kinase receptors by Crizotinib ultimately results in decreased proliferation of cells that carry the genetic mutation and tumor survivability, thus making it multi-target anti-cancer drug. The active sites of both the protein targets have almost 87% similarity as shown in Fig. 4.

In the fourth case study we considered two protein targets, viz. mast/stem cell growth factor receptor and vWF glycoprotein. The active sites of both the protein targets have almost 69% similarity as shown in Fig. 5. Imatinib was identified as common ligand

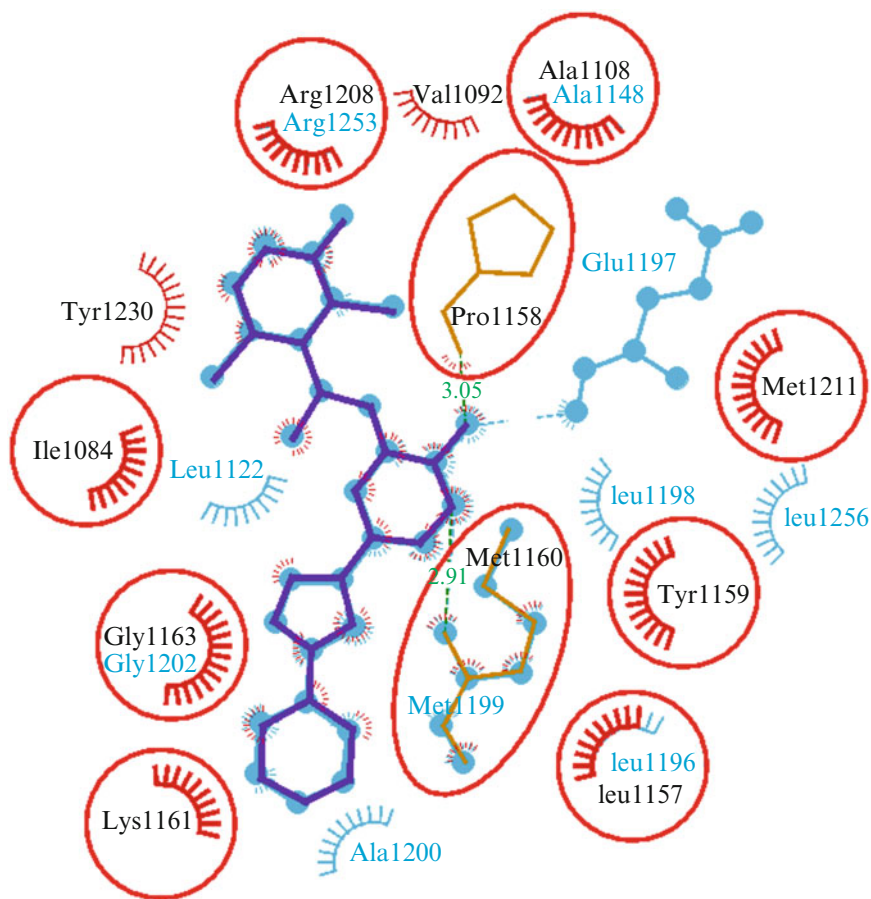


Fig. 4 The interaction patterns of active site residues of proteins hepatocyte growth factor receptor, ALK tyrosine kinase receptor, and ligand crizotinib (ball and stick model representation shown in purple color) are illustrated in which nonpolar and polar residues are labelled black in the former case and deep-skyblue in latter case. Hydrogen bonds are shown in green color and deep-skyblue color for HFGR and ALK receptors, respectively

molecule in the final output list of the protocol (Table 1). Imatinib (DB00619) mesylate is a known protein-tyrosine kinase inhibitor that inhibits the Bcr-Abl tyrosine kinase, the constitutive abnormal tyrosine kinase created by the Philadelphia chromosome abnormality in chronic myeloid leukemia (CML) [71, 72].

Finally in the fifth case study we considered DDR1 and DDR2 receptor tyrosine kinase and Ephrin A2 (EphA2) receptor protein kinase. The active sites of the two protein targets differ significantly as shown in Fig. 6 with 57% binding pocket overlap. The protocol was able to identify Dasatinib among the best binders via screening and atomic level docking using the protocol as shown in Table 1. Dasatinib is an anti-cancer drug which was approved for use against kinases. Experimentally dasatinib is known to bind DDR1 and DDR2, Ephrins, and GFR kinases [73, 74].

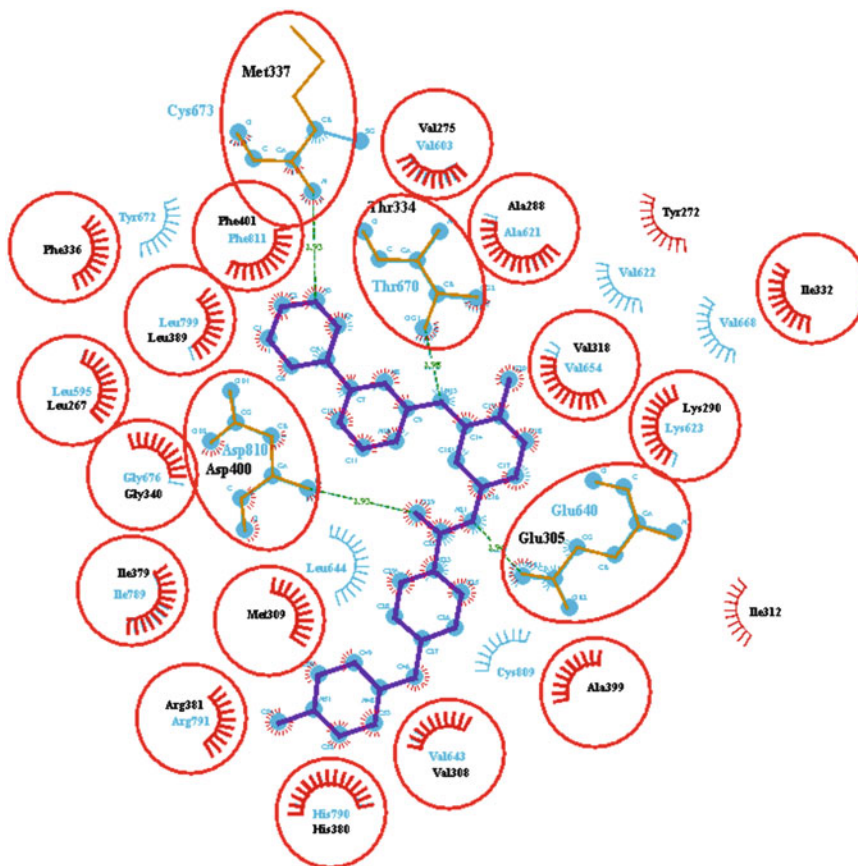


Fig. 5 The interaction patterns of active site residues of proteins Mast/stem cell growth factor receptor Kit and vWF glycoprotein with ligand imatinib (ball and stick model representation shown in purple color) are illustrated in which nonpolar and polar residues are labelled black in the former case and deep-skyblue in latter case. Hydrogen bonds are shown in green color

The active site similarity of the protein targets considered for case study varies from 0.57 to 0.93. As it can be inferred from Table 1, the methodology was able to capture the known common FDA drugs as one of the top molecules for the selected two proteins in each case study.

4 Web Interface

A free web-server implementing the protocol presented is available at the following url: <http://www.scfbio-iitd.res.in/multitarget/>.

The starting point of the MTLTD methodology are two target proteins whose sequences and/or their tertiary structure information are known. The web-server provides options to input either sequences or structures (Fig. 7). In case of sequences, tertiary

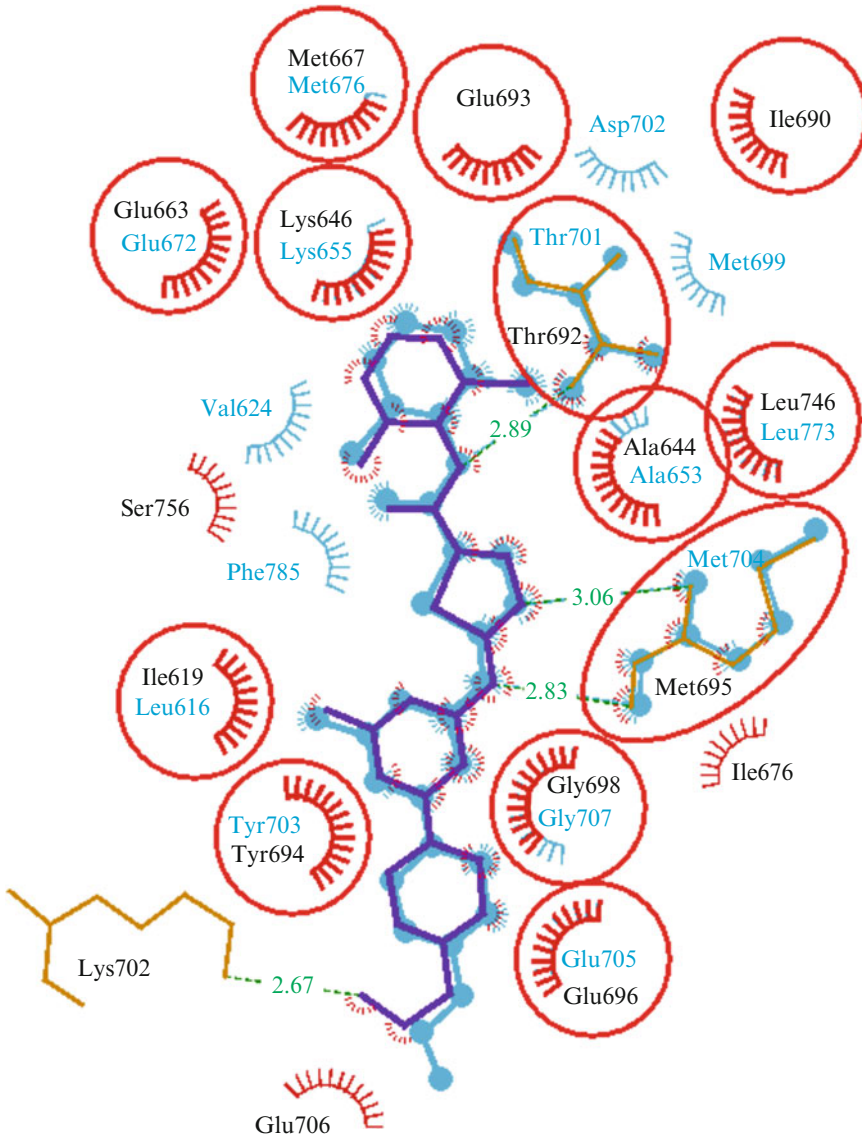


Fig. 6 The interaction patterns of active site residues of proteins DDR1 and DDR2 receptor tyrosine kinases and Ephrin A2 (EphA2) receptor protein kinase with ligand dasatinib (ball and stick model representation shown in purple color) are illustrated in which nonpolar and polar residues are labelled black in the former case and deep-skyblue in latter case. Hydrogen bonds are shown in green color

structure and active site are predicted using Bhageerath-H⁺ and AADS, respectively. The methodology then follows the flowchart given in Fig. 1. User can control various aspects of the methodology using optional parameters for docking and database selection. The job can be submitted by clicking on “find common hits” button. Once the job is submitted a unique job identifier number is generated which user can use to access the results and the status

a

The screenshot shows the MTL D web-server interface for sequence input. At the top, there is a logo on the left and the title "Multi Target Ligand Design" in large blue letters, with the subtitle "Supercomputing Facility for Bioinformatics & Computational Biology, IIT Delhi" below it. A navigation bar contains "Home", "About", "Contact", and "SCFBio". Below the title, there are two radio buttons: "Sequence" (selected) and "Structure". On the right, a green box displays "Job ID:-1518755203" and "Use this Job ID to retrieve the result later", with a "Search Results" button and a text input field "Enter your job Id. Press enter ke". The main area has two large text input fields labeled "Input 1" and "Input 2". Below these are "Options" for "Database" (set to "Million molecule"), "Docking Options" (including "Partial Charge", "Minimization Steps", "Dock", and "Scaffold Optimization"), and an "Email ID.(Optional)" field. A button labeled "Identify Common hits" is at the bottom. The footer contains copyright information: "©2018-2019, Prof B. Jayaram & Co-workers. All rights reserved. Web tool design by Mangreet Singh."

b

The screenshot shows the MTL D web-server interface for structure input. The layout is identical to screenshot (a), but the "Structure" radio button is selected. The "Input 1" and "Input 2" fields are replaced by "Upload file in .pdb format" buttons with "Click to upload" icons. The "Options" and "Identify Common hits" button remain the same. The footer is also identical to screenshot (a).

Fig. 7 Screenshots of the Multi-Target Ligand Design (MTLD) web-server available at SCFBio for identifying common lead molecules for two druggable protein targets. Webpages displaying the options provided to input (a) sequences and (b) structures for both protein targets are shown

of the job later. If an (optional) email address has been specified in the submission form, the result page URL is also sent to the user by email once the job is completed. The result web page features a JSmol applet for the visualization of the predicted binding poses, structure of the common small molecules identified, and the input protein structure(s) available at the web browser (Fig. 8).

The structure(s) of the molecules found as common leads for both input protein targets along with their best binding docked pose(s) are also provided at the completion of the job for both visualization and for downloading. In cases where the input protein target(s) was sequence, the tertiary structure of the input protein is also available for visualization and downloading at the result webpage. The result page also displays all the input parameters selected at the time of submission for user reference. The user has the independence for submitting multiple jobs to extend the work on multiple targets.

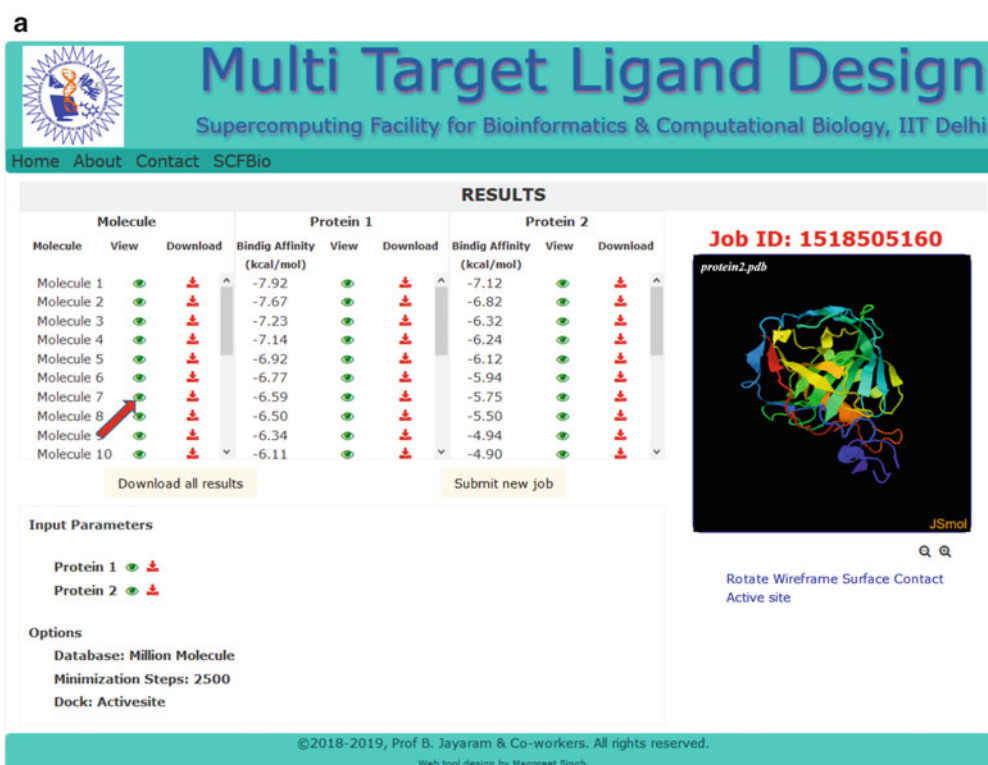



Fig. 8 The result webpage of MTLD. Red arrow highlights the compound selected for visualization. (a) Compound identified among the million database shown as seventh best common hit is displayed in the visualization box. (b) Protein-Ligand docked best pose is displayed in the cartoon (Protein 2) with ball and stick representation (Molecule1). (c) Visualization of 3D structure generated for input Protein 2 sequence. User can also download all the structures in pdf format from the download links provided. The webpage also shows the input parameters used for identifying common hits

b

 **Multi Target Ligand Design**
Supercomputing Facility for Bioinformatics & Computational Biology, IIT Delhi

Home About Contact SCFBio


RESULTS

Molecule			Protein 1			Protein 2		
Molecule	View	Download	Bindig Affinity (kcal/mol)	View	Download	Bindig Affinity (kcal/mol)	View	Download
Molecule 1			-7.92			-7.12		
Molecule 2			-7.67			-6.82		
Molecule 3			-7.23			-6.32		
Molecule 4			-7.14			-6.24		
Molecule 5			-6.92			-6.12		
Molecule 6			-6.77			-5.94		
Molecule 7			-6.59			-5.75		
Molecule 8			-6.50			-5.50		
Molecule 9			-6.34			-4.94		
Molecule 10			-6.11			-4.90		

Download all results Submit new job

Job ID: 1518505160

protein2.pdb



JSmol

Rotate Wireframe Surface Contact Active site

Input Parameters

Protein 1 |

Protein 2 |

Options


Database: Million Molecule

Minimization Steps: 2500

Dock: Activesite

©2018-2019, Prof B. Jayaram & Co-workers. All rights reserved.
Web tool design by Manpreet Singh.

c

 **Multi Target Ligand Design**
Supercomputing Facility for Bioinformatics & Computational Biology, IIT Delhi

Home About Contact SCFBio

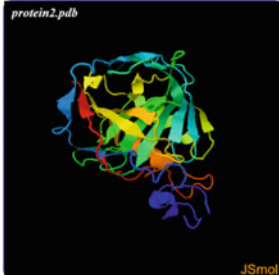
RESULTS

Molecule			Protein 1			Protein 2		
Molecule	View	Download	Bindig Affinity (kcal/mol)	View	Download	Bindig Affinity (kcal/mol)	View	Download
Molecule 1			-7.92			-7.12		
Molecule 2			-7.67			-6.82		
Molecule 3			-7.23			-6.32		
Molecule 4			-7.14			-6.24		
Molecule 5			-6.92			-6.12		
Molecule 6			-6.77			-5.94		
Molecule 7			-6.59			-5.75		
Molecule 8			-6.50			-5.50		
Molecule 9			-6.34			-4.94		
Molecule 10			-6.11			-4.90		

Download all results Submit new job

Job ID: 1518505160

protein2.pdb



JSmol

Rotate Wireframe Surface Contact Active site

Input Parameters

Protein 1 |

Protein 2 |

Options

Database: Million Molecule

Minimization Steps: 2500

Dock: Activesite

©2018-2019, Prof B. Jayaram & Co-workers. All rights reserved.
Web tool design by Manpreet Singh.

Fig. 8 (continued)

In addition to the automated target protein structure prediction or preparation, screening, docking and scoring, alternative sets of parameters are also provided along with sample input files at the web suite. All calculations are run on the server side, thus avoiding any computational power requirements from the user.

5 Scope for Further Improvement

Multi-target based virtual screening methods have been explored extensively in recent years with promising potential towards identifying selective multi-target agents. To refine the protocol, it is possible to introduce more comprehensive structural and physico-chemical features of selective multi-target agents. One such example would be to classify the binding site profiles of different multi-targets to expedite the identification of selective multi-target agents and active compounds. With recent and continuing efforts in protein structure prediction and active site prediction, results are expected to be more accurate and reliable. The methodology currently employs a single level addition of functional groups. To make the list of common hit molecules more selective towards the targeted multiple proteins, a multi-level scaffold extension protocol can be developed. Synthesis constraints can be incorporated to the scaffold extension methodology so that the hits are more viable to chemical synthesis.

6 Conclusions

The MTLD web server aims to provide the scientific community with a free and user-friendly, state-of-the-art software suite to identify common leads for multiple protein targets. The automated setup of preparation of the target input protein(s) and ligand structures, the availability of different docking parameter presets, and the convenient visualization and analysis of predictions make it easy to use. The methodology has been validated on five protein sets with known FDA drugs and has shown promising results. No such web-server similar to MTLD is reported so far. Thus this contribution fills a crucial void in the space of computer-aided drug discovery.

Acknowledgment

Funding from the Department of Biotechnology, Govt. of India, to SCFBio is gratefully acknowledged. A.J. and A.P. are Institute Fellows. R.B. is a DST INSPIRE Fellow.

Author contributions: B.J. conceived the project. A.J., R.B., A.P. carried out the computational development. All authors analyzed the results and wrote the manuscript. M.S. helped in web enabling of the server. All authors have given approval to the final version of the manuscript.

References

1. Winau F, Westphal O, Winau R (2004) Paul Ehrlich – in search of the magic bullet. *Microbes Infect* 6:786–789. <https://doi.org/10.1016/j.micinf.2004.04.003>
2. Koutsoukas A, Simms B, Kirchmair J et al (2011) From in silico target prediction to multi-target drug design: current databases, methods and applications. *J Proteome* 74:2554–2574. <https://doi.org/10.1016/j.jprot.2011.05.011>
3. DiMasi J (2001) New drug development in the United States from 1963 to 1999. *Clin Pharmacol Ther* 69:286–296. <https://doi.org/10.1067/mcp.2001.115132>
4. Adams CP, Brantner VV (2006) Estimating the cost of new drug development: is it really \$802 million? *Health Aff* 25:420–428. <https://doi.org/10.1377/hlthaff.25.2.420>
5. Ashburn TT, Thor KB (2004) Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov* 3:673–683. <https://doi.org/10.1038/nrd1468>
6. Hopkins AL (2008) Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol* 4:682–690. <https://doi.org/10.1038/nchembio.118>
7. Hopkins AL (2009) Drug discovery: predicting promiscuity. *Nature* 462:167–168. <https://doi.org/10.1038/462167a>
8. Apfel B, Blair JA, Gonzalez BZ et al (2008) NIH public access. *Nat Chem Biol* 4:691–699. <https://doi.org/10.1038/nchembio.117>. Targeted
9. Simon Z, Peragovics Á, Vigh-Smeller M et al (2012) Drug effect prediction by polypharmacology-based interaction profiling. *J Chem Inf Model* 52:134–145. <https://doi.org/10.1021/ci2002022>
10. Briansó F, Carrascosa MC, Oprea TI, Mestres J (2011) Cross-pharmacology analysis of G protein-coupled receptors. *Curr Top Med Chem* 11:1956–1963. <https://doi.org/10.2174/156802611796391285>
11. Paolini GV, Shapland RHB, Van Hoorn WP et al (2006) Global mapping of pharmacological space. *Nat Biotechnol* 24:805–815. <https://doi.org/10.1038/nbt1228>
12. Oprea TI, Mestres J (2012) Drug repurposing: far beyond new targets for old drugs. *AAPS J* 14:759–763. <https://doi.org/10.1208/s12248-012-9390-1>
13. Durrant JD, Amaro RE, Xie L et al (2010) A multidimensional strategy to detect polypharmacological targets in the absence of structural and sequence homology. *PLoS Comput Biol* 6:e1000648. <https://doi.org/10.1371/journal.pcbi.1000648>
14. Boran ADW, Iyengar R (2010) Systems approaches to polypharmacology and drug discovery. *Curr Opin Drug Discov Devel* 13:297–309. <https://doi.org/10.1126/scisignal.2001965.Introduction>
15. Reddy AS, Zhang S (2013) Polypharmacology: drug discovery for the future. *Expert Rev Clin Pharmacol* 6:41–47. <https://doi.org/10.1586/ecp.12.74>
16. Zimmermann GR, Lehár J, Keith CT (2007) Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug Discov Today* 12:34–42. <https://doi.org/10.1016/j.drudis.2006.11.008>
17. Timson D (2017) Dicoumarol: a drug which hits at least two very different targets in vitamin K metabolism. *Curr Drug Targets* 18:500–510. <https://doi.org/10.2174/1389450116666150722141906>
18. Giordano S, Petrelli A (2008) From single- to multi-target drugs in cancer therapy: when aspecificity becomes an advantage. *Curr Med Chem* 15:422–432. <https://doi.org/10.2174/092986708783503212>
19. Kropf D, Scheuenpflug J, Erb-Zohar K et al (2017) Pharmacokinetics and safety of letermovir, a novel anti-human cytomegalovirus drug, in patients with renal impairment. *Br J Clin Pharmacol* 83:1944–1953. <https://doi.org/10.1111/bcp.13292>
20. Goldner T, Hewlett G, Ettischer N et al (2011) The novel anticytomegalovirus compound AIC246 (Letermovir) inhibits human

- cytomegalovirus replication through a specific antiviral mechanism that involves the viral terminase. *J Virol* 85:10884–10893. <https://doi.org/10.1128/JVI.05265-11>
21. Razonable R, Melendez D (2015) Letermovir and inhibitors of the terminase complex: a promising new class of investigational antiviral drugs against human cytomegalovirus. *Infect Drug Resist* 8:269. <https://doi.org/10.2147/IDR.S79131>
 22. Chou S (2017) A third component of the human cytomegalovirus terminase complex is involved in letermovir resistance. *Antivir Res* 148:1–4. <https://doi.org/10.1016/j.antiviral.2017.10.019>
 23. Neuber S, Wagner K, Goldner T et al (2017) Mutual interplay between the human cytomegalovirus terminase subunits pUL51, pUL56, and pUL89 promotes terminase complex formation. *J Virol* 91:e02384–e02316. <https://doi.org/10.1128/JVI.02384-16>
 24. Lin H-H, Zhang L-L, Yan R et al (2017) Network analysis of drug–target interactions: a study on FDA-approved new molecular entities between 2000 to 2015. *Sci Rep* 7:12230. <https://doi.org/10.1038/s41598-017-12061-8>
 25. Van der Schyf CJ (2011) The use of multi-target drugs in the treatment of neurodegenerative diseases. *Expert Rev Clin Pharmacol* 4:293–298. <https://doi.org/10.1586/ecp.11.13>
 26. Das T, Sa G, Saha B, Das K (2010) Multifocal signal modulation therapy of cancer: ancient weapon, modern targets. *Mol Cell Biochem* 336:85–95. <https://doi.org/10.1007/s11010-009-0269-0>
 27. Gupta SC, Prasad S, Kim JH et al (2011) Multitargeting by curcumin as revealed by molecular interaction studies. *Nat Prod Rep* 28:1937. <https://doi.org/10.1039/c1np00051a>
 28. Gupta SC, Patchva S, Koh W, Aggarwal BB (2012) Discovery of curcumin, a component of golden spice, and its miraculous biological activities. *Clin Exp Pharmacol Physiol* 39:283–299. <https://doi.org/10.1111/j.1440-1681.2011.05648.x>
 29. Tan W, Lu J, Huang M et al (2011) Anti-cancer natural products isolated from Chinese medicinal herbs. *Chin Med* 6:1–15. <https://doi.org/10.1186/1749-8546-6-27>
 30. Gupta SC, Kim JH, Prasad S, Aggarwal BB (2010) Regulation of survival, proliferation, invasion, angiogenesis, and metastasis of tumor cells through modulation of inflammatory pathways by nutraceuticals. *Cancer Metastasis Rev* 29:405–434. <https://doi.org/10.1007/s10555-010-9235-2>
 31. Hoda N, Naz H, Jameel E et al (2016) Curcumin specifically binds to the human calcium–calmodulin-dependent protein kinase IV: fluorescence and molecular dynamics simulation studies. *J Biomol Struct Dyn* 34:572–584. <https://doi.org/10.1080/07391102.2015.1046934>
 32. Srinivas NR (2010) Baicalin, an emerging multi-therapeutic agent: pharmacodynamics, pharmacokinetics, and considerations from drug development perspectives. *Xenobiotica* 40:357–367. <https://doi.org/10.3109/00498251003663724>
 33. Jayaram B, Singh T, Mukherjee G et al (2012) Sanjeevini: a freely accessible web-server for target directed lead molecule discovery. *BMC Bioinformatics* 13:S7. <https://doi.org/10.1186/1471-2105-13-S17-S7>
 34. Meng X-Y, Zhang H-X, Mezei M, Cui M (2011) Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided-Drug Des* 7:146–157. <https://doi.org/10.2174/157340911795677602>
 35. Holdeman R, Nehrt S, Strome S (1998) MES-2, a maternal protein essential for viability of the germline in *Caenorhabditis elegans*, is homologous to a *Drosophila* Polycomb group protein. *Development* 125:2457–2467
 36. Lionta E, Spyrou G, Vassilatis D, Cournia Z (2014) Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem* 14:1923–1938. <https://doi.org/10.2174/1568026614666140929124445>
 37. Konc J, Hodošček M, Ogrizek M et al (2013) Structure-based function prediction of uncharacterized protein using binding sites comparison. *PLoS Comput Biol* 9:e1003341. <https://doi.org/10.1371/journal.pcbi.1003341>
 38. Kumar P, Kaalia R, Srinivasan A, Ghosh I (2018) Multiple target-based pharmacophore design from active site structures. *SAR QSAR Environ Res* 29:1–19. <https://doi.org/10.1080/1062936X.2017.1401555>
 39. Ploemen JHTM, Johnson WW, Jespersen S et al (1994) Active-site tyrosyl residues are targets in the irreversible inhibition of a class Mu glutathione transferase by 2-(S-glutathionyl)-3,5,6-trichloro-1,4-benzoquinone. *J Biol Chem* 269:26890–26897
 40. Ramsay RR, Majekova M, Medina M, Valoti M (2016) Key targets for multi-target ligands designed to combat neurodegeneration. *Front*

- Neurosci 10. <https://doi.org/10.3389/fnins.2016.00375>
41. Csermely P, Agoston V, Pongor S (2005) The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol Sci* 26:178–182. <https://doi.org/10.1016/j.tips.2005.02.007>
 42. Puls LN, Eadens M, Messersmith W (2011) Current status of Src inhibitors in solid tumor malignancies. *Oncologist* 16:566–578. <https://doi.org/10.1634/theoncologist.2010-0408>
 43. Stella GM, Luisetti M, Inghilleri S et al (2012) Targeting EGFR in non-small-cell lung cancer: lessons, experiences, strategies. *Respir Med* 106:173–183. <https://doi.org/10.1016/j.rmed.2011.10.015>
 44. Yildirim MA, Goh KI, Cusick ME et al (2007) Drug-target network. *Nat Biotechnol* 25:1119–1126. <https://doi.org/10.1038/nbt1338>
 45. Zhang W, Pei J, Lai L (2017) Computational multitarget drug design. *J Chem Inf Model* 57:403–412. <https://doi.org/10.1021/acs.jcim.6b00491>
 46. Knox C, Law V, Jewison T et al (2011) DrugBank 3.0: a comprehensive resource for “Omics” research on drugs. *Nucleic Acids Res* 39:D1035–D1041. <https://doi.org/10.1093/nar/gkq1126>
 47. Jayaram B, Bhushan K, Shenoy SR et al (2006) Bhageerath: an energy based web enabled computer software suite for limiting the search space of tertiary structures of small globular proteins. *Nucleic Acids Res* 34:6195–6204. <https://doi.org/10.1093/nar/gkl789>
 48. Jayaram B, Dhingra P, Mishra A et al (2014) Bhageerath-H: a homology/ab initio hybrid server for predicting tertiary structures of monomeric soluble proteins. *BMC Bioinformatics* 15:S7. <https://doi.org/10.1186/1471-2105-15-S16-S7>
 49. DasGupta D, Kaushik R, Jayaram B (2015) From Ramachandran maps to tertiary structures of proteins. *J Phys Chem B* 119:11136–11145. <https://doi.org/10.1021/acs.jpcc.5b02999>
 50. Kaushik R, Singh A, Jayaram B (2018) Where informatics lags chemistry leads. *Biochemistry* 57:503–506. <https://doi.org/10.1021/acs.biochem.7b01073>
 51. Singh A, Kaushik R, Mishra A et al (2016) ProTSAV: a protein tertiary structure analysis and validation server. *Biochim Biophys Acta* 1864:11–19. <https://doi.org/10.1016/j.bbapap.2015.10.004>
 52. Singh T, Biswas D, Jayaram B (2011) AADS – an automated active site identification, docking, and scoring protocol for protein targets based on physicochemical descriptors. *J Chem Inf Model* 51:2515–2527. <https://doi.org/10.1021/ci200193z>
 53. Mukherjee G, Jayaram B (2013) A rapid identification of hit molecules for target proteins via physico-chemical descriptors. *Phys Chem Chem Phys* 15:9107. <https://doi.org/10.1039/c3cp44697b>
 54. Irwin JJ, Shoichet BK (2005) ZINC – a free database of commercially available compounds for virtual screening. *J Chem Inf Model* 45:177–182. <https://doi.org/10.1021/ci049714+>
 55. NCI Database Download Page. <http://cactus.nci.nih.gov/ncidb2/download.html>
 56. Wishart DS, Knox C, Guo AC et al (2008) DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res* 36:D901–D906. <https://doi.org/10.1093/nar/gkm958>
 57. Wishart DS, Feunang YD, Guo AC et al (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 46:D1074–D1082. <https://doi.org/10.1093/nar/gkx1037>
 58. Gupta A, Gandhimathi A, Sharma P, Jayaram B (2007) ParDOCK: an all atom energy based Monte Carlo docking protocol for protein-ligand complexes. *Protein Pept Lett* 14:632–646. <https://doi.org/10.2174/092986607781483831>
 59. Jain T, Jayaram B (2005) An all atom energy based computational protocol for predicting binding affinities of protein-ligand complexes. *FEBS Lett* 579:6659–6666. <https://doi.org/10.1016/j.febslet.2005.10.031>
 60. Berman HM (2000) The Protein Data Bank. *Nucleic Acids Res* 28:235–242. <https://doi.org/10.1093/nar/28.1.235>
 61. Liu Z, Su M, Han L et al (2017) Forging the basis for developing protein–ligand interaction scoring functions. *Acc Chem Res* 50:302–309. <https://doi.org/10.1021/acs.accounts.6b00491>
 62. Yeturu K, Chandra N (2008) PocketMatch: a new algorithm to compare binding sites in protein structures. *BMC Bioinformatics* 9:543. <https://doi.org/10.1186/1471-2105-9-543>
 63. Nagarajan D, Chandra N (2013) PocketMatch (version 2.0): a parallel algorithm for the detection of structural similarities between protein ligand binding-sites. In: 2013 national conference on parallel computing technologies

- (PARCOMPTECH). IEEE, pp 1–6. <https://doi.org/10.1109/ParCompTech.2013.6621397>
64. Sim L, Jayakanthan K, Mohan S et al (2010) New glucosidase inhibitors from an ayurvedic herbal treatment for type 2 diabetes: structures and inhibition of human intestinal maltase-glucoamylase with compounds from *Salacia reticulata*. *Biochemistry* 49:443–451. <https://doi.org/10.1021/bi9016457>
 65. Roig-Zamboni V, Cobucci-Ponzano B, Iacono R et al (2017) Structure of human lysosomal acid α -glucosidase—a guide for the treatment of Pompe disease. *Nat Commun* 8. <https://doi.org/10.1038/s41467-017-01263-3>
 66. Wallace AC, Laskowski RA, Thornton JM (1995) Ligplot – a program to generate schematic diagrams of protein ligand interactions. *Protein Eng* 8:127–134. <https://doi.org/10.1093/protein/8.2.127>
 67. Bledsoe RK, Madauss KP, Holt JA et al (2005) A ligand-mediated hydrogen bond network required for the activation of the mineralocorticoid receptor. *J Biol Chem* 280:31283–31293. <https://doi.org/10.1074/jbc.M504098200>
 68. Colucci JK, Ortlund EA (2013) X-ray crystal structure of the ancestral 3-ketosteroid receptor-progesterone-mifepristone complex shows mifepristone bound at the coactivator binding interface. *PLoS One* 8:1–12. <https://doi.org/10.1371/journal.pone.0080761>
 69. Cui JJ, Tran-Dubé M, Shen H et al (2011) Structure based drug design of crizotinib (PF-02341066), a potent and selective dual inhibitor of mesenchymal-epithelial transition factor (c-MET) kinase and anaplastic lymphoma kinase (ALK). *J Med Chem* 54:6342–6363. <https://doi.org/10.1021/jm2007613>
 70. Huang Q, Johnson TW, Bailey S et al (2014) Design of potent and selective inhibitors to overcome clinical anaplastic lymphoma kinase mutations resistant to crizotinib. *J Med Chem* 57:1170–1187. <https://doi.org/10.1021/jm401805h>
 71. Mol CD, Dougan DR, Schneider TR et al (2004) Structural basis for the autoinhibition and STI-571 inhibition of c-kit tyrosine kinase. *J Biol Chem* 279:31655–31663. <https://doi.org/10.1074/jbc.M403319200>
 72. Zhou M, Dong X, Baldauf C et al (2011) A novel calcium-binding site of von Willebrand factor A2 domain regulates its cleavage by ADAMTS13. *Blood* 117:4623–4631. <https://doi.org/10.1182/blood-2010-11-321596>
 73. Murray CW, Berdini V, Buck IM et al (2015) Fragment-based discovery of potent and selective DDR1/2 inhibitors. *ACS Med Chem Lett* 6:798–803. <https://doi.org/10.1021/acsmchemlett.5b00143>
 74. Heinzlmeir S, Kudlinzki D, Sreeramulu S et al (2016) Chemical proteomics and structural biology define EPHA2 inhibition by clinical kinase drugs. *ACS Chem Biol* 11:3400–3411. <https://doi.org/10.1021/acscmbio.6b00709>



Computational Method for Prediction of Targets for Breast Cancer Using siRNA Approach

Atul Tyagi, Mukti N. Mishra, and Ashok Sharma

Abstract

The increasing incident of breast cancer, which is a leading cause of women's death in both developed and developing countries, demands the development of novel and efficient therapies. One of the major challenges is to design drugs that can specifically target the genes or proteins responsible for breast cancer, as gene and chemotherapy both are suffering from the drug specificity issues. Several recent studies have highlighted the potential of RNA interference (RNAi)-mediated targeted silencing of breast oncogenes, which can be exploited to develop cancer cell-/target-specific therapeutic molecules. However, one of the bottlenecks of RNAi-based gene therapy is to identify the RNAi sequences for efficient and targeted suppression of oncogenes. In this chapter, we discuss the development and application of a web-based database, BOSS (<http://bioinformatics.cimap.res.in/sharma/boss/index.php>), for selection of potential RNAi based on the sequences that have been used and validated for RNAi-mediated suppression of breast oncogenes. This database includes the latest information regarding used RNAi molecules that can be cost-effective and less time-consuming.

Keywords Breast cancer, Gene silencing, Mammary cancer, Oncogene, RNAi, shRNAs, siRNAs

1 Introduction

Breast cancer, which is the result of uncontrolled proliferation and differentiation of breast cells [1–10], is one of the major causes of women's death across the world. Mutations in the oncogenes result in gain or loss of function that may contribute to the uncontrolled proliferation and differentiation known as malignant phenotype. These mutations are generally the consequences of spontaneous mutations, environmental factors, viral infections, etc. Despite recent advances in early detection and [11] therapeutic strategies (chemotherapy, surgical, and radiation interventions), malignant breast cancer remains incurable. This is because of the toxicity and/or the lack of specificity of the current therapies and development of drug resistance by cancer cells. To overcome these issues, cancer cell-/target-specific efficient anticancer drugs are required. Since amplification or overexpression of breast oncogenes is the major mechanism through which oncogenes participate in the

cancer development [3], targeted suppression of these oncogenes can help to achieve the goal.

Nucleic acid therapy is based on: (1) introducing oligonucleotides or plasmids to express target genes for adding, correcting, or replacing the oncogenes in transformed cells and (2) RNAi-mediated targeted suppression of oncogenes. RNAi-mediated specific suppression of oncogenes is one of the most promising strategies for targeted gene knockdown. RNAi is a eukaryotic regulatory process for gene expression and antiviral defense. These small RNAs come under two distinct classes, small interfering RNAs (siRNAs) and microRNAs (miRNAs). The sizes of si- and miRNAs, which can vary between 19 and 24 nucleotides in length, are generated by processing of longer and complex double-stranded RNAs (dsRNAs) by ribonuclease III type DICER enzymes. Once generated, small RNA molecules base pairs with their complementary target RNA to direct the RNAi-mediated degradation.

Small interfering RNAs (siRNAs) have been achieved in mammalian systems following transfection of synthetic RNA molecules with overhanging 3' ends indicating that siRNA may be used as a powerful tool to block the expression of target genes specifically [12]. The factors that determine siRNA efficacy are thermodynamic stability of the duplex at the 5' antisense end, GC content, and the ability to form internal hairpins in the RISC (RNA-induced silencing complex), which promotes the degradation of target mRNA. Although siRNAs and short hairpin RNAs (shRNAs) have become a standard tool for targeted gene silencing, selection of siRNA/shRNA sequences for efficient and targeted suppression of oncogenes is still a time-consuming and laborious task. A siRNA/shRNA database can provide a virtual platform to compare the suppression efficiency of earlier reported siRNAs/shRNAs for designing of efficient siRNA.

1.1 Earlier siRNA Databases

In the past decade, a considerable number of databases and prediction servers have been developed to understand the features of siRNAs and to implement these features to predict potent siRNAs. Many databases for siRNA have been described in the literature. Some examples are RNAi (siRNA) libraries and their specificity (RNAiAtlas), database for ta-siRNA regulatory pathways (tasiRNAdb) [13], an innovative and comprehensive resource for identification of siRNA-mediated mechanisms in human-transcribed pseudogenes (pseudoMap) [14], an online resource to publish and query data from functional genomics high-throughput siRNA screening projects (HTS-DB) [15], a web-based tool for siRNA sequence design and analysis and an open-access siRNA database (siR) [16], HIV-specific siRNA (HIVsirDB) [8], human-specific siRNA database (HuSiDa) [17], siRNA data of mammalian RNAi experiments (siRecords) [9], and viral-specific siRNAs (VIRsiRNAdb) [18]. The VIRsiRNAdb database provides information on 1358 experimentally

validated siRNAs pertaining to 42 important human viruses. Many databases are also available in the literature like siRNAmoD, chemically modified siRNAs [19], and DSTHO (database of siRNAs targeted at human oncogenes) for targeting human oncogenes [7]. However, only the DSTHO database provides limited information on siRNAs for human oncogenes including few breast oncogenes. Although several (>100) siRNAs have been reported after development of DSTHO, there is no database available for the analysis of the latest breast oncogene siRNAs. Therefore, it is difficult for the researchers to search and analyze the data from the literature. In this chapter, we describe the details of the BOSS database, which exclusively includes the information for all the reported breast oncogenic specific siRNAs.

2 BOSS Database

RNA interference (RNAi) is an important technique for targeted gene silencing that may lead to promising novel therapeutic strategies for breast cancer. Therefore, identification of such molecules having high oncogene specificity is the need of an hour. Several exciting studies demonstrated the potentials of different siRNAs to inhibit the gene expression of breast oncogene and therapeutic molecules against cancer [10, 20–37]. Thus, it is important to collect and compile siRNAs in order to explore their potentiality against breast cancer (Fig. 1). Here, we present a database named as breast oncogenic specific siRNAs (BOSS, <http://bioinformatics>.

BOSS: A Database of Breast Oncogenic Specific siRNAs
CSIR-Central Institute of Medicinal and Aromatic Plants

Home Search Advance Search Browse Tools Download Online Submission Help Related links About us Contact

BOSS database, will provide the comprehensive details of the experimentally validated breast cancer specific breast cancer genes at one platform, to help researchers working in the field of siRNA based anticancer therapeutic development.

The RNA interference pathway

siRNA miRNA

ATP

miRISC

Endonucleolytic cleavage

Breast oncogene sustains cell growth and division

Cancer cell

Cell growth in normal gene

Normal cell

Data architecture

The BOSS database contains the following seventeen fields for each siRNAs entry: (1) BOSS id (2) pubmed id, (3) sequence, (4) siRNA name, (5) target gene, (6) GC content, (7) length of siRNA, (8) cell types, (9) year, (10) siRNA source (siRNA/siRNAs), (11) position of siRNA (siRNA/siRNAs), (12) Text objective, (13) Text method, (14) NCBI Accession no., (15) Biological inhibition, (16) Transfection Reagent and (17) Test time.

BOSS database provide three options

Search option: User can search in any field and display any field.
Advance search option: It allows the user to search more complex queries. **Browse option:** It assists user to search breast oncogenic siRNAs in BOSS DB.

Major tools at BOSS DB:

BOSS database has two tools e.g. siRNAMAP, BLAST

Fig. 1 Breast oncogenic specific siRNA database

cimap.res.in/sharma/boss/) based on current status of siRNA-mediated repression of oncogenes in different breast cancer cell lines. BOSS is the resource of experimentally verified breast oncogenic siRNAs, collected from research articles and patents. The BOSS database contains information on 865 breast oncogenic siRNA entries. Each entry provides comprehensive information of a siRNA that includes its sequence, name of siRNA, target gene, types of cells, inhibition value, etc. Additionally, some useful tools like siRNAMAP and BOSS BLAST were also linked with the database. siRNAMAP can be used for the selection of the best siRNA design for user target, while the BOSS BLAST tool can predict those siRNA sequences that user siRNA can target effectively. In this database, activities of most of the breast oncogenic specific siRNAs have been tested using MCF-7 cell lines (~400 entries), and MCF-7 and MDA-MD-231 cell lines were used for 46% and 18% siRNA entries, respectively. In order to facilitate the users, several web-based tools have been integrated that include search option, advance option, and browsing option (Fig. 2).

2.1 Materials

For this study, Tyagi et al. [10] collected and compiled 551 breast oncogenic siRNA entries. Each entry provides comprehensive information of a siRNA that includes its sequence, name of the siRNA, target gene, types of cells, inhibition value, etc. Additionally, some useful tools like siRNAMAP and BOSS BLAST were also linked with the database. siRNAMAP can be used for the selection of the best siRNA design for user target, while BOSS BLAST tool can predict those sequences that user siRNA can target effectively. In this database, activities of most of the breast oncogenic specific siRNAs have been tested using MCF-7 cell lines (~221 entries). In order to facilitate the users, several web-based tools have been integrated that include search option, advance option, and browsing options.

2.2 Data Architecture

BOSS database contains the following 17 fields for each siRNA entry (Fig. 2): (1) BOSS ID, (2) PubMed ID, (3) sequence, (4) siRNA name, (5) target gene, (6) GC content, (7) length of siRNA, (8) cell types, (9) year, (10) siRNA source (siRNA/shRNA), (11) position of siRNA, (12) test objective, (13) test method, (14) NCBI Accession no., (15) biological inhibition, (16) transfection reagent, and (17).

2.3 Search

This search option is the easiest way to search BOSS. It allows users to search in any field of the database. In the search field, please type any keyword such as PMID, siRNA name, NCBI accession, target gene, target site sequence, inhibition value, transfection reagent, etc. User can select different fields to be displayed. The keyword should be without spaces. User can choose default fields or select "All." It will display a list of BOSS database-related to keyword, with all the selected fields (Fig. 3).

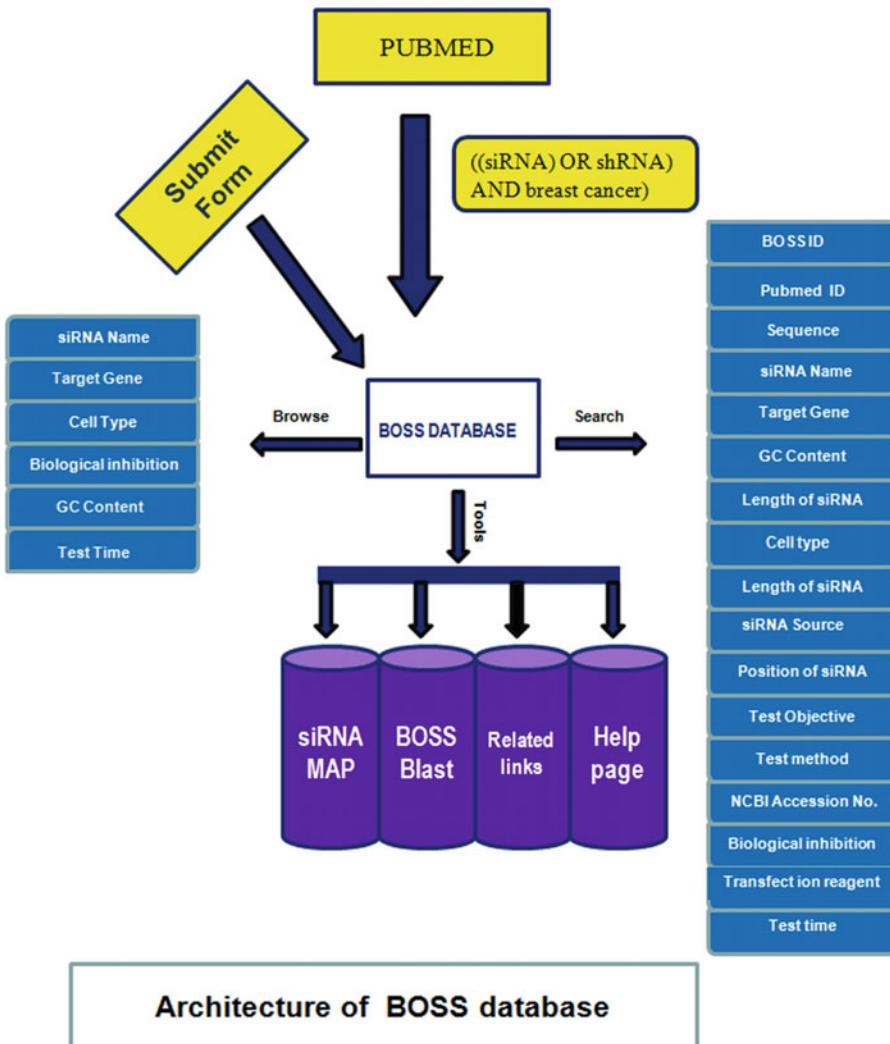


Fig. 2 Data architecture of breast oncogenic specific siRNA database (Tyagi et al. [10])

2.4 Advance Search

This option allows users to perform advance search by making many combinations. User can submit two or more than two queries with conditions. It provides more filtered results. To perform an advanced search, user can enter or select queries and press the submit button. It will display a list of BOSS database related information that satisfies the search criteria (Fig. 4).

2.5 Browse

BOSS database has powerful browsing facility tools that allow a user to browse data using various options. A short description of web interfaces designed for browsing is as follows: This option allows the user to browse BOSS DB based upon some main categories like siRNA name, target gene, cell type, inhibition, GC content, and test time of breast oncogenic specific siRNAs. User can select any of the



Fig. 3 Search page of breast oncogenic specific siRNA database

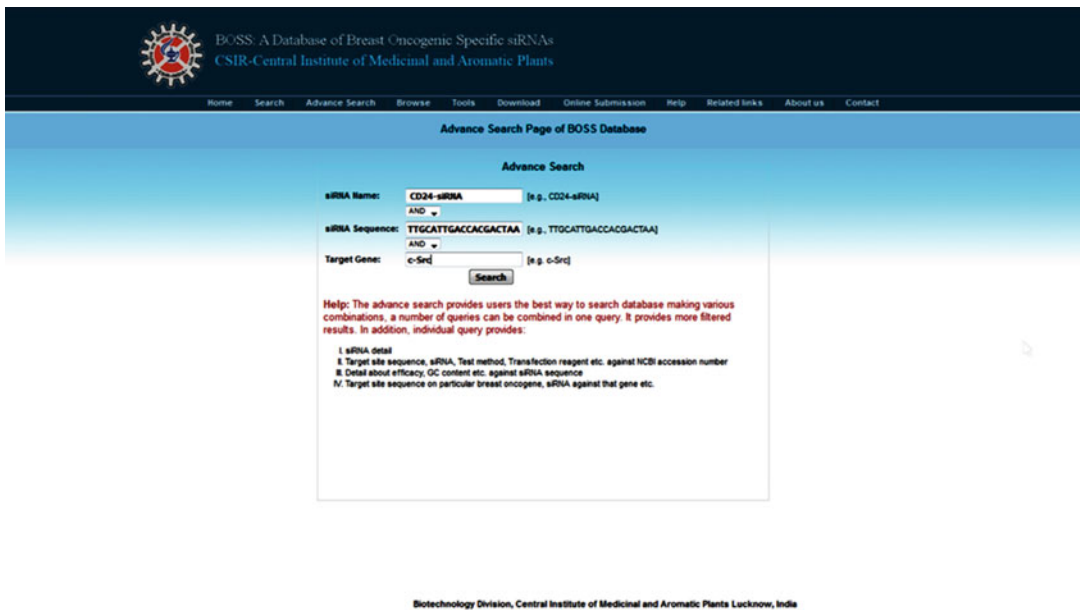


Fig. 4 Advance search page of breast oncogenic specific siRNA database

categories and click button. A list of breast oncogenic specific siRNA related to category is displayed. This option allows the user to browse BOSS DB based upon some main categories like siRNA name, target gene, cell type, inhibition, GC content, and test time of breast oncogenic specific siRNAs. User can select any of the categories and click button. A list of breast oncogenic specific siRNA related to different categories is displayed (Fig. 5).

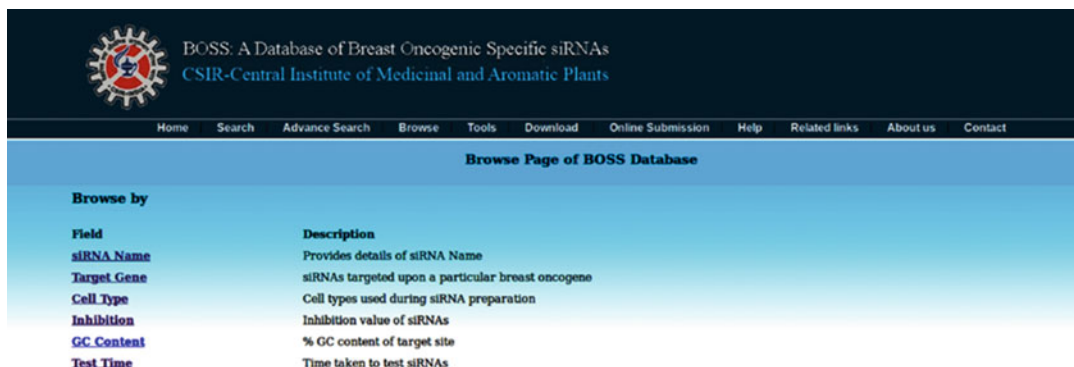


Fig. 5 Browsing page of breast oncogenic specific siRNA database



Fig. 6 siRNA tool page of breast oncogenic specific siRNA database

2.6 Web-Based Tools

BOSS database allows the users to take advantage of useful tools like siRNAMAP and BOSS BLAST. In order to facilitate the user, two tools have been integrated in this database that include:

(1) siRNAMAP for mapping siRNAs on target sequence: The siRNAMAP maps the BOSS database siRNAs on the user nucleotide sequence. It will provide the lists of siRNA from BOSS DB which are complementary to the user-provided sequences. (2) BOSS BLAST for BLAST search against database: User can use this tool for similarity-based search of any query sequences with those present in the BOSS database. By doing this, user can confirm whether a given siRNA sequence or similar siRNA sequence has already been reported or not. User can submit a query sequence in single FASTA format in the search field and press the submit button. It will display all siRNAs similar to query sequence. The server provides option to modify different parameters like scoring matrices, gap penalty, word size, etc.

BOSS database has also an online submission form for the users. In this database user can add new siRNA-/shRNA-related information within specified fields, and subsequently entries are added to the database after proper validation (Fig. 6).

3 Application of BOSS Database

BOSS database will provide the comprehensive details of the experimentally validated breast oncogene specific siRNAs at one platform. It will help researchers to compare the sequence and efficiency of different siRNA molecules and design and/or select potent siRNA for suppression of breast oncogenes. The above-stated approaches will also help in the identification of potential siRNA available in the database. BOSS database includes all the known siRNAs specific for different breast oncogenes that have been experimentally validated. This platform will help to select more than one siRNA molecule targeting different oncogenes. These siRNAs can be used simultaneously to knock down different breast oncogenes if required.

Acknowledgment

A.T. is thankful to ICMR, New Delhi, India, for ICMR-SRF fellowship.

Glossary

BOSS	Breast oncogenic specific siRNAs database
siRNAs	Small interfering RNAs
shRNAs	Short hairpin RNAs
RNAi	RNA interference

References

1. Sledge GW, Miller KD (2003) Exploiting the hallmarks of cancer: the future conquest of breast cancer. *Eur J Cancer* 39:1668–1675
2. Croce CM (2008) Oncogenes and cancer. *N Engl J Med* 358:502–511
3. Osborne C, Wilson P, Tripathy D (2004) Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. *Oncologist* 9:361–377
4. Timmons L, Fire A (1998) Specific interference by ingested dsRNA. *Nature* 395:854–854
5. Zheng Y, Liu Y, Jin H et al (2013) Scavenger receptor B1 is a potential biomarker of human nasopharyngeal carcinoma and its growth is inhibited by HDL-mimetic nanoparticles. *Theranostics* 3:477–486
6. Mishra MN, Mishra MN, Vangara KK et al (2014) Transcriptional targeting of human liver carboxylesterase (hCE1m6) and simultaneous expression of anti-BCRP shRNA enhances sensitivity of breast cancer cells to CPT-11. *Anticancer Res* 34:6345–6351
7. Dash R, Moharana SS, Reddy AS et al (2006) DSTHO: database of siRNAs targeted at human oncogenes: a statistical analysis. *Int J Biol Macromol* 38:65–69
8. Tyagi A, Ahmed F, Thakur N et al (2011) HIVsirDB: a database of HIV inhibiting siRNAs. *PLoS One* 6:e25917
9. Ren Y, Gong W, Zhou H et al (2009) siRecords: a database of mammalian RNAi experiments and efficacies. *Nucleic Acids Res* 37:D146–D149
10. Tyagi A, Semwal M, Sharma A (2017) A database of breast oncogenic specific siRNAs. *Sci Rep* 7:8706
11. Jemal A, Bray F, Center MM et al (2011) Global cancer statistics. *CA Cancer J Clin* 61:69–90
12. Elbashir SM, Harborth J, Lendeckel W et al (2001) Duplexes of 21-nucleotide RNAs

- mediate RNA interference in cultured mammalian cells. *Nature* 411:494–498
13. Mazur S, Csucs G, Kozak K (2012) RNAiAtlas: a database for RNAi (siRNA) libraries and their specificity. *Database (Oxford)* 2012:bas027
 14. Zhang C, Li G, Zhu S et al (2014) tasiRNAdb: a database of ta-siRNA regulatory pathways. *Bioinformatics* 30:1045–1046
 15. Saunders RE, Instrell R, Rispoli R et al (2013) HTS-DB: an online resource to publish and query data from functional genomics high-throughput siRNA screening projects. *Database (Oxford)* 2013:bat072
 16. sIR: siRNA information Resource, a web-based tool for siRNA sequence design and analysis and an open access siRNA database. *BMC Bioinformatics*, Full Text, <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-8-178>
 17. Truss M, Swat M, Kielbasa SM et al (2005) HuSiDa—the human siRNA database: an open-access database for published functional siRNA sequences and technical details of efficient transfer into recipient cells. *Nucleic Acids Res* 33:D108–D111
 18. Thakur N, Qureshi A, Kumar M (2012) VIR-siRNAdb: a curated database of experimentally validated viral siRNA/shRNA. *Nucleic Acids Res* 40:D230–D236
 19. Dar SA, Thakur A, Qureshi A et al (2016) siRNAmoD: a database of experimentally validated chemically modified siRNAs. *Sci Rep* 6:20031
 20. Liang Y, Gao H, Lin S-Y et al (2010) siRNA-based targeting of cyclin E overexpression inhibits breast cancer cell growth and suppresses tumor development in breast cancer mouse model. *PLoS One* 5:e12860
 21. Garrido P, Osorio FG, Morán J et al (2015) Loss of GLUT4 induces metabolic reprogramming and impairs viability of breast cancer cells. *J Cell Physiol* 230:191–198
 22. Qin B, Cheng K (2010) Silencing of the IKKε gene by siRNA inhibits invasiveness and growth of breast cancer cells. *Breast Cancer Res BCR* 12:R74
 23. Xu D, Kang H, Fisher M et al (2004) Strategies for inhibition of MDR1 gene expression. *Mol Pharmacol* 66:268–275
 24. Luo X-G, Zou J-N, Wang S-Z et al (2010) Novobiocin decreases SMYD3 expression and inhibits the migration of MDA-MB-231 human breast cancer cells. *IUBMB Life* 62:194–199
 25. Shaker H, Harrison H, Clarke R et al (2017) Tissue factor promotes breast cancer stem cell activity in vitro. *Oncotarget* 8:25915–25927
 26. Wu J, Richer J, Horwitz KB et al (2004) Progesterin-dependent induction of vascular endothelial growth factor in human breast cancer cells: preferential regulation by progesterone receptor B. *Cancer Res* 64:2238–2244
 27. Liang B, Wang X-J, Shen P-H et al (2013) Synuclein-γ suppression mediated by RNA interference inhibits the clonogenicity and invasiveness of MCF-7 cells. *Oncol Lett* 5:1347–1352
 28. Han G, Fan B, Zhang Y et al (2008) Positive regulation of migration and invasion by vasodilator-stimulated phosphoprotein via Rac1 pathway in human breast cancer cells. *Oncol Rep* 20:929–939
 29. Ji X, Lu H, Zhou Q et al (2014) LARP7 suppresses P-TEFb activity to inhibit breast cancer progression and metastasis. *Elife* 3:e02907
 30. Jang J-Y, Choi Y, Jeon Y-K et al (2008) Suppression of adenine nucleotide translocase-2 by vector-based siRNA in human breast cancer cells induces apoptosis and inhibits tumor growth in vitro and in vivo. *Breast Cancer Res BCR* 10:R11
 31. Aletaha M, Mansoori B, Mohammadi A et al (2017) Therapeutic effects of bach1 siRNA on human breast adenocarcinoma cell line. *Biomed Pharmacother* 88:34–42
 32. Sun L, Cai L, Yu Y et al (2007) Knockdown of S-phase kinase-associated protein-2 expression in MCF-7 inhibits cell growth and enhances the cytotoxic effects of epirubicin. *Acta Biochim Biophys Sin* 39:999–1007
 33. Salceda S, Tang T, Kmet M et al (2005) The immunomodulatory protein B7-H4 is overexpressed in breast and ovarian cancers and promotes epithelial cell transformation. *Exp Cell Res* 306:128–141
 34. Toy EP, Lamb T, Azodi M et al (2011) Inhibition of the c-fms proto-oncogene autocrine loop and tumor phenotype in glucocorticoid stimulated human breast carcinoma cells. *Breast Cancer Res Treat* 129:411–419
 35. Li Z, Meng Q, Pan A et al (2017) MicroRNA-455-3p promotes invasion and migration in triple negative breast cancer by targeting tumor suppressor EI24. *Oncotarget* 8:19455–19466
 36. US 7615627 B2—Rna interference mediated inhibition of aurorakinase B and its combinations as anticancer therapy. *The Lens*, <https://www.lens.org/lens>
 37. Soni A, Akcakanat A, Singh G et al (2008) eIF4E knockdown decreases breast cancer cell growth without activating Akt signaling. *Mol Cancer Ther* 7:1782–1788

Part V

Special Topics



Historeceptomics: Integrating a Drug's Multiple Targets (Polypharmacology) with Their Expression Pattern in Human Tissues

Timothy Cardozo

Abstract

Historeceptomics is a new, integrative informatics approach to describing the mechanism of action of drugs in a holistic, in vivo context. The approach is based on leveraging emerging big data sources in genomics and chemistry to incorporate tissue specificity into mechanism of action descriptions. New insights into drug mechanism of action, drug repurposing, and prediction of adverse effects may be possible, including the design and development of multi-target drugs or drug combinations. The critical elements still under development include: (1) defining the tissue ensemble associated with specific human diseases, (2) appreciating the pattern or partitioning of the expression of drug targets (receptors) across and outside of these ensembles, and (3) informatics methods to integrate direct drug-receptor data with receptor expression data in tissues. Maturation of this field may enable the complementary field of tissue-targeted drug delivery, enabling novel concepts in drug design and development for unmet medical needs.

Keywords Adverse events, Big data, Chemical genomics, Drug mechanism of action, Drug optimization, Systems pharmacology

1 Introduction

Most drugs in clinical use exhibit moderate to high affinities to more than one receptor, a phenomenon termed “polypharmacology.” Ironically, almost all drugs are colloquially classified and identified for discussion purposes according to their primary receptor or target, e.g., “HMG-CoA reductase inhibitor.” This inconsistency is representative of a potentially serious deficiency in drug design and discovery, namely, a tendency to ignore off-targets in designing and discovering drug candidates. This deficiency may be compounded by a conceptually parallel ignorance of the in vivo context of drug action, namely, a failure to appreciate the variable expression and function of drug targets across the different tissues of the human body. Undoubtedly, these deficiencies are a consequence of an historical absence of reliable technologies or data necessary to take polypharmacology or cross-tissue dimensions into account. The recent explosion of big data in chemistry,

pharmacology, and genomics, however, has raised the question as to whether sufficient data is now available to make progress in these areas. Extensive genomic [1], and even proteomic [2], molecular profiling of tissues is available. Here, we review the basic questions and existing work related to this problem and to what degree it can be incorporated into multi-target drug discovery and design. First, we explore whether specific tissue ensembles can be defined for specific human diseases: for example, central nervous system tissues underlie psychiatric diseases, but what is the relevant tissue ensemble for diabetes? Second, we review the scant research into a question that might have been considered an assumption and ignored, namely, are drug targets actually normally expressed in the tissues of the disease for which they are indicated as treatment? Then, we review the data sources that can be used to develop polypharmacology and holistic in vivo cross-tissue-inclusive signatures for the mechanisms of action (MOA) of drugs and drug candidates. Finally, we will review the technologies for generating such signatures and their performance in producing impactful results.

2 Which Tissues Are Responsible for a Disease?

There are three dimensions to defining disease-relevant tissues. One simple definition is to equate disease-relevant tissue with diseased tissue. This is the most common, unspoken, operative definition and has been invoked largely in the context of drugs for cancer. Under this definition, the question of whether a drug target is expressed in a disease-relevant tissue is relatively simple: in many cases, the drug was developed *because* the target receptor was noted to be overexpressed (e.g., Her-2/neu in breast cancer [3]) or overactive (e.g., B-raf in melanoma [4]) in the disease tissue and exhibited a pro-cancer function. Drugs were usually then designed or discovered to target that receptor. For example, the leukemia drug Gleevec (imatinib) targets the bcr-abl fusion protein kinase [5] that is a marker of chronic myelogenous and other leukemias. Nevertheless, cancer drugs designed for a specific cancer often were found to be effective in other cancers, exhibiting tissue specificity in the sense of specificity across different cancers or cancer cell lines [6]. Indeed, this observation and the interest in matching drugs for different cancers have made screening of cancer cell line panels, such as the NCI-60 [7], a common step in cancer drug discovery. Little to no work has focused on non-cancer tissues in cancer patients, despite growing evidence that host (noncancerous) immune and other tissues are crucial players in the establishment, maintenance, and progression of several cancers [8]. Some cancer drugs, such as angiogenesis inhibitors [9], are indeed targeted at host tissues, but only tissues with direct physiological and mechanistic interactions with the cancer are so targeted. This limited approach to the in vivo context of cancer

drugs is illustrated by the absence of widespread, industrial drug discovery interest in targeting tumor-infiltrating lymphocytes (TILs; a tissue) for many years until the mechanism of co-receptor blockade of these TILs was understood, leading to highly effective immunotherapies for previously treatment-resistant cancers [10]. The trajectory of research under the umbrella of a narrow definition of disease-relevant tissues as only diseased tissues thus begs the question of how many other drug discovery opportunities have been missed because the mechanisms of tissues influencing cancer establishment, maintenance, or progression have not yet been understood. Furthermore, the cancer-centric nature of this definition may have held back tissue-specific drug discovery and design approaches in other fields. This is changing to some degree as the urgency of certain unmet medical needs have forced creative approaches into this area. For example, some promising approaches to non-addictive pain medication (to prevent opioid addiction) target inflamed tissue over the CNS [11]. Equating disease-relevant with diseased tissue is intuitive and single-target, but, even so, may be significantly underutilized in drug design.

Another definition that more closely approaches the *in vivo* context of drug action is to define disease-relevant tissues as those mechanistically involved in the disease in human subjects suffering from the disease. This is more complex than the cancer-centric, single-tissue, single-target view of diseased tissues and has the advantage of broadening the drug discovery to the entire ensemble of tissues mechanistically interacting with the tissue that produces the prominent symptoms of the disease. In cancer, as discussed above, those tissues are, at least, noncancerous vascular, immune, and metabolic tissues, and important advances in cancer drug discovery have been made targeting these tissues (angiogenesis inhibitors, co-receptor blockade [12]). However, even these approaches adhere to a single-tissue, single-target view, and rationales for multi-target therapies intended to attack synergistic vulnerabilities *between* tissue-target pairs in cancer have been few and far between. For example, synergistic beneficial effect might be expected for a multi-target drug simultaneously targeting angiogenesis in the vasculature, DNA replication in the cancer tissue, and co-receptor blockade in the TILs.

Beyond cancer, the definition of disease-relevant tissues as the mechanistic ensemble in the disease patient is similarly rarely exploited, perhaps because of the primitive means of targeting drugs to specific tissues (*see* Subheading 6). Overall, this definition opens more avenues to improvements in drug discovery via multi-target tissue-specific drugs, but these avenues remain underexplored. Remarkably, despite the importance of this question, only one group has attempted to systematically map diseases to tissues they affect (disease-tissue matrix) independently of molecular information. Lage et al. surveyed disease-relevant literature in PubMed

to create a disease-tissue covariation matrix of high-confidence associations of >1000 diseases to 73 tissues [13]. Even so, these investigators only focused on heritable diseases and only performed this task in the service of a larger investigation into the connection between tissue-specific pathology and causal gene expression. Nevertheless, the Lage disease-tissue matrix, although published 10 years ago prior to major revolutions in molecular profiling and informatics, remains the only such matrix ever developed. Other matrices are derived from network analysis of gene-expression data across multiple tissues [14], which may skew to the molecular from the full in vivo context, but may nevertheless be useful. In particular, work in this area is showing that different neighborhoods of the human interactome may represent specific, different tissues, suggesting that modules or pathways differentiate tissues or differentiate along with tissue differentiation. This appears to hold true in both multi-genetic [15] and hereditary [16] diseases. Novel informatics methods for this purpose are also being introduced [17]. Still, the paucity of published disease-tissue matrices is indicative of the extreme underutilization of the tissue-specific approach to drug design and discovery in non-cancer diseases, to say nothing of multi-target approaches to these multi-tissue, disease constructs.

3 Are Drug Targets Expressed in Disease-Relevant Tissues?

Perhaps the most general definition of disease-relevant tissue specifies tissues that are biologically relevant to the disease whether or not they originate from patients who have the disease. This is a less intuitive question: is the baseline presence of drug targets in normal tissues known to be biologically/mechanistically relevant to a disease predictive of response to a drug? The presence of the drug target in the tissue of a person without the disease is not necessarily predictive of its presence and druggability in a person with the disease. This assumption was challenged, posed as a question and answered only recently by Kumar and associates who profiled 345 drug targets of 406 diseases in 32 tissues [18]. They found that 87% of drug targets are expressed in the non-diseased, disease-relevant tissues in non-diseased subjects. The data overwhelmingly established that expression of a putative drug target in non-cancer, disease-relevant tissue in vivo is a predictive factor for eventual success of the drug. The finding further diminishes the hypothesis that viable drug targets are *induced* by the disease, calling into question the applicability of common in vitro drug sensitivity prediction approaches, which, in addition to being underwhelming in performance, notoriously find prominent drug targets upregulated by the disease or the drug itself [19]. Furthermore, the divergence between the expression pattern of drug targets in diseased cells in vitro and the Kumar findings in vivo emphasize the importance

of assessing drug targets and especially any multi-target drug concept, in the holistic, cross-tissue *in vivo* context.

4 Tissue Specific and Polypharmacology Data

High-throughput, molecular profiling of human cells and tissues, both normal and diseased, is now routine and abundant. An enormous amount of genetic (DNA), expression (RNA), and proteomic (protein) data on just about every cell and tissue from many sources is available publicly [1]. Major genomic data sets include GWAS, eQTL, GTEx, ENCODE, dbGAP, BioGPS, the Cancer Cell Line Encyclopedia (CCLE), GEO, Expression Atlas, TIGER, and Road-Map Epigenomics databases. Major proteomic data sets include the Human Protein Atlas, Human Proteome Map, and ProteomicsDb. Specific composite data sources of high value include the Allen Brain Atlas, TISSUES, Project Achilles, MotifMap, TRANSFAC, ChIP-X, the Molecular Signatures Database, PhosphoMouse, and CircNet. The amount of data in these sources is already overwhelming, covers just about every disease and phenotype and is growing exponentially. The prospects for extracting profiles of the molecular variability of a drug target across tissues or cells have never been better.

High-throughput drug screening programs have produced extensive data on drug-target affinities. ChEMBL, the Psychoactive Drug Screening Program, and PubChem are the most prominent of these sources. A growing number of large-scale computational molecular docking efforts have added exponentially to this trove of data, driven by the interest of cloud-based computing service companies like Google and Amazon in demonstrating large-scale computing tasks [20]. All possible chemical structures may be considered, not just those synthesized into existence [21]. Although the space of all drug candidates measured against all drug targets in the human genome is immeasurably large, and enormous amount of polypharmacology data is available, to the point that polypharmacology of approved drugs may often be reliably appreciated.

5 Does Assessing the Tissue Pattern of Expression of the Full Polypharmacology (Primary and Off-Targets) of Drugs Add Value?

The data may now exist to infer polypharmacologic ensembles of drugs computationally and project this ensemble of targets across their presence in tissues of the body to approximate the *in vivo* context, but *in silico*. Drugs used to treat human diseases and chemicals that cause diseases or adverse effects indisputably act on specific cellular components that vary between body tissues, and

most drugs exert their beneficial and adverse effects through their combined action on several different molecular targets. The concept of “historeceptomics” signatures of the actions of drugs [20] was introduced to generate signatures integrating polypharmacology with the in vivo, cross-tissue context by collecting the ensemble of all the receptors upon which a drug acts projecting it to the differential level of expression of these receptors across organs/tissues. Remarkably, this is the only report on drug mechanisms in the literature based on the postulate that drug action depends on *both* affinity and receptor abundance in tissues [22]: there are thousands of publications linking drug effects to specific targets/receptors with no mention of the tissue expression pattern of targets.

Historeceptomics (Fig. 1) maps both diseases/phenotypes and polypharmacologic drugs into a common, but previously unexplored, *target-tissue space* [20]. This means that instead of diseases being stratified according either to deficiencies in a single-gene product (e.g., thalassemia \leftrightarrow hemoglobin) or a deficiency in certain tissues (e.g., breast cancer), they are thought of as deficiencies in a

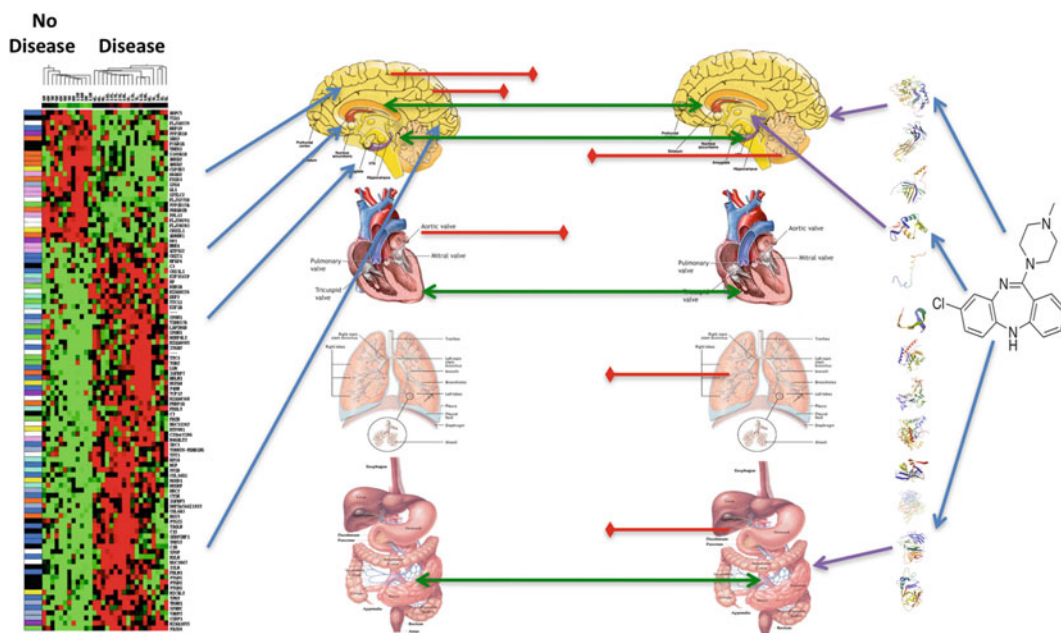


Fig. 1 Historeceptomics. *Far right:* A drug has many targets (blue arrows) that determine its action in the body. These are then projected to the tissues in which they are expressed (purple arrows) to form *target-tissue molecular signature of the polypharmacologic, in vivo action of the drug*. *Far left:* Disease is profiled by genomics/proteomics/etc. across the tissues of the human body to form *target-tissue molecular signatures of the disease*. Where target-tissue pairs are activated by the drug but are unrelated to the disease (red block arrows), these could be adverse effects to be eliminated from the drug action. Where target-tissue pairs are active in the disease but unmatched to the drug target-tissue profile, these are opportunities for multi-target drug improvement

holistic pattern of many specific gene products in many specific tissues (e.g., many genes in many tissues in metabolic syndrome). In turn, drug action is described as a highly specific pattern of modulation of several specific gene products in several specific tissues. The critical integrative data structure is the tissue atlas of the human body, which allows polypharmacologic actions of drugs to be easily matched to Omics-derived biomarkers of complex, untreatable diseases like Alzheimer's, COPD, and schizophrenia by comparing two target-tissue vectors. In addition, combinations of drugs can be matched to complex biomarkers rationally for the first time, and targeted drug delivery can be advanced to use small molecule drugs and be applicable to many diseases (instead of just drug delivery vehicles targeting focal, diseased tissues like cancer), with attendant benefits in precision, efficacy, and safety. This concept renders complex multi-genetic diseases and drugs with complex polypharmacologic mechanisms of action (i.e., most drugs in clinical use today) as easily and simply matched as the previously successful model of single gene biomarker-matched therapy (e.g., Herceptin for Her2/neu + breast cancer).

Generating a tissue-target score: Target-tissue scores are based on a simple postulate—the affinity of a drug for a target is unrelated to its therapeutic effect if the target is not expressed in tissues relevant to the disease. For example, drug X treats schizophrenia and has very high affinity for receptor Y measured in vitro, but receptor Y is only expressed in the bone matrix. Therefore, receptor Y is not a significant target of drug X with respect to schizophrenia. To integrate drug bioactivity with the target gene expression, a key enabling assumption is made to detect differential expression across tissues via normalized gene expression and cross-tissue Z-normalization. This enhances detection of highly expressed outliers in each tissue but misses ubiquitously expressed targets such as actin. The integrated tissue-target score is then calculated by multiplying the Z-score for gene expression for each target identified for a drug in the chemical-biology network by the $-\log(\text{affinity})$ of the drug for the target (Fig. 2).

Thousands of tissue-target scores can thus be generated in this prototype for each individual drug. These scores are normally distributed, so outlier detection can be used to identify the significant scores at a chosen α level. The array of these significant hits is the signature (Fig. 3).

Historeceptomics was used to investigate the unique MOA of the antipsychotic drug clozapine. Most psychiatric diseases do not have a known organic etiology (no genetics, no animal models, etc.). They are simply a classification of symptom and sign patterns. This has greatly inhibited understanding of the mechanism of action of psychiatric drugs, which are very polypharmacologic. The anti-schizophrenia drug clozapine (Clozaril[®]), which is an atypical antipsychotic with greater efficacy for certain psychotic

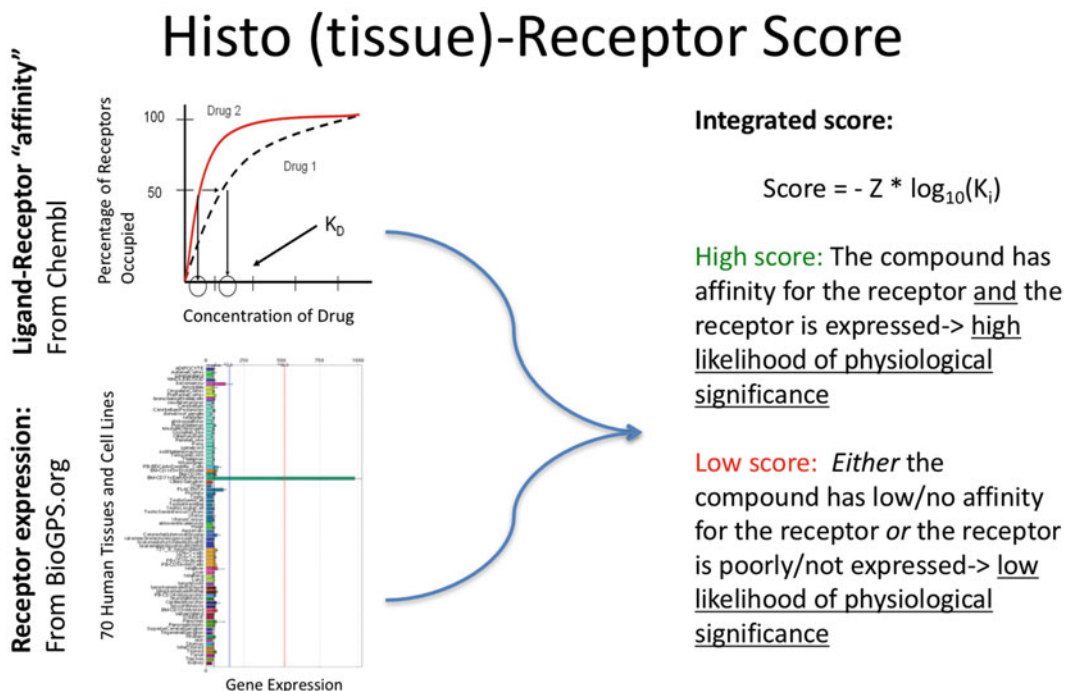


Fig. 2 Calculations of historeceptomics (target-tissue) score. Affinity or computational molecular docking score (upper left) of drug with a target/receptor is integrated with differential expression of that receptor across the tissues of the human body (lower left) using a Z-statistic (center). The resulting scores represent the “product” of the drug’s likelihood of directly interacting with the receptor, specifically and independently in each tissue (right)

symptoms and less propensity for motor side effects as compared to the typical antipsychotic drug chlorpromazine, was studied. The mechanism of action for clozapine’s atypicality/atypia is unknown but can be derived from historeceptomics tissue-target signatures by subtracting the signature of a typical antipsychotic (chlorpromazine) from clozapine’s signature. The 5HT2a receptor in the PFC and D2 receptors in multiple brain regions were found to be common contributors to the MOA of both clozapine and chlorpromazine. These results strongly correlate with what is known about these drugs, including that the hallucinogenic drug, LSD, operates on the 5HT2a receptor in the PFC and produces hallucinations which can resemble those experienced by people with schizophrenia. The clozapine-specific result revealed that M1 and M3 muscarinic receptors in the prefrontal cortex (PFC), the dopamine D4 receptor in the pineal gland, and the histamine H1 receptor in the superior cervical ganglion (SCG) were tissue-target outliers for clozapine and therefore candidates for the mechanism of its atypical effects. Indeed, the pineal gland produces melatonin, which in addition to its well-known mood-stabilizing properties (clozapine has unique mood effects), may be effective against some symptoms of schizophrenia [23] especially negative symptoms

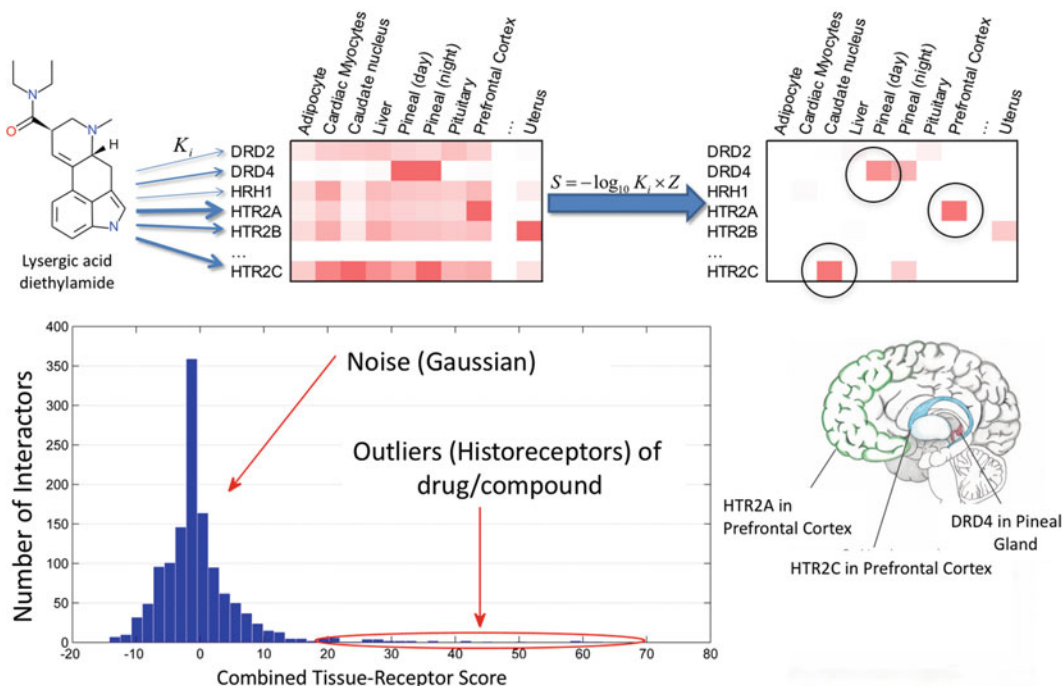


Fig. 3 (a) Historeceptomics signature of action of the drug LSD. Blue arrows represent affinities to targets with thickness proportional to level of affinity. The first heat map shows the expression levels of those targets across tissues. Applying the integration results in massive dimensionality reduction. Significant outliers **(b)** form the tissue-atlas signature **(c)**

against which clozapine has particular efficacy as compared to other neuroleptics. M1/M3 are also recently emerging as promising targets for negative symptoms [24]. Overall, historeceptomics rebalances the view of the MOA of clozapine to reconcile previously divergent data on the action of the drug. Repurposing, adverse effect MOA and target validation all arise from the analysis of a complex and poorly understood but clinically useful drug. As most drugs in clinical use were discovered via phenotypic assays or means, MOAs of most drugs in clinical use are poorly understood and could benefit from this multidimensional view.

6 Tissue Targeting of Drugs

Notably, the historeceptomics concept necessarily invokes targeted drug delivery. A polypharmacologic, in vivo context drug MOA is better understood in terms of the target tissue. Therefore, it stands to reason that beneficial effects of drugs could be amplified with no increase in adverse effects by preferential effects of those specific targets in the specific tissues identified by the method. Crude tissue targeting is already possible via exploitation of in vivo barrier

compartments or altered in vivo environments. For example, several small molecule drugs, such as glycopyrrolate, with potentially serious sedative or other CNS effects, are widely used for peripheral effects because they do not cross the blood-brain barrier. Cancer tissue is sufficiently different and accessible as compared to normal tissues that this is the most advanced area of tissue targeting, where molecular targeting using biologics is already utilized. In addition, the aforementioned strategy of targeting the different pH of inflamed tissue for nonaddictive pain management small molecule chemical compounds is a similar tissue-targeting approach [11]. Nevertheless, historeceptomics reveals avenues that are currently not yet approachable. In the example of clozapine given above, a drug that preferentially or exclusively targets the M1 and/or M3 receptors in the PFC of the brain would be hypothesized to be a breakthrough for psychiatry, but how would this be accomplished?

7 Summary

Incorporating both polypharmacology and the in vivo context of drug action in multi-target drug design is a chemical-biology approach in its infancy but clearly offers pioneering opportunities for progress in drug discovery and design. The explosion of biomedical big data has made such an approach possible, but major current challenges involve:

- Precise definitions of diseases
- Accurate and precise disease-tissue association matrices
- Generating the complete polypharmacologic ensemble of small molecule drugs
- Accurate and precise methods, like historeceptomics, for projecting the complex polypharmacologic ensemble of a (multi-target) drug in a weighted manner into the tissue atlas of the human body
- Precise and effective tissue targeting of drugs

Advances in these areas may greatly enable multi-target drug discovery and design by revealing the exact tissue-target pairs responsible for different aspects of the MOA and for specific adverse effects of drug candidates, which would allow surgical design of multi-target drugs and drug combinations to emphasize beneficial drug actions in specific tissues and eliminate adverse drug action in other tissues. Advances in this area will also therefore increase the urgency of developing creative tissue-targeting drug methods, especially for small molecules and for tissue compartments to which biologics have restricted access.

Disclosure

The author is a co-founder of GeneCentrix, Inc., a company developing historeceptomics software.

References

1. Cardozo T, Gupta P, Ni E, Young LM, Tivon D, Felsovalyi K (2016) Data sources for in vivo molecular profiling of human phenotypes. *Wiley Interdiscip Rev Syst Biol Med* 8:472–484. <https://doi.org/10.1002/wsbm.1354>
2. Uhlen M et al (2015) Proteomics. Tissue-based map of the human proteome. *Science* 347:1260419. <https://doi.org/10.1126/science.1260419>
3. Ross JS, Gray GS (2003) Targeted therapy for cancer: the HER-2/neu and Herceptin story. *Clin Leadersh Manag Rev* 17:333–340
4. Bhatia P, Friedlander P, Zakaria EA, Kandil E (2015) Impact of BRAF mutation status in the prognosis of cutaneous melanoma: an area of ongoing research. *Ann Transl Med* 3:24. <https://doi.org/10.3978/j.issn.2305-5839.2014.12.05>
5. Cortes J (2004) Natural history and staging of chronic myelogenous leukemia. *Hematol Oncol Clin North Am* 18:569–584., viii. <https://doi.org/10.1016/j.hoc.2004.03.011>
6. Yao F et al (2018) Tissue specificity of in vitro drug sensitivity. *J Am Med Inform Assoc* 25:158–166. <https://doi.org/10.1093/jamia/ocx062>
7. Chabner BA (2016) NCI-60 cell line screening: a radical departure in its time. *J Natl Cancer Inst* 108. <https://doi.org/10.1093/jnci/djv388>
8. Malanchi I (2013) Tumour cells coerce host tissue to cancer spread. *Bonekey Rep* 2:371. <https://doi.org/10.1038/bonekey.2013.105>
9. Yadav L, Puri N, Rastogi V, Satpute P, Sharma V (2015) Tumour angiogenesis and angiogenic inhibitors: a review. *J Clin Diagn Res* 9:XE01–XE05. <https://doi.org/10.7860/JCDR/2015/12016.6135>
10. Callahan MK, Postow MA, Wolchok JD (2016) Targeting T cell co-receptors for cancer therapy. *Immunity* 44:1069–1078. <https://doi.org/10.1016/j.immuni.2016.04.023>
11. Spahn V et al (2017) A nontoxic pain killer designed by modeling of pathological receptor conformations. *Science* 355:966–969. <https://doi.org/10.1126/science.aai8636>
12. Waldmann H, Adams E, Cobbold S (2008) Reprogramming the immune system: co-receptor blockade as a paradigm for harnessing tolerance mechanisms. *Immunol Rev* 223:361–370. <https://doi.org/10.1111/j.1600-065X.2008.00632.x>
13. Lage K et al (2008) A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. *Proc Natl Acad Sci U S A* 105:20870–20875. <https://doi.org/10.1073/pnas.0810772105>
14. Kotlyar M, Pastrello C, Sheahan N, Jurisica I (2016) Integrated interactions database: tissue-specific view of the human and model organism interactomes. *Nucleic Acids Res* 44: D536–D541. <https://doi.org/10.1093/nar/gkv1115>
15. Kitsak M, Sharma A, Menche J, Guney E, Ghiassian SD, Loscalzo J, Barabasi AL (2016) Tissue specificity of human disease module. *Sci Rep* 6:35241. <https://doi.org/10.1038/srep35241>
16. Barshir R, Shwartz O, Smoly IY, Yeger-Lotem E (2014) Comparative analysis of human tissue interactomes reveals factors leading to tissue-specific manifestation of hereditary diseases. *PLoS Comput Biol* 10:e1003632. <https://doi.org/10.1371/journal.pcbi.1003632>
17. Mohammadi S, Grama A (2016) A convex optimization approach for identification of human tissue-specific interactomes. *Bioinformatics* 32:i243–i252. <https://doi.org/10.1093/bioinformatics/btw245>
18. Kumar V, Sanseau P, Simola DF, Hurler MR, Agarwal P (2016) Systematic analysis of drug targets confirms expression in disease-relevant tissues. *Sci Rep* 6:36205. <https://doi.org/10.1038/srep36205>
19. Costello JC et al (2014) A community effort to assess and improve drug sensitivity prediction algorithms. *Nat Biotechnol* 32:1202–1212. <https://doi.org/10.1038/nbt.2877>
20. Shmelkov E, Grigoryan A, Swetnam J, Xin J, Tivon D, Shmelkov SV, Cardozo T (2015) Historeceptomic fingerprints for drug-like compounds. *Front Physiol* 6:371. <https://doi.org/10.3389/fphys.2015.00371>

21. Ruddigkeit L, Blum LC, Reymond JL (2013) Visualization and virtual screening of the chemical universe database GDB-17. *J Chem Inf Model* 53:56–65. <https://doi.org/10.1021/ci300535x>
22. Shmelkov E, Grigoryan AV, Swetnam J, Xin J, Shmelkov S, Cardozo T (2015) Historeceptomics fingerprints for drug-like compounds. *Front Physiol* 6:371
23. Morera-Fumero AL, Abreu-Gonzalez P (2013) Role of melatonin in schizophrenia. *Int J Mol Sci* 14:9037–9050. <https://doi.org/10.3390/ijms14059037>
24. Scarr E, Dean B (2008) Muscarinic receptors: do they have a role in the pathology and treatment of schizophrenia? *J Neurochem* 107:1188–1195. <https://doi.org/10.1111/j.1471-4159.2008.05711.x>



Networking of Smart Drugs: A Chem-Bioinformatic Approach to Cancer Treatment

Kavindra Kumar Kesari, Qazi Mohammad Sajid Jamal,
Mohd. Haris Siddiqui, and Jamal Mohammad Arif

Abstract

Increasing rate of cancer incidence has put a global demand of high and advanced level of diagnosis methods. Although surgery (surgical treatment), radiotherapy, and chemotherapy are well-known traditional methods for cancer treatment, the use of novel, advanced, and reliable drug delivery methods (nanodrugs, DNA origami, nanoparticles, and exosomes) are now more applicable to reduce the associated side effects of traditional cancer therapies. Therefore, this chapter will review the existing network of smart drugs by using chem-bioinformatic approach towards cancer treatment. In future, the combination of computational tools in smart drug designing for cancer treatment will be path-breaking.

Keywords DNA origami, Exosomes, In silico, Nanodrugs, Smart drugs

1 Introduction

The growing rate of cancer among young generation is an issue of concern. Lifestyle factors like cigarette smoking, alcohol drinking and caffeine, radiations, illicit drug use, and exposure to different kind of toxicants available in the environment and diet have been found a cause of concern. By 2030, with a projected world population of 8.3 billion, demographic effects alone will give rise to an estimated 21.4 million incident cases and 13.2 million deaths due to cancer [1, 2]. With increasing cases of cancer, demand of advanced technologies for cancer treatment is also increasing. Surgery (surgical treatment), radiotherapy, chemotherapy, and hormone therapy come in our mind first in respect to cancer treatment, as these are more frequently available globally. However, advances in science and technology are showing an amazing world where it keeps moving towards more advanced methods and techniques. The term “smart drug” is more common nowadays to understand the targeted drug delivery systems integrated from various immunological therapies for cancer treatment. It is a method of delivering medication to a patient in a manner that increases the concentration of the medication in the infected organs

or cells, relative to the normal cells [3]. There are different ways of targeted drug delivery systems; however, inverse, double, active, passive, dual, and physical targeting are being used broadly in cancer therapy. The drug-targeted delivery system is to confine, prolong drug properties, fortify a distinct route, target the desired site, reduce side effects of the drugs, and prolong drug interaction with the diseased tissue [4]. “Smart drugs” specifically target cancer cells and not healthy ones as also presented in Fig. 1. On the other hand, it can normally reduce side effects of chemotherapy and radiation.

The carrier of smart drugs could be nanoparticles, liposomes, exosomes, and nanotubes, where the applications of all the above should have to adopt some smart ways to target the cells [5]. An implementation of computational models to target the cancer cell will provide a mechanistic approach to understand the treatment of cancer cells. The possible pathway for the loading of drugs and their delivery through various types of carriers and the uptake in cells by targeting cancer cells are presented in Fig. 1. The purpose of this pathway is to show the drug-targeted delivery system which may increase the therapeutic efficacy by controlling the toxic effects. The delivery of drugs such as tyrosine, DNA origami, and small molecules by using an appropriate carrier to malignant tissues is increased and the normal tissue remains unaffected. The targeted

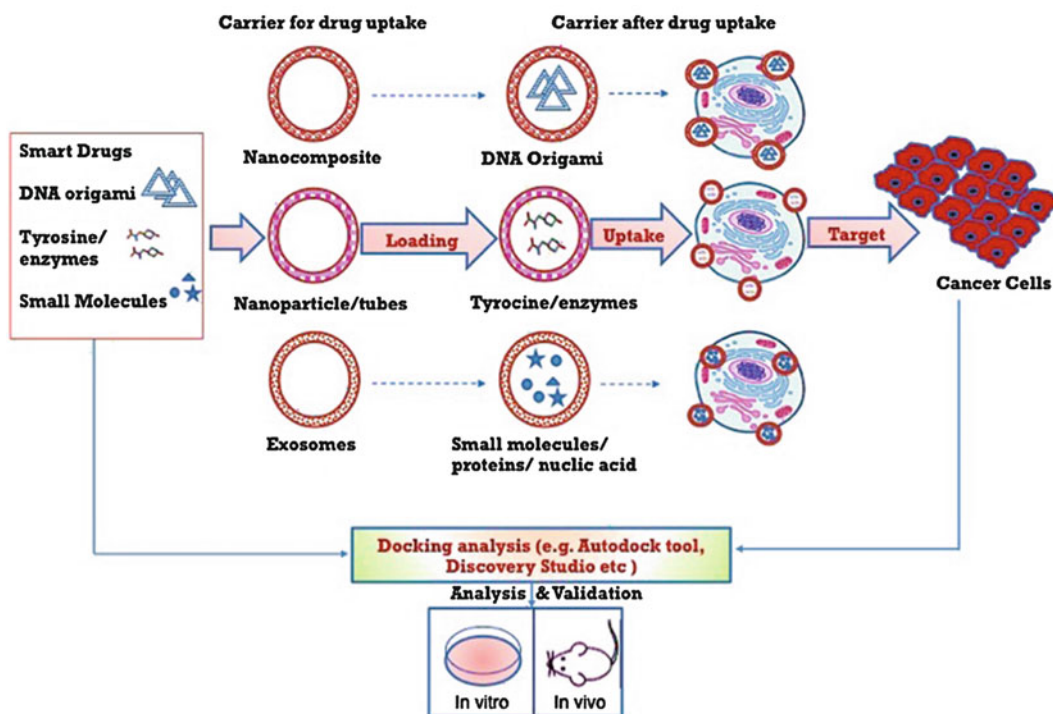


Fig. 1 The possible pathways of smart drug uptake and targeting of cancer cells through different carriers

therapy is being developed rapidly with a significant improvement to target only cancer cells while protecting the normal cells and tissues. The use of nanoparticles, nanodrugs, nanotubes, and nanocarriers are playing significant role in nanomedicine and shows very promising role in problem solving issues in medicine, specifically to diagnose, treat, and prevent diseases [6, 7]. This chapter will explore and discuss some of the advanced methods currently used in targeted drug delivery system.

2 Exosomes as Smart Drug Carrier

Exosomes are a subpopulation of extracellular vesicles (ECV). It can be introduced as a family of nanoparticles due to their basic diameter in the range of 30–120 nm and biogenesis that it is secreted by most of the cell types of body [7, 8]. They can be isolated from several types of extracellular fluids, including saliva, amniotic fluid, blood, urine, and cerebrospinal fluid [8–10]. The delivery of drugs to target the cancer cells, which may be affected due to chemical or physical factors, was proven to significantly influence cell-to-cell communication through different signal transduction systems. It has been found that ECV plays an essential role in cell-to-cell communication by carrying their contents, including proteins, metabolites, RNAs, DNAs, and lipids [11–13]. Several researchers have reported that exosomes have been found one of the distinct cellular entities specifically capable of carrying cargos like RNA, proteins, lipids, etc., to be shared between the cells [8, 11]. Johnsen et al. [14] have reviewed well and reported that how components of successful exosome-based drug delivery can be performed. Most of the studies showed an approach for loading therapeutic cargo into exosomes by transfecting the exosome donor cell to overexpress a certain gene product that the cell will package into the exosome lumen or membrane for secretion [15, 16]. Several other studies investigated it by introducing miRNAs into exosomes using expression vectors or pre-miRNAs [15–19]. The possible pathway(s) how exosomes carry the drugs and target the cancer cells has been indicated in Fig. 1 and the computational model for the identification of affected site and recovery of cells has been explored in Fig. 2. Exosomes can be used as carrier for smart drug or biosensor, mainly due to their involvement in cell-to-cell communication [20], but also in immunological responses [21] and in rescue from apoptosis [22]. Exosome-mediated miRNA transfer plays an important role in the radiation responses and intercellular signaling pathway between irradiated cells and by stander effects [23–25].

Since exosomes can be detected in blood, urine, saliva, and other biological fluids, the identification of exosomal biomarkers related to cancer can lead to the generation of novel non-invasive

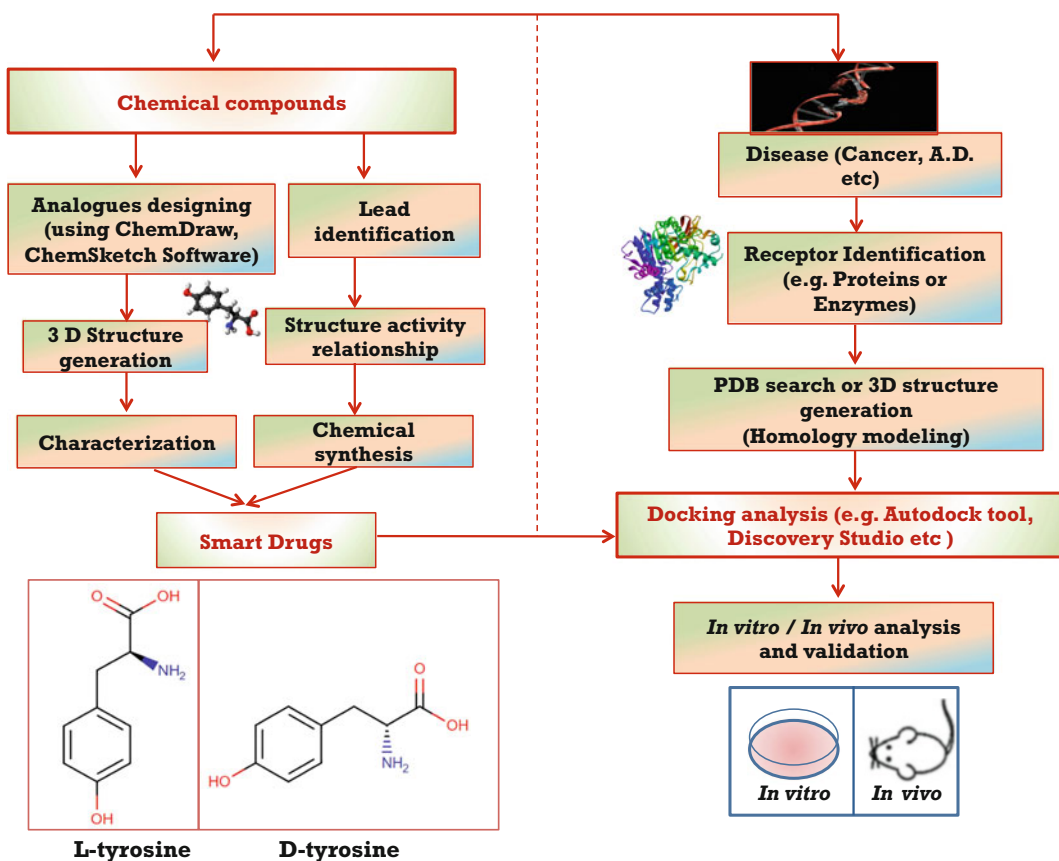


Fig. 2 Flow chart reflecting the computational approach to find out novel smart drug compounds

methods to follow mutagenic-induced changes [26]. The proposed use of exosomes as biosensors could be a novel tool for both cancer detection and therapeutic implementation. There is no understanding of how mutagens can affect the exosomes and exosome-mediated signaling between tumor cells and their microenvironment, and identifying factors affecting such communication seems critical for better diagnosis and treatment of human malignancies.

3 Liposomes in Targeting Drug Delivery

Liposomes are tidy bubbles or closed vesicles [27] and this has similar composition as like cell membrane. In 1906, Ehrlich derived the concept of a “magic bullet” in respect to liposomes for targeting tumor sites [28]. The first study on liposomes as drug carrier was initiated in 1970 [29], and later in 1974 liposomes were recommended as drug carriers in cancer chemotherapy. In the last decade, liposomal systems have well advanced in terms of improved drug delivery potential for cancer therapy [30–32]. In addition, it

has been reported that liposomes are potentially strong to target the cancer cells by delivering the drugs of low molecular weight, proteins, imaging agents, nucleic acids, and peptides [33, 34]. After the approval of small-molecule therapeutics, liposomes were investigated to deliver macromolecules such as plasmid DNA, antisense oligonucleotides, and siRNA to disease-targeted sites. Liposomes potentially encapsulate both hydrophilic and hydrophobic drugs, because their membranes are composed of natural and/or synthetic lipids or amino acid surfactants. Due to this, it has good biocompatibility and biodegradability properties and makes it an attractive carrier for drug delivery. Food and Drug Administration (FDA) has approved the liposome formulations in the form of Doxil and DaunoXome (DOX) for cancer therapy [35–37]. Doxil is a brilliant example of clinical use for anthracycline anticancer agent active against a wide variety of rigid tumors [38, 39]. For the delivery of DOX, several researchers have reported new liposomal formulations [40–43], and they are mainly used in “active loading” and “passive loading” methods, where active loading method has been reported in majority to achieve small encapsulation efficiencies of the drug. Several other formulations include Myocet [44], DaunoXome [45], Marqibo [46], and DepoCyt [47]. The exact mechanism of drug delivery through liposomes is unclear but [45] it has been suggested that the internalization of drugs like DaunoXome by cells is via endocytosis. The endolysosomal pathway has a significant advantage of targeted intracellular drug delivery because not only it does endocytosis to allow for macromolecular internalization but it also activates receptor and lysosome-specific localization. Liposomal drug delivery system is very effective and safe for cancer treatment process because it has the capability of reducing the toxic side effects of chemotherapeutic agents while enhancing their anti-tumor efficacy. The possible ways of drug initialization by liposomal system are presented in Fig. 1.

4 DNA Origami: Advancement in Drug Delivery

The delivery of drug molecules specifically to the tumor site is an exigent requirement to avoid side effects during cancer therapy [48]. Several DNA-based nanostructures, namely tetrahedral, icosahedral [49], nanotubes [50–53], squares, and triangles [54], have been developed recently for *in vitro* and *in vivo* drug delivery applications. Udomprasert and Kangsamaksin [55] reviewed that DNA Origami could be a novel biological detector of cancer which plays an important role in cancer therapy. It has a great application in medical science and biological research for drug delivery system. DNA origami has been examined as a most promising agent for drug delivery in cancer therapy [56–61] and as well as a promising candidate to serve as the next-generation drug delivery vehicle

[56, 62]. The application of DNA origami is to target the specific receptor responsible for diseases like neurodegenerative diseases where it may use receptor proteins of trigeminal sensory neurons, which are naturally designed to detect a multitude of environmental stimuli, such as electromagnetic, heat, mechanical, and chemical factors. These natural detectors are supersensitive by applying sensitizing agents (some kind of carrier) as nanopores constructed by a DNA origami technique could be used to regulate the entry of therapeutic drugs into cancer cells. DNA origami nanostructures possess abilities to enhance efficacies of chemotherapy, reduce adverse side effects, and even circumvent drug resistance. DNA origami nanostructures enhanced anticancer activities and circumvented drug resistance. Jiang and colleagues [63] reported that triangular and tubular DNA origami nanostructures with doxorubicin resulted in an increased apoptosis of doxorubicin-resistant breast cancer cells. This mainly depends on structural variations which precisely defined nanoscale shapes, uniform sizes, and obvious biocompatibility. DNA origami shows more impactful action to target the cancer cells with their size and shape variations. With shape variations and good biocompatibility, Jiang et al. [63] reported a drug delivery system based on triangular and tubular DNA origami nanostructures, whereas Halley et al. [64] synthesized a rod-like DNA origami drug carrier. Figure 1 shows the overall pathways that how DNA origami can be used as a carrier or may be directed to treat the cancer cells by targeting through drug delivery-based channel. Therefore, based on the available data, we can say that DNA origami is potentially capable of playing important roles in cancer therapies in the future.

5 Tyrosine as a Smart Drug

Tyrosine is another smart drug, which may play an important role in the modulation of growth factor signaling in cancer therapy (Fig. 3). Tyrosine kinase is a type of tumorigenic protein and has been divided into two groups as receptor tyrosine kinase (RTK) and non-receptor tyrosine kinase (NRTK). Tyrosine kinases are enzymes that catalyze the transfer of the γ phosphate group from adenosine triphosphate to target proteins. Arora and Scholar [65] have reported that activated forms of these enzymes can cause increases in tumor cell proliferation and growth, induce antiapoptotic effects, and promote angiogenesis and metastasis. Further, RTKs are membrane-spanning cell surface proteins that play critical roles in the transduction of extracellular signals to the cytoplasm [66].

The RTKs and NRTKs can be targeted and inhibited by small-molecule, tyrosine kinase inhibitors (TKIs). The inhibition of RTKs in neoplastic cells shows promising therapeutic changes by the

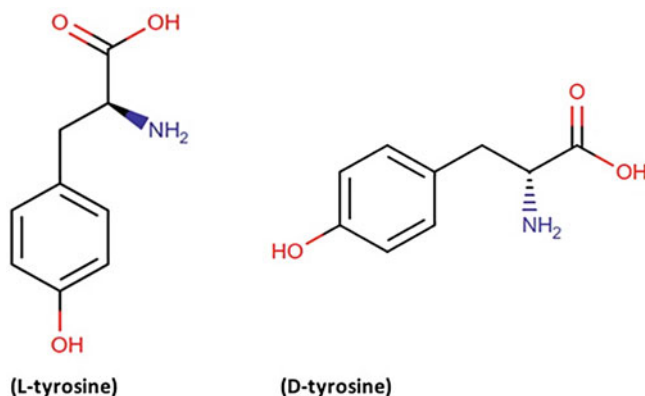


Fig. 3 2D chemical structure of L-tyrosine (DB00135) and D-tyrosine (DB03839) obtained from DrugBank Database

administration of monoclonal antibodies, which interfere RNAs, and/or small kinase inhibitors and impair cell proliferation and survival, thereby inducing arrest of cell growth and apoptosis [67–71]. The ultimate goal of tyrosine is to diagnose cancer by the drug delivery mechanism as also indicated in Figs. 1 and 2.

Shawver et al. [72] have reported that the Herceptin showed antitumor activity against breast cancer providing the proof that RTKs could be used as potential therapeutic targets in cancer therapy. Sierra et al. [73] proposed a functionality of tyrosine that during tumor progression, the hyperactivation of tyrosine kinases leads to the continuous activation of downstream signaling cascades that block cellular apoptosis, promote cellular proliferation, and increase the nutrient/waste interchange by enhancing angiogenesis.

6 Nanodrug Delivery

Nanotechnology has made a significant impact on clinical therapeutics in the last two decades [74]. There are several reports to show that nanoparticles are useful in cancer treatment and can further rely on the enhanced permeability and retention (EPR) effect caused by the leaky tumor for better drug accumulation at the tumor sites [75]. Nanomedicine for cancer therapy covers nano-sized drugs and particles which, when injected into the bloodstream, passively accumulate in tumor tissue through the leaky tumor vasculature [76–78] known as EPR effect. Nanoparticles have the advantage of targeting cancer by simply being accumulated and entrapped in tumors (passive targeting). A number of novel technologies based on nanoparticle–DNA binding and their interactions have been developed and used in molecular diagnosis, gene therapy, sensing, drug delivery, and artificial implants [79–83]. These approaches offer an opportunity for the

development of efficient and low-cost technologies for disease diagnosis and DNA detection with high sensitivity [84, 85]. Meena et al. [86] have investigated that hydroxyapatite (HAP) nanoparticles (NPs) inhibit the growth of MCF-7 breast cancer cells as well as induce apoptosis. The authors reported that HAP NPs induce the production of intracellular reactive oxygen species and activate p53, which may be responsible for DNA damage and apoptosis in a dose-dependent manner.

The structure of the nanoparticles plays an important role in targeting the tumor tissues. DNA origami or nanocarriers are one of the phenomena, which ascertain the specificity of structural importance with respect to nano-delivery materials, while protecting the drug from premature inactivation during its transport. It also protects the drug from degradation and reduces the renal clearance thereby increasing its half-life in the bloodstream. Moreover, it further augments the payload of cytotoxic drugs, improves the solubility of the insoluble drugs, and allows the controlled release of the anticancer drugs [87–89].

Generally, nanoparticles exist in three broad categories: organic, inorganic, and hybrid particles. A polymer-based nanoparticle comes into the category of organic nanoparticles. Both organic and inorganic types of nanoparticles, which may act as nanocarrier for the drug delivery system to target the cells, include carbon nanotubes, fullerene particles, metallic nanoparticles (gold and silver nanoparticles), and many metal oxide species. A combination of organic and inorganic nanomaterials such as peptide- and DNA-functionalized gold nanoparticles [90] and DNA–carbon nanotube arrays are termed as hybrid nanoparticles or engineered nanoparticles. The promising mechanism of nanodrug delivery system has been proposed in Fig. 1. Nanoparticles may help the existing state of medicine in several ways, by providing highly selective and targeted therapeutics, i.e., with increased efficacy and minimizing side effects of current therapeutics, with increased efficiency of diagnostic and prognostic tools and by impacting the development of drugs.

7 Computational Approach and Smart Drug Designing

Bioinformatics influences new drug design outline. The procedures of discovery of novel drugs by utilizing bioinformatics techniques have opened up a new avenue in drug design innovation. In order to find out the smart drug compounds, the below-mentioned methods would be helpful.

7.1 Target Disease Identification

Target distinguishing proof alone may not be adequate keeping in mind the end goal to accomplish a fruitful treatment of the disease. Bioinformatics techniques have been discovered to

computationally screen the potential targets for the drug leads in order to bind or inhibit the proteins. In the current era of computational biology, several software and tools are available to predict the specific targets like TargetHunter, ChEMBL, and many more reported in Table 1 [91–130]. The therapeutic target database (TDD) is another example by which we can identify the drug targets. It is a large annotated database of drugs, cancer drug targets, and their clinical information. Another plausibility is to discover different proteins that control the movement of the target by interacting and forming a complex [131] (Fig. 4). Recently, Jeon et al. [132] have identified 5169 proteins as putative non-drug targets by using TDD database searching.

Kesari et al. [133] have reported that the computational elucidation of melatonin in repair system induced by microwave radiation exposure is a useful tool in different types of therapeutic treatment. Melatonin is a well-known scavenger of ROS which may participate to reduce the oxidative injury. Melatonin is also a useful agent for use as a drug for the delivery to target the cells. Melatonin might use some kind of carriers, like DNA origami, nanotubes, liposomes, and exosomes, to target some specific cells. Kesari et al. [133] selected melatonin, which is reported as an acetylcholinesterase (AChE) and butyrylcholinesterase (BuChE) inhibitor used in the microwave-induced cancer treatment. The knowledge of BuChE structure is essential for understanding its high catalytic efficacy and the molecular basis for the recognition of AChE by other ACh-binding protein (AChE receptors). Therefore, we have selected this enzyme to see the interaction pattern analysis of melatonin. The structural interaction by introducing computational approaches explores the binding/inhibition pattern of melatonin with BuChE enzymes (Fig. 5).

BuChE is known as pseudocholinesterase or non-specific ChE. These ChEs are highly efficient since they are able to cleave more than 10,000 molecules of AChE per second thereby rapidly producing acetate and choline [134].

7.2 Searching of Novel Compounds

We have to distinguish and contemplate the lead compounds that have some action against a disease. These compounds give a beginning stage of refinement of the synthetic/chemical structures. There are a few diverse bioinformatics techniques to recognize the lead compound, for example, high-throughput screening (HTS), virtual screening of structural databases like ZINC [135], PubChem, PDBeChem, and many more reported in Table 1.

7.3 Identification of Receptor Biomolecules

Protein structures are a rich source of structural data about families and superfamilies. It is such distantly evolved proteins that should be perceived as they are destined to show comparable structure and capacity [136].

Table 1
Summary of docking tools and software used for in silico prediction of drugs and targets

S. No.	Tool	Function	Resource	Reference
1	BindingDB	BindingDB is a public, web-accessible database of measured binding affinities, focusing chiefly on the interactions of protein considered to be drug targets with small, drug-like molecules	https://www.bindingdb.org/	Liu et al. [91]
2	BioGRID	BioGRID is an interaction repository with data compiled through comprehensive curation efforts. Our current index is version 3.4.159 and searches 64,826 publications for 1,548,143 protein and genetic interactions, 27,785 chemical associations, and 39,028 post-translational modifications from major model organism species	https://thebiogrid.org/	Stark et al. [92]
3	canSAR	canSAR is an integrated knowledge base that brings together multidisciplinary data across biology, chemistry, pharmacology, structural biology, cellular networks, and clinical annotations and applies machine learning approaches to provide drug-discovery useful predictions	http://cansar.icr.ac.uk/	Tym et al. [93]
4	ChEMBL	ChEMBL is an Open Data database containing binding, functional, and ADMET information for a large number of drug-like bioactive compounds	https://www.ebi.ac.uk/chembl/	Gaulton et al. [94]
5	TargetHunter	TargetHunter therefore provides a promising alternative to bridge the knowledge gap between biology and chemistry, and significantly boost the productivity of chemogenomics researchers for <i>in silico</i> drug design and discovery	http://www.cbligand.org/TargetHunter	Wang et al. [95]

(continued)

Table 1
(continued)

S. No.	Tool	Function	Resource	Reference
6	DINIES	DINIES (Drug–target Interaction Network Inference Engine based on Supervised Analysis) enables us to predict potential interactions between drug molecules and target proteins, based on drug data and omics-scale protein data. The users can use any data as the input, as long as they are represented as the tab-delimited matrices or profiles	http://www.genome.jp/tools/dinies/	Yamanishi et al. [96]
7	ECOdrug	The ECOdrug database contains information on the Evolutionary Conservation Of human Drug targets in over 600 eukaryotic species. The interface allows users to identify human drug targets to 1000+ legacy drugs and explore integrated orthologue predictions for the drug targets, transparently showing the confidence in the predictions both across methods and taxonomic groups	http://www.ecodrug.org/	Verbruggen et al. [97]
8	HitPick	A web server for identification of hits in high-throughput chemical screenings and prediction of their molecular targets	http://mips.helmholtz-muenchen.de/proj/hitpick	Liu et al. [98]
9	iDrug-Target	The web server for predicting the interaction between GPCRs and drugs in cellular networking	www.jci-bioinfo.cn/iDrug-Target/	Xuan et al. [99]
10	idTarget	A web server for identifying biomolecular targets of small chemical molecules with robust scoring functions and a divide-and-conquer docking approach. Identification of biomolecular targets of small	idtarget.rcas.sinica.edu.tw/	Wang et al. [100]

(continued)

Table 1
(continued)

S. No.	Tool	Function	Resource	Reference
		chemical molecules is essential for unraveling their underlying causes of actions at the molecular level		
11	IUPHAR/BPS	Provides expert-curated molecular interactions between successful and potential drugs and their targets in the human genome	www.guidetopharmacology.org/	Southan et al. [101]
12	Open Targets	Open Targets is a public–private initiative to generate evidence on the validity of therapeutic targets based on genome-scale experiments and analysis	https://www.opentargets.org/	Koscielny et al. [102]
13	PharmMapper	An updated integrated pharmacophore matching platform with statistical method for potential target identification	lilab.ecust.edu.cn/pharmmapper/index.php	Wang et al. [103]
14	SuperDRUG2	Collection of drugs (containing 4587 active pharmaceutical ingredients) which include small molecules, biological products, and other drugs. The database is intended to serve as a one-stop resource providing data on: chemical structures, regulatory details, indications, drug targets, side effects, physicochemical properties, pharmacokinetics, and drug–drug interactions	cheminfo.charite.de/superdrug2	Siramshetty et al. [104]
15	SwissTargetPrediction	Predict the targets of a small molecule. Using a combination of 2D and 3D similarity measures, it compares the query molecule to a library of 280,000 compounds active on more than 2000 targets of five different organisms	www.swisstargetprediction.ch/	Gfeller et al. [105]

(continued)

Table 1
(continued)

S. No.	Tool	Function	Resource	Reference
16	TargetNet	Drug–target interactions (DTIs) are central to current drug-discovery processes and public health fields. Analyzing the DTI profiling of the drugs helps to infer drug indications, adverse drug reactions, drug–drug interactions, and drug mode of actions	http://targetnet.scbdd.com	Yao et al. [106]
17	TTD	Therapeutic Target Database is a resource for facilitating bench-to-clinic research of targeted therapeutics. Current coverage for searches based on similarities to the input query includes interconnected information on 3100 target and 34,000 drugs and drug-like compounds, consisting in part of 2500 approved drugs and 18,900 investigational agents	http://xin.cz3.nus.edu.sg/group/ttd/ttd.asp	Chen et al. [107]
<i>Ligand database</i>				
1	PubChem	Collect information on chemical structures, identifiers, chemical and physical properties, biological activities, patents, health, safety, toxicity data, and many others	https://pubchem.ncbi.nlm.nih.gov/	Kim et al. [108]
2	ChemSpider	<i>ChemSpider</i> is a free chemical structure database providing fast text and structure search access to over 63 million structures from hundreds of data sources	http://www.chemspider.com/	Pence and Williams [109]
3	DrugBank	The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug data with comprehensive drug target information	https://www.drugbank.ca/	Wishart et al. [110]

(continued)

Table 1
(continued)

S. No.	Tool	Function	Resource	Reference
4	NCI	The purpose of this service is to provide structures, data, tools, programs, and other useful information to the public. In addition to current and former members of the CADD Group, many others, both individuals and companies or their representatives, have contributed to these services	https://cactus.nci.nih.gov/ncidb2.2/	Voigt et al. [111]
5	ZINC	A free database of commercially available compounds for virtual screening. ZINC contains over 35 million purchasable compounds in ready-to-dock, 3D formats	http://zinc.docking.org/	Irwin et al. [112]
6	PDBChem	Dictionary of chemical components (ligands, small molecules, and monomers) referred to in PDB entries and maintained by wwPDB. It provides comprehensive search facilities for finding a particular component, or determining components in structure entries or vice versa	http://www.ebi.ac.uk/pdbe-srv/pdbechem/	Dimitropoulos et al. [113]
7	PDBbind	The PDBbind database is designed to provide a collection of experimentally measured binding affinity data (K_d , K_i , and IC_{50}) exclusively for the protein–ligand complexes available in the Protein Data Bank (PDB).	http://sw16.im.med.umich.edu/databases/pdbbind/index.jsp	Wang et al. [114, 115]
8	SuperDRUG2	SuperDRUG2 database is a unique, one-stop resource for approved/ marketed drugs, containing more than 4600 active pharmaceutical ingredients. Also, it contains information related to drugs with regulatory details, chemical structures (2D and	http://cheminfo.charite.de/superdrug2/	Siramshetty et al. [104] and Goede et al. [116]

(continued)

Table 1
(continued)

S. No.	Tool	Function	Resource	Reference
		3D), dosage, biological targets, physicochemical properties, external identifiers, side effects, and pharmacokinetic data		
<i>Pathways database</i>				
1	BRENDA	BRENDA is an information system representing one of the most comprehensive enzyme repositories	https://www.brenda-enzymes.org/	Schomburg et al. [117]
2	Reactome	REACTOME is an open-source, open access, manually curated, and peer-reviewed pathway database. Our goal is to provide intuitive bioinformatics tools for the visualization, interpretation, and analysis of pathway knowledge to support basic and clinical research, genome analysis, modeling, systems biology, and education	https://reactome.org/	Fabregat et al. [118] and Milacic et al. [119]
3	KEGG	KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism, and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies	http://www.genome.jp/kegg/	Kanehisa and Goto [120]
4	PANTHER	PANTHER Pathway consists of over 177, primarily signaling, pathways, each with subfamilies and protein sequences mapped to individual pathway components	http://www.pantherdb.org/pathway/	Mi et al. [121]

(continued)

Table 1
(continued)

S. No.	Tool	Function	Resource	Reference
<i>Docking and drug design tools</i>				
1	AutoDock Vina	AutoDock Vina is an open-source program for doing molecular docking	http://vina.scripps.edu/	Trott and Olson [122]
2	AutoDock	AutoDock is a suite of automated docking tools. It is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure	http://autodock.scripps.edu/	Morris et al. [141]
3	DockingServer	DockingServer offers a web-based, easy to use interface that handles all aspects of molecular docking from ligand and protein setup	https://www.dockingserver.com/web	Bikadi and Hazai [123]
4	Click-Docking	Predicts the binding orientation and affinity of a ligand to a target	https://mcule.com/apps/1-click-docking/	NA
5	GLID	Glide reliably finds the correct binding modes for a large set of test cases. It outperforms other docking programs in achieving lower RMS deviations from native co-crystallized structures	https://www.schrodinger.com/glide	Halgren et al. [124]
6	FlexX	Best enrichment tool for structure-based drug design	https://www.biosolveit.de/FlexX/	Rarey et al. [125]
7	GOLD	GOLD is a program for calculating the docking modes of small molecules in protein-binding sites and is provided as part of the GOLD Suite, a package of programs for structure visualization and manipulation (Hermes), for protein–ligand docking (GOLD) and for post-processing (GoldMine) and visualization of docking results	https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/	Jones et al. [126]

(continued)

Table 1
(continued)

S. No.	Tool	Function	Resource	Reference
8	ICM-DOCK	Provides a unique set of tools for accurate individual ligand-protein docking, peptide-protein docking, and protein-protein docking	http://www.molsoft.com/docking.html	Abagyan et al. [127]
<i>Physicochemical properties of prediction software</i>				
1	PhysChem and ADME-TOX	The platform offers a uniform interface for the prediction of ADME, toxicological, and physicochemical property endpoints and offers a single interface for the analysis and interpretation of predicted data	http://www.acdlabs.com/products/percepta/physchem_adme_tox/	ACD [128]
2	SwissADME	A free web tool to evaluate pharmacokinetics, drug-likeness, and medicinal chemistry friendliness of small molecules	www.swissadme.ch/	Daina et al. [129]
3	ChemAxon	Physicochemical properties of the drug molecule	https://chemaxon.com/products/calculators-and-predictors	NA
4	Cyprotex	Cyprotex offers a range of different physicochemical profiling screens including turbidimetric and thermodynamic solubility, chemical stability, pK_a , and lipophilicity ($\log P$ and $\log D$). Cyprotex has partnered with Sirius Analytical to provide pK_a , $\log P$, and dissolution studies using their SiriusT3 technology	http://www.cyprotex.com/physicochemical-profiling	Cyprotex Discovery Ltd. (CDL) [130]

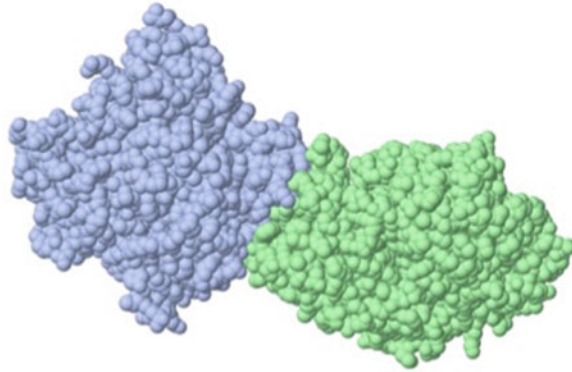


Fig. 4 Schematic diagram shows two proteins whose complexes were obtained from the HDock server (<http://hdock.phys.hust.edu.cn/>) after performing protein–protein interaction

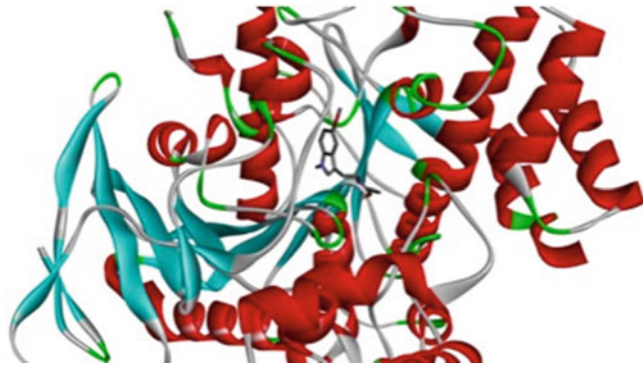


Fig. 5 3D molecular interaction visualization of melatonin with butyrylcholinesterase (BuChE). Graphics generated by Discovery Studio Visualizer 4.5

If we have a better idea after analysis of the available data related to disease that a drug must tie to a specific spot on a specific protein or nucleotide, at that point a drug can be carefully fit to tie at that site. Bioinformaticians regularly demonstrated a few distinct methods like network and pathway analyses using KEGG (Kyoto Encyclopedia of Genes and Genome) database, Reactome Database, BRENDA database, and many others (Table 1). This could also test a huge number of compounds from a database that have accessible structures. The study of Klahan et al. [137] shows that authors have identified 86 differentially expressed genes, including 37 downregulated genes and 49 upregulated genes in lymphovascular invasion-positive patients using pathways analysis method. Recently, Zhu et al. [138] have investigated miR-542-5p as a predictive biomarker and potential target for therapy of breast cancer patients after deep mining of the Gene Expression Omnibus (GEO) database.

7.4 Rational Drug Design Tools

Rational drug design is utilized as a part of the biopharmaceutical research to find out the possibilities to develop a new drug compound [139]. It utilizes an assortment of computational strategies to distinguish novel compounds, plan drug compounds for selectivity, adequacy, and safety, and form drug compounds into clinical trial objects. These strategies fall into a few common classifications—structure-based drug design, ligand-based drug design outline, and homology modeling approach, relying upon how much data is accessible about drug targets and potential novel drug compounds [140]. The software packages like AutoDock and AutoDock Vina tool provided by Molecular Graphics Laboratory (MGL) of the Scripps Research Institute can perform molecular interaction analysis between biomacromolecules and drug compounds or chemical compounds [122, 141]. Also, online molecular interaction tools are available where researchers can easily perform molecular interaction analysis, for example, Docking Server. Many commercial software packages are available for the complete relational drug design, e.g., FlexX, GLID, GOLD, ICM-Dock, Lead Finder, etc. (Table 1). The *in silico* study done by Gao et al. [142] reveals that the selected compound M7594_0037 in their computational analysis exhibited potent anticancer activities against HeLa, A549, and MCF-7 cell lines.

7.5 Drug Compound Refinement

The developed leads compounds using computational and wet lab methods for inspecting the atomic structures to figure out which perspectives are in charge of both the drug compound action and the adverse effects. Quantitative Structure–Activity Relationships (QSAR) can be utilized to discover which part(s) of the molecule associates well with the drug activity or adverse symptoms [143]. The identification of useful functional group(s) in the compound is important keeping in mind the end goal to refine the drug leads.

7.6 Detection of Physicochemical Properties of Drugs

Physicochemical properties of a drug, for example, lipophilicity, dissolvability, pK_a , hydrogen holding, and porousness, importantly affect its pharmacokinetic properties and metabolic destiny in the body [144]. The capacity to get a drug to the right place of the molecular system is a critical factor in its strength. In a more refined way, there is a persistent trade of data between the analysts doing QSAR and ADME-Tox studies, synthesis, and testing [145]. Some computational methods are eligible to find out the physicochemical properties of compounds like SwissADME, ChemAxon, Cyprotex, etc. (Table 1). These methods are utilized regularly and exceptionally effective since they may not depend on the biomolecular premise of the disease which can be extremely hard to decide.

7.7 Drug Validation

Once a drug compound has appeared to be viable by an underlying assay procedure, vigorous testing must be conducted prior to be

given to human patients [146]. Preclinical animal model testing is essential at this stage. Finally, the drug compounds, which are deemed appropriate at this stage, are further tested to clinical trials. In the clinical trials, additional effects might appear and human measurements of dosages are finally determined [147].

8 Multitarget Drug Design

In addition to the above discussions on networking of smart drugs and targets, we mention here about another advanced framework of multitarget drug design, which plays an important role in several cancer or neurological diseases like Alzheimer's disease [148]. There are several methods which may participate to protect the cells by multitargeting with designed drugs. DNA origami is one of the best examples of a designed drug target. However, a positive approach of computational methods in multitarget drug design is surplus for future research [149]. Target-based drug discovery has successfully produced target-specific medicines for cancer, inflammation, diabetes, and central nervous system disorders. In addition, the multitarget drug-design strategy has emerged to address the complexity of the disease networks. Zhang et al. [149] reported that the complexity of disease-related molecular networks, which is robust with many redundant pathways, comes from systemic pathways of multidrug targets. The developments in single-target drugs may not always induce the desired effects to the entire biological system even if they successfully inhibit or activate a specific target [150]. Moreover, the intricacy of the ongoing incurable pathologies has indicated that such single-target drugs are inadequate to accomplish desirable therapeutic effects in complex diseases [151, 152]. More recently, the multitarget drug-design concept has been proposed by researchers particularly for the complex diseases. Network-based approaches have been largely used to integrate, analyze, and visualize the available knowledge on a disease. Moreover, combining different types of information (e.g., drug interactions, and biological signaling pathways) into network models may help to better understand the molecular mechanism of drug actions and to investigate potential drug therapies. Vatali et al. [153] suggested an approach to the case of Triple Negative Breast Cancer (TNBC), a subtype of breast cancer whose biology is poorly understood and that lacks of specific molecular targets. Their "in-silico" findings have been confirmed by a number of *in vitro* experiments, whose results demonstrated the ability of the method to select candidates for drug repurposing. Therefore, it can be summarized that the multitargeted drug designing has great future perspectives towards deadly and complex diseases.

9 Conclusion and Future Prospective

The huge level of information created straightforwardly by the drug-discovery process that have turned out to be openly accessible by the scientific community joined with the disease-based information. Through an implication of database or organization, computational methodologies connecting chemical and molecular level of disease-based information have been explored. Imaginative bioinformatics approaches are as of now affecting the disclosure, preclinical, and clinical periods of the drug revelation process. Furthermore, the difficulties faced by the pharmaceutical industry imply that it is getting noticeably urgent to put additional resources into the bioinformatics assets required to help and facilitate translational drug discovery and revelation. Methodologies include:

1. The advancement of databases and information stockrooms that can file, keep up, and incorporate a lot of drug information disclosure and biomedical information as of now being created.
2. The improvement of vigorous data mining algorithms to empower the examination of huge and complex datasets.
3. Advancement of software to empower the in vivo or in vitro drug discovery for researchers to effortlessly get to and decipher this information.
4. Formal and casual administration systems that empower bioinformaticians to connect up and gain from each other.

These kind of attempts will empower a superior comprehension of how we can utilize genomics and other “omic” ways to deal with group of disease, enhance, analyze, and educate new ways to deal with new chemical compounds or smart drug repositioning. They will enable us to recognize infection and disease biomarkers and hereditary variations which connect well with patients’ data analysis results, and utilize them to enhance remedial methodologies.

References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010) GLOBOCAN 2008, Cancer incidence and mortality worldwide: IARC CancerBase No. 10 [Internet 11 Jan 2010, date last accessed]. *Int J Cancer* 127:2893–2917
2. Siegel RL, Miller KD, Jemal A (2017) Cancer statistics, 2017. *CA Cancer J Clin* 67(1):7–30
3. Derakhshandeh K, Azandaryani AH (2016) Active-targeted nanotherapy as smart cancer treatment. In: Sezer AD (ed) *Smart drug delivery system*. InTech, Croatia. <https://doi.org/10.5772/61791> ISBN: 978-953-51-2247-0, 396 p
4. Dhanasekaran S, Chopra S (2016) Getting a handle on smart drug delivery systems—a comprehensive view of therapeutic targeting strategies. In: Sezer AD (ed) *Smart drug delivery system*. InTech, Croatia, pp 31–62. <https://doi.org/10.5772/61388> ISBN: 978-953-51-2247-0, Chapter 2
5. Archakov AI (2010) Nanobiotechnologies in medicine: nanodiagnosics and nanodrugs. *Biochem Mosc* 4:2–14

6. Feng SS (2006) New-concept chemotherapy by nanoparticles of biodegradable polymers: where are we now? *Nanomedicine (Lond)* 1:297–309
7. Record M, Subra C, Silvente-Poirot S, Poirot M (2011) Exosomes as intercellular signalosomes and pharmacological effectors. *Biochem Pharmacol* 81:1171–1182
8. Vlasov AV, Magdaleno S, Setterquist R, Conrad R (2012) Exosomes: current knowledge of their composition, biological functions, and diagnostic and therapeutic potentials. *Biochim Biophys Acta* 1820:940–948
9. Witwer KW, Buzás EI, Bemis LT, Bora A, Lässer C, Lötvald J et al (2013) Standardization of sample collection, isolation and analysis methods in extracellular vesicle research. *J Extracell Vesicles* 2:18389
10. Jia S, Zocco D, Samuels ML, Chou MF, Chammas R, Skog J et al (2014) Emerging technologies in extracellular vesicle-based molecular diagnostics. *Expert Rev Mol Diagn* 14(3):307–321
11. El Andaloussi S, Mäger I, Breakefield XO, Wood MJA (2013) Extracellular vesicles: biology and emerging therapeutic opportunities. *Nat Rev Drug Discov* 12:347–357
12. De Toro J, Herschlik L, Waldner C, Mongini C (2015) Emerging roles of exosomes in normal and pathological conditions: new insights for diagnosis and therapeutic applications. *Front Immunol* 4:203
13. Krause M, Samoylenko A, Vainio SJ (2015) Exosomes as renal inductive signals in health and disease, and their application as diagnostic markers and therapeutic agents. *Front Cell Dev Biol* 3:65
14. Johnsen KB, Gudbergsson JM, Skov MN, Pilgaard L, Moos T, Duroux M (2014) A comprehensive overview of exosomes as drug delivery vehicles—endogenous nanocarriers for targeted cancer therapy. *Biochim Biophys Acta* 1846(1):75–87
15. Ohno SI, Takanashi M, Sudo K, Ueda S, Ishikawa A, Matsuyama N et al (2012) Systemically injected exosomes targeted to EGFR deliver antitumor microRNA to breast cancer cells. *Mol Ther* 21:185–191
16. Katakowski M, Buller B, Zheng X, Lu Y, Rogers T, Osobamiro O et al (2013) Exosomes from marrow stromal cells expressing miR-146b inhibit glioma growth. *Cancer Lett* 335:201–204
17. Pan Q, Ramakrishnaiah V, Henry S, Fouraschen S, de Ruyter PE, Kwekkeboom J et al (2012) Hepatic cell-to-cell transmission of small silencing RNA can extend the therapeutic reach of RNA interference (RNAi). *Gut* 61:1330–1339
18. Kosaka N, Iguchi H, Yoshioka Y, Hagiwara K, Takeshita F, Ochiya T (2012) Competitive interactions of cancer cells and normal cells via secretory microRNAs. *J Biol Chem* 287:1397–1405
19. Chen L, Charrier A, Zhou Y, Chen R, Yu B, Agarwal K et al (2013) Epigenetic regulation of connective tissue growth factor by microRNA-214 delivery in exosomes from mouse or human hepatic stellate cells. *Hepatology* 59:1118–1129
20. Bang C, Thum T (2012) Exosomes: new players in cell-cell communication. *Int J Biochem Cell Biol* 44:2060–2064
21. Robbins PD, Morelli AE (2014) Regulation of immune responses by extracellular vesicles. *Nat Rev Immunol* 14:195–208
22. Miksa M, Wu R, Dong W, Komura H, Amin D, Ji Y, Wang Z, Wang H, Ravikumar TS, Tracey KJ, Wang P (2009) Immature dendritic cell-derived exosomes rescue septic animals via milk fat globule epidermal growth factor-factor VII. *J Immunol* 183:5983–5990
23. Koturbash I, Zemp FJ, Kutanzi K, Luzhna L, Loree J, Kolb B, Kovalchuk O (2008) Sex-specific microRNAome deregulation in the shielded bystander spleen of cranially exposed mice. *Cell Cycle* 7:1658–1667
24. Xu S, Wang J, Ding N, Hu W, Zhang X, Wang B, Hua J, Wei W, Zhu Q (2015) Exosome-mediated microRNA transfer plays a role in radiation-induced bystander effect. *RNA Biol* 12:1355–1363
25. Dickey JS, Zemp FJ, Martin OA, Kovalchuk O (2011) The role of miRNA in the direct and indirect effects of ionizing radiation. *Radiat Environ Biophys* 50:491–499
26. Magras IN, Xenos TD (1997) Radiation-induced changes in the prenatal development of mice. *Bioelectromagnetics* 18:455–461
27. Hong JS, Vreeland WN, DePaoli Lacerda SH, Locascio LE, Gaitan M, Raghavan SR (2008) Liposome-templated supramolecular assembly of responsive alginate nanogels. *Langmuir* 24:4092–4096
28. Strebhardt K, Ullrich A (2008) Paul Ehrlich's magic bullet concept: 100 years of progress. *Nat Rev Cancer* 8(6):473–480
29. Gregoriadis G, Ryman BE (1972) Fate of protein-containing liposomes injected introrats. An approach to the treatment of storage diseases. *Eur J Biochem* 24:485–491
30. Huwyler J, Drewe J, Krähenbühl S (2008) Tumor targeting using liposomal antineoplastic drugs. *Int J Nanomedicine* 3(1):21–29

31. Elbayoumi T, Torchilin V (2010) Current trends in liposome research. In: Weissig V (ed) *Liposomes*. Humana Press, Totowa, NJ, pp 1–27
32. Sawant RR, Torchilin VP (2010) Liposomes as ‘smart’ pharmaceutical nanocarriers. *Soft Matter* 6(17):4026–4044
33. Laouini A et al (2012) Preparation, characterization and applications of liposomes: state of the art. *J Colloid Sci Biotechnol* 1 (2):147–168
34. Lian T, Ho RJ (2001) Trends and developments in liposome drug delivery systems. *J Pharm Sci* 90(6):667–680
35. Hofheinz RD (2005) Liposomal encapsulated anti-cancer drugs. *Anti-Cancer Drugs* 16 (7):691–707
36. Immordino ML, Dosio F, Cattel L (2006) Stealth liposomes: review of the basic science, rationale, and clinical applications, existing and potential. *Int J Nanomed* 1(3):297–315
37. Barenholz YC (2012) Doxil[®]—the first FDA-approved nano-drug: lessons learned. *J Control Release* 160(2):117–134
38. Herman EH, Rahman A, Ferrans VJ, Vick JA, Schein PS (1983) Prevention of chronic doxorubicin cardiotoxicity in beagles by liposomal-encapsulation. *Cancer Res* 43:5427–5432
39. van Hoesel QG, Steerenberg PA, Crommelin DJ, van Dijk A, van Oort W, Klein S, Douze JM, de Wildt DJ, Hillen FC (1984) Reduced cardiotoxicity and nephrotoxicity with preservation of antitumor activity of doxorubicin entrapped in stable liposomes in the LOU/M Wsl rat. *Cancer Res* 44:3698–3705
40. Silva J et al (2011) DODAB:monoolein-based lipoplexes as non-viral vectors for transfection of mammalian cells. *Biochim Biophys Acta Biomembr* 1808(10):2440–2449
41. Silva JPN et al (2014) Tunable pDNA/DODAB:MO lipoplexes: the effect of incubation temperature on pDNA/DODAB:MO lipoplexes structure and transfection efficiency. *Colloids Surf B: Biointerfaces* 121:371–379
42. da Rocha MEB (2014) Desenvolvimento de uma formulação lipossomal para entrega de um fármaco anticancerígeno. Universidade do Minho
43. Gabizon A, Shmeeda H, Barenholz Y (2003) Pharmacokinetics of pegylated liposomal doxorubicin: review of animal and human studies. *Clin Pharmacokinet* 42:419–436
44. Swenson C, Perkins W, Roberts P, Janoff A (2001) Liposome technology and the development of Myocet (liposomal doxorubicin citrate). *Breast* 10:1–7
45. Forssen EA (1997) The design and development of DaunoXome[®] for solid tumor targeting in vivo. *Adv Drug Deliv Rev* 24:133–150
46. Silverman JA, Deitcher SR (2013) Marqibo[®] (vincristine sulfate liposome injection) improves the pharmacokinetics and pharmacodynamics of vincristine. *Cancer Chemother Pharmacol* 71:555–564
47. Angst MS, Drover DR (2006) Pharmacology of drugs formulated with DepoFoam: a sustained release drug delivery system for parenteral administration using multivesicular liposome technology. *Clin Pharmacokinet* 45:1153–1176
48. Kumar V, Palazzolo S, Bayda S, Corona G, Toffoli G, Rizzolio F (2016) DNA nanotechnology for cancer therapy. *Theranostics* 6 (5):710–725
49. Zhang C, Su M, He Y et al (2008) Conformational flexibility facilitates self-assembly of complex DNA nanostructures. *Proc Natl Acad Sci U S A* 105:10665–10669
50. Yan H, Park SH, Finkelstein G et al (2003) DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science* 301:1882–1884
51. Rothmund PWK, Ekani-Nkodo A, Papadakis N et al (2004) Design and characterization of programmable DNA nanotubes. *J Am Chem Soc* 126:16344–16352
52. Liu D, Park SH, Reif JH et al (2004) DNA nanotubes self-assembled from triple-crossover tiles as templates for conductive nanowires. *Proc Natl Acad Sci U S A* 101:717–722
53. Mathieu F, Liao S, Kopatsch J et al (2005) Six-helix bundles designed from DNA. *Nano Lett* 5:661–665
54. Rothmund PWK (2006) Folding DNA to create nanoscale shapes and patterns. *Nature* 440:297–302
55. Udomprasert A, Kangsamaksin T (2017) DNA origami applications in cancer therap. *Cancer Sci* 108:1535–1543
56. Chhabra R, Sharma J, Liu Y, Rinker S, Yan H (2010) DNA self-assembly for nanomedicine. *Adv Drug Deliv Rev* 62:617–625
57. Zhao YX, Shaw A, Zeng X, Benson E, Nyström AM, Högberg B (2012) DNA origami delivery system for cancer therapy with tunable release properties. *ACS Nano* 6:8684–8691
58. Zhu G, Zheng J, Song E, Donovan M, Zhang K, Liu C et al (2013) Self-assembled, aptamer-tethered DNA nanotrains for targeted transport of molecular drugs in cancer

- theranostics. *Proc Natl Acad Sci U S A* 110:7998–8003
59. Zhang Q, Jiang Q, Li N, Dai L, Liu Q, Song L et al (2014) DNA origami as an *in vivo* drug delivery vehicle for cancer therapy. *ACS Nano* 8:6633–6643
 60. Zhuang X, Ma X, Xue X, Jiang Q, Song L, Dai L et al (2016) A photosensitizer-loaded DNA origami nanosystem for photodynamic therapy. *ACS Nano* 10:3486–3495
 61. Jiang Q, Shi Y, Zhang Q, Li N, Zhan P, Song L et al (2015) A self-assembled dna origami-gold nanorod complex for cancer theranostics. *Small* 11:5134–5141
 62. Pinheiro AV, Han D, Shih WM, Yan H (2011) Challenges and opportunities for structural DNA nanotechnology. *Nat Nanotechnol* 6:763–772
 63. Jiang Q, Song C, Nangreave J et al (2012) DNA origami as a carrier for circumvention of drug resistance. *J Am Chem Soc* 134:13396–13403
 64. Halley PD, Lucas CR, McWilliams EM, Webber MJ, Patton RA, Kural C et al (2016) Daunorubicin-loaded DNA origami nanostructures circumvent drug-resistance mechanisms in a leukemia model. *Small* 12:308–320
 65. Arora A, Scholar EM (2009) Role of tyrosine kinase inhibitors in cancer therapy. *J Pharmacol Exp Ther* 315(3):971–979
 66. Pawson T (2002) Regulation and targets of receptor tyrosine kinases. *Eur J Cancer* 38 (Suppl 5):S3–S10
 67. Corso S, Migliore C, Ghiso E, De Rosa G, Comoglio PM, Giordano S (2008) Silencing the MET oncogene leads to regression of experimental tumors and metastases. *Oncogene* 27:684–693
 68. Normanno N, Bianco C, Damiano V, de Angelis E, Selvam MP, Grassi M, Magliulo G, Tortora G, Bianco AR, Mendelsohn J (1996) Growth inhibition of human colon carcinoma cells by combinations of anti-epidermal growth factor-related growth factor antisense oligonucleotides. *Clin Cancer Res* 2:601–609
 69. McDermott U, Sharma SV, Dowell L, Greninger P, Montagut C, Lamb J, Archibald H, Raudales R, Tam A, Lee D (2007) Identification of genotype-correlated sensitivity to selective kinase inhibitors by using high-throughput tumor cell line profiling. *Proc Natl Acad Sci U S A* 104:19936–19941
 70. Vigna E, Pacchiana G, Mazzone M, Chiriaco C, Fontani L, Basilico C, Pennacchietti S, Comoglio PM (2008) “Active” cancer immunotherapy by anti-Met antibody gene transfer. *Cancer Res* 68:9176–9183
 71. Montagut C, Settleman J (2009) Targeting the RAF-MEK-ERK pathway in cancer therapy. *Cancer Lett* 283:125–134
 72. Shawver LK, Slamon D, Ullrich A (2002) Smart drugs: tyrosine kinase inhibitors in cancer therapy. *Cancer Cell* 1(2):117–123
 73. Sierra JR, Cepero V, Giordano S (2010) Molecular mechanisms of acquired resistance to tyrosine kinase targeted therapy. *Mol Cancer* 9:75
 74. Obaid G, Broekgaarden M, Bulin A-L et al (2016) Photonanomedicine: a convergence of photodynamic therapy and nanotechnology. *Nanoscale* 8(25):12471–12503
 75. Su YC, Burnouf PA, Chuang KH, Chen BM, Cheng TL, Roffler SR (2017) Conditional internalization of PEGylated nanomedicines by PEG engagers for triple negative breast cancer therapy. *Nat Commun* 8:15507
 76. Jain RK, Stylianopoulos T (2010) Delivering nanomedicine to solid tumors. *Nat Rev Clin Oncol* 7:653–664
 77. Petros RA, DeSimone JM (2010) Strategies in the design of nanoparticles for therapeutic applications. *Nat Rev Drug Discov* 9:615–627
 78. Siegler EL, Kim YJ, Wang P (2016) Nanomedicine targeting the tumor microenvironment: therapeutic strategies to inhibit angiogenesis, remodel matrix, and modulate immune responses. *J Cell Immunother* 2 (2):69–78
 79. Waren CWC, Nie SM (1998) Quantum dot bioconjugates for ultrasensitive nonisotopic detection. *Science* 281:2016–2018
 80. Curtis A, Wilkinson C (2001) Nanotechniques and approaches in biotechnology. *Trends Biotechnol* 19:97–101
 81. Sachlos E, Gotoro D, Czernuszka JT (2006) Collagen scaffolds reinforced with biomimetic composite nano-sized carbonate-substituted hydroxyapatite crystals and shaped by rapid prototyping to contain internal microchannels. *Tissue Eng* 12:2479–2487
 82. Vaseashta A, Dimova-Malinovska D (2005) Nanostructured and nanoscale devices, sensors and detectors. *Sci Technol Adv Mater* 6:312–318
 83. Yao VJ, D’Angelo S, Butler KS et al (2016) Ligand-targeted theranostic nanomedicines against cancer. *J Control Release* 240:267–286
 84. Yezhelyev VM, Gao X, Xing Y et al (2006) Emerging use of nanoparticles in diagnosis

- and treatment of breast cancer. *Lancet* 7:657–667
85. Buxton DB, Lee SC, Wickline SA, Ferrari M (2003) Recommendations of the National Heart, Lung, and Blood Institute Nanotechnology Working Group. *Circulation* 108:2737–2742
 86. Meena R, Kesari KK, Rani M, Paulraj R (2012) Effects of hydroxyapatite nanoparticles on proliferation and apoptosis of human breast cancer cells (MCF-7). *J Nanopart Res* 14:712
 87. Danhier F, Feron O, Pr at V (2010) To exploit the tumor microenvironment: passive and active tumor targeting of nanocarriers for anti-cancer drug delivery. *J Control Release* 148:135–146
 88. Davis ME, Chen Z, Shin DM (2008) Nanoparticle therapeutics: an emerging treatment modality for cancer. *Nat Rev Drug Discov* 7:771–782
 89. Peer D, Karp JM, Hong S, Farokhzad OC, Margalit R, Langer R (2007) Nanocarriers as an emerging platform for cancer therapy. *Nat Nanotechnol* 2:751–760
 90. Lin YW, Liu CW, Chang HT (2009) DNA functionalized gold nanoparticles for bioanalysis. *Anal Methods* 1:14–24
 91. Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK (2007) BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Res* 35: D198–D201
 92. Stark C, Breitkreutz B-J, Reguly T, Boucher L, Breitkreutz A, Tyers M (2006) BioGRID: a general repository for interaction datasets. *Nucleic Acids Res* 34:D535–D539
 93. Tym JE, Mitsopoulos C, Coker EA, Razaz P, Schierz AC, Antolin AA, Al-Lazikani B (2016) canSAR: an updated cancer research and drug discovery knowledgebase. *Nucleic Acids Res* 44:D938–D943
 94. Gaulton A, Hersey A, Nowotka M, Bento AP, Chambers J, Mendez D, Leach AR (2017) The ChEMBL database in 2017. *Nucleic Acids Res* 45:D945–D954
 95. Wang L, Ma C, Wipf P, Liu H, Su W, Xie X-Q (2013) TargetHunter: an in silico target identification tool for predicting therapeutic potential of small organic molecules based on chemogenomic database. *AAPS J* 15:395–406
 96. Yamanishi Y, Kotera M, Moriya Y, Sawada R, Kanehisa M, Goto S (2014) DINIES: drug–target interaction network inference engine based on supervised analysis. *Nucleic Acids Res* 42:W39–W45
 97. Verbruggen B, Gunnarsson L, Kristiansson E,  sterlund T, Owen SF, Snape JR, Tyler CR (2018) ECOdrug: a database connecting drugs and conservation of their targets across species. *Nucleic Acids Res* 46:D930–D936
 98. Liu X, Vogt I, Haque T, Campillos M (2013) HitPick: a web server for hit identification and target prediction of chemical screenings. *Bioinformatics* 29:1910–1912 pmid:23716196
 99. Xiao X, Min JL, Lin WZ, Liu Z, Cheng X, Chou KC (2015) iDrug-Target: predicting the interactions between drug compounds and target proteins in cellular networking via benchmark dataset optimization approach. *J Biomol Struct Dyn* 33:2221–2233
 100. Wang JC, Chu PY, Chen CM, Lin JH (2012) idTarget: a web server for identifying protein targets of small chemical molecules with robust scoring functions and a divide-and-conquer docking approach. *Nucleic Acids Res* 40:W393–W399
 101. Southan C, Sharman JL, Benson HE, Faccenda E, Pawson AJ, Alexander SPH (2016) The IUPHAR/BPS Guide to PHARMACOLOGY in 2016: towards curated quantitative interactions between 1300 protein targets and 6000 ligands. *Nucleic Acids Res* 44:D1054–D1068
 102. Koscielny G, An P, Carvalho-Silva D, Cham JA, Fumis L, Gasparyan R, Dunham I (2017) Open targets: a platform for therapeutic target identification and validation. *Nucleic Acids Res* 45:D985–D994
 103. Wang X, Shen Y, Wang S, Li S, Zhang W, Liu X, Li H (2017) PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res* 45:W356–W360
 104. Siramshetty VB, Eckert OA, Gohlke BO, Goede A, Chen Q, Devarakonda P, Preissner S, Preissner R (2018) SuperDRUG2: a one stop resource for approved/ marketed drugs. *Nucleic Acids Res* 46: D1137–D1143
 105. Gfeller D, Grosdidier A, Wirth M, Daina A, Michielin O, Zoete V (2014) SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucleic Acids Res* 42:W32–W38
 106. Yao ZJ, Dong J, Che YJ, Zhu MF, Wen M, Wang NN, Wang S, Lu AP, Cao DS (2016) TargetNet: a web service for predicting potential drug–target interaction profiling via multi-target SAR models. *J Comput Aided Mol Des* 30:413–424

107. Chen X, Ji ZL, Chen YZ (2002) TTD: therapeutic target database. *Nucleic Acids Res* 30:412–415
108. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2016) PubChem substance and compound databases. *Nucleic Acids Res* 44: D1202–D1213
109. Pence HE, Williams AJ (2010) ChemSpider: an online chemical information resource. *Chem Educ* 87:1123–1124
110. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M (2017) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkx1037>
111. Voigt JH, Bienfait B, Wang S, Nicklaus MC (2001) Comparison of the NCI open database with seven large chemical structural databases. *J Chem Inf Comput Sci* 41:702–712
112. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG (2012) ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model* 52:1757–1768
113. Dimitropoulos D, Ionides J, Henrick K (2006) In: Baxevanis AD, Page RDM, Petsko GA, Stein LD, Stormo GD (eds) *Current protocols in bioinformatics*. Wiley, Hoboken, NJ, pp 14.3.1–14.3.3 ISBN: 978-0-471-25093-7
114. Wang R, Fang X, Lu Y, Yang CY, Wang S (2005) The PDBbind database: methodologies and updates. *J Med Chem* 48:4111–4119
115. Wang R, Fang X, Lu Y, Wang S (2004) The PDBbind Database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J Med Chem* 47:2977–2980
116. Goede A, Dunkel M, Mester N, Frommel C, Preissner R (2005) SuperDrug: a conformational drug database. *Bioinformatics* 21:1751–1753
117. Schomburg I, Chang A, Ebeling C, Gremse M, Heldt C, Huhn G, Schomburg D (2004) BRENDA, the enzyme database: updates and major new developments. *Nucleic Acids Res* 32:D431–D433
118. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, Haw R, Jassal B, Korninger F, May B, Milacic M, Roca CD, Rothfels K, Sevilla C, Shamovsky V, Shorser S, Varusai T, Viteri G, Weiser J, Wu G, Stein L, Hermjakob H, D'Eustachio P (2018) The reactome pathway knowledgebase. *Nucleic Acids Res* 46: D649–D655
119. Milacic M, Haw R, Rothfels K, Wu G, Croft D, Hermjakob H, D'Eustachio P, Stein L (2012) Annotating cancer variants and anti-cancer therapeutics reactome. *Cancers (Basel)* 8:1180–1211
120. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30
121. Huaiyu M, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD (2016) PANTHER version 11: expanded annotation data from gene ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkw1138>
122. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem* 31:455–461
123. Bikadi Z, Hazai E (2009) Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock. *J Cheminform* 1:15
124. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* 47:1750–1759
125. Rarey M, Kramer B, Lengauer T, Klebe G (1996) A fast flexible docking method using an incremental construction algorithm. *J Mol Biol* 261:470–489
126. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 267:727–748
127. Abagyan R, Totrov M, Kuznetsov D (1994) ICM—a new method for protein modeling and design—applications to docking and structure prediction from the distorted native conformation. *J Comput Chem* 15:488–506
128. ACD (Advanced Chemistry Development) (2015) Toronto, ON, Canada. www.acdlabs.com
129. Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* 7:42717
130. <http://www.cyprotex.com/physicochemical-profiling>

131. Yan Y, Zhang D, Zhou P, Li B, Huang S-Y (2017) HDock. A web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkx407>
132. Jeon J, Nim S, Teyra J, Datti A, Wrana JL, Sidhu SS, Kim PM (2014) A systematic approach to identify novel cancer drug targets using machine learning, inhibitor design and high-throughput screening. *Genome Med* 6:57
133. Kesari KK, Jamal QMS, Sharma A, Chauhan P, Dhasmana A, Siddiqui MH, Sisodia R, Verma HN (2017) Induction of LPO and ROS production in rat brain exposed to microwaves: computational elucidation of melatonin in repair system. In: Kesari K (ed) *Perspectives in environmental toxicology*. Springer International Publishing, Switzerland, pp 31–46. https://doi.org/10.1007/978-3-319-46248-6_2
134. Choi BW, Ryu G, Park SH et al (2007) Anticholinesterase activity of plastoquinones from *Sargassum sagamianum*: lead compounds for Alzheimer’s disease therapy. *Phytother Res* 21:423–426
135. Irwin JJ, Shoichet BK (2005) ZINC—a free database of commercially available compounds for virtual screening. *J Chem Inf Model* 45(1):177–182
136. Blundell TL, Sibanda BL, Montalvão RW, Brewerton S, Chelliah V, Worth CL, Burke D (2006) Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery. *Philos Trans R Soc Lond B Biol Sci* 361(1467):413–423
137. Klahan S, Wong HS, Tu SH, Chou WH, Zhang YF, Ho TF, Liu CY, Yih SY, Lu HF, Chen SC, Huang CC, Chang WC (2017) Identification of genes and pathways related to lymphovascular invasion in breast cancer patients: a bioinformatics analysis of gene expression profiles. *Tumour Biol* 39:1010428317705573
138. Zhu QN, Renaud H, Guo Y (2018) Bioinformatics-based identification of miR-542-5p as a predictive biomarker in breast cancer therapy. *Hereditas* 155:17
139. Lionta E, Spyrou G, Vassilatis DK, Cournia Z (2014) Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem* 14(16):1923–1938
140. Ferreira LG, Dos Santos RN, Oliva G, Andriacopulo AD (2015) Molecular docking and structure-based drug design strategies. *Molecules* 20:13384–13421
141. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 16:2785–2791
142. Gao J, Wang T, Qiu S, Zhu Y, Liang L, Zheng Y (2016) Structure-based drug design of small molecule peptide deformylase inhibitors to treat cancer. *Molecules* 21:396
143. Cherkasov A, Muratov EN, Fourches D et al (2014) QSAR modeling: where have you been? Where are you going to? *J Med Chem* 57(12):4977–5010. <https://doi.org/10.1021/jm4004285>
144. Camp D, Gavelas A, Campitelli M (2015) Analysis of physicochemical properties for drugs of natural origin. *J Nat Prod* 786:1370–1382
145. Wan H (2013) What ADME tests should be conducted for preclinical studies? *ADMET and DMPK* 1:19–28
146. Stone HB, Bernhard EJ, Coleman CN et al (2016) Preclinical data on efficacy of 10 drug-radiation combinations: evaluations, concerns, and recommendations. *Transl Oncol* 9(1):46–56
147. Fördös I, Fazekas J, Singer J, Jensen-Jarolim E (2015) Translating clinical trials from human to veterinary oncology and back. *J Transl Med* 13:265
148. Simoni E, Bartolini M, Abu IF, Blockley A, Gotti C, Bottegoni G, Caporaso R, Bergamini C, Andrisano V, Cavalli A, Mellor IR, Minarini A, Rosini M (2017) Multitarget drug design strategy in Alzheimer’s disease: focus on cholinergic transmission and amyloid- β aggregation. *Future Med Chem* 9(10):953–963
149. Zhang W, Pei J, Lai L (2017) Computational multitarget drug design. *J Chem Inf Model* 57(3):403–412
150. Lu J-J, Pan W, Hu Y-J, Wang Y-T (2012) Multi-target drugs: the trend of drug research and development. *PLoS One* 7(6):e40262
151. Bolognesi ML (2013) Polypharmacology in a single drug: multitarget drugs. *Curr Med Chem* 20(13):1639–1645
152. Bolognesi ML, Cavalli A (2016) Multitarget drug discovery and polypharmacology. *ChemMedChem* 11(12):1190–1192
153. Vitali F, Cohen LD, Demartini A, Amato A, Eterno V, Zambelli A et al (2016) A network-based data integration approach to support drug repurposing and multi-target therapies in triple negative breast cancer. *PLoS One* 11(9):e0162407

INDEX

A

Absorption, distribution, metabolism, excretion and toxicity (ADME/Tox) 54, 64, 76, 94, 96, 99, 102, 160, 161, 177, 178, 214, 219, 220, 238, 242, 276, 424, 429, 437, 453, 472, 538
Acetylcholine (AChE)..... 73, 96, 101, 158, 203, 206–208, 216–218, 234, 261, 263, 537
Acetylcholinesterase 73, 96, 101, 158, 203, 216, 218, 234, 261, 263, 537
Activity cliffs 12, 19
Activity landscapes 18–19, 57
Activity networks..... 54, 197
ADME/Tox, *see* Absorption, distribution, metabolism, excretion and toxicity
Alzheimer's disease (AD)..... vii, 52, 157, 158, 203–242, 255–340, 367–382, 428, 429, 548
AMBER 38, 45, 451, 452, 474–476
AMBER force field..... 38, 437, 475
Amyloid 157, 210, 213, 235, 257, 259, 260, 273, 279, 331, 372, 379, 381
Amyloid β -A4 protein 97
Amyotrophic lateral sclerosis (ALS)..... 157, 158
Antiamyloidogenic 174, 373, 376–378, 429
Antigen-presenting cells (APCs)..... 395
Antituberculosis 108, 110, 116, 119, 129, 133–135, 137, 139, 140, 142, 143
Artificial intelligence methods..... 64
Artificial neural networks (ANN)..... 29, 40, 59, 61–63, 161, 163
Assembly problem..... 64
Astemizole 52
Atomic force microscopy 378
AutoDock 34, 65, 66, 70, 71, 166, 192, 429–433, 437, 450, 474, 544, 547
AutoDock-Vina 33–37, 66, 67, 429, 473, 474, 544, 547

B

BACE1, *see* β -secretase-1 inhibitors (BACE1)
Bayesian methods..... 61
Bayesian QSAR..... 62
Benzimidazoles..... 138–140, 143, 331
 β -amyloid 157, 158, 211, 272, 273, 287, 288, 295, 305, 308, 324, 327, 328, 428

β -secretase-1 inhibitors (BACE1) 173–175, 213–216, 225, 232, 235, 240, 260, 261, 265, 266, 268, 279, 308, 310, 315, 317, 318, 369, 370, 372–375, 379, 428, 429
 β -secretases 69, 213, 214, 216, 257, 260, 261, 369
Binary kernel discrimination 29, 61
Binding affinity..... 31, 40, 64, 71–73, 139, 267, 325, 367, 375, 386, 450–453, 476, 490, 540
Binding pocket 34, 407, 470, 473, 474, 476, 489, 490, 494
BINDNET..... 30
Bioactive 54, 56, 57, 59, 62, 95, 165, 195, 205, 256, 260, 261, 297, 305, 340, 360, 422, 538
Biological receptors..... 39, 46
Biphasic partition system 40
Blind docking 64, 65, 67, 165
B-Raf..... vii, 355–364, 446, 448, 449, 518
Breaking point..... 376, 377
Breast cancer..... vii, 429, 443–453, 505–512, 518, 522, 523, 534, 535, 546, 548
Butyrylcholinesterase 170, 218, 262, 263, 537, 546

C

CADD, *see* Computer-aided drug design
Cancer 27, 52, 134, 188, 255, 368, 428, 443–453, 466, 488, 505–512, 518, 529–549
CANDO 33, 35
Cardiac hERG channel 52
Central nervous system (CNS)..... 52, 157, 206, 257, 369, 518, 548
CFP10-ESAT6 129–130, 139, 143
Chalcones 135–137, 143
Charge assignment..... 32
CHARMM 45, 432, 451
CHEMBL 4, 169, 172, 195, 205, 214, 406, 413, 437, 521, 537, 538
Chemical databases vi, 3, 4, 7, 17, 98, 101, 160, 165, 167, 173, 425
Chemical genomics 7

- Chemical scaffolding clustering 56
Chemical space vi, 4, 7, 15–19,
62, 78, 162, 195, 382, 406, 421, 425, 439, 470
Cheminformatics vi, 3–22, 39, 172, 437, 541
ChemMaps 16–18
Chemogenomic v, 60, 167, 176, 189, 538
Chemometrics 143, 156, 161, 167
Cholinesterases 94, 216–224,
242, 272, 276, 292, 295, 308, 318, 334
Classical force field potential 37
Classical mechanics 42
Cluspro 34
Combination therapy 358, 423,
424, 444–446, 465, 466, 478
Comparative molecular field analysis (CoMFA) 39,
60, 162, 206, 207
Comparative molecular similarity indices analysis
(CoMSIA) 162
Compound database 4–7, 11,
17, 20, 21, 359–360, 386, 450
Computational chemistry 53
Computational docking vii, 419–439
Computer-aided drug design (CADD) 53–55,
57, 63, 80, 108, 110, 156, 215, 500, 542
Conformational changes 42, 45,
70, 71, 141, 174, 413
Conformational search 31, 65, 71
Consensus diversity plot (CDP) 10–11
Co-receptor blockade 519
Coulomb's electrostatic potential 43
Covalent binding 31, 37, 175
Cytochrome P450 monooxygenase (CYP) 124–125,
137, 193, 421, 425
- D**
- Data-driven 55, 405, 406, 521
Data-mining 4, 30, 40,
56, 60, 81, 82, 406–409, 413, 549
Deazaflavin-dependent nitroreductase
(Ddn) 114–116, 137
Decision trees (DT) 29, 61, 62
Deep learning vi, 64
De novo
design 63, 64, 68, 75, 80, 98, 470, 472
methods 63–64, 470–472
Density functional theory (DFT) 38
Dereplication 189, 194–195
Derringer's desirability function 63
Dielectric screening constant 44
Dihydrofolate reductase (DHFR) 131–133, 425
Dirty compounds 52
Discovery Studio Visualizer 36, 546
Disease pathway v, 51, 367, 488
DNADock 34
DNA gyrase 130–131, 135, 142,
143, 463, 466
DNA minor groove 42
DNA molecule 32
DNA Origami 530, 534–537, 548
DOCK 34, 450
DOCK BLASTER 34
Docking
box 32
protocol 31, 32, 42, 66, 71,
362, 364, 438
Donepezil 158, 169–172,
174, 219, 220, 227, 229, 231, 232, 235, 238,
240, 242, 267–286, 319, 331, 334, 336, 340
Dopaminergic receptor 169
DOT 34
DprE1 112–114, 134, 143
Drug-based similarity inference (DBSI) 30
Drug discovery v–viii, 3, 7,
21, 22, 27, 29, 39, 46, 51–55, 58, 63, 70, 77, 80,
94, 107–143, 156, 176, 188, 189, 197, 204, 308,
355, 386, 407, 421, 427, 439, 449, 465, 468,
487, 488, 500, 518, 519, 526, 538, 541, 549
Druglikeness 7, 94, 95,
102, 160, 428, 429, 434, 473, 545
Drug repurposing viii, 28,
35, 97, 187, 421, 426, 437, 439, 488, 525, 548
Drug targets vi, 3, 30,
31, 33, 51, 53, 56, 70, 71, 73, 108, 206, 207, 317,
408, 426, 448–450, 467, 517–521, 530,
538–541, 547
D-tools 4, 11, 12
Dual inhibitors vii, 175,
190, 191, 232, 265, 298, 334, 355, 360, 363,
421, 424, 429–436, 445–449, 452, 453
Dynamic docking 70
Dynamic systems 32
- E**
- Efficacy 5, 28, 52, 53, 80,
82, 141, 187, 188, 196, 211, 256, 261, 290, 355,
356, 367, 405, 413, 421, 463, 467, 488, 506,
523, 525, 530, 533, 536, 537
Electrostatic interactions 37, 43–45, 242
Emerging chemical patterns (ECP) 62
Ensemble docking 65, 68–70,
165, 374, 375, 439
Ensemble linking 64
Ensemble pharmacophore vii, 386, 387,
389–392
EpiTOP 396, 402–403
Estrogen 6, 210–212, 444
Exosomes 530–533, 537
Extracellular vesicles (ECV) 531

F

Feature net (FN)	60
Flenfluramine	52
Fibrils	210, 216, 222, 229, 234, 235, 239, 272, 307, 308, 378
Fingerprints for ligands and proteins (FLAP)	166
Focused library	7, 63, 360
Force field	33, 37, 38, 43, 45, 72, 361, 374, 391, 412, 432, 437, 451, 474, 475
FORECASTER Suite	35
Fragment-based drug design	95–97
Framework combination	57, 77, 370–372
Free energy	31, 66, 70–72, 137, 166, 451, 452, 469, 472, 476
Free energy landscape (FEL)	469, 472, 476

G

GEMDOCK	35, 438
Gene silencing	506, 507
Genomics	30, 506, 518, 522, 549
Global optimal solution	62
Global Range Molecular Matching (GRAMM)	34
Glutamine synthetase (GS)	127–128, 143
Glutamyl cyclase (GC)	210
Glycogen synthase kinase (GSK)	73, 97, 318
gOpenMol	46
Grepafloxacin	52
GRIID Independent Descriptors (GRIND)	60
GROMACS	45, 388, 391, 451, 452, 474–477
GROMOS	45, 451, 474, 475
Group based QSAR (GQSAR)	62

H

Hesperidin	369, 375, 377–381
Hex Protein Docking	34
High throughput screening (HTS)	196, 205, 329, 450, 537
Histamine N-methyltransferase (HMT)	172
Histamine 3 receptor (H ₃ R)	172
HLA class II binding prediction	402
HLA-DP proteins	396–398, 401
Hologram QSAR	174
Hsp90	vii, 355–364, 429
Human biology	29
Huntington's disease (HD)	155, 157–159, 168, 177
Hybrid QM-MM method	37
5-Hydroxytryptamine receptor	170
Hyphenated methods	37–39

I

Induced fit docking	70
Inducible nitric oxide synthase	170
Infectious diseases	52, 422–424, 426, 429, 439, 460, 488
InhA	122–124, 136, 139, 142, 143, 463, 466, 473, 474
In silico	vi, vii, 53–55, 62, 64, 80–82, 94, 98, 110, 111, 143, 156, 160, 161, 167–169, 177, 178, 204, 208, 214, 222, 233, 235, 237–239, 386, 406, 424, 428, 466, 473, 521, 538, 547
In silico approach	40, 231, 386
Intermolecular interactions	37, 46, 165
INVDock	166
Inverse docking	v, 100, 166, 176, 382
In vitro	vii, 52, 82, 115, 127, 134, 135, 139, 143, 168, 172, 175, 188, 189, 196, 209, 210, 221–223, 225–227, 229, 231–236, 238, 240, 242, 262, 263, 265, 272, 276, 278, 280, 285, 289, 290, 293, 296, 298, 299, 301, 302, 308, 313, 317, 321, 331, 333, 335, 362, 378, 452, 453, 477, 520, 523, 533, 548, 549
In vivo	vii, viii, 52, 56, 82, 196, 206, 210, 262, 278, 280, 301, 302, 308, 310, 327, 452, 453, 477, 517–522, 525, 526, 532, 533, 549
Isocitrate lyase	126–127, 141, 143
I-TASSER	33
Iterations	32

K

Kernels	60
<i>k</i> -nearest neighbor (kNN)	29, 61, 162
KNIME	3–5, 7, 9, 12, 15, 406, 407
KNIME workflow	11, 408, 409
Knowledge-based lead generation	57

L

Lennard-Jones (LJ) potential	42
Ligand-based	28, 30, 31, 55, 74, 94–101, 160, 164, 165, 173, 174, 191, 390, 405, 409, 428, 470–472, 547
approaches	v
design	405
virtual screening	28
Ligand-based drug design (LBDD)	160, 547
Ligand centric	55
Ligand efficiency (LE)	214, 235, 434, 435
Ligand similarity	30
Ligand-target docking	31

- Linear discriminant analysis (LDA)..... 28, 29,
40, 59, 61–63, 161, 163, 169
- ‘Linked,’ ‘fused’ or ‘merged’ framework..... 57
- Logistic regression (LR) 61
- M**
- Machine learning..... 19–21, 28, 40,
55, 68, 72, 82, 98, 161, 163, 168, 192–193, 196,
382, 407, 409, 471, 538
- methods 98, 471
- tools 40, 163
- Macromolecules 29, 31, 34,
35, 40, 140, 431, 436, 439, 474, 533, 547
- Major histocompatibility complex (MHC) 395
- Malaria 28, 73, 188, 419–439, 466
- MARCH-INSIDE (MI) 40, 59, 60, 62
- Markov chain 62, 63
- MARKovian CHEmicals IN Silico Design 62
- Memoquin 159, 261–263, 308, 309
- Merged pharmacophore 386–389, 392, 427
- Metabolic pathway 30, 113, 126
- Micro RNAs (miRNAs) 506, 531
- MODELLER 33, 407, 411
- Model validation 398, 401
- Molecular descriptors..... 7–8, 39,
57–59, 62, 161, 163
- Molecular docking 31–37, 55,
64–79, 100, 107–143, 160, 165, 166, 173–176,
191, 205, 215, 280, 323, 374, 386, 407, 411,
429, 437, 449–450, 468, 470, 473, 474, 521,
524, 544
- Molecular dynamics
- approach 33, 41–45
- simulations..... 42–45, 70,
160, 166, 386, 388, 391, 412, 413, 429, 439,
451, 452, 469, 472, 474
- trajectories 44
- Molecular fingerprint..... 7–9, 11,
12, 15, 16, 100, 167, 176, 189
- Molecular Graphics Laboratory tools (MGLTools) 33, 35
- Molecular interaction..... 42, 62, 94, 406,
409, 540, 546, 547
- Molecular mechanics..... 37, 45, 374, 476
- Molecular mechanics Poisson-Boltzmann surface area
(MM-PBSA) 40, 472, 476
- Molecular modeling..... 143, 206,
223–225, 227–229, 239, 270, 271, 275, 291,
293, 312, 321, 330, 412
- Molecular Operating Environment (MOE)..... 59,
76, 161, 165, 360–362, 407, 410, 437
- Molecular profiles..... 64, 81
- MolFit 34
- MOLS 2.0 34
- Monoamine oxidases 94, 96,
97, 159, 171, 176, 208–210, 242, 268, 269
- MOOP-DESIRE..... 63
- Moving average approach..... 40, 162
- M2 receptor..... 52, 206
- MtFabH 119–120, 134
- Mtk-QSBER..... 40
- Mt-spectral moments 63
- Multidrug resistance tuberculosis
(MDR-TB)..... 108, 464
- Multi kinase (PDGFR-beta, FGFR-1,
and SRC) 62
- Multi-kinase target..... 188
- Multi-objective optimization (MOOP)..... 61–63, 102
- Multiple linear regression (MLR) 40, 59,
161, 409
- Multiple outputs 59, 60
- Multiple target 21, 22, 28–31,
33, 39, 46, 51, 52, 73, 76, 80, 81, 177, 188, 193,
196, 256, 260, 261, 263, 265, 367–369, 385,
386, 396, 405, 413, 436, 447, 449, 450, 452,
465, 467, 468, 472, 477, 488, 498, 517–526
- Multi-serotonin target 62
- Multi-target agents 28, 62, 73,
303, 427, 500
- Multitarget compounds..... 133–143,
168, 177, 225, 230, 231, 236
- Multitarget directed ligands (MTDL) 156,
159, 160, 162, 164, 166, 168–177, 255–340,
367–382
- Multi target drug discovery/design
(MTDD)..... 54–79, 81, 94–96, 98–103
- Multi-targeted drugs (MTD)..... 27–46,
51–83, 93–103, 162, 177, 187–197, 280, 290,
308, 385, 396, 405, 408, 421–429, 439,
443–453, 459–478, 488, 518, 521, 522, 526, 548
- Multi-targeted molecular dynamics
(MTMD) 70, 71
- Multi-target inhibitors 62, 73, 76, 77,
79, 190, 191, 193, 385, 388, 422, 427, 428, 448
- Multi-target ligand design 489–501
- Multi-target QSAR (mt-QSAR) 40, 61
- Multi-target quantitative structure–activity
relationships (mt-QSARs)..... 57–59, 62, 370
- Multi-target screening 424–430,
434–436, 438, 439
- Multi-target structure–activity relationships (mt-SARs)
- Multi-task algorithm 40
- Multi-tasking QSAR 58, 59
- Multi-task learning (MTL)..... 60
- Multivariate analysis 376
- Mur Ligase..... 114, 115
- Mycobacterium tuberculosis*

N

N-acetylglucosamine-1-phosphate uridyltransferase
 (GlmU) 117–118
 NAMD 45, 451, 474, 476
 Nanodrugs 531
 Natural products 82, 135, 187–197,
 221, 261, 304–317, 338, 431, 437
 Neglected tropical diseases (NTDs) 422, 424, 439
 Network-based inference (NBI) 30
 Neurodegenerative diseases 93–103,
 155–178, 203, 259, 284, 296, 306, 321, 325, 534
 Newtonian movement 42
 Newton's equations of motion 42
 Nitroimidazoles 116, 136–138, 143
 Nonamer 399–402
 Nucleic acids 27, 29, 45, 66, 436, 506

O

Off-target v, 52, 54, 93, 98, 172, 176,
 187, 189, 405, 421, 425, 430, 488, 517, 521–525
 Oncogene viii, 447, 505–508, 512
 ONIOM 37–39
 OpenBabel 36, 176, 360, 431, 438
 Open data vii, 405–414, 437, 538
 Open Source program 34, 544

P

Pantothenate synthase (PS) 120–121,
 139–141, 143, 466
 Parallel processing 33
 ParDock 34, 67, 490
 Parkinson's disease (PD) 52, 67,
 93, 155, 157–159, 168, 172, 177, 207
 Partial least squares least squares (PLS) 59,
 60, 161, 214, 400–401
 Pattern recognition 60
 PDB, *see* Protein Data Bank
 Peptides vii, 34, 65,
 114, 125, 140–143, 157, 174, 213, 374,
 379–381, 395–403, 425, 533, 536
 Perturbation theory 40
 Pharmaceutical 7, 8, 81, 82,
 94, 133, 187, 196, 204, 256, 335, 368, 463, 540,
 542, 549
 Pharmacokinetic 28, 52,
 53, 55, 63, 81, 159, 368, 369, 540, 543, 545,
 547, v
 Pharmacological 22, 40, 60,
 61, 63, 64, 81, 93, 110, 173, 231, 265, 269, 270,
 279, 283, 300, 309, 312, 320, 330, 368, 385,
 408, 467
 Pharmacophore 6, 55–57,
 64, 74–80, 82, 99, 101, 102, 164–166, 172–174,

188, 190–192, 231, 269, 307, 311, 321, 327,
 329, 360–362, 371, 372, 387–392, 427, 450,
 467–470, 472, 540
 Pharmacophore mapping v, vi, 30, 99,
 101, 160, 164–165, 167, 173–174
 PharmMapper server 30
 Physicochemical 7, 8, 10, 11,
 16, 52, 55, 59, 75, 98, 101, 102, 133, 137, 160,
 161, 192, 195–196, 205, 215, 218, 376, 490,
 500, 540, 543, 545, 547
 Phytochemical library 374
 PKS13 118–119, 134, 143
 PLS, *see* Partial least squares least squares
 Pocketome 30
 Polypharmacology v, vi, 3–22,
 28, 51–55, 60, 80, 98–100, 160–167, 172,
 176–177, 188, 189, 193–194, 256, 355, 356,
 362, 367, 371, 421, 488, 517–527
 Predictability 58, 371
 Presenilin-1 (PSEN1) 97, 169
 Principal components analysis (PCA) 15–17, 476
 Prion disease (PD) 157, 158, 168, 172, 177
 Promiscuous vii, 28, 52, 72,
 80, 196, 409, 421
 Protein Data Bank (PDB) 30, 37, 110,
 112, 140, 205, 209, 211, 214–216, 218–223,
 228, 229, 231–241, 359, 360, 374, 375, 388,
 410, 411, 425, 428, 431, 432, 436, 450, 452,
 473, 542
 Protein homology modeling 436
 Protein–protein docking 34, 545
 Proteins 27, 51, 108, 157, 187, 205, 256, 355,
 367, 385, 395–403, 406, 420, 461, 487–500,
 518, 531
 Proteochemometrics vi, vii, 395–403
 PS, *see* Pantothenate synthase
 Pterostilbene 312, 313
 Ptp A and B 121–122
 pt-QSPR 40
 PyMol 33, 36, 68, 410,
 432, 433, 438, 475, 476
 Pyridinone 326, 327
 Pyridoxine 5'-phosphate 125–126
 Pyrimidine 139, 329, 358

Q

QM–MM approach 37–39
 QM–Polarized Ligand Docking (QPLD) 39
 Quantitative structure–activity relationship
 (QSAR) vi, vii, 19, 20,
 39–41, 54–63, 82, 94, 98, 99, 102, 135, 137, 156,
 160, 162–163, 165, 168, 169, 172, 174, 207,
 209, 214, 218, 226, 232, 242, 275–276,
 376–378, 392, 396, 409, 427–429, 466, 547

- Quantitative structure–property relationship (QSPR) 490
- Quantum mechanics (QM) 37, 42, 45
- Quinazoline 239, 333
- Quinoline 241, 318, 422
- Quinolones 142–143, 270, 286
- R**
- RA, *see* Rheumatoid arthritis
- Random forests (RF) 29, 61, 66, 161
- Rapacuronium 52
- Reactivity characteristics 56
- Receptor binding site 31, 65, 165, 476
- Receptor flexibility 31, 33, 44, 65, 67, 69
- Receptors 6, 28, 34, 52, 97, 110, 131, 158, 188, 203, 206–207, 263, 357, 374, 390, 443, 473, 491, 517, 533
- Receptor tyrosine kinases (RTK) 357, 444–446, 449, 494, 496, 534, 535
- Resveratrol 188, 196, 297, 298, 302, 306–308, 312, 315, 316, 323, 324
- Rheumatoid arthritis (RA) 52, 53, 255
- Rigid body docking 32, 34, 41
- Rigid body system 32
- Riluzole 159
- Rivastigmine 158, 227, 235, 261, 265, 320, 321, 335, 340
- RMSD, *see* Root mean square deviation
- RNA interference (RNAi) viii, 506, 507
- Robustness 27, 41, 58, 427
- Root mean square deviation (RMSD) 375, 388, 412, 413, 433, 439, 451, 474, 475
- S**
- SAR, *see* Structure–activity relationships
- Scaffold 7, 9–14, 21, 55, 56, 62, 75, 94, 96, 97, 175, 188, 221, 265, 266, 269, 279, 282, 286, 287, 293, 295, 306, 313, 314, 318, 322, 325, 326, 340, 358, 360, 361, 364, 370, 371, 375, 387, 406, 407, 409, 411, 412, 490, 500
- Scaffold-based optimization viii
- Scaffold hopping 62, 192, 371
- Schrödinger suite 39, 388, 389, 410, 411, 414
- Scoring criteria 31
- Scoring functions 32, 64, 66, 67, 69–74, 76, 100, 165, 166, 389, 407, 468, 469, 474, 477, 539
- Screening virtual ligands 63, 64
- Scutellarin 320, 321
- SEA, *see* Similarity ensemble approach
- Serotonin 52, 62, 208, 284, 406
- Serotonin 5HT2B receptor 52
- Serotonin receptor 1A 62
- Serotonin transporter 62
- Shikimate kinase 78, 128–129, 142
- Shogaol 309
- Short hairpin RNAs (shRNAs) 506
- Side effects 28, 51–54, 93, 108, 187, 188, 355, 356, 367–369, 405, 421, 424, 425, 467, 524, 530, 533, 534, 536, 540, 543
- Similarity v, 7, 9, 11–16, 18–21, 30, 40, 56, 57, 60, 69, 96, 99–101, 112, 120, 122, 160, 162, 164, 167, 168, 176, 191–193, 213, 214, 360, 364, 371, 372, 382, 388, 396, 405, 406, 412, 428, 468, 472, 488, 489, 491–493, 495, 540
- Similarity ensemble approach (SEA) v, 164, 167, 405
- Similarity rule 56
- Similarity searches 20, 30, 160, 164, 168, 176, 191–192, 213, 214
- Simulations v, 30, 33, 42–45, 68, 70, 100, 135, 136, 160, 166, 176, 191, 197, 213, 225, 226, 234, 239, 241, 386, 388, 391–392, 407, 412–413, 429, 439, 450–453, 467, 469, 472, 474–477
- Small interfering RNAs (siRNAs) viii, 505–512, 533
- SMART, *see* Structure–multiple activity relationships
- Smart drug designing viii, 536–548
- Smart drugs vii, 529–549
- Solvation effects 31
- Steady state fluorescence 378, 379
- Stitch Server 30
- Stochastic algorithms 33
- Structural features 8, 39, 42, 56, 96, 164, 205, 209, 225, 287, 290, 292, 302, 360
- Structure–activity relationships (SAR) 4, 9, 11–16, 18, 19, 54–57, 82, 171, 210, 221, 230, 313, 329, 331, 371, 409, 412, vi, vii
- Structure-based v, 74, 76, 78, 101, 129, 160, 164–166, 174–176, 190, 358, 374, 390, 406, 409–411, 413, 429, 439, 468, 470, 472, 544, 547
- approaches v, vii, 29, 406
- design 80, 94, 166, 386
- drug design 64, 67, 392, 544, 547
- virtual screening 28
- Structure–multiple activity relationships (SMART) 3, 22
- Substructure mining 63
- Sulfonamides 134–135, 143
- Support vector machines (SVM) 29, 60, 61, 98, 161, 163, 429
- SwissDock 34, 437
- Synthase-kinase 3 73, 97, 259, 318

Systematic algorithms	33	TS, <i>see</i> Transition state	
Systems approach	29–31	TTL8	31
SystemsDock	30	Tuberculosis (TB)	vi, vii, viii, 78, 107, 108, 110–112, 114, 116, 118–128, 130, 131, 133–137, 139–143, 422–424, 429, 459–478
T		Tyrosine kinase inhibitors (TKIs)	427, 447, 493, 494, 534
Tacrine	174, 218–223, 225–227, 231, 235–237, 240, 242, 261, 277, 286–303, 319, 340	U	
Tanimoto coefficients.....	12, 13, 15, 164, 388	UCSF-Chimera	33, 36, 438
Target-based similarity inference (TBSI).....	30	V	
Target center	55, 467	Van der Waals interaction	28, 42, 43, 476
Target deconvolution.....	21, 355	Van der Waals radius	43, 390
Target fishing	19–21, 167, 197, 382	Verlet algorithm	44
Tau protein	157, 158, 176, 257, 259, 262, 292, 309, 321	VinaMPI	33, 67, 406
TBSI, <i>see</i> Target-based similarity inference		Virtual screening (VS)	vi, vii, 4, 6, 7, 11, 17, 19, 22, 28, 63–67, 69, 71, 73, 75, 100–102, 108, 140, 143, 156, 160, 165, 166, 189, 190, 204–205, 214, 222, 355–364, 370, 375–376, 386, 387, 408, 449–450, 452, 466, 473, 474, 476, 500, 537, 542
Terfenadine.....	52	Visual molecular dynamics (VMD).....	45, 46, 410, 475
Thioflavin T assay.....	378, 380	VS, <i>see</i> Virtual screening	
Thiophenes	133–134, 143, 221	W	
3D structures.....	32, 34, 76, 374, 389, 425, 431, 432, 436, 498, 544	Weighted-sum-of-objective-functions (WSOF).....	63
Thymidine monophosphate kinase (TMPK)	126, 140, 143	X	
TINKER.....	45	X-ray structures	208, 221, 222, 231–235, 239, 241, 358–360, 363, 397, 411, 452
Tissue specificity	518	Z	
TKIs, <i>see</i> Tyrosine kinase inhibitors		ZDOCK.....	34
TMPK, <i>see</i> Thymidine monophosphate kinase		ZINC database	490
TOMOCOMD	59		
TOPS-MODE	59		
Toxicity	40, 51–54, 82, 108, 122, 133, 134, 159, 169, 267, 273, 274, 286, 288, 291, 293, 295–297, 301, 303, 306, 308, 311, 315, 317, 330, 333, 368, 372, 425, 437, 453, 463, 465, 477, 478, 488, 505, 541		
Toxicological	28, 63, 545		
Training datasets	470		
Transition state (TS)	9, 37, 44		
Triazolopyrimidine	318		
Trolox	269, 282, 293, 299, 303, 307, 311, 312, 317, 319, 321, 323, 333		