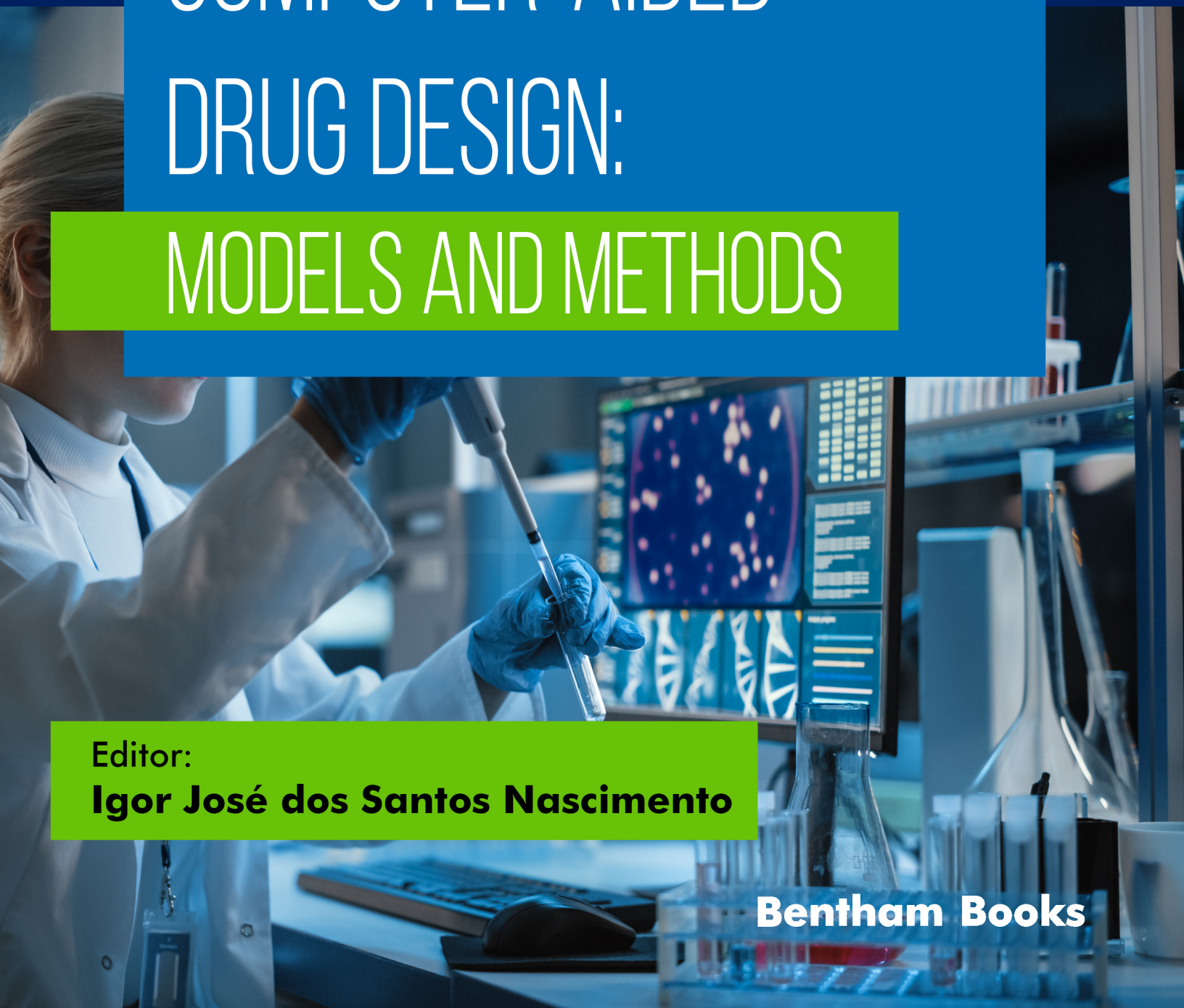
# APPLIED COMPUTER-AIDED DRUG DESIGN:

## MODELS AND METHODS



Editor:
**Igor José dos Santos Nascimento**

Bentham Books

# Applied Computer-Aided Drug Design: Models and Methods

Edited by

**Igor José dos Santos Nascimento**
*Programa de Pós-Graduação em Ciências Farmacêuticas (PPGCF)*
*Departamento de Farmácia*
*Universidade Estadual da Paraíba*
*Campina Grande-PB*
*Brazil*

**Applied Computer-Aided Drug Design: Models and Methods**

Editor: Igor José dos Santos Nascimento

ISBN (Online): 978-981-5179-93-4

ISBN (Print): 978-981-5179-94-1

ISBN (Paperback): 978-981-5179-95-8

© 2023, Bentham Books imprint.

Published by Bentham Science Publishers Pte. Ltd. Singapore. All Rights Reserved.

First published in 2023.

# BENTHAM SCIENCE PUBLISHERS LTD.

## End User License Agreement (for non-institutional, personal use)

This is an agreement between you and Bentham Science Publishers Ltd. Please read this License Agreement carefully before using the ebook/echapter/ejournal (**"Work"**). Your use of the Work constitutes your agreement to the terms and conditions set forth in this License Agreement. If you do not agree to these terms and conditions then you should not use the Work.

Bentham Science Publishers agrees to grant you a non-exclusive, non-transferable limited license to use the Work subject to and in accordance with the following terms and conditions. This License Agreement is for non-library, personal use only. For a library / institutional / multi user license in respect of the Work, please contact: permission@benthamscience.net.

### Usage Rules:

1. All rights reserved: The Work is the subject of copyright and Bentham Science Publishers either owns the Work (and the copyright in it) or is licensed to distribute the Work. You shall not copy, reproduce, modify, remove, delete, augment, add to, publish, transmit, sell, resell, create derivative works from, or in any way exploit the Work or make the Work available for others to do any of the same, in any form or by any means, in whole or in part, in each case without the prior written permission of Bentham Science Publishers, unless stated otherwise in this License Agreement.
2. You may download a copy of the Work on one occasion to one personal computer (including tablet, laptop, desktop, or other such devices). You may make one back-up copy of the Work to avoid losing it.
3. The unauthorised use or distribution of copyrighted or other proprietary content is illegal and could subject you to liability for substantial money damages. You will be liable for any damage resulting from your misuse of the Work or any violation of this License Agreement, including any infringement by you of copyrights or proprietary rights.

### *Disclaimer:*

Bentham Science Publishers does not guarantee that the information in the Work is error-free, or warrant that it will meet your requirements or that access to the Work will be uninterrupted or error-free. The Work is provided "as is" without warranty of any kind, either express or implied or statutory, including, without limitation, implied warranties of merchantability and fitness for a particular purpose. The entire risk as to the results and performance of the Work is assumed by you. No responsibility is assumed by Bentham Science Publishers, its staff, editors and/or authors for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products instruction, advertisements or ideas contained in the Work.

### *Limitation of Liability:*

In no event will Bentham Science Publishers, its staff, editors and/or authors, be liable for any damages, including, without limitation, special, incidental and/or consequential damages and/or damages for lost data and/or profits arising out of (whether directly or indirectly) the use or inability to use the Work. The entire liability of Bentham Science Publishers shall be limited to the amount actually paid by you for the Work.

### General:

1. Any dispute or claim arising out of or in connection with this License Agreement or the Work (including non-contractual disputes or claims) will be governed by and construed in accordance with the laws of Singapore. Each party agrees that the courts of the state of Singapore shall have exclusive jurisdiction to settle any dispute or claim arising out of or in connection with this License Agreement or the Work (including non-contractual disputes or claims).
2. Your rights under this License Agreement will automatically terminate without notice and without the

need for a court order if at any point you breach any terms of this License Agreement. In no event will any delay or failure by Bentham Science Publishers in enforcing your compliance with this License Agreement constitute a waiver of any of its rights.

3. You acknowledge that you have read this License Agreement, and agree to be bound by its terms and conditions. To the extent that any other terms and conditions presented on any website of Bentham Science Publishers conflict with, or are inconsistent with, the terms and conditions set out in this License Agreement, you acknowledge that the terms and conditions set out in this License Agreement shall prevail.

**Bentham Science Publishers Pte. Ltd.**
80 Robinson Road #02-00
Singapore 068898
Singapore
Email: subscriptions@benthamscience.net

# CONTENTS

# PREFACE

The drug discovery and development process is time-consuming and demands a high financial cost. In this way, it is estimated to take approximately 10 to 17 years, costing around 4 billion dollars. This stimulates the advancement of new methodologies that can accelerate the discovery process and increase the probability of a promising molecule. In addition, constant developments in informatics and computations have led to the routine use of high-performance computing in medicinal chemistry. Thus, Computer-Aided Drug Design (CADD) methods emerge, capable of providing critical information for the design of new molecules, essential in any new drug discovery program [1, 2].

In this context, the book "***Applied Computer-Aided Drug Design: Models and Methods***" appears, presenting the computational methods used by researchers and pharmaceutical companies. Each chapter explains a technique with high precision so that readers can apply it in their research.

This first edition is organized into nine chapters, namely:

Chapter 1 "***Ligand- and Structure-Based Drug Design (SBDD and LBDD)"***: Promising Approaches to Discover New Drugs. Here, the reader will have an approach from a historical perspective on strategies used in designing new drugs until the development of LBDD and SBDD strategies, exemplifying important discoveries of commercial drugs.

Chapter 2 "***Quantitative Structure-activity relationship (QSAR) in studying the biologically active molecules***". This chapter will bring the principles and methods of this technique based on LBDD. It will present a historical perspective from the first QSAR models to the most current ones like 6D-QSAR. Furthermore, it provides a great read on protocol validation procedures, which are crucial to successful QSAR studies.

Chapter 3 "***Pharmacophore Mapping: An Important Tool in Modern Drug Design and Discovery***". This chapter approaches a method that can be applied to SBDD and LBDD protocols. The reader will have a historical perspective of the evolution of the method, a presentation of the leading software used, and, in the end, a great background on carrying out a well-validated virtual screening protocol based on pharmacophore. Further, the text addresses successful studies and how their protocols were carried out.

Chapter 4 "***Up-To-Date Developments In Homology Modeling***". Similar to the previous chapters, the readers will have a theoretical basis on the technique, quite explored when there is information about the target without an experimental structure. Homology modeling is a powerful tool for constructing and applying molecular targets in drug design studies. With this, readers can perform this protocol safely and efficiently.

Chapter 5 "***Anticancer Activity of Medicinal Plants Extract and Molecular Docking Studies***". In fact, this is the most used tool by drug developers worldwide. Through this technique, new drugs can be safely planned, or even virtual screenings can be carried out to find new drugs. Thus, the authors will bring the technique's theoretical framework, the method's evolution, computational software, and studies in which the application of molecular docking was vital to finding promising molecules.

Chapter 6 "***FBDD & de novo Drug Design***". In this chapter, the main tools used in Fragment-Based Drug Design (FBDD) and *de novo* Drug Design (DNDD) will be presented,

mainly through in silico approaches. It is essential to highlight that these methods control molecules from scratch, generating critical *hits* that later become optimizable *leads*. In addition, all the theoretical frameworks and important discoveries are applied through these strategies.

Chapter 7 "***Molecular simulation in drug design; an overview of molecular dynamics methods***". Despite being a promising technique, molecular docking has several problems, such as disregarding the flexibility of the active site during simulation. Thus, this chapter will address the molecular dynamics technique, which tries to solve some problems from molecular docking. In fact, with the popularization of computers in drug design, this is the fastest-growing technique, and its application is essential in drug discovery programs. Thus, with great clarity, the authors present the theoretical framework and how to apply it in a design campaign for new drugs.

Chapter 8 "***Quantum Chemistry in Drug Design: density function theory (DFT) and other quantum mechanics (QM)-related approaches***". The application of quantum chemistry (QM) protocols in predicting biological activity or enzymatic mechanism are highlighted in the current drug discovery process. Increasingly, researchers are adopting these tools in their drug development projects. Thus, in this chapter Rodrigues et al. They explored the entire theoretical foundation of QM, focusing on applying Density Functional Theory, providing new insights to medicinal chemists to use in their projects.

Chapter 9 "***Free energy estimations for drug discovery: Background and perspectives***". This chapter is one of the most current and essential in this book. Here are shown energy predictions and applications of perturbation theory in drug design. This approach has gained increasing prominence in medicinal chemistry, mainly for solving some limitations related to classic MD simulations. In this way, an excellent theoretical framework and its application in drug design are shown with updated examples.

I hope that with this book, readers will have new insights and be able to safely apply the protocols shown here, providing new trends that help discover new drugs to improve the quality of life of the world's population.

**Igor José dos Santos Nascimento**
Programa de Pós-graduação em Ciências Farmacêuticas (PPGCF)
Departamento de Farmácia
Universidade Estadual da Paraíba
Campina Grande-PB
Brazil

## REFERENCES

[1] Nascimento IJS, de Aquino TM, da Silva-Júnior EF. The New *Era* of Drug Discovery: The Power of Computer-aided Drug Design (CADD). Lett Drug Des Discov 2022; 19(11): 951-5. [http://dx.doi.org/10.2174/1570180819666220405225817]

[2] Nascimento IJ, de Aquino TM, da Silva-Júnior EF. Drug Repurposing: A Strategy for Discovering Inhibitors against Emerging Viral Infections. Curr Med Chem 2021; 28(15): 2887-942. [http://dx.doi.org/10.2174/1875533XMTA5rMDYp5] [PMID: 32787752]

# List of Contributors


| | |
|---|---|
| **Abha Vyas** | Department of Pharmaceutical Chemistry, L.J. Institute of Pharmacy, L.J. University, Ahmedabad 382 210, India |
| **Anwesha Das** | Department of Medicinal Chemistry, National Institute of Pharmaceutical Education and Research, Ahmedabad, Palaj, Gandhinagar 382355, Gujarat, India |
| **Arijit Nandi** | Department of Pharmacology, Dr. B.C. Roy College of Pharmacy and Allied Health Sciences, Durgapur-713206, West Bengal, India |
| **Blanca Colín-Lozano** | Facultad de Ciencias Químicas, Benemérita Universidad Autónoma de Puebla, Puebla, México |
| **Burak TÜZÜN** | Plant and Animal Production Department, Technical Sciences Vocational School of Sivas, Sivas Cumhuriyet University, Sivas, Turkey |
| **Dharmraj V. Pathak** | Department of Pharmaceutical Chemistry, L.J. Institute of Pharmacy, L.J. University, Ahmedabad 382 210, India |
| **Edeildo Ferreira da Silva-Júnior** | Institute of Chemistry and Biotechnology, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões campus, 5587072-970, Alagoas, Maceió, Brazil<br>Laboratory of Medicinal Chemistry, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões campus, 5587072-970, Alagoas, Maceió, Brazil |
| **Emin SARIPINAR** | epartment of Chemistry, Faculty of Science, Erciyes University, Melikgazi, 38039, Kayseri, Turkey |
| **Esin Aki-Yalcin** | Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Cyprus Health and Social Sciences University, Guzelyurt, Northern Cyprus |
| **Fernando D. Prieto-Martínez** | Instituto de Química, Universidad Nacional Autónoma de México, Ciudad de México, México |
| **Francisco Jaime Bezerra Mendonça-Júnior** | Laboratory of Synthesis and Drug Delivery, Department of Biological Sciences, State University of Paraiba, João Pessoa 58429-500, PB, Brazil |
| **Hardik G. Bhatt** | Department of Pharmaceutical Chemistry, Institute of Pharmacy, Nirma University, Ahmedabad 382 481, India |
| **Igor José dos Santos Nascimento** | Pharmacy Department, Estácio of Alagoas College, Maceió-AL, Brazil<br>Pharmacy Department, Cesmac University Center, Maceió-AL, Brazil<br>Programa de Pós-Graduação em Ciências Farmacêuticas (PPGCF), Departamento de Farmácia, Universidade Estadual da Paraíba, Campina Grande-PB, Brazil |
| **Mallika Alvala** | MARS Training Academy, Hyderabad, India |
| **Muhammed Tilahun Muhammed** | Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Suleyman Demirel University, Isparta, Turkey |
| **Paresh K. Patel** | Department of Pharmaceutical Chemistry, L.J. Institute of Pharmacy, L.J. University, Ahmedabad 382 210, India |

| | |
|---|---|
| **Peng Zhan** | Department of Medicinal Chemistry, Key Laboratory of Chemical Biology (Ministry of Education), School of Pharmaceutical Sciences, Cheeloo College of Medicine, Shandong University, 44 West Culture Road, 250012 Jinan, Shandong, PR China |
| **Rodrigo Santos Aquino de Araújo** | Laboratory of Synthesis and Drug Delivery, Department of Biological Sciences, State University of Paraiba, João Pessoa 58429-500, PB, Brazil |
| **Ricardo Olimpio de Moura** | Programa de Pós-Graduação em Ciências Farmacêuticas (PPGCF), Departamento de Farmácia, Universidade Estadual da Paraíba, Campina Grande-PB, Brazil |
| **Samuel Baraque de Freitas Rodrigues** | Institute of Chemistry and Biotechnology, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões campus, 5587072-970, Alagoas, Maceió, Brazil |
| **Serap ÇETINKAYA** | Department of Molecular Biology and Genetics, Science Faculty, Sivas Cumhuriyet University, Sivas, Turkey |
| **Sneha R. Sagar** | Department of Pharmaceutical Chemistry, L.J. Institute of Pharmacy, L.J. University, Ahmedabad 382 210, India |
| **Thayane Regine Dantas de Mendonça** | Laboratory of Medicinal Chemistry, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões campus, 5587072-970, Alagoas, Maceió, Brazil |
| **Vijeta Kumari** | Laboratory of Natural Product Chemistry, Department of Pharmacy, Birla Institute of Technology and Science, Pilani (BITS Pilani), Pilani Campus, Pilani-333031, Rajasthan, India |
| **Yelzyn Galván-Ciprés** | Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Ciudad de México, México |

# CHAPTER 1

# Ligand and Structure-Based Drug Design (LBDD and SBDD): Promising Approaches to Discover New Drugs

**Igor José dos Santos Nascimento**[1,2,3,*] and **Ricardo Olimpio de Moura**[3]

[1] *Pharmacy Department, Estácio of Alagoas College, Maceió-AL, Brazil*

[2] *Pharmacy Department, Cesmac University Center, Maceió-AL, Brazil*

[3] *Programa de Pós-Graduação em Ciências Farmacêuticas (PPGCF), Departamento de Farmácia, Universidade Estadual da Paraíba, Campina Grande-PB, Brazil*

**Abstract:** The drug discovery and development process are challenging and have undergone many changes over the last few years. Academic researchers and pharmaceutical companies invest thousands of dollars a year to search for drugs capable of improving and increasing people's life quality. This is an expensive, time-consuming, and multifaceted process requiring the integration of several fields of knowledge. For many years, the search for new drugs was focused on Target-Based Drug Design methods, identifying natural compounds or through empirical synthesis. However, with the improvement of molecular modeling techniques and the growth of computer science, Computer-Aided Drug Design (CADD) emerges as a promising alternative. Since the 1970s, its main approaches, Structure-Based Drug Design (SBDD) and Ligand-Based Drug Design (LBDD), have been responsible for discovering and designing several revolutionary drugs and promising *lead* and *hit* compounds. Based on this information, it is clear that these methods are essential in drug design campaigns. Finally, this chapter will explore approaches used in drug design, from the past to the present, from classical methods such as bioisosterism, molecular simplification, and hybridization, to computational methods such as docking, molecular dynamics (MD) simulations, and virtual screenings, and how these methods have been vital to the identification and design of promising drugs or compounds. Finally, we hope that this chapter guides researchers worldwide in rational drug design methods in which readers will learn about approaches and choose the one that best fits their research.

**Keywords:** CADD, Computational methods, Drug design, Drug discovery, Drug Development, Docking, FBDD, LBDD, QSAR, Rational Design, SBDD.

[*] **Corresponding author Igor José dos Santos Nascimento:** Pharmacy Department, Estácio of Alagoas College, Maceió-AL, Brazil; Cesmac University Center, Pharmacy Department, Maceió-AL, Brazil; and Programa de Pós-Graduação em Ciências Farmacêuticas (PPGCF), Departamento de Farmácia, Universidade Estadual da Paraíba, Campina Grande-PB, Brazil; Tel.: (+55)8299933-5457; E-mail: igorjsn@hotmail.com & igor.nascimento@cesmac.edu.br

**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

## INTRODUCTION

The process of designing and developing new drugs is challenging and has evolved constantly in recent years, from empirical approaches related to natural products to the current phase with the use of computers and artificial intelligence [1 - 3]. One of the most significant advances in this area has been high-throughput screening (HTS), in which thousands of compounds can be screened in a few hours. In addition, the growth of genomics, proteomics, metabolomics, and molecular modeling promoted substantial advances in the knowledge of critical biochemical pathways for the P&D of drugs [4 - 6]. Associated with this, the synthetic approach exploring combinatorial chemistry could masterfully explore the available chemical space, supporting the discovery of new molecules [5]. However, the high financial cost and time-related to these approaches have driven researchers to adopt *in silico* methods [7, 8]. In this way, Computer-Aided Drug Design (CADD) emerged and perfected itself, indispensable in any new drug design discovery program [7, 9].

Traditionally, the discovery of a new drug can take between 10 and 15 years, with an investment of approximately US$800 million to US$1.8 billion [10, 11]. In this context, developing new drug design tools has become a constant quest to overcome old paradigms and speed up the discovery process at a lower financial investment [10, 12]. Over time, the scientific community accepted the new paradigm in the rational design of new drugs through CADD [13, 14]. The main reason is constant failures in the clinical evolution of prototypes identified and designed through classical techniques [13]. Thus, this paradigm shift facilitated the identification of new drugs, designing drugs with optimal physicochemical properties, and evaluating their potential *in silico* before they were synthesized [13]. With this, virtual screenings (VS) are increasingly explored, finding drug candidates in libraries of thousands of compounds. In addition, this method can be used in scaffolds identification as a starting point in molecular modeling studies, further confirming the *in silico* methods and rational design in the new era of P&D of drugs [15].

CADD can usually be divided into Structure-Based Drug Design (SBDD), and Ligand-Based Drug Design (LBDD) approaches. The researcher's choice between these approaches is related to the availability of key information about the clinical condition or known compounds against the same [16, 17]. For an SBDD protocol, the main requirement is the knowledge and availability of the target related to the explored clinical condition, in which the ligands are designed to interact with the target in question [18, 19]. On the other hand, in LBDD, there is no information about the target, but there are ligands of known activity against the clinical

condition in question, and new molecules can be designed based on the production of pharmacophoric models or Quantitative Structure-Activity Relationship studies (QSAR) [20]. Traditionally, SBDD is preferred by the scientific community mainly due to the easy access to software and the wide availability of experimental structures of biological targets [21 - 23].

Currently, CADD methods using SBDD or LBDD approaches are vital in discovering new molecules and identifying critical information in drug design. Thus, this chapter will present a historical perspective on the evolution of drug design methods to CADD approaches. We hope that this chapter will guide drug developers in deciding on the type of strategy in their studies, increasingly promoting scientific advances in rational drug design.

## DRUG DESIGN AND DISCOVERY: PAST AND TODAY METHODS AND OTHER APPROACHES

Strategies used in drug design and discovery have improved over the years [24]. In a historical context, each strategy was responsible for numerous discoveries. However, the improvement of methods made the process faster and more effective in the search for innovative molecules until the arrival of computational methods [25]. The following topics will address the evolution of the methods and their historical importance.

### Natural Compounds (NC)

Before any study of rational drug design, Natural Compounds (NC) were the primary sources of drugs explored. During the last five decades, NCs were the target of isolation or total syntheses, as they presented high biological potential and challenging structural complexity. The discovery of numerous NCs against threatening diseases like cancer and infectious diseases has increased the interest in discovering new revolutionary NCs [26]. Indeed, most drugs introduced into the pharmaceutical market since 1994 are NCs or modified synthetic analogs, highlighting their potential for many years [27].

Traditionally, drug discovery by NCs starts with testing the extract of interest in vitro or in vivo assays. After demonstrating the pharmacological effect, the responsible compounds are then isolated [27, 28]. These compounds can be modified from then on to improve their pharmacological effect [27]. In a more current approach, drug repurposing using known NCs is used to find new potentials for available structures [29, 30]. Examples of natural compounds include Artemisinin **(1)**, Atropine **(2)**, Metformin **(3),** and Quinine **(4)** (Fig. **1**). It is essential to highlight that these molecules were useful as molecular scaffolds that led to important clinical discoveries, which highlights the role of NCs in the

CHAPTER 2

# Quantitative Structure-activity Relationship (QSAR) in Studying the Biologically Active Molecules

**Serap ÇETINKAYA[1], Burak TÜZÜN[2,*]** and **Emin SARIPINAR[3]**

[1] *Department of Molecular Biology and Genetics, Science Faculty, Sivas Cumhuriyet University, Sivas, Turkey*

[2] *Plant and Animal Production Department, Technical Sciences Vocational School of Sivas, Sivas Cumhuriyet University, Sivas, Turkey*

[3] *Department of Chemistry, Faculty of Science, Erciyes University, Melikgazi, 38039, Kayseri, Turkey*

**Abstract:** Recently, many new methods have been used in the research and development of a new drug. In this article, QSAR, which is one of the usable areas of artificial intelligence during molecule research, and the analysis and formulation studies related to the suitability of this area are discussed. It is explained how a model to be created is prepared and calculation formulas for how to verify this model are shown. Examples of the most recent 4D-QSAR calculations are given.

**Keywords:** Molecular Modelling, Pharmacophore, QSAR, Quantitative Structure-activity Relationship, Validation.

## INTRODUCTION

Quantitative structure-activity relationship (QSAR) analysis uses the molecular structure of a compound or ligand to predict its biological activity. It presupposes that similar biological activities are retained in similar molecular structures [1]. It also uses known biological activity data to predict unknown activities. This approach has been adapted to diverse but related scientific disciplines [2-5], including the design of new chemical entities (NCEs) [5, 6] with high biological potentials.

* **Corresponding author Burak TÜZÜN:** Plant and Animal Production Department, Technical Sciences Vocational School of Sivas, Sivas Cumhuriyet University, Sivas, Turkey; E-mail: theburaktuzun@yahoo.com

**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

QSAR is a systematic multi-step process (Fig. **1**), made up of dataset preparation, selection, and generation of molecular descriptors; derivation of mathematical or statistical models; model training and validation using a training dataset; and model testing on a test dataset [7 - 10].

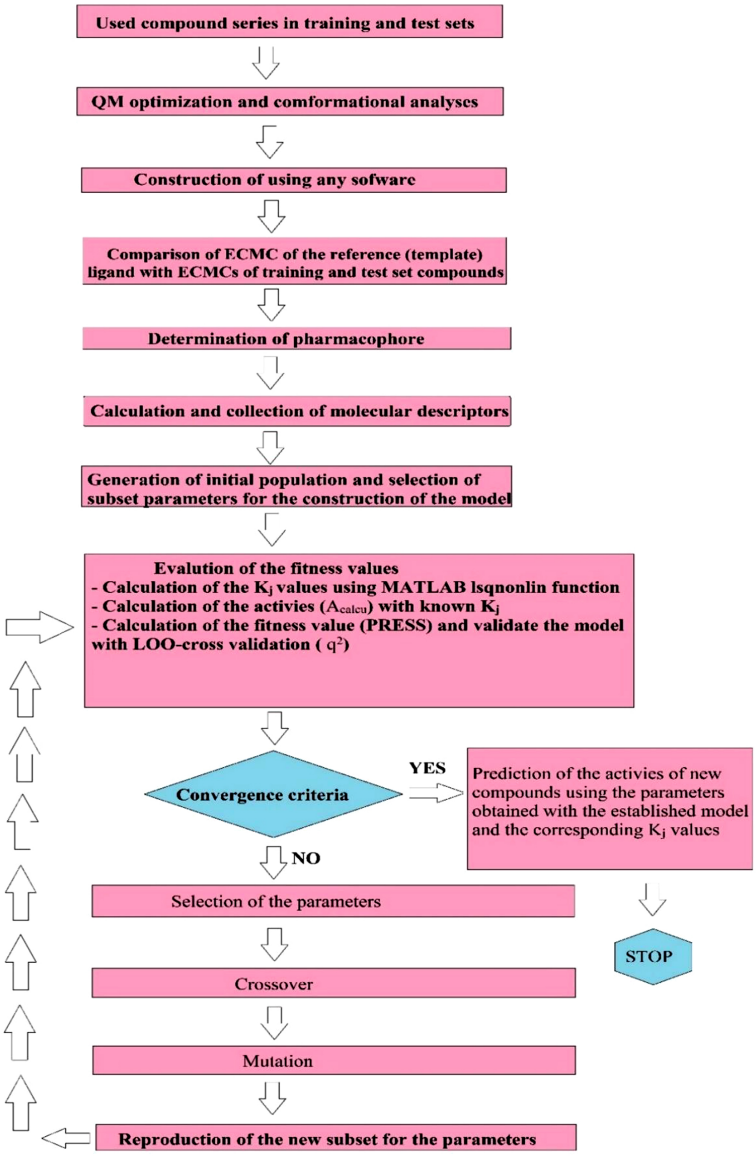



**Fig. (1).** Outlines of a QSAR model development.

In order to create a consistent QSAR model, it is central to utilize high-quality data that have been derived from bioassays, and to use an adequate number of compounds. Biological data are preferred to have been produced in a single laboratory [5, 11].

Selection and generation of molecular identifiers form the second step. Here the selection of appropriate descriptors, describing structural variations, is important. Various methods, such as machine learning techniques (*e.g.*, forward selection) and evolutionary algorithms (*e.g.*, genetic algorithm [11]), are utilized for descriptor/variable/feature selection.

A suitable mathematical or statistical model must be chosen to define the correlation between relevant descriptors and biological activities. The model can be linear partial least squares (PLS) [12], multiple linear regression (MLR) [13] or nonlinear. The selected model is then trained on a randomly chosen dataset, and the rest is used as test compounds. Model training often involves validation procedures, for example, exclusion cross validation (LOOCV) [14]. The training process is reiterated in order to reach an acceptable performance. The final step involves the testing process [11].

The concept of QSAR was first envisioned by Free, Wilson, Hansch and Fujita in 1964 [15, 16]. Subsequently, a new 3D-QSAR method, named comparative molecular field analysis (CoMFA) [17], has been worked out to overwhelm general 3D-QSAR problems. It has provided the basis for the development of multidimensional (nD) QSARs.

## QSAR's Use

QSAR should not be seen as an academic tool that allows for the subsequent rationalization of data. It aims to derive molecular structure relationships between biology and chemistry for a valid reason. Models can be developed from these relationships and are thought to be predictive with common sense, luck, and expertise. A QSAR model can have many practical commitments [18, 19]:

- Rational estimation of biological activity and physicochemical properties.
- Understand and rationalize the action mechanisms of a wide variety of chemicals.
- Cost-effective product development.
- Minimization of the production time.
- Elimination of the ethical concerns.
- Spurring of “green” chemistry.

# CHAPTER 3

# Pharmacophore Mapping: An Important Tool in Modern Drug Design and Discovery

**Dharmraj V. Pathak[1], Abha Vyas[1], Sneha R. Sagar[1], Hardik G. Bhatt[2]** and **Paresh K. Patel[1,*]**

[1] *Department of Pharmaceutical Chemistry, L.J. Institute of Pharmacy, L.J. University, Ahmedabad 382 210, India*

[2] *Department of Pharmaceutical Chemistry, Institute of Pharmacy, Nirma University, Ahmedabad 382 481, India*

**Abstract:** Computer-Aided Drug Design (CADD) has become an integral part of drug discovery and development efforts in the pharmaceutical and biotechnology industry. Since the 1980s, structure-based design technology has evolved, and today, these techniques are being widely employed and credited for the discovery and design of most of the recent drug products in the market. Pharmacophore-based drug design provides fundamental approach strategies for both structure-based and ligand-based pharmacophore approaches. The different programs and methodologies enable the implementation of more accurate and sophisticated pharmacophore model generation and application in drug discovery. Commonly used programmes are GALAHAD, GASP, PHASE, HYPOGEN, ligand scout *etc.* In modern computational chemistry, pharmacophores are used to define the essential features of one or more molecules with the same biological activity. A database of diverse chemical compounds can then be searched for more molecules which share the same features located at a similar distance apart from each other. Pharmacophore requires knowledge of either active ligands and/or the active site of the target receptor. There are a number of ways to build a pharmacophore. It can be done by common feature analysis to find the chemical features shared by a set of active compounds that seem commonly important for receptor interaction. Alternately, diverse chemical structures for certain numbers of training set molecules, along with the corresponding $IC_{50}$ or *Ki* values, can be used to correlate the three-dimensional arrangement of their chemical features with the biological activities of training set molecules. There are many advantages in pharmacophore based virtual screening as well as pharmacophore based QSAR, which exemplify the detailed application workflow. Pharmacophore based drug design process includes pharmacophore modelling and validation, pharmacophore based virtual screening, virtual hits profiling, and lead identification. The current chapter on pharmacophores also describes case studies and applications of pharmacophore mapping in finding new drug molecules of specific targets.

[*] **Corresponding author Paresh K. Patel:** Department of Pharmaceutical Chemistry, L.J. Institute of Pharmacy, L.J. University, Ahmedabad 382 210, India; Tel: +91 9712151531; E-mail: pareshpharmacist@gmail.com

**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

**Keywords:** Features, Ligand Based, Pharmacophore Query, Pharmacophore, Structure Based, Virtual Screening.

## INTRODUCTION

Currently engaged in creating a new medicine, drug design and development is a costly and time-consuming process [1]. From foundational research to commercial products, a new medicine requires 10 to 14 years of research and billions of dollars *via* several preclinical and clinical phases [2]. With the amazing advancement of computational resources, computer aided drug design (CADD) and discovery technologies are highly valued all over the world. Designing small lead and drug-like molecules with expected multitarget actions increasingly employs both ligand and structure-based methods. CADD has advanced significantly in recent years, boosting the comprehension of multiple and complicated biological processes, allowing for the fast development of novel pharmacologically active drugs [3]. One such CADD tool employed in drug design and discovery is pharmacophore mapping or pharmacophore modeling. In the late 19th century, *Paul Ehrlich* was the first who propose that certain groups inside a molecule (phoros) are responsible for a molecule's biological activity (pharmacon), giving rise to the idea of “pharmacophores” [4, 5]. The pharmacophore theory postulates that a collection of shared properties that engage a group of contrasting locations on a biological target can explain how a class of chemicals recognizes that target on a molecular level [6]. In the contemporary drug discovery process, the pharmacophore approach serves as a helpful bridge between medicinal chemistry and computational chemistry, both in virtual screening (VS) and library design for effective hit finding and in the optimization of lead compounds to final therapeutic candidates.

### Definitions of Pharmacophore

As per the IUPAC definition, “A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response.”

Apart from the official IUPAC definition, other similar definitions have also been given in the literature. “A pharmacophore does not represent a real molecule or a real association of functional groups, but a purely abstract concept that accounts for the common molecular interaction capacities of a group of compounds with their target structure.”

"A pharmacophore is defined by pharmacophoric descriptors, including H-bonding, hydrophobic, and electrostatic interaction sites, defined by atoms, ring centers, and virtual points.

"A pharmacophore can be considered the largest common denominator shared by a set of active molecules". This definition discards a misuse often found in the medicinal chemistry literature, which consists of naming as pharmacophores simple chemical functionalities such as guanidine, sulphonamides, or dihydroimidazoles (formerly imidazolines), or typical structural skeletons such as flavones, phenothiazines, prostaglandins, or steroids [5, 7].

To describe unique functional groups or chemical classes with biological activity, scientists frequently use the terms "pharmacophore" or "pharmacophoric group. In this context, the word "pharmacophore" is used in conjunction with the concept of "privileged structures," which refers to the alternative idea of structure and function. The chemical scaffolds and retroactive examination of medicinal molecules' chemical structures allowed for the identification of a few structural motifs that are frequently linked to bioactive compounds. Evans *et al.* referred to these patterns as "privileged structures" to describe substructures that bestow activity against a number of different targets [8]. Dihydropyridines, arylethylamines, *N*-arylpiperazines, diphenylmethane derivatives, biphenyls, pyridazines, sulphonamides and benzodiazepines are a few well-known instances of the advantaged structures [7 - 10].

## Pharmacophore: History

The pharmacophore was first envisioned by Paul Ehrlich, the pioneer of chemotherapy, and that idea has remained unchanged for the past 100 years [11]. Langley, who coined the phrase "receptive substance," first proposed the notion that bioactive compounds interact with receptors in 1878 [12]. Paul Ehrlich, meanwhile, coined the word "receptor" a few years down the line [13], as well as introduced the term "pharmacophore". In conjunction with Emil Fischer's lock-and-key concept, it tends to be evident but not the properties of a molecule, the "key", are equally significant aimed at biological action [14]. Biological activity can be dramatically altered by small changes in some parts of a molecule, while minor changes in others can do the same. Modern drug discovery and development is based on Langley, Ehrlich, and Fischer's concepts. As soon as they were confirmed according to the earliest protein-ligand complex crystal structures half a decade later, they established a new paradigm [15]. Before the development of computers and modelling software, simple pharmacophores were documented in the literature and recognised as tools for the discovery of novel compounds. Modest 2D models remained first proposed in the 1940s based on the

# CHAPTER 4

# Up-to-Date Developments in Homology Modeling

**Muhammed Tilahun Muhammed**[1,*] and **Esin Aki-Yalcin**[2]

[1] *Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Suleyman Demirel University, Isparta, Turkey*

[2] *Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Cyprus Health and Social Sciences University, Guzelyurt, Northern Cyprus*

**Abstract:** Homology modeling is used to predict protein 3D structure from its amino acid sequence. It is the most accurate computational approach to estimate 3D structures. It has straightforward steps that save time and labor. There are several homology modeling tools under use. There is no sole tool that is superior in every aspect. Hence, the user should select the most appropriate one carefully. It is also a common practice to use two or more tools at a time and choose the best model among the resulting models.

Homology modeling has various applications in the drug design and development process. Such applications need high-quality 3D structures. It is widely used in combination with other computational methods including molecular docking and molecular dynamics simulation. Like the other computational methods, it has been influenced by the involvement of artificial intelligence. In this regard, homology modeling tools, like AlphaFold, have been introduced. This type of method is expected to contribute to filling the gap between protein sequence release and 3D structure determination.

This chapter sheds light on the history, relatively popular tools and steps of homology modeling. A detailed explanation of MODELLER is also given as a case study protocol. Furthermore, homology modeling's application in drug discovery is explained by exemplifying its role in the fight against the novel Coronavirus. Considering the new advances in the area, better tools and thus high-quality models are expected. These, in turn, pave the way for more applications of it.

**Keywords:** Computer Aided Drug Design, 3D Structure, Drug Discovery, Homology Modeling, Modeller, Molecular Modeling.

[*] **Corresponding author Muhammed Tilahun Muhammed:** Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Suleyman Demirel University, Isparta, Turkey; E-mail: muh.tila@gmail.com

**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

## INTRODUCTION

Proteins' 3D (3-dimensional) structures play a crucial role in defining their functions [1]. Hence, investigations about protein structures have an important contribution to understanding the mechanism of diseases [2]. Thus, knowledge about protein 3D structure has a vital role in rational drug design and discovery [3]. As a result, a number of Nobel Prizes have been awarded to researchers in this area. Scientists have been awarded the prize for elucidating the structures of myoglobin, lysozyme, integral membrane protein, HIV (human immunodeficiency virus) protease, ion channels, RNA (ribonucleic acid) polymerase, and GPCR (G protein-coupled receptor). In addition to this, the prize was awarded to researchers who pioneered in using X-ray crystallography, NMR (nuclear magnetic resonance), and Cryo-electron microscopy (Cryo-EM) for protein structure determination [4].

The quality of protein 3D structures solved has been improved as the available techniques improved [5]. Together with this, the experimental methods are not applicable to solving the structure of each protein. In this regard, NMR is used to solve the 3D structure of relatively small molecules, which are dissolvable [6]. Similarly, X-ray crystallography is used to solve protein 3D structures in a crystal state [7]. Cryo-EM is preferred to large macromolecule complexes with low resolution [8]. In addition to this, the experimental methods take a long time, labor and resource [9]. As a result, the experimental protein 3D structure determination could not keep pace with the protein sequence release. Consequently, the gap between the protein sequences available and the experimentally solved protein structures has been widening. Hence, computational protein 3D structure prediction methods can play a substantial role in filling this gap [10].

Homology (comparative) modeling is protein 3D structure prediction from its amino acid sequence. Homology modeling is used when the query sequence and templates selected share a common ancestor. In comparative modeling, there is just sequence similarity without shared ancestral history [11]. Homology modeling yields 3D structures with better reliability than the other computational structure prediction approaches [12, 13]. In addition to this, it has straightforward steps that take relatively less time. Hence, homology modeling is used to generate high quality structures that have the potential to convert the applications of the other computational methods in case they require 3D structures [14].

In this chapter, the brief history and the general procedures of homology modeling are presented. Homology modeling tools that are widely used these days are also given. Together with this, a case study protocol with MODELLER is included.

Furthermore, applications of homology modeling in the drug discovery process is summarized with a special focus on the latest ones. So, this chapter is expected to provide updated information on homology modeling.

## BRIEF HISTORY OF HOMOLOGY MODELING

The idea of protein structure prediction has a long history since 1894 when Emin Fischer suggested that a protein's 3D structure determines its function [11]. Thereafter, Christian Anfinsen suggested that among the possible conformations, the native conformation has the lowest energy. In the 1970s, he proposed that the structure of a protein is determined by its amino acid sequence in a particular physiological condition [15]. This is the basis for the concept of homology modeling. The α-lactalbumin 3D structure, which was built based on the structure of lysozyme in 1969, is considered the first homology model [16]. After this time, various homology modeling programs and servers were developed. In this regard, MODELLER was revealed in 1993 [17]. In the same year, the concept of a server for automated homology modeling was introduced through SWISS-MODEL [18]. The milestones in the history of homology modeling are summarized in Fig. (**1**).

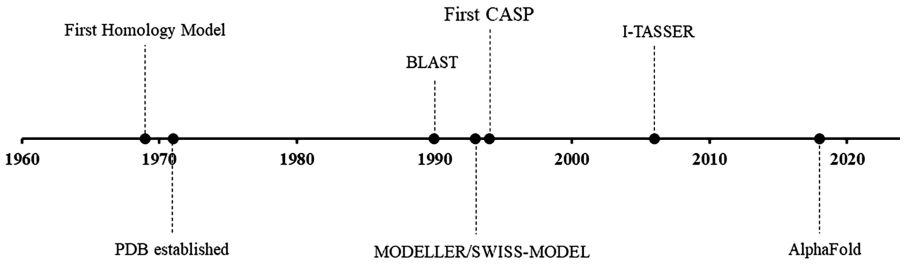



**Fig. (1).** Milestones in the history of homology modeling.

## HOMOLOGY MODELING PROCEDURE

Homology modeling has straightforward major steps (Fig. **2**). General information about each step is presented in this section.

### Identification and Selection of templates

In the first step of the process, the target (query) sequence is used to identify template structures in the worldwide PDB (https://www.wwpdb.org/) or other structural databases [19]. First, the protein basic local alignment search tool (BLASTp) search is performed by using the target sequence as a query and PDB as a database in NCBI (national center for biotechnology information) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) [20, 21]. BLASTp search gives the 3D structures inside the PDB with high identity and coverage of the query. In case

CHAPTER 5

# Anticancer Activity of Medicinal Plants Extract and Molecular Docking Studies

**Serap ÇETINKAYA**[1] and **Burak TÜZÜN**[2,*]

[1] *Department of Molecular Biology and Genetics, Science Faculty, Sivas Cumhuriyet University, Sivas, Turkey*

[2] *Plant and Animal Production Department, Technical Sciences Vocational School of Sivas, Sivas Cumhuriyet University, Sivas, Turkey*

**Abstract:** Molecular docking involves the interaction of a molecule with another place, usually in the protein structure, and simulating the placement of the molecule in the protein structure with certain score algorithms, taking into account many quantities, such as the electro-negativity of atoms, their positions to each other, and the conformation of the molecule to be inserted into the protein structure. Finally, the activity of the molecule with the highest percentage by mass against various cancer proteins was investigated according to the GC-MS results made on some medicinal and aromatic plants in order to set an example of molecular docking calculations.

**Keywords:** Activity, Aromatic plants, Cancer proteins, Molecular docking, Medicinal.

## INTRODUCTION

Molecular docking involves the interaction of a molecule with another place, usually in the protein structure, and consists of simulating the placement of the molecule in the protein structure with certain score algorithms, taking into account many quantities such as the electro-negativity of the atoms, the positions of the atoms to each other, and the conformation of the molecule to be inserted into the protein structure [1, 2].

The docking process plays an important role in explaining the receptor-ligand, enzyme-ligand relationship. Finding suitable antagonist and agonist compounds has an important place in enzyme inhibition studies [3].

* **Corresponding author Burak TÜZÜN:** Plant and Animal Production Department, Technical Sciences Vocational School of Sivas, Sivas Cumhuriyet University, Sivas, Turkey; E-mail: theburaktuzun@yahoo.com

**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

Molecular docking studies are the theoretical method that plays an important role in determining whether millions of synthesized compounds are effective drug substances [4]. It is impossible to study each of the millions of chemical substances individually *in vitro*, and molecular docking studies have a very important role in selecting the most effective substances.

By converting the effect of a chemical substance on a protein structure into a numerical value, it saves money by preventing *in vitro* and *in vivo* studies of molecules that are impossible to be effective [5, 6]. It also lays the groundwork for the modification of the molecule with the correct estimation of the binding modes of the relevant molecule, and creates a strategic infrastructure by guiding the synthesis of molecules that are likely to be more effective.

The docking process plays an important role in explaining the receptor-ligand, enzyme-ligand relationship [7, 8]. It enables the comparison of the activities of molecules against proteins in studies to inhibit the enzyme in finding suitable antagonist compounds.

In addition to all these important and useful features of molecular docking studies, it also needs to be supported by molecular dynamics. Because the molecule clamped into the protein structure may have achieved good coupling and high scores in the first place, but both the enzyme and the relevant molecule in the solvent are in interaction [9]. This dynamic and synergetic state means that the chelating molecule cannot stay in the docked place for a long time, and its effect will be limited as it is related to the residence time in the attached area. Due to this situation, molecular docking calculations in computational chemistry are supported by molecular dynamics and the problem is solved.

## Computer Aided Drug Design (CADD)

*In silico* methods are increasingly used for the development of new drugs. Computer-aided drug design (CADD) [10, 11] is a discipline that uses computational methods to simulate drug-receptor, drug-enzyme interactions. Calculations made by examining the 3-D properties of chemical molecules accelerate the optimization process of precursor compounds [12]. Thus, the success rate in drug research and development (R&D) studies increases, R&D costs decrease and R&D period shortens [13].

Computer-aided drug design programs require knowledge of ligands and receptors; bioinformatics develops depending on tools, applications and databases. If a target (receptor) exists, its 3D structure (by x-ray or NMR) together with its ligand must be known; If there is no experimental data, the 3D structure

of the target molecule is tried to be created by homology modeling based on the sequence data.

There are two basic approaches to drug design: ligand-based and receptor-based molecular design methods.

### *Ligand-based Approach*

The second and more branched approach to drug discovery is the ligand-based route. The general assumption of ligand-based methods [14, 15] is that the active site of a target protein may have similar atoms, functional groups, or moieties to have 11 similar functional properties. Nitrogen atoms in a Histidine residue at a particular position in the protein sequence must make a Hydrogen bond interaction with a polar Hydrogen atom on the ligand for the protein to lose its biological function (also called "inhibition") [16]. Of course, the change in the properties of a protein cannot be brought about by a single interaction on a single atom. However, if this approach is embodied for an entire molecule that has several interactions with more than one amino acid in the binding gap, the desired switch of function can be established. One of the first uses of ligand-based methods is seen as structure-activity relationship (SAR) studies, a method that has been used for decades [17].

However, the problem with the activity of small molecules in the body is that it cannot be predicted with sufficient accuracy. The reasons behind this disadvantage are [18]:

i) A full quantum-mechanical description of a ligand *(i.e*., accurate calculation of the partial charges on each of its atoms) cannot be made,

ii) Actual activity depends on numerous factors such as: the character of the target, its environment, and the interactions established between a target and the ligand.

Perhaps the most promising avenues in a ligand-based approach are 3D pharmacophore modeling or 3D quantitative structure-activity relationship (QSAR) methods. Pharmacophore modeling encompasses the discovery of the spatial arrangement of pharmacophore groups in a molecule because that molecule is considered biologically active or relevant. The term "pharmacophore" was first defined by Schueler in the 1960s [19, 20] as functionalities in a molecule that determine its biological activity [21].

These chemical groups responsible for the activity of a drug molecule can be searched for and compared with desired activities through chemical libraries

# CHAPTER 6

# FBDD & *De Novo* Drug Design

**Anwesha Das[1,†], Arijit Nandi[2,†], Vijeta Kumari[3]** and **Mallika Alvala[4,*]**

[1] *Department of Medicinal Chemistry, National Institute of Pharmaceutical Education and Research, Ahmedabad, Palaj, Gandhinagar 382355, Gujarat, India*

[2] *Department of Pharmacology, Dr. B.C. Roy College of Pharmacy and Allied Health Sciences, Durgapur-713206, West Bengal, India*

[3] *Laboratory of Natural Product Chemistry, Department of Pharmacy, Birla Institute of Technology and Science, Pilani (BITS Pilani), Pilani Campus, Pilani-333031, Rajasthan, India.*

[4] *MARS Training Academy, Hyderabad, India*

**Abstract:** Fragment-based drug or lead discovery (FBDD or FBLD) refers to as one of the most significant approaches in the domain of current research in the pharmaceutical industry as well as academia. It offers a number of advantages compared to the conventional drug discovery approach, which include – 1) It needs the lesser size of chemical databases for the development of fragments, 2) A wide spectrum of biophysical methodologies can be utilized for the selection of the best fit fragments against a particular receptor, and 3) It is far more simpler, feasible, and scalable in terms of the application when compared to the classical high-throughput screening methods, making it more popular day by day. For a fragment to become a drug candidate, they are analyzed and evaluated on the basis of numerous strategies and criteria, which are thoroughly explained in this chapter. One important term in the field of FBDD is *de novo* drug design (DNDD), which means the design and development of new ligand molecules or drug candidates from scratch using a wide range of *in silico* approaches and algorithmic tools, among which AI-based platforms are gaining large attraction. A principle segment of AI includes DRL that finds numerous applicabilities in the DNDD sector, such as the discovery of novel inhibitors of BACE1 enzyme, identification and optimization of new antagonists of DDR1 kinase enzyme, and development and design of ligand molecules specific to target adenosine A2A, *etc*. In this book chapter, several aspects of both FBDD and DNDD are briefly discussed.

**Keywords:** Artificial Intelligence, Autoencoder, Deep Learning, *De Novo* Drug Design, Drug Development, Drug Discovery, Evaluation Criteria, Expansion, Fragment-based Fragment to Lead, Hotspot analysis, *In silico*, Lead Optimization, Machine Learning, Molecular Docking, Optimization, Pharmacokinetic Properties, Property Prediction, Synthetic Accessibility.

[*] **Corresponding author Mallika Alvala:** MARS Training Academy, Hyderabad, India;
E-mail: mallikaalvala@yahoo.in
[†] *Authors with equal contribution.*

**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

## INTRODUCTION

Since the last two decades, FBDD or FBLD has become one of the most triumphant methodologies in the area of early-stage drug development in the pharmaceutical industry as well as academia [1]. FBDD constitutes the screening of numerous molecules with lower molecular weights against clinically significant biological targets as these smaller fragments may fit into one or multiple binding sites of the protein and can act as potential beginning points in case of lead development. For fragment development, the physicochemical, pharmacokinetic, and toxic properties must be considered. One of the most popular methods, structure-based fragment screening, firstly employs a combination of multiple techniques, such as biophysical methods (thermophoresis, surface plasmon resonance [SPR], and differential scanning fluorimetry [DSF] *etc*.), later the employment of experimentations like X-ray crystallography or NMR, optimizes and structurally characterizes the fragments. Followed by that, further analytical stages like fragment growth also require the structural characterization of the screened hit fragments. The entire workflow of FBLD includes a massive high-throughput screening of all fragments that ultimately leads to the lead compound, and this approach is known as fragment-to-ligand optimization (F2L approach).

FBDD is referred to as one of the most attractive, effective, and popular approaches for chemical space exploration for perfectly fitting into the binding site of a biological target. While in the case of classical high-throughput screening (HTS), the screening of large libraries of complex molecules takes place against a target [2], in the case of FBDD, smaller libraries of lesser complex molecules that make fragments of larger drug-like molecules are usually screened against the target binding site for evaluating their binding efficiencies [3]. In spite of having lower potency than the larger drug-like compounds obtained *via* HTS, the fragments are considered potential starting points for designing larger drug-like molecules with higher affinity towards the target using the prior knowledge of the targets. This downside-up approach yields lead compounds with higher affinity and specificity, where a greater range of chemical space can be explored. Another advantage of FBDD includes that it requires lower expenses and lesser time for drug development through FBDD approach [4]. For example, Vemurafenib (Zelboraf™) is the first FBDD-derived drug that took only 6 years in all phases of the drug discovery pipeline before it went to FDA approval [5]. NMR can also be used in FBDD; for example, Bruker's Ligand Observed NMR is one of the most popular techniques for FBDD [6]. In the case of the computationally derived FBDD-approach, numerous tools can be employed for rationally designing a molecule. For example, AutoGrow4 is a genetic algorithm-based open source platform that can predict and design ligands computationally [7, 8]. Moreover, LigBuilder employs computational approaches to design ligands that can bind to

multiple targets, multiple binding sites of a single target, or multiple conformations of a single target, thus forming a multi-target directed ligand (MTDL) [9]. This way, FBDD offers numerous attractive opportunities in the domain of drug discovery.

On the other hand, *De novo* drug design (DNDD) refers to the design of novel molecules that perfectly fit into a protein's binding site using several computational algorithms and approaches [10]. The meaning of the word "*De Novo*" is "starting from scratch or from the beginning", which implies that in DNDD, novel chemicals can be designed without any prior information of the starting point [11]. Among the several advantages of DNDD, such as larger chemical space exploration, new intellectual property containing compound design, time- and cost-effective development of novel chemical entities, and the strength of newer improved therapies as well as therapeutics, *etc*., it shows one major disadvantage or challenge of synthesizability [12]. In this book chapter, several aspects of both FBDD and DNDD are briefly discussed.

## TYPES OF DRUG DESIGN

DNDD can be defined as a drug designing methodology where new chemical entities (NCE) can be found from scratch from either the information related to the enzyme/receptor/biological target or its already known ligands having a strong inhibitory activity or good binding affinity towards the enzyme [13 - 25]. Needless to say, the main workflow behind the DNDD approach is - 1) A proper description and demonstration of the target's active binding site, 2) Pharmacophore modeling of the binding ligands, 3) Construction or generation of ligands by sampling, and 4) Evaluation of the constructed ligands. Principally, there are two types of DNDD approaches, namely, structure or receptor-based drug design (SBDD) and ligand-based drug design (LBDD).

### Structure or Receptor-based Drug Design (SBDD)

SBDD is based on the three-dimensional structure (3D) of the biological target, where its structure is elucidated mainly by three methods, viz, electron microscopy, Nuclear magnetic resonance (NMR), and X-ray crystallography [26, 27]. Principally, SBDD starts with the determination of the receptor's active site. It is one of the most significant steps in SBDD as the reduction in the higher number of generated conformers and structures improves the specificity and selectivity towards the ligand. This specificity and tightness of the ligand binding at the receptor's active site are governed by the shape of the ligand molecule and its physical and chemical properties (non-covalent interactions, such as

# CHAPTER 7

# Molecular Simulation in Drug Design: An Overview of Molecular Dynamics Methods

**Fernando D. Prieto-Martínez[1,*], Yelzyn Galván-Ciprés[2]** and **Blanca Colín-Lozano[3]**

[1] *Instituto de Química, Universidad Nacional Autónoma de México, Ciudad de México, México*

[2] *Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Ciudad de México, México*

[3] *Facultad de Ciencias Químicas, Benemérita Universidad Autónoma de Puebla, Puebla, México*

**Abstract:** Molecular interaction is the basis for protein and cellular function. Careful inhibition or modulation of these is the main goal of therapeutic compounds. In the pharmaceutical field, this process is referred to as pharmacodynamics. Over the years, there have been several hypotheses attempting to describe this complex phenomenon. From a purely biophysical point of view, molecular interactions may be attributed to pairwise contributions such as charge angles, torsions, and overall energy. Thus, the computation of binding affinity is possible, at least in principle. Over the last half of the past century, molecular simulation was developed using a combination of physics, mathematics, and thermodynamics. Currently, these methods are known as structure-based drug design (SBDD) and it has become a staple of computer-aided drug design (CADD). In this chapter, we present an overview of the theory, current advances, and limitations of molecular dynamics simulations. We put a special focus on their application to virtual screening and drug development.

**Keywords:** Drug Design, Enhanced Sampling, Molecular Interaction, Molecular Simulation.

## INTRODUCTION

Traditional methods for drug development often involve a multidisciplinary approach; the process usually begins by selecting what is known as a drug target and the consequent study of its biochemistry. Then comes the molecular design followed by organic synthesis, and subsequently, pre-clinical *in vitro*, *ex vivo,* and *in vivo* studies are carried out, when possible, depending on the task at hand. After

[*] **Corresponding author Fernando D. Prieto-Martínez:** Instituto de Química, Universidad Nacional Autónoma de México, Ciudad de México, México; E-mail: ferdpm4@hotmail.com

**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

gathering enough information, the research team can now decide which compounds can be considered as leads and a pharmacophore is then identifiable. Later, the ADMET (absorption, distribution, metabolism, elimination and toxicity) properties are optimized for clinical trials to finally market the best compound as a drug [1]. It should be noted, however, that this route represents a challenging, long, and expensive process that, on average, takes up to 15 years. Since 2019, the average cost of developing a drug can go from $161 to $4,540 million dollars, being anticancer drugs the most expensive to develop with a cost that goes between $944 and $4,540 million dollars [2].

Interestingly, around 90% of clinical drug development fails due to poor ADMET properties: absorption, solubility, permeability, efficacy, metabolism, excretion and high toxicity [3].

As a result of the above, there are several more novel strategies for discovering new drug candidates, such as: optimization of existing drugs, drug repurposing, systematic biological assays, use of available biological information, rational drug design and computer-aided drug design (CADD) also called *in silico* drug discovery methods.

The term *in silico* comes from Latin 'in silicon' and it refers to performed by using computers or *via* computer simulation. A mathematician Pedro Miramontes from the Universidad Nacional Autónoma de México (UNAM), who was the son of Luis Ernesto Miramontes Cárdenas, responsible for the synthesis of the active pharmaceutical ingredient (API) norethisterone of the first anticontraceptive pill, presented in his talk "DNA and RNA Physicochemical Constraints, Cellular Automata and Molecular Evolution" the term "*in silico*" to explain biological experiments performed *via* computer simulation [4].

CADD provides a complement to explain and predict biological activities and it comprises various methods, such as QSAR, virtual high-throughput screening, pharmacophore modeling, fragment-based screening, molecular docking and molecular dynamics simulations (MDS). It is worth mentioning that in recent years, these techniques have been put in the spotlight since they considerably reduce time and costs in all stages of drug development from the initial lead design to final stage clinical trials. Particularly in 2021, *in silico* drug discovery methods have gained popularity, as they have made it possible to optimize research work even remotely, which is a very useful tool in complex scenarios such as the COVID-19 pandemic.

Molecular recognition processes arise from pairwise interactions. Physical descriptions of these are possible using potential energy functions. In the literature, these functions are referred to as force fields, serving as angular stones

in molecular mechanics and other related methods. In CADD, these approaches are grouped under structure-based (SBDD) approaches where computational resources are used to make numerical simulations of molecular phenomena. As of today, molecular docking has become the most prominent method for SBDD efforts, mostly due to its ease of implementation, flexibility, and overall prompt results.

Nevertheless, even with such positive attributes, there is no denying that molecular docking is prone to erratic or even aberrant results. Moreover, the technique has been trivialized in recent years. This has led to what we may call 'literature flooding', as evidenced in the trend for keyword docking (Fig. **1**) in recent years. Of course, the causes for this are multifactorial; still, a conserved tendency seems to be an overreliance on docking scores.

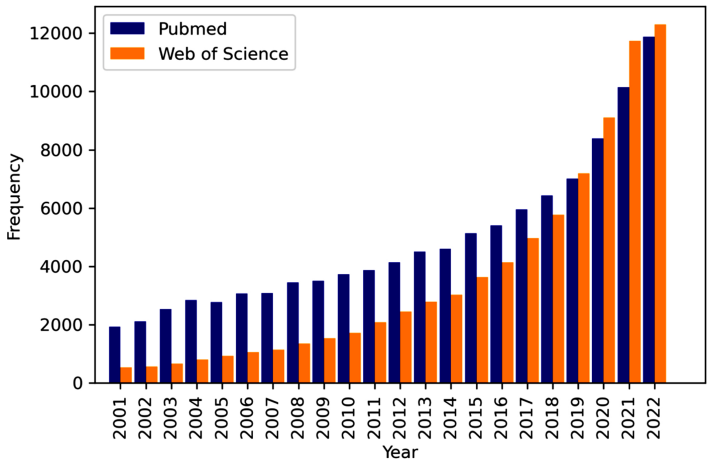



**Fig. (1).** Trends for "molecular docking" as a search query from two major academic search engines.

The truth of the matter is that the development of these tools has proven to be a complex task. Even with recent implementations of artificial intelligence, the development of universal scoring functions remains daunting. This raises the need for more exhaustive methods, such as MM-PBSA/GBSA and free energy perturbations (see Chapter 9).

For decades, one of the main problems was the computing power needed to solve the equations of motion for the *N*-atoms composing the system. In the beginning, a rather small system, *i.e.*, a couple of thousand atoms, could take months to simulate the movement for a couple of hundred picoseconds. Recently, thanks to technological advances, it has been possible to build powerful workstations that are on par with the last generation high computing clusters (HPC). In sharp contrast, on today's hardware, even a rather "discrete" workstation can increase up

# CHAPTER 8

# Quantum Chemistry in Drug Design: Density Function Theory (DFT) and Other Quantum Mechanics (QM)-related Approaches

**Samuel Baraque de Freitas Rodrigues[1], Rodrigo Santos Aquino de Araújo[2], Thayane Regine Dantas de Mendonça[3], Francisco Jaime Bezerra Mendonça-Júnior[2], Peng Zhan[4]** and **Edeildo Ferreira da Silva-Júnior[1,3,*]**

[1] *Institute of Chemistry and Biotechnology, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões campus, 5587072-970, Alagoas, Maceió, Brazil*

[2] *Laboratory of Synthesis and Drug Delivery, Department of Biological Sciences, State University of Paraiba, João Pessoa 58429-500, PB, Brazil*

[3] *Laboratory of Medicinal Chemistry, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões campus, 5587072-970, Alagoas, Maceió, Brazil*

[4] *Department of Medicinal Chemistry, Key Laboratory of Chemical Biology (Ministry of Education), School of Pharmaceutical Sciences, Cheeloo College of Medicine, Shandong University, 44 West Culture Road, 250012 Jinan, Shandong, PR China*

**Abstract:** Drug design and development are expensive and time-consuming processes, which in many cases result in failures during the clinical investigation steps. In order to increase the chances to obtain potential drug candidates, several *in silico* approaches have emerged in the last years, most of them based on molecular or quantum mechanics theories. These computational strategies have been developed to treat a large dataset of chemical information associated with drug candidates. In this context, quantum chemistry is highlighted since it is based on the Schrödinger equation with mathematic solutions, especially the Born-Oppenheimer approximation. Among the Hartree-Fock-based methods, the Density Functional Theory (DFT) of Hohenberg-Kohn represents an interesting and powerful tool to obtain accurate results for electronic properties of molecules or even solids, which in many cases are corroborated by experimental data. Additionally, DFT-related methods exhibit a moderate time-consuming cost when compared to other *ab initio* methods. In this chapter, we provide a deep overview focused on the formalism behind DFT, including historical aspects of its development and improvements. Moreover, different examples of the application of DFT in studies involving GABA inhibitors, or catalytic mechanisms of enzymes, such as RNA-dependent RNA polymerase (RdRp) of SARS-CoV-2, and different proteases

[*] **Corresponding author Edeildo Ferreira da Silva-Júnior:** Institute of Chemistry and Biotechnology, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões campus, 5587072-970, Alagoas, Maceió, Brazil; & Laboratory of Medicinal Chemistry, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões campus, 5587072-970, Alagoas, Maceió, Brazil; Tel.: +55-82-9-9610-8311; E-mail: edeildo.junior@iqb.ufal.br

**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

associated impacting diseases, such as malaria, Chagas disease, human African trypanosomiasis, and others. Moreover, the role of metal ions in catalytic enzymatic mechanisms is also covered, discussing iron-, copper-, and nickel-catalyzed processes. Finally, this chapter comprises several aspects associated with the elucidation of catalytic mechanisms of inhibition, which could be used to develop new potential pharmacological agents.

**Keywords:** Catalytic Mechanism, Copper, Hydrolase, Nickel, Protease, Quantum Mechanics.

## INTRODUCTION TO THE HISTORY OF QUANTUM CHEMISTRY (QC)

Currently, there is a constant rise in the need for growing efficiency in a drug design and discovery campaign or even during the lead optimization since the central idea is reducing the time costs, yielding more effective drugs in the pipeline. Thus, this increased necessity requests accurate software for processing a large amount of information in a limited time, overstimulating software upgrades, and the development of novel protocols using well-known programs. In this context, different methods have been developed to treat a large dataset of chemical information, aiming to fill the lack that emerged during the development of a new drug. Well, before discussing how some computer-aided methods can help to elucidate essential information for designing drugs, we need a better understanding of what formalism these methods are based on, as well as their possible applicability. In the next pages, this chapter will lead the reader on a journey from the emergence of the most important computer methods and their current utilization focused on drug design and development, starting from the *Schrödinger equation.*

The main point of *quantum chemistry* (QC) is the obtainment of solutions for the Schrödinger equation to accurately determine the chemical properties of atoms and even more complex molecular systems. Then, we typically are searching solutions for stationary states that could involve different approximation methods. Thus, QC methods depend not only on computer advances but also on the development of new theories or methodologies. Currently, there are several methods involving QC for solving chemical problems associated with molecules, among them, the *ab initio* Hartree-Fock (HF) has been used to provide great approximated solutions for *many-electron problems*. Its theory treats the electrons individually, moving in an average field for all other electrons and nuclei, which allows the generation of a set of electron-coupled equations. Years later, *semi-empirical methods* emerged to reduce computational time-consuming. Otherwise, chemical problems that previously were treated with HF approximation are currently frequently treated by using the *Density Functional Theory* (DFT) calculations, resulting in values even closer to the experimental data. It has been

used to study the electronic properties of molecules and solids. Furthermore, the development of more precise exchange correlation functionals and efficient algorithms of numerical integration has contributed to the development of the DFT method.

In 1927, Max Born and J. Robert Oppenheimer formulated the *Born-Oppenheimer approximation*, which assumes that the nuclei are much heavier than electrons and, as a consequence, they move more slowly, making this theory considered the heart of QC [1]. Considering this approximation, its main problem still remains in solving the non-relativistic time-independent Schrödinger equation:

$$\widehat{H} \mid \Phi\rangle = \varepsilon \mid \Phi \rangle$$

in which,

$$\widehat{H} = -\sum_{i=1}^{N} \frac{\hbar^2}{2m} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A e^2}{4\pi\epsilon_0 r_i A} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{e^2}{4\pi\epsilon_0 r_{ij}}$$

Where $m$ represents the electron mass, $Z_A$ means the atomic number of the nucleus $A$, $r_{ij}$ is the distance between $i$ and $j$ electrons, whereas $r_{iA}$ means the distance between electron $i$ and nucleus $A$. Finally, $N$ and $M$ represent the number of electrons and nuclei in the system, respectively. The above equation expresses the electronic term for the molecular *Hamiltonian operator* $\widehat{H}$. Since the electrons in a molecule are considered moving faster than nuclei, the second term of this equation (kinetic energy of the nuclei) can be neglected. Moreover, the repulsion between the nuclei (the last term) is taken to be constant. Thusly, the remaining terms are called the *electronic Hamiltonian* [2]:

$$\widehat{H}_{elec} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}}$$

Then, HF approximation has an essential role in the development of modern QC concepts [2]. Douglas Hartree's methods were guided by some earlier semi-empirical methods of the early 1920s set in the old quantum theory of Bohr [3]. In general, HF approximation substitutes the many-electron problem with a one-electron problem, considering the electron-electron repulsion term as an average way [2]. In this context, the *Self-Consistent-Field* (SCF) method is used as a procedure for solving the HF equation. In essence, this approach creates an initial guess for the spin orbitals, from which it can calculate the average field seen by each electron, solving the eigenvalue equation for a new set of spin orbitals.

**CHAPTER 9**

# Free Energy Estimation for Drug Discovery: Background and Perspectives

**Fernando D. Prieto-Martínez**[1,*] and **Yelzyn Galván-Ciprés**[2]

[1] *Instituto de Química, Universidad Nacional Autónoma de México, Ciudad de México, México*

[2] *Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Ciudad de México, México*

**Abstract:** Drug development is a remarkably complex subject, with potency and specificity being the desired traits in the early stages of research. Yet, these need careful thought and rational design, which has led to the inclusion of multidisciplinary efforts and non-chemistry methods in the ever-changing landscape of medicinal chemistry. Computational approximation of protein-ligand interactions is the main goal of the so-called structure-based methods. Over the years, there has been a notable improvement in the predictive power of approaches like molecular force fields. Mainstream applications of these include molecular docking, a well-known method for high-throughput virtual screening. Still, even with notable success cases, the search for accurate and efficient methods for free energy estimation remains a major goal in the field. Recently, with the advent of technology, more exhaustive simulations are possible in a reasonable time. Herein, we discuss free energy predictions and applications of perturbation theory, with emphasis on their role in molecular design and drug discovery. Our aim is to provide a concise but comprehensive view of current trends, best practices, and overall perspectives in this maturing field of computational chemistry.

**Keywords:** Alchemistry, Computer-aided Drug Design, Free Energy Methods and Simulation.

## INTRODUCTION

As of today, drug development is a multidisciplinary field where the areas of competence go beyond pharmacology or organic synthesis. Within such a context, the first question we must answer is; what exactly do we mean by the drug? The International Union for Pure and Applied Chemistry (IUPAC) has the following definition:

[*] **Corresponding author Fernando D. Prieto-Martínez:** Instituto de Química, Universidad Nacional Autónoma de México, Ciudad de México, México; E-mail: ferdpm4@hotmail.com

**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

"Any substance which, when absorbed into a living organism, may modify one or more of its functions" [1].

Starting from it, we may add that for our intent and purposes, a drug shall be any "small" molecule (*i.e.*, with a molecular weight < 650 Da) with an intended target, rationally designed or optimized and with the clear goal of having a therapeutic use.

Since the dawn of civilization, along with agriculture, mankind has indirectly learned of the medicinal potential of some plant species. In a historical context, it has been a long and slow transition from mystical to therapeutic. A paragon example is perhaps salicylic acid; a natural product which led to the development of one of the most well-known drugs: aspirin. It is very remarkable that this rather "simple" structure is, in fact, a prodrug from a metabolite present in the willow tree bark, which came to be associated with a Nobel Prize in physiology while also being one of the most commercially successful drugs of all time [2].

At first, drug design relied on mimicking endogenous ligands of known targets. This may seem rather straightforward, but quite the opposite is true. Returning to the aspirin example, its active ingredient, salicylic acid, had its mode of action identified and described until the latter half of the XX century. Thus, since the development and maturing of pharmacology, approaches towards drug design have become more systematic and less form of art and chance. A prime example to consider is the development of angiotensin converting enzyme (ACE) inhibitors. It was between the decades of 1960 and 1970 that John Vane's group actively researched the cause of hypertension. Studying the effect of the venom from a Brazilian viper (*Bothrops jararaca*) *in vitro*, Vane recognized the importance of ACE as a major regulator of blood pressure [3]. This led to the involvement of Squibb, specifically David Cushman and Miguel Ondetti, who characterized several peptides as antihypertensive agents and hypothesized that ACE was a zinc metallopeptidase [4]. From here, the main challenge to overcome was oral bioavailability, so the group turned to recently described carboxypeptidase A inhibitors, hypothesizing structural similarities towards ACE (Fig. **1**). This rationale led to the eventual synthesis of captopril [5]. Nonetheless, such a decision proved to be fortuitous, as the crystal structures of ACE (published almost 30 years later) showed that its catalytic domain is actually unrelated to that of carboxypeptidase A [6].

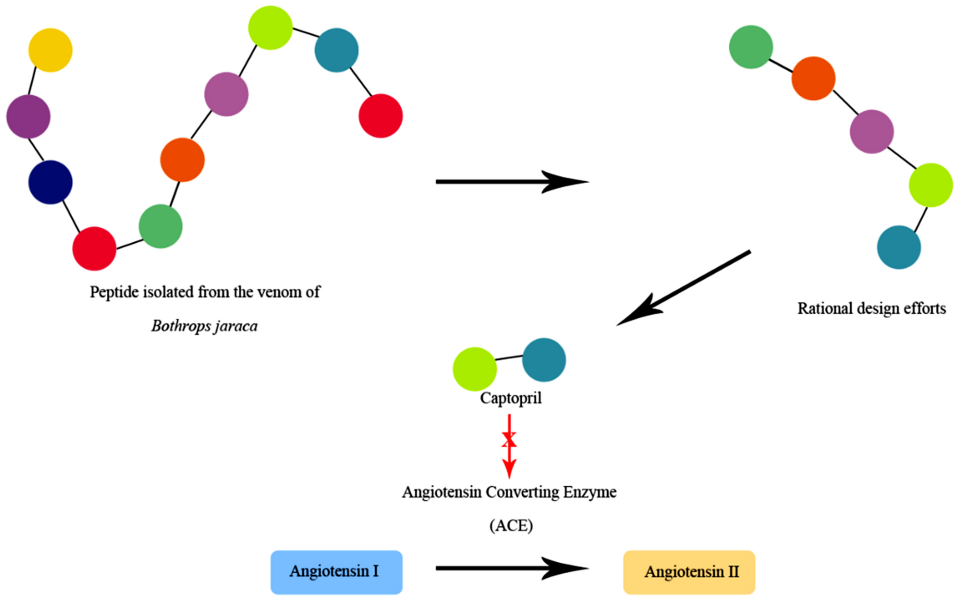



**Fig. (1).** Schematic view of the optimization process leading to the development of captopril.

From here, it becomes clear that drug development endeavours involve a great deal of complexity. For instance, during the 1980s, combinatorial chemistry was seen as a promising venue to tackle molecular diversity. The proposal of synthesizing hundreds of compounds in a systematic and rather efficient way was very appealing, leading to the development of high-throughput screening methods. A framework where thousands of compounds could be quickly evaluated using a combination of biochemical assays and robotic machinery [7]. While HTS has had notable success cases, the overall rate of developed drugs from it is rather discrete. Additionally, HTS campaigns gave rise to a phenomenon known as frequent hitters or pan-assay interference compounds (PAINS). Said designation is given to compounds showing "promiscuity" to a broad range of proteins or more generally to false positives due to interference with assay elements [8]. Indeed, this situation revealed that drug development cannot be solved by brute force, as it demands both critical and creative thinking [9].

Thus, the industry turned to state-of-the-art methodologies. Parallel to the development of HTS, there was a significant shift towards other computational modelling techniques. Early examples of this include the pharmacophore elucidation studies during the late 1960s and 1970s [10]. Nonetheless, pharmacophore models proved insufficient tools, as no information on the target is obtained. The pressing need for methods capable of predicting the binding

# SUBJECT INDEX



**Igor José dos Santos Nascimento (Ed.)**
**All rights reserved-© 2023 Bentham Science Publishers**

## D



## E



## F



## G



## H

## I



## L



## M



## N

## O



## P



## Q



## R



## S

## T



## V



## W

## Igor José dos Santos Nascimento

Professor Igor José dos Santos Nascimento has a degree in pharmacy and a Ph.D. in chemistry from the Federal University of Alagoas – Brazil, with experience in medicinal chemistry and drug development. He is a professor at the graduate level in the disciplines of medicinal chemistry and pharmacology and postgraduate level in disciplines related to drug design at the pharmacy department of State University of Paraíba - Brazil. He has approximately 30 articles and 11 book chapters published in high-impact journals, with his research focused on the application of computational tools in drug discovery and development, with experience in molecular docking, molecular dynamics simulations, density functional theory (DFT), pharmacophore models, and homology modeling. During his research studies, he performed these methods to optimize the pharmacokinetic (PK) and pharmacodynamic (PD) profile of new potential drug prototypes against Neglected Tropical Diseases (NTDs), viruses, bacteria, cancer, and inflammation.