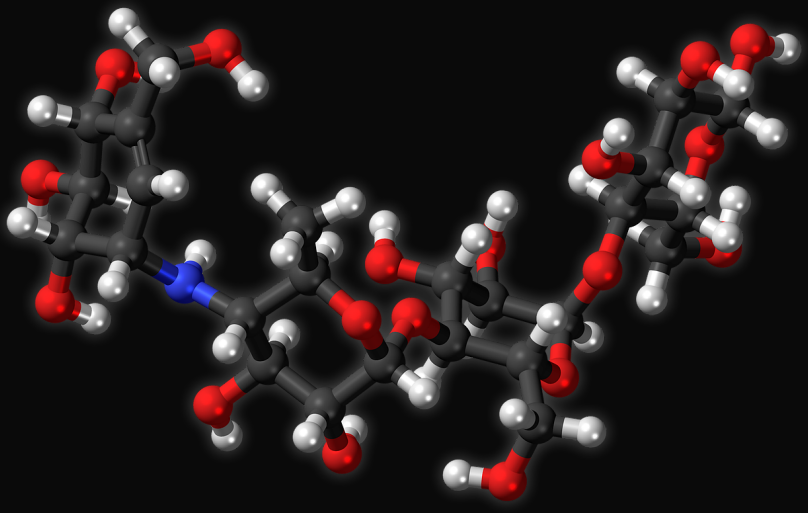

A TEXTBOOK OF

COMPUTER AIDED DRUG DESIGN

Editors:

Dr. Kishor Danao

Dr. Deweshri Nandurkar

A TEXTBOOK OF COMPUTER AIDED DRUG DESIGN

Editors:

Dr. Kishor Danao
Assistant Professor
Dadasaheb Balpande College of Pharmacy, Nagpur, Maharashtra

Dr. Deweshri Nandurkar
Associate Professor
Dadasaheb Balpande College of Pharmacy, Nagpur, Maharashtra

CREATIVE BOOKS

Website: www.creativebooks.co.in
Email: creativebook11@gmail.com

Head Office: C-91/1 Yadav Nagar Samaypur Badli, Delhi - 110042

First Published by Creative Books 2025
Copyright © Creative Books
Title: A Textbook of Computer Aided Drug Design
ISBN: 978-81-968406-9-3

No part of this book may be reproduced or transmitted in any form whatsoever, electronic, or mechanical, including photocopying recording, or by any informational storage or retrieval system without the expressed written, dated and signed permission from the editor(s).

LIMITS OF LIABILITY/DISCLAIMER OF WARRANTY: The editor(s) and publisher of this book has used their best efforts in preparing this material. The editor(s) and publisher make no representation or warranties with respect to the accuracy, applicability, or completeness of the contents. They disclaim any warranties (expressed or implied), or merchantability for any particular purpose. The editor(s) and publisher shall in no event be held liable for any loss or other damages, including but not limited to special, incidental, consequential, or other damages. The information presented in this publication is compiled from sources and freely available on the internet believed to be accurate, however, both the publisher and editor(s) assume no responsibility for errors or omissions. The information in this publication is not intended to replace or substitute professional advice. The strategies outlined in this book may not be suitable for every individual, and are not meant to provide individualized advice or recommendations.

The advice and strategies found within may not be suitable for every situation. This work is sold with the understanding that neither the editor(s) nor the publisher is held responsible for the results accrued from the advice in this book. Nothing contained herein is to be constructed as medical advice.

All disputes are subject to Delhi jurisdiction only.

About Authors

Dr. Kishor R. Danao is 17 year of teaching experience in Dadasaheb Balpande College of Pharmacy, Nagpur, Maharashtra. He obtained his B. Pharm. and M. Pharm. from RTMNU, Nagpur in Pharmaceutical Chemistry. He has completed Ph. D. from Oriental University, Indore. Also, He has to his credit certified course in Drug Design & Patenting (DDP) from BII, India. He has published more than 36 research papers in national and international reputed journals. He has credited 3 books. He has published 13 chapter in cited books and also, 3 Indian patents in his credit. His interesting working area is on the computer aided drug design. He is life time member of APTI. His main interest & mission of life are to train students for better pharmacy career.

Dr. Deweshri Nandurkar is working as Associate Professor, Department of Pharmaceutical chemistry, Dadasaheb Balpande College of Pharmacy, Besa, Nagpur. She has completed her B. Pharm. from Department of Pharmaceutical Sciences, Nagpur University, Nagpur and M. Pharm. in Pharmaceutical Chemistry from Kamla Nehru College of Pharmacy, Butibori, Nagpur. Afterword's, she has completed her Ph. D. from Department of Pharmaceutical Sciences, Nagpur University, Nagpur. Her Research area of interest involves computer assisted rational designing and synthesis of various pharmacophores. She had published 20 research papers in National and International Journal. She has credited 5 chapters and 2 patents in her credit. She attended more than 25 national and international conferences. She is a life member of APTI, ACT and MSPC council. She received two gold medals for being University topper in M. Pharm. Pharmaceutical Chemistry branch 2012 examination. Her primary interest and aim in life are to prepare students for a successful pharmacy career.

CREATIVE BOOKS

Preface

We are pleased to introduce the first edition of ***"A Textbook of Computer Aided Drug Design"*** for Bachelor and Masters Pharmacy students, which is in accordance strictly with the syllabus of Pharmacy Council of India (PCI). The authors tried their best to present the book in a more illustrated manner, including up-to-date information on bioinformatics tools for the discovery and development of new drug molecules. Major objectives of this book is to provide a thorough background in those who want to explore themselves in the field of Computer Aided Drug Design. For better understanding, the book has been divided into 10 units. Unit 1 and 2 gives an insight into the introduction to drug discovery and development from traditional medicine and lead discovery. Unit 3 and 4 focusses on the analogue based drug design quantitative structure-activity relationship studies. Unit 5 and 6 integrates the knowledge of 3D QSAR. Unit 7 and 8 explains the knowledge of molecular docking and virtual screening techniques. Unit 9 and 10, the concept of molecular modelling and informatics.

CREATIVE BOOKS

Table of Contents

Chapter 1: *Drug Discovery and Development* 1

Chapter 2: *Lead Discovery* 23

Chapter 3: *Analog Based Drug Design* 43

Chapter 4: *Quantitative Structure Activity Relationship* 67

Chapter 5: *Physicochemical Parameters of QSAR* 91

Chapter 6: *3D QSAR* 107

Chapter 7: *An Overview of Virtual Screening Techniques* 131

Chapter 8: *Molecular Docking* 153

Chapter 9: *Informatics and Methods in Drug Design* 191

Chapter 10: *Molecular Modelling* 215

CRE&TIVE BOOKS

Chapter 9

Informatics and Methods in Drug Design

Ruchi Shivhare[1]*, Sneha Agrawal[2], Kirti Korde[3], Ujwala Mahajan[4]

*[1]*Pharmaceutical Chemistry Deoartment, Dadasaheb Balpande College of Pharmacy, Besa, Nagpur; [2]Bharati Vidyapeeth's College of Pharmacy, CBD Belapur, Navi Mumbai; [3] Dadasaheb Balpande College of Pharmacy, Besa; [4]Principal, Dadasaheb Balpande College of Pharmacy, Besa*

**shivharer4@gmail.com*

INTRODUCTION

CREATIVE BOOKS

A key area of pharmaceutical science is drug design, sometimes referred to as rational drug design, which focuses on creating novel drugs by using biological targets as knowledge. This interdisciplinary field combines pharmacological, biological, and chemical principles to create compounds with therapeutic benefits. A key area of pharmaceutical science is drug design, sometimes referred to as rational drug design, which focuses on creating novel drugs by using biological targets as knowledge. This interdisciplinary field combines pharmacological, biological, and chemical principles to create compounds with therapeutic benefits [1].Structure-based drug design (SBDD), which is predicated on the target biomolecule's three-dimensional structure, is one of the fundamental methods in drug design. By using methods like nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography, scientists can create molecules that precisely fit into the target site by obtaining detailed images of the molecular structure [2]. In cases where the target's structure is unknown, ligand-based drug design (LBDD) is a noteworthy alternative. Using this approach, novel drug designs are created using data

from established ligands, or compounds that bind to their target. In order to predict the biological activity of novel compounds, computational techniques—such as quantitative structure-activity relationship, or QSAR—are essential [3]. By significantly reducing the time and expense involved in drug development, computational drug design has completely changed the industry. Finding promising candidates is made easier by computer-aided drug design (CADD), which simulates and analyses the interactions between medications and their targets [4].

The field of "chemoinformatics" handles and analyses chemical data by fusing information science and chemistry. Because it makes the effective management and analysis of massive datasets of chemical compounds possible, it is essential to the modern process of drug discovery and development. The main goal of chemoinformatics is to transform chemical data into insightful knowledge that can direct choices regarding the creation and advancement of novel medications [5].Building and maintaining chemical databases is one of the main uses of chemoinformatics, wide need of chemoinformatics depicts in Figure no.1. These databases make it easier to find and analyse pertinent data by storing details about the composition, characteristics, and actions of chemical compounds [6]. Chemoinformatics tools also make it possible to compare and visualize chemical structures, which is crucial for finding patterns and connections in chemical datasets.

CREATIVE BOOKS

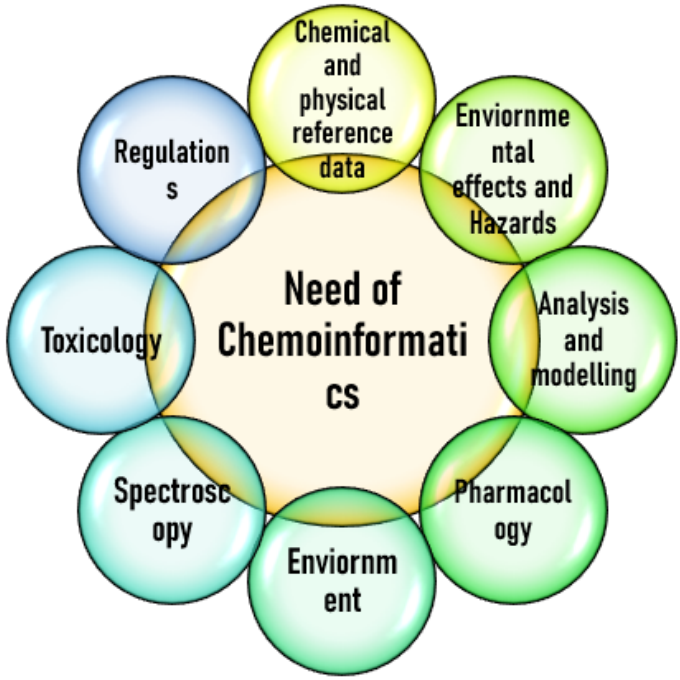

Figure 1: Need of Chemoinformatics

An important multidisciplinary field called bioinformatics brings together information technology, computer science, and biology to analyses and interpret biological data. When it comes to drug design, bioinformatics is crucial for comprehending the molecular causes of illnesses, locating possible targets for drugs, and creating novel therapeutics. Large-scale genomic, proteomic, and metabolomic data analysis has been made possible by the integration of bioinformatics into drug design, which has completely changed the procedure. Through these analyses, new drug targets can be found by gaining insight into the genetic and molecular causes of diseases. For example, focusing on particular biomolecules that are essential to the pathophysiology of a disease is made possible by the identification of genes linked to the disease and the proteins that result from them [7]. Drug targets are identified and validated, which is one of the major contributions of bioinformatics to drug design. The comparison and analysis of genetic sequences to find genes linked to diseases is made easier by bioinformatics databases

and tools like GenBank and BLAST (Basic Local Alignment Search Tool)[8] . Additionally, by analysing the functional connections between various biomolecules, bioinformatics techniques like gene expression profiling and protein-protein interaction networks aid in identifying crucial nodes that can be the target of pharmaceuticals [9]. Chemoinformatics and bioinformatics approaches are diverse and serve a wide range of tasks, including drug discovery, materials science, and chemical research. Here are some significant chemoinformatic techniques:

CHEMOINFORMATICS TECHNIQUES

1. Molecular Descriptors in Chemoinformatics: Molecular descriptors are quantitative illustrations of molecular attributes that describe molecules' chemical structures and physical properties. These characteristics are essential for creating quantitative structure-activity relationship (QSAR) models and doing virtual screening in drug discovery [10].Molecular descriptors are of various types like 1D, 2D, 3D and 4D. Simple counts, such as molecular weight or atom count, are examples of 1D descriptors.2D Descriptors: Incorporate connectivity and topological indices, like the Wiener index or the number of rings. Geometric characteristics like surface area and molecule volume should be included in 3D descriptors. For flexible molecules, 4D descriptors take into consideration the various conformations that a molecule can take on. Electronic Descriptors: Terms like polarizability and dipole moment that describe the electronic structure.

2. Quantitative Structure-Activity Relationship (QSAR): One important chemoinformatics technique for predicting the biological activity of chemical substances based on their chemical structure is the Quantitative Structure-Activity Relationship (QSAR). In order to guide the design of new medication with most wanted qualities, researchers can enhanced value and predict how changes in molecular structure can affect biological activity, which is made possible by the use of QSAR models [11,12].QSAR include data preparation, feature selection, model development (using statistical or machine learning methods), and validation. The general procedures for modeling QSAR shown Figure no. 2.

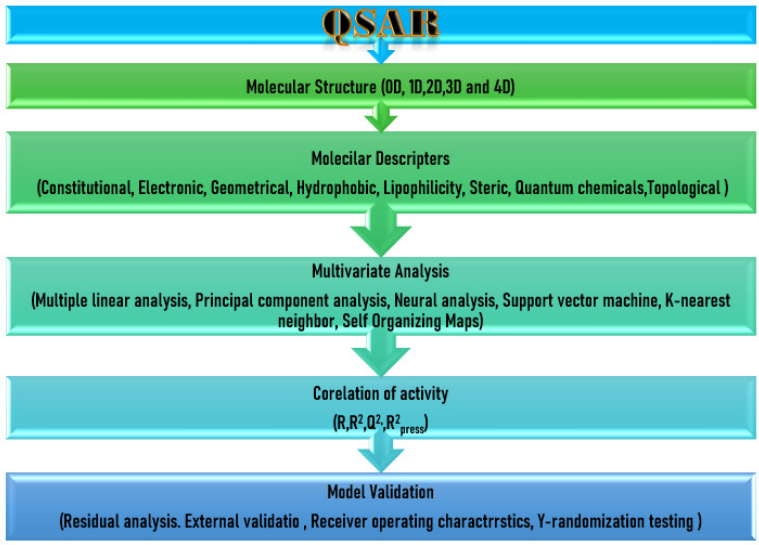

Figure 2: The general procedures for modeling QSAR

Essential Elements and Procedures of QSAR:[12]

- **Data Collection:** The first step in the procedure is compiling a list of chemical compounds that have been shown to have biological activity.
- **Calculation of Molecular Descriptors:** For every chemical, a set of numerical values representing different molecular properties (such hydrophobicity and electronic effects) is determined.
- **Model Development:** Statistical or machine learning techniques are used to simulate the link between the biological activity and the descriptors. Neural networks, support vector machines, and multiple linear regressions are common methods.
- **Model Validation:** Using approaches such as cross-validation or external validation with an independent

dataset, the QSAR model's predictive capacity is evaluated.

- **Prediction and Optimization:** Next, the biological activities of novel or unexplored substances are predicted using the verified model. Additionally, it can direct lead compound optimization by recommending structural changes that will boost activity.

3 **Molecular Docking:** Molecular docking is a computational technique in drug design that determines the most effective position of a molecule (ligand) when it interacts to a target protein, receptor, or enzyme. The fundamental goal of molecular docking is to know how the ligand interacts with the target, predict binding affinity and discover probable binding sites. This information is critical for rational drug design, as it allows researchers to modify the structure of therapeutic candidates for improved efficacy and fewer adverse reactions. Figure no. 3 Diagram illustrating the Molecular Docking Process. Different methods of molecular docking are as follows [13,14].

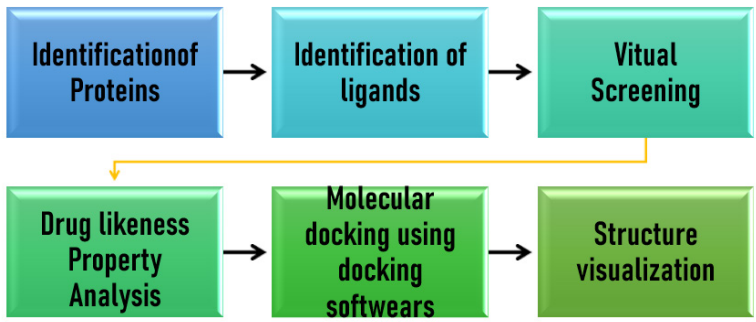

Figure no.3: Flow chart showing the steps of Molecular docking

- **Rigid Docking:** Assumes that both the ligand and the receptor are rigid, and focuses on finding the optimum fit based on shape complementarily.

- **Flexible Docking**: Allows for conformational flexibility

of the ligand, receptor, or both in order to better simulate biological circumstances. This technique can account for variations in the geometry of the ligand or protein during binding.

- **Scoring Functions**: Predict the binding affinity and stability of the ligand-protein complex. Empirical, knowledge-based, and force field-based techniques are examples of common scoring functions. These functions assess interactions such as hydrogen bonding, hydrophobic effects, van-der Waals forces, and electrostatics.

- **Docking algorithms:** Monte Carlo simulations and grid-based approaches are all examples of common docking algorithms. These algorithms search the conformational space of the ligand and protein to determine the most effective binding.

Various applications of molecular docking are as follows:

- **Lead Identification and Optimization:** Molecular docking is often utilized in the early phases of drug discovery to identify possible lead compounds that have a strong binding affinity for a target protein. It also contributes to the optimization of these leads by estimating the effects of chemical changes on binding affinity.

- **Structure-Based Drug create (SBDD):** To create new medications, this technique uses the target protein's 3D structure. Docking aids in understanding the interaction between the target and small molecules, allowing for the development of more effective and selective medications.

- **Virtual Screening:** Docking can screen vast libraries of chemicals to select those with the greatest biological activity potential, saving time and money on experimental testing.

- **Predicting Drug Metabolism and Toxicity:** Researchers can predict metabolic pathways and probable toxicological effects by docking drug candidates with metabolism-related enzymes (such as cytochrome P450 enzymes).

4 Data Mining and Machine Learning: Data mining and machine learning (ML) approaches are getting increasingly important in drug research, providing tools for managing and evaluating huge datasets. These techniques aid in discovering patterns, forecasting results, and enhancing candidates for drugs [15, 16].

- Supervised Learning is the process of training models on labeled data to predict drugs' biological activity or toxicity. Common techniques employed include support vector machines (SVM), decision trees, and neural networks.

- Unsupervised learning: Clustering and principal component analysis (PCA) are methods for discovering natural groupings in data, which can be beneficial for classifying medications and identifying novel therapeutic candidates.

- Deep learning: Namely convolutional neural networks (CNNs) and recurrent neural networks (RNNs), has been utilized for tasks including protein structure prediction and molecule creation.

- Predictive modeling: ML models predict several pharmacokinetic and pharmacodynamic aspects of medications, including absorption, distribution, metabolism, elimination, and toxicity. These models help to optimize medication candidates for higher efficacy and safety.

5. Pharmacophore Modeling: A pharmacophore is the spatial arrangement of atoms or groups of a molecule required for optimal interactions with a given biological target, resulting in the desired biological response. It comprises key properties such as hydrogen bond acceptors and donors, aromatic rings, hydrophobic areas, and positively or negatively charged groups [17,18] .

CRESTIVE BOOKS

Modeling Process:

I. Identifying Pharmacophore Features.

Data collection involves using a set of active chemicals that are known to interact with a specific biological target.

- Feature Extraction: We identify key molecular traits that are required for biological activity.

II. Pharmacophore Hypothesis Generation:

- Common pharmacophore hypotheses. Alignment: Aligns active chemicals to discover shared pharmacophoric components. This step frequently employs software tools such as Catalyst or Phase.
- Hypothesis testing involves comparing the developed pharmacophore model to inactive drugs to ensure its specificity.

CREATIVE BOOKS

III. Validation and refinement:

- Validation: The pharmacophore model is validated by predicting the activity of novel drugs.
- Refinement: The pharmacophore model is iteratively refined using more experimental data and computational predictions.

Pharmacophore modeling has various applications like virtual screening, lead optimization, de novo drug design, and understanding receptor-ligand Interactions. These applications demonstrate the versatility and relevance of pharmacophore modeling in the early stages of drug discovery and development, allowing for fast identification and optimization of therapeutic candidates [19].

6. ADMET Prediction:

The abbreviation ADMET stands for Absorption, Distribution, Metabolism, Excretion, and Toxicity. These features are crucial for establishing a drug's efficacy and safety. ADMET prediction is utilizing computational models to anticipate how a drug will react in the human body, which is critical for optimizing drug candidates and decreasing late-stage failures in drug development [20,21].

Important Characteristics of ADMET Prediction:

- Absorption: Indicates how well the drug is absorbed into the blood stream from the point of administration. This includes calculating characteristics like as solubility, permeability, and intestinal absorption.
- Distribution: Determines how a medicine is distributed throughout the body's tissues and organs. This includes estimating the volume of distribution and protein binding affinity.
- Metabolism: Determines how a medicine is metabolized, primarily in the liver, and identifies possible metabolites. This helps to understand the drug's biotransformation and potential drug-drug interactions.
- Excretion: Determines how a drug and its metabolites are removed from the body, usually through renal or hepatic pathways.
- Toxicity: Predicts probable hazardous consequences such as cytotoxicity, genotoxicity, hepatotoxicity, and cardiotoxicity, among others.

Applications of ADNET predictions are like in drug screening, lead optimization, **regulatory compliance**, and **personalized medicine etc.**

CREATIVE BOOKS

7. Molecular Dynamic and Simulation:

Molecular Dynamics (MD) is a computational modeling method that investigates the physical activities of atoms and molecules as time passes. It uses classical mechanics to anticipate the time-dependent actions of a molecular system by computing the trajectories of particles using Newton's equations of motion [22,23].

Important Characteristics of Molecular Dynamics:

- **Initial Setup:** The simulation begins with an initial configuration of atoms, including their locations and velocities. Force fields describe the system's potential energy by accounting for a variety of interactions, including bond stretching, angle bending, van der Waals forces, and electrostatic interactions.
- Simulation Process: MD simulations compute forces and update particle locations and velocities iteratively. Time steps are typically femtoseconds to nanoseconds long, enabling for comprehensive investigation of molecular movement.
- Analysis of Trajectories: The trajectories that are obtained offer valuable information about the thermodynamics, kinetics, and structural dynamics of the system.

The adaptability and influence of molecular dynamics simulations across a range of domains offer fundamental comprehension and useful implementations. Numerous uses, such as drug development, materials science, enzymatic mechanisms, protein folding and dynamics, and insights into the dynamics of biological membranes, such as the diffusion of lipids and proteins [24,25,26].

8. Structural Bioinformatics:

The investigation and forecasting of the three-dimensional structures of biomolecules, such as proteins, nucleic acids, and their complexes, is the focus of the subfield of structural

bioinformatics. Clarifying these structures is essential to understanding how biomolecules interact, what their functions are, and the molecular mechanisms that underlie biological activities. The techniques of structural bioinformatics are as follows [27,28,29]:

I. Homology Modeling: This technique uses the sequence similarity of a target protein to one or more identified structures (templates) to predict the three-dimensional structure of the protein. It entails creating a model of the target protein's backbone, fine-tuning side chains, and matching the target sequence with the template structures. When the target protein's structure is unknown but comparable proteins with known structures exist, homology modeling is particularly helpful.

II. Structural Alignment: Comparing the structures of two or more proteins in order to find structurally comparable regions and evolutionary relationships is known as structural alignment. Through the identification of functional domains and conserved motifs, this technique can shed light on the evolution and function of proteins. Structural alignment is often achieved with tools like TM-align and DALI.

These techniques are essential to structural bioinformatics because they enable scientists to forecast and examine the structure-function correlations of biomolecules, which are crucial for comprehending biological processes and the development of new drugs.

A key component of modern drug development is the use of bioinformatics techniques, which help scientists quickly and cheaply find and refine novel therapeutic possibilities. Technologies such as Genomics and Transcriptomics Analysis are employed to examine transcriptomic and genomic data in order to pinpoint genes and pathways linked to certain illnesses. Large-scale analyses of proteins and metabolites in living systems are the focus of proteomics and metabolomics research. Scholars can find illness biomarkers and possible targets for pharmaceuticals by examining the proteome and metabolome [30,31].

CHEMICAL, BIOCHEMICAL, AND PHARMACEUTICAL DATABASES

A. Chemical Databases:

Chemical databases are informational databases pertaining to chemical substances. They are essential at every step of the drug discovery process, from lead optimization to initial compound identification. These databases give researchers the ability to predict biological activities, optimize the properties of possible drug candidates, and identify them. Now a day's various chemical database was used like PubChem, ChemSpider, ChEMBL, ZINC and BindingDB (Table 1).

- **PubChem:** Chemical compounds and their actions against biological assays are listed in the free PubChem database. One essential component of the National Library of Medicine (NLM) is the National Centre for Biotechnology Information (NCBI), which ensures it. PubChem is widely used in drug design and discovery processes for a variety of purposes, including identifying bioactive compounds, conducting Structure-Activity Relationship (SAR) studies, repurposing drugs, virtual screening, and data integration [32].

- **ChemSpider:** ChemSpider is a crowdsourcing community for chemists who provide their data, expertise, and knowledge to the improvement and curation of the database. It is more than just a search engine built on terabytes of chemistry data. Therefore, ChemSpider is similar to Wikipedia in that it promotes community involvement and contributions. Access to more than 67 million structures, attributes, and related data can be found in the free chemical structure database ChemSpider. The Royal Society of Chemistry (RSC) is responsible for its upkeep. It combines information from a variety of sources, such as government, business, and academic databases [33].

CREATIVE BOOKS

- **ChEMBL:** The majority of the content in ChEMBL is comprised of bioactivity data that is routinely taken from the literature on medicinal chemistry. In addition to a plethora of other uses, these data allow researchers to pinpoint tool compounds for possible therapeutic targets, explore the SAR data that is accessible for a target, look into phenotypic information linked to related compounds, and pinpoint potential off-target effects of particular chemotypes. With some significant new areas included in recent releases, ChEMBL's scope and data types have gradually expanded to provide a more comprehensive view, partly in response to user feedback. substances undergoing clinical trials, information from patents, direct depositions regarding undiagnosed illnesses, and data on agrochemicals. ChEMBL is constantly evolving to meet the diverse needs of users and applications. The resource's core consists of selected molecule-target interaction data from published literature. However, as previously reviewed, it has expanded in terms of data content, annotation, and infrastructure. The European Bioinformatics Institute (EBI) maintains ChEMBL, a manually curated database of bioactive molecules with druglike properties [34].

- **ZINC and Zinc 15:** Free databases of commercially available compounds intended for virtual screening are ZINC and ZINC15. The Irwin and Shoichet Laboratories in the University of California, San Francisco (UCSF) Department of Pharmaceutical Chemistry are in-charge of maintaining these databases. ZINC and ZINC15 are essential tools for drug discovery and design for a number of reasons. Virtual Screening, Lead Compound Discovery, Chemical Diversity, Structure-Based Drug Design (SBDD), and Data Integration [35].

BindingDB: A public database called BindingDB focuses on the interactions between proteins and small molecules by

offering measured binding affinities. Its goal is to assist in systems pharmacology, computational chemistry, and medicinal chemistry. BindingDB is a valuable resource in drug design for several reasons: Analyzing Protein-Ligand Interactions, Optimizing Binding Affinities, Structure-Activity Relationship (SAR) Analysis, Validating Computational Predictions, and Drug Repurposing [36].

Table 1: Chemical databases for drug design

Sr. No.	Chemical Databases	Information in databases
1	PubChem	Chemical structures and identifiers Chemical and physical properties Biological activities Patents Health and safety data Toxicity data Pharmacological actions · Bioassay results
2	ChemSpider	Chemical structures and identifiers Physical and chemical properties Spectral data Synonyms and trade names Vendor and purchasability information Links to relevant literature and patents Toxicity and safety information

CREATIVE BOOKS

3	ChEMBL	Bioactive molecules and their drug-like properties Biological activities (e.g., IC50, EC50, Ki, Kd) Target information (e.g., protein targets, molecular pathways) Assay data (experimental methods and conditions) Compound structures and properties Pharmacokineticand pharmacodynamic data
4	ZINC and Zinc 15	Chemical structures and identifiers Physical and chemical properties Availability and vendor information 3D structures and conformers Annotated biological activities
5	BindingDB	Binding affinities (e.g., Kd, Ki, IC50) Protein and small molecule structures Experimental conditions for binding assays Chemical structures and identifiers Biological targets, including proteins, DNA, and RNA Links to relevant scientific literature

B. Biochemical Databases:

Information about biological macromolecules, metabolic pathways, and their interactions can be found in biochemical databases. They are critical to comprehending both the molecular mechanisms

underlying drug action and the biochemical underpinnings of disease. PDB (Protein Data Bank), KEGG (Kyoto Encyclopaedia of Genes and Genomes), and UniProt are some biochemical databases utilized nowadays.

- **PDB (Protein Data Bank):** The Protein Data Bank (PDB) is a useful tool for drug development. It offers three-dimensional structural information of big biological molecules, such as nucleic acids and proteins, which are crucial targets for the development of new drugs. To design medications that can bind to particular targets with high affinity and selectivity, researchers need to understand the shapes and interactions of these molecules, which is made possible by the data they collect. PDB provides information about 3D structures of proteins, nucleic acids, and complex assemblies and experimental data from X-ray crystallography, NMR spectroscopy, and cryo-electron microscopy. This data bank is utilize in drug design for structure-based drug design through molecular docking and virtual screening, understanding molecular mechanisms of drug action and identifying binding sites and optimizing drug-target interactions [37].

- **KEGG (Kyoto Encyclopedia of Genes and Genomes):** KEGG is a comprehensive database resource for understanding the high-level functions and utilities of biological systems, such as cells, organisms, and ecosystems, using molecular-level data. Kanehisa Laboratories developed and maintains KEGG, which integrates genomic, chemical, and systemic functional information. KEGC also provides information about Pathway maps integrating genomic, chemical, and systemic functional information and databases on genes, proteins, compounds, reactions, and diseases. KEGG is available online at https://www.kegg.jp/ and includes tools for browsing and analyzing its

vast data collection. The KEGG Mapper tool is especially beneficial for pathway mapping and data visualization [38].

- **UniProt:** Another essential tool for bioinformatics and drug design is UniProt. It is an extensive database that contains functional data and protein sequences, offering important insights into the interactions, structures, and biological roles of proteins. Drug design for protein identification and characterization, validation of drug targets, structural insights, pharmacogenomics, and data integration are all aided by UniProt. In conclusion, UniProt is vital to drug discovery and design because it provides information about proteins that is necessary for target identification, validation, structural analysis, and personalized medicine [39].

Databases such as the Target Central Resource Database (TCRD) offer comprehensive information on target proteins, encompassing their sequences, structures, genomic information, and related medical conditions. It aids in locating and confirming possible therapeutic targets. Although it offers useful data on ongoing and completed clinical trials, including drug interventions, outcomes, and side effects, ClinicalTrials.gov. is not solely a drug design database. This keeps scientists abreast of the most recent advancements in drug testing and efficacy.

OVERVIEW OF APPLICATIONS OF DATABASES IN DRUG DESIGN:

- Drug discovery is the process of finding novel therapeutic candidates and comprehending how they work.
- Research in Chemical Biology: Examining how tiny molecules interact with biological systems.
- Predicting ligand-protein interactions and improving therapeutic candidates are two aspects of computational drug design.

- The process of identifying possible pharmacological targets and confirming their biological significance is known as target validation.
- Clinical Trial Support: Tracking current trials and investigating the safety and efficacy of medications.

These databases are essential resources for scientists working in the biotechnological and pharmaceutical fields because they facilitate the effective creation and refinement of novel treatments.

CONCLUSION

The development and optimization of new therapies may now be done more quickly and cheaply mainly to the combining of chemoinformatics and bioinformatics tools, which has completely changed the field of drug design. Both disciplines use large-scale biological data and computational techniques to speed up the drug development process, from lead optimization to target identification.

CREATIVE BOOKS

Quantitative structure-activity relationships (QSAR), molecular docking, pharmacophore modeling, and other techniques are included in the field of cheminformatics. These methods make it possible to identify possible therapeutic candidates, anticipate the biological actions of chemical substances, and characterize and analyze them. To comprehend medication efficacy and specificity, for example, molecular docking offers insights on the binding affinity and orientation of ligands to their receptors. The utilization of QSAR models facilitates the rational design of medications with desired features by enabling the prediction of biological activity of compounds based on their chemical structures. However, bioinformatics methods like systems biology, proteomics and metabolomics, genomes and transcriptomics analysis, and molecular dynamics are crucial for comprehending the underlying mechanisms of illness and determining prospective therapeutic targets. Analyses of transcriptomics and genomes are essential for identifying genes and pathways linked to disease, whereas proteomics and metabolomics offer detailed perspectives on the profiles of proteins and metabolites, revealing biomarkers

and potential treatment targets. By forecasting the stability and structural changes of drug-target complexes, molecular dynamics simulations provide comprehensive insights into the dynamic behavior of biomolecules and help guide drug design. These chemoinformatics and bioinformatics techniques work together to provide a thorough, interdisciplinary approach to drug discovery. They aid in the discovery of new therapeutic targets, the screening of enormous compound libraries, the improvement of lead compounds, and the comprehension of pharmacological mechanisms. The adoption of a comprehensive approach speeds up the process of developing new drugs and increases the chances of success by concentrating on compounds that have advantageous pharmacokinetic and pharmacodynamic characteristics.

Databases related to chemicals, biochemistry, and pharmaceuticals are essential drug design and informatics tools. They furnish the requisite information for diverse computational techniques that accelerate the identification and advancement of novel pharmaceuticals. By integrating sophisticated informatics techniques with these databases, researchers can effectively identify and optimize possible drug candidates for clinical application.

CREATIVE BOOKS

Future developments in the domains of bioinformatics and chemoinformatics should significantly improve our capacity to create safer, more potent medications, ultimately leading to better patient outcomes and the advancement of customized medicine.

REFERENCE

1. *Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. Infant and child Development, 20(3), 251-271.*

2. *Anderson, A. C. (2003). The process of structure-based drug design. Chemistry & biology, 10(9), 787-797.*

3. *Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., ... & Tropsha, A. (2014). QSAR modeling: where have you been? Where are you going to? Journal of Medicinal Chemistry, 57(12), 4977-5010.* https://doi.org/10.1021/jm4004285

4. *Jorgensen, W. L. (2004). The many roles of computation in drug discovery. Science, 303(5665), 1813-1818. https://doi.org/10.1126/science.10963613] "What is bioinformatics? An introduction and overview."*

5. *Brown, F. K. (2005). Chemoinformatics: What is it and how does it impact drug discovery. Annual Reports in Medicinal Chemistry, 40, 281-291.* https://doi.org/10.1016/S0065-7743(05)40074-1

6. *Gasteiger, J., & Engel, T. (2003). Chemoinformatics: A Textbook. Wiley-VCH.*

7. *Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research, 28(1), 27-30.* https://doi.org/10.1093/nar/28.1.27

8. *Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215(3), 403-410.* https://doi.org/10.1016/S0022-2836(05)80360-2

9. *Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., ... & Wanker, E. E. (2005). A human protein-protein interaction network: a resource for annotating the proteome. Cell, 122(6), 957-968.* https://doi.org/10.1016/j.cell.2005.08.029

10. *Todeschini, R., & Consonni, V. (2008). Handbook of molecular descriptors. John Wiley & Sons.*

11. *Niazi, S. K., & Mariam, Z. (2023). Recent advances in machine-learning-based chemoinformatics: a comprehensive review. International Journal of Molecular Sciences, 24(14), 11488.*

12. Patel, H., & Shah, P. (2023). *Advances in QSAR Modeling for Drug Design: Modern Approaches and Applications. Current Topics in Medicinal Chemistry, 23(2), 110-123. doi:10.2174/1568026623666230102094703.*

13. *Morris, G. M., & Lim-Wilby, M. (2008). Molecular docking. Methods in Molecular Biology, 443, 365-382.*

CREATIVE BOOKS

https://doi.org/10.1007/978-1-59745-177-2_19

14. *Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein-ligand docking and virtual drug screening with the AutoDock suite.* Nature Protocols, *11(5), 905-919.* https://doi.org/10.1038/nprot.2016.051

15. *Burbidge, R., Trotter, M., Buxton, B., & Holden, S. (2001). Drug design by machine learning: support vector machines for pharmaceutical data analysis. Computers & chemistry, 26(1), 5-14.*

16. *Schuster, D., & Wolber, G. (2010). Pharmacophore modeling and virtual screening. Methods in Molecular Biology, 672, 361-384.* https://doi.org/10.1007/978-1-60761-839-3_14

17. *Langer, T., & Hoffmann, R. D. (2006). Pharmacophores and Pharmacophore Searches. Wiley-VCH.* https://doi.org/10.1002/3527607633

18. *Yang, S. Y. (2010). Pharmacophore modeling and applications in drug discovery: Challenges and recent advances. Drug Discovery Today, 15(11-12), 444-450.* https://doi.org/10.1016/j.drudis.2010.03.013

19. *Guner, O. F. (Ed.). (2000). Pharmacophore Perception, Development, and Use in Drug Design. International University Line.*

20. *Ekins, S., Xu, J. J., & Wang, B. (2010).* Computational Toxicology: Risk Assessment for Pharmaceutical and Environmental Chemicals. *John Wiley & Sons.*

21. *Van de Waterbeemd, H., & Gifford, E. (2003). ADMET in silico modelling: Towards prediction paradise?* Nature Reviews Drug Discovery, *2(3), 192-204.* https://doi.org/10.1038/nrd1032

22. *Shaw, D. E., et al. (2010). Atomic-level characterization of the structural dynamics of proteins.* Science, *330(6002), 341-346.* https://doi.org/10.1126/science.1187409

23. Dror, R. O., et al. (2011). Pathway and mechanism of drug binding to G-protein-coupled receptors. Proceedings of the National Academy of Sciences, 108(32), 13118-13123. https://doi.org/10.1073/pnas.1112053108

24. Frenkel, D., & Smit, B. (2002). Understanding Molecular Simulation: From Algorithms to Applications (2nd ed.). Academic Press.

25. Marrink, S. J., & Tieleman, D. P. (2013). Perspective on the Martini model. Chemical Society Reviews, 42(16), 6801-6822. https://doi.org/10.1039/C3CS60093A

26. 26. Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. Nature Structural Biology, 9(9), 646-652.

27. Fiser, A., & Sali, A. (2003). Modeller: Generation and refinement of homology-based protein structure models. Methods in Enzymology, 374, 461-491. https://doi.org/10.1016/S0076-6879(03)74020-8

28. Vakser, I. A. (2014). Protein-protein docking: From interaction to interactome. Biophysical Journal, 107(8), 1785-1793. https://doi.org/10.1016/j.bpj.2014.08.033

29. 29. Krissinel, E., & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallographica Section D: Biological Crystallography, 60(12), 2256-2268.

30. Zhang, B., Wang, J., Wang, X., Zhu, J., Liu, Q., Shi, Z., ... & Li, Y. (2011). A comprehensive evaluation of RNA-seq data reveals novel heterogeneity in transcription levels in breast cancer cell lines. Nucleic Acids Research, 39(2), 526-538.

31. Hanash, S. M., Pitteri, S. J., & Faca, V. M. (2008). Mining the plasma proteome for cancer biomarkers. Nature, 452(7187), 571-579.

32. Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., ... & Bryant, S. H. (2016). PubChem

CREATIVE BOOKS

substance and compound databases. Nucleic acids research, 44(D1), D1202-D1213.

33. *Pence, H. E., & Williams, A. (2010). ChemSpider: an online chemical information resource.*

34. *Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., ... & Leach, A. R. (2017). The ChEMBL database in 2017. Nucleic acids research, 45(D1), D945-D954.*

35. *Sterling, T., & Irwin, J. J. (2015). ZINC 15–ligand discovery for everyone. Journal of chemical information and modeling, 55(11), 2324-2337.*

36. *Gilson, M. K., Liu, T., Baitaluk, M., Nicola, G., Hwang, L., & Chong, J. (2016). BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. Nucleic acids research, 44(D1), D1045-D1053.*

37. *Burley, S. K., Berman, H. M., Kleywegt, G. J., Markley, J. L., Nakamura, H., & Velankar, S. (2017). Protein Data Bank (PDB): the single global macromolecular structure archive. Protein crystallography: methods and protocols, 627-641.*

38. *Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic acids research, 45(D1), D353-D361.*

39. *UniProt Consortium. (2019). UniProt: a worldwide hub of protein knowledge. Nucleic acids research, 47(D1), D506-D515.*

A TEXTBOOK OF COMPUTER AIDED DRUG DESIGN

Dr. Kishor Danao

Dr. Deweshri Nandurkar

ABOUT THE BOOK

The ***"A Textbook of Computer Aided Drug Design"*** is intended for Bachelor and Masters Pharmacy students and is based on the syllabus of Pharmacy Council of India (PCI). The authors did their utmost to deliver the book in a more visually appealing manner.

This book provides up-to-date information on bioinformatics tools for the discovery and development of new drug molecules. It discusses a range of computational applications, including three-dimensional modelling of protein structures, protein-ligand docking, and molecular dynamics simulation of protein- ligand complexes for identifying desirable drug candidates. Moreover, it presents structure and ligand-based drug design tools to optimize known drugs and guide the design of new molecules. The book also describes methods for identifying small-molecule binding pockets in proteins, and summarizes the databases used to explore the essential properties of drugs, drug-like small molecules and their targets. In addition, the book highlights various tools to predict the absorption, distribution, metabolism, excretion (ADME) and toxicity (T) of potential drug candidate.

CREATIVE BOOKS

978-81-968406-9-3

Rs. 595/-