JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Computer aided drug designing as tool leading in modern day pharmaceutical industry: A review

Santosh Kumar Mohato[1*], Vikrant Jain[1],Pranay Guru[1], Dharmendra Kumar Singh[2]
[1]Department of Chemistry, Madhyanchal Professional University, Bhopal 462044, India
[2]Department of Chemistry BBMK College Dhanbad, India

Abstract

Computer-aided drug design (CADD) is a multidisciplinary field at the intersection of chemistry, biology, and computational science, which plays a pivotal role in modern drug discovery. This abstract provides an overview of CADD, its methodologies, and its significant impact on pharmaceutical research.CADD leverages computational techniques to expedite the drug discovery process. It involves the rational design of molecules with high potential for therapeutic efficacy and minimal side effects. The primary goal of CADD is to identify lead compounds, optimize their binding to specific biological targets, and predict their pharmacokinetic properties. This approach significantly reduces the time and cost associated with traditional trial-and-error drug development. Key components of CADD include molecular modeling, virtual screening, and quantitative structure-activity relationship (QSAR) analysis. Molecular modeling techniques, such as molecular docking and molecular dynamics simulations, enable researchers to visualize the interaction between potential drug candidates and target proteins at the atomic level. Virtual screening involves the rapid assessment of large chemical libraries to identify promising compounds for further investigation. QSAR models provide insights into the relationship between a molecule's structure and its biological activity. CADD has transformed drug discovery by accelerating the identification of novel drug candidates and repurposing existing drugs for new indications. It has been instrumental in the development of therapies for various diseases, including cancer, infectious diseases, and neurological disorders. Moreover, CADD contributes to the optimization of drug candidates to enhance their safety and efficacy profiles. In conclusion, computer-aided drug design is a vital component of modern pharmaceutical research. Its integration of computational techniques with experimental methods expedites the drug discovery process,

reduces costs, and increases the likelihood of success in developing new therapeutic agents. As computational technologies continue to advance, CADD will play an increasingly pivotal role in shaping the future of drug development.

Keywords: Medicinal Chemistry, Computer-Aided Drug Design, Molecular docking, Quantitative Structure Activity Relationship

Introduction

The process of developing new medications often takes a very long time and is very expensive. This is partly because of the necessity that new therapeutic entities be shown to be both effective and safe in clinical trials before being released into the market. It has been estimated that the process of moving from the stage of target evaluation to the stage of regulatory clearance for a single innovative small molecule may take up to 14 years and need an expenditure of more than one billion dollars [1,2]. It is common knowledge that the pharmaceutical sector produces a considerable number of goods that do not meet customer expectations. According to some estimations, the probability of a molecule developing into a useful medicine is somewhere between one and two out of every 10,000 molecules that are investigated [3].

The enlargement of the chemical space has resulted in the creation of a "drug-like" environment, which has become one of the most critical challenges facing the pharmaceutical and medical industries. The entire number of seconds that comprise all living things in the universe is less than the estimated number of tiny molecules, which is roughly 1060 trillion [4]. This means that the estimated number of small molecules exceeds the total number of seconds. From an experimental point of view, the exploration of a chemical domain with such a wide extent presents a considerable set of problems. The discipline of high-throughput screening, often known as HTS, has seen considerable advancements in recent years, which has made it possible to examine the on-target activity of hundreds of thousands of compounds per week [5]. However, it is essential to keep in mind that the number of possible candidates for a particular biological entity places a ceiling on the total number of compounds that may be put through the testing process. Medicinal chemists have the option of using a "virtual environment" as a way of transferring the candidate selection issue outside the constraints of the laboratory in order to circumvent this limitation. This allows the problem to be moved outside of the lab entirely. Utilising computers to carry out "virtual" screenings of molecules was one

of the first suggestions that was made, and this was done before any actual tests were carried out in the lab. The process that is referred to as "high-throughput virtual screening" (HTVS) is one of the principal uses of computational techniques in the pharmaceutical business [6]. The processing capability of the infrastructure that is used for this purpose is the primary determinant of how well virtual screening can be carried out. It provides an alternative to the planning and carrying out of real testing that is both noticeably quicker and more cost-effective. Recently, researchers at Oak Ridge National Laboratory used the SUMMIT supercomputer to carry out a GPU-accelerated virtual screening against the SARS-CoV-2 major protease [11]. This was done in order to speed up the screening process. This effective demonstration illustrates the possibility for analysing an enormous number of chemicals on a daily basis utilising legitimate software and hardware combinations [7,8,9]. The number of compounds that might be evaluated each day ranges from millions to billions. In recent years, both academic institutions and private companies have made substantial investments in the research and development of these methodologies. As a direct consequence of this, they have evolved into an essential element of the contemporary drug development pipeline, especially in the first stages of drug discovery.

The Use of Computational Methods in the Drug Discovery Process

According to the guidance provided by the United States Food and Drug Administration (US FDA), the process of developing a new medication may be broken down into five separate stages [12]. This segmentation results in enhanced clarity and understanding for the audience. The "discovery and development" stage is the first step in the process. It is also often referred to simply as "stage 1." This stage includes a variety of tasks, such as hit-to-lead (H2L) conversion, lead creation, and lead optimisation. In the first phase, promising drug candidates are narrowed down based on the pharmacokinetic (PK) or pharmacodynamic (PD) features they exhibit. After that, certain molecules are selected on the basis of the ideal activity profile that they exhibit in opposition to the target. Molecules that are hit often have activity in the micromolar (M) range and exhibit reduced selectivity, particularly in terms of their on-target potency. In spite of these restrictions, the hit compounds are very important since they steer drug design teams in the right direction and provide good starting points for further modifications [13,14].

In what follows, we'll go over several tactics for getting the most out of your website visitors in terms of generating leads. At this time, the hit compounds are being optimised by having different changes made to them in order to improve both their activity on the target and their selectivity. During this whole procedure, the pharmacokinetic and pharmacodynamic characteristics of the drugs are under close observation [15]. The compounds that are produced as a consequence of this procedure, which are often referred to as "lead" compounds, have substantial potency in the nanomolar range while also demonstrating moderate selectivity. The second main step consists of performing experiments on animal and organoid models to establish whether or not the chemicals are both safe and effective before moving on to the third primary phase, which consists of conducting clinical studies on humans. In the third and final step, extensive human clinical studies will be carried out. The process of testing is broken down into three separate parts, which are referred to as Phase I, Phase II, and Phase III respectively. Each phase is distinguished by the particular goals it aims to accomplish and the growing number of people who take part in it. After a successful conclusion of the clinical phase III study, the pharmaceutical business will be able to go on to the fourth step of medication development. After that, the business may go through with the process of submitting a formal request for commercialization to the relevant regulatory bodies, such as the European Medicines Agency (EMEA) for Europe and the Food and Drug Administration (FDA) for the United States. Post-market pharmaceutical safety monitoring is the following step, which is also the fifth and last phase of the process.

In spite of the significant financial investment and time commitment required, preclinical and clinical testing are essential components of the process of developing new drugs. Because of the sophisticated nature of the data about the safety and effectiveness that these studies produce, they are very necessary. The use of computational design tools has shown to be quite successful in improving the procedures that are involved in moving from hit discovery to lead compound development. These technologies not only permit a large increase in the daily volume of virtual compounds that can be evaluated, but they also make it possible to conduct an in-depth analysis of the patterns present in the chemical data that is being considered [16]. In addition to this, they make it possible to create such entities in a sensible way. Drug design groups are now able to include visual analysis of proteins, ligands, and physiologically relevant complexes into their routine workflows as a result of the integration of spectroscopic technologies and the fast development of computer graphics [17]. This group of techniques is referred to as "computer-aided drug design" (CADD), which

comes from the fact that computational methods may be used to great advantage in the process of developing fresh molecular candidates.

Computer-aided design and drafting (CADD):

When solving a pharmaceutical problem, a computational chemist may choose one of many approaches, the primary distinction between which being the quantity of data at their disposal (Figure 1). The existence of experimental structural data relevant to the target of interest is an important consideration that must be taken into account [16]. There are a variety of different approaches one might use to get this information. However, the nuclear magnetic resonance (NMR), X-ray crystallography (XR), and cryogenic electron microscopy (cryo-EM) techniques are the ones that are the most applicable [18]. To make the most of the information they have access to, scientists often make use of a wide variety of computational approaches such as molecular docking and molecular dynamics. A group of techniques that are applied in the sector are referred to together as "structure-based drug design" (SBDD), which is an abbreviation. In situations when there is a lack of accessible experimental data on the desired three-dimensional structure, CADD specialists have the option of choosing between two major techniques. The first thing that has to be done is to develop a computer model of the thing that's being looked at using SBDD methods. This kind of model is often referred to as the homology model. After that, the structural dependability of this model is proven with the help of several close homologs [19]. In recent years, protein structure prediction has undergone a substantial transition, which can be mostly ascribed to the advent of AlphaFold [20], which has now been improved to version 2.0. This change was made possible by advancements in computing power and increased data storage capacity. The approach, which was created by DeepMind, makes use of several artificial intelligence (AI) methods in order to forecast the three-dimensional structure of a biological entity. The goal of the method is to achieve this prediction. This forecast is derived from the sequence of the item as well as a confidence score that is connected to the individual functional components of the object. This method's dependence on the homology models developed by scientists on an as-needed basis presents a further limitation of its use. In addition to this, it is only capable of forecasting a single conformational state for the targets that are wanted, which is often the state when the target is inactive. However, the development of AphaFill [21] has helped to alleviate some of these worries to some degree. Additionally, it is important to

remember that the AlphaFold database does not include all proteins in its scope of coverage. For example, the database does not yet include information on a great deal of viral proteins [22].

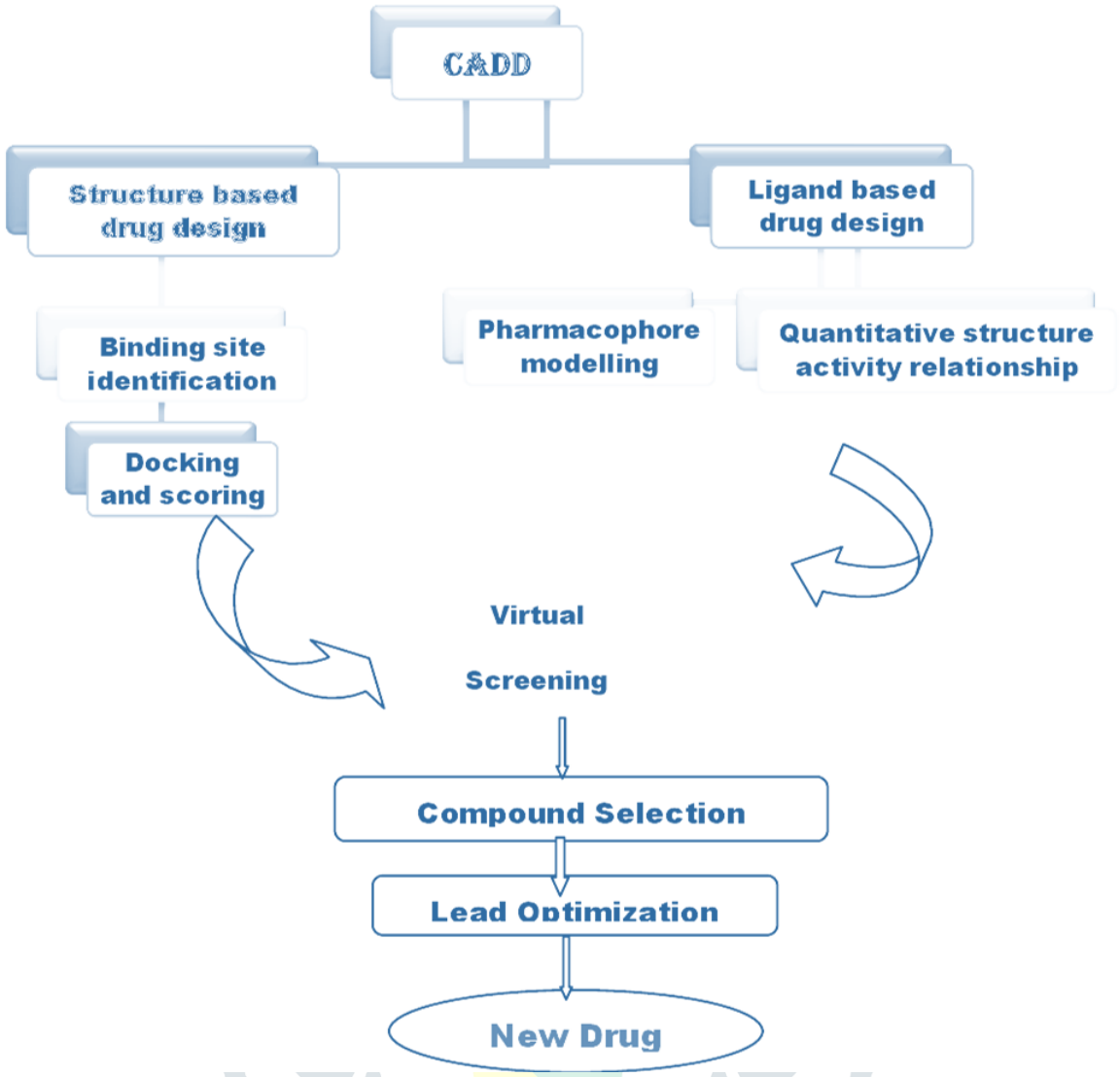

Figure 1. The key computational approaches that are used by scientists working in the field of computer-aided drug design (CADD) to create brand-new medicines are shown in a condensed and simplified. It has been shown how crucial it is to have access to any relevant information on the target structure. Although they are linked in many ways, the ideas of quantitative structure-activity relationship (QSAR) and quantitative structure-property relationship (QPR) have certain key distinctions that set them apart from one another. The combination of quantum mechanics and molecular mechanics is sometimes referred to by the acronyms QM/MM. This is a frequent practise.

The CADD scientist must make a difficult choice since there is insufficient information available on the structural components of the living thing. They need to decide whether to rely simply on the data gained from tested ligands or to extract sufficient information from these ligands in order to create reliable

quantitative structure-activity relationship (QSAR) models [23]. If they decide to rely primarily on the data received from tested ligands, they must decide whether to rely exclusively on the data obtained from tested ligands. What is now known as the "ligand-based drug design" (LBDD) approach was the foundation on which the first rational drug design approaches were built. Methods such as matching molecular pair analysis [25] and pharmacophore search [24] are included in this category of research approaches. Despite the fact that the development of SBDD procedures has led to a decrease in the use of these approaches over the course of time, they continue to see widespread use. Recent developments in computer science, in conjunction with the explosive increase in the use of machine learning (ML) and artificial intelligence (AI) methods, have offered experts in computational-aided drug design (CADD) with a new and powerful resource [26]. This new and powerful resource has enabled CADD specialists to develop new and more effective drugs. The inclusion of a large amount of data related to the target and ligands within a particular context increases the possibility of successfully predicting pharmacologically relevant molecular characteristics utilising these approaches. In addition to this, these "computational brains" have the potential to manufacture totally new chemical structures [27] by using a cutting-edge algorithm that has been devised and perfected in recent years. This is something that has been accomplished in recent times.

Ligand-Based Drug Design (LBDD).

Utilising structural data derived only from compounds that have been put through testing on the target is the technique that has shown to be the most successful in the early stages of rational drug design. The major goal of utilising these strategies is to uncover patterns within the data that can be extrapolated to influence the succeeding phases of drug development. This may be accomplished by comparing the results of previous iterations of the drug development process. Quantitative structure-activity relationship (QSAR) models are applied in order to establish connections between distinct chemical moieties and the pharmacological effects that are associated with each of those moieties. Scientists are able to recognise patterns and trends with the use of these models, which enables them to discover a relationship between the two variables. There are a few different approaches that may be used to determine interactions based on ligands. Cheminformatics [28], ligand-based pharmacophore search, and Free-Wilson analysis [29] are some of the approaches that fall under this category. Hansch, Hammet, and Taft [30] are credited with being the ones who derived a number of well-known equations that are used in QSAR modelling. The approaches that have been

discussed up to this point are still put to use; nevertheless, they all have one major drawback in common: they cannot be generalised. When using a ligand series that is substantially congeneric to one another, it is possible to acquire trustworthy findings. If, on the other hand, the ligand series is not very congeneric, then a considerable quantity of data gleaned from experiments is essential. Because they only focus on a two-dimensional depiction of the molecules under investigation, these techniques do not take into consideration the conformational flexibility of ligands. In LBDD strategies, there has been a great focus placed on taking into account the conformational characteristics of ligands. This demonstrates how important structure-based approaches are becoming overall, as well as the relevance of "three-dimensionality." This may be viewed as an example in the development of "3D pharmacophores" [31], which has been going on recently. These models provide realistic "3D-QSAR" models by making use of the atomic and structural features of the substances being modelled.

Quantitative structure-activity relationship (QSAR)

Ability to establish a link between chemical modifications and relevant biological functions was one of the original requirements for becoming a medicinal chemist. Medicinal chemists are responsible for developing new treatments for medical conditions. According to the postulated idea, much attention was not given to the target in the earlier phases of the design process; rather, the emphasis was placed predominantly on ligand small molecules. The method of designing these molecules required making modifications to their characteristics purely on the basis of observations made from the outcomes of tests. Quantitative structure-activity relationship (QSAR) modelling was the prevalent name for the study in question [32]. Significant methodological progress was made in this technique throughout the later half of the 20th century, which is largely responsible for its rising popularity in the field of drug design [33,34,35]. These developments may be traced back to the time period. During this time period, the word "cheminformatics" came into popular use to refer to the integration of computational techniques and the instruments associated with those methods with other methodological approaches. According to Gasteiger and Engel's definition of the word, it refers to the use of methodologies from the field of informatics to the solution of chemical issues [36]. Over the course of the last several decades, the area of data analysis for chemical data has seen tremendous expansion. This expansion has ranged from the development of fundamental tools to sophisticated cheminformatics suites and packages in widely used programming languages [37]. The creation of RDKit

[38], a widely used and adaptable cheminformatics software package for Python, which is regularly referenced and examined in academic papers [39,40,41], is of special relevance in this particular situation. [39] RDKit [38] is a software programme for Python that is extensively used. The RDKit library provides a wide variety of functions that are applicable to a variety of different kinds of work. These activities include, but are not limited to, molecular clustering, the search for substructures, the fragmentation of compounds, the control of chemical reactions, and the examination of structural and form similarities [42, 43].

Because of the importance of molecular modelling [43,44], more and more attention is being paid to the portrayal of chemical entities in three dimensions. Alongside the more conventional two-dimensional representation, this method is now being investigated on a regular basis. Because of the importance of conformer creation and prioritisation in chemical research and drug development [45], contemporary cheminformatics systems, such as RDKit, have included a variety of methods for generating conformers and ranking them in order of preference.

In spite of the fact that SBDD techniques have seen a meteoric rise in use within the pharmaceutical business, 3D-QSAR and cheminformatics are still widely applied [46,47]. In addition, the use of genuine machine learning methods has made it easier to integrate these tools. This has led to the creation of algorithms that are capable of independently locating structural patterns within chemical data and producing unique QSAR models [48].

The practise of structure-based drug design (SBDD)

Since the beginning of the 2000s, there has been an increase in the number of three-dimensional protein and nucleic acid structures, which has spurred a change in the computational drug design methodology. These methods have now been updated to include procedures that take into account the three-dimensional interaction properties of molecules in respect to the target. Because of the scientists' prior knowledge of the biological entity that was the subject of the inquiry, they had a substantial advantage in the scenario that they found themselves in. With this newfound information, the researchers were able to develop novel chemical compounds by capitalising on the particular characteristics of the binding site. A subset of methods known as "structure-based drug design" (SBDD) [49] is one of the regularly used strategies in the area of computational drug development. SBDD is an abbreviation for "structure-based drug design."

Additionally, the use of cryo-electron microscopy (cryo-EM) technology has substantially enlarged the scope of single-particle electron cryo-microscopy (SBDD), which is an acronym for single-beam deep-diffraction. Because of this progress, it is now possible to resolve complicated systems [18,50] accurately, even if they were previously thought to be impossible to examine experimentally. The most common SBDD procedures are shown in Figure 2, which offers a quick summary of these methods.

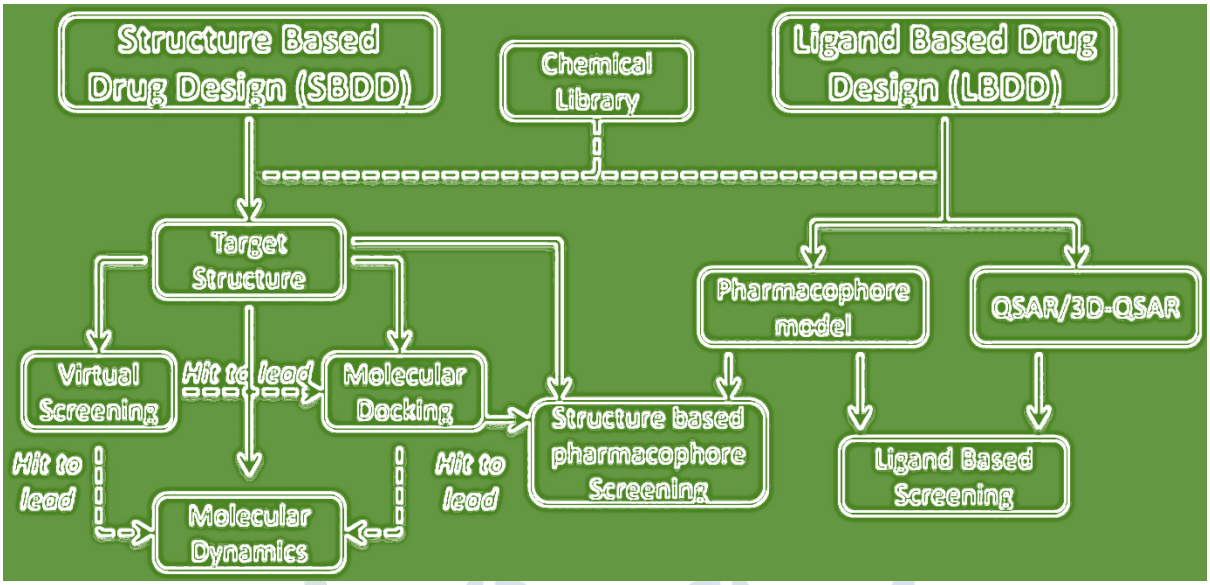

Figure 2: The primary SBDD approaches are broken down into their respective categories in accordance with explanation of their functions. The amount of processing power that is available is of utmost significance given the circumstances of this situation. It permits a differentiation to be made between techniques that is approximative and is based on the number of molecules that are screened in a given day while using the same computer infrastructure. When the main objective is to test tiny molecules against a specific biological target, this aspect also carries a substantial amount of weight in the evaluation process. In the context of this discussion, many abbreviations, including quantum mechanics/molecular mechanics (QM/MM), free-energy perturbation (FEP), thermal titration molecular dynamics (TTMD), and artificial intelligence (AI), are used.

Conclusion

Vital Component of Modern Pharmaceutical Research: Computer-aided drug design (CADD) is an indispensable element of contemporary pharmaceutical research. Traditionally, drug discovery relied heavily on trial-and-error experimentation, which was time-consuming, expensive, and often led to many failures. CADD has revolutionized this process by offering a systematic and data-driven approach.

Integration of Computational Techniques: CADD seamlessly integrates computational techniques with experimental methods. This synergy is critical because it allows researchers to leverage the power of computers to model and predict molecular interactions. Through molecular modeling, CADD enables scientists to visualize how potential drug molecules interact with specific biological targets, such as proteins or enzymes. This insight is invaluable for designing and selecting compounds with a higher likelihood of success.

Expedited Drug Discovery Process: One of the most notable advantages of CADD is its ability to expedite the drug discovery process. By using computational simulations and algorithms, researchers can quickly sift through vast chemical libraries to identify potential lead compounds. This reduces the time required to move from the initial drug discovery phase to clinical trials. In a field where time is often a critical factor, this acceleration is invaluable.

Cost Reduction: CADD also brings about significant cost savings in drug development. Traditional drug discovery can be incredibly expensive due to the extensive laboratory work involved, including the synthesis and testing of numerous compounds. CADD reduces the need for as many physical experiments by guiding researchers toward the most promising candidates. This not only saves money but also minimizes resource wastage.

Increased Likelihood of Success: By combining computational insights with experimental validation, CADD enhances the likelihood of success in developing new therapeutic agents. It helps in identifying compounds that have a higher probability of effectively interacting with the target biomolecule, thereby increasing the chances of developing safe and efficacious drugs.

Future of Drug Development: As computational technologies continue to advance at an unprecedented pace, CADD is poised to play an even more pivotal role in shaping the future of drug development. Artificial intelligence, machine learning, and big data analytics are increasingly being integrated into CADD workflows, enabling more accurate predictions and smarter decision-making in drug discovery.

In sum, computer-aided drug design is not just a tool but a transformative force in the pharmaceutical industry. It combines the strengths of computation and experimentation, streamlines drug discovery, saves resources, and holds the promise of delivering innovative and life-saving medications more efficiently than

ever before. As technology evolves, CADD will remain at the forefront of innovation in the pharmaceutical field, paving the way for a future with more effective and accessible drugs.

References

1. Martin, L.; Hutchens, M.; Hawkins, C. Clinical trial cycle times continue to increase despite industry efforts. Nat. Rev. Drug Discov. 2017, 16, 157. [Google Scholar] [CrossRef] [PubMed]
2. Simoens, S.; Huys, I. R&D Costs of New Medicines: A Landscape Analysis. Front. Med. 2021, 8, 760762. [Google Scholar] [CrossRef]
3. The Pharmaceutical Industry in Figures. Available online: https://www.efpia.eu/media/602709/the-pharmaceutical-industry-in-figures-2021.pdf (accessed on 2 March 2023).
4. Bohacek, R.S.; McMartin, C.; Guida, W.C. The art and practice of structure-based drug design: A molecular modeling perspective. Med. Res. Rev. 1996, 16, 3–50. [Google Scholar] [CrossRef]
5. Szymański, P.; Markowicz, M.; Mikiciuk-Olasik, E. Adaptation of High-Throughput Screening in Drug Discovery—Toxicological Screening Tests. Int. J. Mol. Sci. 2011, 13, 427–452. [Google Scholar] [CrossRef][Green Version]
6. Moro, S.; Bacilieri, M.; Deflorian, F. Combining ligand-based and structure-based drug design in the virtual screening arena. Expert Opin. Drug Discov. 2007, 2, 37–49. [Google Scholar] [CrossRef]
7. Petrović, D.; Scott, J.S.; Bodnarchuk, M.S.; Lorthioir, O.; Boyd, S.; Hughes, G.M.; Lane, J.; Wu, A.; Hargreaves, D.; Robinson, J.; et al. Virtual Screening in the Cloud Identifies Potent and Selective ROS1 Kinase Inhibitors. J. Chem. Inf. Model. 2022, 62, 3832–3843. [Google Scholar] [CrossRef]
8. Luttens, A.; Gullberg, H.; Abdurakhmanov, E.; Vo, D.D.; Akaberi, D.; Talibov, V.O.; Nekhotiaeva, N.; Vangeel, L.; De Jonghe, S.; Jochmans, D.; et al. Ultralarge Virtual Screening Identifies SARS-CoV-2 Main Protease Inhibitors with Broad-Spectrum Activity against Coronaviruses. J. Am. Chem. Soc. 2022, 144, 2905–2920. [Google Scholar] [CrossRef]
9. Gorgulla, C.; Boeszoermenyi, A.; Wang, Z.-F.; Fischer, P.D.; Coote, P.W.; Das, K.M.P.; Malets, Y.S.; Radchenko, D.S.; Moroz, Y.S.; Scott, D.A.; et al. An open-source drug discovery platform enables ultra-large virtual screens. Nature 2020, 580, 663–668. [Google Scholar] [CrossRef]
10. Gupta, G. Racing the Clock, COVID Killer Sought among a Billion Molecules. Available online: https://blogs.nvidia.com/blog/2020/05/26/covid-autodock-summit-ornl/ (accessed on 2 March 2023).

11. LeGrand, S.; Scheinberg, A.; Tillack, A.F.; Thavappiragasam, M.; Vermaas, J.V.; Agarwal, R.; Larkin, J.; Poole, D.; Santos-Martins, D.; Solis-Vasquez, L.; et al. GPU-Accelerated Drug Discovery with Docking on the Summit Supercomputer. In Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, Virtual Event, 21–24 September 2020; pp. 1–10. [Google Scholar] [CrossRef]
12. FDA. The Drug Development Process. Available online: https://www.fda.gov/patients/learn-about-drug-and-device-approvals/drug-development-process (accessed on 2 March 2023).
13. Zhu, T.; Cao, S.; Su, P.-C.; Patel, R.; Shah, D.; Chokshi, H.B.; Szukala, R.; Johnson, M.E.; Hevener, K.E. Hit Identification and Optimization in Virtual Screening: Practical Recommendations Based on a Critical Literature Analysis. J. Med. Chem. 2013, 56, 6560–6572. [Google Scholar] [CrossRef][Green Version]
14. Hughes, J.P.; Rees, S.; Kalindjian, S.B.; Philpott, K.L. Principles of early drug discovery. Br. J. Pharmacol. 2011, 162, 1239–1249. [Google Scholar] [CrossRef][Green Version]
15. Keseru, G.M.; Makara, G.M. Hit discovery and hit-to-lead approaches. Drug Discov. Today 2006, 11, 741–748. [Google Scholar] [CrossRef]
16. Leelananda, S.P.; Lindert, S. Computational methods in drug discovery. Beilstein J. Org. Chem. 2016, 12, 2694–2718. [Google Scholar] [CrossRef][Green Version]
17. Dauter, Z.; Wlodawer, A. Progress in protein crystallography. Protein Pept. Lett. 2016, 23, 201–210. [Google Scholar] [CrossRef]
18. Benjin, X.; Ling, L. Developments, applications, and prospects of cryo-electron microscopy. Protein Sci. 2020, 29, 872–882. [Google Scholar] [CrossRef]
19. Cavasotto, C.N.; Phatak, S.S. Homology modeling in drug discovery: Current trends and applications. Drug Discov. Today 2009, 14, 676–683. [Google Scholar] [CrossRef]
20. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021, 596, 583–589. [Google Scholar] [CrossRef] [PubMed]
21. Hekkelman, M.L.; de Vries, I.; Joosten, R.P.; Perrakis, A. AlphaFill: Enriching AlphaFold models with ligands and cofactors. Nat. Methods 2023, 20, 205–213. [Google Scholar] [CrossRef]

22. David, A.; Islam, S.; Tankhilevich, E.; Sternberg, M.J. The AlphaFold Database of Protein Structures: A Biologist's Guide. J. Mol. Biol. 2021, 434, 167336. [Google Scholar] [CrossRef]
23. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR Modeling: Where Have You Been? Where Are You Going to? J. Med. Chem. 2014, 57, 4977–5010. [Google Scholar] [CrossRef][Green Version]
24. Yang, S.-Y. Pharmacophore modeling and applications in drug discovery: Challenges and recent advances. Drug Discov. 2010, 15, 444–450. [Google Scholar] [CrossRef]
25. Kramer, C.; Fuchs, J.E.; Whitebread, S.; Gedeck, P.; Liedl, K.R. Matched Molecular Pair Analysis: Significance and the Impact of Experimental Uncertainty. J. Med. Chem. 2014, 57, 3786–3802. [Google Scholar] [CrossRef] [PubMed]
26. Zhong, F.; Xing, J.; Li, X.; Liu, X.; Fu, Z.; Xiong, Z.; Lu, D.; Wu, X.; Zhao, J.; Tan, X.; et al. Artificial intelligence in drug design. Sci. China Life Sci. 2018, 61, 1191–1204. [Google Scholar] [CrossRef] [PubMed]
27. Mouchlis, V.D.; Afantitis, A.; Serra, A.; Fratello, M.; Papadiamantis, A.G.; Aidinis, V.; Lynch, I.; Greco, D.; Melagraki, G. Advances in De Novo Drug Design: From Conventional to Machine Learning Methods. Int. J. Mol. Sci. 2021, 22, 1676. [Google Scholar] [CrossRef]
28. Begam, B.; Kumar, J.S. A Study on Cheminformatics and its Applications on Modern Drug Discovery. Procedia Eng. 2012, 38, 1264–1275. [Google Scholar] [CrossRef][Green Version]
29. Kubinyi, H. Free Wilson Analysis. Theory, Applications and its Relationship to Hansch Analysis. Quant. Struct. Relatsh. 1988, 7, 121–133. [Google Scholar] [CrossRef]
30. Silakari, O.; Singh, P.K. QSAR: Descriptor calculations, model generation, validation and their application. In Concepts and Experimental Protocols of Modelling and Informatics in Drug Design; Elsevier: Amsterdam, The Netherlands, 2021; pp. 29–63. [Google Scholar] [CrossRef]
31. Ragno, R.; Esposito, V.; Di Mario, M.; Masiello, S.; Viscovo, M.; Cramer, R.D. Teaching and Learning Computational Drug Design: Student Investigations of 3D Quantitative Structure–Activity Relationships through Web Applications. J. Chem. Educ. 2020, 97, 1922–1930. [Google Scholar] [CrossRef]

32. Dearden, J.C. The History and Development of Quantitative Structure-Activity Relationships (QSARs). Int. J. Quant. Struct. Relatsh. 2016, 1, 1–44. [Google Scholar] [CrossRef][Green Version]
33. TSoares, T.A.; Nunes-Alves, A.; Mazzolari, A.; Ruggiu, F.; Wei, G.-W.; Merz, K. The (Re)-Evolution of Quantitative Structure–Activity Relationship (QSAR) Studies Propelled by the Surge of Machine Learning Methods. J. Chem. Inf. Model. 2022, 62, 5317–5320. [Google Scholar] [CrossRef]
34. Golbraikh, A.; Wang, X.S.; Zhu, H.; Tropsha, A. Predictive QSAR Modeling: Methods and Applications in Drug Discovery and Chemical Risk Assessment. In Handbook of Computational Chemistry; Springer: Dordrecht, The Netherlands, 2012; pp. 1309–1342. [Google Scholar] [CrossRef]
35. Verma, J.; Khedkar, V.M.; Coutinho, E.C. 3D-QSAR in Drug Design-A Review. Curr. Top. Med. Chem. 2010, 10, 95–115. [Google Scholar] [CrossRef]
36. Gasteiger, J.; Engel, T. Chemoinformatics: A Textbook; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2003. [Google Scholar]
37. Muchmore, S.W.; Edmunds, J.J.; Stewart, K.D.; Hajduk, P.J. Cheminformatic Tools for Medicinal Chemists. J. Med. Chem. 2010, 53, 4830–4841. [Google Scholar] [CrossRef]
38. Landrum, G. RDKit: Open-Source Cheminformatics. 2010. Available online: https://www.rdkit.org/ (accessed on 2 March 2023).
39. Dalke, A.; Hert, J.; Kramer, C. mmpdb: An Open-Source Matched Molecular Pair Platform for Large Multiproperty Data Sets. J. Chem. Inf. Model. 2018, 58, 902–910. [Google Scholar] [CrossRef] [PubMed]
40. Bolcato, G.; Heid, E.; Boström, J. On the Value of Using 3D Shape and Electrostatic Similarities in Deep Generative Methods. J. Chem. Inf. Model. 2022, 62, 1388–1398. [Google Scholar] [CrossRef] [PubMed]
41. Spiegel, J.O.; Durrant, J.D. AutoGrow4: An open-source genetic algorithm for de novo drug design and lead optimization. J. Cheminform. 2020, 12, 25. [Google Scholar] [CrossRef]

42. Riniker, S.; Landrum, G.A. Similarity maps-a visualization strategy for molecular fingerprints and machine-learning methods. J. Cheminform. 2013, 5, 43. [Google Scholar] [CrossRef][Green Version]
43. Daré, J.K.; Freitas, M.P. Is conformation relevant for QSAR purposes? 2D Chemical representation in a 3D-QSAR perspective. J. Comput. Chem. 2022, 43, 917–922. [Google Scholar] [CrossRef]
44. Nikonenko, A.; Zankov, D.; Baskin, I.; Madzhidov, T.; Polishchuk, P. Multiple Conformer Descriptors for QSAR Modeling. Mol. Inform. 2021, 40, 2060030. [Google Scholar] [CrossRef]
45. Günther, S.; Senger, C.; Michalsky, E.; Goede, A.; Preissner, R. Representation of target-bound drugs by computed conformers: Implications for conformational libraries. BMC Bioinform. 2006, 7, 293. [Google Scholar] [CrossRef][Green Version]
46. Bosc, N.; Atkinson, F.; Felix, E.; Gaulton, A.; Hersey, A.; Leach, A.R. Large scale comparison of QSAR and conformal prediction methods and their applications in drug discovery. J. Cheminform. 2019, 11, 4. [Google Scholar] [CrossRef][Green Version]
47. Neves, B.J.; Braga, R.C.; Melo-Filho, C.C.; Moreira-Filho, J.T.; Muratov, E.N.; Andrade, C.H. QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. Front. Pharmacol. 2018, 9, 1275. [Google Scholar] [CrossRef][Green Version]
48. Kwon, S.; Bae, H.; Jo, J.; Yoon, S. Comprehensive ensemble in QSAR prediction for drug discovery. BMC Bioinform. 2019, 20, 521. [Google Scholar] [CrossRef][Green Version]
49. Anderson, A.C. The Process of Structure-Based Drug Design. Chem. Biol. 2003, 10, 787–797. [Google Scholar] [CrossRef] [PubMed][Green Version]
50. Pavan, M.; Bassani, D.; Sturlese, M.; Moro, S. From the Wuhan-Hu-1 strain to the XD and XE variants: Is targeting the SARS-CoV-2 spike protein still a pharmaceutically relevant option against COVID-19? J. Enzyme Inhib. Med. Chem. 2022, 37, 1704–1714. [Google Scholar] [CrossRef]