Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

Structure-Based Drug Design: In-Depth Review

Mustafa Sabri Cheyad [1] * , Alaa Hussein J Al-Qaisi [2] , Ahmed Ahmed [3]

Department of Chemistry, College of Science, Al-Nahrain University, Baghdad, Iraq

*** Corresponding Author:** Mustafa Sabri, Department of Chemistry, College of Science, Al-Nahrain University, Baghdad, Al-Jadriah, Iraq, Telephone: +9647715059845, E-mail : mustafafallen@gmail.com

Abstract

In the last decades the development of computational approaches has undergone a major breakthrough. The growing fields of virtual screening and X-ray crystallography have a huge impact on structure-based drug design (SBDD) in developing new therapeutic agents. Drug discovery is a challenging and time-consuming process, Consequently SBDD has given a significant rule in reducing these challenges and there is a new compound is found every year, which is not introduced into the market, which indicates that the electronic information of these possible compounds needs to be stored in a database and better understand the known chemistry. In the past decade, the role of molecular modeling in sedate plan, which has been improved, increase in computing resources and molecular models, and software development, finally introduced a competitive advantage in the first early stages of drug discovery. The process of bringing effective drugs to market requires a lot of money and time investment, in this review which shows a brief advance of SBDD in the current stages of drug discovery.

Keywords: SBDD, Structure-Based Drug Design, Molecular Docking, Computational Stimulation.

Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

Introduction

Drugs designed to treat or alleviate disease Symptoms; a clear understanding of disease pathways and related processes are essential for selecting therapeutic targets. Therefore, in the past, we used programs such as Gene Go and KEGG to construct a pathway map of hypoxia-inducible factors in brain injury to identify key signals or transcriptional pathways, specific protein-protein interactions and the relationship between upstream and downstream proteins. Practice has proved that this background knowledge helps us choose key therapeutic targets (Marc et al,2000); Computer-aided drug design is a process that facilitates calculation resources and methods which have been used to design and discover new and feasible therapeutic agents, Drug discovery process (Gowri, 2015). In the past few years, computer-aided drug design (CADD) has developed rapidly, enhancing the ability to perceive multiple and difficult biological processes with the help of these calculation tools, new pharmacologically active agents can now be discovered in a short time. (Sam and Xiaoqin, 2014).

Protein structure determination

To determine the protein structure of structure-based drug design (SBDD), before studying relationship between receptors and ligands, the target structure must first be obtained, there are two main methods for determining protein structure by physical methods: X-ray diffraction and NMR (Gowri, 2015). The resolved protein structures could be found on (PDB) Protein Data Bank (www.rcsb.org/), however, for proteins that have not yet been resolved or are very difficult to isolate, modeling methods can be used. Ongoing advances in computational innovation and X-ray crystallography give an increasingly

compelling strategy to characterizing the key cooperation among ligands and binding sites of the target, computational methods, and techniques for sedate revelation can be generally divided into atomic docking and pharmacophore modeling. These techniques are regularly utilized to create models for virtual screening of VEGF (VEGFR TKI) receptor tyrosine kinase inhibitors. (Gerhard and Thierry, 2005), A pharmacological model is a virtual compound substance containing steric hindrance, hydrophobic and electronic properties, and is relied upon to interface with the objective restricting pocket, These qualities incorporate hydrophobic regions, hydrogen bond givers, ionizable groups and hydrogen bond acceptors.

Molecular Docking:

Docking is a decisive step for SBDD and helps to visualize the interaction mode and binding energy of the ligand – receptor (encoded as 1CRK downloaded from PDB) and performed using molecular operating environment (MOE) as shown in figure 1. complex, since the creation of the first docking program; researchers can now use more than 50 different docking programs.

SBDD has no prescribed tools, but the different algorithms and scoring functions of each software will make the docking results different (Gerhard, 2006).

before docking, a series of preparations should be completed, including the 3D structure of the target protein and ligand. In addition, it has been shown that the ligand conformation initially input affects the determination of the binding posture, and its effect on accuracy is 10-50%. Ligands usually bind to macromolecules in a low-energy conformation, but there is no correlation between binding affinity and strain energy, which indicating that in some cases a wide range of conformational rearrangements can be tolerated (Sundarapandian et al., 2011). An example of docking we perfumed is the docking of dipyridamole and creatine kinase (encoded as 1CRK downloaded from PDB) and performed using molecular operating environment (MOE) as shown in figure 1.

Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

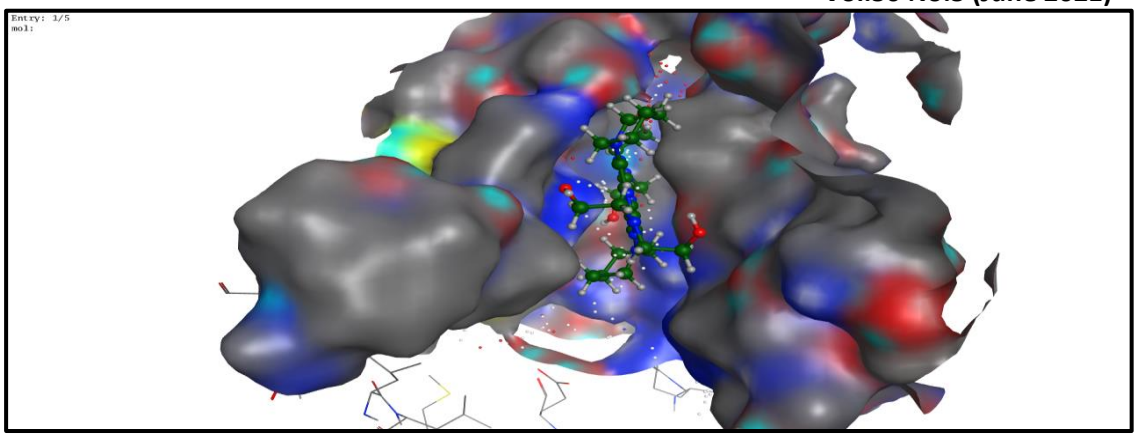


Figure 1: 3D binding modes of creatine kinase with dipyridamole

Molecular docking helps to investigate the interactions of ligand-receptor in order to identify the active sites of the target proteins. The most energy-stable geometric structure of the ligand-receptor complex is obtained by docking. The energy of interaction can be characterized by different scoring functions, for example, dock score, average force score potential, and electrostatic steric score and score. This score helps to indicate the binding modes of affinity of the ligand to the receptor site (Doug et al., 1996) as shown in Figure 2 where shown the ligand-receptor interactions of dipyridamole and creatine kinase.

STRUCTURE-BASED DRUG DESIGN

The utilization of three-dimensional (3D) protein crystal structure data in the advancement of new bioactive atoms, the approach of (SBDD), is an effective, mature and highly attractive procedure utilized by scholarly and pharmaceutical research centers around the world (Sachin and Mariam, 2007). As an inventive and information driven methodology, the fundamental requirements of structure-based examination is a significant comprehension of the spatial and vitality viewpoints that influence the binding affinity of protein ligand buildings.

Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

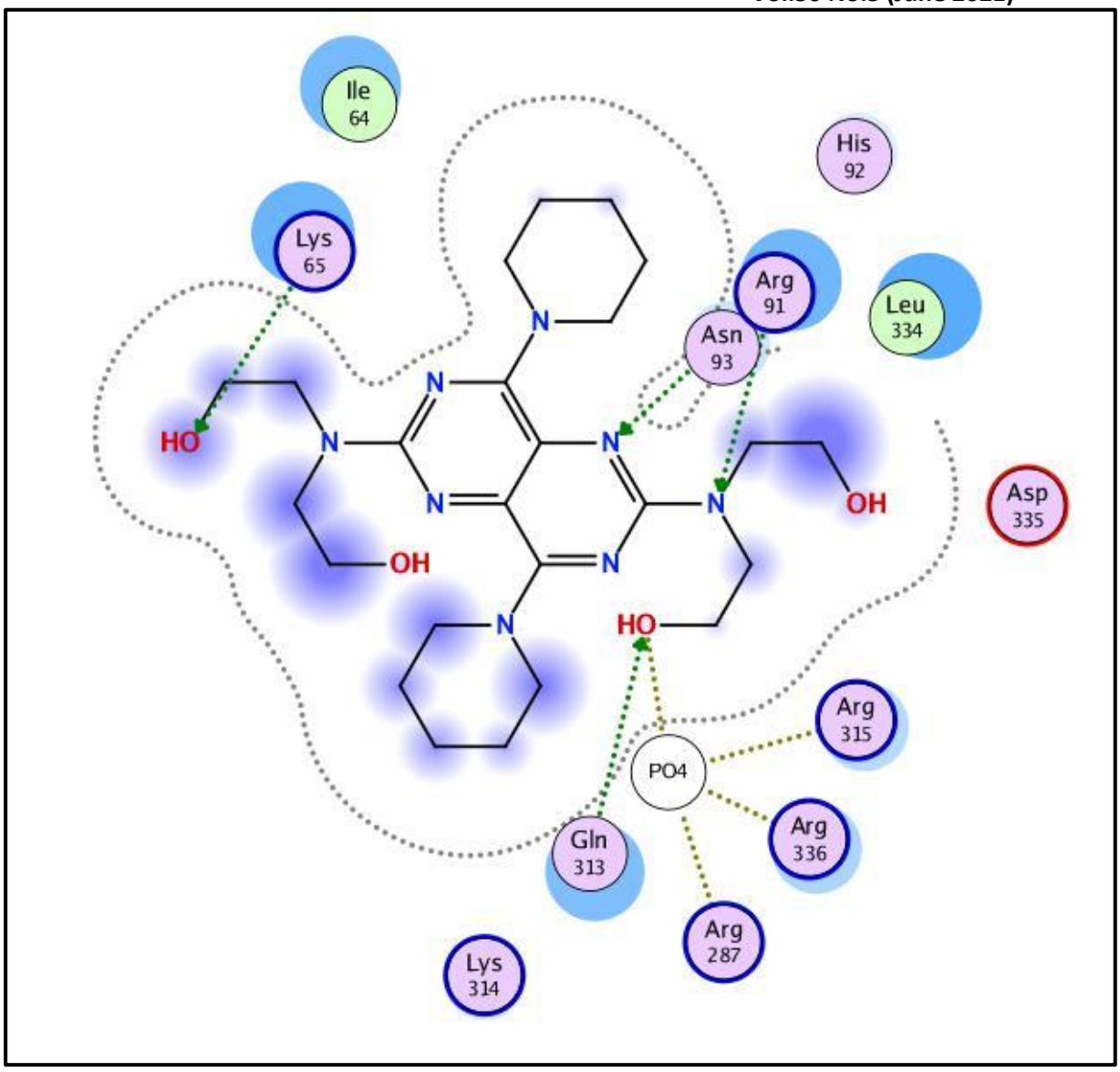


Figure 2: 2D binding modes of creatine kinase with dipyridamole

Taking into account the chemical properties and shape of explicit target protein active sites are known, and the conceivable intermolecular associations between the ligand and the protein in its dynamic site have been characterized, this certified data can be legitimately utilized for ID Optimization of new ligands and lead compounds (Adriano et al., 2009). This opens up additional opportunities that can encourage the search for lead particles and breaking point the quantity of exacerbates that should be assessed experimentally (Claudio and Andrew, 2007).

The Process Analysis of Structure-Based Drug Design

Knowledge of parasite biology and the availability of entire proteomics and genomes that provide an insightful information for drug

costly and tedious medication

screening objective. Direct screening and advancement consider (Montserrat and Jordi, 2009).

Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

design (Najib et al., 2005). a large amount of information that generated by these experiments leads to the need for reliable and powerful data processing methods in order to analyze and identify related-molecular targets. This circumstance has prompted the advancement of coordinated methods dependent on the information on various disciplines, for example, science, biology and informatics (Maria et al., 2019). in this specific situation, chemical genomics is an appealing and helpful methodology for characterizing and assessing substance and organic spaces. To build the effectiveness of early revelation and accelerate the screening procedure, compound genomics techniques are broadly used to allot need to the most encouraging atomic receptors for additional exploration, empowering determination of potential preceding costly and tedious medication screening objective. Direct screening and advancement consider (Montserrat and Jordi, 2009). situation, chemical genomics is an appealing and helpful methodology for characterizing and assessing substance and organic spaces. To build the effectiveness of early revelation and accelerate the screening procedure, compound genomics techniques are broadly used to allot need to the most encouraging atomic receptors for additional exploration, empowering determination of potential preceding
which provides insightful information into related ligand-receptor interactions, which can be used to enhance the biological properties of the synthesized leads.

SBDD - Tropical Diseases Strategies

Since the discovery of antiparasitic drugs is not entirely driven by commerce, other methods of developing new drugs have been developed. It has been shown that many drugs used to treat many diseases were first developed for parasitic infections. This method, which is known as label expansion, has standalone advantages, such as short-term development cost and time reduction. Since the drug was first studied for other

SBDD - TROPICAL DISEASES Relationship

In the modern medical chemistry, SBDD has become an essential part (Fernan et al., 2008). During the last decade, algorithms, high-performance computers, expertise and methods have been developed and transformed into SBDD approaches represent tools that have a significant impact on drug discovery (Rafael et al., 2008). They combine information from target receptors, so these knowledge-driven methods require a lot of information about the target topology under study, such as NMR or X-ray structures or robust homology modeling (Douglas et al., 2004).

In this case, public well-known databases such as ExPASy, protein Data Bank (PDB), InterPro and Relibase are essential invaluable data sources for analyzing and searching the 3D structure of target proteins. The molecular characterization phenomenon depends on the nature and characteristics of the binding sites cavity, which is greatly influenced by the presence of amino acids in the binding sites cavity. The arrangement of amino acids within the binding site dictates the physicochemical and structural constraints which any putative ligand must meet (Antoine et al., 2017). As for that, in-depth analysis of the target's binding pocket such as stereo electronic properties

Fragment-based lead discovery and SBDD

For this method complements the traditional method and is a new pattern in order to generate lead synthesized compounds in the process of drug development. In this process, the potential high-affinity ligands are identified, optimized, and linked based on the structure-activity relationship (SAR) between the structure and the molecule. SAR benefits in using techniques such as NMR and various other techniques that concerns the relationship between small chemical fragments or groups and their

indications and has been already approved for treatment, the drug shows promising toxicity and pharmacokinetics characteristics. On the contrary, some pharmaceutical companies in the medical industry are fully aware of the medical risk of unexpected toxic effects and are subjected to allow further research on their drugs to discover neglected diseases (Solomon and Alan, 2006).

the proliferation of structural and genomic (eg, proteomics, metabolomics and genomics) obtained data for such protein targets, SBDD became a very imperative and powerful tool for discovering lead compounds, lead synthesized compounds can be optimized to accelerate drug development and also take part of marketable drugs development.

protein targets active sites (Harren et al., 2013).

The standards criteria for choosing chemical fragments are based on the fact that their chemical group functions should be simpler than drug structures , weak in potency but effective in combination, high in purity and good in water solubility, and follow the "three-point method" with a molecular weight of <300, ClogP <3, and hydrogen holding The quantity of contributors, acceptors and rotatable bonds ought to be 3 individually (David, 2007).

Structure-Based Virtual Screening

This is a stimulation technique that includes rapidly looking through a huge library of compound structures to recognize potential medication competitors (hits) (protein receptors, catalysts) that are destined to tie to drug targets, and afterward coordinating the hit docking protein

targets and use of the scoring function to evaluate the probability that the medication up-and-comer will tie to the protein focus with highly affinity (Tiejun et al., 2012). The fundamental favorable position of this screening is that by extraordinarily diminishing the quantity of mixes utilized for trial assessment of its activity, the hit rate can be expanded, along these lines expanding the achievement pace of in vitro explores. Structure-based virtual screening has been greatly utilized in academic communities and pharmaceutical organizations for premature medication discovery.

Lead Optimization and in Silico Structure-Based

ADMET modeling

ADMET represents absorption, distribution, digestion, excretion and toxicology of medications. This strategy is also termed pharmacokinetic modeling based on physiology. It is utilized in drug examination, advancement, and health threat assessments to foresee the ADMET of mixes in people. There are various instruments accessible for imagining and foreseeing ADMET information: ADMET parameter predictor, metabolism predictor, the predictor of metabolic stability, cytochrome P450 substrate predictor and physiological-based pharmacokinetic (PBPK) displaying programming (Suzanne et al., 1996). ADMET boundaries Its performance and pharmacological influence movement of the compound as a medication dependent on the energy of its presentation to tissues and how

Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

the necessary hits have been established thro ugh virtual screening; this strategy can be use d depending on the choice of lead mixes as a r esult of their high affinity to the objective and explaining its characteristics of absorption , d istribution, digestion, excretion and toxicity (ADMET) and its physicochemical attributes, which increase the probability of achievement in clinical stage study.

Various calculation methods could be used to achieve this optimization, counting quantitative structure-action connections, pharmacophores, closeness look, homology models, databases, information mining, organize examination, and information investigation apparatuses. In any case, similar to every strategy, this technique has its own area (Ekins et al., 2007). Figure 3 shows the general setrategy for CADD.

the human body is dealt with or gotten, in this manner giving key bits of knowledge into the properties of drug compounds in the body.

This strategy assumes an essential job in drug discovery and improvement since it assists with choosing compounds in the beginning Discovery phases and significantly decreases the late use of drugs during preclinical clinical stage trials, so it has end up being Cost-powerful and time-productive innovation (David, 2007.

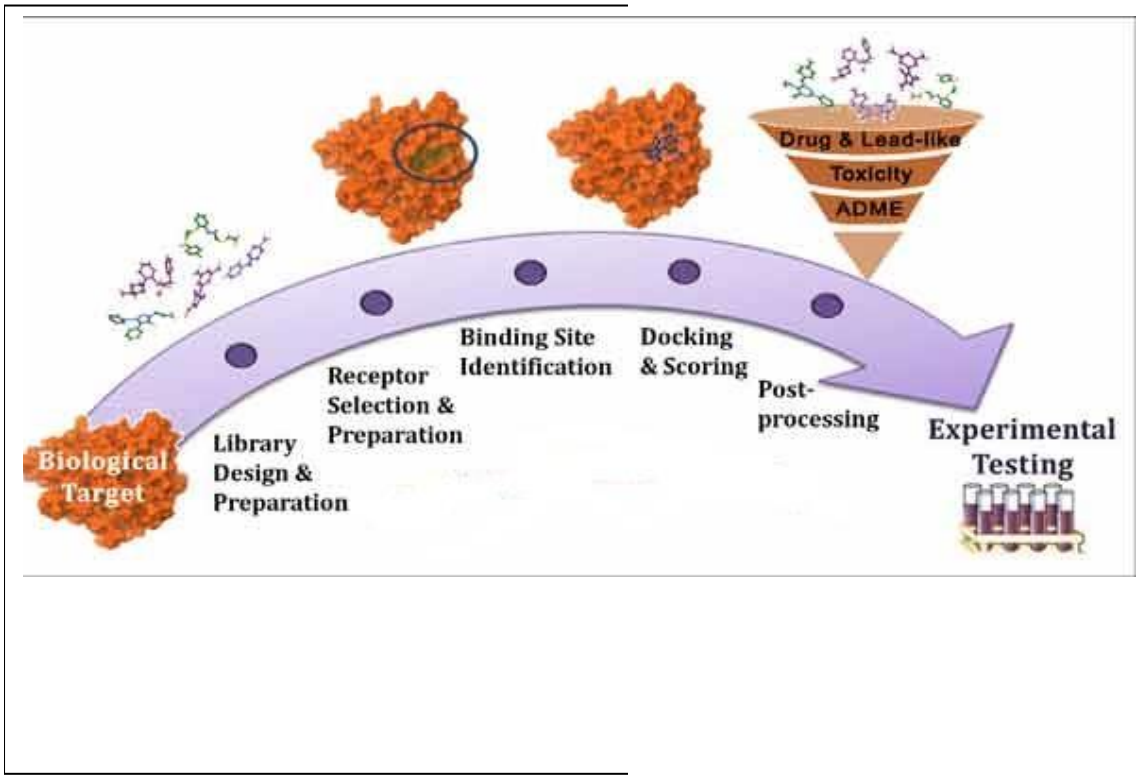


Figure 3 : General Setrategy for Computer-Aided Drug Design

Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

Restricting Site Examination

The selection of the binding active site of the target macromolecule can provide clear information about protein ligand interaction phenomena, post-docking kinetics, hydrogen bond formation, complex free energy, etc., so that the best pharmacophore of the new ligand can be calculated (Lei et al., 2017). The binding site in the target macromolecule can be determined experimentally, for example using site-directed mutagenesis or X-ray crystallography. The massive protein literature co-crystallizing with its substrates or known inhibitors provides useful information about binding sites.

Ligand Library Preparation

Ligands or drug-like compounds can be selected from natural ligands, public knowledge bases or commercial resources to construct libraries (Maria, 2017). In addition, in order to improve accuracy, ligands should be screened to limit the limitations of "Lipinski's 5 rules" (log P value, molecular weight, number of hydrogen bond donors and acceptors) (Jianling and Suzanne, 2009).

ADMET risk score (absorption, distribution, metabolism, excretion and toxicity) and risk parameter size, water solubility, charge, carcinogenicity, acute rat toxicity, distribution, elevated serum levels glutamate oxaloacetate aminotransferase, liver toxicity and The inhibitory effect of midazolam 3A4 oxidation also needs to consider ligands before docking), which helps to obtain bioactive ligands with pharmacokinetic properties and drug safety (Gregor et al., 2015). At the same time, the synthetic accessibility of these compounds should be valued; some programs (SYLVIAXT) provide visual parameters to measure the difficulty of synthesis. In addition, if no quantitative structure-activity relationship (QSAR) studies have been reported, the number of candidates can be further reduced.

Docking and Scoring Functions

Docking is a decisive step for SBDD and helps to visualize the interaction mode and binding energy of the ligand-receptor complex. Since the creation of the first docking program, researchers can now use more than 50 different docking programs. SBDD has no prescribed tools, However, the different algorithms and scoring functions of each software will make the docking results different. Before docking, a series of preparations should be completed, including the 3D structure of the target protein and ligand (Emanuele and Paul, 2004).

In addition, it has been shown that the ligand conformation initially input affects the determination of the binding posture, and its effect on accuracy is 10-50%. Ligands usually bind to macromolecules in a low-energy conformation, But there is no correlation between strain energy and binding affinity, which indicates that a wide range of conformation rearrangements may be tolerated in some cases (Mohamed, 2013). In this case, it is usually necessary to obtain multiple initial ligand conformations for each ligand. Some software can generate 3D structures and different conformations of ligands for docking (Jie and Renxiao, 2015).

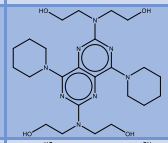
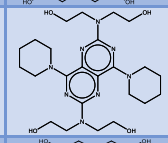
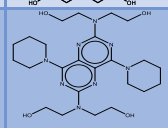
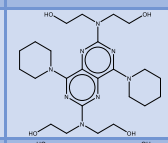
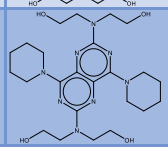
For all molecular docking procedures, an important issue is to develop an energy

Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

scoring function to quickly and accurately describe and estimate the fitness of protein-ligands. The scoring function is mainly used for analysis for after docking. Since the early 1990s, This published hundreds of scoring functions. In theory, the scoring function can be represented by the binding affinity between the target and ligand the macromolecule by calculating non-covalent interactions and providing basic features such as speed and accuracy scoring functions. The scoring function is usually divided into three basic types

Type: Force field analysis, experience scoring function and knowledge-based scoring function (Pui et al., 2015). The table 1 shown the scoring function data of docking dipyridamole and creatine kinase, the greater S values (in minus) the greater the affinity binding between the ligand and the receptor in case of dipyridamole and creatine kinase the highest S value was -8.3919 which indicates great affinity of binding between dipyridamole and creatine kinase.

Table 1 : The S scoring functions of docking dipyridamole and creatine kinase.

	mol	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine
1		1	1	-8.3919	2.0930	-43.4000	-92.4914	-10.9939	-54.7367
2		1	1	-7.8254	1.5798	-50.8109	-91.9310	-12.0774	-44.3770
3		1	1	-7.8133	3.9545	-45.7732	-51.5978	-10.3034	-47.6188
4		1	1	-7.4766	1.7105	-44.4261	-70.2046	-12.1737	-30.4587
5		1	1	-7.4307	2.3881	-45.0870	-92.2869	-10.2434	-47.0243

Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

CONCLUSION

As more and more bio molecular 3D structures are identified through experimental or computational methods, the SBDD method is timesaving, cost-effective, and productive, and has been commonly used to design and discover new lead compounds that related disease targets. We reviewed the progress, challenges and prospects of the SBDD method. Preparation of protein structures, identification of binding sites, preparation of ligand libraries, molecular docking and sequencing are key steps of SBDD.

In the past few decades, several software and network services supporting the SBDD method have been developed based on different scoring functions. With the help of these molecular docking programs, the binding active mode and binding free energy of small molecule ligands in the receptor binding site can be predicted and estimated, and then lead compounds can be quickly selected from low-cost ligand databases for comparison. Experimental techniques like HTS. Although the SBDD method has been used for drug research and design, challenges remain. Due to the limitations of each scoring function, the current

scoring function cannot perfectly estimate the accurate combined free energy.

Conflict of Interest

The authors declare that they have no conflict of interest.

REFERENCES

Marc A, Marti R, Ashley C, Stuart, Andras F, Roberto S, Francisco M, Andrej S. Comparative Protein structure modeling of genes and Genomes. Annu. Rev. Biophys. Biomol. Struct. 2000;29:291-325.

Gowri S. A study on various computer-aided drug design methodologies. J. Chem. Pharm. Res. 2015; 7(3):262-267.

Sam ZG, Xiaoqin Z. Challenges, Applications, and Recent Advances of Protein-Ligand Docking in Structure-Based Drug Design. Molecules. 2014;19:10150-10176.

Gerhard W, Thierry L. Ligand Scout: 3-D Pharmacophores Derived from Protein-Bound Ligands and Their Use as Virtual Screening Filters. J. Chem. Inf. Model. 2005; 45:160-169.

Sundarapandian T, Shalini J, Yuno L, Songmi K, and Keun W L. Dynamic Structure-Based Pharmacophore Model Development: A New and Effective Addition in the Histone Deacetylase 8 (HDAC8) Inhibitor Discovery. Int. J. Mol. Sci. 2011; 12: 9440-9462.

Gerhard K. Virtual ligand screening: strategies, perspectives and limitations, Drug Discovery Today. 2006;11: 580-594.

Doug B, Jonathan G, Andrew S, and Peter S. Identification of Common Functional Configurations Among Molecules. J. Chem. Inf. Comput. Sci. 1996; 36: 563-571.

Sachin SN, Mariam SD. Pharmacophore Refinement and 3D-QSAR Studies of Histamine H3 Antagonists. QSAR Comb. Sci. 2007; 26: 744-753.

Adriano DA, Livia BS, Donald J. Structure-Based Drug Design Strategies in Medicinal Chemistry, Current Topics in Medicinal Chemistry. 2009; 9:771-790.

Claudio NC, Andrew JW. Ligand Docking and Structure-based Virtual Screening in Drug Discovery. Current Topics in Medicinal Chemistry. 2007; 7: 1006-1014.

Najib ME, Peter JM, Gaelle B, Matthew B,Jonathan C, Gautam A, et al. Comparative Genomics of Trypanosomatid Parasitic Protozoa. Science . 2005;309:404-409.

Maria B, Bilal A, Sangdun C. A Structure-Based Drug Discovery Paradigm. Int. J. Mol. Sci. 2019;20: 2783-2800.

Montserrat C, and Jordi M, A chemogenomic approach to drug discovery: focus on cardiovascular. Drug Discovery Today. 2009;14:479-485.

Fernan A, Bissan A, Martin A, Matthew B, Frederick SB, Robert KC, et al. Genomic-scale prioritization of drug Targets: the TDR Targets database. Nature Reviews Drug Discovery. 2008;7:900-907.

Journal of Chemistry Letters
Published by the Chemical Society of Japan
Online ISSN 03667022, 1348071
Vol.50 No.5 (June 2021)

Rafael VC, Glaucius O, Adriano DA. Virtual Screening and Its Integration with Modern Drug Design Technologies. Current Medicinal Chemistry. 2008;15:37-46.

Douglas BK, Helene D, John RF, Jurgen B. Docking And Scoring In Virtual Screening For Drug Discovery: Methods And Applications. Nature Reviews Drug Discovery. 2004;3:935-949.

Antoine D, Marie C B, Vivienne BG, Patricia MP, Diana M, Ioannis X, et al. Drug Design Workshop: A Web-Based Educational Tool To Introduce Computer-Aided Drug Design to the General Public. J. Chem. Educ. 2017;94:335-344.

Solomon N, Alan H. Innovative lead discovery strategies for tropical diseases, Nature Reviews Drug Discovery. 2006;5:941-955.

Harren J, Glyn W, David CR, Christopher WM. The rule of three for fragment-based drug discovery: where are we now?. Nature Reviews Drug Discovery. 2013;12: 5-7.

David SW, Improving Early Drug Discovery through ADME Modelling.Drugs R D. 2007;8:349-362.

Tiejun C, Qingliang L, Zhigang Z, Yanli W, Stephen HB. Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. J The AAPS. 2012;14:133-141.

Suzanne BS, Philip JH, Robert PM, Stephen WF, Discovering Hig h-Affini ty Ligands for Proteins:SAR by NMR. Science. 1996;274:1531-1534.

David SW, Improving Early Drug Discovery through ADME Modelling,, Drugs R D 2007; 8 (6): 349-362

Lei p, christopherL G, Fernando AP, Claudio FG, Graciela Ll. Identification of the Tolfenamic Acid Binding Pocket in Prbp from Liberibacter asiaticus. Frontiers in Microbiology. 2017;8:1-12.

Lulia ML, Maria K, Alexandru CL. Proteomics for Drug Discovery: Methods and Protocols, Protocols, Methods in Molecular Biology (pp255-265).Springer New York, 2017.

Jianling W, Suzanne S. Recent Advances in Physicochemical and ADMET Profiling in Drug Discovery. Chemistry & Biodiversity. 2009;6:1887-1899. Gregor J,Marko A, Tihomir T, Comparative evaluation of several docking tools for docking small molecule ligands to DC-SIGN. J Mol Model. 2015;21:164-176.

Emanuele P, Paul SC. Conformational Analysis of Drug-Like Molecules Bound to Proteins: An Extensive Study of Ligand Reorganization upon Binding. J. Med. Chem. 2004;47:2499-2510.

Mohamed AA, Khaled ME, Bin W, John JB, Mohamed FA, Mohamed HH, et al. Computationally Assisted Assignment of Kahalalide Y

Ekins S, Mestres J, Testa B. In silico

pharmacology for drug discovery: applications to targets and beyond. British Journal of Pharmacology. 2007;152:21-37.

Jie L, Renxiao W. Classification of Current Scoring Functions. J. Chem. Inf. Model. 2015;55:475-482.

Pui MH, Shang L, Chi TV, Hisa HL, Yiu WK, Simon MY. Recent advances in structure-based drug design and virtual screening of VEGFR tyrosine kinase inhibitors. Methods. 2015;71:85–9.

Configuration Using an NMR-Constrained Conformational Search, J. Nat. Prod. 2013;76:178−185.