International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 07, July 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

Computational Approaches in Drug Discovery: An Overview

A.Jainul Fathima[1], G.Muruga Boopathi[2], P.Selvam[3]

[1]Research Scholar, [2]Associate Professor, [3]Research Faculty

[1, 2]Department of Computer science and engineering

[3] International Research Centre

[1,2,3]Kalasalingam University, Krishnankoil, (India)

ABSTRACT

Drug discovery and development process involves a series of events that include target identification and validation, lead identification and optimization, pre-clinical pharmacology and toxicology. Computer-aided drug discovery (CADD) tools can be used to automate and speed up these process and to reduce the research and development cost. Today CADD has become an essential tool in drug development. The bioinformatics research has made available a significant amount of data sources like biological structures, ligand databases, and various computational tools that can be used in various phases of the drug discovery and development pipeline. This paper gives an overview of computational methods used in different stages of drug discovery. In this review, both structure-based and ligand-based drug discovery methods are discussed. Developments in virtual high-throughput screening, prediction of protein–ligand interaction using docking tools are reviewed.

Keywords: ADME, Docking, High-throughput screening, lead optimization.

I. INTRODUCTION

The bioinformatics research worldwide is keenly interested in automating and speeding the drug discovery and development process. Bringing a drug to the market is a long termexpensive process. It is estimated that the cost associated with developing and bringing a drug to the market has increased nearly 150% in the last decade. The cost is now estimated to be a confounding $2.6 billion dollars. The probability ofa failure in the drug discovery and development pipeline is highand 90% of the drugs entering clinical trials fail to get FDAapproval and reach the consumer market. Approximately 75%of the cost is due to failures that happen along the drugdiscovery and design pipeline [1].In recent years faster high-throughput screening (HTS) experiments, which can evaluate thousands of molecules with automation tools, human labor associated with screening of compounds is no longer necessary.However, HTS is still expensive and requires a lot of resourcesof targets and ligands. These resources are frequently not available in academic settings.

Additionally, many pharmaceuticalcompanies are now looking for ways that can avoid screening of ligands that have no possibility of showing success. Therefore,computer-aided drug discovery (CADD) tools are getting a lotof attention in the pharmaceutical industry. CADD technologies are powerful tools that can reduce thenumber of ligands that need to be screened in experimentalanalyses. The most popular balancing approach to HTS isthe use of virtual (i.e., in silico) HTS.[2] Computer-aided drugdiscovery and design not only reduces the costs associated withdrug discovery by ensuring that best possible lead compoundenters animal studies, but it may

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

also reduce the time it takes fora drug to reachmarket.. CADD tools identifylead drug molecules for testing, can predict effectiveness andpossible side effects, and assist in improving bioavailability ofpossible drug molecules. Many studies show how CADD can influence the development of novel drugs [3]

The first section gives an introduction to computer aided drug design and development with a brief explanation about structure and ligand based methods and various computer software tools available for research. The next section deals with literature review. The third section focuses on major steps involved in docking process and summarizes about various docking tools. The last section gives an idea of docking can be applied to various application area for lead optimization.

II. COMPUTER AIDED DRUG DISCOVERY

The power of CADD is reviewed in various case studies. This section briefly discusses the various research for drug development using computational methods

This paper [4] researchers used CADD tools to screen for inhibitors of tyrosine phosphate, an enzyme involved in diabetes. While applying virtual screening approach yielded 365 compounds out of which 127 shows high inhibition. Followed by this the same group performed HTS against the same target which gives 81 compounds with inhibition. Comparative study in this paper effectively display the power of CADD .In another research [5], dorzolamine was discovered using CADD tools have passed through clinical trials and become a carbonic anhydraze inhibitor. Captopril was approved as an antihypertensive drug [6]. Three therapeutics saquinar,ritonavir and indinavir were approved for the treatment of human immuno deficiency virus(HIV)[7].The group at Biogen idec identified 87 hits, the best hit being identical in structure to lead compound discovered through the traditional HTS approach [8].

CADD is capable of increasing the hit rate of novel drug compounds because it uses much targeted search than traditional approach. This approach explains the molecular basis of drug activity and also to predict possible derivatives that would improve the activity. The three major purpose of CADD includes filter large compound libraries to predict active compounds, Optimize lead compound to increase its affinity and to calculate drug-likeness using ADMET properties [9],to design a novel compound from lead compound.

2.1 Structure Based and Ligand Based Cadd

CADD can be classified into structure based and ligand based methods [10]. The following table gives the difference between structure based and ligand based approach

Table 1: Difference Between Two General Categories: Structure Based and Ligand Based

STRUCTURE BASED APPROACH	LIGAND BASED APPROACH
1. Knowledge of protein target structure to calculate interaction energy	Knowledge of ligand structure which helps in construction of predictive, QSAR models
2. Preferred when high resolution structural data of protein target is available	Preferred when no or little structural information is available
3. The goal is to design compounds that bind tightly to the target	The goal is to virtual screening of ligand database
4. Techniques include Molecular Dynamics,De nova design and Pharmacophore modeling	Techniques include QSAR, High throughput screening, and Pharmacophore modeling

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

The successful CADD application will allow to identify multiple lead compounds, Identification of lead is often followed by several steps of lead optimization and lead identification using CADD approach.

2.2 Computational Software used in Drug Design

The various stages involved in computational drug design includes Homology modelling, Active binding site prediction, Molecular graphic system, Virtual screening and docking. There are large number of tools available for researchers [11]. The following table gives the various software available for every stages in drug discovery

TABLE 2: SOFTWARE TOOLS AVAILABLE FOR CADD

COMPUTATIONAL METHODS	SOFTWARE AVAILABLE FOR STUDY
Homology modelling [12]	Insight, Prime, LOOK, Sybyl, -DS Modeller, Prime, LOOK, ICM, Sybyl, MODELLER, MOE, SWISS-MODEL, Raptor X, LOMETS, Phyre, I-Tasser
Active Binding Site prediction [13]	CASTp, POOL, PASS, Pocket-Finder, 3DLigandSite, LIGSITE, meta Pocket, FINDSITE, Site-hound
Molecular Graphic System [14]	Avogadro, Chemlab, Athena, Maestro, Jmol, PyMOL, UCSF chimera, VMD, Vimol, Webmol, Zeus
Virtual screening databases [15]	PubChem, MMsINC, ZINC, ZincPharmer, 4SC discovery, therapeutic target database, drug Bank, ChemSpider, ChEMBL
Docking software [16]	PyRx, Autodock Vina, Dock Blaster, Vis3d, Schrodinger, GOLD, Libdock, FlexX, Glide, Fred, ICM
Molecular dynamic simulation software [17]	Gromacs, Amber, CHARM, Gromos, ADF, Desmond, NWChem

The next section, explains the major steps involved in molecular docking for predicting the protein-ligand interaction

III. PREDICTION OF BINDING POTENTIAL OF PROTEIN-LIGAND INTERACTION: MOLECULAR DOCKING

Molecular Docking is the process of predicting the binding potential or the intermolecular interaction between 2 molecules computationally. In this process, the large molecule is the protein receptor. The micro molecule is the Ligand molecule which can be acted as an inhibitor. So, the Docking process involves the following steps

3.1 Protein preparation

Three-dimensional structure of the Protein should be retrieved from Protein data bank (PDB) [18]. PDB has more than 81,000 protein structures. When protein structure is not available the structure is predicted through homology modeling and the structure is optimized. The retrieved structure should be pre-processed. This should admit removal of the water molecules from the cavity, stabilizing the charges, filling the missing residues, generation the side chains etc. according to the parameters available.

3.2 Prediction of Active binding site and characterization

Interaction of protein-ligand is a prerequisite for drug activity. It is possible only when the high-affinity binding sites are identified. The protein molecule has lots of active sites. Various computational tools like POCKET, SURFNET, and Q-SITEFINDER are available for binding site prediction [19].

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

3.3 Ligand Preparation

Ligands are retrieved from ligand databases such as ZINC, Pub Chem, DrugBank, and Chem DB or can be sketched using chem sketch tool. Drug-likeness is evaluated using LIPINSKY'S RULE OF 5 to eliminate nondrug like, unstable or unfavorable compounds. The active drug should satisfy the ADMET property [20]. 1. Should have more than five hydrogen bond donors, 2. Should have less than ten hydrogen bond acceptor, 3. Molecular mass less than 500 Da, 4. The logp value should not over 5. Molar refractivity should be between 40-130. If two or more conditions are violated, poor absorption can be expected. The ligands are represented as SMILES (Simplified Molecular Input Line System) format.

3.4 Molecular Docking

Ligand is docked against the protein and interaction are analyzed. Docking can be separated into search algorithm and Scoring function. The following table 3 summarizes the various docking software tools with their benefits and drawbacks

3.4.1 Search algorithm for Protein-Ligand Docking

Docking methods can be classified as rigid-body docking and flexible docking application depending on the degree to which they consider ligand and protein flexibility during the docking process. The various algorithm applied for docking include Monte Carlo, Fragmentbased, Genetic algorithms, Systematic searches

3.4.2 Scoring Function for evaluation protein-ligand complexes

Scoring function attempts to predict target-ligand binding affinities for hit to lead and lead to drug optimization. This can be grouped into four types: force-field, empirical, Knowledge-based and consensus scoring function

Table3: Characteristics For Widely Used Protein-Ligand Docking Tools

S.No	Docking Tool	Platform supported	Docking Approach	Scoring function	Benefits	Drawbacks	Company Designed
1	Auto Dock	Unix,Mac, Linux	Lamarckin genetic algorithm, Simuated Annealing	Force-Field Methods	Small Binding sites	Not suitable for highly polar ligands	The Scripps Research Institute
2	Glide	Unix, LinuxIBM	Monte Carlo Sampling	Glide score, Glide comp	Flexible Ligands	Slow speed	Schrodinger Inc.
3	GOLD	Linux,IBM ,Sun,Windows	genetic algorithm	Gold Score,Chem Score	Small hydrophobic ligands	Ranking ligands in large cavities	Cambridge Crystallographic Data Centre
4	FRED	Unix,Linux ,Windows, Mac	Gaussian Shape fitting	Screen Score,Gaussian shape score	Large binding sites, High speed	Small polar buried ligands	Open Eye scientific Software
5	SLIDE	Unix,Linux	Induced	Multistage	Slide	Sensitivity to	Protein

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

		,Windows, Mac	fit Docking	indexing	chain flexibility	input coordinates	Structural Analysis Laboratory
6	QXP	Unix,Mac, Linux	Monte Carlo perturbation	template fitting and building pseudo-receptors	Reliable,Easy to use	Sensitivity to input coordinates	Boston De Novo Design

3.5 Molecular docking Applications

Docking provides a various tools for drug design and analysis. Various application of docking includes Virtual Screening, Lead Optimization, Prediction of biological activity, Blind docking, Protein-Protein interation,Searching of lead structures for protein targets, Protein engineering. Docking can also be used to predict and optimize drug metabolism and pharmacokinetics and toxicity properties

3.5.1 Molecular docking of Protein-ligand using Auto Dock: Screen Shot

Auto dock is the most widely used docking tool used for binding protein and ligand. The following screenshot shows the working environment of Autodock and output for protein ligand interaction

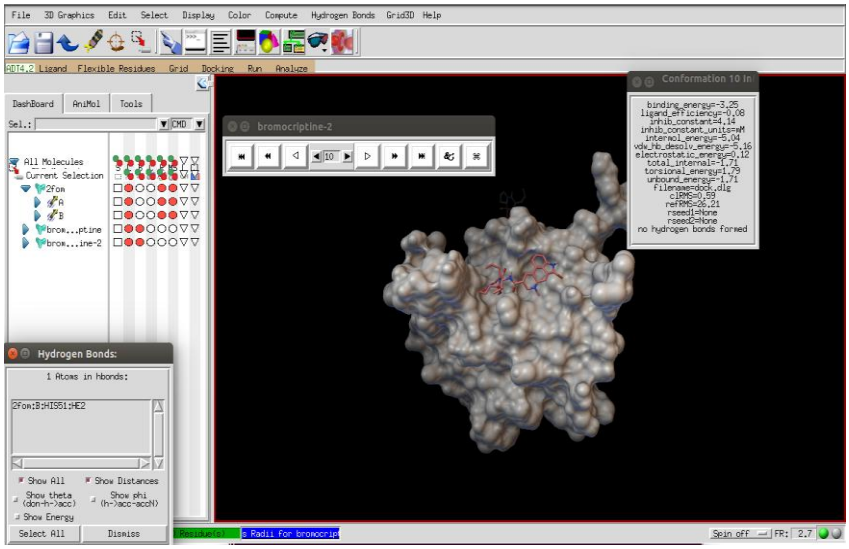

Fig 1 : Screenshot of Docking using Autodock tool

VI. CONCLUSION

The wide variety of computational tools used in drug discovery and development suggests that there are no basically superior techniques. The performance of methods varies greatly with the target protein, available data, and available resources [21]. With the significant increase in the number of drug targets, computational methods such as protein structure prediction methods, virtual high-throughput screening, and docking methods have been used to speed up the drug discovery process, and are regularly used in the pharmaceutical industry. These approaches are conventional and are now a valuable part of the drug discovery pipeline and have shown great promise and success. It is inexpensive and faster to computationally predict and filter large molecular databases and to select the most likely molecules to be optimized. Only the molecules predicted to have the desired biological activity will be screened in vitro. This saves money and time because the risk of binding resources on

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

possibly unsuccessful compounds that would otherwise be tested in vitro is reduced. The field of CADD is continuously developing with improvements being made in each and every area. Some of the focus areas are scoring functions,search algorithms for molecular docking and virtual screening, optimization of hits, and assessment of ADME properties of possible drug candidates. With the current successes, there is a promising future for computational methods to aid in the discovery of many more drugs in the future.

REFERENCES

[1] Tollman, P. A Revolution in R&D: How genomics and genetics aretransforming the biopharmaceutical industry,2001.

[2] Yang, L.; Wang, W.; Sun, Q.; Xu, F.; Niu, Y.; Wang, C.; Liang, L.;Xu, P. Bioorg. Med. Chem. Lett.2016,26,2801–2805.doi:10.1016/j.bmcl.2016.04.067.

[3] Gregory Sliwoski, Sandeep kumar Kothiwale, Jens Meiler and Edward W. Lowe Eric L. Barker, Computational Methods in Drug Discovery, Pharmaco logical Reviews January 2014, 66 (1) 334-395; DOI: https://doi.org/10.1124/pr.112.007336.

[4] Doman TN, McGovern SL, Witherbee BJ, Kasten TP, Kurumbail R, Stallings WC, Connolly DT, and Shoichet BK(2002) Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. J Med Chem 45:2213–2221

[5] Vijaya krishnan R(2009) Structure-based drug design and modern medicine. J Postgrad Med 55:301–304.

[6] Talele TT, Khedkar SA, and Rigby AC(2010) Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. Curr Top Med Chem 10:127–141.

[7] Jorgensen, W. L. Science 2004, 303, 1813–1818. doi:10.1126/science.1096361

[8] Craig, J. C.; Duncan, I. B.; Hockley, D.; Grief, C.; Roberts, N. A.; Mills, J. S. Antiviral Res. 1991, 16, 295–305. doi:10.1016/0166-3542(91)90045-S

[9] Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia, M. A. J. Am. Chem. Soc. 1995, 117, 1181–1182. doi:10.1021/ja00108a056

[10] Anderson, A. C. Cell Chem. Biol. 2003, 10, 787–797. doi:10.1016/j.chembiol.2003.09.002

[11] Yildirim, M. A.; Goh, K.-I.; Cusick, M. E.; Barabasi, A.-L.; Vidal, M. Nat. Biotechnol. 2007, 25, 1119–1126. doi:10.1038/nbt1338

[12] Vyas, V. K.; Ukawala, R. D.; Ghate, M.; Chintha, C. Indian. J. Pharm. Sci. 2012 doi:10.4103/0250-474X.102537

[13] Beilstein J. Org. Chem. 2016, 12, 2694–2718. 2712 21. Carpenter, E. P.; [12]. Beis, K.; Cameron, A. D.; Iwata, S. Curr. Opin. Struct. Biol. 2008, 18, 581–586. doi:10.1016/j.sbi.2008.07.001

[14] Trott, O.; Olson, A. J. J. Comput. Chem. 2010, 31, 455–461. doi:10.1002/jcc.21334

[15] Brooijmans, N.; Kuntz, I. D. Annu. Rev. Biophys. Biomol. Struct. 2003, 32, 335–373. doi:10.1146/annurev.biophys.32.110601.142532

[16] McGann, M. J. Chem. Inf. Model. 2011, 51, 578–596. doi:10.1021/ci100436p

[17] Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. J. Comput. Chem. 1998, 19, 1639–1662. doi:10.1002/(SICI)1096-987X(19981115)19:143.0 .CO;2-B

[18] Chaudhary KK, Mishra N (2016) A Review on Molecular Docking: Novel Tool for Drug Discovery. JSM Chem 4(3): 1029.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

[19] Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Proteins: Struct., Funct., Bioinf. 2003, 52, 609–623. doi:10.1002/prot.10465

[20] Kramer, B.; Rarey, M.; Lengauer, T. Proteins: Struct.,Funct., Bioinf. 1999, 37, 228–241. doi:10.1002/(SICI)1097-0134(19991101)37:23.0.C O;2-8

[21] Computational methods in drug discovery. Available from: https://www.researchgate.net/publication/311631714_Computational_methods_in_drug_discovery [accessed Jul 2, 2017].