USE OF COMPUTER IN DRUG DESIGN AND DRUG DISCOVERY: A REVIEW

Md. Mofizur Rahman[1*], Md. Rezaul Karim[2], Md. Qamrul Ahsan[3], Abul Bashar Ripon Khalipha[1], Mohammed Raihan Chowdhury[4] and Md Saifuzzaman[5]

[1]Department of Pharmacy, Bangladesh University, Dhaka, Bangladesh

[2]Department of Pharmacy, Manarat International University, Dhaka, Bangladesh

[3]Department of Pharmacy, Southern University, Bangladesh

[4]Department of Pharmacy, University of Asia pacific, Dhaka, Bangladesh

[5]Department of Pharmacy, Khulna University, Dhaka, Bangladesh

Abstract

Drug design through computer, a recent, very effective technique in modern arena. Now a days Computer Aided Drug Design (CADD) technologies are used in nanotechnology, molecular biology, biochemistry etc. The main benefit of the CADD is cost effective in research and development of drugs. There are wide ranges of software are used in CADD, Grid computing, window based general PBPK/PD modeling software, PKUDDS for structure based drug design,APIS, JAVA, Perl and Python, CADD as well as software including software libraries. There are different techniques used in CADD visualization, homology, molecular dynamic, energy minimization molecular docking, QSAR etc. Computer aided drug design is applicable in Cancer disease, transportation of drug to specific site in body, data collections and storages of organics and biologicals. Conformational properties and energetics of small molecules and DNA cleavage, molecular diagnostics based on fluorescences are focusing using this technique.

Corresponding author:

[*]Md. Mofizur Rahman

Department of Pharmacy, Bangladesh University, Dhaka-1207, Bangladesh,

Tel: +8801911605139, E-mail: rmfi02@yahoo.com

Introduction

Computer Aided Drug Design (CADD) and Delivery Systems offers an in-depth discussion of the computer-assisted techniques used to discover, design, and optimize new, effective, and safe drugs. Recent technological developments in biochemistry, biomedical science, and nanotechnology have made computer-aided drug design and delivery systems possible on a molecular basis. This in-depth treatise covers this pioneering advances.[1]

The objective of drug design is to find a chemical compound that can fit to a specific cavity on a protein target both geometrically and chemically. [2] It is generally recognized that drug discovery and development are very time and resources consuming processes. There is an ever growing effort to apply computational power to the combined chemical and biological space in order to streamline drug discovery, design, development and optimization. In biomedical arena, computer-aided or in silico design is being utilized to expedite and facilitate hit identification, hit-to-lead selection, optimize the absorption, distribution, metabolism, excretion and toxicity profile and avoid safety issues. The development of any potential drug begins with years of scientific study to determine the biochemistry behind a disease, for which pharmaceutical intervention is possible. The result is the determination of specific receptors (targets). In the post genomic era, computer-aided drug design (CADD) has considerably extended its range of applications, spanning almost all stages in the drug discovery pipeline, from target identification to lead discovery, from lead optimization to preclinical or clinical trials [3]. One method that was quickly adopted by industry was the use of combinatorial chemistry and HTS. In HTS, large libraries of compounds are screened against drug targets to identify lead compounds that can modulate a particular outcome. However, setting up a combinatorial chemistry program and HTS is costly and not able to address the specific needs of many biological (drug target) systems. Compounds identified in such screenings are not always amenable to further medicinal chemistry development, with poor ADME (absorption, distribution metabolism and elimination) properties. Although these methods have increased the rate at which lead compounds can be identified, there has not been a commensurate increase in the rate of introduction of

new chemical entities (NCE) into the world drug market. As an attractive alternative, in silico methods show promise in identifying new lead compounds faster and at a fraction of the cost of combinatorial approaches and HTS. The addition of computer aided drug design technologies to the R&D approaches of a company, could lead to a reduction in the cost of drug design and development by up to 50%. [4]

Drug discovery and computer aided drug design

It is estimated that a typical drug discovery cycle, from lead identification through to clinical trials, can take 14 years with cost of 800 million US dollars. In the early 1990s, rapid developments in the fields of combinatorial chemistry and high-throughput screening technologies have created an environment for expediting the discovery process by enabling huge libraries of compounds to be synthesized and screened in short periods of time. However, these concerted efforts not only failed to increase the number of successfully launched new molecular entities, but seemingly aggravated the situation. Hit rates are often low and many of these identified hits fail to be further optimized into actual leads and preclinical. Among the late stage failures, 40–60% was reportedly due to absorption, distribution, metabolism, excretion and toxicity (ADME/Tox) deficiencies. Collectively, these issues underscore the need to develop alternative strategies that can help remove unsuitable compounds before the exhaustion of significant amount of resources.[5] The more recent foundations of CADD were established in the early 1970s with the use of structural biology to modify the biological activity of insulin [6] and to guide the synthesis of human haemoglobin ligands.[7] At that time, X-ray crystallography was expensive and time-consuming, rendering it infeasible for large-scale screening in industrial laboratories.[8] Over the years, new technologies such as comparative modeling based on natural structural homologues have emerged and began to be exploited in lead design.[9] These, together with advances in combinatorial chemistry, high-throughput screening technologies and computational infrastructures, have rapidly bridged the gap between theoretical modeling and medicinal chemistry. Numerous successes of designed drugs were reported, including Dorzolamide for the treatment of cystoid macular edema[10], Zanamivir for therapeutic or prophylactic treatment of influenza infection[11], Sildenafil for

the treatment of male erectile dysfunction [12], and Amprenavir for the treatment of HIV infection[13].

Software used in CADD

Among the software programs are applications programmed in Grid computing, window based general PBPK/PD modeling software, PKUDDS for structure based drug design, APIS, JAVA, Perl and Python, CADD as well as software including software libraries. Showed in figure 1 and Figure 2.
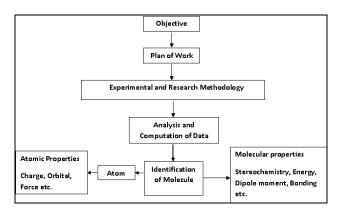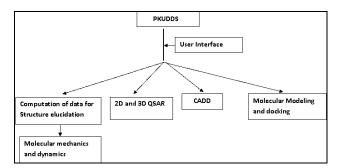

Figure: 1 An overview of the GaussDal database structure for storing molecular properties[14]
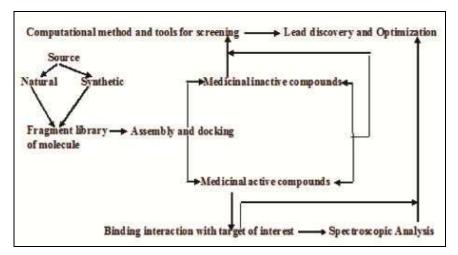

Figure: 1 A flow chart for Peking University Drug Design System (PKUDDS) [14]

Figure: Drug discovery

Visualization

Rasmol, VMD, Molscript, Raster 3D are tools used to optimize ligand or chemical compound and target molecule. For the depiction and exploration of biological macromolecule structures,

Rasmol, computer program written for molecular graphics is used.

Homology and homology modeling programs

Most drug targets are proteins so it is important to know their 3D structures in detail. It estimated that the human body has five lacks to one million proteins, but the 3D structure is known for only a small fraction of these. Homology modeling is used to predict the 3D structure of proteins[15]. Homology modeling is nothing but similarity searching for drug analogs. It starts with promising drug molecule.Molecular modeling[16] is a science of representing molecular structures numerically and simulating their behaviors with the equations of quantum and classical physics. There are two computational tools for similarity searching and sequence alignment such as BLAST, FASTA and for multiple sequence alignments ClastalW ClastalX.

There are two Homology Modeling Programs.[17] They are Swiss Model, Modeller Swiss Model makes it quick and easy to submit a target sequence and get back an automatically generate a comparative model, provided an empirical structure with >30%

sequence identity exist to use as a template. Modeller is used for homology or comparative modeling of protein 3D structure. The user provides an alignment of a sequence to be modeled with known related structures and modeler automatically calculates a model containing all non hydrogen atoms.

Molecular dynamics

Molecular dynamics is a study of movement of molecule. Every molecule has its own frequency of vibration. It can oscillate position one to two through zero, where the molecule has high potential energy at one and two position and least at zero position.[18]

Energy minimization

It is also called energy optimization or geometry optimization; it is used to compute the equilibrium configuration of molecules and solids. By this technique we can only obtain a final state of system that corresponds to minimum of potential energy. In Energy minimization one can obtain a molecule with least energy state i.e. zero energy state. In this state molecule get equilibrium configuration. Energy minimization tools are GAMESS Ghemical PS13 TINKER Ghemical can be used or PS13 For quantum mechanical calculations. If proteins are used, a program such as PyMol can be used to identify ligand binding pockets, together with the DeepView PDB viewer to investigate the amino acid sequences of the protein. To transfer files between programs, Open Babel might be useful or even required to interconvert the file formats.[19]

Docking

In the field of molecular modeling docking is a method which predicts the preferred orientation of one molecule to second, when bound to each other to form a stable complex. Docking represents ligand binding to its receptor or target protein. Docking is used to identify and optimize drug candidates by examine & modeling molecular interactions between ligand and target macromolecules. In the docking multiple ligand conformations and orientations are generated & the most appropriate ones are selected. There are several docking tools are presently available they are ArgusDock DOCK FRED eHITS AutoDock FTDock. Scoring methods are used torank the affinity of

ligands to bind to the active site of a receptor. In virtual high throughput screeningcompounds are docked into an active site and then scored to determine which once morelikely to bind tightly to the target macromolecule.[20]

QSAR

A descriptor is a molecular property that QSAR can calculate. QSAR provides a wide variety of descriptors that you can use in determining new QSAR relationships. There is a limited number of datasets and little information regarding the training and validation used by previous researchers. Tetko et al. suggested the use of SMILES or .sdf files on a website to promote the calculation of additional parameters by other drug discovery scientists. The self-organizing molecular field analysis (SoMFA) test set, which represents the steroid set used to construct the first comparative molecular field analysis (CoMFA), can be downloaded from the Richards group's web site. This information facilitates a more-rapid evaluation of the SoMFA program.[21]

Application of computer in drug design

Anticancer agent

The sequencing of the human genome represents one of the major scientific endeavours of this century. A major aspect of the utilization of this information will be the provision of small molecules which will recognize selected sequences, perhaps with the goal of switching off particular genes as in cancer chemotherapy. For some time antibiotics such as netropsin have been known to bind preferentially to sequences rich in A-T pairs. A variant based on this research has been to try to design a bioreductive ligand based upon netropsin. The idea of bioreductive anti-cancer agents *statts* with the fact that tumours receive less blood and hence less oxygen than normal tissue. Thus it becomes possible, at least in principle, to contemplate having a ligand which can exist in two forms, oxidized and reduced, and if the redox potential is appropriate to be in the oxidized form in normal tissue but reduced in tumours. If only the reduced form will bind to the macromolecular target and cause cell death, then differentiation in action between cells which it is desirable to desttoy and normal cells is achievable, with concomitant reduction in side-effects. A second starting point for sequence selective

ligands is an organometallic molecule with chiral properties. The propeller- like ruthenium tris-phenanthroline complexes do show differential binding between A-T and G-C sequences and moreover may exhibit a preference for purine 3', *5′* pyrimidine sites in DNA.[22]

Target Enzyme

If an enzyme structure is known then designing inhibitors which will block activity in the test-tube should be a relatively straightforward problem. More spice to such a challenge is added if we at the same time attempt to make the ligand bioreductive as outlined above [23]. The published work has taken dihydrofolate reductase as the target enzyme, but current activity is being focussed on thymidylate synthetase. The binding free energy of the inhibitor to the enzyme is a crucial quantity: strong binding is essential.

Drug Transport

Sceptics quite rightly point out that designing an enzyme inhibitor which will work in the test-tube is one thing; getting a compound which will work in a cell is another. Transport across the biological membrane is essential. Compounds must be soluble enough in the lipid to get into the membrane, but not so soluble that they remain there. Within the pharmaceutical industry the partition coefficient between water and n-octanol is used as a guide to membrane transport. The free energy perturbation technique just described *can* also be adapted to compute partition coefficients.[24]

More excitingly, however, it is becoming possible to model biological membranes. Starting with crystal structures of membranes involving DMPC (1,2-dimyristoyl-sn-glycero-3-phosphoryclh oline) a highly realistic simulation is possible, involving a hydrated lipid bilayer. After very long molecular dynamics simulations the resulting membrane model is in agreement with all the available experimental data; lead p u p separation; order parameters and diffusion coefficients 'Ibis model *can* be used as the 'solvent' in calculations of partition coefficients which should be considerably more realistic than experimental values in n-octanol. Furthermore it will be W b l e to introduce cholesterol and protein into the model membrane to produce a truer simulation of how a given drug is transported into a cell.[25]

Structure determination of protein

One of the major contempomy scientific aims is to use the 'abundance of gene and hence protein sequences to predict the three-dimensional structure of proteins: going from primary to tertiary structure. Were this routinely pwsible the choice of drug target where the architecture of the binding site is known would increase from a handful of cases to many thousands. The currently favoured and only successful methods are all based upon finding similarities and homologies between the protein of known sequence but unknown topology and known structures from three-dimensional databases.

Generally sequences are compared with scoring matrices being used to ascertain just how similar a short length of polypeptide in the unknown is in comparison with a known case. One successful prediction that of the important small protein big endothelin, was made using not the identities of amino acids in the sequence, but their properties, notably their hydrophobicities. The property profile is smoother than an identity specification where each amino acid can be one of twenty.[26]

Where the similarity is low the use of colour graphics to permit the human eye to detect similarities has many advantages although it is inevitably subjective. This approach, using the computer program CAMELEON, has recently been used to predict the structure of the interleukin-4 receptor . It is believed that the folding topology of the beta sheets of IL4R is the same as that seen in the crystal structure of 0 4 , despite sequence identity being low. Each domain of the IL4R monomer was aligned with 0 4 using single residue hydropathy properties. Loops were added from a database of immunoglobulins *so* as to connect the sheets; side-chains were added using a side-chain rotamer library and the unsolvated structure energy-minimized using molecular dynamics. The whole structure was thus placed in an 8A shell of water and unconstrained molecular dynamics canied out for *60* ps. Finally the whole structure was minimized. Assuming that the IL4 receptor acts in the same way as growth hormone receptor as a h e r , one molecule of IL4 was docked to a pair of receptor proteins. The docking shows which portion of IL4 binds to the receptor: in this prediction notably the D helix. On the basis of this it should be possible to design mimetics of the crucial parts of the IL4 D helix which would interfere with the biochemical consequence of the cytokine binding to its receptor, leading to antagonists with potential medicinal applications.[25]

Biochemical Transformation

Where no knowledge about the macromolecular target in atomic detail exists, then it is still possible to utilize computer-aided design techniques. A popular idealized approach would be to compute the energy profile of a biochemical transformation which it would be desirable to inhibit; locate the transition state or intermediate and then create a stable mimic of these unstable transients recognised by the enzyme responsible for catalysing the reaction and would hence act as an inhibitor. Such a mimic should be Only two logical steps are necessary: find the transient structure and secondly design a stable mimic. The former task is probably best achieved by using a combination of quantum and molecular mechanics. A recent review suggests that the combined potential method used by Bash et a1 for the triosephosphate isomerase reaction is probably the technique likely to be followed in the future. The second stage of the process invokes the introduction of the idea of molecular similarity, a quantitative measure of just how similar one molecule is to another. Perhaps the most important aspect of similarity is similarity of shape and secondly similarity of molecular electrostatic potential, both of which can be represented by gaussian functions which introduce major computational gabiS in the calculation of similarity indices, of which several different types may be defined. [27]

Molecular similarity

Much more striking has been the achievement of similarity measures in structure-activity relationships and in quantitative structure-activity relationships. *Good* et a1[28] considered the series of steroids for which binding affinity data are available and which was the set studied in the earliest comparative molecular field three-dimensional structure-activity work. The cross validated correlation coefficients obtained from the statistical analysis compare well with those obtained using the more commonly used matrices of similarities at grid points in the space surrounding the molecules which of coume demand massive matrices of perhaps thousands of points. In addition there is no need for arbitrariness about the extent of molecular 'surface' or the size of the three-dimensional box into which the molecules have to be placed. Although in its infancy

molecular *similarity* matrices to seem to have a lot to offer in *QSAR* and in the optimisation of molecular structures for particular biological effects.

Molecular dissimilarity

'Ihe molecular dissimilarity between a pair of molecules *can* be defined as (1 - Similarity). Similarity has a range of values 0 to 1 with unity representing identity. The interest in dissimilarity is in the comparison of chiral forms of the same molecule. The dissimilarity *can* be used as a 'chirality coefficient', a number which gives a range of values of chirality rather than this being an all-or-none property. Currently there is a great deal of research into producing pure chiral forms of compounds for use as pharmaceutical agents: the more active form being termed the eutomer and the less active the distomer, with their ratio being the eudismic ratio. For an homologous series of compunds we have shown that their is a direct correlation between the eudismic ratio and the chirality coefficient. [25]

Current advanced in computer aided drug design[29]

Data accessibility is critical for the success of a drug discovery and development campaign. Huge amounts of organic molecules, biological sequences and related information have been accumulated in scientific literature and case reports. These data are collected and stored in a structured way in a number of databases showed below. Every year, hundreds of biological databases are described. At the same time, computational algorithms are actively developed to facilitate the design of combinatorial libraries. So Computer aided drug design focuses on these.

Some small molecule databases

PubChem http://pubchem.ncbi.nlm.nih.gov/

ACD http://www.mdli.com

ZINC http://zinc.docking.org/

LIGAND http://www.genome.jp/ligand/

DrugBank http://www.drugbank.ca/

ChemDB http://cdb.ics.uci.edu/

Some biological databases

Type Name URL

DNA sequences GenBank http://www.ncbi.nlm.nih.gov/Genbank/

DDBJ http://www.ddbj.nig.ac.jp/

EMBL http://www.embl-heidelberg.de/

Protein sequences Swiss-Prot http://www.expasy.ch/sprot/

PIR http://pir.georgetown.edu/

Protein structures PDB http://www.rcsb.org/pdb

Gene expression ArrayExpress http://www.ebi.ac.uk/microarray-as/ae/

GEO http://www.ncbi.nlm.nih.gov/geo/

CIBEX http://cibex.nig.ac.jp/index.jsp

2D gel electrophoresis SWISS-2DPAGE http://www.expasy.ch/ch2d/

GELBANK http://gelbank.anl.gov/

Mass spectrometry OPD http://bioinformatics.icmb.utexas.edu/OPD/

GPMDB http://www.thegpm.org/GPMDB/index.html

Metabolomics HMDB http://www.metabolomics.ca

MDL Metabolite database http://www.mdl.com/products/predictive/metabolite/index.jsp

METLIN http://metlin.scripps.edu/

Protein^protein interactions BIND www.bind.ca/

HPRD http://www.hprd.org/

IntAct www.ebi.ac.uk/intact/

Transcriptional regulation TRANSFAC http://www.biobase-

international.com/pages/index.php?id¼transfac

TRED rulai.cshl.edu/TRED/

Post-translational modifications dbPTM dbptm.mbc.nctu.edu.tw/

RESID www.ebi.ac.uk/RESID/

Biological pathways KEGG www.genome.jp/kegg/

BioCarta http://www.biocarta.com/

Ligand selectivity

The discovery of new molecular entities for drug intervention is a highly combinatorial science due to the diversity of protein targets, as well as huge variability of possible lead candidates. The theoretical number of natural proteins is approximately 250 000, while the number of real organic compounds with molecular weight <2000 Da is more than 1060. Due to the astronomically huge chemical space, the cost required for systematic studies can be extraordinarily high. Computational tools are increasingly used as a cost-effective way for the selection, modeling, analysis and optimization of potential lead candidates. This section surveys the computational methods that have been developed for the prediction of ligand selectivity.

Receptor-based techniques

The availability of a protein target structure is usually helpful in identifying potential ligand interactors. Such approaches usually involve explicit molecular docking of ligands into the receptor binding site, producing a predicted binding mode for each candidate compound. Predicting the preferred binding poses of ligands within a protein active site is difficult. First, the location and geometry of the binding site must be known, which may not always be addressed by X-ray crystallography or NMR studies. Second, the method must find the correct positioning of a compound in the active site of the protein. Third, the system must evaluate the relative goodness-of-fit or how well the compound can bind to the receptor in comparison with other compounds. An early venture was described by Platzer and colleagues, on calculating the relative standard free energy of binding of substrates to a-chymotrypsin.

Ligand-based techniques

Central to screening procedures based on ligands is the Similarity Property Principle, which asserts that molecules with similar structures are likely to share similar properties. This forms the basis for many ligand-based screening efforts where molecular structure and property descriptors of interacting molecules are extrapolated to search for other molecules with similar characteristics . For this, various machine learning techniques

have been described, including the use of decision trees , recursive partitioning, artificial neural networks (ANN) and support vector machines (SVM).

Assessment of ADME/Tox deficiencies

The disposition of a pharmaceutical compound may be described by its pharmacokinetic or ADME properties. In order to exert a pharmacological effect in tissues, a compound has to penetrate various physiological barriers, such as the gastrointestinal barrier, the blood–brain barrier and the microcirculatory barrier, to reach the blood circulation. It is subsequently transported to its effector site for distribution into tissues and organs, degraded by specialized enzymes, and finally removed from the body via excretion. In addition, genetic variation in drug metabolizing enzymes implies that some compounds may undergo metabolic activation and cause adverse reactions or Tox in humans. Accordingly, the ADME/Tox properties of a compound directly impact its usefulness and safety.

Limitation of computer aided drug design

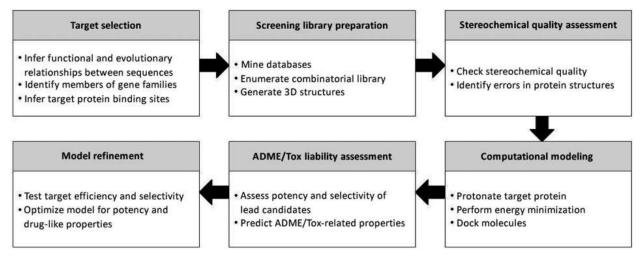Lengthy, expensive, and intellectually inelegant

Future research activities

1. Development and application of computational techniques for prediction of free energies of binding and salvation
2. Development and application of new methods for carbohydrate computational chemistry
3. Biomolecular simulation studies of proteins, sugars and DNA
4. QM/MM studies of the condensed phase
5. Homology/similarity modelling to obtain 3-dimensional structures for proteins we are interested in as targets for drug design, to design mutations or to study potential interactions with other proteins or nucleic acids
6. Designing lead drug structures and molecules which bind to enzyme active sites, to DNA in the minor groove or to tRNA and ribozymes

7. Designing novel molecular diagnostics based on new approaches to fluorescence using exciplexes

8. Re-designing proteins for molecular engineering, for example to produce variants of Green Fluorescent Protein that can be specifically chemically labelled to register enzyme action, such as the action of caspase 3 inside cells undergoing apoptosis

9. Designing molecules with novel chemical activities such as DNA cleavage

10. Understanding the conformational properties and energetics of small molecules

11. Determining high-resolution structures of chemically modified nucleic acids or of DNA: drug complexes using full distance geometry restraints combined with high- field NMR structural determinations of nucleic acid structures.

12. Understanding how families of ligands dock into binding sites of macromolecules

Procedure of computer aided drug design

Computer-based Design is Target specific and structure-based, Fast and automatic, Very low cost, high success rate. Computer aided drug design is the process which facilitate computational methods and resources that are used in design and discovery of new therapeutic solutions .Several new technologies have been developed and applied in drug R & D to shorten the research cycle and to reduce the expenses. Computer-aided drug design (CADD) is one of such evolutionary technologies.[30] CADD technologies including molecular modeling and simulation have become promising in drug discovery. Recently, CADD has even been used in designing highly selective ligands for a certain target that shares very similar structures with many proteins, which is difficult to be done by other methods. One such example is the rational design of selective inhibitors of p90 ribosomal protein S6 kinase. [31] In the post genomic era, owing to the dramatic increase of small molecule and biomacromolecule information, CADD tools have been applied in almost every stage of drug R & D, greatly changing the strategy and pipeline for drug discovery. [30] CADD, from its traditional application of lead discovery and optimization, has extended toward two directions: upstream for target

identification and validation, and downstream for preclinical study (ADMET prediction). Target identification and validation is the first key stage in the drug discovery pipeline. However, identification and validation of druggable targets from among thousands of candidate macromolecules is still a challenging task. [32] Numerous technologies for addressing the targets have been developed recently. Genomic and proteomic approaches are the major tools for target identification. For example, a proteomic approach for identification of binding proteins for a given small molecule involves comparison of the protein expression profiles for a given cell or tissue in the presence or absence of the given molecule. This method has not been proved very successful in target discovery because it is laborious and time-consuming. [33] Therefore, complementary to the experimental methods, a series of computational (insilico) tools have also been developed for target identification. They can be cataloged into sequence-based approach and structure-based approaches. Hence computational approaches to drug design fall into two general categories: those that do not assume information on the structure of the target macromolecule, and the structure-based approaches that do make use of such information.[34] Structure based approaches are not yet applicable because the structure of the target macromolecule is unknown; in these cases, quantitative structure-activity relationship (QSAR) techniques provide the best approach to rational drug design. Traditional (two-dimensional) QSAR methods attempt to correlate biological activity with local features of atoms, whole molecular properties (e.g. charge) and substituent effects (e.g. fragment hydrophobicity indices). New developments in traditional QSAR continue to appear in the literature (e.g. the OASIS program) (10). [35]Most interest in this field, however, now focuses on three-dimensional QSAR. Recent examples of molecules to which this approach has been applied include HIV-1 protease [36] and the cholecystokinin-A receptor.[37]

Target selection	Screening library preparation	Stereochemical quality assessment
• Infer functional and evolutionary relationships between sequences • Identify members of gene families • Infer target protein binding sites	• Mine databases • Enumerate combinatorial library • Generate 3D structures	• Check stereochemical quality • Identify errors in protein structures

Model refinement	ADME/Tox liability assessment	Computational modeling
• Test target efficiency and selectivity • Optimize model for potency and drug-like properties	• Assess potency and selectivity of lead candidates • Predict ADME/Tox-related properties	• Protonate target protein • Perform energy minimization • Dock molecules

A roadmap for structure-based screening campaign, comprising of (i) target selection (ii) library preparation and (iii) stereochemical quality assessment, ADME/Tox assessment and computational optimization.[29]
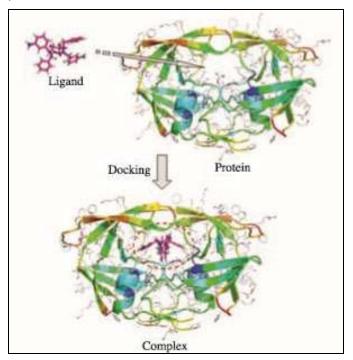

Fig. Docking Process

Conclusion

CADD is now widely recognized as a viable alternative and complement to high-throughput screening. The search for new molecular entities has led to the construction of high quality datasets and design libraries that may be optimized for molecular

diversity or similarity. Conversely, advances in molecular docking algorithms, combined with improvements in computational infrastructure, are enabling rapid improvement in screening throughput. Propelled by increasingly powerful technology, distributed computing is gaining popularity for large-scale screening initiatives. Combined with concerted efforts towards the design of more detailed physical models such as solubility and protein solvation, these advancements will, for the first time, allow the realization of the full potential of lead discovery by design.

Very Recent the European Union funded WISDOM (World-wide In Silico Docking on Malaria) project which analyzed over 41 million malaria-relevant compounds in 1 month using 1700 computers from 15 countries, and the Chinese funded Drug Discovery Grid (DDGrid) for anti-SARS and anti-diabetes research with a calculation capacity of >1 Tflops per second.

References

1. Ahindra Nag ,Baishakhi Dey, 2011. Computer-Aided Drug Design and Delivery Systems. The McGraw-Hill Companies, Inc.
2. Baldi A et al. Computational Approaches for Drug Design and Discovery: An Over View, IP: 117.204.64.189 August 25, 2010.
3. Yun Tang et al. New technologies in computer-aided drug design: Toward target identification and new chemical entity discovery. Vol. 3, No. 3, 2006.
4. Werner J.Geldenhuys et al. Optimizing the use of open-source software applications in drug discovery, Volume 11, Number 3/4 February 2006.
5. Song et. al., 2009. Recent advances in computer-aided drug design. Briefings in bioinformatics. VOL 10. NO 5. Pp.579-591.
6. Blundell TL, Dodson GG, Mercola D, et al. The structure, chemistry and biological activity of insulin. AdvProteinChem 1972;26:279–402.
7. Beddell CR, Goodford PJ, Norrington FE, et al. Compounds designed to fit a site of known structure in human haemoglobin. Br J Pharmacol 1976; 57:201–9.
8. Congreve M, Murray CW, Blundell TL. Structural biology and drug discovery. Drug DiscovToday 2005; 10:895–907.

9. Blundell TL. Structure-based drug design. Nature 1996; 384: 23–6.

10. Grover S, Apushkin MA, Fishman GA. Topical dorzolamide for the treatment of cystoid macular edema in patients with retinitis pigmentosa. Am J Ophthalmol 2006; 141:850–8.

11. Von Itzstein M, Wu WY, Kok GB, et al. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature 1993; 363:418–23.

12. Terrett NK, Bell AS, Brown D, et al. Sildenafil (ViagraTM), a potent and selective inhibitor of Type 5 CGMP phosphodiesterase with utility for the treatment of male erectile dysfunction. BioorgMed Chem Lett 1996; 6:1819–24.

13. Goodgame JC, Pottage JC Jr, Jablonowski H, et al. Amprenavir in combination with lamivudine and zidovudine versus lamivudine and zidovudine alone in HIV-1-infected antiretroviral-naive adults. Amprenavir PROAB3001 International Study Team. Antivir Ther 2000; 5:215–25.

14. Suman Bala *et al, 2010. Overview of different software used in drug design and drug discovery: A review. IJPT,Sep-2010,Vol. 2 ,Issue No.3,367-392.

15. Qian Yang et al. Modeling the binding modes of Kv1.5 potassium channel and blockers Journal of Molecular Graphics and Modelling 27, 178–187, 2008.

16. Bing Wu et al. In silico predication of nuclear hormone receptors for organic pollutants by homology modeling and molecular docking Toxicology Letters 191, 69–73, 2009.

17. Jadhav Ramulu and P. Goverdhan. Pharma tutor. Computer Aided Drug Design an Emerging Tool for Research and Drug Development. PHARMATUTOR-ART-1101. Page 1-16.

18. Luis Fernando Saraiva Macedo Timmers et al. Molecularmodeling, dynamics and docking studies of PurineNucleoside Phosphorylase from Streptococcus pyogenes Biophysical Chemistry 142, 7–16, 2009.

19. Werner J.Geldenhuys et al. Optimizing the use of open-source software applications in drug discovery DDT Volume 11, Number 3/4 February 2006.

20. Perdo and Hui lei, 2010. A systematic review on CADD: Docking and Scoring. JMPI. Page 47-51.

21. Tetko, I.V. et al. Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program. J. Chem. Inf. Comput. Sci. 42, 1136–1145, 2002.

22. LS. Haworth, A.H. Elcock, A. Rodger and W.G. Richards. J. Sequence selective binding to the DNA major groove: tris (1,lO phenanhline) metal complexes binding to poly(dGdC) and poly (dAdT). *Biomol. Struct. and Dynamics* 9,23 (1991).

23. Bioreductive anti-cancer drugs. W.G. Richards and C.A. Reynolds, in "Theoretical Biochemistry and Molecular Biophysics" (eds. D.L. Beveridge *8t* R. Lavery), Adenine Press, Schenectady (1990).

24. Free energy calculations of pharmaceutically important properties. P.M. King, C.A. Reynolds, J.W. &ex, G.A. Worth and W.G. Richards. *Mol. Simulation* 5,265 (1990).

25. W. G. RICHARDS, 1994. Computer-aided drug design. *Pure* & *Appl. Chern.,* Vol. 66, No. 8, pp. 1589-1596.

26. A theoretical study of the structure of big endothelin. M.C. Menziani, M. Cocchi, P.G. De Benedetti, R.G. Gilbert, W.G. Richards, M. Zamai, V.R. Caiolfa. J. *ChimJhys.* 88,2687 (1991).

27. A linear molecular similarity index. C.A. Reynolds, C. Burt and W.G. Richards. *Quant. Struct.Act. Relat.* 11, 34 (1992).

28. Structure-activity relationships from molecular si.milarity matrices. A.C. Good, Sung-Sau So and W.G. Richards. *J. Med. Chem.* 36,433 (1993).

29. Chun Meng Song et. al., 2009. Recent advances in computer-aided drug design. Briefings in bioinformatics. VOL 10. NO 5. pp.579-591.

30. Jorgensen, W.L. et al. The many roles of computation in drug discovery. Science 303, 1813- 1818, 2004.

31. Cohen, M.S. et al. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. Science308, 1318–1321, 2005.

32. Hajduk, P.J. et al. Predicting protein druggability. Drug Discovery Today 10, 1675–1682, 2005.

33. Huang, C.M. et al. Proteomics reveals that proteins expressed during the early stage of Bacillus anthracis infection are potential targets for the development of vaccines and drugs. Genomics Proteomics Bioinformatics 2, 143–151, 2004

34. Robert C Jackson et al. Update on computer-aided drug design, Current Opinion in Biotechnology , 6:646-651, 1995.

35. Mekenyan OG et al. A new development of the OASIS computer system for modeling molecular properties. Computer Chemistry, 18:173-187, 1994.

36. Oprea TI et al. 3-D QSAR of HIV-1 protease inhlbltors, 2. Predictive power using limited exploration of alternate binding modes. Med Chem, 37:2206-2215, 1994.

37. Tokarski JS et al. 3-D molecular shape analysis-QSAR of a series of cholecystokinin-A receptor antagonists. J Meal Chem, 37:3639-3654, 1994.