Bulletin of Pharmaceutical Research 2014;4(3):118-23
An Official Publication of Association of Pharmacy Professionals
ISSN: 2249-6041 (Print); ISSN: 2249-9245 (Online)

REVIEW ARTICLE

EFFICIENT WAY OF DRUG DESIGNING: A COMPREHENSIVE REVIEW ON COMPUTATIONAL TECHNIQUES

Vikas Sharma[1] and Vipin Kumar[2]*

[1]*Department of Pharmaceutical Chemistry, Institute of Pharmaceutical Sciences, Kurukshetra University, Kurukshetra-136 119, Haryana, India*
[2]*Department of Pharmacy, School of Chemical Sciences and Pharmacy, Central University of Rajasthan, Ajmer-305 801, Rajasthan, India*

**E-mail*: vipkumar@curaj.ac.in
Tel.: +91 1463 238729.

Received: Nov 10, 2014 / Revised: Dec 09, 2014 / Accepted: Dec 10, 2014

As demand of novel drugs is increasing day by day, the need of efficient, inexpensive techniques arises which can help in drug design and discovery. In order to achieve the goal, computational techniques like docking, pharmacophore modeling, homology modeling are employed by researchers around the globe in search of novel potent drugs. In this review, relevant computational techniques are highlighted which will help new generations to carry out hassle free research in an efficient manner.

Key words: Computational, Docking, Methodology, Modeling, Drug design.

INTRODUCTION

The serendipitous part of drug discovery can now be overcome by rational designing of drugs with the help of computational techniques like molecular modeling, docking, virtual screening etc. that can identify promising candidates prior to synthesis. The structural aspects of active site (drug targets) composition and the orientation of different amino acids can be used in drug design and discovery. Nowadays, with the help of computational techniques, a specific potent lead molecule against a particular disease can be designed on the basis of nature of interactions like enzyme-substrate, drug-protein and drug-nucleic acid interactions that may provide a conceptual framework for designing the desired potency and specificity of potential drug leads for a given therapeutic target (Nantasenamat *et al* 2009; Aparoy *et al* 2012).

The active and time consuming process of drug discovery and development *via* traditional approaches like synthesis and evaluation of a potent medicinally active compound can be replaced by computational methods. The most important aspect of computer aided drug design is its crucial role in drug discovery that can save time and cost involved in drug development process. It can now easily be assessed from the literature survey that computational methods have become an interdisciplinary science with the involvement of different scientific fields like pharmacology, molecular biology, chemistry etc. (Aparoy *et al* 2012).

In present review, an attempt was made to describe different computational techniques with an idea to achieve wide circulation which in turn would be beneficial for new researchers where they can understand and design their research more efficiently. The literature study reveals the impact of computational approaches in drug designing and a wide number of publications are available on docking, mapping, homology modeling etc. (Teif, 2005; Todeschini and Consonni, 2008; Congreve and Marshall, 2010; Baron *et al* 2010; Yang, 2010; Kumar, 2011; Sharma *et al* 2011; Alberts *et al* 2013; **Figure 1**). **Table 1** enumerates different terms used in computational studies.

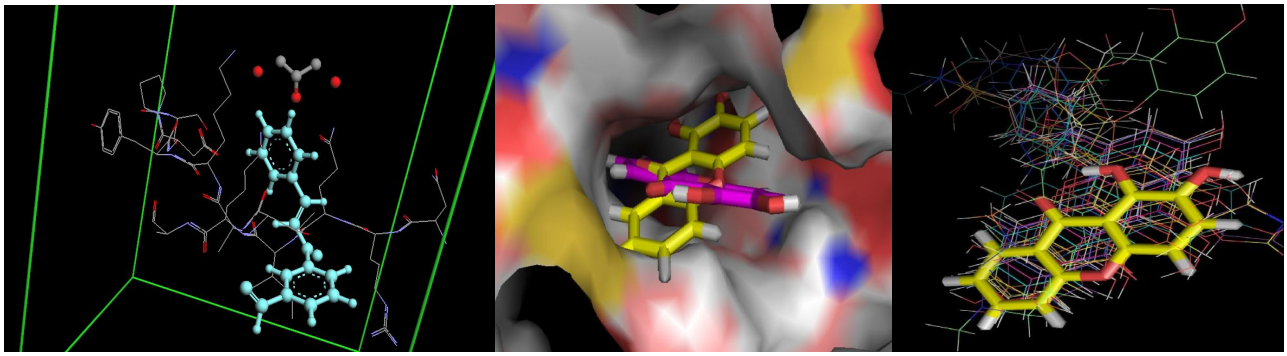

Fig. 1. Images of molecular docking studies

Table 1. General terms used in computational drug design

S. No.	Computational term	Definition
1	Pharmacophore	It is a molecular framework that carries (phoros) the essential features responsible for a drug's (pharmacon) biological activity'
2	Active site	The active site is a small port in an enzyme that where substrate molecules bind and undergo a chemical reaction.
3	Descriptor	The molecular descriptor is the final result of a logic and mathematical procedure which transforms chemical information encoded within a symbolic representation of a molecule into a useful number or the result of some standardized experiment
4	Ligand	A ligand is a substance (usually a small molecule) that forms a complex with a biomolecules to serve a biological purpose.
5	Receptor	It is a protein molecule usually found embedded within the plasma membrane surface of a cell that receives chemical signals from outside the cell.
6	Affinity	The tendency or strength of ligand binding to receptor is known as affinity.
7	Dataset	A set of biologically active molecules.
8	Training Set	The dataset used to develop the 3D-model.
9	Test Set	The dataset used to evaluate the developed 3D model.

Computational techniques

Drug Design can be classified mainly in two different categories *viz.* a) Structure based drug design (SBDD) and b) Ligand based drug design (LBDD).

Structure based drug design (SBDD)

In structure based drug design, an inhibitor for a particular target has been developed on the basis of structural information of the drug target like receptor structure (mostly protein). If the structure of receptor is not available, the receptor structure can be predicted by homology modeling. Homology modeling usually refers to as comparative modeling in which on the basis of known amino acid sequences of a protein, a model of protein can be constructed and the structure is comparable with the 3D-structure of similar homologous protein (template).

Docking:

In this approach, ligands are 'docked' against the structures of bimolecular targets and a certain score (usually referred to as 'docking score') has been given to each orientation a ligand docked in the active site. This score can then be used to evaluate the potential of ligand-protein affinity which ultimately leads to prediction of biological effectiveness of a ligand against the particular protein. Docking is employed only when the structure of target's active/binding site is available and orientation/conformation of a ligand is predicted in the active site of desired target.

Process

There are three types of computational docking *viz.* a) *Rigid body docking* in which both receptor and ligand are rigid; b) *Flexible ligand docking* in which receptor is rigid while ligand is flexible; c) *Flexible docking* in which both receptor and ligand are flexible. Most commonly used docking type is flexible docking (Mohan *et al* 2005).

Different algorithms used in the process of docking are Monte Carlo, genetic algorithm, fragment-based, molecular dynamics etc and on the basis of these algorithms, different programs were developed (free and commercial purpose).

Ligand Preparation:
It is the initial step for molecular docking studies where geometrical refining of 3D structures takes place. In most docking programs, different tautomeric and conformations were generated by employing search algorithms. All the generated conformers were minimized and prepared ligands will be used to be docked against biomolecular target (Mohan *et al* 2005).

Protein preparation:
Structures of protein/biomolecules evaluated by X-ray crystallographic technique are available on protein data bank and it could easily be downloaded in text format from their website [http://www.rcsb.org]. The selected chains of the biomolecules were edited for missing hydrogens and to assign proper bond orders. Later on all the H-bonds were optimized and all polar hydrogen were displayed. Later on, minimization of resulted protein structure was carried out (Mohan *et al* 2005).

Receptor grid generation:
From the prepared protein/biomolecule, the cocrystallized ligand was separated from its active site. The active site is generally represented as an enclosing box at the centroid of workspace ligand. Following this protocol, a grid centered on the ligand was generated using the default settings of desired software. All ligands were docked into this grid structure (Kaushik *et al* 2012).

Docking and scoring:
On a defined receptor grid, flexible docking was performed using appropriate module of desired software. The module analyses the protein-ligand interaction on the basis of different interactions between them like vander waals, hydrogen bonding and electrostatic interactions. Furthermore, the binding free energy of a protein-ligand complex by adding up individual contributions from different types of interactions is predicted by using force field scoring functions. However, it is relevant to note that due to availability of various docking programs, the scoring function of each program might be different with respect to another and lack of universal scoring function results in non-uniform results by different docking programs (Mohan *et al* 2005).

Homology modeling:
Structural information of any biomolecule can be accessed from Protein Data Bank (PDB), (http://www.pdb.org). However, when structure of a desired target is not available, in that case 'Homology modeling' can be used to predict 3D protein structure in order to provide an insight into the functioning of protein (Cavasotto and Phatak, 2009).
This methodology is derived from a general observation that 'proteins with similar sequences have similar structures'. The models can be generated for a homologous sequence (target) that bears template significant sequence of more than 30% (Cavasotto and Phatak, 2009).

Process
1) Identification of a template (known 3D structure)
2) Sequence alignment of target and template proteins
3) Building the model
4) Validation of the generated model
These steps can be repeated again and again in order to build a satisfactory model (Cavasotto and Phatak, 2009).
The first step in this procedure is to get the sequence of amino acids for the desired protein. Subsequently, by using BLAST (Basic Local Alignment Search Tool) algorithm sequence alignment is conducted where most similar amino acid sequence is identified that resembles the query sequence. Quality of a developed model can be assessed by the quality of sequence alignment and template structure.
Quality of developed model can be ascertained by estimating the local quality of the predicted structure on the basis of tests like mean force potential, GROMOS empirical force field energy, QMEAN etc. Furthermore, stereochemistry of developed model and analysis of the ramachandran plot is always of great worth.

Ligand based drug design (LBDD)
This approach is particularly useful when 3D structure of the receptor is not available and it relies on the knowledge of ligands that bind to the desired target. Two most prominent techniques used in this approach are a) 3D quantitative structure activity relationships (3D QSAR); b) Pharmacophore modeling.
The above techniques are used to develop models with predictive ability suitable for lead identification and optimization (Acharya *et al* 2011). In quantitative structure activity relationship (QSAR), a correlation between experimentally determined biological activity

and calculated properties of molecules is derived. These exhibit a particular squared predictive correlation coefficient (r^2); and model with r^2 value close to 1 will be designated as best model. These QSAR models relationships can then be used to predict the activity of new analogs.

3D-QSAR:

Comparative Molecular Field Analysis (CoMFA)

CoMFA is categorized as 3D QSAR computational technique in which experimental activities (log units of KI or IC_{50}) and the 3D structures of the molecules are incorporated in the study. For ease in calculation of electrostatic energy, charges should be added to the molecules so that electrostatic energy can be determined. A common substructure with good alignment is considered as an important aspect in CoMFA analysis. The common substructure should have the same conformation in all molecules overlaid in a single binding mode by adjusting their internal torsional angles. Molecular fields around each conformation calculated and the fields, generally electrostatic (sp_3-hybridized carbon with +1 charge) and steric energy (Lennard-Jones potential) are measured at the lattice points of a regular Cartesian 3D grid (lattice spacing ~ 2 Å).

Pharmacophore modeling

Pharmacophore modelling is intended for the virtual screening of compounds for lead discovery (Kaushik *et al* 2012). Pharmacophore modeling proposed to be a powerful technique which can easily classify a group of molecules/ligands into active or inactive compounds.

The key feature identified by pharmacophore modeling will provide new insights in drug design and discovery. Currently, various automated pharmacophore generators have been developed, including commercially available softwares (Barnum *et al* 1996; Li *et al* 2000; Martin, 2000; Jones *et al* 2000; Triopos; Dixon *et al* 2006; Goodsell *et al* 1996; Cozzini *et al* 2008; Wolber and Langer, 2005; Sakkiah *et al* 2009; Bansal *et al* 2011; Balasubramanian and Vijaya Gopal, 2012), as shown in **Table 2**.

In general, pharmacophore generation from multiple ligands (usually called training set compounds) involves two main steps 1) creating the conformational space for each ligand in the training set to represent conformational flexibility of ligands, and 2) aligning the multiple ligands in the training set and determining the essential common chemical features to construct pharmacophore models.

Table 2. Softwares available for drug designing

S. No.	Name of software	Applications in drug design
1	HipHop	QSAR
2	HypoGen	QSAR
3	DISCO	QSAR
4	GASP	QSAR
5	GALAHAD	QSAR
6	PHASE	QSAR
7	MOE	QSAR
8	LigandScout	Pharmacophore modeling
9	Sanjeevani	docking
10	AutoDock	docking
11	SLIDE	docking
12	Surflex	docking
13	ICM	docking
14	GLIDE	docking

Process

Biological activity data:

A dataset of more than 25 derivatives with a particular biological activity can be considered for pharmacophore modeling. The IC_{50} (*i.e.* concentration in μM required for 50% inhibition of enzyme activity) of all derivatives was converted into pIC_{50} (–log IC_{50}) which will be used as data input for a particular software. The dataset was divided randomly into the training (A %) and test sets (B %); where A/B represents the percentage of dataset.

Training set was used to generate pharmacophore models and prediction of the activity of test set was used as a method to validate the proposed models.

Ligand preparation:

This step is similar to the step that is described before in molecular docking studies where ligands with correct chiralities and with different conformations were developed using different programs.

Creation of pharmacophoric sites:

A set of pharmacophore features like hydrogen-bond acceptor, hydrogen-bond donor, hydrophobic group, aromatic ring etc. were used to create pharmacophoric sites for developing a pharmacophore model.

Searching common pharmacophore:
In this step, a specific module of software analyzes the pharmacophores from all conformations of the training set molecules and assembles together the pharmacophores having identical sets of features with very similar spatial arrangements. A common pharmacophore is generated if a given group is found to contain at least one pharmacophore from each ligand. A tree-based partitioning technique was then used to identify common pharmacophores, in which similar pharmacophores were grouped together according to their intersite distances.

Scoring hypothesis
In this step, common pharmacophore hypothesis were examined using a scoring function to yield the best alignment of the active ligands. The quality of alignment was measured by survival score.

Generation of 3D-QSAR model
3D-QSAR models were developed for a set of ligands by using the method of structure alignment. The pharmacophore model partitions space into a grid of uniformly sized cubes, and each molecule was characterized by a set of binary-valued (0 or 1) independent variables that encode the occupancy of these cubes by various atom classes and pharmacophore feature types. 6 N occupancies (N = no. of atom classes) of the cubes and atom classes are the independent variables used in the QSAR model. These binary-valued (0 or 1) independent variables are used as parameters to develop correlation with biological activity.

Validation of pharmacophore model
External validation is considered to be a conclusive proof for judging predictability of a model as validation is a crucial aspect of pharmacophore design, particularly when the model is built for the purpose of predicting activities of compounds in external test series. The main target of any QSAR modelling is that the developed model should be robust enough to be capable of making accurate and reliable predictions of biological activities of new compounds. In studies, the prediction of the activity of test set molecules used as a method to validate the developed pharmacophore model.

CONCLUSION

Regardless of the achievements of computational techniques in past decades, there is some room of improvement and several key challenges needs to be addressed in this field. Some of the challenges includes a) Modeling of ligand and protein flexibility b) high computing cost c) unavailability of a universal methodology; as different softwares implies their own methodology (like use of different algorithms) for pharmacophore modeling. An ideal program for processing a particular computational technique should bear the following characteristics: a) keeping low-energy conformations of ligand molecules b) efficient generation of all conformations c) less time consuming. Another important point which normally ignored by researchers is "*selection of proper training set compounds*". It has been established that dataset size, type of ligand and their chemical diversity may significantly affect quality of generated pharmacophore model. Moreover, the supremacy of computation techniques in drug design and discovery is well-recognized by researchers from interdisciplinary fields like chemistry, molecular biology etc. and success stories in drug discovery is increasing day by day.

REFERENCES

Acharya C, Coop A, Polli JE, Mackerell AD Jr. Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr. Comput. Aided Drug Des.* 2011;7(1):10-22. [DOI: 10.2174/157340911793743547]

Alberts B, Bray D, Hopkin K, Johnson AD, Lewis J, Raff M, Roberts K, Walter P. Essential Cell Biology, 4th edition, Garland Science, New York: 2013; 91.

Aparoy P, Reddy KK, Reddanna P. Structure and ligand based drug design strategies in the development of novel 5-LOX inhibitors. *Curr. Med. Chem.* 2012;19(22):3763-78. [DOI: 10.2174/092986712801661112]

Balasubramanian R, Vijaya Gopal R. Design and *in silico* analysis of ring-A monosubstituted chalcones as potential anti-inflammatory agents. *Bull. Pharm. Res.* 2012;2:70-7.

Bansal H, Sharma A, Sharma V, Kumar V. Pharmacophore modeling studies on xanthones as monoamine oxidase-A inhibitors. *Bull. Pharm. Res.* 2011;1(1):15-21.

Barnum D, Greene J, Smellie A, Sprague P. Identification of common functional configurations among molecules. *J. Chem. Inf. Comput. Sci.* 1996;36(3):563-71. [DOI: 10.1021/ci950273r]

Baron R, Setny P, and McCammon JA. Water in cavity-ligand recognition. *J. Am. Chem. Soc*. 2010;132 (34):12091-7. [DOI: 10.1021/ja1050082]

Cavasotto CN, Phatak SS. Homology modeling in drug discovery: current trends and applications. *Drug Discov. Today* 2009;14(13-14):76-683. [DOI: 10.1016/j.drudis.2009.04.006]

Chemical Computing Group, http://www.chemcomp.com

Congreve M, Marshall F. The impact of GPCR structures on pharmacology and structure-based drug design. *Br. J. Pharmacol.* 2010;159(5): 986-96. [DOI: 10.1111/j.1476-5381.2009.00476.x]

Cozzini P, Kellogg GE, Spyrakis F, Abraham DJ, Costantino G, Emerson A, Fanelli F, Gohlke H, Kuhn LA,Morris GM, Orozco M, Pertinhez TA, Rizzi M, Sotriffer CA. Target flexibility: an emerging consideration in drug discovery and design. *J. Med. Chem.* 2008;51(20):6237-55. [DOI: 10.1021/jm800562d]

Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3DQSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J. Comput. Aided Mol. Des.* 2006;20(10-11):647-71. [DOI: 10.1007/s10822-006-9087-6]

Glide, Version 5.6 (2010) Schrodinger. LLC, New York

Goodsell DS, Morris GM, Olson AJ. Automated docking of flexible ligands: applications of AutoDock. *J. Mol. Recognit.* 1996;9(1);1-5.

Jain J, Bansal SK, Chowdhury P, Sinha R, Tripathi U, Malhotra M. In silico pharmacophore validation of anticonvulsant activity of (E) (±)-3-menthone derivatives. *Bull. Pharm. Res.* 2013;3(3):146-56.

Jones G, Willet P, Glen RC. GASP: genetic algorithm superimposition program. In Pharmacophore Perception, Development, and Use in Drug Design Guner OF. (ed.). International University Line, La Jollan, CA: 2000; 85-106.

Kaushik U, Sharma V, Kumar V. Computation of pharmacophore models for the prediction of mitogen-activated protein kinase activated protein kinase-2 inhibitory activity of pyrrolopyridines. *Med. Chem. Res.* 2012;21:3777-84. [DOI: 10.1007/s00044-011-9910-z]

Kumar V. Topological models for the prediction of tyrosine kinase inhibitory activity of 4-anilinoquinazolines. *Bull. Pharm. Res.* 2011;1(2):53-9.

Li H, Sutter J, Hoffman R. HypoGen: an automated system for generating 3D predictive pharmacophore models. In Pharmacophore Perception, Development, and Use in Drug Design. Guner OF. (ed.). International University Line, La Jollan, CA: 2000; 171-89.

Martin YC. DISCO: what we did right and what we missed. In Pharmacophore Perception, Development, and Use in Drug Design. Guner OF. (ed.). International University Line, La Jollan, CA: 2000; 49-68.

Mohan V, Gibbs AC, Cummings MD, Jaeger EP, DesJarlais RL. Docking: successes and challenges. *Curr. Pharm. Des.* 2005;11:323-333. [DOI: 10.2174/138161205338210 6]

Nantasenamat C, Isarankura-Na-Ayudhya C, Naenna T, Prachayasittikul V. A practical overview of quantitative structure-activity relationship. *Excli. J.* 2009;8:74-88.

Sakkiah S, Krishnamoorthy N, Gajendrarao P, Thangapandian S, Lee Y, Kim S, Suh J-K, Kim H-H, Lee KW. Pharmacophore mapping and virtual screening for SIRT1 activators. *Bull. Korean Chem. Soc.* 2009;30(5):1152-6. [DOI: 10.5012/bkcs.2009.30.5.1152]

Sharma V, Wakode SR, Lather V, Mathur R, Fernandes MX. Structure based rational drug design of selective phosphodiesterase-4 ligands as anti-inflammatory molecules. *Bull. Pharm. Res.* 2011;1(2):33-40.

Teif VB. Ligand-induced DNA condensation: choosing the model. *Biophys. J.* 2005;89(4):2574-87. [DOI: 10.1529/biophysj.105.063909]

Todeschini R, Consonni V. Handbook of Molecular Descriptors. John Wiley & Sons, USA: 2008.

Tripos Inc., http://www.tripos.com

Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.* 2005;45(1): 160-9. [DOI: 10.1021/ci049885e]

Yang S-Y. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov. Today* 2010;15(11-12):444-50. [DOI: 10.1016/j.drudis.2010.03.013]

http://www.molsoft.com/docking.html

http://www.tripos.com/index.php?family=modules,SimplePage,,,&page=surflex_dock&s=0

http://www.scfbio-iitd.res.in/sanjeevini/sanjeevini.jsp

*****