

COMPUTER-AIDED DRUG DESIGN - A NEW APPROACH IN DRUG DESIGN AND DISCOVERY

Satyajit Dutta^{*1}, Sovan Sutradhar², Kapil Sachan³

¹Assistant Professor, ²Post Graduate Research Scholar, Department of Pharmacy, IIMT College of Medical Sciences, 'O'Pocket, Ganga Nagar, Mawana Road, Meerut-250001, Uttar Pradesh, India.

³Post Graduate Research Scholar, School of Pharmaceutical Sciences, Shobhit University, NH-58, Modipuram, Meerut - 250 110, Uttar Pradesh, India.

*Email: sanku6@gmail.com

ABSTRACT

Two distinct approaches are possible in the area of computer-aided drug design. If the molecular structure of the target macromolecule is known the methods are obvious and direct and have achieved a high level of sophistication. That area may be extended by using computational techniques to predict protein structure like Cancer-Causing H-Ras p21 Mutant Protein and to simulate drug-receptor interactions. CADD methods are heavily dependent on bioinformatics tools, applications and databases. As such, there is considerable overlap in CADD research and bioinformatics. When the only lead is a set of known active compounds or knowledge of a biochemical transformation which is to be interrupted, then the path is less direct. Currently favored tactics include the use of molecular similarity methods and the employment of neural networks. Recent advances include the prediction of the relative potency of different chiral forms of drugs.

Keywords: Bioinformatics, Computer-Aided Drug Design, Combinatorial library, Virtual High-Throughput Screening.

INTRODUCTION

Advances in the fields of biochemistry, molecular biology and cell biology, facilitated by developments in genomics and proteomics, are producing a large number of novel biological targets that may be exploited for therapeutic intervention. To facilitate the discovery of novel therapeutic agents, rational drug design methods in combination with structural biology offer great potential. The Computer-Aided Drug Design (CADD) Center was created to foster collaborative research between biologists, biophysicists, structural biologists and computational scientists¹. The major goal of the CADD Center is to initiate these collaborations leading to the establishment of research projects to discover novel chemical entities with the potential to be developed into novel therapeutic agents. Computer-aided drug design (CADD) is an exciting and diverse discipline where various aspects of applied and basic research merge and stimulate each other². The latest technological advances (QSAR/QSPR, structure-based design, combinatorial library design, cheminformatics & bioinformatics); the growing number of chemical and biological databases; and an explosion in currently available software tools are providing a much improved basis for the design of ligands and inhibitors with desired specificity³.

Structure-based drug design

Within many of the rational drug design projects in the group, computer-aided methods, such as virtual screening and *de novo* design techniques, play an important role. These projects involve the School's computational chemistry group, led by Dr Richard Bryce⁴. His group also develops new computational approaches, with a focus on molecular dynamics, solvation, hybrid QM/MM methods and carbohydrate modeling.

NMR spectroscopy in conjunction with molecular modeling and other spectroscopic methods allows investigations to be made into molecular mechanisms of ligand-target recognition at the atomic level. This provides a description of the region of target and drug surfaces involved in the interaction, and important contacts between drug and target responsible for affinity and specificity. The structural analysis of DNA/RNA-ligand interactions by high-field NMR is crucial to define structure-function correlations. This information is a necessary component in the design of novel therapeutics and in prediction of interactions of drugs with the targets.

For example, the group can analyze the shapes of macrolide antibiotics bound to ribosomes⁵. Also over the years, the group has studied details of binding of ligands to the minor groove of DNA, such as Hoechst 33258, or to tRNA. NMR methods are also used by the group to study interactions of proteins with ligands. There is 300 MHz instrumentation in the School, and the group has shared usage of 500 MHz high-field instruments housed in the Department of Chemistry. The group collaborates extensively with Professor Gareth Morris, inventor and pioneer of many modern NMR techniques, thereby bringing novel techniques to bear on red biological problems⁵.

Bioinformatics in Computer-Aided Drug Design

A few years ago, the National Institutes of Health (NIH) created the Biomedical Information Science and Technology Initiative (BISTI) to examine the current state of bioinformatics in the United States. BISTI's working definition of bioinformatics included its use in biomedical research, in particular for drug discovery and development programs. Bioinformatics was seen as an emerging field with the potential to significantly improve



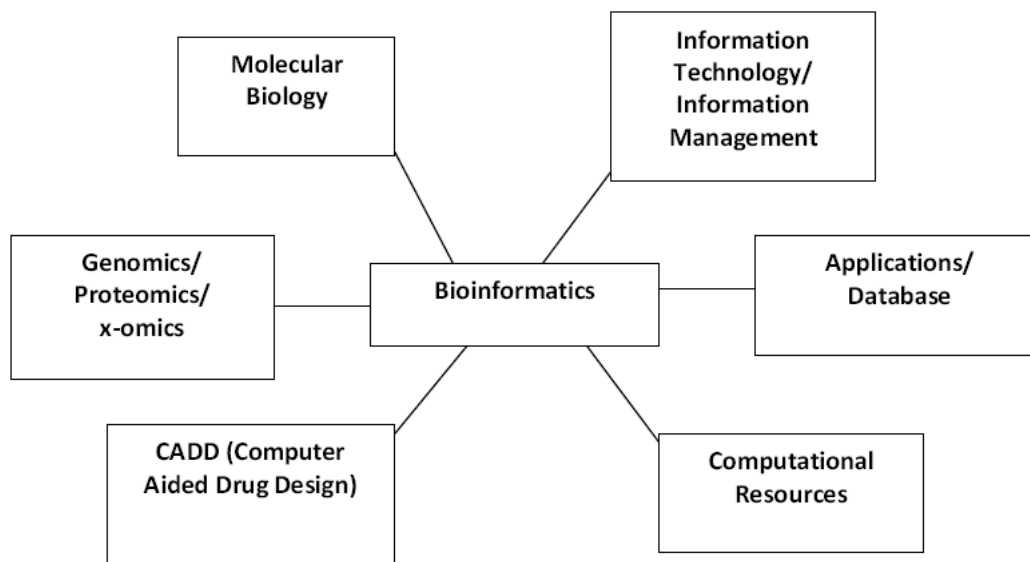
how drugs are found, brought to clinical trials and eventually released to the marketplace.

Computer-Aided Drug Design (CADD) is a specialized discipline that uses computational methods to simulate drug-receptor interactions. CADD methods are heavily dependent on bioinformatics tools, applications and

databases. As such, there is considerable overlap in CADD research and bioinformatics.

Bioinformatics Hub

Bioinformatics can be thought of as a central hub that unites several disciplines and methodologies⁶.



On the support side of the hub, Information Technology, Information Management, software applications, databases and computational resources all provide the infrastructure for bioinformatics. On the scientific side of the hub, bioinformatics methods are used extensively in molecular biology, genomics, proteomics, other emerging areas (*i.e.* metabolomics, transcriptomics) and in CADD research.

There are several key areas where bioinformatics supports CADD research.

Virtual High-Throughput Screening (vHTS):

Pharmaceutical companies are always searching for new leads to develop into drug compounds. One search method is virtual high-throughput screening. In vHTS, protein targets are screened against databases of small-molecule compounds to see which molecules bind strongly to the target. If there is a “hit” with a particular compound, it can be extracted from the database for further testing. With today’s computational resources, several million compounds can be screened in a few days on sufficiently large clustered computers. Pursuing a handful of promising leads for further development can save researchers considerable time and expense. ZINC is a good example of a vHTS compound library.

Sequence Analysis:

In CADD research, one often knows the genetic sequence of multiple organisms or the amino acid sequence of proteins from several species. It is very useful to determine how similar or dissimilar the organisms are based on gene or protein sequences. With this

information one can infer the evolutionary relationships of the organisms, search for similar sequences in bioinformatics databases and find related species to those under investigation. There are many bioinformatics sequence analysis tools that can be used to determine the level of sequence similarity.

Homology Modeling:

Another common challenge in CADD research is determining the 3-D structure of proteins. Most drug targets are proteins, so it’s important to know their 3-D structure in detail. It’s estimated that the human body has 500,000 to 1 million proteins. However, the 3-D structure is known for only a small fraction of these. Homology modeling is one method used to predict 3-D structure. In homology modeling, the amino acid sequence of a specific protein (target) is known, and the 3-D structures of proteins related to the target (templates) are known. Bioinformatics software tools are then used to predict the 3-D structure of the target based on the known 3-D structures of the templates⁷. MODELLER is a well-known tool in homology modeling, and the SWISS-MODEL Repository is a database of protein structures created with homology modeling.

Similarity Searches:

A common activity in biopharmaceutical companies is the search for drug analogues. Starting with a promising drug molecule, one can search for chemical compounds with similar structure or properties to a known compound. There are a variety of methods used in these searches, including sequence similarity, 2D and 3D shape similarity, substructure similarity, electrostatic similarity and others.



A variety of bioinformatic tools and search engines are available for this work⁸.

Drug Lead Optimization:

When a promising lead candidate has been found in a drug discovery program, the next step (a very long and expensive step) is to optimize the structure and properties of the potential drug. This usually involves a series of modifications to the primary structure (scaffold) and secondary structure (moieties) of the compound. This process can be enhanced using software tools that explore related compounds (bioisosteres) to the lead candidate. OpenEye's WABE is one such tool. Lead optimization tools such as WABE offer a rational approach to drug design that can reduce the time and expense of searching for related compounds⁹.

Physicochemical Modeling:

Drug-receptor interactions occur on atomic scales. To form a deep understanding of how and why drug compounds bind to protein targets, we must consider the biochemical and biophysical properties of both the drug itself and its target at an atomic level. Swiss-PDB is an excellent tool for doing this. Swiss-PDB can predict key physicochemical properties, such as hydrophobicity and polarity that have a profound influence on how drugs bind to proteins.

Drug Bioavailability and Bioactivity:

Most drug candidates fail in Phase III clinical trials after many years of research and millions of dollars have been spent on them. And most fail because of toxicity or problems with metabolism. The key characteristics for drugs are Absorption, Distribution, Metabolism, Excretion, Toxicity (ADMET) and efficacy-in other words bioavailability and bioactivity. Although these properties are usually measured in the lab, they can also be predicted in advance with bioinformatics software¹⁰.

Benefits of CADD

CADD methods and bioinformatics tools offer significant benefits for drug discovery programs¹¹.

Cost Savings:

The Tufts Report suggests that the cost of drug discovery and development has reached \$800 million for each drug successfully brought to market. Many biopharmaceutical companies now use computational methods and bioinformatics tools to reduce this cost burden. Virtual screening, lead optimization and predictions of bioavailability and bioactivity can help guide experimental research. Only the most promising experimental lines of inquiry can be followed and experimental dead-ends can be avoided early based on the results of CADD simulations.

Time-to-Market:

The predictive power of CADD can help drug research programs choose only the most promising drug

candidates. By focusing drug research on specific lead candidates and avoiding potential "dead-end" compounds, biopharmaceutical companies can get drugs to market more quickly.

Insight:

One of the non-quantifiable benefits of CADD and the use of bioinformatics tools is the deep insight that researchers acquire about drug-receptor interactions. Molecular models of drug compounds can reveal intricate, atomic scale binding properties that are difficult to envision in any other way. When we show researchers new molecular models of their putative drug compounds, their protein targets and how the two bind together, they often come up with new ideas on how to modify the drug compounds for improved fit. This is an intangible benefit that can help design research programs.

CADD and bioinformatics together are a powerful combination in drug research and development. An important challenge for us going forward is finding skilled, experienced people to manage all the bioinformatics tools available to us, which will be a topic for a future article.

More Computer-aided Drug Design Developed Will Speed up Drug Development

Researchers in Germany report an advance toward the much awaited era in which scientists will discover and design drugs for cancer, arthritis, AIDS and other diseases almost entirely on the computer, instead of relying on the trial-and-error methods of the past.

In the report, Michael C. Hutter¹² and colleagues note that computer-aided drug design already is an important research tool. The method involves using computers to analyze the chemical structures of potential drugs and pinpoint the most promising candidates. Existing computer programs check a wide range of chemical features to help distinguish between drug-like and nondrug materials. These programs usually cannot screen for all features at the same time, an approach that risks overlooking promising drug-like substances.

In the new study, researchers describe a more gradual and efficient system. Their new program uses an initial quick screen for drug-like features followed immediately by a second, more detailed screen to identify additional drug-like features. They applied this new classification scheme to a group of about 5,000 molecules that had previously been screened for drug-like activity. The new strategy was more efficient at identifying drug-like molecules "whereby up to 92 percent of the nondrugs can be sorted out without losing considerably more drugs in the succeeding steps," the researchers say.

Computer-Aided Drug Design for Cancer-Causing H-Ras p21 Mutant Protein:

GTP-bound mutant form H-Ras (Harvey-Ras) proteins are found in 30% of human tumors. Activation of H-Ras is due to point mutation at positions 12, 13, 59 and/or 61 codon.



Mutant form of H-Ras proteins is continuously involved in signal transduction for cell growth and proliferation through interaction of downstream-regulated protein Raf (Fig. 1).

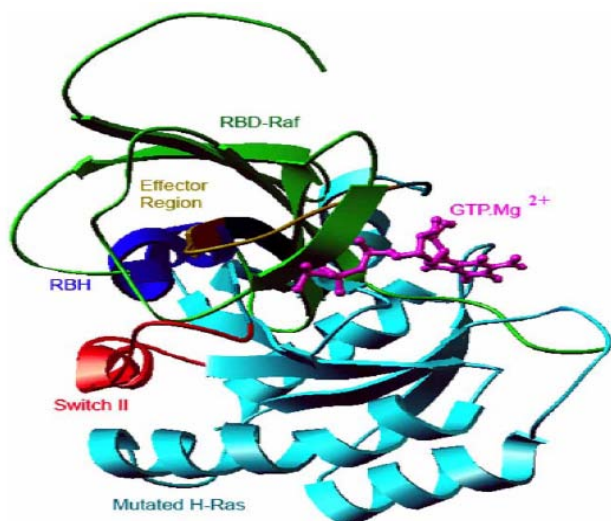


Figure 1: Mutated H-Ras-Ras Binding Domain (RBD) Complex Structure; the regions of effector loop and switch II in mutated HRas are colored in yellow and red respectively. GTP.Mg²⁺, an activator of H-Ras is colored in pink. In Raf, Ras binding domain (RBD) is colored in blue. Mutated H-Ras and Raf are represented in “Ribbon” model and GTP.Mg²⁺ is in “Ball and Stick” model.

It has been reported the virtual screening of lead compounds for H-Ras P21 mutant protein from ChemBank and DrugBank databases using LigandFit and DrugBank-BLAST. The analysis resulted in 13 hits which were docked and scored to identify structurally active leads that make similar interaction to those of bound complex of H-Ras P21 mutant-Raf. This approach produced two different leads, 3-Aminopropanesulphonic acid (docked energy -3.014 kcal/mol) and Hydroxyurea (docked energy -0.009 kcal/mol) with finest Lipinski’s rule-of-five. Their docked energy scores were better than

the complex structure of H-Ras P21 mutant protein bound with Raf (1.18 kcal/mol). All the leads were docked¹³ into effector region forming interaction with ILE36, GLU37, ASP38 and SER39 (Fig. 2-4).

From this *in silico* study and previously reported experimental data in literature, we conclude that hydroxyurea and 3-aminopropanesulphonic acid would be an effective drug to inhibit function of mutant H-Ras P21 protein, which will in turn arrest the process of cell growth and proliferation of the cancer cell (Fig. 5). Further, the two ligand molecules can be incorporated into the drug development phases or clinical trial¹⁴.

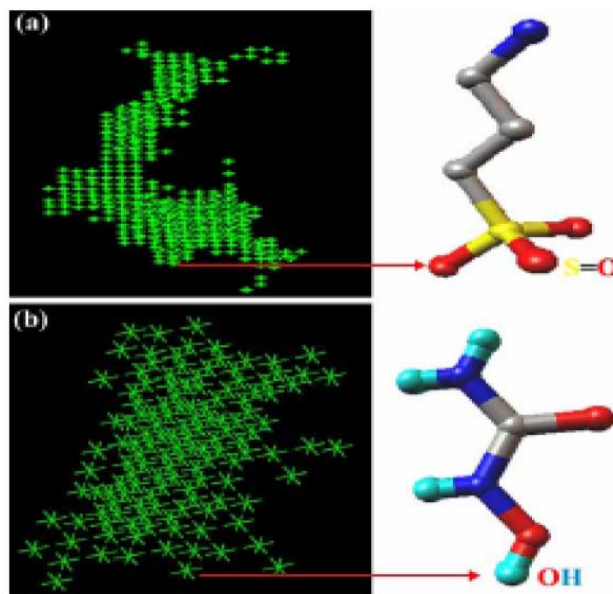


Figure 3: A schematic representation of the grid systems for binding site-I and II (Green boundary, Jacks representation). (a) Binding site-I is having 575 atoms and (b) 135 atoms in binding site-II that were embraced within effector region of Ras. The ligand molecule such as 3-Aminopropanesulphonic acid was matching exactly with the position of binding site-I and hydroxyurea matching with binding site-II in grid system. These two ligand molecules were screened by LigandFit docking algorithm.

| | Effector Region | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--------|-----------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|---|---|---|
| | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | | | | |
| Ki-Ras | T | I | Q | L | I | Q | N | H | F | V | D | E | Y | D | P | T | I | E | D | S | Y | R | K | Q | V | V | I | D | G | E | T | C | L | L | D |
| H-Ras | T | I | Q | L | I | Q | N | H | F | V | D | E | Y | D | P | T | I | E | D | S | Y | R | K | Q | V | V | I | D | G | E | T | C | L | L | D |
| N-Ras | T | I | Q | L | I | Q | N | H | F | V | D | E | Y | D | P | T | I | E | D | S | Y | R | K | Q | V | V | I | D | G | E | T | C | L | L | D |
| RAR1A | T | V | Q | F | V | Q | G | I | F | V | E | K | Y | D | P | T | I | E | D | S | Y | R | K | Q | V | E | V | D | C | Q | Q | C | M | - | - |
| TC21 | T | I | Q | F | I | Q | S | Y | F | V | T | D | Y | D | P | T | I | E | D | S | Y | T | K | Q | C | V | I | D | D | R | A | A | R | L | D |
| RIN | T | M | L | F | I | S | H | Q | F | P | D | Y | H | D | P | T | I | E | D | A | Y | K | T | Q | V | R | I | D | N | E | P | A | Y | L | - |

Figure 2: Ras effector-interaction sequences; the multiple sequence alignment of Ras family of proteins that are involved in interactions with Raf and other candidate Ras effector targets. Effector region of Ras family shows complete identity with the equivalent sequences of Rap- 1A; TC21 and RIN may represent a common interaction site for all Ras-GTP binding proteins. This multiple sequence alignment was carried out by using online ClustalW tool and the figure was generated by BioEdit.

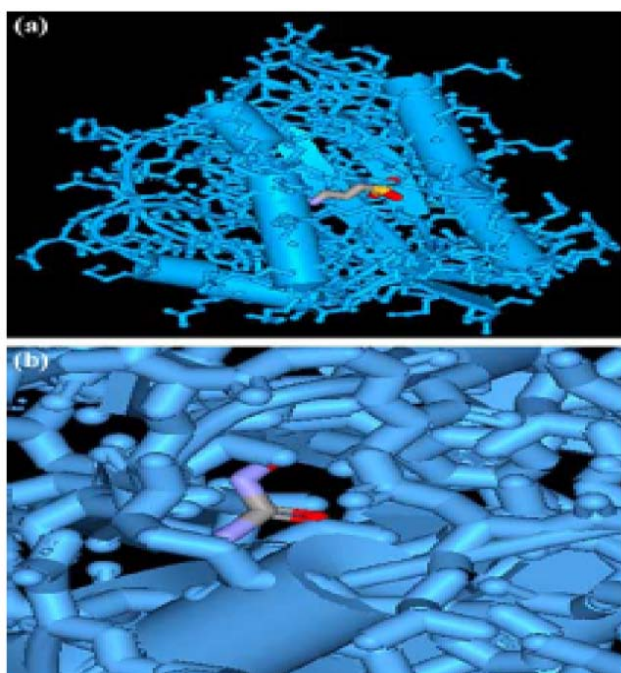


Fig. 4: Illustration of docked structure for (a) 3-aminopropanesulphonic acid in ligand binding site-I and (b) Hydroxyurea in ligand binding site-II of mutated H-Ras. This protein structure is represented as both “secondary structure schematics” and “ball and stick” model (Colored in Cyan) and ligands are in “Ball and Stick” model (Colored by CPK, Red-Oxygen, Yellow- Sulfur, Light grey-Carbon, Nitrogen-Blue, Hydrogen-Cyan). LigandFit docking algorithm was used for virtual-screening of ligands from ChemBank database and used to dock with the H-Ras. This figure was generated by DS Visualizer.

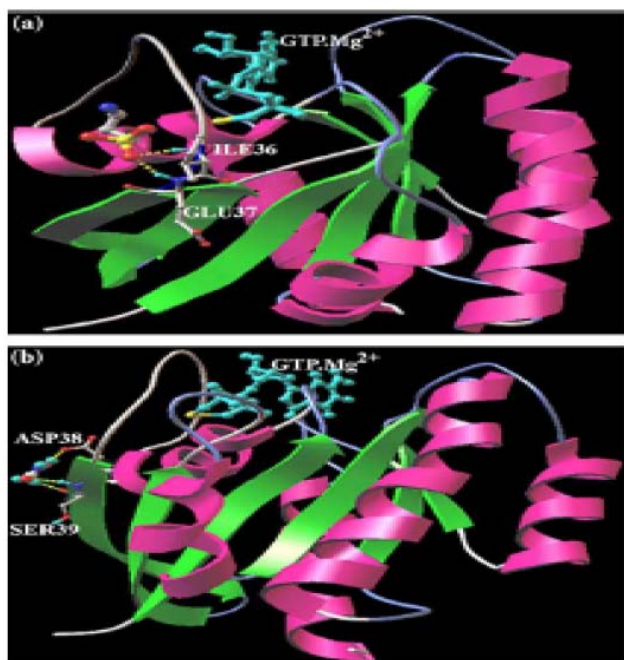


Fig. 5: Representation of re-docked structure for (a) 3-aminopropanesulphonic acid and (b) Hydroxyurea in effector region of mutated H-Ras. This mutated protein structure is represented as in a ribbon model (Colored by secondary structure, Helix- Pink, Sheet-Green) and ligand molecules are represented as “Ball and Stick” model

(Colored by CPK, Red-Oxygen, Yellow- Sulfur, Light grey-Carbon, Nitrogen-Blue, Hydrogen-Cyan). Residues such as ILE36, GLU37 of H-Ras are involved in hydrogen bond formation with 3-aminopropanesulphonic acid whereas residues ASP38, SER39 involved in hydrogen bond with hydroxyurea. Hydrogen bonds are indicated by yellow dot line and GTP.Mg²⁺ in Cyan color. These re-docking processes were carried out by using AutoDock 3.0.

CONCLUSION

Computer-Aided Drug Design is a natural outgrowth of theoretical chemistry, the traditional role of which involves the creation and dissemination of a penetrating conceptual infrastructure for the bioinformatics, chemical sciences, particularly at the atomic and molecular levels. The mathematical sciences have been indispensable allies and have provided the vital tools for that role. Computer-Aided Drug Design is central to rational drug design, it contributes to the selection and synthesis of new materials, and it guides the design of catalysts. New quantum mechanical techniques underlie the understanding of electronic properties of materials and have advanced the level of precision at which molecules of at least moderate size can be modeled¹⁵.

The marketed growth of Computer-Aided Drug Design inevitably has involved a substantial investment of skilled human resources and to expensive computing resources. Both of these types of commodities are relatively scarce and are subject to competition between alternative scientific and technological disciplines. Reports show that the heavy dependence of these chemistry codes on mathematical software such as LINPACK and EISPACK¹⁶. The productivity of these computational resources, broadly construed, must be an issue for continual analysis and informed action by policy makers. In particular, the strong mathematical flavor of Computer-Aided Drug Design leads to a natural examination of the efficacy of links between mathematical and the chemical sciences, and to the past, present, and future roles of interdisciplinary research at the interface between these subjects. The issues constitute basis concerns for the present study.

REFERENCES

1. Irwin J, Lorber D M, McGovern S L, Wei B, Shoichet B K. Docking and Drug Discovery. *Computational Nanoscience and Nanotechnology*. 2002; 2: 50-51.
2. Taft C A, Silva V B da, Silva C H T de P da. Current topics in computer-aided drug design. *Journal of Pharmaceutical Sciences*. 2008; 97 (3): 1089 – 1098.
3. Bernard D, Coop A, MacKerell A D Jr. Computer-Aided Drug Design: Structure-Activity Relationships of Delta Opioid Ligands. *Drug Design Reviews*. 2005; 2: 277-291.
4. Richards W G. Computer-aided drug design. *Pure and Applied Chemistry*. 1994; 66 (8): 1589-1598.

5. Douglas K, Bichenkova E, Bryce R, Barber J. Computer-aided drug design and computational chemistry. School of Pharmacy and Pharmaceutical Sciences. University of Manchester. 2006.
6. Casey R. Bioinformatics in Computer-Aided Drug Design. BeyeNETWORK. 2005.
7. Park D S, Kim J M, Lee Y B, Ahn C H. QSID Tool: a new three-dimensional QSAR environmental tool. *Journal of Computer-Aided Drug Design*. 2008; 22 (12): 873-883.
8. Good A C, So S S, Richards W G. Structure-activity relationships from molecular similarity matrices. *Journal of Medicinal Chemistry*. 1993; 36: 433.
9. Sams-Dodd F. Target-based drug discovery: Is something wrong? *Drug Discovery Today*. 2005; 10: 139-147.
10. Dauber-Osguthorpe P, Maunder C M, Osguthorpe D J. Molecular dynamics: Deciphering the data. *Journal of Computer-Aided Drug Design*. 1996; 10 (3): 177-185.
11. Perun T J, Propst C L. Computer-aided drug design: methods and applications. Marcel Dekker, INC., New York, USA. 1989: 2-15.
12. MacKerell A D Jr. Empirical Force Fields for Biological Macromolecules: Overview and Issues. *Journal of Computational Chemistry*. 2004; 25: 1584-1604.
13. Jayakanthan M, Wadhwa G, Mohan T M, Arul L, Balasubramanian P, Sundar D. Computer-Aided Drug Design for Cancer-Causing H-Ras p21 Mutant Protein. *Letters in Drug Design and Discovery*. 2009; 6: 14-20.
14. Valencia A, Chardin P, Wittinghofer A, Sander C. The ras protein family: Evolutionary tree and role of conserved amino acids. *Biochemistry*. 1991; 30: 4637-4648.
15. Anderson A C, Wright D L. The Design and Docking of Virtual Compound Libraries to Structures of Drug Targets. *Current Computer-Aided Drug Design*. 2005; 1 (1): 103-127.
16. Vangrevelinghe E, Rudisser S. Computational Approaches for Fragment Optimization. *Current Computer-Aided Drug Design*. 2007; 3(1): 69-83.

