Computer Aided versus Wet Lab Drug Discovery

S. K. BAGGA*, V. GARG, A. V. SINGH, R. KEITH, S. AHUJA, S. GOSWAMI AND R.V.S.V. VADLAMUDI
American-Indian Pharmaceutical Education and Research Group (AI-PEAR-GP)
7734, Effingham Square, Alexandria VA, USA.

To discover and develop a new drug it takes $ 800 million and 10–15 years and even then only one third of those reaching the market pay back, the rest of them piggy ride on the successful ones. To be in the market the leading Companies need to spend $2 billion per year and to be successful at least 4 new drugs must be generated. But at present the average is around one per year and even for that there is no guarantee that it will realize the minimum threshold sales level of $ 350 million, in the second year post-launch. It is hyped around that with new technologies involving Combinatorial Libraries, High-throughput screening and use of molecular targets based on genomics and proteomics will reduce the time and the costs involved for new drug development, but this does not seem to be. The main advantage of these new techniques is that drugs discovered would be more target and disease specific. As a part of monthly discussion AI-PEAR-GP decided to discuss this issue and members exchanged views on the SWOT analysis of wet lab drug discovery and computer aided drug discovery. A concise report on discussion on this topic is presented here.

Computer-aided drug design (CADD) is an important field that did not come into being until about 10 years ago as it was highly dependent upon faster computing system. Drug design usually requires study of larger molecules than traditional computational biology, making its computational demand even greater. It was not until a decade ago that the promises of computer-aided drug design became a reality, signaling such successes as the design of an HIV protease inhibitor at Dupont Merck[1]. Unlike many other drugs, which are usually found by assaying a large number of candidates, the HIV protease inhibitor was developed completely on a computer, by studying theoretically the molecular properties favorable to such an inhibitor, and then designing a molecule to meet the necessary requirements. This concept is thought to be the starting-point of a new generation of drug-design, a much more efficient and viable alternative to the current lab-based methods. Many scientists believe that the first cures for cancer and AIDS will come from a computer, not a laboratory.

*For correspondence
E-mail: sbagga_pear@yahoo.com

CADD is applicable to only molecular targets whose structures are well defined and X-ray data is available. Data from a molecular target that is crystallized with a ligand bound to it gives more information for designing a drug. That excludes many drug targets whose structures yet to be known from being amenable to CADD. Doman and colleagues compared the performance of HTS (WLDD) and molecular docking (CADD) in searches for inhibitors of protein tyrosine phosphatase 1B (PTP1B), a target for type 2 diabetes and their findings indicate that HTS and docking could be complementary techniques for lead discovery[2]. In the HTS experiments, 400,000 compounds from a corporate chemical library were screened against human recombinant PTP1B. Those that inhibited PTP1B at 300 μM were chosen for further evaluation, and 85 showed IC_{50} values lower than 100 μM, a cut-off value for the molecule to be considered a hit, which represents a hit rate of 0.021%.

To test the ability of molecular docking, 235,000 commercially available compounds were computationally screened against the known active site in the crystal structure of PTP1B. In this case, 365 of the highest-scoring molecules in terms of potential for binding were investigated

further in inhibition assays (using the WLDD assay), and 127 (34.8%) compounds showed IC_{50} values lower than 100 μM. So, docking enriched the hit rate by 1,700-fold over random screening. One argument against this very high hit rate could be that this hit rate represents hits out of the selected 365 compounds. In the case of HTS, 0.021% is out of the total 400 000 compounds screened. Instead, if one considered 85 hits out of those selected for further study, the percentage would have been different. Since the number of compounds chosen for further study was not given, it is necessary to express the 127 hits obtained with molecular docking method as % of the total 235 000 compounds studied. The hit rate then will be 0.054%, which is 2.6 fold greater than that of HTS method.

Analysis of the hit lists showed that hits obtained through both methods were dissimilar from the natural ligand phosphotyrosine and that is important, as phosphate-containing molecules are not good candidates for drug development. When the drug-like qualities of these hits were compared by considering how many of the hits would pass the Lipinski Rule-of-5 criteria showed that, on average, the docking hits passed 3.49 out of 4 rules compared to 2.73 for the HTS hits, giving rise to an unexpectedly large difference. This comparative data certainly encourages the use of docking and structure-based discovery approaches.

On the other hand WLDD is applicable to all sorts of targets and many resources can be screened against pure as well as crude target preparations. Although it is cumbersome it has the best track record as far as numbers of new drug introductions are counted. Furthermore, there is no escape from WLDD. Even the best design from CADD has to be synthesized and screened to see whether the predicted structures are active in reality. CADD can serve as a very powerful complimentary tool to WLDD. Although the question of whether structure-based methods can compete with random screening cannot be completely answered by the above mentioned single study, which has several gaps such as the differences between the databases that were screened by each technique, it does suggest that the best strategy could be to use both techniques.

CADD is a very big asset to the drug discovery process, but as mentioned above there are quiet a few limitations to this process, which are ultimately dealt with WLDD. Though CADD can potentially save time, manpower and money it still has to go a long way, as there is very small number of available crystal data. In addition, even the results of CADD need to be validated using WLDD.

CADD:

This technique requires a 3D structure as a template using simulations based on the basic principles of Chemistry and Biophysics. Several programs have been developed to facilitate simulations of NOVEL drug. Programs like CHARMM (Chemistry at HARvard Molecular Mechanics), AMBER (Assisted Model Building for Energy Refinement) uses the basic rules of Atomic and Molecular Interactions. Molecular Dynamics and Monte Carlo simulations have been long used for the simulations and studying interactions between ligands and receptors. The overall theory of designing new drugs is based on the information about how biomolecules react and interact, how the Biological System would react to the drug and how effective would the drug be.

WLDD:

A thorough understanding of the biochemistry of the receptors and other molecular targets is the key step. Information about properties of the proteins becomes extremely handy in designing compatible drugs. For instance if the target is a membrane protein, the drug candidate would preferably be a hydrophobic type, similarly other post translational modifications related to the target proteins such as glycosylation or phosphorylation would also give some idea about what type of target chemical moieties to be used for screening.

The major limitation of CADD is the availability of 3D structures for the target protein for drug discovery. To expedite the process of prediction of 3D structures, NIH has sanctioned grants to NINE consortiums. The basic aim of these consortiums is to develop high-throughput screening and determination of structures for the target proteins from various organisms. This is a step towards more number of 3D structures that can be used for more efficient drug discovery and development process. It is true that the person doing CADD needs to have a through knowledge of integrated science (chemistry, biochemistry, biophysics, mathematics and statistics.) and not just one specialization as the CADD apart from 3D structure mainly depends on the interpretation of the result that is displayed by the computer.

Although there are so many programs we are yet to find one program that can predict the right molecule for the target. This is because of a paucity of data from the wet lab that can be used by the programs for reference or there may be lots of basic data but not in the digital form. The availability of the basic data from the wet lab has become a

major hindrance in the development of good programs.

At the WLDD scenario the world is moving towards higher throughput screening of the compounds at every stage of drug development. This has helped the pharma industry in many ways, as they are able to screen many targets simultaneously. In addition to the technical differences that exist between CADD and WLDD, availability of resources is a major issue for CADD. In case of WLDD, there are different departments within an industry to carry out different experiments and the availability of manpower is very high and training costs are minimal. However, for CADD resource availability is rare due to the skill sets that are required and the training cost is very high compared to that of WLDD. Infact many major companies have people working on CADD but the success rate is very low.

To make CADD a success in the future the pharma industries have to store and integrate their data from various departments in order to make this data useful for future reference. Although many major pharma companies are in the process of accomplishing this integration, Indian pharma companies are lagging behind on this front. In fact, it can be viewed that the Indian pharma companies can integrate the data more easily as they have yet to start when compared to the major multinationals, which have built many database islands for as many departments and now facing a major integration issue.

Summary:

For using CADD one needs to know the molecular structure of the target protein. X-Ray crystallographic structure of the target protein preferably co-crystallized with the ligand bound to it offers the most directly usable information. If such information is not available about the target, homology modeling may provide ideas about the structure of the target that is useful in designing ligands. Using the 3D structure of the target protein, both known chemical structures and hypothetical structures can be evaluated for best fit/binding with the active site of the target protein. Hypothetical or designed structures need to be synthesized and all molecules, existing or newly designed shall be tested in the wet lab for determining IC50 values. CADD is expected to increase chances of identifying lead molecules several folds over that of random screening.

WLDD is the empirical one where one screens a variety of samples (chemical libraries and natural product extract libraries) on targets using the currently available HTS platforms. Hits are an outcome of screening randomly a large number of chemical compounds often numbering up to tens and hundreds of thousands, Hits need to be further evaluated to identify leads that can be developed into developable compounds. Here the process is not constrained by lack of information about the molecular structure of either the target protein or the sample that is being screened. WLDD as a result has universal applicability while CADD can be applied to only selected cases.

In conclusion, the strength, weakness, opportunity and threat analysis of CADD and WLDD is as given below;

Strengths:

Trained manpower and knowledge base, good network of research laboratories, rich biodiversity, well developed base pharmaceuticals industries, access to intellectual resources of NRIs in this area. Extensive clinical trials and research ~ access to vast and diverse disease populations, Biodiversity ~ India's human gene pools offer an exciting opportunity for genomics and gene based drug design.

Weaknesses:

Missing link between research and commercialization, lack of venture capital, relatively low R and D expenditure by industry, image of Indian industry -doubts about ability of Indian products to meet international standards of quality.

Opportunities:

Large market, export potential, base for contract research for international companies due to rising costs of R and D abroad, large number of patients covering wider range of diseases.

Threats:

Danger of anti-*in-silico* propaganda gaining ground, IPR policies.

The views provided here are absolutely personnel and are not influenced by any other individual or thoughts.

REFERENCES

1. Lam, P.Y.S., Jadhav, P.K., Eyermann, C.J., Hodge, C.N., Ru, Y., Bacheler, L.T., Meek, J.L., Otto, M.J., Rayner, M.M., Wong, Y.N., Chang, C-H., Weber, P.C., Jackson, D.A., Sharpe, T.R. and Erickson- Viitanen, S., **Science,** 1994, 262, 380.
2. Kirkpatrick, P., **Nature Reviews Drug Discovery,** 2002, 1, 570.

About AI-PEAR-GP

Cyber-based American-Indian Pharmaceutical Education And Research Group is offering a common platform for

expressing ideas on any aspect of pharmaceutical sciences and its allied subjects. Drug development needs a close interaction of several allied scientific areas and the American-Indian Pharmaceutical Education And Research Group is open to members of the Bioinformatics, Biotechnology, Clinical Research, Immunology, Life Sciences, Pharmacology, SAS-Programming and Pharmaceutical industry or Academia. One of the unique features of the American-Indian Pharmaceutical Education And Research Group is the monthly discussion and during December 2002 – January 2003, these group members have exchanged views on Computer Aided versus Wet Lab Drug Discovery. A concise report on this topic is presented here. In future, the American-Indian Pharmaceutical Education And Research Group has plans to conduct online mini-symposia and chat sessions. To join AI-PEAR-GP, please visit http://groups.yahoo.com/group/ai_pear_gp/ or write to sbagga_pear@yahoo.com.