Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2016, 8(8):1092-1099

Review Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Re-thinking drug discovery: *In silico* method

Deepali Gangrade*, Gauri Sawant and Ashish Mehta

Vivekanand Education Society College of Pharmacy, Chembur (E), Mumbai (MS), India

ABSTRACT

Insilico methods are the leading-edge potential tools for assessing ADME properties. These Machine learning methods have ability in allocating diverse structures and complex mechanisms, are appropriate for prediction of biological activity and therapeutic potency. Insilico is simply; Latin- in silicon (i.e. Performed using computer simulation). These newer Insilico approaches has led to easier and broader discovery of new drug , which in turn affect the success and time for carrying out Clinical trials. The In silico techniques like molecular docking, QSAR, Virtual High throughput screening, Pharmacophore, Fragment based screening are explained in this review. Efforts have been directed at broadening of application scopes and improvement of predictive performance of these methods. Here the progresses and performances as well as challenges of scrutinizing Insilico method by molecular docking of Tea leaves extracted as anti-malarial (Gallocatecin) in correlation with PLANTS® software has been illustrated as a case study.

Keywords: Insilico, computer techniques, anti-malarial (Gallocatecin), Machine learning methods, PLANTS®

INTRODUCTION

Drug discovery and development is an intense, lengthy and an interdisciplinary endeavour. It is considered as a linear, consecutive process that starts with target and lead discovery, followed by lead optimization and pre-clinical in vitro and in vivo studies to determine if such compounds satisfy a number of pre-set criteria for initiating clinical development. [1]

Drugs are indispensable for the treatment and cure of diseases. There has been a plethora of new diseases being discovered. Hence, ideal drugs are always in great demand. To meet the challenges of ideal drugs, an efficient method of drug development is demanding. The process of drug development is challenging, time consuming, expensive, and requires consideration of different aspects. To accomplish these challenges, several multidisciplinary approaches are required for the process of drug development; collectively these approaches would form the basis of rational drug design. [2]

With the advent of genomics, proteomics, bioinformatics and technologies like crystallography, NMR, the structures of more and more protein targets are becoming available. So there is a need for computational tools that can identify and analyse active sites and suggest potential drug molecule that can bind to these sites. [4]

Drug design is an integrated developing discipline which portends an era of tailored drug. It involves study of effects of biologically active compounds on the basis of molecular interactions in terms of molecular structure or its physicochemical properties involved. It studies processes by which the drugs produce their effects, how they react with the protoplasm to elicit a particular pharmacological effect or response, how they are modified or detoxified, metabolised or eliminated by the organism.

Drug discovery process-

The Development process or 'pipeline' consists of a number of distinct vital steps. It starts by selecting a disease, with further continuation with target hypothesis, lead, compound screening, and lead optimization, preclinical & clinical trials. The following chart shows pictorial representation of pipeline.

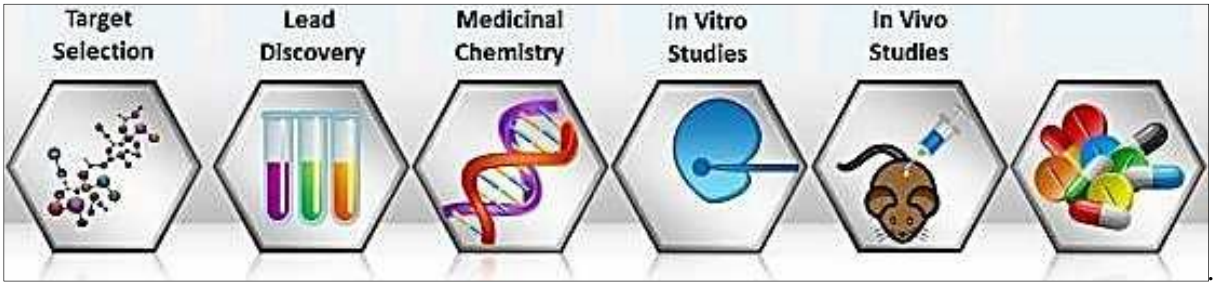


Fig. 1: Pipeline of drug discovery [2]

Cost of innovation

In 2001 Pharmaceutical research and manufacturers of America (PhRMA) estimated the cost at US$802 million over a period of 11 years from the initial research stage to the successful marketing of a new drug4. More recent estimates by DiMasi at the Tufts Center for Study of Drug Development (CSDD) that was published in 2003 put the average cost at US$802 million spread over 12 years, while the Boston Consulting Group estimates the cost as $880 million over 15 Boston Consulting Group estimates the cost as $880 million over 15 ranges from $800 million to $1.8 billion. These estimates are averages and there is significant variation in both time and cost averages and there is significant variation in both time and cost drug being developed and the nature and scope of the clinical trials required to gain regulatory approval.[4-5]

Need for modern in silico techniques[1]

- These techniques offer the advantage of delivering new drug candidates more quickly and at a lower cost.
- They increase the chance of success in many stages of the discovery process.
- They facilitate accessing huge amount of data generated.
- They transform the massive complex biological data into workable knowledge.

In SILICO

Latin- in silicon (i.e. performed using computers or via computer simulation)

Pedro Miramontes, a mathematician from National Autonomous University of Mexico (UNAM), presented the report "DNA and RNA Physicochemical Constraints, Cellular Automata and Molecular Evolution." In his talk, Miramontes used the term "in silico" to characterize biological experiments carried out entirely in a computer.

Types of Insilico approaches

1) **MOLECULAR DOCKING**

Docking is the computational determination of binding affinity between molecules (protein structure and ligand). Given a protein and a ligand find out the binding free energy of the complex formed by docking them. Docking or Computer aided drug designing can be broadly classified as;

Receptor based methods:

Uses the 3D structure of the target receptor to search for the potential candidate compounds that can modulate the target function. These involve molecular docking of each compound in the chemical database into the binding site of the target and predicting the electrostatic fit between them. Receptor based method has been successfully applied in many targets[1].

Ligand based methods:

In the absence of the structural information of the target, ligand based method make use of the information provided by known inhibitors for the target receptor. Structures similar to the known inhibitors are identified from chemical databases by variety of methods, some of the methods widely used are similarity and substructure searching, pharmacophore matching or 3D shape matching.[1].

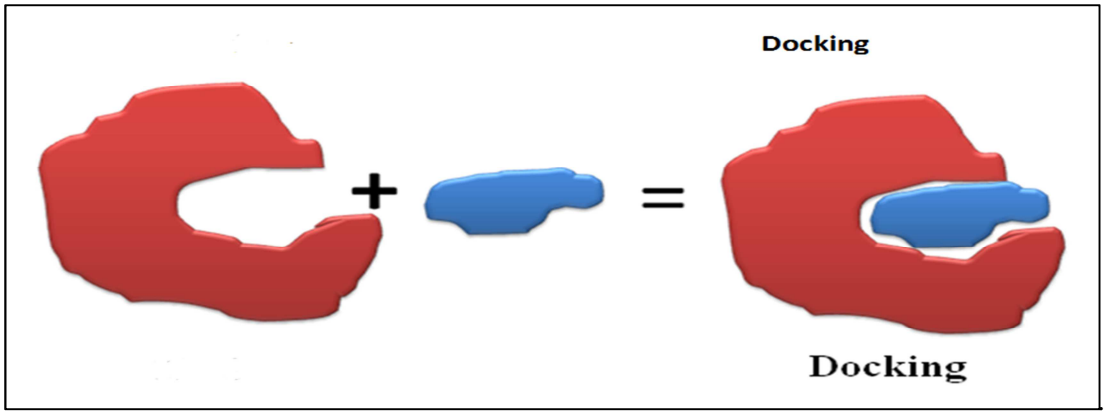


Fig.2: Shows the process of molecular docking of a certain drug in the target protein [22]

Steps involved-

Receptor Preparation:

•Dependent on docking program used•Structureselection•Siteselection•Addcharges•Often have to add hydrogens, some programs more sensitive to positions than other•Remove/include waters, cofactors, metals•Pre-dockingrefinement•Remember to consider missing residues or atoms[2].

Ligand preparation:

•Input structures (extract from PDB, draw, convert from SMILES)•Add bond orders•Generate isomers if chiralcenters•Calculate charges

–Predict pka'sfor each potential charged atom

–Generate a structure for each charge combination for a given ph range (e.g., 5-9)

•Minimize structures

–Generally using a molecular mechanics forcefield

•For Screening, can download public sets from ZINC (available compounds) or pubchem.[12].

Commercially available softwares-

–AutoDock(Art Olsen, David Goodsell, Scripps), UCSFDOCK(KuntzGroup), Glide(Schrodinger), GOLD(CCDC), FlexX(BiosolveIT),ICM (Molsoft),Surflex(Tripos).[12]

2) **VIRTUAL HIGH THROUGHPUT SCREENING**

Virtual screening is a computational method where large libraries of compounds are assessed for their potential to bind specific sites on target molecules such as proteins, and well-matched compounds tested. By using computers, it deals with the quick search of large libraries of chemical structures in order to identify those structures which are most likely to bind to a drug target, typically a protein receptor or enzyme .[1]

Virtual screening has become an integral part of the drug discovery process. Walters, et al. define virtual screening as "automatically evaluating very large libraries of compounds" using computer program. [1] VS focuses on questions like how can we filter down the enormous chemical space of over 1060 conceivable compounds to a manageable number that can be synthesized, purchased, and tested. More practical VS scenarios focus on designing and optimizing targeted combinatorial libraries and enriching libraries of available compounds from in-house compound repositories or vendor offerings. It is less expensive than High Throughput Screening, Faster than conventional screening, scanning a large number of potential drugs like molecules in very less time. [3]

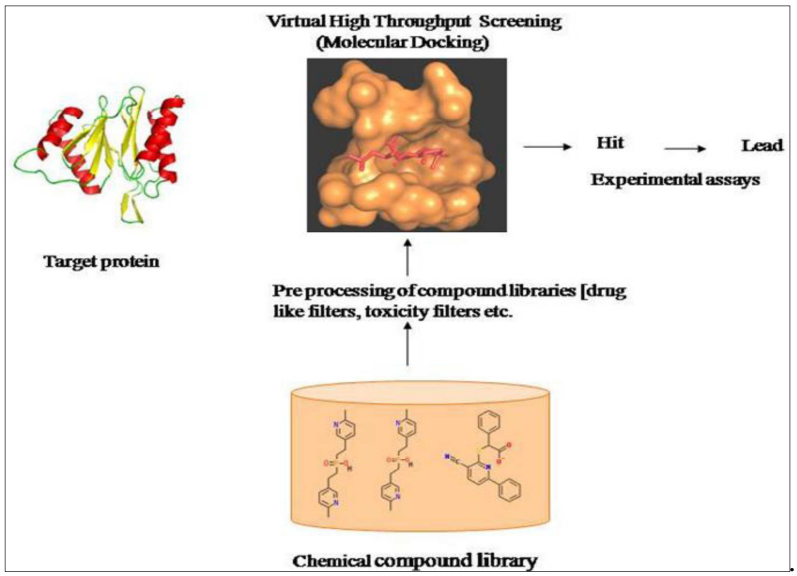


Fig.3 virtual screening[22]

QSAR(Quantitative structure-activity relationship)

QSAR is statistical approach that attempts to relate physical and chemical properties of molecules to their biological activities. The aim of QSAR is the prediction of molecular properties from their structure without the need to perform the experiment using invitro or invivo. It saves times and resources. Various descriptors like molecular weight, number of rotatable bonds, LogP etc. are commonly used. Many QSAR approaches are in practice based on the data dimensions. It ranges from 1D QSAR to 6D QSAR. The methods called quantitative structure-activity relationship (QSAR) are based on the assumption that the activity, or the property, for instance the toxic effect, is related to the chemical structure through a certain mathematical algorithm, or rule. [1][16]

Table no.1 Types of QSAR

Dimension	Methods
1D-QSAR	Affinity correlates with pK_a, molecular volume etc.
2D-QSAR	Affinity correlates with structure motifs.
3D-QSAR	Affinity correlates with a 3D-structure of the ligand.
4D-QSAR	Ligands are represented as an ensemble of conformers, orientations, protonation states, tautomers and stereoisomers.
5D-QSAR	Like 4D, with additional consideration of different induced-fit models
6D-QSAR	Like 5D, with additional consideration of different solvation scenarios

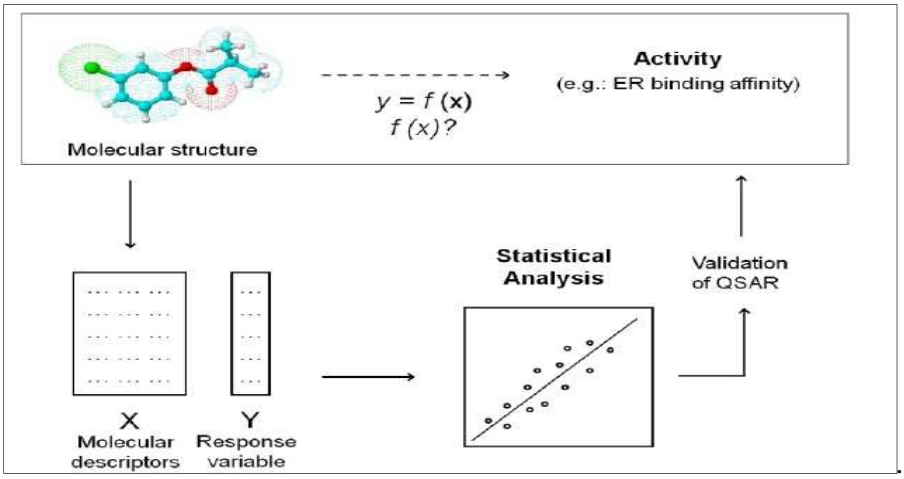


Fig.4 QSAR depiction[22]

For instance, it is well known that if in the chemical compound there are certain groups, like an aromatic amine, or an epoxide, there is a higher probability that the chemical compound is genotoxic. The basic assumption is that there

is a mathematical function of the chemical properties which is related to the effect. Thus, the effect called y is a function called f of the chemical properties, called x. mathematically, y = f(x). For each chemical compound calculate a series of parameters, called chemical descriptors. Then find an algorithm that provides a quite accurate value, similar to the real experimental value. The final step is to check if the so-obtained algorithm is capable to predict the property values for other chemicals, not used to build up the model. This last phase is called validation of the QSAR. Indeed, it is very important to generate a model which is working not only for the chemical substances used within the training set, but also for other chemicals. The challenge is to define the correct statistical properties of the model.[1]

4) PHARMACOPHORE MAPPING

It is the process of deriving a 3D pharmacophore. A pharmacophore is a set of features together with their relative spatial orientation that are thought to be capable of interaction with a particular biological target such as Hydrogen bond donors and acceptors, positively and negatively charged groups, hydrophobic regions and aromatic rings. It depends on atomic properties rather than element types, it does not depend on specific chemical connectivity. It has conformational flexibility and mapping the different combinations of pharmacophoric groups in the molecule.

A Pharmacophore map can be generated by superposition of active compounds to identify their common features. Based on the pharmacophore map either de novo design or 3D database searching can be carried out. Frequently small molecules with very different 2D structures displace each other from a binding site on macromolecules. Even more often, mono modification of the structure of an active molecule renders it inactive. Such structure bioactivity relationships are an indirect probe of the 3D structure and chemical properties of the macromolecular recognition site for the ligands. [2][17]

The goal of pharmacophore mapping is to transform such 2D structure-activity information into the 3D requirements for binding to the target biomolecule. This allows one to search 3D databases for other molecules that match these 3D properties or to design new active molecules. A pharmacophore map identifies the bioactive confirmation of each active molecule and indicates how to superimpose, compare in 3D, the various active compounds. The map identifies which types of points match in what conformation of the compounds. The decisions as to the required points and the bioactive conformations are interdependent. i.e. the choice of one affects the choices available for the other. A pharmacophore features include hydrogen bond acceptor atoms, hydrogen bond donor atoms, hydrogen bond donor site, hydrogen bond acceptor site, and hydrophobic centers.[1]

For example-
2ZIS-NH8903 OVERLAID ON PHARM-b HYPOTHESIS. HY- HYDROPHOBIC, RA- RING AROMATIC, HBA-HYDROGEN BOND ACCEPTOR, ZB- ZINC BINDER. [18]

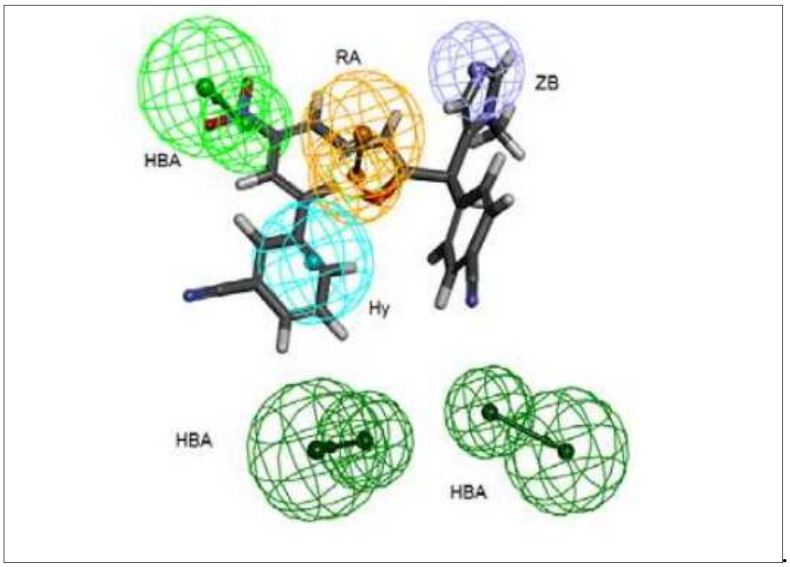


Fig.5 pharmacophore mapping [18]

5) FRAGMENT BASED SCREENING

Fragment-based lead discovery (also referred to as needles, shapes, binding elements, seed templates or scaffolds) is a new lead discovery approach in which much lower molecular weight (120–250Da) compounds are screened

relative to HTS campaigns. Fragment-based hits are typically weak inhibitors (10µM–mM), and therefore need to be screened at higher concentration using very sensitive biophysical detection techniques such as protein crystallography and NMR as the primary screening techniques, rather than bioassays. Compared with HTS hits, these fragments are simpler, less functionalized compounds with correspondingly lower affinity. However, fragment hits typically possess high 'ligand efficiency' (binding affinity per heavy atom) and so are highly suitable for optimization into clinical candidates with good drug-like properties [1, 19].

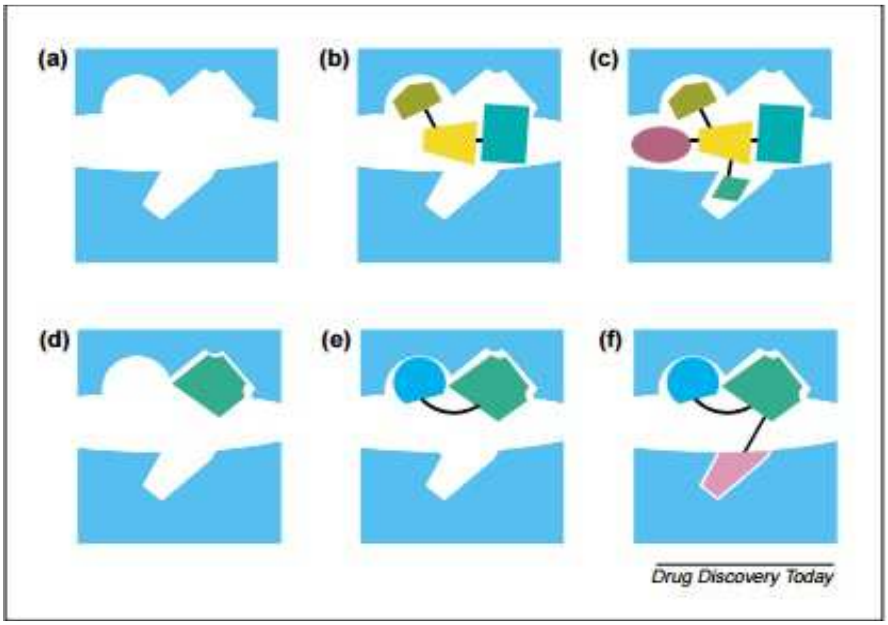


Fig.6 Schematic representation of 'drug-like'HTS hits and fragments as start points for drug discovery

There are now increasing numbers of examples appearing in the literature that demonstrate that fragment-based discovery can identify quality leads for targets where HTS has not succeeded . The second benefit, establishing that a fragment-based approach increases drug discovery efficiency, will by necessity take longer to establish. It can be argued that published fragment-based leads with high ligand efficiency and good lead-like physical properties are higher quality leads than most HTS derived leads, but ultimately this is a subjective judgment, and probably the best assessment of the quality of a lead is the ability to progress it efficiently into a clinically successful compound. Although we have already seen the first clinically successful compound from this approach, further such successes over the next few years will be required before the full potential of this new lead discovery approach can be established. [19]

Tea leaves extracted as anti-malaria based on molecular docking PLANTS [5]

Aim:

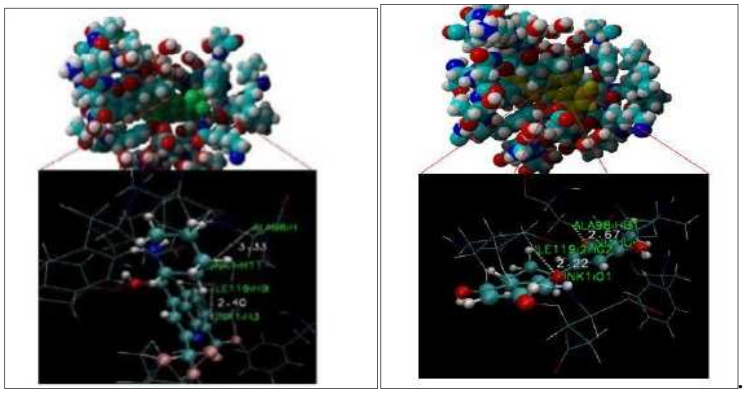
To find natural compounds having potential as anti-malarial agents which are more potential than mefloquine.

Theory:

Malaria, a form of P.Falcifarum, is an infectious disease which is often occurred. Mefloquineas a synthetic drug with anti-malarial activity is selective inhibitor with lactate dehydrogenase mechanism. Inhibition of glycolysis is needed for cell survival. Meanwhile, gallocatecinis kind of flavonoids contained in tea leaves extract (Camellia sinensis).Based on molecular docking, gallocatecinhas more potent anti-malarial activity than mefloquine.

Preparation starts with performing molecular docking with PLANTS. Ligand preparation is applied using Marvin sketch by drawing gallocatecincompound and then it is optimized. YASARA program is used for protein preparation. Removing docking protocol (including water if essential) is not required. After all of preparation is completed, docking PLANTS is applied. Decrease in score indicates bond stability with protein.

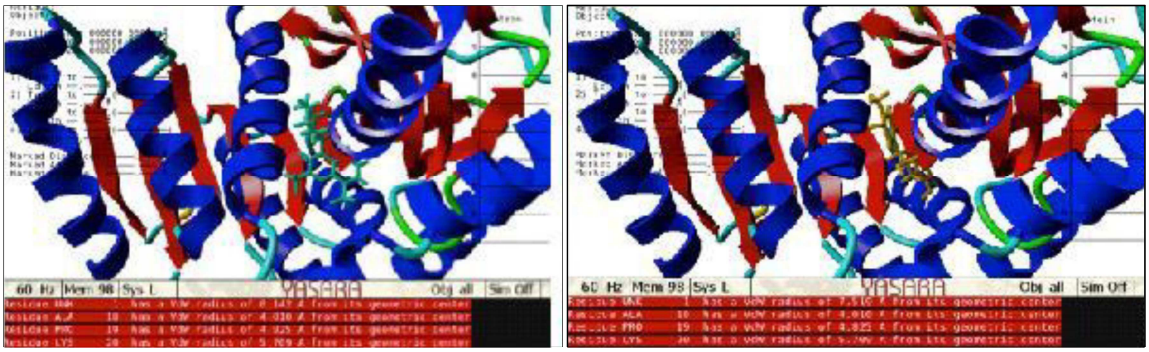
DOCKING SCORES BY USING PLANTS SOFTWARE:



7(a) **7(b)**

Binding energy of Gallocatecin to 1CET: **-95.396**

Binding energy of Mefloquine to 1CET: **-71.4036**



7 (c) **7 (d)**

Fig.7 (c) Interaction of mefloquinewith enzyme lactatdehydrogenase
Fig.7 (d) Interaction of gallocatecinwith enzyme lactatdehydrogenase

From figure 7(c) & 7(d), it can be seen that mefloquine has distance about 8142Å.It is calculated from the centre or midpoint binding site NADH. It is assumed that the central pocket is the most stable bond. Closer the distance of the ligand to pocket, more stable bond between ligand and active amino acid.

Gallocatecinis known to have shorter distance 7519Å than mefloquine. Moreover, NADH binding site, in fact, is a protein containing many amino acids; the active amino acids of lactate dehidrogenaseare Ala 98 and Ile 119. Assuming that the distance is different between ligand and the receptor, the closer distance, the more stable the bond. From figure 1.1, we can see that the distance of gallocatecinbond with the amino acid Ala 98 and Ile 119 respectively 2.67Å & 2.22Å while the distance between mefloquine are 3.33Å & 2.40Å. It means that gallocatecin have more stable bond than that mefloquine to bind to lactate dehydrogenase.

CONCLUSION

Gallocatecinhas smaller energy than mefloquine to bind to 1CET. Thus, it can be concluded that Gallocatecinh as more potential anti-malaria activity than mefloquine, based on docking molecular. [20]

Applications[1]

- It can be used to analyse the target structures for possible binding/ active sites.
- Generation of potential molecules.
- Investigate for their drug likeness.
- Dock these molecules with the target.
- Rank them according to their binding affinities.
- Optimization the molecules to improve binding characteristics.

Limitations

- "Sequence implies the Structure and Structure implies the Function"[21]
- Selected protein structures from databases such as PDB, FSSP, SCOP or CATH after removing proteins with high sequence similarity act as structural templates for the alignment.[6]
- These computational models often represent only fractions of the full length of desired protein.[6]

Oftenly a drug's market price is high, this is not because of the manufacturing cost of the sole drug but the price of failed drugs is also added to it. In such scenarios, in silico studies can be of great help as they can reduce the production cost and thereby the marketed price.

They represent a way for industry to spend less for toxicological research, or can be used to save animals to be used for experiments.

The real challenge is not to identify the best method to protect human beings and environment. The challenge is to take advantage of all the contributions that each approach, in vivo, in vitro, and Insilico offer.

REFERENCES

[1] Science, C., & Engineering, S. (**2013**). *Insilico Methods in Drug Discovery - A Review*, 3(5), 680–683.
[2]. Mandal, S., & Mandal, S. K. (**2009**). *European Journal of Pharmacology*, 625(1-3), 90–100.
[3] Discovery, D. (**2002**). *Chemoinformatics and Drug Discovery*, 566–600
[4] Bharath, E. N., Manjula, S. N., & Vijaychand, A. (**2011**). In silico drug design - tool for overcoming the innovation deficit in the drug discovery process, 3(2), 6–10.
[5] Dimasi, J. A., Hansen, R. W., & Grabowski, H. G. (**2003**). The price of innovation : new estimates of drug development costs, 22, 151–185. http://doi.org/10.1016/S0167-6296 (02)00126-1.
[6] Dror, O., & Shulman-Peleg, a. (**2004**). Predicting molecular interactions in silico: I. *Current Medicinal*, 1–57.
[7] Carr, R. a E., Congreve, M., Murray, C. W., & Rees, D. C. (2005). *Drug Discovery Today*, 10(14), 987–992. http://doi.org/10.1016/S1359-6446(05)03511-7
[8] Mishra, P., Tripathi, V., & Yadav, B. S. (**2010**). Insilco QSAR modeling and drug development process, 1(December), 37–40.
[9]Mishra, V., Kashyap, S., & Hasija, Y. (**2015**) *Journal of Taibah University for Science*, 9(1), 20–26. http://doi.org/10.1016/j.jtusci.2014.04.007
[10]Nicolotti, O., Benfenati, E., Carotti, A., Gadaleta, D., Gissi, A., Mangiatordi, G. F., & Novellino, E. (**2014**). *Drug Discovery Today*, 19(11), 1757–1768. http://doi.org/10.1016/j.drudis.2014.06.027
[11]Potshangbam, A. A. M. A., Tanneeru, K., Reddy, B. M., & Guruprasad, L. (**2011**). *Bioorganic & Medicinal*, 21(23), 7219–7223
[12]Sanschagrin, P. (**2010**). An Introduction to Molecular Docking What is Docking ?
[13]Chen, Y. C. (**2015**). Beware of docking! Trends in Pharmacological Sciences, 36(2), 78–95. http://doi.org/10.1016/j.tips.2014.12.001
[14]Selzer, D., Neumann, D., Neumann, H., Kostka, K.-H., Lehr, C.-M., & Schaefer, U. F. (**2015** *European Journal of Pharmaceutics and Biopharmaceutics*, 95, 68–76. URL: http://doi.org/10.1016/j.ejpb.2015.05.002
[15]Tao, L., Zhang, P., Qin, C., Chen, S. Y., Zhang, C., Chen, Z., Chen, Y. Z. (**2015**). *Advanced Drug Delivery Reviews*, 86, 83–100.
[16] "Chapter 8 : Introduction to QSAR", *Introduction to the Principles of Drug Design*, By H. J. Smith, Hywel Williams, page no. 213-241
[17] G. Wolber, "Structure-Based 3D Pharmacophores : An Alternative to Docking ? Abstract & Outline."
[18] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3816728/
[19] R. a E. Carr, M. Congreve, C. W. Murray, and D. C. Rees, *Drug Discov. Today*, 10, 14, 987–992, **2005**.
[20] M. Tegar and H. Purnomo, *Procedia Environ. Sci.*, vol. 17, pp. 188–194, 2013.
[21] C. L. Gupta, S. Akhtar, and P. Bajpai, *IN SILICO PROTEIN MODELING* :," pp. 513–515, **2014**.
[22]URL: http://www.authorstream.com/templates/powerpoint-diagrams-drug-discovery-ppt-design-7332/
[23]URL: http://www.intechopen.com/books/an-integrated-view-of-the-molecular-recognition-and-toxinology-from-analytical-procedures-to-biomedical-applications/computer-based-methods-of-inhibitor-prediction
[24] J.Drews, *Drug Discovery : a historical perspective*, *Science* 287 (**2000**), 1960-1964.
[25] S.Ekins, B.J. Ring , J.Grace,D.J. Mcrobie-Belle,S.A. Wrighton , J.*Pharmacol. Toxicol.Methods* 44 (2000) 313-324.