Open access Journal International Journal of Emerging Trends in Science and Technology
Impact Factor: 2.838 DOI: http://dx.doi.org/10.18535/ijetst/v3i03.09

Methods for Protein Structure Prediction and Its Application in Drug Design Using Hidden Markov Model

Authors
Nidhi Katiyar[1], Ravindra Nath[2]
[1]Dr A.P.J. Abdul Kalam University
Email: *Nidhi26kanpur@gmail.com*
[2]University Institute of Engineering Technology, CSJM University Kanpur India
Email: *rnkatiyar@gmail.com*

Abstract
Drug design and drug innovation are critical importance in human fitness. To design a drug must successfully to the compound target from the substitute structures present in the organism. Many traditional methods are used to design a drug in the laboratory. Now day computational methods have become a major role in the drug design. A structure-based drug design is so complemented because structure-based drug design uses the 3-dimensional structure of protein. To design the candidate drug that is predicted to bind with high affinity and selectivity to the target. For prediction the new drug structure many methods are used like artificial neural networks (ANN), fuzzy neural networks and hidden Markov Model (HMM). All of these methods require the identification of peptide binding (chain of amino acid) cores for model building. HMM modeling has become more popular in the all area of applications from last several years because the models are very rich in mathematical structure and also theoretical structure. HMM also play an important role in trans-membrane region prediction and trans-membrane topology prediction in drug design. A computational base Hidden Markov Model became recently important among bioinformatics research and many software tools are based on them.

Keywords: *Structure based drug design (SBDD); Hidden Markov Model (HMM), Protein structure prediction (PSP); Drug Design (DD); 3-dimensional structure (3DS)*

Introduction

To design a new, effective and safe drug has become increasingly sophisticated. The design of new drug is also "competitiveness" and "high cost" in the present market. Drug design and drug discovery is common uses 3D structure of the target macromolecule (DNA, RNA and protein). For most modern drug discovery projects start with protein identification and verification to find and verified drug target. Two methods are used for three-dimensional structure-based drug design. One is determined experimentally by using either x-ray crystallography other method is nuclear magnetic resonance NMR spectroscopy. Now days computational approaches have become a major part of structure based drug design. Structure-based drug design utilizes the 3-dimensional structure of a protein target to design the drug candidate that is calculating to bind with high affinity and selectivity to the target [1]. For most computer approaches are now being developed to reduce the cost and cycle time for discovering a new drug. In order to appreciate the drug target directed in silico approaches in drug discovery and development [2].

The research and development cost to designing a new drug is also very high. 3D

Structure based drug design incorporated a number of software tools into the ligand design process. These tools started to design a new

algorithm to discover and analyse paralogs and orthologs in genome databases. It also analysis multiple sequence alignments (in order to uncover family specific sequence motifs), to docking post-processing tools (able to abstract family specific interaction patterns from docking calculations), and to algorithms for the generation of receptor-specific scoring functions (to be used in virtual screening) or combinatorial library design [3].

In modern drug design many drugs are design using with the effects of choosing of biological macromolecules for example these enzyme are deoxyribonucleic acid (DNA) ribonucleic acid (RNA), glycoproteins, hormones, receptors and transcription factors, which are regarded as drug targets. It is known that in most of the cases, drugs exert their functions by interacting with their targets mainly by non-covalent bonds such as van der Waals interactions, the same hydrogen bond interactions, and electrostatic interactions. Only in few instances are covalent interactions formed [4].

Recent data shows that to discover a new drug would take more than 10 years and cost more than 200 million U.S. dollars. In this stage research and development (R&D) took 10 years, for the marketing period would only left 10 years. If we reduce R & D time as less as possible and cost which would spend on it. Resulting that a longer precious time and cost for the exclusive marketing would left. After more than 150 years of drug design, development and discovery of a new drug is still a long and expensive process while it has become much more competitive.

2. Structure Based Drug Design

2.1 Surface Representation

Protein structures are used without modification, downloaded from the Protein Databank (PDB) of internet. It has different heteroatoms, which includes all waters and cofactors, are ignored. Hydrogen molecules are normally not used in the purpose of surface areas. Using this precedence here and analytic representations of the macromolecular surface is generated. It involves first constructing the 3D weighted Delaunay tessellation of the bio molecule and then subtracting the alpha shape complex. [8]

2.2 Receptor Structure Based Drug Design

We are focus on those chemicals that are more possible to be drug leads, which is fulfilled by rational drug design approaches for improving the efficiency to discover the drug design. If we have an exact drug target and its 3D structure are known, receptor structure based drug design can be conducted. With the rolling of molecular biology, X-ray crystallography and NMR techniques, the structures of many drug targets have been determined. More structures of drug targets can be modelled using homology-based methods. Molecular modelling techniques are first applied to infer the mechanism of interaction between the target and its ligands it is based on the 3D structure of the macromolecule receptor [9-12]. We are divided such techniques basically two types, namely the “whole-molecule method” and the “connection method”. The first method mainly trusts on the molecular docking technique is called whole-molecule method. It searches an entire 3D structure database of small molecules to find supposed drug for a specific therapeutic target. In this course, docking single or multiple small molecules in single or multiple conformations to the receptor binding sites of the target is attempted, in edict to find the best supposed ligand-receptor complex conformation [13-20]. In second method such as a “Connection methods” work increasingly like building a house by bricks. This is based on the greedy search method often used in mathematical optimization techniques. Many drug design tools have been used to developed, such optimization technique example as CLIX [21], LUDI [22], CAVEAT [23], LEGEND [24], and MCDNLG [25].

2.3 Trans-Membrane Region Prediction (TMHMM)

Different servers TMHMM, SOSUI, HMMTOP and TMpred servers were accessed to validate the TM region [26,27]. TMHMM, a new membrane

protein topology prediction method, is based on a hidden Markov model [28].

3. Hidden Markov Model

A Hidden Markov Model (HMM) is a stochastic model that uses the statistical properties of observed real world data. A good HMM accurately models the real world source of the observed data and has the ability to simulate the source. Machine Learning techniques based on HMMs have been successfully useful to applications including speech recognition, optical character recognition, and as we will examine problems in computational biology [29].

Hidden Markov model used for applications to more complex process, including speech recognition and computational gene finding. A generalized Hidden Markov Model (HMM) contains of a finite set of states, an alphabet of output symbols, a set of state transition probabilities and a set of emission probabilities. The emission probabilities require the distribution of output symbols that may be released from each state.

It is important to note that in most cases of HMM use in bioinformatics a fictitious inversion occurs between causes and properties when dealing with emissions. For example, one can synthesise a (known) polymer sequence that can have different (unknown) features along the sequence. In an HMM one must select as emissions the monomers of the sequence, because they are the only known data, and as internal states the features to be estimated. An excellent case is provided by the polypeptides, for which it is just the amino acid sequence that causes the secondary structures, while in an HMM the amino acids are expected as emissions and the secondary structures are supposed as internal states [31-33].

3.1 Three Main Problem of HMM

The three main types of problems occurring in the use of Hidden Markov Models are:

A) Evaluation problem: Given the observations sequence O = O_1, O_2…….O_T, and a model lemda = (A,B, π), how do we efficiently compute P(O| lemda).

B) Decoding problem: Given the observations sequence O = O_1, O_2…….. O_T, and a model lemda = (A, B, π), how do we choose a corresponding state sequence.

C) Learning problem: How do we adjust the model parameters lemda = (A, B, π). To maximise P (O| lemda).

For example

The set of internal states is *U* = ('*c*','*n*') where '*c*' and '*n*' stand for the coding and non-coding internal states and the set of emissions is the set of the four DNA bases: *X* = (' *A*','*T*','*C*','*G*')

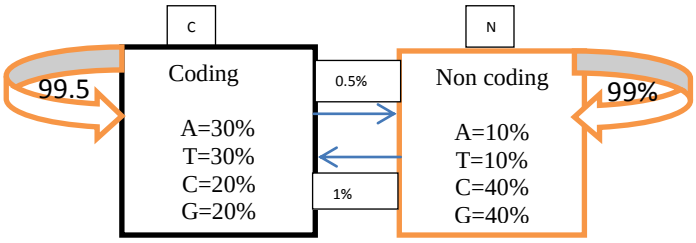

Figure 1: An Example of HMM

a) The square boxes represents the internal states 'c' (coding) and 'n' (noncoding), inside the boxes there are the probabilities of each emission ('A', 'C', 'T' and 'G') for each state; outside the boxes four arrows are labelled with the corresponding transition probabilities.

ATTACGTTGACATTAGCAATATCATAGAA
CAAATCATCGGGGCAGGATACCGCCGACC
TGCAGGG
ccccccccccccccccccccccccccccccccccccccccccc
ccccccccccnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnn

b) The first row is a sample sequence of 65 observed emissions and the second row is one of the likely of sequences of internal states.

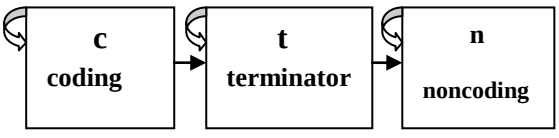

Figure 2. An example of HMM with three internal states. The square boxes represent the

internal states 'c' (coding),'t' (terminator) and 'n' (non coding). The arrows indicate the possible transitions.

Another use of HMMs can also be considered as special instances of Machine Learning Techniques that are often alternatively used for similar applications. A list of such numerous techniques could include, besides HMMs, also: Decision Tree, Support Vector Machines (SVM), Artificial Neural Networks (ANN), Clustering, Genetic Algorithms, Association Rules and Fuzzy Sets etc. Obviously each one of these techniques has advantage and disadvantage, often depending on the problem at hand. In somewhat rough terms we can say that the merits of HMMs in bioinformatics are demonstrated by their wide use. Other popular technique in bioinformatics are ANNs and SVMs. Certainly a detailed comparison of the main techniques, either at conceptual or at benchmark level. In the other hand most available comparisons are too sharply focused on very narrow subjects. In general terms we can say that the main advantages of HMMs are often the ease of use, the fact that they typically require much smaller training sets, and that the observation of the inner structure of the model provides often a deeper understanding of the phenomenon. Among the main drawbacks of HMMs is often their greater computational cost. We note that frequently hybrid models are designed combining some of the above techniques, typically with results better than with stand-alone techniques. For example, HMMs are also used for bioinformatics predictions together with the so-called Support Vector Machine (SVM), a technique based on the Vapnik-Chervonenkis theory that produces decision surfaces in multidimensional spaces.

3.2 Major Bioinformatics Application

In bioinformatics, many algorithms based on HMMs have been useful to biological sequence analysis, as gene finding and protein family classification. A technical description of HMMs and their applications to bioinformatics are the followings [35,36].

3.2.1 Genetic Mapping

One of the earliest applications of HMMs in bioinformatics (or even the first, as far as we know) has been the use of a nonstationary HMM for genetic mapping i.e. the estimation of some kind of distance between loci of known (or at least presumed) demand along the chromosome [37].

3.2.2 Gene Finding

The term "gene finding" indicates the act of finding genes within a DNA sequence, but is often used with a more common meaning of labeling DNA tracts, for example labelling them as coding, intergenic, introns, etc. In this last sense gene finding can be considered a special case (the most important in bioinformatics) of the more general action known as sequence labeling (also for non-DNA sequences) [38].

3.2.3 Secondary Structure of Protein Prediction

HMMs are also employed to predict the secondary structure of a protein (i.e. the type of the local three-dimensional structure, usually alpha-helix, beta-sheet, or coil), an important step for predicting the inclusive three-dimensional structure. Asai et al. [39] first used a simple HMM for the secondary structure prediction, while Goldman et al. [40] in the HMM approach exploited broken some evolutionary information contained in protein sequence alignments.

3.2.4 Signal Peptide Prediction

Signal peptide prediction, i.e., the determination of the protein purpose address contained in the peptide first tract is often of dominant importance both for diseases enquiry and for drug design [41].

3.2.5 Transmembrane Protein Prediction

It is well known that a direct amount of the complete 3D structure of a trans-membrane protein is now possible only in very few cases. On the other hand, for many useful purposes (such as

drug design), it is already very useful to simply know at least the trans-membrane protein topology (i.e., whether a tract is cytoplasmatic, extracellular, or trans-membrane); and to this end a number of models are presented to predict such topology. The secondary structure of the trans-membrane areas of most proteins (the helical trans-membrane proteins) is of alpha helix type; important exceptions are the so-called beta-barrels (bundles of trans-membrane beta-sheet structures), restricted to the outer membrane of Gram-negative bacteria and of mitochondria [42,43].

3.2.6 RNA Secondary Structure Prediction

The non-coding RNA builds stable and physiologically important secondary structures (typically absent in coding RNA). Such structures are usually stabilised by palindromic tracts, so that predicting the secondary RNA structures essentially amounts to identifying palindromic sequences [44].

4. Methods for Protein Structure Prediction and its Application in Drug Design

In the biological process protein is an essential component. Proteins are responsible for different biochemical reaction like catalyzing and regulations etc. In structure-based drug design three dimensional structure of the protein needs to be determined experimentally. Two main approaches in determination of protein 3D structure are: Ab initio prediction and comparative modeling.

4.1 Comparative Modelling

The amino acid sequence of a protein is known as its primary structure, while local confirmation in this sequence namely alpha helices, beta sheets, and random coils are known as secondary structures. 3D structure of proteins is known as tertiary structure [26]. Comparative protein structure modelling constructs a three-dimensional model of a given protein sequence based structure. It is supported out in four sequential steps: finding known structures (templates) related to the sequence to be modelled (target), aligning the target sequence with the templates, building the model, and assessing the model. Therefore, comparative modelling is only applicable when the target sequence is detectably related to a known protein structure [45, 46].

4.2 3D Structure Generation by Using Modeller

It is a computer program used for comparative modelling of protein 3D structures. The alignment of a sequence to be modelled is delivered with known related structures. Modeller automatically calculates a model containing all non-hydrogen atoms. Modeller implements comparative protein structure modelling by approval of spatial restraints. The sequences requires of known 3D structure and the target having more than 35% of similarity in homology modelling [47].

4.3 Validation

The best validation is the process of evaluating reliability for 3-dimensional atomic models of large biological molecules such as proteins and nucleic acids chains common sense, biological knowledge and results from analytical tools. These validations provide 3D coordinates for each atom in the molecule come from structural biology experiments such as x-ray crystallography or nuclear magnetic resonance (NMR). Most alteration involves changing the alignment. PROCHECK [48, 49] is used to calculate the main chain torsion angles, i.e. the Ramachandran plot [50] for our predicted structures. Three models were predicted using different templates among those the one that shows the good resolution factor and R-factor. It was used as a template and evaluated by Procheck performing full geometric analysis with a resolution of 1.5 Å. The validation for structure models found from the three software tools was performed by using PROCHECK [51].

4.4 Drug Development Based on Protein Structure

The main object of drug design is to find mostly small, drug molecule that tightly binds to the

target protein. It moderating its function or competing with natural substrates of the protein. Such a drug can be best found on the basis of knowledge of the protein structure.

If the spatial shape of the site of the protein is known, to which the drug is supposed to bind, then docking methods can be applied to select suitable lead compounds that have the potential of being refined to drugs.

4.5 Protein Structures Modelling

The detection of selectivity sites within protein families and the computational search for putative selective ligands by virtual screening require, in most cases, of the ability to generate high quality homology models for some of the members in the protein family. Recent studies have shown that while homology models are useful in virtual screening, improvements in their quality are still required [52].

The procedure starts with the known structure of one or more templates, from which several preliminary homology models of the target are generated. Ligands are then docked into an averaged binding-site representation of the binding site models, and new homology models are obtained considering explicitly the docked ligands by transforming the ligand information into user-defined restraints. Ligand supported homology models are selected as the ones that best explain the observed ligand-binding affinities [53,54].

4.6 Protein Flexibility in Docking

Developments in our capacity to model selectivity are necessarily is based on our skill to faithfully model the protein-ligand recognition process. Protein flexibility is essential in this process. First attempts to consider protein flexibility in docking used changed energy functions with soft van der Waals interactions (soft docking) [55].

4.7 Docking

Docking is a process which predicts the preferred location of one molecule to a record when certain to each other to form stable complex knowledge of the favourite locations in turn may be used to predict the binding power of association or binding affinity between two molecules [56]. Docking is frequently used to predict the binding locations of small molecules drug applicants to protein targets in knowledge to in turn predict the affinity and activity of the small molecule [57]. The improvement and application of a range of molecular docking algorithms based on different search methods were observed in the last few years. This method has had several recent successes in drug discovery.

5. Conclusion

Computational approaches for protein structure prediction are silent in the presently stage of progress. It has homology-based prediction technique that become exclusively helpful in an environment where the different others techniques can be used. Using computational techniques and algorithms found better performance with experimental and functional determination of protein. The HMM are also working to predict the secondary structure of a protein (i.e. the type of the local three-dimensional structure, usually alpha-helix, beta-sheet, or coil). It also predict overall three-dimensional structure of proteins. HMM is also used to predict the Trans-membrane region and Trans-membrane topology of protein structure. We determine the method on known drug targets and find that the method which is largely successful methods and techniques for new drug design using computers and computational techniques. Computational methods provide the advantage of new drug design candidate faster and at a lesser price.

References

1. Afergan E, Najajreh Y, Gutman D, Epstein H, Elmalak O, et al. (2010) 31P-NMRand Differential Scanning Calorimetry Studies for Determining Vesicle's Drug Physical State and Fraction in Alendronate Liposomes. J Bioanal Biomed 2: 125-131.

2. Marshall, G.R., Computer-aided drug design, Annu Rev PharmacolToxicol, 27 (1987) 193-213.
3. Takeda-Shitaka, M.; Takaya, D.; Chiba, C.; Tanaka, H.; Umeyama,H. Protein structure prediction in structure based drug design. *Curr.Med. Chem.* 2004, *11*, 551-558.
4. Loew, G.H., Villar, H.O. and Alkorta, I., Strategies for indirect computer-aided drug design, Pharm Res, 10 (1993) 475-86.
5. Langer, T. and Hoffmann, R.D., Virtual screening: an effective tool for lead structure discovery? C*urr Pharm Des*, 7 (2001) 509-27.
6. Kenny, B.A., Bushfield, M., Parry-Smith, D.J., Fogarty, S. and Treherne, J.M., The application of high-throughput screening to novel lead discovery, P*rog Drug Res*, 51(1998) 245-69.
7. Jackson, R.C., Update on computer-aided drug design, CurrOpinBiotechnol, 6 (1995) 646-51.
8. Coleman, R. G.; Burr, M. A.; Souvaine, D. L.; Cheng, A. C. An Intuitive Approach to Measuring Protein Surface Curvature. *Proteins*2005, *61*, 1068-1074.
9. Veselovsky, A.V. and Ivanov, A.S., Strategy of computer-aided drug design, C*urrDrugTargets Infect Disord*, 3 (2003) 33-40.
10. Vedani, A., [Computer-Aided Drug Design: An Alternative to Animal Testing in the Pharmacological Screening], A*ltex*, 8 (1991) 39-60.
11. Ooms, F., Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry, Curr Med Chem, 7 (2000) 141-58.
12. Cohen, N.C. and Tschinke, V., Generation of new-lead structures in computer-aided drug design, P*rog Drug Res*, 45 (1995) 205-43.
13. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E., A geometric approach to macromolecule-ligand interactions, J *MolBiol*, 161 (1982) 269-88.
14. Lybrand, T.P., Ligand-protein docking and rational drug design, C*urrOpinStruckBiol*, 5(1995) 224-8.
15. Jones, G. and Willett, P., Docking small-molecule ligands into active sites, C*urrOpin Biotechnology*, 6 (1995) 652-6.
16. Goodsell, D.S., Morris, G.M. and Olson, A.J., Automated docking of flexible ligands: applications of Auto Dock, J *MolRecognition*, 9 (1996) 1-5.
17. Nussinov, R. and Wolfson, H.J., Efficient computational algorithms for docking and for generating and matching a library of functional epitopes II. Computer vision-based techniques for the generation and utilization of functional epitopes, C*omb ChemHighThroughput Screen*, 2 (1999) 261-9.
18. Abagyan, R. and Totrov, M., High-throughput docking for lead generation, C*urrOpinChemBiol*, 5 (2001) 375-82.
19. Schneider, G. and Bohm, H.J., Virtual screening and fast automated docking methods, D*rug Discover Today*, 7 (2002) 64-70.
20. Taylor, R.D., Jewsbury, P.J. and Essex, J.W., A review of protein-small molecule docking methods, J *Compute Aided Mol Des*, 16 (2002) 151-66.
21. Lawrence, M.C. and Davis, P.C., CLIX: a search algorithm for finding novel ligand scapable of binding proteins of known three-dimensional structure, P*roteins*, 12 (1992)31-41.
22. Bohm, H.J., LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads, J *Compute Aided Mol Des*, 6 (1992) 593-606.
23. Lauri, G. and Bartlett, P.A., CAVEAT: a program to facilitate the design of organic molecules, J *Compute Aided Mol Des*, 8 (1994) 51-66.

24. Nishibata, Y. and Itai, A., Confirmation of usefulness of a structure construction program based on three-dimensional receptor structure for rational lead generation, J *MedChem*, 36 (1993) 2921-8.
25. Gehlhaar, D.K., Moerder, K.E., Zichi, D., Sherman, C.J., Ogden, R.C., et al., De novo design of enzyme inhibitors by Monte Carlo ligand generation, J *Med Chem*, 38 (1995) 466-72.
26. Jung S, Bae SE, Son HS (2011) Validity of Protein Structure Alignment Method Based on Backbone Torsion Angles. J Proteomics Bioinform 4: 218-226.
27. Krogh A, Larsson B, Von HG, Sonnhammer EL (2001) Predicting trans-membrane protein topology with a hid-den Markov model: application to complete genomes. J MolBiol 305: 567-80.
28. Hirokawa T, Boon CS, Mitaku S (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14: 378-379.
29. Ganesh CS, Manas RD, Mukta R, Pradeep D (2009) Homology Modelling and Functional Analysis of LPG2 Protein of Leishmania Strains. J Proteomics Bioinform 0: 032-050.
30. Baldi, P. & Brunak S. "Bioinformatics - The Machine Learning Approach". Massachusetts Institute of Technology, 1998.
31. Krogsgaard-Larsen, P., Liljefors, T. and Madsen, U., Textbook of drug design and discovery, 3rd edn, Taylor & Francis, London; New York, 2002, xviii, 572 pp.
32. Birney, E. "Hidden Markov Models in Biological Sequence Analysis". IBM Journal of Research and Development Volume 45, Numbers 3/4, 2001.
33. Eddy SR. Hidden Markov models. Curr Opin StructBiol 1996; 6:361-5.
34. Choo KH, Tong JC, Zhang L. Recent applications of Hidden Markov Models in computational biology. Genomics Proteomics Bioinformatics 2004; 2: 84-96.
35. Pedersen JS, Hein J. Gene finding with a hidden Markov model of genome structure and evolution. Bioinformatics 2003; 19: 219-27.
36. Eddy SR. Hidden Markov models. Curr Opin Struct Biol 1996; 6:361-5.
37. Choo KH, Tong JC, Zhang L. Recent applications of Hidden Markov Models in computational biology. Genomics Proteomics Bioinformatics 2004; 2: 84-96.
38. Lander ES, Green P. Construction of multilocus genetic linkage maps in humans. ProcNatl AcadSci U S A 1987; 84: 2363-7.
39. Krogh A, Mian IS, Haussler D. A hidden Markov model that finds genes in E. coli DNA. Nucleic Acids Res 1994; 22: 4768-7.
40. Asai K, Hayamizu S, Handa K. Prediction of protein secondary structure by the hidden Markov model. ComputApplBiosci 1993; 9: 141-46.
41. Goldman N, Thorne JL, Jones DT. Using evolutionary trees in protein secondary structure prediction and other comparative sequence analyses. J MolBiol 1996; 263: 196-208.
42. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 2003; 12: 1652-62.
43. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting trans-membrane protein topology with a hidden Markov model: application to complete genomes. J MolBiol 2001; 305: 567-80
44. Liu Q, Zhu YS, Wang BH, Li YX. A HMM-based method to predict the trans-membrane regions of beta-barrel membrane proteins. ComputBiolChem 2003; 27: 69-76.

45. Yoon B-J, Vaidyanathan PP. HMM with auxiliary memory: a new tool for modeling RNA secondary structures. Proceedings of theThirty-Eighth Asilomar Conference on Signals, Systems, and Computers, Monterey, CA, 2004; 2: 1651-5.
46. Wheeler KE, Zemla A, Jiao Y, AliagaGoltsman DS, Singer SW, et al. (2010)Functional Insights from Computational Modeling of Orphan Proteins Expressed in a Microbial Community. J Proteomics Bioinform 3: 266-274.
47. Sahay A, Shakya M (2010) Insilico Analysis and Homology Modeling of Antioxidant Proteins of Spinach. J Proteomics Bioinform 3: 148-154.
48. Sali A, Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. J MolBiol 234: 779-815.
49. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereo chemical quality of protein structures. J Appl Crystallography 26: 283-291.
50. Nath M, Goel A, Taj G, Kumar A (2010) Molecular Cloning and Comparative Insilico Analysis of Calmodulin Genes from Cereals and Millets for Understanding the Mechanism of Differential Calcium Accumulation. J Proteomics Bioinform 3:294-301
51. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. J MolBiol 7: 95-99.
52. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8: 477-486.
53. McGovern, S. L.; Shoichet, B. K. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. *J. Med. Chem.* 2003, *46*, 2895-2907
54. Evers, A.; Gohlke, H.; Klebe, G. Ligand-supported homology modeling of protein binding-sites using knowledge-based potentials. *J. Mol. Biol.* 2003, *334*, 327-345.
55. Evers, A.; Klebe, G. Ligand-supported homology modeling of g protein-coupled receptor sites: models sufficient for successful virtual screening. *Angew. Chem. Int. Ed. Engl.* 2004, *43*, 248-251.
56. Ferrari, A. M.; Wei, B. Q.; Costantino, L.; Shoichet, B. K. Soft docking and multiple receptor conformations in virtual screening. *J. Med. Chem.* 2004, *47*, 5076-5084.Abdelouahab C, Abderrahmane B (2008) Comparative Study of the Efficiency of Three Protein-Ligand Docking Programs. J Proteomics Bioinform 1: 161-165.
57. Girija CR, Karunakar P, Poojari CS, Begum NS, Syed AA (2010) Molecular Docking Studies of Curcumin Derivatives with Multiple Protein Targets for Procarcinogen Activating Enzyme Inhibition. J Proteomics Bioinform 3: 200-203.