Research and Reviews: Journal of Medicinal Chemistry

Protein Structure Prediction and Its Application in Drug Design by using Computational Methods

Satya Varali*

Department of Human Genetics, Andhra University, India

Commentary

Received: 15/May/2015
Revised: 29/May/2015
Accepted: 2/June/2015

*For Correspondence

Satya Varali, Department of Human Genetics, Andhra University, India

Keywords: Protein structure expectation; SOPMA; TMHMM; Comparative displaying; Modeler; Drug outline; Docking

ABSTRACT

Medication outline and medication disclosure are of discriminating significance in human social insurance. Computational methodologies have turn into a noteworthy piece of structure based medication outline. Structure-based medication outline uses the three dimensional structure of a protein focus to plan hopeful medications that are anticipated to tie with high natural inclination and selectivity to the objective. In this survey computational systems for expectation of the protein structure are depicted and their utilization towards the medication outline is talked about.

INTRODUCTION

Most advanced medication revelation ventures begin with protein target identification and check to get a checked medication target. For structure-based medication outline the three-dimensional structure of the protein needs to be resolved tentatively by utilizing either x-ray crystallography or NMR spectroscopy [1]. While both routines are progressively being connected in a high-throughput way, structure determination is not yet a clear process. X-ray crystallography is restricted by the trouble of getting a few proteins to frame gems, and NMR must be connected to generally little protein atoms. Proteins are fundamental to organic procedures [2]. They are in charge of catalyzing and controlling biochemical responses, transporting atoms, the science of vision and of the photosynthetic transformation of light to development, and they frame the premise of structures, for example, skin, hair, and tendon. Protein capacity can be seen in wording of its structure. To be sure, the three-dimensional structure of a protein is firmly identified with its organic capacity. Proteins that perform comparative capacities have a tendency to demonstrate a noteworthy level of auxiliary homology [3-6]. The three-dimensional design of the essential structure is characterized as the tertiary structure, depicting the fold of the protein.

Important Structure Investigation

Amino corrosive succession examination [7] gives essential insightinto the structure of proteins, which thusly enormously encourages the comprehension of its biochemical and cell capacity. Endeavors to utilize computational techniques in foreseeing protein structure construct just in light of grouping data began 30 years prior [8-10]. Be that as it may, just amid the most recent decade, has the presentation of new computational methods, for example, protein fold acknowledgment and the development of succession and structure databases because of cutting edge high-throughput innovations prompted an increment in the achievement rate of forecast systems [11].

Grouping recovery database looks

Succession similitude looking is a critical stride in dissecting recently decided protein successions. Regularly, extensive grouping databases, for example, the non-excess (nr) database at the NCBI [12-15] (combination of Gen-Bank, EMBL and DDBJ databases) or genome groupings are checked for DNA or amino corrosive arrangements that are like an objective succession. Arrangements of the objective succession are developed for every database section, ordinarily utilizing element programming calculations [16]. Scores got from these arrangements are utilized to recognize measurably noteworthy matches. Generally, quests were completed utilizing projects for pairwise arrangement examinations like FASTA [17] or BLAST [18-20]. Then again, the relationship between successions of homologous proteins can be perceived via pairwise arrangement correlations.

Protein area distinguishing proof

After protein disclosure, there are numerous inquiries that are connected with protein's general personality, putative capacity and naturally huge locales recognizable proof [21]. To answer these inquiries, various databases and apparatuses have been modified. Most proteins are made from a limited number of developmentally monitored modules or spaces. Protein spaces are particular units of three-dimensional protein structures, which regularly convey a discrete sub-atomic capacity, for example, the coupling of a particular kind of atom [22-25]. These areas fluctuate long from between around 25 amino acids up to 500 amino acids. The direct utilitarian and basic determination of the considerable number of proteins in a living being is restrictively immoderate and time intensive in view of the relative shortage of 3D auxiliary data consequently essential grouping investigation is liked to recognize larger part of protein area families [26]. A couple of thousand preserved spaces, which cover more than 66% of known protein arrangements have been distinguished and portrayed in writing [27].

Auxiliary structure forecast

The imperative ideas in auxiliary structure expectation are distinguished as: buildup conformational affinities, arrangement edge impacts, snippets of hydrophobicity, position of insertions [28-30] and deletions in adjusted homologous succession, snippets of preservation, auto-connection, deposit proportions, optional structure input impacts, and sifting [31]. The early techniques for auxiliary structure forecast are experienced an absence of information and Predictions were performed on single arrangements instead of groups of homologous groupings, and there were moderately few known 3D structures from which to determine parameters [32-35].

Late changes

The accessibility of vast groups of homologous arrangements reformed optional structure expectation [36]. Customary techniques, when connected to a group of proteins instead of a solitary grouping demonstrated significantly more precise at distinguishing center auxiliary structure components. The blend of arrangement information with modern processing methods, for example, neural systems [37-40] has led to exact nesses well in abundance of 70%. Despite the fact that this appears a little rate build, these forecasts are really significantly more valuable than those for single succession [39], since they have a tendency to foresee the center precisely. Additionally, the breaking point of 70-80% may be an element of auxiliary structure variety inside homologous proteins. SOPMA (Self Optimized Prediction Method from Alignment) was utilized for expectation of auxiliary structure components like alpha helix, augmented strand, beta turn and irregular loops as far as rate for all the groupings. These components were considered as info parameters for self-sorting out maps for further investigation [40].

Transmembrane area forecast

Distinctive servers TMHMM, SOSUI [47], HMMTOP and TM were gotten to approve the TM area. TMHMM, a new film protein topology expectation system, is in view of a hidden Markov model [41].

Transmembrane topology forecasts

In a study directed by Kumar et al. the transmembrane topology of AHA1 was anticipated from the amino corrosive arrangement [42] by averaging the aftereffects of four diverse prescient calculations: DAS, HMMTOP, TMHMM [43] and TMPRED [44]. The exactness of the forecast was evaluated by utilizing the same calculations on 3b8c protein and the outcomes were contrasted and the topology characterized in the 3.6Ao structures (3b8c).

Tertiary structure

Knowing a protein's 3-dimensional structure (Tertiary Structure) helps us to comprehend its usefulness and gives intends to arranging analyses and medication plan. The Brookhaven Protein Data Bank (PDB) [45] is the store for those structures. Records including particle coordinates which are suited for perception by graphical atom viewers like rasmol can be acquired at this site. PDB is additionally searchable with an arrangement as an inquiry, e.g. with the BLAST administration situated at NCBI with a polypeptide as an inquiry. Tertiary structure of a protein is fabricate by pressing of its auxiliary structure components to shape discrete spaces or independent collapsing units [46].

Relative displaying

Homology or relative protein structure displaying develops a three-dimensional model of a given protein grouping in light of its closeness to one or more known structures [47]. It is done in four successive steps: discovering known structures (layouts) identified with the arrangement to be demonstrated (target), adjusting the objective succession to the formats, fabricating the model, and surveying the model [48]. Subsequently, relative demonstrating is just relevant when the objective arrangement is recognizably identified with a known protein structure.

3D structure era by utilizing Modeler

Modeler is a PC program for near displaying of protein three-dimensional structures. Arrangement of a grouping to be displayed is given known related structures and modeler naturally computes a model containing all non-hydrogen iotas. Modeler [49] actualizes similar protein structure demonstrating by fulfillment of spatial restrictions. Format distinguishing proof and arrangement

Format recognizable proof is an essential step. It lays the foundation by distinguishing proper homologues of known protein structure, called layout, which are adequately like the objective grouping to be displayed. Layout arrangement were chosen by a basic pursuit presents the objective grouping to projects BLASTP look alongside default parameters was performed against the Brook Heaven Protein Data Bank (PDB) [50].

Model building and refinement

Despite the fact that the hypothesis behind building a protein homology model is muddled, utilizing accessible projects is generally simple. A few demonstrating projects [45] are accessible, utilizing distinctive routines to develop the 3D structures. In section coordinating systems, the objective is partitioned into short fragments, and arrangement is done over sections instead of over the whole protein.

CONCLUSION

Computational routines for protein structure expectation are still in the phase of improvement and techniques like homology-based forecast turn out to be particularly useful in a situation where the systems can be utilized as a part of show with exploratory procedures for structure and capacity determination of protein. The utilization of PCs and computational routines saturates all parts of medication disclosure today and structures the center of structure-based medication plan. Accessibility of protein 3D structures, superior processing, information administration programming and web are encouraging the entrance of immense measure of information created and changing the huge complex natural information into workable learning in current medication disclosure process. Computational apparatuses offer the benefit of conveying new medication applicants all the more rapidly and at a lower expense.

REFERENCES

1. Chen B. Reconstruction of Local Biochemical Reaction Network Based on Human Chromosome 9 Sequence Data. J Proteomics Bioinform. 2011;4: 087- 090..
2. Asthana M, et al. Isolation, Cloning and In silico Study of Hexon Gene of Fowl Adenovirus 4 (FAV4) Isolates Associated with Hydro Pericardium Syndrome in Domestic Fowl. J Proteomics Bioinform. 2011;4: 190-195.
3. Mazzara S, et al. Application of Multivariate Data Analysis for the Classification of Two Dimensional Gel Images in Neuroproteomics. J Proteomics Bioinform. 2011;4: 016-021.
4. Kanamoto T, et al. Proteomic Analyses of Proteins Differentially Expressed in Recurrent and Primary Pterygia. J Proteomics Bioinform. 2011;4: 058-061.
5. Rajkumar R, et al. Characterization of Rat Odorant Binding Protein Variants and its Post-Translational Modifications (PTMs): LC-MS/MS analyses of Protein Eluted from 2D-Polyacrylamide Gel Electrophoresis. J Proteomics Bioinform. 2011;4: 210-217.
6. Lu CH, et al. Detection of Abundant Proteins in Multiple Myeloma Cells by Proteomics. J Proteomics Bioinform. 2010;3: 005-009.
7. Nagano K. Logical analysis of the mechanism of protein folding. I. Predictions of helices, loops and beta-structures from primary structure. J Mol Biol. 1973;75: 401–420.
8. Amit KB, et al. Exploring the Interplay of Sequence and Structural Features in Determining the Flexibility of AGC Kinase Protein Family: A Bioinformatics Approach. J Proteomics Bioinform. 2008;1: 077-089
9. Needleman SB and Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol. 1970;48: 443–453
10. Neha S, et al. Comparative In silico Analysis of Partial Coat Protein Gene Sequence of Zucchini Yellow Mosaic Virus Infecting Summer Squash (Cucurbita pepo L.) Isolated From India. J Proteomics Bioinform. 2011;4: 068-073
11. Neha G, et al. Primer Designing for Dreb1A, A Cold Induced Gene. J Proteomics Bioinform. 2008;1: 028-035
12. Dangre DM, et al. An in Silico Molecular Evolutionary Analysis of Selected Species of Phoma : A Comparative Approach. J Proteomics Bioinform. 2009;2: 295-309
13. Varsale AR, et al. Cheminformatics. J Proteomics Bioinform. 2010;3: 253-259
14. Sonnhammer EL, et al. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol. 1998;6: 175–182
15. Asthana M, et al. Isolation, Cloning and In silico Study of Hexon Gene of Fowl Adenovirus 4 (FAV4) Isolates Associated with Hydro Pericardium Syndrome in Domestic Fowl. J Proteomics Bioinform. 2011;4: 190-195

16. Mishra A, et al. Molecular Cloning and In silico Analysis of Functional Homologues of Hypersensitive Response Gene(s) Induced During Pathogenesis of Alternaria Blight in Two Genotypes of Brassica. J Proteomics Bioinform. 2010;3: 244-248
17. Allam AR, et al. Computational Analysis of Mutations in Colon Cancer Genes Reveals a Possible Role of Micro Satellite in Mutagenesis. J Proteomics Bioinform. 2008;S1: S041-S045
18. Afergan E, et al. 31P-NMR and Differential Scanning Calorimetry Studies for Determining Vesicle's Drug Physical State and Fraction in Alendronate Liposomes. J Bioanal Biomed. 2010;2: 125-131
19. Chan HS and Dill K .The protein folding problem. Physics Today. 1993.
20. Jung S, et al. Validity of Protein Structure Alignment Method Based on Backbone Torsion Angles. J Proteomics Bioinform. 2011;4: 218-226
21. Wilcox GL, et al.Neural network analysis of protein tertiary structure. Tetrahedron Computer Methodology. 1190;3: 191–211
22. Gomase VS, et al. Prediction of MHC Binding Peptides and Epitopes from Groundnut Bud Necrosis Virus (GBNV). J Proteomics Bioinform. 2008;1: 188-205
23. Kabsch W and Sander C. How good are predictions of protein secondary structure. FEBS Letters. 1983;155: 179-182.
24. Vaseeharan B and Valli SJ. In silico Homology Modeling of Prophenoloxidase activating factor Serine Proteinase Gene from the Haemocytes of Fenneropenaeus indicus. J Proteomics Bioinform. 2011;4: 053-057
25. Ganesh CS, et al. Homology Modeling and Functional Analysis of LPG2 Protein of Leishmania Strains. J Proteomics Bioinform. 2009; 2: 032-050
26. Kumar S, et al. In Silico Identification of Putative Proton Binding Sites of a Plasma Membrane H+-ATPase Isoform of Arabidopsis Thaliana, AHA1. J Proteomics Bioinform. 2009;2: 349-359
27. Tusnady GE and Simon I. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J Mol Biol. 1998;283: 489-506
28. Min XJ. Evaluation of Computational Methods for Secreted Protein Prediction in Different Eukaryotes. J Proteomics Bioinform. 2010;3: 143-147.
29. Subramanian R, et al. Comparative Modeling and Analysis of 3-D Structure of EMV2, a Late Embryogenesis Abundant Protein of Vigna Radiata (Wilczek). J Proteomics Bioinform. 2008;1: 401-407
30. Halperin I, et al. Principles of docking: an overview of search algorithms and a guide to scoring functions. Proteins; 2002;47: 409-443
31. Yadav BS, et al. Predicting Secondary Structure of Oxidoreductase Protein Family Using Bayesian Regularization Feed-forward Backpropagation ANN Technique. J Proteomics Bioinform. 2010;3: 179-182
32. Murty USN, et al. An In Silico Approach to Cluster CAM Kinase Protein Sequences. J Proteomics Bioinform. 2009;2: 097-107
33. Selvaraj D, et al. Molecular Characterization and Phylogenetic Analysis of BZIP Protein in Plants. J Proteomics Bioinform. 2010;3: 230-233
34. Liang K and Wang X. Protein Secondary Structure Prediction using Deterministic Sequential Sampling. J Data Mining in Genom Proteomics. 2011;2:107
35. Sechi M, et al. Design of novel bioisosteres of beta-diketo acid inhibitors of HIV-1 integrase. Antivir Chem Chemother. 2005;16: 41-46
36. Kaushik SH, et al. In-silico Interaction Studies of Quinazoline Derivatives for their Inhibitory Action on Both Wild and Mutant EGFRs. J Proteomics Bioinform. 2009;2: 126-130
37. Chikhi A and Bensegueni A. In Silico Study of the Selective Inhibition of Bacterial Peptide Deformylases by Several Drugs. J Proteomics Bioinform. 2010;3: 061-065

38. Bagchi P, et al. Pharmaco-Informatics: Homology Modeling of the Target Protein (GP1, 2) for Ebola Hemorrhagic Fever and Predicting an Ayurvedic Remediation of the Disease. J Proteomics Bioinform. 2009;2: 287-294
39. Prakash N, et al. Molecular Docking Studies of Antimalarial Drugs for Malaria. J Comput Sci Syst Biol. 2010;3: 070- 073
40. Kumar PM, et al. An Insight into Structural and Functional Characteristics of Plasmodium falciparum Farnesyltransferase (PfFT) 3d7: Comparative Modeling and Docking Studies. J Proteomics Bioinform. 2010;3: 305-309
41. Wheeler KE, et al. Functional Insights from Computational Modeling of Orphan ProteinsExpressed in a Microbial Community. J Proteomics Bioinform. 2010;3: 266-274
42. Sahay A and Shakya M. In silico Analysis and Homology Modelling of Antioxidant Proteins of Spinach. J Proteomics Bioinform. 2010;3: 148-154
43. Shakil S and Khan AU. Interaction of 2009 CTX - M Variants with Drugs and Inhibitors: a Molecular Modeling and Docking Study. J Proteomics Bioinform. 2010;3: 130-134
44. Marti-Renom MA, et al. Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct. 2000;29: 291-325
45. Suresh N and Vasanthi NS. Pharmacophore Modeling and Virtual Screening Studies to Design Potential Protein Tyrosine Phos-phatase 1B Inhibitors as New Leads. J Proteomics Bioinform. 2010;3: 020-028
46. Nath M, et al. Molecular Cloning and Comparative In silico Analysis of Calmodulin Genes from Cereals and Millets for Understanding the Mechanism of Differential Calcium Accumulation. J Proteomics Bioinform. 2010;3: 294-301
47. Ramachandran GN, et al. Stereochemistry of polypeptide chain configurations. J Mol Biol. 1963;7: 95-99
48. Laskowski RA, et al. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR. 1996;8: 477-486
49. Sunil K, et al. Prediction of 3-Dimensional Structure of Cathepsin L Protein of Rattus Norvegicus. J Proteomics Bioinform. 2008;1: 307-314
50. Luthy R, et al. Assessment of protein models with three-dimensional profiles. Nature. 1992;356: 83-85
51. Seenivasagan R, et al. Comparitive Modeling of Viral Protein R (VpR) from Human Immunodeficiency Virus 1 (HIV 1). J Proteomics Bioinform. 2008;1: 073-076
52. Seetharaaman B and Krishnan VM. In Silico Analysis of Alkaline Shock Proteins in Enterobacteria. J Proteomics Bioinform. 2008;S1: S021-S037
53. Eswar N, et al. Comparative Protein Structure Modeling using MODELLER. Curr Protoc Bioinformatics. 2006;61: 5630
54. Fiser A, et al. Modeling of loops in protein structures. Protein Sci. 2008;9: 1753-73
55. Ramón A, et al. Integration and Prediction of PPI Using Multiple Resources from Public Databases. J Proteomics Bioinform. 2008;1: 166-187
56. Enright AJ, et al. Protein interaction maps for complete genomes based on gene fusion events. Nature. 1999;402: 86-90
57. Ramanathan K, et al. Structure Based Drug Designing for Diabetes Mellitus. J Proteomics Bioinform. 2010;3: 310-313
58. Abdelouahab C and Abderrahmane B. Comparative Study of the Efficiency of Three Protein-Ligand Docking Programs. J Proteomics Bioinform. 2008;1: 161-165
59. 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.