International Journal of Theoretical & Applied Sciences, 14(1): 61-67(2022)

ISSN No. (Print): 0975-1718
ISSN No. (Online): 2249-3247

Application of Homology Modeling; A Molecular Visualization Method

Piyusha Sharma[1*], Anjana Thakur[2] and Neha Kumari[2]
[1]Assistant Professor, Department of Life Sciences,
Desh Bhagat University, Mandi Gobindgarh, 147301 (Punjab), India.
[2]Research Scholar, Department of Biotechnology,
Shoolini Institute of Life Science and Business Management (Himachal Pradesh), India.

(Corresponding author: Piyusha Sharma)*
(Received 27 April 2022; Accepted 24 June 2022)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Diverse Research in areas like protein engineering, human genetics, structure-based drug design and analysis of protein function knowledge of the three-dimensional structure of proteins is a prerequisite. It is ultimately dictated by protein sequence and is typically necessary to comprehend the mechanism of protein function. It takes a lot of time and doesn't always work with all proteins, particularly membrane proteins, to determine the structure of a protein using experimental techniques like X-ray crystallography or NMR spectroscopy. Protein modelling aims to predict a three-dimensional structure from its sequence with an accuracy using a technique called homology modelling also known as comparison modelling or knowledge-based modelling. The current review offers a methodical evaluation of the efficacy of frequently employed (and commercially accessible) homology modelling software for therapeutically important proteins, evaluating both the sequence alignments and the created 3D models. The *ab-initio* approach and homology modelling can be grouped into two extreme categories for theoretical structure prediction. One objective of the first method is to predict folds using physical chemistry concepts. A protein sequence's three-dimensional structure can be predicted using a second method, which principally bases its prediction on the protein sequence's similarity to other proteins with known structures.

Keywords: Homology modeling, X-ray crystallography, NMR spectroscopy, databases, 3D structure.

I. INTRODUCTION

A precise definition of homology is having a shared evolutionary ancestor. As a result, homology cannot be partial and is a qualitative representation of the nature of the relationship between two or more objects. There is either an evolutionary connection or there isn't. Normally, a homology claim must remain a hypothesis. Sequence or three-dimensional similarities, the relationships between which may be quantitatively defined, are examples of supporting data indicating a homologous relationship. The fact that the three-dimensional structures of a group of proteins that are thought to be homologous are more conserved than their fundamental structures is a key finding in homology modelling. By using homologues with extremely little sequence similarity, this observation has been exploited to create models of proteins. An effort is made to create homologous protein models of an unknown. The ability to accurately predict protein structure, which was not even possible a few years ago, has been made possible by recent advances in homology modelling, particularly in the detection of distant homologues, alignment of sequences with template structures, modelling of loops and side chains, and error detection. The ongoing efforts to solve protein structures, which can be time-consuming and frequently challenging, will encourage the development of a variety of new computational techniques that can close the knowledge gap and advance our understanding of the relationship between protein structure and function. Homology modelling has been made easier and more efficient by fully automated workflows and servers, which also enable users without specialised computational knowledge to create accurate protein models and have simple access to modelling outcomes, their display, and interpretation [34]. The description and inclusion of software created for protein modelling are part of the current review.

II. REVIEW OF LITERATURE

The idea behind homology modelling is that homologous proteins' 3D structures are more conserved than their sequences [35]. The pioneering homology modelling articles were published in the late 1970s [8]. Finding an appropriate template is the first step in the procedure; Insertions, deletions, and substitutions of residues are made after an alignment is created [30]. Homology modeling is applied for constructing protein 3D structures exploiting its primary sequence available in databases and utilizing prior knowledge gained from structural similarities with other proteins [11]. Studying the underlying theory and implementation techniques of the database's structure, storage, design, maintenance

provides the ease in processing and analysis of the data contained within the database [37]. By using online computing resources, high-resolution protein 3D structures can be created for research. The modeller should be able to access and use the atomic coordinates for creating homology models through the accessible experimental 3D-structure sources [10].

Software for homology modelling. There are numerous tools and servers for homology modelling that are designed to create a whole model using query sequences. In order to create constraints on atomic distances, dihedral angles, and other parameters, MODELLER employs the query structures. These constraints are then paired with statistical distributions obtained from several homologous structure pairings in the PDB. Sequences and structures are combined by Modeller 0 to provide a complete alignment that may be manually modified and inspected with molecular graphics software. By wrapping the software tool MODELLER in a graphical user interface, Accelrys Software Inc.'s programme for modelling protein homology is known as DSModeler [25]. Given a template(s) and a sequence alignment, DSModeler creates protein homology models. A molecular force field, a database of crystal structures in the PDB, and distance constraints determined from the template are used to forecast the structures. Disulfide bonds and cis-prolines are examples of restrictions and ligand structures that can be introduced into the model building stage. A technique that incorporates knowledge-based potentials from well-known crystal structures generates loops from scratch. Software called PROFIT was created using protein fold recognition (threading) technology by Proceryon Biosciences [27]. Sequence alignments are generated utilizing potential in terms for atom pair and protein-solvent interactions that are based on information. To generate an alignment of the target sequence with a template structure, these potentials are used with dynamic programming and frozen approximation. Optimized gap penalties, gap restrictions, and potentials are all employed to evaluate how well the alignment would be [5]. The ExPASy Web server provides access to SWISS-MODEL, a fully automated system for protein structure homology modelling. In 1993, Manuel Peitsch launched SWISS-MODEL. A sequence alignment and a PDB file for the template are provided as input. The knowledge-based homology model is created using the ProModII tool after these are transmitted via a server [23]. Complete backbone and side chain building, loop creation, and model quality assurance, including packing, are all components of model construction [29]. In order to create a composite template, PrISM does homology modelling utilising alignment, choosing each secondary structure from the best suitable template. Loop modelling is done using ab- initio techniques, and side-chain dihedrals are either obtained from the template or predicted based on the structure's main chain torsion angles using a neural network approach [36]. For constructing homology models, COMPOSER makes use of a variety of template structures. The primary goal of the biannual Critical Assessment Approaches for Protein Structure Prediction (CASP) experiments is to establish benchmarking criteria for the protein structure prediction techniques used by various internet servers and applications. They keep track of advancements in the field of modelling protein structure from sequence. These experiments' primary goals are to ensure the models' overall quality, predictability, and evaluation of the parameters offered by various tools [39]. A service for comparative modelling that predicts protein structure is called Homer (Homology ModellER). It creates a model structure using a single template structure and an alignment in FASTA format (PDB format)**.** The CaspR web server is based on a set of common software applications that are popular among protein crystallography and bioinformatics researchers. The T-COFFEE software is the first step in the procedure to create a trustworthy multiple alignment. The provision of a dependability index [CORE index] for each place in the alignment is a distinctive feature of T-COFFEE. Better multiple alignments are produced by T-COFFEE, which integrates structure and sequencing data, and MODELLER's homology models benefit as a result [4]. An automated service for modelling comparative protein structures is called ModWeb. It accepts one or more sequences in the FASTA format and generates models for them using the best Protein Data Bank template structures. A free online homology modelling server is Phyre2 Protein Homology/analogy Recognition Engine 2. To represent those portions of the proteins in question that have no discernible similarity to recognised structures, Phyre2 uses a new ab-initio folding simulation termed Poing [22]. In order to satisfy spatial constraints (dihedral and distances), FASPR is a very helpful tool for protein structure modelling and protein design due to its excellent accuracy, speed, and determinacy for modelling the side-chains of both native and non-native main-chain conformations [12]. Geno3D models comparative protein structures. Homology or comparative protein structure modelling is where Geno3D is most frequently employed. Geno3d accepts input in a manner similar to FASTA, but all that is required is a single letter of the code. The output is in PDB format, which may be viewed in any molecular modelling programme. SPORulate server provides several tools, including SWISS-MODEL, CPH models, and SDSC1, are made available through this server to assist the user in submitting the sequence for homology modelling. For secondary structure prediction and Fold identification, it also contains a number of programmes. A technology called Wloop the loop homology modelling server is used to forecast the backbone structures of protein loops based on their sequences and flank backbone structures [19].

Modeled structures through homology modeling. Over the years, many homology models have been developed. Antibodies and other proteins important in human biology and medicine have been used as targets. Nearly 50% of all drug development programmes target GPCRs since they are the largest family of signalling receptors in cells [33]. To evaluate the accuracy of structural predictions and the relevance of those

predictions to chemotaxis research, homology models of the chemotaxis proteins CheW from *E. coli* and *T. maritima were* built. The sequence for modelling *E. coli* CheW was found in the UniProt database (Entry ID: P0A964), and it was modelled using the T. maritime CheW structure as a model that was found in the Protein Data Bank (PDB ID: 1K0S) [2]. Using the programme MODELLER9v4, the structure of SRCR1 (residues 95-203 on DMBT1, renumbered as 1-109 in this article) was developed based on the knowledge of the structure of M2bp SRCR domain (PDB ID: 1BY2) [3]. The homologous X-ray crystal structure of the Hepatitis C viral helicase served as a basis for creating the 3D model of the Classical Swine Fever virus (CSFV) helicase. The model's geometry, fold recognition, and compliance with the standards necessary as a member of the Flaviviridae virus family were successfully assessed in silico [31]. Exo-inulinase from *Penicillium* sp. TN-88(BAC16218) and endo-inulinase from *Penicillium* sp. TN-88(BAA19132) structural modelling and active site research are reported. For structural modelling, the exo-inulinase sequence from *Penicillium* sp. TN-88(BAC16218) was obtained from NCBI (http://www.ncbi.nlm.nih.gov/) and uploaded to Swiss model (http://swissmodel.expasy.org). The RCSB Protein Data Bank was used to model the structures of the obtained sequences (PDB)[26]. A novel strategy based on additional data from the template structures has been developed to address the shortcomings of standard homology modelling. The strategy uses the space filled by ligands or substrates in the template structures to prevent the target protein from folding [32]. In addition to being necessary for healthy T-cell development, the BCL11B transcription factor has recently been linked to the pathogenesis of T-cell acute lymphoblastic leukaemia (T-ALL), which is brought on by TLX over expression or Atm deficiency. Based on the high-resolution crystal structure of Egr1 (Zif268) in association with DNA, structural homology modelling was utilised to simulate BCL11B tandem ZF2-ZF3 zinc finger binding to a common GC-rich DNA oligonucleotide sequence. Using SWISS-MODEL to calculate the structural model, a high-quality structure with a model-template C-root mean square deviation of 2.9 Angstroms was produced [9]. Using comparative modelling, a structural model of mature *L. mexicana* CPB2.8 CTE was created in order to get the protein's 3D structure for covalent docking investigations and subsite residue identification. When compared to its template structure, the final homology model of CPB2.8 CTE had a C RMSD score of 0.699. By superimposing the two protein structures, the homology model was then structurally compared to bovine cathepsin B (BtCatB, PDB ID 1QDQ).By using the coordinates of the human MLN64 protein (1EM2) from the Brookhaven Protein Database (BPD), an automated comparative protein modelling server (Swiss-Model) at the University of Geneva was able to create the three-dimensional structure of STAR proteins (cytosolic steroidogenic acute regulatory protein). ClustalW was used to align a total of 13 sequences, including STAR and STARD3 from humans to worms/insects. Using the ConSurf web server, the scores were projected onto a STAR homology model's three-dimensional structure. The potential cholesterol recognition/interaction amino acid consensus (CRAC) domain and steroid binding pocket were shown in relation to the inner mitochondrial membrane using the three-dimensional structure of human CYP11A1 [6]. *Osmium basilicum*, a plant that has been studied using homology modelling, has many medicinal uses, including decreasing blood pressure, acting as an antispasmodic, and cleansing blood. Homology modelling was used to establish the protein's 3D structure. To locate the most appropriate templates for homology modelling, a BLAST P search against the Protein Data Bank (PDB) was conducted using default parameters. The homology model was predicted using Protein Structure Prediction Server (PS)2 based on software MODELLER [13]. From the Protein Knowledgebase of UniProt (http://www.uniprot.org), the amino acid sequence of human MK5 (Swiss-Prot: Q8IW41) was obtained. BLAST via UniProt was used to find close homologues having accessible crystal structures in the Protein Databank (http://www.pdb.org) [16].The three-dimensional structure of MP was created using homology modelling and molecular dynamic (MD) methods (Marine alkaline protease). The MP target sequence (accession number ACY25898) was retrieved from the protein database of the National Center for Biotechnology Information (NCBI). On the BLAST web service (http://blast.ncbi.nlm.nih.gov), a sequence similarity search for this protease was conducted against sequences from the Protein Data Bank (PDB) database. MODELLER (version 9.9), a programme used for homology modelling of protein three-dimensional structures, was used to construct MP's tertiary structure. The predicted structures were stored in PDB format and organised by scores derived from GA341 and discrete optimised protein energy (DOPE) scoring. The PROCHECK tool (https://saves.mbi.ucla.edu/) is used to analyse the energy and stereochemical properties of the modelled protein structures from the Structural Analysis and Verification of Protein (SAVES) website. The PROCHECK tool is also used to create a Ramachandran plot for each target in order to determine whether the residues were located in an energetically advantageous area. With the use of SWISS-PDB Viewer, energy minimization and loop creation for residues in the Ramachandran plot's forbidden zones were carried out. Cancer cells' survival and antiapoptotic mechanisms are crucially activated by the phosphatidylinositol 3-kinase/AKT signalling pathway. A total of 22 putative PH domain inhibitors were found using specialised docking software. Using UNITY (Tripos, L.P.), a three-dimensional pharmacophore search was conducted based on the hydrogen-bonding pattern between the ligand, inositol (1,3,4,5)-tetrakisphosphate, and the PH domain of AKT (1H10) [21]. The cystic fibrosis transmembrane conductance regulator protein (CFTR), which causes the disease, was modelled using homology theory (CF). Using

MODELLER, a homology model of the open-channel state of CFTR (without the R-domain) was created [24]. For bubaline Pregnancy associated glycoprotein 2, the first 3D structure and potential activities have been proposed using homology modelling and comparative genomics methods. Then, in order to create the 3D model of buPAG2, the query sequence and template structure were supplied as inputs in MODELLER9v10 [7]. Actin-binding proteins called filamins help to organise the actin-based cytoskeleton in cells. Human filamin was modelled using homology modelling with Modeller 9v5. The final model's reliability was demonstrated through Verify 3D and PROCHECK evaluations. Human filamin isoforms A, B, and C's FASTA sequences were downloaded from Uniprot. The FASTA tool was used to query this sequence against the PDB. Given a protein sequence (target), comparative modelling predicts the 3-D structure of the filamin A, B, and C isoforms based mostly on the alignment to the template (structure determined experimentally) [14].A human stomach lipase (PDB ID: 1HLG) was used as a template to generate the Arabidopsis thaliana lipase homology model (NP 179126). The stereochemical quality and side chain environment of this model was then evaluated [15]. To study the interaction between ligands and substrates, natural substrates such as tributyrin, trioctanoin, and triolen were docked into the model. Swiss-PDB Viewer 4.0.1 was used to create the homology model. This programme has a user-friendly interface that enables simultaneous analysis of many proteins. Using the relevant software, the stereochemical and amino acid environment quality of the *Arabidopsis thaliana* lipase homology model was validated. The pathophysiology of illness involves the enzyme coagulase significantly. The crystal structure of the *Staphylococcus aureus* coagulase is yet unknown. In order to create a three-dimensional model of the coagulase in *S. aureus*, homology modelling is being used in the research. The NCBI database provided the coagulase sequence from *S. aureus* (Accession Number: CAC 84776.1). The Protein Data Bank database was queried using the query sequence from *S. aureus* coagulase to determine the corresponding protein structure that could be utilised as a template (or templates) by the BLAST tool. On the basis of the alignment of the target and template sequences of coagulase, the 3D homology models were constructed using the crystal structure coordinates of the templates. Discovery Studio by Accelerys carried out the procedures (San Diego, CA, USA). Understanding coagulase's native conformation and the mechanism of coagulase action requires an understanding of the three-dimensional structure of the enzyme [20]. For the *Pisum sativum* sieve element occlusion 1 (Ps.SEO1) (forisomes) protein, a homology-based three-dimensional model was created. The 3D structure of Ps.SEO1 was modelled using a region of amino acids (residues 320 to 456) that is well conserved in all known members of the forisomes proteins. Using the Protein Homology/analogY Recognition Engine (PHYRE) web service, the structural prediction was carried out. The thioredoxin-fold containing protein [Structural Classification of Proteins (SCOP) code d1o73a_], a member of the *Glutathione peroxidase* family, was chosen as a template for modelling the spatial structure of Ps.SEO1 based on investigations of local sequence alignment. Comparison of the primary sequence, better match quality, and alignment precision were used to make the decision. Motif 1 (EVF) is conserved in Ps.SEO1, Viciafaba (Vf.For1), and Medicagotruncatula (MT.SEO3); motif 2 (KKED), which is present in all forisome proteins, is also conserved; and motif 3, which is present in Ps.SEO1 and Vf.For1. All known forisomes proteins were shown in the alignment studies to have conserved amino acids (shown by asterisks) and conservative amino acid modifications (marked by dots) along the modelled amino acid stretch [28]. Target sequences for homology modelling included the HIV2 (UniProt entry P04584), SIVagm (UniProt entry Q02836), and SIVmac (UniProt entry P05897) capsid protein sequences. As structural models, the experimental crystal structures of HIV1 (PDB ID 1AFV), N-MLV (PDB ID 1U7K), and B-MLV (PDB ID 3BP9) were employed. The MUSCLE algorithm was used to accomplish multiple sequence alignment between the target and template sequences. The sequence similarity between the capsid proteins of HIV and SIV, which are both members of the lentivirus retroviral genus, ranges from 60 to 86%. The N-MLV and B-MLV capsid proteins share 97% of their sequence among themselves, and MLV is a member of the gammaretroviral genus. Only about 13% of the capsid protein sequences of lenti viruses and gammaretro viruses are identical. The MODELLER programme Version 9 was used to build model structures of the HIV-2, SIVagm, and SIVmac capsid proteins based on target-template sequence alignments. 1. For each protein, 2000 models were created, and the final model was chosen based on the modeler's intended function. The UCSF Chimera program's MatchMaker command was used to carry out the structural sequence alignments. This alignment is based on a combination of secondary structure correspondence and residue identity/similarity. While the latter is calculated using the Kabsch and Sander algorithm, the former is derived using the Needleman-Wunsch algorithm and the BLOSUM-62 residue similarity matrix [18]. The coupling of homology modelling and chemical cross-linking mass spectrometry provides the side-view crossover configuration, provides two basal cylinders view consolidating the crucial functions of the anchoring domains made up of the ApcE PB loop and ApcD which is necessary for regulating photochemical activity [17].

III. CONCLUSION

The most comprehensive method for predicting a protein's three-dimensional structure from its amino acid sequence is homology modelling. This technique creates realistic 3D models. We observed that homology modelling is significant because it discovers linkages between sequence patterns and structural characteristics and further illustrates how proteins have developed. It creates assumptions regarding a protein's function, forecasts how a sequence will fold, and builds a model by comparison with an existing structure with a comparable sequence. It aids in the study of how mutations affect structure, functions and forecasts the impact of a novel mutation on either. Additionally, it creates completely new proteins with inventive functionalities (protein engineering). In the creation of drugs, it is frequently employed. Based on a survey of the literature, we have created Table 1, which includes a list of the applications and programmes used in homology modelling. This table lists software programmes along with a description of the programme and a link to its website. For automatic protein modelling, tools like Geno3d, Swiss Model, CHP models, and Homology are employed. For loop modelling, Wloop is employed. Programming is used in conjunction with Profit, CaSpR, and Phyre 2. The basic goal of homology modelling is to accurately anticipate a structure from its sequence, matching the results of experiments.

Table 1: Tools for Homology Modeling (Comparative Modeling).

Sr. No.	Program	Website address	Program description
1.	Geno3d	http://pbil.ibcp.fr/	Automatic modeling of protein three-dimensional structure
2.	Swiss Model	http://www.expasy.org/swissmod/SWISS-MODEL.html	An automated knowledge-based protein modeling server; first approach and optimize
3.	CPHmodels	Http://www.cbs.dtu.dk/services/CPHmodels/	Automated neural-network based protein modeling server
4.	Modeller	http://salilab.org/	A program for automated protein Homology Modeling
5.	Amber	http://amber.scripps.edu/	Similar package as CHARMm. Developed by Kollaman's group at UCSF
6.	Homology	http://www.accelrys.com/	Automatic Homology Modeling module. The software suite also has Modeller, SeqFold modules,
7.	Wloop	http://psb00.snv.jussieu.fr/wloop/	The Loop Homology Modeling Server
8.	What-If Server	http://www.cmbi.kun.nl/gv/servers/WIWWWI/	V.Friend's What-IF Homology Modeling Server
9.	SPORulate	http://cgat.ukm.my/spores/Predictory/sporulate/s_predict_metaserver.html	Send jobs by 'SPORulation' (meta server) to selected servers available above using the respective server's default values.
10.	Phyre2	www.sbg.bio.ic.ac.uk/phyre2/html	Alignment of hidden Markov models via HHsearch
11.	PROFIT	http://www.proceryon.com/	Used in combination with dynamic programming
12.	COMPOSER	www.tripos.com/data/SYBYL/composer 072505.pdf	Templates are used to provide an average frame work form building the structure
13.	CASP	http://predictioncenter.org	Ensure overall quality of the modles
14.	ModWeb	http://salilab.org/modweb	Set of non-redundant chains extracted from structure in the PDB
15.	CaspR	http://igs-server.cnrs-mrs.fr/Caspr/index.cgi	Used with the protein crystallography and bioinformatics communities

Docking is often utilised to provide a more thorough explanation of knowledge that has already been discovered through experimental research. However, the docking tool can be used more broadly as a way to describe the shape and interfacial characteristics of a protein without attempting to relate the results to experimental data. Docking is not flawless, but we demonstrate here that variations in its level of repeatability can be instructive in and of themselves. Here are few examples of homology modelling being successfully used in drug development. Homology models have aided in the formulation of a number of effective pharmacological drugs in the lack of experimental structures for drug target proteins. The ease and speed with which homology models can be constructed is one of their benefits. Additionally, these models could provide evidence in favour of medicinal chemists' assumptions about how to produce physiologically active molecules during the crucial early conceptual stage of a drug discovery project. One of the advantages of this approach is the design of molecules that are specifically targeted at particular therapeutic target proteins. Such selective substances can even be used to learn more about the physiological function of brand-new medication targets. Being in its infancy, in silico protein structure-based prediction of metabolism and toxicity of small compounds, particularly CYP inhibition and induction and hERG inhibition, may only be able to classify. The homology modelling technique offers one way to fill the gap until comprehensive experimental structures of proteins that are significant from a pharmacological standpoint are available. The integration of AI has played a significant role in the development of homology modeling accuracy. It is evident that the accuracy of results obtained through the homology models can be

improved by adding new developed modules. The improved version of modeling tools will enhance the efficiency of homology modeling. Protein homology detection based on sequence has become one of the most sensitive and exact methods for predicting protein structure. Despite the progress, weakly similar proteins with different evolutionary histories still make homology identification highly difficult [1].

ACKNOWLEDGEMENT

I extend my sincere thanks to Dr. Geeta Puri, Director SILB, Dr. Shalini Sharma, HOD Dept of Biotechnology and Dr. Deepinderpal Singh for giving me proper guidance throughout the course of study. I acknowledge the hard work of Research scholars Ms. Neha and Ms.Anjana.

REFERENCES

[1]. Bhattacharya, S., Roche, R., Shuvo, M. H. and Bhattacharya, D. (2021). Recent advances in protein homology detection propelled by inter-residue interaction map threading. *Frontiers in Molecular Biosciences*, *8*: 643752.

[2]. Cashman, D.J., Ortega, D.R., Zhulin, I.B. and Bautry, J. (2013). Homology modeling of CheW coupling protein of the chemotaxis signaling complex. *Public library of Sciences*, *8*(8): e70705.

[3]. Chu, Y., Li, J., Wu, X., Hua, Z. and Wu, Z. (2013). Identification of human immunodeficiency virus type 1 (HIV-1) gp 120 binding sites on scavenger receptor cysteine rich 1 (SCR) domain of gp 340. *Journal of Biomedical Science, 20*(1): 44.

[4]. Claude, J.B., Suhre, K., Notredama, C., Claverie, J.M. and Abergel, C. (2013). CaspR: a web server for automated molecular replacement using homology modeling. *Oxford Journals*, (32): 606-609.

[5]. Domingues, F.S., Koppensteiner, W.A., Jaritz, M., Prlic, A., Weichenberger, C., Wiederstein, M., Floeckner, H., Lackner, P. and Sippl, M.J. (1999). Sustained performance of knowledge-based potentials in fold recognition. *Proteins Supplements*, (3): 112–120.

[6]. Fan, J. and Papodopoulos, V. (2013). Evolutionary origin of the mitochondrial cholesterol transport machinery reveals a universal mechanism of steroid hormone biosynthesis in animals. *Pubic Library of Science One, 8*(10): e76701.

[7]. Ganguly, B. and Prasad, S. (2012). Homology modeling and functional annotation of bubaline pregnancy associated glycoprotein. *Journal of Animal Science Biotechnology*, *3*(1): 13.

[8]. Greer, J. (1980). Model for haptoglobin heavy chain based upon structural homology. *Proceeding of the National Academic Science USA, 77*(6): 3393-3397.

[9]. Gutierrez, A., Kentsis, A., Sanda, T., Holmfeldt, L., Chen, S.C., Zhang, J., Protopopov, A., Chin, L., Dhalberg, S.E., Neuberg, D.S., Silverman, L.B., Winter, S.S., Hunger, S.P., Sallan, S.E. and Zha, S. (2011). The BCL11B tumor soppresor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia. *Blood, the Journal of the American Society of Hematology*, *118*(15): 4169-4173.

[10]. Haddad, Y., Adam, V., and Heger, Z. (2020). Ten quick tips for homology modeling of high-resolution protein 3D structures. *PLoS computational biology*, *16*(4): e1007449.

[11]. Hameduh, T., Haddad, Y., Adam, V. and Heger, Z. (2020). Homology modeling in the time of collective and artificial intelligence. *Computational and Structural Biotechnology Journal*, *18*: 3494-3506.

[12]. Huang, X., Pearce, R., and Zhang, Y. (2020). FASPR: an open-source tool for fast and accurate protein side-chain packing. *Bioinformatics*, *36*(12): 3758-3765.

[13]. Idrees, S., Nadeen, S., Kanwal, S., Eshsan, B., Yousaf, A., Nadeen, S. and Rajoka, M.I. (2011).*In silico* sequence analysis, homology modeling and function annotation of *Ocimumbasilicum* hypothetical protein GICT28_OCIBA. *International Journal Bioautomation*, *16*(2): 111-118.

[14]. Kumar, S. (2011). Homology modeling and consensus protein disorder prediction of human filamin. *Bioinformation 6*(10): 366-369.

[15]. Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S. and Thornton, J. M. (2018). PDBsum: Structural summaries of PDB entries. *Protein science*, *27*(1): 129-134.

[16]. Lindin, I., Wuxiuer, Y., Kufareva, I., Abagyan, R., Moens, U., Sylte, I. and Ravna, A.W. (2013). Homology modeling and ligand docking of mitogen-activated protein kinase-activated proteins kinase 5MK5. *Theoretical Biology and Medical Modelling*, *10*(1): 1-21.

[17]. Liu, H., Zhang, M. M., Weisz, D. A., Cheng, M., Pakrasi, H. B. and Blankenship, R. E. (2021). Structure of cyanobacterialphycobilisome core revealed by structural modeling and chemical cross-linking. *Science advances*, *7*(2): eaba5743.

[18]. Maillard, P.V., Zoete, V., Michielin, O. and Trono, D. (2011). Homology – based identification of Capsid determination that protect HIV 1 from human TRIM5ά restriction. *Journal of Biological Chemistry*, *286*(10): 8128-8140.

[19]. Mifsud, D. and Bateman, A. (2002). Membrane-bound progesterone receptors contain a cytochrome b5 like ligand-binding domain. *Genome Biology*, *3*(12): 0068.1-0068.5.

[20]. Mohamed, N.A. and Mohamad, R. (2012). Homology modeling of coagulase in Staphylococcus aureus. *Bioinformation*, *8*(9): 412-414.

[21]. Moses, S.A., Ali, M.A., Zuohe, S., Guny , L.D., Zhou, L.L., Lemos, R., Ihle, N., Mash, E.A., Powis, G. and Meuillet, E.J. (2009). In vitro and in vivo activity of novel small molecule inhibitors targeting the pleckstrin homology domain of protein kinase B/AKT. *Cancer Research*, *69*(12): 5073-5081.

[22]. Nema, V. and Pal, S.K. (2013). Exploration of freely available web-interface for comparative homology modeling of microbial protein. *Bioinformatics, 9* (15): 796-801.

[23]. Peitsch, M.C. (1995). ProMod: Automated knowledge-based protein modeling tool PDB Quarterly Newsletter, *72:* 4

[24]. Rahman, K.S., Cui, G., Harvey, S.C. and McCarty, N.A. (2013). Modeling the conformational

changes underlying channel opening in CFTR. *Public Library of Science One*, *8*(9): e74574.
[25]. Sali, A. and Blundell, T.L. (1993). Computational protein modeling by satisfaction of spatial restraints. *Journal of Molecular Biology*, (234): 779-815.
[26]. Singh, P.K. and Shukla, P. (2012). Molecular modeling and docking microbial inulinase towards perceptine enzyme- substrate interactions. *Indian Journal of Microbiology*, *52*(3): 373-380.
[27]. Sippl, M.J. and Weitckus, S. (1992). Detection of native-like models for amino acid sequences of unknown three-dimensional structure in a data base of known protein conformations *Proteins: Structure, Function, and Bioinformatics*, *13*(3): 258-271.
[28]. Tuteja, N. and Umate, P. (2010). Conserved thioredoxin fold is present in *Pisum sativum* L. sieve element occlusion -1 protein. *Plant signal Behavior*, *5*(6): 623-628.
[29]. Van Gunsteren, W., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott W.R.P. and Tironi, IG (1996). Biomolecular simulation: the GROMOS96 manual and user guide. *Vdf Hochschulverlag AG an der ETH Zürich, Zürich*, *86:* 1-1044.
[30]. Venselaar, H., Jooster, R.P., Vroling, B., Baakman, C.A.B., Hekkelman, M.L., Krieger, E. and Vriend, G. (2010). Homology modeling and spectroscopy , a never ending love story. *European Biophysics Journal*, *39*(4): 551-563.
[31]. Vlachakis, D., Kontopoulos, D.G. and Kossida, S. (2013). Space constrained homology modeling: The paradigm of the RNA- dependent RNA polymerase of dengue (Type II)virus. *Computational and mathematical methods in medicine 2013*.
[32]. Vlachakis, D., Kossida, S. (2013). Molecular modeling and pharmacophore elucidation study of the classic swine fever virus helicase as a promising pharmacological target. *Peer J*, (1): e85.
[33]. Vyas, V.K., Ukawala, R.D., Ghaye, M., and Chintha, C. (2012). Homology Modeling a Fast Tool for Drug Discovery: Current perspective. *Indian Journal of Pharmaceutical Science*, *74*(1): 1-7.
[34]. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R. and Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*, *46*(1): 296-303.
[35]. Wieman, H., Tondel, K., Anderssen, E., and Drablos, F. (2004). Homology modeling of targets for drug design. *Mini – Reviews in Medicinal Chemistry*, (4): 793-804.
[36]. Yang, A.S., and Honig, B. (1999). Sequence to structure alignment in comparative modeling using PrISM. *Proteins: Structure, Function, and Bioinformatics*, *37*(S3): 66-72.
[37]. Yang, J., Li, Y., Liu, Q., Li, L., Feng, A., Wang, T. and Lyu, J. (2020). Brief introduction of medical database and data mining technology in big data era. *Journal of Evidence-Based Medicine*, *13*(1): 57-69.
[38]. Zemla, A., Venclovas, Moult J. and Fidelis, K. (2001). Processing and evaluation of predictions in CASP4. *Proteins: Structure, Function, and Bioinformatics, 45*(S5): 13-21.