Review Article

# Homology Modeling a Fast Tool for Drug Discovery: Current Perspectives

V. K. VYAS*, R. D. UKAWALA, M. GHATE AND C. CHINTHA
Department of Pharmaceutical Chemistry, Institute of Pharmacy, Nirma University, Ahmedabad-382 481, India



**Major goal of structural biology involve formation of protein-ligand complexes; in which the protein molecules act energetically in the course of binding. Therefore, perceptive of protein-ligand interaction will be very important for structure based drug design. Lack of knowledge of 3D structures has hindered efforts to understand the binding specificities of ligands with protein. With increasing in modeling software and the growing number of known protein structures, homology modeling is rapidly becoming the method of choice for obtaining 3D coordinates of proteins. Homology modeling is a representation of the similarity of environmental residues at topologically corresponding positions in the reference proteins. In the absence of experimental data, model building on the basis of a known 3D structure of a homologous protein is at present the only reliable method to obtain the structural information. Knowledge of the 3D structures of proteins provides invaluable insights into the molecular basis of their functions. The recent advances in homology modeling, particularly in detecting and aligning sequences with template structures, distant homologues, modeling of loops and side chains as well as detecting errors in a model contributed to consistent prediction of protein structure, which was not possible even several years ago. This review focused on the features and a role of homology modeling in predicting protein structure and described current developments in this field with victorious applications at the different stages of the drug design and discovery.**

**Key words: Drug discovery, GPCRs, homology modeling, ligand design, loop structure prediction, model validation, sequence alignment**

The prediction of the 3D structure of a protein from its amino acid sequence remains a basic scientific problem. This can often achieved using different types of approaches and the first and most accurate approach is "comparative" or "homology" modeling[1]. Homology modeling methods use the fact that evolutionary related proteins share a similar structure[2,3]. Determination of protein structure by means of experimental methods such as X-ray crystallography or NMR spectroscopy is time consuming and not successful with all proteins, especially with membrane proteins[4]. Currently, experimental structure determination will continue to increases the number of newly discovered sequences which grows much faster than the number of structures solved. Currently, 79,356 experimental protein structures are available in the Protein Data Bank (PDB)[5], http://www.rcsb.org/pdb (February 2012). Homology modeling is only the method of choice to generate a reliable 3D model of a protein from its amino acid sequence as notably shown in several meetings of the bi-annual critical assessment of techniques for protein structure prediction (CASP)[6]. Homology modeling is used to search the conformation space by minimally disturbing those existing solutions, i.e., the experimentally solved structures. Homology modeling technique relaxes the tough requirement of force field and enormous conformation searching, because it deals with the calculation of a force field and replaces it in large part, with the counting of sequence identities[7]. The method is based on the fact that structural conformation of a protein is more highly conserved than its amino acid sequence, and that small or medium changes in sequence normally result in little variation in the 3D structure[8]. The process of homology modeling consists of the various steps depicted in fig. 1[9]. These steps may be repeated until suitable models were built. Homology modeling is helpful in molecular biology, such as hypotheses about the drug design[10], ligand binding site[11,12], substrate specificity[13,14], and function annotation[15]. It can also provide starting models for solving structures

***Address for correspondence**
E-mail: vicky_1744@yahoo.com

from X-ray crystallography, NMR and electron microscopy[16,17]. The conformational constancy of homology models of channels may be assessed by subsequent molecular dynamics simulations[18,19]. Homology modeling provides structural insight of protein although quality depends on sequence similarity with the template structure[20]. Quality of model is directly linked with the identity between template and target sequences, as a rule that, models built over 50% sequence similarities are accurate enough for drug discovery applications, those between 25 and 50% identities can be helpful in designing

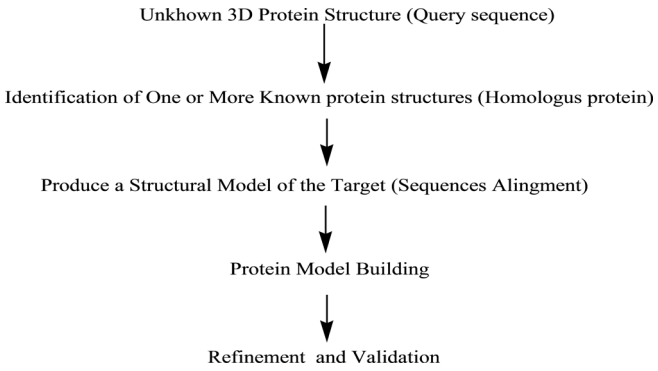


**Fig. 1: Homology modeling process**

of mutagenesis experiments and those in between 10 and 25% are tentative at superlative[21–23]. In the present communication, we reviewed recent advances in the homology modeling methods, and reported some applications of homology modeling to the drug discovery process.

## STEPS IN HOMOLOGY MODELLING

Development of homology model is a multi steps process, that can be summarized in following way (1) identification of template; (2) single or multiple sequence alignments; (3) model building for the target based on the 3D structure of the template; (4) model refinement, analysis of alignments, gap deletions and additions, and (5) model validation[24].

**Template (fold) recognition and alignment:**
This is the initial step in which the program/server compare the sequence of unknown structure with known structure stored in PDB (fig. 2). The most popular server is BLAST (Basic Local Alignment Search Tool)[23] (http://www.ncbi.nlm.nih.gov/blast/). A search with BLAST against the database for

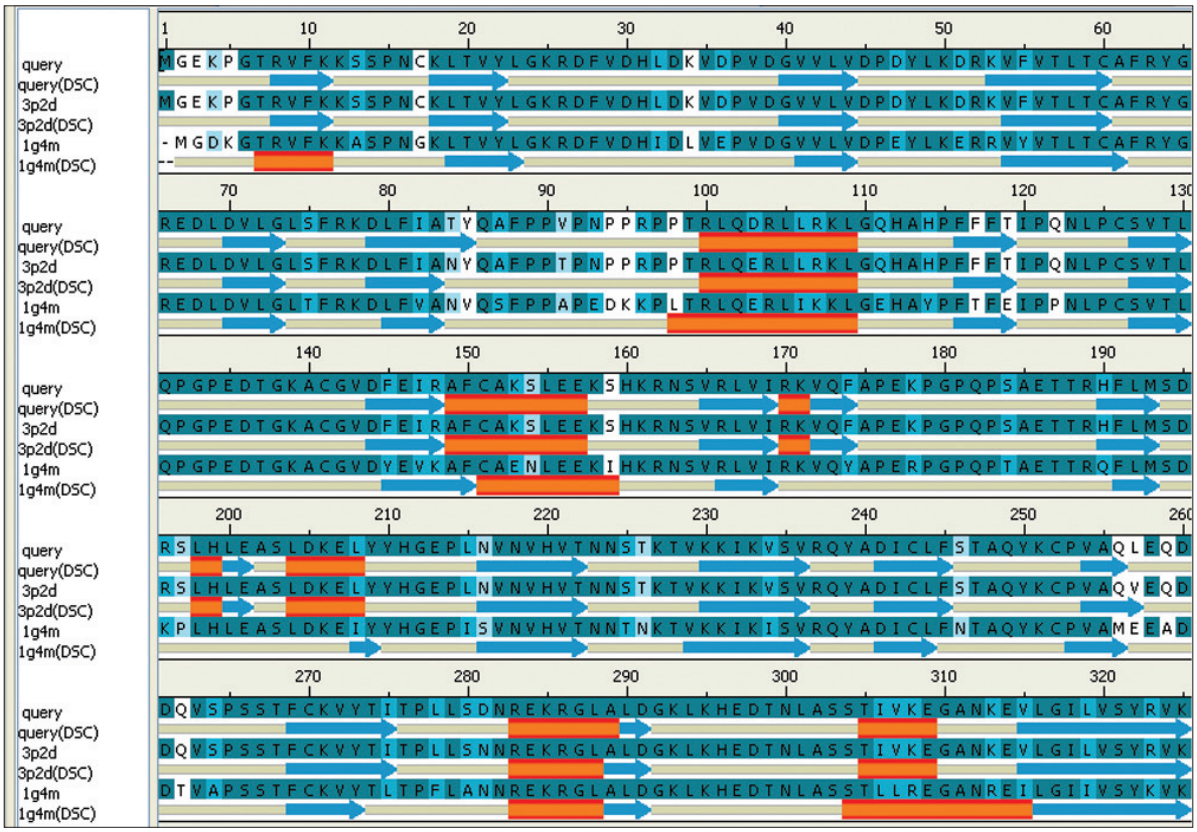


**Fig. 2: Multiple sequence alignment of β-Arrestin family member (query is experimentally derived sequence taken from UNIPROT (ID: P32121) aligned with sequences of PDB entry codes 3P2D and 1G4M. Identical residues, conserved residues are indicated in the form of secondary structure using Discovery Studio Visualiser 2.5)**

optimal local alignments with the query, give a list of known protein structures that matches the sequence. BLAST cannot find a template when the sequence identity is well below 30%; homology hits from BLAST are not reliable. The sequence alignment is more sensitive in detecting evolutionary relationships among proteins and genes[25–27]. The resulting profile–sequence alignment properly align approximately 42-47% of residues in the 0-40% sequence identity range, this number is approximately double than that of the pair wise sequence methods[28,29]. Alignment errors are the main cause of deviations in comparative modeling even when the correct template is chosen. In recent years, significant progress has been made in the development of sensitive alignment methods based on iterative searches, e.g. PSI-BLAST[30], Hidden Markov Models (HMM), e.g. SAM[31],HMMER[32] or profile-profile alignment such as FFAS03[33], profilescan[34] and HHsearch. Multiple alignments are typically heuristic[35] well known as progressive alignment. Progressive alignments are simple to perform and allow large alignments of distantly related sequences to be constructed. This is implemented in the most widely used programs (ClustalW[36] and ClustalX[37]). Alignment of divergent protein sequences can be performed with high accuracy using ClustalW[36] program. ClustalW includes many features like assigning individual weights to each sequence in a partial alignment and amino acid substitution matrices are varied at different alignment stages according to the divergence of the sequences to be aligned. Specific importance is given to residue-specific gap penalties in hydrophilic regions which encourage new gaps in potential loop regions. HMMs[31,32] are a class of probabilistic models that are generally applicable to time series or linear sequence. Profile HMM are very effective in detecting conserved patterns in multiple sequences. The SATCHMO algorithm in the LOBSTER package simultaneously constructs a similarity tree and compares multiple sequence alignments of each internal node of the tree using HMMs. A new HMM, SAM-T98 ID known for finding remote homologs of protein sequences. The method begins with a single target sequence and iteratively builds a HMM from the sequence and homologs found using the HMM for database search. This is also used in construction of model libraries automatically from sequences. The LAMA[38] program aligns two multiple sequence alignments, first by transforming them into profiles and then comparing these two with each other by the Pearson correlation coefficient. The COMPASS[39] program was developed to locally align two multiple sequence with assessment of statistical significance, which compare two profiles by constructing a matrix of scores for matching every position in one profile to each position in the other profile, followed by either local or global dynamic programming to calculate the optimal alignment. T-Coffee uses progressive alignment as optimization technique[40]. T-Coffee can merge heterogeneous data in alignments. 3D Coffee incorporates a link to the FUGUE[41] threading package, which carries out sequence alignment using local structural information. Probabilistic-based program PROBCONS uses BAliBASE[42], which is a most accurate method available for multiple alignments. In simple words PROBCONS is like T-Coffee, but it uses probabilities instead of the heuristic algorithms. HOMSTRAD is exclusively based on sequences with known 3D structures and PDB files. Katoh *et al*.[43] extended the HOMSTRAD by incorporating a large number of close homologues, as found by the BLAST search, which tend to increase the accuracy of the alignment[44]. Sadreyev and Grishin[39] reported that the accuracy of profile alignments can be increased by including confident homologues with the help of COMPASS program. Further knowledge on different programs and server for sequence alignment can be gained by the surfing the URL's provided in the Table 1.

**Model building:**

After the target–template alignment, next step in the homology modeling is the model building. A variety of methods can be used to build a protein model for the target. Generally rigid-body assembly[45–47], segment matching[48], spatial restraint[49], and artificial evolution[50] are used for model building. Rigid-body assembly model building relies on the natural dissection of the protein structure into conserved core regions, variable loops that connect them and side chains that decorate the backbone. Model accuracy is based on the template selection and alignment accuracy. Accordingly, significant modeling method allows a degree of flexibility and automation, making it easier and faster to obtain good models. Segment matching based on the construction of model by using a subset of atomic positions from template structures as guiding positions, and by identifying and assembling short. All-atom segments that match the guiding positions can be obtained either by scanning all the known protein structures. In addition

**TABLE 1: SEQUENCE ALIGNMENT PROGRAMS AND THEIR WEB SERVER SITES**

| Program | Internet address |
|---|---|
| Expresso | http://www.tcoffee.org/ |
| PROMALS3D | http://prodata.swmed.edu/promals3d/ |
| 3D-Coffee | http://igs-server.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi |
| MUSCLE | http://www.drive5.com/muscle/ |
| PROBCONS | http://probcons.stanford.edu/ |
| PRALINE | http://ibivu.cs.vu.nl/programs/pralinewww/ |
| VAST, Cn3D | http://www.ncbi.nlm.nih.gov/Structure |
| ClustalW | http://www.ebi.ac.uk/clustalw/ |
| SAM | http://www.cse.ucsc.edu/research/compbio/sam.html |
| GENEWISE | http://www.sanger.ac.uk/Software/Wise2/ |
| MAFFT | http://align.bmr.kyushu-u.ac.jp/mafft/online/server/ |
| T-Coffee | http://www.tcoffee.org/ |
| PROMALS | http://prodata.swmed.edu/promals/ |
| SPEM | http://sparks.informatics.iupui.edu/Softwares-Services_files/spem.htm |
| PROBE | ftp://ncbi.nlm.nih.gov/pub/neuwald/probe1.0/ |
| BLOCKS | http://www.blocks.fhcrc.org/ |
| PSI-BLAST | http://www.ncbi.nlm.nih.gov/BLAST/newblast.html |

to that it includes those protein structures that are not related to the sequence being modeled[51], or by a conformational search restrained by an energy function[52,53]. Modeling by satisfaction of spatial restraints based on the generation of many constraints or restraints on the structure of target sequence, using its alignment to related protein structures as a guide. Generation of restraints is based upon the assumption the corresponding distances between aligned residues in the template and the target structures are similar.

**Model refinement:**

Model refinement is a very important task that requires efficient sampling for conformational space and a means to accurately identify near-native structures[54]. Homology model building process evolves through a series of amino acid residue substitutions, insertions and deletions. Model refinement is based upon tuning alignment, modeling loops and side chains. The model refinement process will usually begin with an energy minimization step using one of the molecular mechanics force fields[55,56] and for further refinement, techniques such as molecular dynamics, Monte Carlo and genetic algorithm-based sampling can be applied[57,58]. Monte Carlo sampling focused on those regions which are likely to contain errors, while allowing the whole structure to relax in a physically realistic all-atom force field, can significantly improve the accuracy of models in terms of both the backbone conformations and the placement of core side chains. The accuracy of alignment by modeling strongly depends on the degree of sequence similarity. Misalignment of the models some time results into the errors which may be hard to remove at the later stages of refinement[59].

**Loop modeling:**

Homologous proteins have gaps or insertions in sequences, referred to as loops whose structures are not conserved during evolution. Loops are considered as the most variable regions of a protein where insertion and deletion often occur. Loops often determine the functional specificity of a protein structure. Loops contribute to active and binding sites. The accuracy of loop modeling is a major factor in determining the usefulness of homology models for studying protein-ligand interactions[60]. Loop structures are more difficult to predict than the structure of the geometrically highly regular strands and helices because loops exhibit greater structural variability than strands and helices. Length of a loop region is generally much shorter than that of the whole protein chain. Modeling a loop region possess challenges, which are not likely to be present in the global protein structure. Modeled loop structure has to be geometrically consistent with the rest of the protein structure[61].

**Loop prediction methods:**

Loop prediction methods can be evaluated in determining their utilities for: (1) backbone construction; (2) what range of lengths are possible; (3) how widely is the conformational space searched; (4) how side chains are added; (5) how the conformations scored (i.e., the potential energy function) and (6) how much has the method been tested. Most of the loop construction methods were tested only on native structures from which the loop to be built[62,63]. But in reality homology modeling is more complicated process requiring several choices to be made in building the complete structure. The available programs for loop structure prediction along with their web addresses are given in Table 2.

**Database methods:**

Database methods of loop structure prediction measure the orientation and separation of the backbone segments, flanking the region to be modeled, and then search the PDB for segments of the same length that span a region of similar size and orientation. In current years, as the size of the PDB

has increased, database methods have continued to attract attention. Database methods are suitable for the loops of up to 8 residues[64].

**Construction methods:**

The main alternative to database methods is construction of loops by random or exhaustive search mechanisms. Moult and James[65] performed a systematic search to predict loop conformations up to 6 residues long. They found various useful concepts in loop modeling by construction: (1) the use of a limited number of Φ, ψ pairs for construction; (2) construction from each end of the loop simultaneously; (3) discarding conformations of partial loops that span the remaining distance with those residues left to be modeled; (4) using side-chain clashes to reject partial loop conformations and (5) the use of electrostatic and hydrophobic free energy terms in evaluating predicted loops[66].

**Scaling-relaxation method:**

In scaling-relaxation method a full segment is sampled and its end-to-end distance is measured. If this distance is longer than the segment needs, then the segment is scaled in size so that it fits the end-to-end distance of the protein anchors, which result in very short bond distances, and unphysical connections to the anchors. From there, energy minimization is performed on the loop, slowly relaxing the scaling constant, until the loop is scaled back to full size[67,68].

**Molecular mechanics/molecular dynamics:**

Other loop prediction methods build chains by sampling Ramachandran conformations randomly, keeping partial segments as long as they can complete the loop with the remaining residues to be built[69]. These methods are capable of building longer loops since they spend less time in unlikely conformations searched in the grid method. These methods are based on Monte Carlo or molecular dynamics simulations with simulated annealing to generate many conformations, which can then be energy minimized and tested with some energy function to choose the lowest energy conformation for prediction[68,70].

**Side-chain modeling:**

Side-chain modeling is an important step in predicting protein structure by homology. Side-chain prediction usually involve placing side chains onto fixed backbone coordinates either obtained from a parent structure or generated from *ab initio* modeling simulations or a combination of these two. Protein side chains tend to exist in a limited number of low energy conformation called rotamers. In side-chain prediction methods (Table 3), rotamers are selected based on the preferred protein sequence and the given backbone coordinates, by using a defined energy function and search strategy. The side-chain quality can be analyzed by root mean square deviation (RMSD) for all atoms or by detecting the fraction of correct rotamers found[71–73].

**Model validation:**

Each step in homology modeling is reliant on the former processes. Therefore, errors may be accidentally introduced and propagated, thus the model validation and assessment of protein is necessary for interpreting them (Table 4). The protein model can be evaluated as a whole as well as in individual regions[74]. Initially, fold of a model can be assessed by a high sequence similarity with the template. One basic necessity for a constructed model is to

**TABLE 2: LOOP MODELING PROGRAM**

| Loop prediction methods | Internet address |
|---|---|
| BRAGI | http://bragi.gbf.de/index.html |
| BTPRED | http://www.biochem.ucl.ac.uk/bsm/btpred/ |
| RAMP | http://www.ram.org/computing/ramp/ramp.html |
| CONGEN | http://www.congenomics.com/congen/doc/index.html |
| Drawbridge | http://www.cmpharm.ucsf.edu/cohen/ |
| Swiss-PDB Viewer | http://spdbv.vital-it.ch/ |

**TABLE 3: SIDE CHAIN MODELING PROGRAM**

| Side-chain prediction methods | Internet address |
|---|---|
| RAMP | http://www.ram.org/computing/ramp/ramp.html |
| SCWRL | http://www.fccc.edu/research/labs/dunbrack/scwrl |
| Segmod/CARA | http://www.bioinformatics.ucla.edu/~genemine |
| SMD | http://condor.urbb.jussieu.fr/Smd.html |

**TABLE 4: MODEL ASSESSMENT AND VALIDATION PROGRAM**

| Program | Internet address |
|---|---|
| PROCHECK | http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html |
| WHATCHECK | http://www.sander.embl-heidelberg.de/whatcheck/ |
| ProsaII | http://www.came.sbg.ac |
| VERIFY3D | http://www.doe-mbi.ucla.edu/Services/Verify_3D/ |
| ERRAT | http://www.doe-mbi.ucla.edu/Services/Errat.html |
| ANOLEA | http://www.fundp.ac.be/pub/ANOLEA.html |
| Probe | http://kinemage.biochem.duke.edu/software/probe.php |

have good stereochemistry[75]. The most important factor in the assessment of constructed models is scoring function. The programs evaluate the location of each residue in a model with respect to the expected environment as found in the high-resolution X-ray structure[76]. Techniques used to determine misthreading in X-ray structures can be used to determine alignment errors in homology models. Errors in the model are very much common and most attention is needed towards refinement and validation. Errors in model are usually estimated by (1) superposition of model onto native structure with the structure alignment program Structal[77] and calculation of RMSD of Cα atoms[78]; (2) generation of Z-score, a measure of statistical significance between matched structures for the model, using the structure alignment program CE, scores four indicate good structural similarity and (3) development of a scoring function that is capable of discriminating good and bad models. Statistical effective energy functions[79] are based on the observed properties of amino acids in known structures. A variety of statistical criteria derived for various properties such as distributions of polar and apolar residues inside or outside of protein, thus detecting the misfolded models[80]. Solvation potentials can detect local errors and complete misfolds[81]; packing rules have been implemented for structure evaluation[82]. A model is said to be valid only when a few distortions in atomic contacts are present. The Ramachandran plot is probably the most powerful determinant of the quality of protein[83,84], when Ramachandran plot quality of the model is comparatively worse than that of the template, then it is likely that error took place in backbone modeling. WHAT_CHECK determines Asn, His or Gln side chains need to be rotated by 180° about their C2, C2 or C3 angle, respectively. Side chain torsion angles are essential for hydrogen bonding, sometimes altered during the modeling process. Conformational free energy distinguishes the native structure of a protein from an incorrectly folded decoy. A distinct advantage of such physically derived functions is that they are based on well-defined physical interactions, thus making it easier to learn and to gain insight from their performance. In addition, *ab-initio* methods showed success in recent CASP. One of the major drawbacks of physical chemical description of the folding free energy of a protein is that the treatment of solvation required usually comes at a significant computational expense. Fast solvation models such as the generalized born and a variety of simplified scoring schemes[85] may prove to be extremely useful in this regard. A number of freely available programs can be used to verify homology models, among them WHAT_CHECK (Table 4) solves typically crystallographic problems[86]. The validation programs are generally of two types: (1) first category (e.g. PROCHECK and WHATIF) checks for proper protein stereochemistry, such as symmetry checks, geometry checks (chirality, bond lengths, bond angles, torsion angles models[80]. Solvation) and structural packing quality and (2) the second category (e.g.VERIFY3D and PROSAII) checks the fitness of sequence to structure and assigns a score for each residue fitting its current environment. GRASP2 is new model assessment software developed by Honig[87]. For example, gaps and insertions can be mapped to the structures to verify that they make sense geometrically. It is suggested that, manual inspection should be combined with existing programs to further identify problems in the model.

## SOFTWARE FOR HOMOLOGY MODELING

Several programs and servers are available for homology modeling that are planned to build a complete model from query sequences. MODELLER developed by Andrej Sali and colleagues[88,89], SwissModel[90,91], RAMP, PrISM[92], COMPOSER[64,93] CONGEN+2[94,95] and DISGEO/Co-nsensus[96,97] are some of the examples. Homology modeling techniques are described in a number of available programs, both in the commercial and public area (Table 5). A comparative study of available modeling programs and servers (Table 6) for high-accuracy homology modeling has been captured in some excellent publications[98,99]. The authors tried to evaluate several characteristic of the homology modeling programs, including (1) the reliability; (2) the speed by which the programs build models and (3) the similarity of the structure.

### MODELLER:

MODELLER uses the query structures to construct constraints on atomic distances, dihedral angles, and so forth, these are then combined with statistical distributions derived from many homologous structure pairs in the PDB. MODELLER combines the sequences and structures into a complete alignment which can then be examined using molecular graphics programs and edited manually[88,89].

**SwissModel:**

SwissModel is accessible via a web server that accept the sequence to be modeled, and then delivers the model by an electronic mail[100]. In contrast to Modeller, SwissModel follows the standard protocol of homologue identification, sequence alignment, determining the core backbone and modeling loops and side chains. SwissModel will search a sequence database of proteins in the PDB with BLAST, and will attempt to build a model for any PDB hits[90,91].

**TABLE 5: SERVER AND PROGRAMS USEFUL IN HOMOLOGY MODELING**

| Programs Name | WWW address | Availability |
|---|---|---|
| SWISS-MODEL* | http://swissmodel.expasy.org/ | Academically free |
| MODELLER** | http://salilab.org/modeller/ | Academically free |
| ExPASy * | http://www.expasy.ch/tools/ | Academically free |
| BLAST* | http://blast.ncbi.nlm.nih.gov/Blast.cgi | Academically free |
| SCHRODINGER** | http://www.schrodinger.com | Commercial |
| WHATIF* | http://swift.cmbi.kun.nl/whatif/ | Academically free |
| SYBYL** | http://www.tripos.com | Commercial |
| SNPWEB* | http://modbase.compbio.ucsf.edu/LS-SNP/ | Academically free |
| ICM | http://www.molsoft.com/homology.html | Academically free |
| SEA * | http://bioinformatics.burnham.org/sea/ | Academically free |
| SCWRL* | http://www1.jcsg.org/prod/scripts/scwrl/serve.cgi | Academically free |
| EVA* | http://rostlab.org/cms/index.php?id=94 | Academically free |
| VERIFY3D* | http://nihserver.mbi.ucla.edu/Verify_3D/ | Academically free |
| MOE** | http://www.chemcomp.com/software.htm | Commercial |
| GOLD** | http://www.ccdc.cam.ac.uk/products/life_sciences/gold/ | Commercial |
| PROCHECK* | http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/ | Academically free |

*Server, **Program

**PrISM:**

PrISM performs homology modeling using alignment to builds a composite template by selecting each secondary structure from the most appropriate template. *Ab initio* methods are used for loop modeling and side-chain dihedrals are taken either from the template or predicted structure based on main chain torsion angles and a neural network algorithm[92].

**COMPOSER:**

COMPOSER uses multiple template structures for building homology models. If a target sequence is related to more than one template (of different sequence) then all templates are used to provide an average framework for building the structure[64,93].

**CONCEN:**

CONCEN develops distance restraints from the template structure and the sequence alignment of the target and template for atoms. These atoms include all backbone atoms and side-chain atoms of the same chemical type and hybridization state. No homologous atoms are defined in the loop regions[94,95].

**Critical assessment of techniques for protein structure prediction (CASP):**

CASP experiments are biannual and their main aim is to set benchmarking standards to the protein structure prediction methods followed by various online servers and software. They monitor the state of the art in modeling protein structure from the sequence. The main objective behind these experiments is to ensure overall quality of the models, accuracy of prediction and evaluating the parameters provided by the various tools. The independent assessors evaluate predictions using a battery of numerical criteria[101]. There are several conclusions drawn from the CASP such as, the comparative modeling remained most accurate technique for protein structure modeling when compared with others. Although majority of predictions were again closer to the template than to the real structure,

**TABLE 6: COMPARISON OF SOFTWARE FOR HOMOLOGY MODELING**

| Modeling program | Potential energy | Search method | Description |
|---|---|---|---|
| MODELLER | CHARMM | Spatial restraints | Modeling by satisfying spatial restraints |
| SwissModel | GROMOS | Rigid-body assembly | Web server using rigid-body assembly with loop modeling |
| COMPOSER | | Rigid-body assembly | Use of multiple template structures for building homology model |
| 3D-JIGSAW | Mean-field minimization methods | Rigid-body assembly | Web server using rigid-body assembly with loop modeling |
| PrISM | | Rigid-body assembly | Most appropriate template is used for each segment of the target to be built |
| CONGEN | CHARMM | Rigid-body assembly | Distance constraints derived from known structure and alignment. |

there has been improvement in some cases. However, accurate modeling by using a single template is not possible and model refinement has always been a challenge till date. Drastic changes are being done to the algorithm to meet the standards by various tools. Eighteen successful refinements of model coordinates to a value closer to the experimental structure were observed in CASP 7. Model refinement has been identified as an area in which further developments are required to be done[102–105]. Readily available models for a given sequence, such as those generated by automated servers often form the basis for the input to model refinement methods. Previous attempts have included molecular dynamics[106], Monte Carlo[107] and knowledge-based techniques[108]. However automated prediction servers have found to improve their algorithms and models predicted were closer to the experimentally determined. The function prediction category (FN) was introduced in the 6th CASP (Table 7), where predictions for gene ontology molecular function terms, enzyme commission numbers and ligand-binding site residues were evaluated. These services, EVA[109], LiveBench[110] and CASP are very useful for the protein structure prediction community, giving clear observations of the development and the need for further progression within the field. Results of several CASP experiments and evaluations are made publicly available through the prediction centre website (http://predictioncenter.org). The latest advances in structure prediction and assessment of model quality are to be evaluated by CASP 10th in the year 2012.

## APPLICATIONS OF HOMOLOGY MODELING

Homology modeling is widely used in structure based drug design process. The importance of homology modeling is increasing as the number of available crystal structures increases. There are several other common applications of homology models: (1) studying the effect of mutations[111]; (2) identifying active and binding sites on protein (useful for ligand design)[112]; (3) searching for ligands of a given binding site (database mining)[113]; (4) designing novel ligands of a given binding site; (5) modeling substrate specificity[114]; (6) predicting antigenic epitopes[115]; (7) protein–protein docking simulations[116]; (8) molecular replacement in X-ray structure refinement[117]; (9) rationalizing known experimental observations[118] and (10) planning new computational experiments with the provided models. Typical applications of a homology model in drug discovery require a very high accuracy of the local side chain positions in the binding site. A very large number of homology models have been built over the years. Targets have included antibodies[119] and many proteins involved in human biology and medicine[120,121].

**TABLE 7: CASPS RESULTS IN THE FUNCTION PREDICTION CATEGORY (FN)**

| Rank | CASP ID | Name |
|---|---|---|
| 1 | FN096 | ZHANG |
| 2 | FN339 | I-TASSER_FUNCTION |
| 3 | FN315 | FIRESTAR |
| 4 | FN242 | SEOK |
| 5 | FN035 | CNIO-FIRESTAR |
| 6 | FN110 | STERNBERG |
| 7 | FN104 | JONES-UCL |
| 8 | FN094 | MCGUFFIN |
| 9 | FN113 | FAMSSEC |
| 10 | FN114 | LEE |

### Case study of G-protein coupled receptors (GPCRs):

GPCRs constitute the largest family of signalling receptors in the cell and therefore being target for nearly half of all drug discovery programs. In the year 2000 only a single crystal structure was available, bovine rhodopsin (bRho) (PDB code 1f88, 1l9h), before which bacterio-rhodopsin was used for modeling. Recently, the appearance of crystal structures of four new GPCRs (Opsin, 3cap; $\beta_2$ adrenergic ($\beta_2$-AR), 2rh1; turkey b1 adrenergic ($\beta_1$-AR), 2vt4; human A2A adenosine receptor, 3eml) brings a broader template diversity for *in silico* modeling. The newly crystallized $\beta_2$-AR has been already investigated as an alternative template to model other Class-A GPCRs for drug discovery applications[122]. Analyzing the bovine rhodopsin structure with the human $\beta_2$-adrenergic receptor (2rh1) gives basis for understanding some facts and drawbacks of the modeling techniques used earlier for GPCR's[123]. Arrangement of the seven trans-membrane helix segments is generally correctly represented, and significant differences was observed in the relative orientation and shifts of the helices with regard to the centre of the receptor. Most deviations are observed for helices III, V and the extracellular loop ECL2, which connects helices IV and V, while ECL2 is forming a β-sheet structure in rhodopsin. $\beta_2$-adrenergic receptor contains an unexpected additional α-helical segment and a second disulfide bridge that might stabilize the more solvent exposed conformation. Consequently, specific interactions between the ligand

molecule and side chains forming the binding pocket are only partially reproduced by a comparative model based on rhodopsin. A novel ligand-steered homology modeling method was presented recently[124], in which the information about known ligands is explicitly used to shape and optimize the binding site through a docking-based stochastic global energy minimization procedure[125–128]. This method is useful to reduce the uncertainty in modeling the binding site, as both the ligand and receptor are held flexible during modeling. A combination of homology modeling and molecular dynamics studies on known inhibitors crystallized with other homologous proteins was used to shape and optimize the binding site of the ribosomal S6 kinase 2 (RSK2), target for human breast and prostate cancer. Subsequent docking reported two low micromolar inhibitors[129]. Homology models have proved to be an important source to rationalize SAR data and predict binding modes of compounds like cannabinoid receptor-2[130,131], human adenosine A2A receptor[132] and alpha-1-adrenoreceptors[133].

**Homology model-based ligand design:**

The Applications of homology modelling in ligand designing is given in Table 8. Watts *et al*.[134] generated two homology models of the gastric $H^+/K^+$-ATPase in the $E_1$ and $E_2$ conformations of its catalytic cycle based on templates provided by its related P-type ATPase ($Ca^{2+}$-ATPase). Generated models were based on the CLUSTALW alignment, *G*-factor ranks values above −0.5 as positive candidates for homology models. In this study, the values were found to be exceeding −0.5 and ranged from −0.34 to −0.05. Correlation between the results of ligand docking and existing mutagenesis information for the protein showed that the models are realistic and could reveal an insight into the binding mechanism for a class of site-specific reversible inhibitors of the gastric $H^+/K^+$-ATPase. A 3D model of the $AT_1$ receptor was constructed[135] using X-ray structure of bovine rhodopsin as template. The site-directed mutagenesis data was also taken into account. Docking based alignment was used for the development of a 3D-QSAR model for several non-peptide antagonists. The results of this study confirmed the binding hypothesis and reliability of the model. Cavasotto *et al*.[136] developed a ligand-steered homology modeling approach followed by docking-based virtual screening to model melanin-concentrating hormone receptor 1 (MCH-R1). MCH-R1 is a GPCR and a target for obesity. Ligand-steered homology modeling method was applied to shape and optimize the binding site, which reduced the uncertainty in its structural characterization by homology modeling. The authors reported that the absence of solid experimental evidence has led them to use small-scale virtual screening for model validation. They evaluated the accuracy of the models by estimating the ability to discriminate binders and nonbinders in a virtual screening of known MCH-R1 antagonists seeded within a GPCR class A ligand library. Wiest *et al*.[137] constructed 3D models of class I histone deacetylases, HDAC1, HDAC2, HDAC3, and HDAC8 for understanding of differences between the isoforms of class-I HDAC. A series of HDAC inhibitors were docked to understand the similarities and differences between the binding modes. The results of the study helped in design of novel HDAC inhibitors. 3D structure of Fyn kinase was modeled[138] using the Sybyl-Composer. Rosmarinic acid was evaluated as a new Fyn kinase inhibitor using immunochemical and *in silico* methods. In the process to identify possible active site of homology model for docking of the ligands, PDB data was searched for solved crystal structures of tyrosine kinases in complex with ligands. The crystal structure of Lck complexed with staurosporine (1QPJ) was spatially aligned with model of the Fyn. Active sites of both structures were carefully analyzed, and the most important differences in the residue positions were identified. The study reported that the rosmarinic acid binds to the second "non-ATP" binding site of the Fyn tyrosine kinase.

**TABLE 8: APPLICATIONS OF HOMOLOGY MODELING RELEVANT TO LIGAND DESIGN**

| Protein structure | Program/Server | Reference |
|---|---|---|
| Gastric $H^+/K^+$-ATPase | MODELLER, PROCHECK, PROFIT, AUTODOCK3.0, CLUSTALW, ProsaII3.0 | [134] |
| $AT_1$ receptor | MODELLER, AUTODOCK, PROCHECK, MACROMODEL | [135] |
| Melanin-concentrating hormone receptor 1 (MCH-R1) | MODELLER | [136] |
| Histone deacetylases (HDACs) | MODELER 7, WHAT IF, AutoDock 3.0, BLAST | [137] |
| Fyn tyrosine kinase | SYBYL, BLASTP, GOLD, FlexX | [138] |
| Acetyl CoA carboxylase | SWISS-MODEL, CLUSTALW, PROCHECK, Sybyl 7.0 | [139] |
| Human histamine $H_4$ receptor ($H_4R$) | MOE | [140] |
| Human dopamine ($D_{2L}$ and $D_3$) receptors | MODELLER 9.2, SYBYL7.2, PROCHECK | [141] |
| Dopamine $D_2$ receptor | Clustal X 2.09, BLAST, MODELER, SYBYL | [142] |

Yang *et al*.[139] constructed homology models of the carboxyl-transferase domain of acetyl-coenzyme A carboxylase from sensitive and resistant foxtail and used these models as templates to study the molecular mechanism and stereochemistry-activity relationships of aryloxyphenoxypropionates (APPs). Further docking analysis using the Insight II program indicated that the binding model of highly active compounds was similar to that in the crystal structure of enzyme-ligand complexes. A 3D protein model of the human histamine $H_{4R}$ has been constructed by Leurs *et al*.[140] using rhodopsin as template. The derived computational model was able to explain the experimental data obtained for several mutant receptors at the fundamental atomic level. All simulations were performed using Amber99 force field in MOE 2006.08. Reichert *et al*.[141] reported homology models for both human $D_{2L}$ and $D_3$ receptors in complex with haloperidol using MODELLER 9.2. Further structure and ligand-based approach was explored for a class of $D_2$-like dopamine receptor ligands. 3D-QSAR analyses were performed to explore the intermolecular interactions of a large library of ligands with dopamine $D_2$/$D_3$ receptors. Chen *et al*.[142] employed molecular dynamics simulation techniques to identify the predicted $D_2$ receptor structure. Homology models of the protein were developed on the basis of crystal structures of four available receptor crystals. Clustal X 2.09 program was used for sequence alignment and MODELLER program was used to model $D_2$ receptor. Docking studies revealed the possible binding mode and five other residues (Asp72, Val73, Cys76, Leu183 and Phe187) which were responsible for the selectivity of the tetralindiol derivatives. The result of this study revealed that constructed novel models can be used to design new protease antagonists.

**Structure-based homology modeling:**

The structure based homology modelling studied are given in Table 9. Nowak *et al*.[143] constructed rhodopsin-based homology model of 5-$HT_{1A}$ serotonin receptor. The crystal structure of bovine rhodopsin was used as a template structure. Modeller was used to produce 400 models and a cyclohexylarylpiperazine derivative was docked to all the 400 receptor models using FlexX. Ligand binding mode in the 5-$HT_{1A}$ receptor was analyzed based on top-scored ligand-receptor complexes. The main objectives of this study was to validate the model with a decreased conformational flexibility, which encoded the information on a shape of binding site and the spatial arrangement of specific interaction points within the binding pocket. The anaplastic lymphoma kinase (ALK) is a receptor tyrosine kinase normally expressed in neural tissues during embryogenesis is a valid target for anticancer therapy. In the absence of a resolved crystal structure of ALK, Passerini and colleagues[144] generated homology models of the ALK kinase domain in different conformational states. The authors observed that mutation of the leucine residue in ALK to a smaller threonine residue, which was found sufficient to allow binding of the inhibitors inside the ATP pocket and consequently, inhibition of mutated ALK. Evers *et al*.[145] modeled alpha1A receptor based on the X-ray structure of bovine rhodopsin (template). The authors applied a modified version of the MOBILE approach (modelling binding sites including ligand information explicitly), which modeled protein by homology including information about bound ligands as restraints, thus resulting in more relevant geometries of protein binding sites. Virtual screening study identified putative alpha1A receptor antagonists. The authors mentioned that

**TABLE 9: APPLICATIONS OF STRUCTURE-BASED HOMOLOGY MODELING**

| Protein | Program/Server | Reference |
|---|---|---|
| Serotonin 5-$HT_{1A}$ Receptor | MODELLER 7v7, SYBYL 7.0, FlexX | [143] |
| Anaplastic lymphoma kinase (ALK) | MODELLER 6v2, FlexX | [144] |
| Alpha1A adrenergic receptor | MOE, MOBILE, GOLD2.0, Sybyl 6.92 | [145] |
| Human CCR5 chemokine receptor | MODELLEER 6.2, NMRCLUST, $GridHex_{38}$ | [146] |
| Carbonic Anhydrase IX | MODELLER 9v1, PROCHECK, GOLD | [147] |
| Leishmanial Farnesyl Pyrophosphate Synthases | MODELLER 7.0, SWISS-PROT, BLAST, Sybyl 6.9, InsightII | [148] |
| hASIC1a ion channel | MODELLER 9v4, DOT 2.0, SHAKE | [149] |
| Aminergic GPCRs Dopamine ($D_2$, $D_3$, and $D_4$), serotonin (5-$HT_{1B}$, 5-$HT_{2A}$,5-$HT_{2B}$, and 5-$HT_{2C}$), histamine ($H_1$), and muscarinic ($M_1$) receptors | MODELLER, Sybyl, Prime, ICM | [150] |
| Cytochrome $P_{450}$ sterol 14α-demethylase | AMBER 8.0, CPHmodels2.0, FlexX/GOLD, PROCHECK, SYBYL 7.0, PMF, DOCK | [151] |
| Dopamine ($D_3$) receptor | Modeller 9v2, GOLD | [152] |
| Human galactose-1-phosphateuridylyltransferase | MODELLER, PROCHECK | [153] |
| Human serum carnosinase | FlexX, STRIDE | [154] |

among the 80 top-scored hits, 37 revealed affinity below 10 μM, with 24 compounds binding in the submicromolar range. Chemokine receptors (CCRs) are the members of the GPCRs, identified as potential target systems for preventing virus-cell fusion. The authors[146] modeled a 3D structure of the human CCR5 from the bovine rhodopsin by incorporating extensive molecular dynamics simulations (MD), flexible docking of a synthetic antagonist and soft protein-protein docking with the large (70 KD) natural agonists using some novel docking protocol. The results of this combined modeling, dynamics and docking study provides new structural insights into CCR5/chemokine interactions, which may be useful in the rational design of HIV-1 entry blockers. Tuccinardi *et al*.[147] developed a homology model of carbonic anhydrase (CA) IX using the X-ray structure of murine CA XIV as a template. CA IX constitutes an interesting target for cancer therapy. The authors docked one twenty four CA IX inhibitors, and the best poses were used for developing a receptor-based 3D-QSAR model. The results of the study suggested structural peculiarities which can be useful for the design of new CA IX active and CA IX/CA II selective ligands. Avery *et al.*[148] generated highly refined homology models of *Leishmania donovani* farnesyl pyrophosphate synthase (*Ld*FPPS) and *Leishmania major* FPPS (*Lm*FPPS) enzyme using *Trypanosoma cruzi* FPPS (*Tc*FPPS) as a reference structural homologue. The authors suggested that highly refined model along with the validated docking and scoring algorithms could be utilized to identify hits with novel scaffolds as antileishmanial agents. Pietra *et al*.[149] constructed homology models of acid-sensing cation-permeable, ligand-gated ion channel (*h*ASIC1) using the crystal structure of the *c*ASIC1 channel as a template and the known sequence of *h*ASIC1a. ASIC channels are under intense scrutiny for their ability in sensing proton gradients. Psalmotoxin 1 (PcTx1) - a peptide isolated from the venom of the aggressive Trinidad chevron tarantula (*Psalmopoeus cambridge*) is a inhibitor of ASIC. The results of the study showed the way to *in silico* search for improved peptides, for blocking ASIC1a channels. Yuriev *et al*.[150] constructed homology models of dopamine ($D_2$, $D_3$ and $D_4$), serotonin (5-$HT_{1B}$, 5-$HT_{2A}$, 5-$HT_{2B}$ and 5-$HT_{2C}$), histamine ($H_1$), and muscarinic ($M_1$) receptors using $\beta_2$-adrenergic receptor. The authors performed induced fit docking for binding site optimization and virtual screening of known ligands and decoys. The study addressed the required modeling of extracellular loop 2, which is implicated in ligand binding. Feng *et al*.[151] generated homology models of cytochrome $P_{450}$ sterol 14R-demethylases (CYP51s) from *Penicillium digitatum* (PD-CYP51). The preceding 3D structure of PD-CYP51 was further subjected to a molecular dynamic (MD) study to reduce steric clashes and obtain converged 3D modeling structure of PD-CYP51. After active site generation docking-based virtual screening was performed using FlexX/GOLD. The seven new hit compounds with comparable inhibitory activities were identified. Zhang *et al*.[152] constructed homology model of human $D_3$ receptor using the X-ray crystal structure of the human $\beta_2$AD receptor (PDB entry: 2RH1, resolution 2.4 Å) as the template structure. The authors performed combined computational study to investigate the agonist binding to the $D_3$ receptor, which is important for the design of potent $D_3$ receptor agonists. Marabotti *et al*.[153] constructed homology models of human galactose-1-phosphate uridylyltransferase (GALT). The genetic disorder called "classical galactosemia" or "galactosemia I" is associated with the impairment of GALT. Mutation is associated with genetic galactosemia. The authors analyzed the impact of this mutation both on enzyme-substrate interactions as well as on inter-chain interactions. It was concluded from the study that constructed model will be useful for characterization of all galactosemia-linked mutations at a molecular level. Inhibition of serum carnosinase may be a useful therapeutic approach in the treatment of diabetic nephropathy. Vistoli *et al*.[154] constructed homology models of human serum carnosinase on the basis of β-alanine synthetase structure. Homology model was validated by docking a few histidine-containing dipeptides, later on, molecular dynamics (MD) simulations were used to examine the effects of citrate ions on the activity of serum carnosinase.

**Loop structure prediction:**

The homology modelling is very good tool for prediction of loop structures, the exaples are given in Table 10. Loops frequently resolved the functional specificity of a protein environment, thus contribute to active and binding sites. Zhang *et al*.[155] developed homology models of the lysophosphatidic acid ($LPA_4$) receptor based on the X-ray crystal structures of photoactivated bovine rhodopsin (PDB code 1U19). GOLD, was employed to dock LPA molecule into the homology models of the receptors. It was observed that three non-conserved amino acid residues engaged

**TABLE 10: APPLICATIONS OF HOMOLOGY MODELING RELEVANT TO LOOP STRUCTURE PREDICTION**

| Protein | Program/Servers | Reference |
|---|---|---|
| Lysophosphatidic acid $LPA_4$ receptor | SYBYL v7.3, GOLD 3.1, SCWRL 3.0 | [155] |
| Farnesyl pyrophosphate synthase | PSI-BLAST, Swiss-PDBViewer v3.7, WHATCHECK, AMBER 7.0 | [156] |
| Varicella zoster virus thymidine kinase (VZV TK) | SYBYL 6.8, AUTODOCK 3.0, SHAKE, AMBER, PROCHECK | [157] |
| Glycogen synthase kinase (GSK3/SHAGGY) | FASTA, INSIGHT-II, PROFILE-3D | [158] |
| Hemoglobin-binding protein HgbA | PROCHECK, WHAT IF, Modeller | [159] |

in hydrogen bonding interactions with the polar head group of the LPA molecule. These hydrogen bonding patterns were found to contribute significantly to the recognition of LPA within the $LPA_4$ receptor. Turjanski *et al*.[156] modeled the structure of *Trypanosoma cruzi* fanesyl pyrophosphate synthase (*Tc*FPPS) based on the structure of the avian FPPS which share 36% identity and 50% similarity with the sequence of *Tc*FPPS. The authors constructed the model using Swiss PDBViewer version 3.7. The authors modeled the interaction of *Tc*FPPS with isopentenyl pyrophosphate and dimethylallyl pyrophosphate. Based on the study, authors have proposed specific role for the third $Mg^{2+}$ in closing of the protein active site based on molecular dynamics simulations (MD). Scapozza *et al*.[157] constructed homology models of *Varicella Zoster* virus thymidine kinase (VZV TK) based on herpes simplex virus type 1 thymidine kinase (HSV-1 TK) structure as template. Acyclovir and ganciclovir were docked in the constructed model to investigate the predictivity of these model as well as the characteristics of the binding with other substrates. It was found that there are slight differences in the way VZV TK binds the substrates in respect with HSV-1 TK. Missing loops in the VZV TK was modeled using the loop search routine of SYBYL 6.8. The study suggested that differences could be exploited for future ligand design in order to obtain more selective drugs. Li *et al*.[158] built homology models for a glycogen synthase kinase (GSK3)/SHAGGY-like kinase based on the known crystal structure of glycogen synthase kinase-3β (Gsk3 β PDB code:1I09). Initial module of GSK-3β was obtained with the help of FASTA program. Binding pocket of GSK3/SHAGGY-like kinase was determined by binding site search module. Several variable regions (loops) were constructed using loop searching algorithm. Optimization of structure was done by INSIGHT-II and PROFILE-3D. The authors[159] constructed homology models of hemoglobin-binding protein HgbA from *Actinobacillus pleuropneumoniae* using BtuB, FepA, FhuA and FecA of *Escherichia coli* as template structure. Using HMM the authors assigned β strands to regions of predicted HgbA amino acid sequence, culminating in a structure-based multiple sequence alignment to BtuB, FepA, FhuA, and FecA. 3D model generated from this alignment provides an overall topology of HgbA and identifies extracellular loop regions.

**Miscellaneous applications of homology modeling for protein structure prediction:**

Pillai *et al*.[160] provided a structural basis for the biological functions of human Smad5 by building a model of the DNA-binding domain of it. The authors reported similarities and differences between the human Smad family members using the constructed model. Gellert *et al*.[161] constructed homology models of cucumber mosaic virus (CMV) strains R, M and Trk7, tomato aspermy virus (TAV) strain P and peanut stunt virus (PSV) strain Er, using Fny-CMV CP subunit B as a template. Models were analyzed by the PROCHECK program and electrostatic potential calculations were applied to all models. Guo *et al*.[162] developed 3D model of human phosphate mannose isomerase based on the known crystal structure of mannose-6-phosphate isomerase (PDB code: 1PMI). The homologous protein was searched by the FASTA program and 3 reference proteins were taken i.e. mannose-6-phosphate isomerase from *C. albicans, B. subtilis* and human. The sequence alignment was done based on identification of structurally conserved regions (SCRs). This study facilitated the understanding of the mode of action of the ligands and guided further genetic studies. Reddanna *et al*.[163] generated homology models of human 12R-LOX structure, based upon rabbit reticulocyte 15-Lipoxygenase 1LOX as a template. The 3D model was built using Modeller and BLAST, ClustalW for sequence alignment, AMBER 3.0 for refinement and PROCHECK was used for validation of themodel. The authors[164] built 3D structure of the chorismate synthase (CS) from *S. flexneri* with the cofactor FMN and using MODELLER 6 v.2. The AMBER 8.0 program and AMBER 2003 force field were used for molecular simulation. CS is a valid target for antibacterial drugs. Hirashima *et al*.[165] suggested homology models of octopamine receptor ($OAR_2$) of *Periplanata americana* using DS Modeling 1.1 using rhodopsin as a template. BLAST and PSI-BLAST

were used for alignment, and docking study was performed using LigandFit. These models can be used in scheming new leads for $OAR_2$ receptors. Kayastha *et al*.[166] constructed homology models of α-amylase from germinated mung beans (*Vigna radiata)* using the automated Swissmodel server where two known structure of amylase AMY1 and AMY2 were chosen as a templates. The sequence identity between target and templates is over 65%. The Ramachandran Z-score for the model is -1.132. Serrano *et al*.[167] presented model for the 3D structure of the C-terminal 19 kDa fragment of *P. vivax* MSP-1, based on the known crystal structure of *P. cynomolgy* $MSP\text{-}1_{19}$. The presence of a main binding pocket was determined by CASTp and the combination of GOLD docking and scoring functions was used to observe the interaction between Akt PH (protein kinase β pleckstrin) domain and its inhibitors. Park *et al*.[168] constructed homology models for yeast β-glycosidase using BLAST, PSI BLAST and MODELLER. Furthermore, the author performed virtual screening with docking simulations to study the effects of ligand solvation in the binding free energy function which results in 13 novel α-glucosidase inhibitors. Wang *et al*.[170] used human cytochrome $P_{450}$ 2C8 (CYP2C8) as template and created human cytochrome $P_{450}$ 2C11 (CYP2C11) and human cytochrome $P_{450}$ 2C13 (CYP2C13) models. The authors demonstrated the pattern of testosterone binding with various human cytochrome $P_{450}$ enzymes. Rigid structure and induced-fit docking proposed that testosterone binds in both CYP2C11 and CYP2C13. These results demonstrated the binding of substrate to CYPs. Kotra *et al*.[171] modeled two new epidermal growth factor receptors (EGFR) taking SYK tyrosine kinase coordinates as template. Mutation was incorporated at G695S and L834R to develop the new receptor structures. This study determined the receptor-inhibitor interactions and thus provides rational approach to design and development of potent inhibitors. Construction of homology models of dipeptide epimerase suggested novel enzymatic functions[172]. Docking of dipeptide library against the binding site of these models was performed using Glide. The dipeptide library was prepared by using Ligprep. Homology modeling of trihydroxynaphthalene reductase (3HNR) was performed with 17b-HSDcl as a template, which possesses 58% identical residues. Molecular modelling package WHATIF was used along with the CHARMM program for macromolecular simulations. The study of 3HNR explained the binding modes of the ligand with the model. Chavatte *et al*.[173] constructed models for both human melatoninergic ($MT_1$ and $MT_2$) receptors by homology modeling using the X-ray structure of bovine rhodopsin as template. Models were checked using Ramachandran plots to assess the quality of structure. No residue lies in the disallowed part of the plots and very few residues are in the less favourable regions. Thus, constructed models can be explored at an atomic level for the melatoninergic receptors. Yao *et al*.[174] studied the two metabolic pathway for gliclazide by building homology models of human cytochrome $P_{450}$ 2C9 (CYP2C9) and human cytochrome $P_{450}$ 2C19 (CYP2C19) enzyme. Structural optimization was performed using molecular mechanics and molecular dynamics simulations. To further check the reliability of the 3D structure of models, the automated molecular docking was performed using docking program Insight II. It was found that affinity for methylhydroxylgliclazide pathway found to be more than 6β-hydroxylgliclazide with respect to both refined model of CYP2C19 and crystal structure of CYP2C9.

## CONCLUSION

Structure-based drug design techniques were hampered in the past by the lack of a crystal structure for the target protein. In this instance, now a day the best option is building a homology model of the entire protein. The main aim of homology modeling is to predict a structure from its sequence with an accuracy that is similar to the results obtained experimentally. Homology modeling provides a feasible cost-effective alternative method to generate models. Homology modeling studies are fastened through the use of visualization technique, and the differential properties of the proteins can be discovered. The role and reliability of homology model building will continue to grow as the number of experimentally determined structures increases. Homology modeling is a powerful tool to suggest modeling of ligand-receptor interactions, enzyme-substrate interactions, mutagenesis experiments, SAR data, lead optimization, loop structure prediction and to identify hits. Homology modeling strongly relies on the virtual screening and successful docking results. Various examples of the successful applications of homology modeling in drug discovery are described in this review. These recent advances should help to improve our knowledge of understanding the role of homology modeling in drug discovery process.

## REFERENCES


1. Cavasotto CN, Phatak SS. Homology modeling in drug discovery: current trends and applications. Drug Discov Today 2009;14:676-83.
2. Chandonia JM, Brenner SE. Implications of structural genomics target selection strategies: Pfam5000, whole genome, and random approaches. Proteins 2005;58:166-79.
3. Vitkup D, Melamud E, Moult J, Sander C. Completeness in structural genomics. Nat Struct Biol 2001;8:559-66.
4. Floudas CA, Fung HK, McAllister SR, Mönnigmann M, Rajgaria R. Advances in protein structure prediction and de novo protein design: A review. Chem Eng Sci 2006;61:966-88.
5. Westbrook J, Feng Z, Chen L, Yang H, Berman HM. The Protein Data Bank and structural genomics. Nucleic Acids Res 2003;31:489-91.
6. Tramontano A, Morea V. Assessment of homology-based predictions in CASP5. Proteins 2003;53:352-68.
7. Johnson MS, Srinivasan N, Sowdhamini R, Blundell TL. Knowledge-based protein modeling. Crit Rev Biochem Mol Biol 1994;29:168.
8. Lesk AM, Chothia CH. The response of protein structures to amino-acid sequence changes. Philos Trans R Soc London Ser B 1986;317:345-56.
9. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct 2000;29:291-325.
10. Zhou Y, Johnson ME. Comparative molecular modeling analysis of-5-amidinoindole and benzamidine binding to thrombin and trypsin: specific H-bond formation contributes to high 5-amidinoindole potency and selectivity for thrombin and factor Xa. J Mol Recognit 1999;12:235-41.
11. Jung JW, An JH, Na KB, Kim YS, Lee W. The active site and substrates binding mode of malonyl-CoA synthetase determined by transferred nuclear Overhauser effect spectroscopy, site-directed mutagenesis, and comparative modeling studies. Protein Sci 2000;9:1294-303.
12. De Rienzo F, Fanelli F, Menziani MC, De Benedetti PG. Theoretical investigation of substrate specificity for cytochromes P450 IA2, P450 IID6 and P450 IIIA4. J Comput-Aided Mol Des 2000;14:93-116.
13. Ginalski K, Rychlewski L, Baker D, Grishin NV. Protein structure prediction for the male-specific region of the human Y chromosome. Proc Natl Acad Sci 2004;101:2305-10.
14. Talukdar AS, Wilson DL. Modeling and optimization of rotational C-arm stereoscopic X-ray angiography. IEEE Trans Med Imaging 1999;18:604-16.
15. Ceulemans H, Russell RB. Fast fitting of atomic structures to low-resolution electron density maps by surface overlap maximization. J Mol Biol 2004;338:783-93.
16. Teichmann SA, Chothia C, Gerstein M. Advances in Structural Genomics. Curr Opin Struct Biol 1999;9:390-9.
17. Capener CE, Shrivastava IH, Ranatunga KM, Forrest LR, Smith GR, Sansom MS. Homology modelling and molecular dynamics simulation studies of an inward rectifier potassium channel. Biophys J 2000;78:2929-42.
18. Law RJ, Capener C, Baaden M, Bond PJ, Campbell J, Patargias G. *et al*. Membrane protein structure quality in molecular dynamics simulation. J Mol Graph Model 2005;24:157-165.
19. Oshiro C, Bradley EK, Eksteowicz J, Evensen E, Lamb ML, Lanctot JK, *et al*. Performance of 3D-database molecular docking studies into homology models. J Med Chem 2004;47:768-71.
20. Hilbert M, Bohm G, Jaenicke R. Structural relationships of homologous proteins as a fundamental principle in homology modeling. Proteins 1993;17:138-51.
21. Takeda-Shitaka M, Takaya D, Chiba C, Tanaka H, Umeyama H. Protein structure prediction in structure based drug design. Curr Med Chem 2004;11:551-8.
22. Francoijs CJ, Klomp JP, Knegtel RM. Sequence annotation of nuclear receptor ligand-binding domains by automated homology modeling. Protein Eng 2000;13:391-4.
23. Wong WC, Stroh SM, Eisenhaber F. Not all transmembrane helices are born equal: Towards the extension of the sequence homology concept to membrane proteins. Biol Direct 2011;6:1-30.
24. Joo K, Lee J, Lee J. Methods for accurate homology modeling by global optimization. Methods Mol Biol 2012;857:175-88.
25. Lindahl E, Elofsson A. Identification of related proteins on family, superfamily and fold level. J Mol Biol 2000;295:613-25.
26. Park J, Karplus K, Barrett C, Hughey R, Haussler D, Hubbard T, *et al*. Sequence comparisons using multiple sequences detect three times as many remote homologues as pairwise methods. J Mol Biol 1998;284:1201-10.
27. Sauder JM, Arthur JW, Dunbrack RL Jr. Large-scale comparison of protein sequence alignment algorithms with structure alignments. Proteins 2000;40:6-22.
28. Gribskov M, Luthy R, Eisenberg D. Profile analysis. Methods Enzymol 1990;183:146-59.
29. Skolnick J, Fetrow JS. From genes to protein structure and function: novel applications of computational approaches in the genomic era. Trends Biotechnol 2000;18:34-9.
30. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, *et al*. Gapped BLAST and PSIBLAST: A new generation of protein database search programs. Nucl Acids Res 1997;25:3389-402.
31. Karplus K, Barrett C, Hughey R. Hidden Markov models for detecting remote protein homologies. Bioinformatics 1998;14:846-56.
32. Eddy SR. Profile hidden Markov models. Bioinformatics 1998;14:755-63.
33. Jaroszewski L, Rychlewski L, Li Z, Li W, Godzik A. FFAS03: A server for profile – profile sequence alignments. Nucleic Acids Res 2005;33:W284-8.
34. Marti-Renom MA, Madhusudhan MS, Sali A. Alignment of protein sequences by their profiles. Protein Sci 2004;13:1071-87.
35. Feng DF, Doolittle RF. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J Mol Evol 1987;25:351-60.
36. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673-80.
37. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25:4876-82.
38. Pietrokovski S. Searching databases of conserved sequence regions by aligning protein multiple-alignments. Nucl Acids Res 1996;24:3836-45.
39. Sadreyev RI, Grishin NV. Quality of alignment comparison by COMPASS improves with inclusion of diverse confident homologs. Bioinformatics 2004;20:818-28.
40. Notredame C, Higgins DG, Heringa J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 2000;302:205-17.
41. Shi J, Blundell TL, Mizuguchi K. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J Mol Biol 2001;310:243-57.
42. Do CB, Mahabhashyam MS, Brudno M, Batzoglou S. ProbCons: probabilistic consistency-based multiple alignment of amino acid sequences. Genome Res 2005;15:330-40.
43. Katoh K, Kuma KI, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 2005;33:511-8.
44. Thompson JD, Plewniak F, Poch O. A comprehensive comparison of multiple sequence alignment programs. J Mol Biol 1999;27:2682-90.
45. Collura V, Higo J, Garnier J. Modeling of protein loops by simulated annealing. Protein Sci 1993;2:1502-10.
46. Browne WJ, North ACT, Phillips DC, Brew K, Vanaman TC, Hill RC. A possible three-dimensional structure of bovine α-lactalbumin based on that of hen's eggwhite lysozyme. J Mol Biol 1969;42:65-86.
47. Greer J. Comparative modelling methods: application to the family of the mammalian serine proteases. Proteins 1990;7:317-34.
48. Levitt M. Accurate modelling of protein conformation by automatic

segment matching. J Mol Biol 1992;226:507-33.
49. Rost B. Twilight zone of protein sequence alignments. Protein Eng 1999;12:85-94.
50. Xiang Z. Advances in homology protein structure modeling. Curr Protein Pept Sci 2006;7:217-27.
51. Claessens M, Van-Cutsem E, Lasters I, Wodak S. Modelling the polypeptide backbone with 'spare parts' from known protein structures. Protein Eng 1989;2:334-45.
52. Holm L, Sander C. Database algorithm for generating protein backbone and side-chain co-ordinates from a C alpha trace application to model building and detection of co-ordinate errors. J Mol Biol 1991;218:183-94.
53. Bruccoleri RE, Karplus M. Conformational sampling using high-temperature molecular dynamics. Biopolymers 1990;29:1847-62.
54. Van-Gelder CW, Leusen FJ, Leunissen JA, Noordik JH. A molecular dynamics approach for the generation of complete protein structures from limited coordinate data. Proteins 1994;18:174-85.
55. Zhu J, Fan H, Periole X, Honig B, Mark AE. Refining homology models by combining replica-exchange molecular dynamics and statistical potentials. Proteins 2008;72:1171-88.
56. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, *et al*. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179-97.
57. Das R, Qian B, Raman S, Vernon R, Thompson J, Bradley P, *et al.* Structure prediction for CASP7 targets using extensive all-atom refinement with Rosetta. Proteins 2007;69(S8):118-28.
58. Han R, Leo-Macias A, Zerbino D, Bastolla U, Contreras-Moreira B, Ortiz AR. An efficient conformational sampling method for homology modeling. Proteins 2008;71:175-88.
59. Misura KM, Baker D. Progress and challenges in high-resolution refinement of protein structure models. Proteins 2005;59:15-29.
60. Lee J, Lee D, Park H, Coutsias EA, Seok C. Proteinloop modeling by using fragment assembly and analytical loop closure. Proteins 2010;78:3428-36.
61. Michalsky E, Goede A, Preissner R. Loops in proteins (LIP)—a comprehensive loop database for homology modelling. Protein Eng 2003;16:979-85.
62. Jones TA, Thirur S. Using known substructures in protein model building and crystallography. EMBO J 1986;5;819-22.
63. Blundell TL, Sibanda BL, Sternberg MJE, Thornton JM. Knowledge-based prediction of protein structures and the design of novel molecules. Nature 1987;326:347-52.
64. Sutcliffe MJ, Haneef I, Carney D, Blundell TL. Knowledge based modeling of omologous proteins, part I: three-dimensional frameworks derived from the simultaneous superposition of multiple structures. Prot Eng 1987;5:377-84.
65. Moult J, James MN. An algorithm for determining the conformation of polypeptide segments in proteins by systematic search. Proteins 1986;1:146-63.
66. Pellequer JL, Chen SW. Does conformational free energy distinguish loop conformations in proteins? Biophys J 1997;73:2359-75.
67. Fine RM, Wang H, Shenkin PS, Yarmush DL, Levinthal C. Predicting antibody hypervariable loop conformations. IL Minimization and molecular dynamics studies of MCPC603 from many randomly generated loop conformations. Proteins 1986;1:342-62.
68. Sowdhamini R, Rufino SD, Blundell TL. A database of globular protein structural domains: clustering of representative family members into similar folds. Fold Des 1996;1:209-20.
69. Shenkin PS, Yarmush DL, Fine RM, Wang HJ, Levinthal C. Predicting antibody hypervariable loop conformation. I. Ensembles of random conformations for ringlike structures. Biopolymers 1987;26:2053-85.
70. Shirai H, Nakajima N, Higo J, Kidera A, Nakamura H. Conformational sampling of CDR-H3 in antibodies by multicanonical molecular dynamics simulation. J Mol Biol 1998;278:481-96.
71. Mendes J, Baptista AM, Carrondo MA, Scares CM. Improved modeling of side-chains in proteins with rotamer-based methods: A flexible rotamer model. Proteins 1999;37:530-43.
72. Liang S, Grishin NV. Side-chain modeling with an optimized scoring function. Protein Sci 2002;11:322-31.
73. Canutescu AA, Shelenkov AA, Dunbrack RL Jr. A graph-theory algorithm for rapid protein side-chain prediction. Protein Sci 2003;12:2001-14.
74. Mart-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct 2000;29:291-325.
75. Hillisch A, Pineda LF, Hilgenfeld R. Utility of homology models in the drug discovery process. Drug Disc Today 2004;9:659-69.
76. Petrey D, Honig B. Protein structure prediction: Inroads to biology. Mol Cell 2005;20:811-9.
77. Gerstein M, Levitt M. Comprehensive assessment of automatic structural alignment against a manual standard, the scop classification of proteins. Protein Sci 1998;7:445-56.
78. Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. EMBO J 1986;5:823-6.
79. Sippl M. Knowledge-based potentials for proteins. Curr Opin Struct Biol 1995;5:229-35.
80. Baumann G, Froemmel C, Sander C. Polarity as a criterion in protein design. Prot Eng 1989;2:329-34.
81. Holm L, Sander C. Evaluation of protein models by atomic solvation preference. J Mol Biol 1992;225:93-105.
82. Gregoret LM, Cohen FE. Novel method for the rapid evaluation of packing in protein structures. J Mol Biol 1990;211:959-74.
83. Laskowski RA, MasArthur MW, Moss DS. Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283-91.
84. Hooft RW, Sander C, Vriend G, Abola E. Errors in protein structures. Nature 1996;381:272.
85. Petrey D, Honig B. Free energy determinants of tertiary structure and the evaluation of protein models. Protein Sci 2000;9:2181-91.
86. Huang E, Subbiah S, Levitt M. Recognizing native folds by the arrangement of hydrophobic and polar residues. J Mol Biol 1995;252:709-20.
87. Petrey D, Honig B. GRASP2: visualization, surface properties, and electrostatics of macromolecular structures and sequences. Methods Enzymol 2003;374:492-509.
88. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. Mol Biol 1993;234:779-815.
89. Sanchez R, Sali A. Evaluation of comparative protein structure modeling by MODELLER-3. Proteins 1997;1:50-8.
90. Peitsch MC. ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. Biochem Soc Trans 1996;24:274-9.
91. Peitsch MC. Large scale protein modelling and model repository. Proc Int Conf Intell Syst Mol Biol 1997;5:234-6.
92. Yang AS, Honig B. Sequence to structure alignment in comparative modeling using PrISM. Proteins 1999;37:66-72.
93. Sutcliffe MJ, Hayes FR, Blundell TL. Knowledge based modeling of homologous proteins. Part II: Rules for the conformations of substituted side chains. Prot Eng 1987;1:385-92.
94. Li H, Tejero R, Monleon D, Bassolino-Klimas D, Abate-Shen C, Bruccoleri RE, *et al*. Homology modeling using simulated annealing of restrained molecular dynamics and conformational search calculations with CONGEN: application in predicting the three-dimensional structure of murine homeodomain Msx-1. Protein Sci 1997;6:956-70.
95. Bower MJ, Cohen FE, Dunbrack RL. Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. J Mol Biol 1997;267:1268-82.
96. Havel TF, Snow ME. A new method for building protein conformations from sequence alignments with homologues of known structure. Mol Biol 1991;217:1-7.
97. Havel TF. Predicting the structure of the fiavodoxin from *Escherichia coli* by homology modeling, distance geometry, and molecular dynamics. Mol Simul 1993;10:175-210.
98. Nayeem A, Sitkoff D, Krystek S Jr. A comparative study of available software for high-accuracy homology modeling: From sequence

alignments to structural models. Protein Sci 2006;15:808-24.
99. Wallner B, Elofsson A. All are not equal: A benchmark of different homology modeling programs. Protein Sci 2005;14:1315-27.
100. Guex N, Peitsch MC. SWISS-MODEL and the Swiss- PdbViewer: an environment for comparative protein modeling. Electrophoresis 1997;18:2714-23.
101. Zemla A, Venclovas, Moult J, Fidelis K. Processing and evaluation of predictions in CASP4. Proteins 2001;45(Suppl 5):13-21.
102. Valencia A. Protein refinement: a new challenge for CASP in its 10th anniversary. Bioinformatics 2005:21;277.
103. Moult J. A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction. Curr Opin Struct Biol 2005;15:285-9.
104. Kryshtafovych A, Venclovas C, Fidelis K, Moult J. Progress over the first decade of CASP experiments. Proteins 2005;61(Suppl 7):225-36.
105. Tress M, Ezkurdia I, Graña O, López G, Valencia A. Assessment of predictions submitted for the CASP6 comparative modeling category. Proteins 2005;61(Suppl 7):27-45.
106. Lee MR, Tsai J, Baker D, Kollman PA. Molecular dynamics in the endgame of protein structure prediction. J Mol Biol 2001;313:417-30.
107. Bradley P, Malmström L, Qian B, Schonbrun J, Chivian D, Kim DE, *et al*. Free modeling with Rosetta in CASP6. Proteins 2005;61(Suppl 7):128-34.
108. Qian B, Ortiz AR, Baker D. Improvement of comparative model accuracy by free-energy optimization along principal components of natural structural variation. Proc Natl Acad Sci USA 2004;101:15346-51.
109. Eyrich VA, Martí-Renom MA, Przybylski D, Madhusudhan MS, Fiser A, Pazos F, *et al*. EVA: continuous automatic evaluation of protein structure prediction servers. Bioinformatics 2001;17:1242-3.
110. Rychlewski L, Fischer D. LiveBench-8: the large-scale, continuous assessment of automated protein structure prediction. Protein Sci 2005;14:240-5.
111. Boissel JP, Lee WR, Presnell SR, Cohen FE, Bunn HF. Erythropoietin structure-function relationships. Mutant proteins that test a model of tertiary structure. J Biol Chem 1993;268:15983-93.
112. Sheng Y, Sali A, Herzog H, Lahnstein J, Krilis S. Modelling, expression and site-directed mutagenesis of human β2- glycoprotein I. Identification of the major phospholipid binding site. J Immunol 1996;157:3744-51.
113. Ring CS, Sun E, McKerrow JH, Lee GK, Rosenthal PJ, Kuntz ID, Cohen FE. Structure-based inhibitor design by using protein models for the development of antiparasitic agents. Proc Natl Acad Sci U S A 1993;90:3583-7.
114. Xu LZ, S´anchez R, Sali A, Heintz N. Ligand specificity of brain lipid binding protein. J Biol Chem 1996;271:24711-9.
115. Sali A, Matsumoto R, McNeil HP, Karplus M, Stevens RL. Three dimensional models of four mouse mast cell chymases. Identification of proteoglycan-binding regions and protease-specific antigenic epitopes. J Biol Chem 1993;268:9023-34.
116. Vakser IA. Evaluation of GRAMM low-resolution docking methodology on the hemagglutinin-antibody complex. Proteins 1997;29(Suppl 1): 226-30.
117. Howell PL, Almo SC, Parsons MR, Hajdu J, Petsko GA. Structure determination of turkey egg-white lysozyme using Laue diffraction data. Acta Crystallogr 1992;48:200-7.
118. Guenther B, Onrust R, Sali A, O'Donnell M,Kuriyan J. Crystal structure of the δ subunit of the clamp-loader complex of *E. coli* DNA polymerase III. Cell 1997;91:335-45.
119. Mas MT, Smith KC, Yarmush DL, Aisaica K, Fine RM. Modeling the anti-CEA antibody combining site by homology and conformational search. Proteins 1992;14:483-98.
120. Sahasrabudhe PV, Tejero R, Kitao S, Furuichi Y, Montelione GT. Homology modeling of an RNP domain from a human RNA-binding protein: Homology-constrained energy optimization provides a criterion for distinguishing potential sequence alignments. Proteins 1998;33:558-66.
121. Weber T. Evaluation of homology modeling of HIV protease. Proteins 1990;7:172-84.
122. Costanzi S. On the applicability of GPCR homology models to computer aided drug discovery: a comparison between in silico and crystal structures of the beta2-adrenergic receptor. J Med Chem 2008;51:2907-14.
123. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, *et al*. High-resolution crystal structure of an engineered human β2-adrenergic G protein coupled receptor. Science 2007;318:1258-65.
124. Katritch V, Rueda M, Lam PC, Yeager M, Abagyan R. GPCR 3D homology models for ligand screening: Lessons learned from blind predictions of adenosine A2a receptor complex. Proteins 2010;78:197-211
125. Cavasotto CN, Abagyan RA. Protein flexibility in ligand docking and virtual screening to protein kinases. J Mol Biol 2004;337:209-25.
126. Kovacs JA, Cavasotto CN, Abagyan R. Conformational sampling of protein flexibility in generalized coordinates: application to ligand docking. J Comput Theor Nanosci 2005;2:354-61.
127. Monti MC, Casapullo A, Cavasotto CN, Napolitano A, Riccio R. Scalaradial, a dialdehyde-containing marine metabolite that causes an unexpected noncovalent PLA2 inactivation. Chembiochem 2007;8:1585-91.
128. Monti MC, Casapullo A, Cavasotto CN, Tosco A, Dal Piaz F, Ziemys A, *et al*. The binding mode of petrosaspongiolide M to the human group. IIA phospholipase A(2): exploring the role of covalent and noncovalent interactions in the inhibition process. Chemistry 2009;15:1155-63.
129. Nguyen TL, Gussio R, Smith JA, Lannigan DA, Hecht SM, Scudiero DA, *et al*. Homology model of RSK2 N-terminal kinase domain, structure-based identification of novel RSK2 inhibitors, and preliminary common pharmacophore. Bioorg Med Chem 2006;14:6097-105.
130. Diaz P, Phatak SS, Xu J, Astruc-Diaz F, Cavasotto CN, Naguib M. 6-Methoxy-N-alkyl isatin acylhydrazone derivatives as a novel series of potent selective cannabinoid receptor 2 inverse agonists: design, synthesis, and binding mode prediction. J Med Chem 2009;52:433-44.
131. Diaz P, Phatak SS, Xu J, Fronczek FR, Astruc-Diaz F, Thompson CM, *et al*. 2,3-Dihydro-1-benzofuran derivatives as a novel series of potent selective cannabinoid receptor 2 agonists: design, synthesis, and binding mode prediction through ligand-steered modeling. Chem Med Chem 2009;4:1615-29.
132. Michielan L, Bacilieri M, Schiesaro A, Bolcato C, Pastorin G, Spalluto G, *et al*. Linear and nonlinear 3D-QSAR approaches in tandem with ligand-based homology modeling as a computational strategy to depict the pyrazolo-triazolo-pyrimidine antagonists binding site of the human adenosine A2A receptor. J Chem Inf Model 2008;48:350-63.
133. Li M, Fang H, Du L, Xia L, Wang B. Computational studies of the binding site of alpha1Aadrenoceptor antagonists. J Mol Model 2008;14:957-66.
134. Kim CG, Watts JA, Watts A. Ligand docking in the gastric H+/K+-ATPase: homology modeling of reversible inhibitor binding sites. J Med Chem 2005;48:7145-52.
135. Tuccinardi T, Calderone V, Rapposelli S, Martinelli A. Proposal of a new binding orientation for non-peptide at1 antagonists: Homology modeling, docking and three-dimensional quantitative structure-activity relationship analysis. J Med Chem 2006;49:4305-16.
136. Cavasotto CN, Orry AJ, Murgolo NJ, Czarniecki MF, Kocsi SA, Hawes BE, *et al*. Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. J Med Chem 2008;51:581-8.
137. Wang D, Helquist P, Wiect NL, Wiest O. Toward selective histone deacetylase inhibitor design: homology modeling, docking studies, and molecular dynamics simulations of human class I histone deacetylases. J Med Chem 2005;48:6936-47.
138. Jelic D, Mildner B, Kostrun S, Nujic K, Verbanac D, Cÿulic O, *et al.* Homology modeling of human fyn kinase structure: discovery of rosmarinic acid as a new fyn kinase inhibitor and *in silico* study of its possible binding modes. J Med Chem 2007;50:1090-100.

139. Zhu X, Zhang L, Chen Q, Wan J, Yang G. Interactions of aryloxyphenoxypropionic acids with sensitive and resistant acetyl-coenzyme A carboxylase by homology modeling and molecular dynamic simulations. J Chem Inf Model 2006;46:1819-26.
140. Jongejan A, Lim HD, Smits RA, de Esch IJ, Haaksma E, *et al*. Delineation of agonist binding to the human histamine H4 receptor using mutational analysis, homology modeling, and ab initio calculations. J Chem Inf Model 2008;48:1455-63.
141. Wang Q, Mach RH, Luedtke RR, Reichert DE. Subtype selectivity of dopamine receptor ligands: insights from structure and ligand-based methods. J Chem Inf Model 2010;50:1970-85.
142. Wang Y, Su Z, Hsieh C, Chen C. Predictions of binding for dopamine D2 receptor antagonists by the SIE method. J Chem Inf Model 2009;49:2369-75.
143. Nowak M, Kołaczkowski M, Pawłowski M, Bojarski AJ. Homology modeling of the serotonin 5-HT1A receptor using automated docking of bioactive compounds with defined geometry. J Med Chem 2006;49:205-14.
144. Gunby RH, Ahmed S, Sottocornola R, Gasser M, Redaelli S, Mologni L, *et al*. Structural insights into the ATP binding pocket of the anaplastic lymphoma kinase by site-directed mutagenesis, inhibitor binding analysis, and homology modeling. J Med Chem 2006;49:5759-68.
145. Evers A, Klabunde T. Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the Alpha1A adrenergic receptor. J Med Chem 2005;48:1088-97.
146. Fano A, Ritchie DW, Carrieri A. Modeling the structural basis of human CCR5 chemokine receptor function: from homology model building and molecular dynamics validation to agonist and antagonist docking. J Chem Inf Model 2006;46:1223-35.
147. Tuccinardi T, Ortore G, Rossello A, Supuran CT, Martinelli A. Homology modeling and receptor-based 3D-QSAR study of carbonic anhydrase IX. J Chem Inf Model 2007;47:2253-62.
148. Mukherjee P, Desai PV, Srivastava A, Tekwani BL, Avery MA. Probing the structures of leishmanial farnesyl pyrophosphate synthases: homology modeling and docking studies. J Chem Inf Model 2008;48:1026-40.
149. Pietra F. Docking and MD simulations of the interaction of the tarantula peptide psalmotoxin-1 with ASIC1a channels using a homology model. J Chem Inf Model 2009;49:972-7.
150. McRobb FM, Capuano B, Crosby IT, Chalmers DK, Yuriev E. Homology modeling and docking evaluation of aminergic G protein-coupled receptors. J Chem Inf Model 2010;50:626-37.
151. Zhang Q, Li D, Wei P, Zhang J, Wan J, Ren Y, *et al*. Structure-based rational screening of novel hit compounds with structural diversity for cytochrome P450 sterol 14r-demethylase from *Penicillium digitatum*. J Chem Inf Model 2010;50:317-25.
152. Cui W, Wei Z, Chen Q, Cheng Y, Geng L, Zhang J, *et al*. Structure-based design of peptides against G3BP with cytotoxicity on tumor cells. J Chem Inf Model 2010;50:380-7.
153. Marabotti A, Facchiano AM. Homology modeling studies on human galactose-1-phosphate uridylyltransferase and on its galactosemia-related mutant Q188R provide an explanation of molecular effects of the mutation on homo- and heterodimers. J Med Chem 2005;48:773-9.
154. Vistoli G, Pedretti A, Cattaneo M, Aldini G, Testa B. Homology modeling of human serum carnosinase, a potential medicinal target, and MD simulations of its allosteric activation by citrate. J Med Chem 2006;49:3269-77.
155. Li G, Mosier PD, Fang X, Zhang Y. Toward the three-dimensional structure and lysophosphatidic acid binding characteristics of the LPA4/p2y9/GPR23 receptor: A homology modeling study. J Mol Graphics Modell 2009;28:70-9.
156. Sigman L, Sanchez VM, Turjanski AG. Characterization of the farnesyl pyrophosphate synthase of *Trypanosoma cruzi* by homology modeling and molecular dynamics. J Mol Graphics Modell 2006;25:345-52.
157. Spadola L, Novellino E, Folkers G, Scapozza L. Homology modelling and docking studies on Varicella Zoster Virus Thymidine kinase. Eur J Med Chem 2003;38:413-9.
158. Xiao J, Li Z, Sun M, Zhang Y, Sun C. Homology modeling and molecular dynamics study of GSK3/SHAGGY-like kinase. Comput Biol Chem 2004;28:179-88.
159. Pawelek PD, Coulton JW. Hemoglobin-binding protein HgbA in the outer membrane of *Actinobacillus pleuropneumoniae*: homology modelling reveals regions of potential interactions with hemoglobin and heme. J Mol Graph Model 2004;23:211-21.
160. Hariharan R, Pillai MR. Homology modeling of the DNA-binding domain of human Smad5: A molecular model for inhibitor design. J Mol Graph Model 2006;24:271-7.
161. Gellert A, Salanki K, Naray-Szabo G, Balazs E. Homology modelling and protein structure based functional analysis of five cucumovirus coat proteins. J Mol Graph Model 2006;24:319-27.
162. Xiao J, Guo Z, Guo Y, Chu F, Sun P. Computational study of human phosphomannose isomerase: Insights from homology modeling and molecular dynamics simulation of enzyme bound substrate. J Mol Graph Model 2006;25:289-95.
163. Aparoy P, Leela T, Reddy RN, Reddanna P. Computational analysis of R and S isoforms of 12-Lipoxygenases: Homology modeling and docking studies. J Mol Graph Model 2009;27:744-50.
164. Zhou H, Singh NJ, Kim KS. Homology modeling and molecular dynamics study of chorismate synthase from *Shigella flexneri*. J Mol Graph Model 2006;25:434-41.
165. Hirashima A, Huang H. Homology modeling, agonist binding site identification, and docking in octopamine receptor of *Periplaneta americana*. Comput Biol Chem 2008;32:185-90.
166. Tripathi P, Leggio LL, Mansfeld J, Ulbrich-Hofmann R, Kayastha AM. α-Amylase from mung beans (*Vigna radiata*) Correlation of biochemical properties and tertiary structure by homology modeling. Phytochemistry 2007;68:1623-31.
167. Serrano ML, Perez HA, Medina JD. Structure of C-terminal fragment of merozoite surface protein-1 from *Plasmodium vivax* determined by homology modeling and molecular dynamics refinement. Bioorg Med Chem 2006;14:8359-65.
168. Park H, Hwang KY, Oh KH, Kim YH, Lee JY, Kim K. Discovery of novel α-glucosidase inhibitors based on the virtual screening with the homology-modeled protein structure. Bioorg Med Chem 2008;16:284-92.
169. Wang H, Cheng JD, Montgomery D, Cheng KC. Evaluation of the binding orientations of testosterone in the active site of homology models for CYP2C11 and CYP2C13. Biochem Pharmacol 2009;78:406-13.
170. Kotra S, Madala KK, Jamil K. Homology models of the mutated EGFR and their response towards quinazolin analogues. J Mol Graph Model 2008;27:244-54.
171. Kalyanaraman C, Imker HJ, Fedorov AA, Fedorov EV, Glasner ME, Babbitt PC, *et al*. Discovery of a dipeptide epimerase enzymatic function guided by homology modeling and virtual screening. Structure 2008;16:1668-77.
172. Brunskole M, Stefane B, Zorko K, Anderluh M, Stojan J, Rizner TL, *et al*. Towards the first inhibitors of trihydroxynaphthalene reductase from *Curvularia lunata*: Synthesis of artificial substrate, homology modelling and initial screening. Bioorg Med Chem 2008;16:5881-9.
173. Farce A, Chugunov AO, Loge C, Sabaouni A, Yous S, Dilly S, *et al*. Homology modeling of MT1 and MT2 receptors. Eur J Med Chem 2008;43:1926-44.
174. Yao Y, Han W, Zhou Y, Li Z, Li Q, Chen X, *et al*. The metabolism of CYP2C9 and CYP2C19 for gliclazide by homology modeling and docking study. Eur J Med Chem 2009;44:854-61.


**Accepted 26 February 2012**
**Revised 24 February 2012**
**Received 20 June 2011**
**Indian J. Pharm. Sci., 2012, 74 (1): 1-17**