

RESEARCH ARTICLE

Computational Analysis of the 3-D structure of Human GPR87 Protein: Implications for Structure-Based Drug Design

Mukta Rani¹, Anuradha Nischal¹, Ganesh Chandra Sahoo², Sanjay Khattri^{1*}

Abstract

The G-protein coupled receptor 87 (GPR87) is a recently discovered orphan GPCR which means that the search of their endogenous ligands has been a novel challenge. GPR87 has been shown to be overexpressed in squamous cell carcinomas (SCCs) or adenocarcinomas in lungs and bladder. The 3D structure of GPR87 was here modeled using two templates (2VT4 and 2ZIY) by a threading method. Functional assignment of GPR87 by SVM revealed that along with transporter activity, various novel functions were predicted. The 3D structure was further validated by comparison with structural features of the templates through Verify-3D, ProSA and ERRAT for determining correct stereochemical parameters. The resulting model was evaluated by Ramachandran plot and good 3D structure compatibility was evidenced by DOPE score. Molecular dynamics simulation and solvation of protein were studied through explicit spherical boundaries with a harmonic restraint membrane water system. A DRY-motif (Asp-Arg-Tyr sequence) was found at the end of transmembrane helix3, where GPCR binds and thus activation of signals is transduced. In a search for better inhibitors of GPR87, *in silico* modification of some substrate ligands was carried out to form polar interactions with Arg115 and Lys296. Thus, this study provides early insights into the structure of a major drug target for SCCs.

Keywords: GPCR - GPR87 - GPR95 - DRY motif - threading - molecular dynamics - squamous cell carcinomas

Asian Pac J Cancer Prev, 14 (12), 7473-7482

Introduction

G protein-coupled receptors (GPCRs) are still a large superfamily of membrane bound signaling proteins that hold great pharmaceutical interest as a promising drug target (Costanzi et al., 2012). GPCRs accounts that more than 2% of all genes encoded by the human genome. According to a recent analysis, 40-50% of all current drug targets, with 26 out of the 100 top selling drugs are exposed by targeting different GPCRs family (Hans et al., 2007). GPCRs is also the best target for inflammatory mediators, therefore it's provides a link between chronic inflammation and cancer disease. In addition, GPCRs intellectual have most important roles in tumor-induced angiogenesis and metastasis, which might involve in GPCR-guided migration of cancer cells to their target organs. Therefore, interfering with GPCRs and their downstream targets its potency provides an opportunity to develop a new technique for cancer diagnosis, prevention and treatment (Robert et al., 2007). GPCRs serve as a key signal transduction conduits by connecting extracellular inputs with diverse cellular responses (Audet et al., 2012). Although molecular biology and bioinformatics approaches builds the identification of orphan GPCRs

which are responsible for the search of their endogenous ligands which has been a significant challenge. The invention of endogenous ligands has given birth to the reverse pharmacology approach, which uses orphan GPCRs as a target to identify their ligands. This technique is magnificently used for the deorphanization of about 300 GPCRs over two decades (Chung et al., 2007). Due to their characteristic structure and specific ligand binding ability, many GPCRs have been extensively used as a target for therapeutic drug development and designing. Cancer is a deregulated multiplication of cells with the consequence of an abnormal increase of the cell number in particular organs or tissues. At a later stage, these cells are distributed via the hematopoietic and lymphatic systems throughout the body where they can colonize in distant tissues and cause metastasis.

A drug which has been launched in the market has consent through the seven basic steps are disease selection, target selection, lead compound identification, lead optimization, preclinical trials, clinical trial testing and pharmacogenomic optimization (Roderick, 2006). Orphan GPCR like GPR87 (also known as GPR95) is recently discovered orphan receptor and classified in the P2Y₁₂ subgroup that contains P2Y₁₂, P2Y₁₃, P2Y₁₄, CysLT₁, and

¹Department of Pharmacology and Therapeutics, King George's Medical University, (Erstwhile C.S.M. Medical University), Lucknow, ²Biomedical Informatics Centre, Rajendra Memorial Research Institute of Medical Sciences, Patna, India *For correspondence: drskhattri@gmail.com, mukta.rani1@gmail.com

CysLT₂ receptors (Nonaka et al., 2005). GPR87 is located on human chromosome 3q25 region (Wittenberger et al., 2001) in which the given P2Y receptor genes are clustered. GPR87 are found in neoplastic cells of different location with lung carcinomas (squamous), cervical squamous carcinomas, head and neck squamous cell carcinomas (larynx, pharynx, tonsil and tongue), urinary bladder carcinomas (Glatt et al., 2008; Gugger et al., 2008) and also in the placenta (Wittenberger et al., 2001). Human G-protein coupled receptor 87 (hGPR87) plays a crucial role in intracellular signal transduction (Wittenberger et al., 2001) and normal ovarian development as well as ovarian cancer progression. It was shown to be overexpressed in squamous cell carcinoma (SCCs) or adenocarcinoma in the lung and bladder carcinomas (Gugger et al., 2008). The two most promising reason for targeting this protein are first- GPR87 is the only example of the over expressed GPCRs that can be correlated on a mutation-based level to squamous cell carcinoma of the lung. Second- it shows the highest fold level of the over expressed GPCR. GPR87 is a novel target gene of p53 it's over expression in testicular cancer as a probable downstream effect of chemotherapy-induced p53 (Kerley-Haniilton et al., 2005). Orphan GPR87 plays crucial role in SCCs and it will certainly lead to the identification of novel cancer therapeutic targets. Recent research confirm that the GPR87 is an LPA receptor (Ochiai et al. 2013).

Till date no crystal structure information either from X-ray crystallography or NMR is available for hGPR87 protein are reported. In the present study, different homology models of GPR87 protein with 358 amino acids was constructed by using multiple templates through different homology modeling servers. According to the Mobarec et al. reported that, use of multiple templates for modeling of GPCRs increases its model quality, in terms of higher predictive power in flexible ligand-rigid protein docking experiments (Mobarec et al., 2009). Molecular dynamics (MD) simulations studies with explicit solvent treatments which have been most frequently employed to assess the stability of protein. *In silico* modification of some substrate ligands has been carried out to search for better inhibitors of GPR87 protein. In the absence of crystal structure, our study provides an early insight into the detailed 3Dstructure of a major drug target or developing new inhibitors for GPR87 which will clearly be a challenging work in the near future.

Materials and Methods

Sequence alignment

The amino acid sequences of GPR87 from *Homo sapiens* (GenBank ID: Q9BY21) was retrieved from National Centre for Biotechnology Information (NCBI). Multiple sequence alignment between P2Y receptors, group I: P2Y₁, P2Y₂, P2Y₄, P2Y₆ and P2Y₁₁; group II: P2Y₁₂, P2Y₁₃ and some other pharmacologically related receptors like CysLT₁ and CysLT₂ were retrieved from NCBI. Alignment was performed by using the ClustalW algorithm and cladogram was assessed by TreeView algorithm (Larkin et al., 2007) and maximum likelihood method from www.phylogeny.fr (Dereeper et al., 2008).

Translated protein sequence of GPR87 protein has 358 amino acid was taken for further study.

Molecular modeling

BLASTP (Altschul et al., 1990) search was performed against Protein Data Bank (PDB) to find suitable templates for homology modeling. There was no single template available in PDB that could be used to model GPR87, therefore two templates were selected. Based on the maximum identity with high score and lower e-value crystal structure of squid rhodopsin by intracellularly extended cytoplasmic region with 3.70Å resolution (PDB ID: 2ZIIY) (Shimamura et al., 2008) and turkey β₁-adrenergic G protein-coupled receptor with 2.70Å resolution (PDB ID: 2VT4) (Warne et al., 2008) were taken as a best templates have sequence similarity 21% from C to N-terminal ends. By using several homology servers like Geno3D (Combet et al., 2002), Swiss-Model (Arnold et al., 2006), Phyre (Kelley et al., 2009), I-Tasser (Zhang, 2008) and ESyPred3D (Lambert et al., 2002) servers. Dali server (Holm et al., 2008) has been accessed to compare and validate the template for GPR87 and find that the 2VT4 and 2ZIIY are best templates and their sequence identity (21%) for constructing the 3Dstructure of GPR87 by using threading method. The threading method uses specialized schemes to relate protein sequences with library of known structures. Even in cases where there is no clear sequence homology but they have been shown to be able to identify the likely protein fold (Shailza et al., 2011). The use of several templates (two) generally increases the model accuracy as it combines information from multiple template structures. The utility of multiple templates in comparative modeling relies on the accuracy of their multiple structure alignment (Pandian et al., 2011). The coordinates obtained from 3D modeled structures were visualized using the program PyMOL (DeLano et al., 2002). All the different modeled structures of GPR87 protein which were generated by different servers are validated through DOPE score (Discovery Studio2.5) which shows the lowest DOPE scores are the best modeled structures which use for further analysis (Table 1).

Model validation

The analysis of conformational correctness and reliability was carried out by using Ramachandran plot through PROCHECK (Laskowski et al., 1993) for validate the modeled structure of GPR87. The initial model that showed problems in conformation of the different loop region, which was further subjected to loop modeling and validated using side chain and loop refinement protocol of DS2.5 (Accelrys Software, 2003) and also by webserver like ERRAT plot (Colovos et al., 1993). ERRAT plot which gives a measurement of structural errors at each residue of modeled structure. This process was repeated iteratively until most of the amino acids were below 95% cutoff value and residue which lies above 95% cutoff value were subjected to loop modeling in MODELLER. Finally, GPR87 model showing the best PROCHECK and ERRAT plot was then subjected to native protein folding energy evaluation using ProSA server (Wiederstein et al., 2007).

Table 1. The Comparative Analysis of Various 3-D Models of GPR87 Protein Which were Built by Different Servers Like Geno3D, SwissModel, Phyre, I-Tasser and ESyPred3D

Model No.	Servers	No. of residues	Ramachandran plot quality (%)				DOPE Score	RMSD of the main chain with the template (Å)	Verify_3D	Errat Scores (Overall quality factor)
			Core	Allowed	Generously allowed	Disallowed				
2VT4	From PDB	274	92.90%	7.00%	0.10%	0.00%	-163838.2383	1.535Å	72.28%	97.994
2ZIY	From PDB	370	70.40%	27.10%	2.10%	0.30%	-44993.51563	1.169Å	68.73%	81.977
A	Geno3D	32	74.10%	25.90%	0.00%	0.00%	-2041.300659	2.949Å	0.00%	100
B	SWISS MODEL (2ZIY_A)	293	74.70%	21.20%	3.30%	0.70%	-35901.44531	1.795Å	35.71%	68.772
C	SWISS MODEL (2VT4)	286	92.10%	6.80%	0.40%	0.80%	-36460.64844	1.994Å	41.46%	71.223
D	PHYRE	347	79.10%	15.60%	4.40%	0.90%	-39222.22265	2.011Å	32.18%	53.412
E	I-Tasser	358	77.90%	15.70%	3.90%	2.40%	-46755.23828	1.049Å	53.20%	89.714
F	ESyPred3D	359	86.00%	10.20%	2.50%	1.30%	-39711.86719	2.161Å	39.12%	43.807

*Predicted number of TMs; **Numbers in bracket are the lengths of TMs

Transmembrane helices prediction and function assignment of GPR87 protein

Different servers like DAS, TMHMM, HMMTOP, TMpred, TopPred, SOSUI, SPLIT, and Predictprotein (PHD) servers were accessed to predict and validate the transmembrane helical region (TM region) of GPR87 protein (Hofmann et al., 1993; Claros et al., 1994; Cserzo et al., 1994; Hirokawa et al., 1998; Tusnády et al., 1998; Krogh et al., 2001; Juretic et al., 2002; Jones et al., 2004; Rost et al., 2004). To know the novel functions of GPR87 protein by using a SVMProt server with the aim of support vector machine learning techniques (SVM) which classifies protein into functional families from its primary sequences i.e. amino acids (Cai et al., 2003).

Molecular dynamic simulations

Molecular dynamic (MD) simulations were carried out using the CHARMM (Brooks et al., 1993) force field by module of standard dynamic cascade protocol in DSv2.5. The binding of GPR87 model in presence of explicit solvent was analyzed by molecular dynamics simulation. The protein was solvated using explicit spherical boundary with harmonic restraint solvation protocol of DS2.5 was used. The protein atoms were parameterized using the CHARMM force field (Brooks et al., 1993). All bond lengths involving hydrogen atoms were fixed using a SHAKE algorithm (Ryckaert et al., 1977). Simulations were carried out at 300K with 7000 steps (7ns) in steepest descent and adopted basis Newton-Raphson (ABNR) minimization technique, until the root-mean square deviation (RMSD) was less than 0.001 kcal/Mol/Å R.M.S gradient. The whole system was then equilibrated for 500ps at the target temperature of 300K using 'Berendsen Coupling Bath' and finally the production run was carried out for another 500ps for the canonical ensemble (NVT) using Leapfrog Verlet dynamics Integrator. In all the three stages (Heating, Equilibration and Production) of simulation, time step of 1fs was used and only the backbone of GPR87 protein was kept rigid.

Normal mode analyses on GPR87 protein 3-D structure

Normal mode analysis (NMA) was conducted to analyze the intrinsic motions of the GPR87 model. The eInémo online server (<http://igs-server.cnrs-mrs.fr/elne/mo/index.html>) (Neudecker et al., 1988) is a part of the Elastic Network Model, which provides fast simple tool to compute, visualize and analyze low-

frequency normal modes of biological macromolecules. The structural model of *hGPR87* in dot pdb (*.pdb) format was submitted to NMA analysis; while another 3Dstructure complex was submitted as reference of the desired conformational change. The key parameters which are used in computation that includes: DQMIN=-100, DQMAX=100, DQSTEP=20 and NRBL="auto". A total of 100 normal modes with the lowest frequencies were requested. The normal mode theory is based on the harmonic approximation of the potential energy function, around a minimum energy conformation. This approximation allows an analytical solution of the motion by diagonalizing Hessian matrix. Hessian matrices are used in large-scale optimization problems within Newton-type methods because they are coefficient of the quadratic term of a local Taylor expansion function i.e.

$$y=f(x+\Delta x)\approx f(x)+J(x)\Delta x+\frac{1}{2}\Delta x^T H(x)\Delta x$$

Where J is the Jacobian matrix, which is a vector (gradient) for scalar-valued functions. The full Hessian matrix can be difficult to compute in practice; in such situations, quasi-Newton algorithms have been developed that use approximations to the Hessian. The most well-known quasi-Newton algorithm is the BFGS algorithm (Binnmore et al., 2001).

Active site prediction

Pocket-Finder or Q-site finder is a molecule-binding site prediction server based on a Ligsite algorithm (Hendlich et al., 1997). It works by scanning a probe radius 1.6Å along all gridlines of grid resolution 0.9Å surrounding the protein. The probe also scans cubic diagonals. Grid points are defined to be part of a site when the probe is within range of protein atoms followed by free space followed by protein atoms. Grid points are only retained if they are defined to be part of a site at least five times (Hendlich et al., 1997). The different active sites in *hGPR87* protein also predicted through CASTp server (Dundas et al., 2006) and compare their results with the pocket finder result. The different binding sites and functional residues were identified and stored for further ligand protein interaction analysis.

Docking with substrate ligands

Protein-ligand interaction i.e. docking study was done by LigandFit protocol of DS2.5 (Venkatachalam et al., 2003). The substrate ligands were used i.e. 4-[(2S)-3-(tert-butylamino)-2-hydroxypropyl] Oxy}-3H-indole-

2-carbonitrile (P32), 2-Hydroxymethyl-6-octylsulfanyl-Tetrahydro-Pyran -3,4,5 triol (SOG), palmitic acids (PLM) and retinal (RET) were obtained from the template crystal structure (PDB ID: 2VT4, 2ZIY) (Shimamura et al., 2008; Warne et al., 2008) to dock with GPR87 refined model. The LigandFit docking algorithm combines a shape comparison filter with a Monte Carlo conformational search to generate docked poses consistent with the binding site shape. Among these poses, the most suitable docking mode with a high dock score and a high fitness score from consensus functions was finally selected for analysis. Interaction study with some other protein i.e. EMR2 in *H sapiens* with chondroitin 4-sulphate is reported (Rani et al., 2011), TLR4 in *H sapiens* (Das et al., 2012) and GAPDH protein in *Leishmania* were already been reported (Sahoo et al., 2013). GPCR-membrane models would provide a significant tools for studying receptor behavior and its modulation by small drug-like ligands, a relevant issue for drug discovery (Sadiq et al., 2013).

Results

Sequence alignment and phylogenetic analysis

The translated amino acid sequences of G-protein coupled receptor 87 from *Homo sapiens* consists of 358 amino acids (Q9BY21) taken from NCBI. Multiple sequence alignment between two groups of P2Y receptors, group I: P2Y₁, P2Y₂, P2Y₄, P2Y₆ and P2Y₁₁ and group II: P2Y₁₂, P2Y₁₃ and some other pharmacologically related receptors like CysLT₁ and CysLT₂ were obtained from NCBI. Multiple sequence alignment was performed to calculate a distance matrix by using ClustalW and cladogram was constructed by using TreeView algorithm is shown in Figure 1A. By maximum likelihood method clustering of GPR87 of *Homo sapiens* with *Mus musculus* showed 89% sequence similarity and with P2Y receptors, CysLT₁ and CysLT₂ receptors producing similarity in between 17-38% Figure 1A. In sequence alignment there are total eight conserved residues were examined in two groups of P2Y receptors and CysLT receptors are C114, R139, P146, W166, Y226, P267, H269 and P311 predicted by ClustalW server.

For the finding of the best template by using BLASTP, three significant hits i.e. 2VT4/A, 2ZIY/A and 3KJ6/A are found, which shows 16-21% sequence identity with GPR87 are shown in Figure 1B. It shows that total 41 amino acids are highly conserved with these three templates sequences. In GPR87 protein, it comprises two different domains that consist of 1-226 residues in the first domain, similarly 227-358 in second domain. The disulfide bond lies in the first domain of GPR87 predicted by Dipro/SCRATCH protein predictor program which also shown in Figure 1B. In disulfide bonds there are 78 amino acids involved in between 114-192 amino acids in length.

Model building and validation

The 3Dstructure of orphan GPR87 protein was built with threading method on the basis of two different templates i.e. PDBID 2VT4 and PDBID 2ZIY by using different homology modeling serves. The comparative

analysis of various 3Dstructures which were built by using Geno3D, Swiss-Model, Phyre, I-Tasser and ESyPred3D are summarized in Table 1. The model generated by I-TASSER (model E) is the only model which lays the entire 358 amino acids. For model validation, Ramachandran plot (the Φ/Ψ distribution of the backbone conformation angles for each residue of the 3Dstructure) statistics of model 'C' revealed that the highest 92.1% residues in the core region or 6.8% in allowing region, 0.4% lies in generously allowed region and 0.8% residue in disallowed region is approximate equal to the validation of template structure i.e. 2VT4, but the model structure lay only 286 residues instead of 358. Based on the scores of the Ramachandran plot (validation), the exact length of residues and RMSD value of 3Dstructure in comparison with template main chain model 'E' is the best model are plotted in Figure 2. In model 'E', 2.4% residues lie in disallowed regions which includes Phe3, Leu29, Asn33, Thr74, Lys110, Asp149, Gln183, Asp187, Trp203, Ile241, Glu296 and Lys290 were refined through loop refinement and side chain refinement protocol of DSv2.5 and their RMSD of the main chain with their template is 1.049Å.



Figure 1. A) Phylogenetic tree of hGPR87 of *H sapiens* with P2Y receptors. Group I: P2Y₁, P2Y₂, P2Y₄, P2Y₆ and P2Y₁₁; and group II: P2Y₁₂, P2Y₁₃ and Some Other Pharmacologically Related Receptors Like CysLT₁ and CysLT₂ were Obtained from <http://www.phylogeny.fr/> with Maximum Likelihood Method. The tree is formed with 0.99 maximum sequence difference values; B) Multiple Sequence Alignment of 3 Different Templates their PDB IDs: 3KJ6/A 2R4R/A and 2VT4/A Taken from BLASTP Which Shows 16-21% Sequence Similarity with GPR87 Protein and Find 41 Conserved Residues. The disulphide bridges are comprised in 114-192 residues represented in blue color

All the models were validated by DOPE scores (Discrete Optimized Protein Energy) of DSv2.5 and model 'E' has the highest DOPE score is -46755.238281 which also proof that it is the best model structure of *hGPR87*. The DOPE score is the scoring function of protein which can be viewed as a conformational energy and measures the relative stability of a conformation with respect to other conformations in the same protein (Shen et al., 2006).

The secondary structure of GPR87 proteins (α -helices and β -sheets) and their comparisons with two templates (2VT4|A, 2ZIY|A) which contains 7 transmembrane helices (7TM) and 8 small α -helices, 7TM and 10 small α -helices in 2VT4 and 7TM and 12 small α -helices and 2 β -sheets have been observed are represented in Figure 3A. The superimposition of the *hGPR87* model (orange) over with its highest closest homolog (2VT4) (cyan) resulted in a root mean square deviation (RMSD) of 1.286Å represented in Figure 3B. Therefore, we influenced that

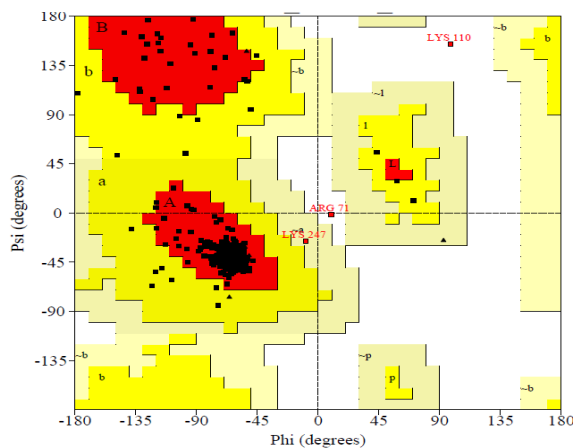


Figure 2. Ramachandran Plot of Modeled 3Dstructure (model E) of GPR87 Protein in *H sapiens* Calculated by the PROCHECK Program

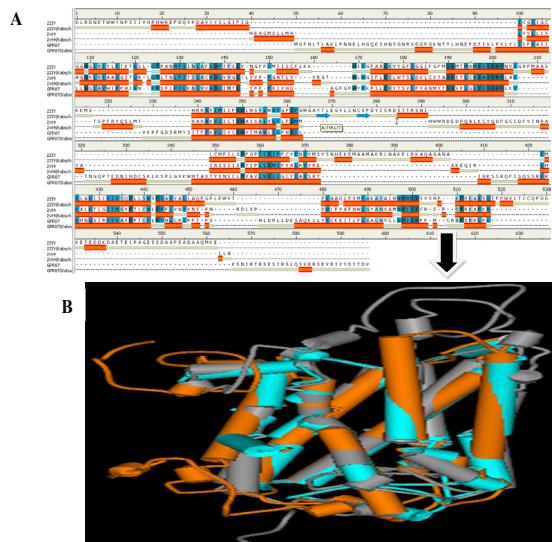


Figure 3. A) Sequence Alignment of GPR87 Protein with Two Templates i.e. 2ZIY and 2VT4. Comparison of their secondary structure is also shown, that represents orange board lines is helical, blue arrow shows β -sheets and rests are loop regions; and B) Superimposed of GPR87 Model (orange) with Two Templates are 2VT4 (cyan) and 2ZIY (grey) Crystal Structures

we had reasonable and reliable confirmation of modeled 3Dstructure of GPR87 protein for further analysis.

Different models of GPR87 were further validated by comparison with structural features with their template by using Verify-3D; ProSA and ERRAT servers were used for determining the stereo-chemical parameters of the model structure. ProSA is widely used to check 3D models of protein structures for potential errors. ProSA energy profile of modeled GPR87 in comparison with 2VT4 is presented in Table 1. Verify-3D shows that 53.20% of the residues have an averaged 3D - 1D score greater than 0.09 and ERRAT plot gives an overall quality factor of 89.714 for the modeled structure of GPR87 proteins (Table 1). The quality assessment of the model via ProSA revealed that models 'B' and 'E' built by Swissmodel and I-Tasser are matched with the NMR region of the plot with Z-score -1.96 and -1.97 respectively which indicates good quality of the model. Furthermore, we considered I-Tasser modeled structure for further study which consists of 7 transmembrane helices (7TM) and 8 small α -helices prepared by using DS2.5 is shown in Figure 4. In GPR87 there are sequence conflicts in amino acids Ser154 which usually of unknown origin and they may originate from sequencing errors or may be yet uncharacterized natural polymorphism (SNP) reported in UniProtKB (Ota et al., 2004). In Figure 5, schematic representation of 3Dstructure of GPR87 protein is shown using the one-

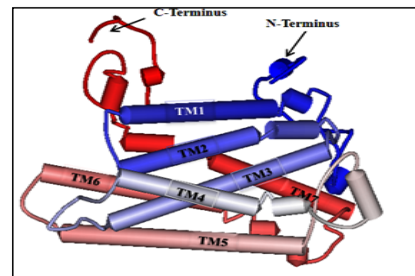


Figure 4. Ribbon Representation of the Modeled Structure of GPR87 Protein Consists of 7 Transmembrane Helices and 8 small α -helices. The rest of the structure is depicted as loops and N-terminal shows blue color and forwarding C-terminal that shows in red color. The figure was prepared using DSv2.5

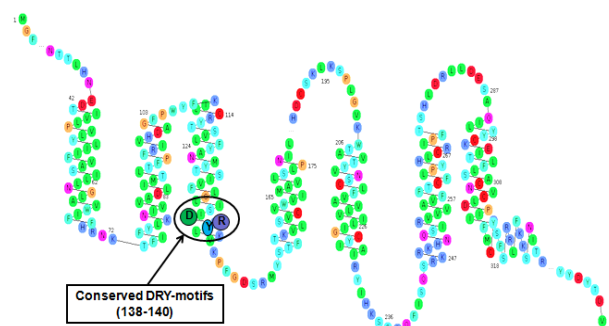


Figure 5. Schematic Representation of the Amino Acid Sequence of GPR87 used in the Computational Model. Note that the C- and N-terminus ends are shortened and also the loop between helix 2 and 3 has been shortened. The black letter amino acids indicates conserved Asp-Arg-Tyr (DRY) motif is located at the beginning of the second intracellular loop and at the end of helix 3 where the G-protein binds and thus activation signals are transduced

letter residue abbreviation in snake plot. Furthermore, the DRY-motif (an Asp- Arg-Tyr sequence) at the end of helix 3 is highlighted (138-140aa) it's a significant position where the G-protein binds and thus activation signals are transduced. The DRY-motif of GPR87 is located at the beginning of the second intracellular loop which is useful to find novel lead compounds which can be designed on the basis of ligand protein interaction (docking). The DRY-motif patterns are also present in all three different templates i.e. 2VT4|A, 2ZIY|A and 3KJ6|A are shown in the sequence alignment (Figure 2). Similarly, the same hypothesis which have already reported that the C-terminus tail which provide the main driving force for stable interaction between the receptor and β -arrestin and the DRY-motif in these receptors appears to play important roles in β -arrestin binding as well as in maintaining the receptor conformations (Kyeong et al., 2008).

Transmembrane helices (TMs) prediction

There are twelve different transmembranes prediction servers i.e. DAS, TMHMM, HMMTOP, TMpred, TopPred, SOSUI, SPLIT, Predictprotein (PHD), MEMSAT, Wave TM, HMM-TM and SMART servers are used to predict the position and number of transmembrane regions in *hGPR87* protein are summarized in Figure 4 and 5. The transmembrane domain of orphan GPCR87 is composed of 6-7 transmembrane helices (TMs) are listed in Table 2. These seven TMs were then modeled by using the known structure of β 1-adrenergic GPCR (PDBID 2VT4) as a template. The extracellular loops on the transmembrane domain of GPR87 are relatively long and therefore its 3Dstructure is modeled on the basis of appropriate templates. The comparative analysis of transmembrane helices prediction programs showed that the lowest range and higher range of transmembrane helices in first TM is 42-71 residues, 74-99 in second TM, 111-149 in third TM, 155-180 in fourth TM, 206-232 in fifth TM, 252-278 in sixth TM and 292-322 in seventh TM. In case of TopPred and SOSUI servers both predict only six transmembrane helices at position second and sixth there is missing one transemebrane residues of GPR87 protein respectively.

Protein function assignment of GPR87 protein

Drug design can be applied to arrest the post-translational modification of GPR87 which will inactivate functional characteristics of that particular protein. It shows six different N-glycosylation sites revealed by

using MOTIF SCAN but on PROSCAN server only one site at 129-132 (NITV) is present are shown in Table 3. In GPR87 protein of *H sapiens*, one N-myristoylation site is present at 224-229 amino acids (GCYIAI). The GPCR87 protein belongs to Zn metallo- β -lactamase family having an E-value of 0.00015. It is a member of rhodopsin-like GPCR superfamily which varies in between 44-322 residues and rhodopsin-like GPCR superfamily (7tm_1) which lays involvement of 59-314 amino acids (Table 3). The closest structural domain lies in lactamase_B domain i.e. metallo- β -lactamase superfamily which shows 10-158 residues by MOTIFSCAN and 5-148 through the Intrepro tool of SWISSMODEL. The family metallo- β -lactamases which degrade all β -lactam drugs including carbapenems and requires zinc as a cofactor (Varsha, 2008). These enzymes are not inhibited by routine β -lactamase inhibitors, so these compounds are being evaluated to inhibit these types of enzymes.

Table 3. Different Post-translational Modification Sites in GPR87 Protein of *H sapiens**

Results	Motif scan Pattern	Proscan result Pattern	Swiss-model result Pattern
ASN_GLYCOSYLATION (N-glycosylation site)	4→7(NLTL) 25→28(NRSD) 33→36(NTTL) 72→75(NKTS) 129→132(NITV) 251→254(NQSI)	129→132 (NITV)	NO
Rhodopsin-like GPCR superfamily			44→68 76→97 121→143 157→178 207→230 250→274 296→322
Rhodopsin-like GPCR superfamily (7tm_1)	59→314		59→314
(G_PROTEIN_RECEP_F1_2)	59→314		59→314
(G_PROTEIN_RECEP_F1_1)	127→143		127→143
Purinoreceptor, Family (P2 purinoreceptor Family)			90→101 114→131 138→148 225→236 264→275
Purinoreceptor, Family (P2Y ₁₃ purinoreceptor Family)			133→144 181→196
Unintegrated, (Family A G protein-coupled receptor-like)			19→350
MYRISTYL (N-myristoylation Site)	224→229 (GCYIAI)		224→229 (GCYIAI)
Lactamase_B (Metallo-beta-lactamase superfamily)	10→158		5→148

*The motifs were predicted by using three different servers like Motif Scan, Proscan and Swiss Model server

Table 2. The Predicted Number and Locations of Transmembrane Helices (TMs) of GPR87 Protein in Homo Sapiens

Servers	N*	TM1	TM2	TM3	TM4	TM5	TM6	TM7
DAS	7	42→68 ^{(27)**}	78→95 ⁽¹⁸⁾	111→137 ⁽²⁷⁾	157→177 ⁽²¹⁾	210→229 ⁽²⁰⁾	255→269 ⁽¹⁵⁾	300→317 ⁽¹⁸⁾
TMHMM	7	47→69 ⁽²³⁾	76→98 ⁽²³⁾	118→137 ⁽²⁰⁾	157→179 ⁽²³⁾	207→229 ⁽²³⁾	256→278 ⁽²³⁾	298→317 ⁽²⁰⁾
HMMTOP	7	43→67 ⁽²⁵⁾	74→92 ⁽¹⁹⁾	119→137 ⁽¹⁹⁾	160→179 ⁽²⁰⁾	206→230 ⁽²⁵⁾	253→271 ⁽¹⁹⁾	298→317 ⁽²⁰⁾
TMpred	7	47→68 ⁽²²⁾	76→99 ⁽²⁴⁾	119→137 ⁽¹⁹⁾	160→178 ⁽¹⁹⁾	206→229 ⁽²⁴⁾	253→271 ⁽¹⁹⁾	298→317 ⁽²⁰⁾
TopPred	6	42→62 ⁽²¹⁾		117→137 ⁽²¹⁾	160→180 ⁽²¹⁾	210→230 ⁽²¹⁾	252→272 ⁽²¹⁾	298→318 ⁽²¹⁾
SOSUI	6	43→65 ⁽²³⁾	76→98 ⁽²³⁾	127→149 ⁽²³⁾	157→179 ⁽²³⁾	210→232 ⁽²³⁾		296→318 ⁽²³⁾
SPLIT	7	42→69 ⁽²⁸⁾	75→98 ⁽²⁴⁾	117→138 ⁽²²⁾	156→178 ⁽²³⁾	208→230 ⁽²³⁾	255→270 ⁽¹⁶⁾	297→319 ⁽²³⁾
PHD	7	46→69 ⁽²⁴⁾	74→95 ⁽²²⁾	116→136 ⁽²¹⁾	158→175 ⁽¹⁸⁾	209→229 ⁽²¹⁾	254→271 ⁽¹⁸⁾	300→318 ⁽¹⁹⁾
MEMSAT	7	44→68 ⁽²⁵⁾	76→97 ⁽²²⁾	121→143 ⁽²³⁾	157→178 ⁽²²⁾	207→230 ⁽²⁴⁾	250→274 ⁽²⁵⁾	296→322 ⁽²⁷⁾

*Predicted number of TMs; **Numbers in bracket are the lengths of TMs

Different unknown and hidden functions of GPR87 were predicted by using machine learning technique like statistical support vector machine based classifier i.e. SVMProt. The GPR87 protein has been assigned to transmembrane (99.2%), 7 transmembrane receptor [rhodopsin family and chemoreceptor (99.1%)], zinc-binding (94.7%) functional family by SVMProt. Analysis of glycoprotein M (matrix protein) by SVMProt shows that it may be a transmembrane region protein that participates in Type II (general) secretory pathway (IISP) (58.6%) and various activities.

Molecular dynamic simulations

Based on intrinsic dynamic, structural stability and improved relaxation of the modeled GPR87 protein, energy minimization by steepest descent methods of standard dynamic cascade protocol of DSv2.5, which is an important step for the convergence of free MD simulations, were calculated. The effect of solvent on the stability of binding of the GPR87 was studied by MD simulation of a fully hydrated model is shown in Figure 6. The radius of the sphere is 40Å and salt concentration is 0.145 is fixed. The total energy and temperature were found to remain steady with little fluctuation during the production stage of 500ps. During the production stage of 500ps following the initial heating and equilibration phases, total energy and simulation temperature were found to remain steady with slight fluctuation. The snapshots of steepest descent dynamics trajectory at 0 to 8000 of the production runs are shown in Figure 7. The result indicates that the variation in total energy is -15782.28127 KJ/Mol to -19888.33249 KJ/Mol calculated by the steepest descent method at 6500ps dynamics trajectory was retrieved and further minimized to full convergence. Van der Waals energy did not show much variation both by steepest descent method (-2344.01480 KJ/Mol to -2312.75325 KJ/Mol). The

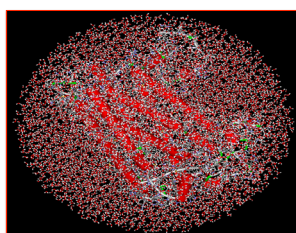


Figure 6. Overview of GPR87 Receptor Solvated in the Hydrated Lipids Bilayer. The backbone of the receptor is represented in red helix, water is in red/white and the internal water molecules are displayed as spheres

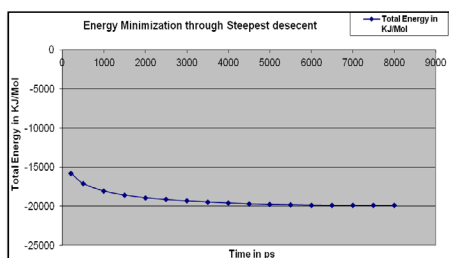


Figure 7. Comparisons and calculate energy versus time plot by using Steepest descent of energy minimization protocol of DSv2.5 software. The X-axis is Time (ps) and Y-axis Total energy in KJ/Mol

electrostatic energy and RMS gradient of GPR87 protein is between -16107.48364 KJ/Mol to -20528.43254KJ/Mol and RMS gradient varies between 3.31585 KJ/Mol to 0.09989 KJ/Mol respectively.

Normal mode analyses on GPR87 3Dstructure

Normal Mode Analyses (NMA) program was conducted to analyze the intrinsic motions (vibrations and thermal properties) of the GPR87 protein modeled structure. The structural model of GPR87 in apo-form was submitted for NMA analysis; while the GPR87/2VT4 model complex was submitted as a reference of the desired conformational change. Essential features of the top ten low-frequency normal modes, including its frequency, amplitude, collectivity of atom movements and the overlap with observed conformational changes are illustrated in Table 4. The two lowest-frequency normal modes (mode 7 and 8 in Table 4) are illustrated in Figure 8 one feature showing bending of the N-terminal domain towards the transmembrane domain, the other feature showing rotation of the N-terminal domain on the top of the transmembrane domain. In order to confirm the outcomes of NMA analysis, we also conducted the principal component analysis (PCA) on GPR87 in apo-form to detect its major conformational motions during MD simulation. Each snapshot of the GPR87 structure was fitted with a reference configuration to remove translational and rotational motions.

Analysis of active sites

The potential ligand binding sites (LBSs) of orphan GPCR87 protein of human have been found by CASTp server. Total sixty three different possible binding sites were retrieved and top ten different sites are represented in Figure 9. The frequently involved amino acid residues involved in forming the active sites are Lys36, Asn37, Leu45, Cys49, Tyr58, Ala64, Ile55, Leu62, His693, Lys81, Phe94, Gly105, Cys114, Phe121, Gly148, Thr158, Ser173, Thr183, Gly200, Val201, Lys204, Val210, Arg245, Phe261, Leu281, Tyr292, Lys295, Phe301, Phe325, Arg329, Lys337, Arg345, Tyr352 and Val358. These are the key residues of GPR87 protein which is responsible for protein-ligand interaction studies of drug designing

Table 4. The Different Normal Mode Analysis of GPR87 Protein Returned by the eINémo Server

Mode ^a	Frequency	Collectivity ^b	Cumulative Overlap ^c	Amplitude (dq)
7	1	0.2784	0.038	-201.7625
8	1.27	0.2067	0.058	-143.9066
9	1.5	0.3139	0.103	-225.1856
10	1.73	0.376	0.195	-316.2696
11	1.81	0.1335	0.196	33.1312
12	2.21	0.4038	0.226	181.4111
13	2.25	0.2636	0.235	-97.38002
14	2.53	0.3778	0.244	-98.6877
15	2.64	0.4753	0.257	-117.6877
16	3.05	0.3058	0.298	213.9873

^aOnly the 10 normal modes with lowest frequencies are displayed here; ^bThe level of collectivity indicates the percentage of residues that are involved in a certain normal mode; ^cThe level of overlap measures the similarity between a desired conformational change and that of a certain normal mode

approaches. All these structural, functional and proteomic analysis about human GPR87 protein of will lead to identification of novel lead compounds for squamous cell carcinoma.

Protein-ligand interaction study

After knowing the active sites in GPR87 protein, we analyze the specific substrate ligands of the template were docked with both the 3D model and crystal structures optimally to estimate the site of reaction on the ligand molecule and their binding affinity with the target protein. There are sixteen different binding sites were detected in GPR87 protein and also in 2VT4 but 2ZIY have fourteen sites detected by using find sites from receptor cavities protocol of DSv2.5.

The highest ligandfitness score has been detected as 61.414 with P32 and low as 56.952 with palmitic acids (PLM) having binding affinity for GPR87 protein are listed in Table 5. The ten different conformations (poses) have been observed during docking study in DS (Accelrys). The groove contains several positively charged side chains that interact directly with the negatively charged groups on P32, SOS, PLM and retinal are shown in Figure 10 A-D. The GPR87 belongs to the members of the opsin family, which have been considered to be typical members of the

rhodopsin superfamily which share several motifs mainly 7TMs, GPCRs of the rhodopsin superfamily. Mostly the opsin that binds with chromophore, such as 11-cis-retinal but in case of retinal the binding affinity is 60.675 and there is an absence of H-bonding. The most significant function of opsins other than the photoisomerases is split into two steps: light absorption and G-protein activation. The interactions between all the four substrate molecules and GPR87 protein, the frequently involve amino acids in the formation of H-bond are Arg115, Ser118, Tyr122, Pro198, Gly200, Val201, Lys202, Tyr268 and Lys296 and their atom involves OG, HH, O, HN, HN, O, HZ1, HZ3, OH and HZ3 revealed ion-pair interactions with substrate ligands atoms are O18, O20, O2, O1, N5, N17 and O16 respectively. For the same the crystal structure interaction shows the 69.723 dock score with P32 with PDBID 2ZIY and 45.726 with PDBID: 2VT4 and interacts frequently with Lys305, Pro312, Trp332, Val333 and Arg284, Arg345 and Arg350. Our docking results suggest that, GPR87 can bind close to the active site with similar binding energy with template crystal structures so, we concluded that these substrate ligands and their various analogues can be the best suitable lead compounds for the development of a novel class of selective drugs for cancer in order to

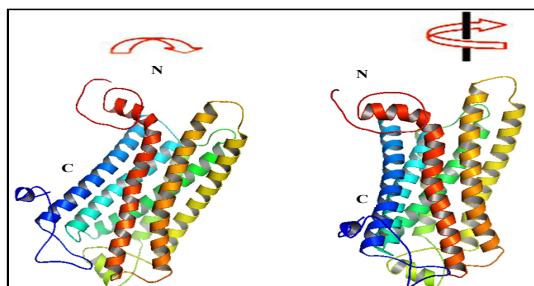


Figure 8. The Two Different Normal Modes (i.e., modes 7 and 8) of GPR87 Protein with the Lowest Frequencies. Left: bending of the N-terminal domain towards the transmembrane domain. Right: rotation of the N-terminal domain on the top of the transmembrane domain

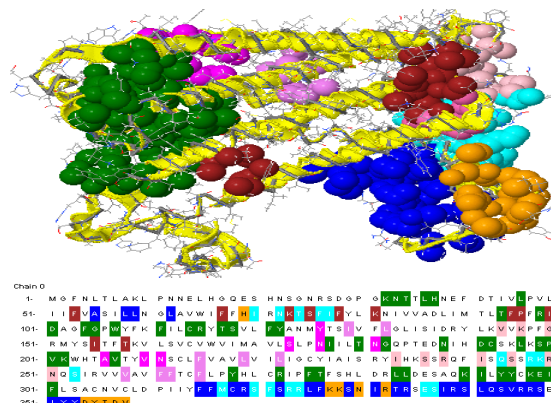


Figure 9. Prediction of Ligand Binding Sites (LBSs) of GPR87 Protein by CASTp Server

Table 5. Binding Free Energies and Total Number of H-bonds (in all 10 poses) Calculated by Using Ligandfit Protocol of DS2.5 for Four Main Substrate Ligands from Crystal Structure PDB ID: 2VT4 and 2ZIY Templates used for Construct the 3-D Structure of hGPR87 Protein

Inhibitors	Dock Score/Total No. of H-bond in all 10 Pose	3-D structure	GPR87 Residues	GPR87 Atom	Inhibitors Atom	Distance (Å)	Angle (Degree)	<i>in situ</i> Final Energy
P32	61.414/11		Tyr122 Tyr122	OH	O1	3.05	11.84	19.471
4-[[2S)-3-(tert-butylamino)-2-hydroxypropyl]oxy}-3H-indole-2-carbonitrile			Lys296	HH	O1	2.12	152.87	
			Val201	HZ3	N17	1.79	116.14	
			Tyr122	O	N5	3.19	150.87	
SOG	58.289/7		Tyr122	HH	O20	2.41	131.51	26.424
			Ser118	OG	O18	3.2	110.91	
			Tyr122	HH	O20	2.45	131.2	
			Tyr268	OH	O16	3	77.52	
			Pro198	O	O20	3.2	80.37	
			Gly200	HN	O20	2.48	115.41	
PLM	56.952/14		Tyr268	OH	O16	2.68	71.31	20.25
			Lys202	HZ1	O2	2.44	102.8	
			Lys202	HZ3	O2	2.42	80.79	
RET	60.675/0		Lys202	HZ1	O2	2.32	101.2	7.271
			No H-bond is formed					

*The motifs were predicted by using three different servers like Motif Scan, Proscan and Swiss Model server

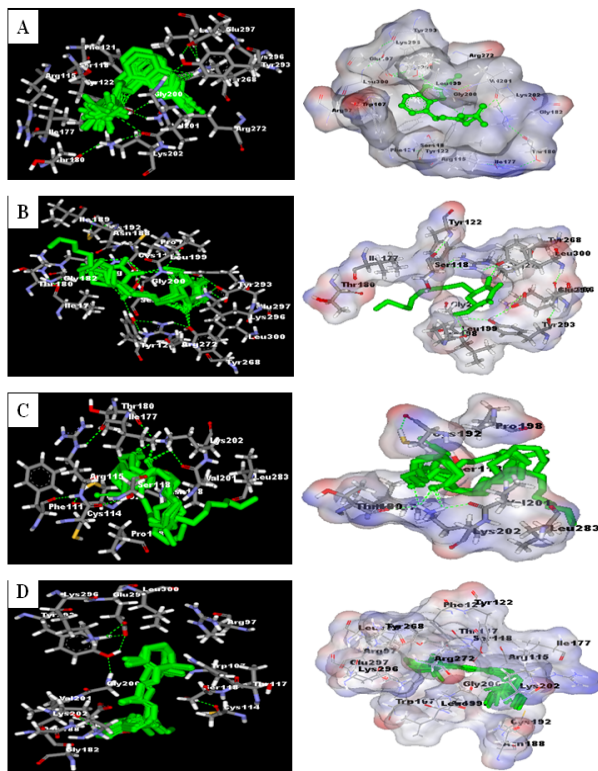


Figure 10. Interaction of GPR87 Protein with Substrate Ligands of Template Crystal Structure (PDB ID: 2VT4 and 2Z1Y). In the above figure, A) represents P32; B) is SOG; C) is palmitic acids; and D) is retinal. It shows ten different conformations (poses) have been observed during docking study. The green color is the substrate ligands, protein is in atomic color, green dotted lines represented as H-bonding and GRASP model of GPR87 protein shown in electrostatic potential energy red is positive and blue is negative and white are neutral. All figures generated by using DSv2.5

control squamous cell carcinoma.

Discussion

In conclusion, this is the first comprehensive study highlighting the structural features of a potential drug target i.e. orphan GPR87 protein. The structural aspects of the active site and the different predicted motifs help to understand the functional details of the protein. On the basis of modeled structure, predicted binding sites and functions, high throughput screening of various compounds may be carried out to find appropriate leads against squamous cell carcinoma, which can aid in the design and development of novel drug molecules against this drug target GPR87. In the absence of crystal structure, our study provides an early insight into the structure of the major drug target GPR87 and facilitating the inhibitor design. As the 3-D structure of GPR87 protein is known from this study, novel lead compounds can be designed on the basis of ligand protein interaction (docking) of available anti-cancer drugs and designing various analogues of the presently available drugs or defining a novel candidate on the basis of interaction with different binding sites of GPR87 protein in *H sapiens*. It would be helpful in the future if computational methods could provide guidelines as to what chemical manipulations

of ligands will be useful to control protein function. In conclusion, substrate-binding prediction by docking studies and molecular dynamics may contribute to the identification of metabolic routes to drug development and thus aid in the designing of new anti-cancer drugs.

Acknowledgements

An author expresses their gratitude to Head of Department of Pharmacology and Therapeutics, C. S. M. Medical University, Lucknow, India and BIC Centre, RMRIMS, Patna, ICMR to provide all facilities required for this work. Mukta Rani also thanks to Council for Scientific and Industrial Research (CSIR), Government of India for her present fellowship under its CSIR-SRF scheme, ACK.No. 181098/2K12/1.

References

- Altschul SF, Gish W, Miller W, et al (1990). Basic local alignment search tool. *J Mol Biol*, **215**, 403-10.
- Arnold K, Bordoli L, Kopp J, et al (2006). The swiss-model workspace: a web-based environment for protein structure homology modeling. *Bioinformatics*, **22**, 195-201.
- Audet M, Bouvier M (2012). Restructuring G-protein-coupled receptor activation. *Cell*, **151**, 14-23.
- Bokoch MP, Zou Y, Rasmussen SG, et al (2010). Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature*, **463**, 108-12.
- Brooks BR, Brucoleri RE, Olafson BD, et al (1993). CHARMM: a program for macromolecular energy minimization and dynamics calculations. *J Comput Chem*, **4**, 187-217.
- Cai CZ, Han LY, Ji ZL, et al (2003). SVM-Prot: web-based support vector machine software for functional classification of a protein from its primary sequence. *Nucleic Acids Res*, **31**, 3692-7.
- Chung S, Funakoshi T, Civelli O (2008). Orphan GPCR research. *Brit J Pharmacol*, **153**, 339-46.
- Claros MG, von Heijne G (1994). TopPred II: An improved software for membrane protein structure predictions. *Comput Appl Biosci*, **10**, 685-6.
- Colovos C, Yeates TO (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*, **2**, 1511-9.
- Combet C, Jambon M, Deléage G, et al (2002). Geno3D: An automated protein modeling Web server. *Bioinformatics*, **18**, 213-4.
- Costanzi S (2012). Homology modeling of class a G protein-coupled receptors. *Methods Mol Biol*, **857**, 259-79.
- Cserzo M, Wallin E, Simon I, et al (1994). Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: Application of the Dense Alignment Surface (DAS) method. *J Mol Biol*, **243**, 388-96.
- Das S, Rani M, Pandey K, et al (2012). Combination of paromomycin and miltefosine promotes TLR4-dependent induction of antileishmanial immune response *in vitro*. *J Antimicrob Chemother*, **67**, 2373-8.
- Dereeper A, Guignon V, Blanc G, et al (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res*, **36**, 465-9.
- Dundas J, Ouyang Z, Tseng J, et al (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acid Res*, **34**, 116-8.
- Glatt S, Halbauer D, Heindl S (2008). *hGPR87* contributes to viability of human tumor cells. *Int J Cancer*, **122**, 2008-16.

- Gugger M, White R, Song S (2008). GPR87 is an over expressed G-protein coupled receptor in squamous cell carcinoma of the lung. *Dis Markers*, **24**, 41-50.
- Hans BO, Petrine W, Anders AJ (2007). Structure, pharmacology and therapeutic prospects of family C G-protein coupled receptors. *Current Drug Targets*, **8**, 169-84.
- Hendlich M, Rippmann F, Barnickel G (1997). LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *J Mol Graph Model*, **15**, 359-63.
- Hirokawa T, Boon-Chieng S, Mitaku S (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics*, **14**, 378-9.
- Hofmann K, Stoffel W (1993). TMbase: a database of membrane spanning proteins segments. *Biol Chem Hoppe-Seyler*, **374**, 166-70.
- Holm L, Kaariainen S, Rosenstrom P, et al (2008). Searching protein structure databases with DALI-Litev.3. *Bioinformatics*, **24**, 2780-1.
- Jones DT, Taylor WR, Thornton JM (1994). A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry*, **33**, 3038-49.
- Juretic D, Zoranic L, Zucic D (2002). Basic charge clusters and predictions of membrane protein topology. *J Chem Inf Comput Sci*, **42**, 620-32.
- Kerley-Haniilton JS, Pike, AM, Li N, et al (2005). A p53-do mutant transcriptional response to cisplatin in testicular germ cell tumor-derived human embryonal carcinoma. *Oncogene*, **24**, 6090-100.
- Kelley LA, Sternberg MJ (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc*, **4**, 363-71.
- Krogh A, Larsson B, von Heijne G, et al (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, **305**, 567-80.
- Kyeong MK, Marc GC (2008). Complementary roles of the DRY motif and C-terminus tail of GPCRS for G protein coupling and β -arrestin interaction. *Biochem Bioph Res Comm*, **366**, 42-7.
- Lambert N, Leonard X, De Bolle, ED (2002). ESYPred3D: prediction of proteins 3D structures. *Bioinformatics*, **18**, 1250-6.
- Larkin MA, Blackshields G, Brown NP, et al (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947-8.
- Laskowski RA, MacArthur MW, Moss DS, et al (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst*, **26**, 283-91.
- Mobarec JC, Sanchez R, Filizola M (2009). Modern homology modeling of G-protein coupled receptors: which structural template to use? *J Med Chem*, **52**, 5207-16.
- Nonaka Y, Hiramoto H, Fujita N (2005). Identification of endogenous surrogate ligands for human P2Y₁₂ receptors by *in silico* and *in vitro* methods. *Biochem Biophys Res Commun*, **337**, 281-8.
- Ochiai S, Furuta D, Sugita K, et al (2013). GPR87 mediates lysophosphatidic acid-induced colony dispersal in A431 cells. *Eur J Pharmacol*, **715**, 15-20.
- Ota T, Suzuki Y, Nishikawa T, et al (2004). Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat Genet*, **36**, 40-5.
- Pandian S, Shylajanacyar MD, Murugesan R (2011). Multiple templates-based homology modeling enhances structure quality of AT1 receptor: validation by molecular dynamics and antagonist docking. *J Mol Model*, **17**, 1565-77.
- Rani M, Dikhit MR, Sahoo GC, et al (2011) Comparative domain modeling of human EGF-like module EMR2 and study of interaction of the fourth domain of EGF with chondroitin 4-sulphate. *J Biomedical Res*, **25**, 100-10.
- Rasmussen SG, Choi HJ, Rosenbaum DM, et al (2007). Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*, **15**, 383-7.
- Robert TD, Silvio JG (2007). G-protein-coupled receptors and cancer. *Nature Publishing Group*, **7**, 79-94.
- Roderick EH (2006). 3D Structure and the Drug Discovery Process. *RSC Publishing group*, **13**, 9780854043514.
- Rost B, Yachdav G, Liu J (2004). The predict protein server. *Nucleic Acids Research*, **32**, 321-6.
- Ryckaert JP, Ciccotti G, Berendsen HJC (1977). Numerical integration of the cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J Comp Phys*, **23**, 327-41.
- Sadiq SK, Guixa-Gonzalez R, Dainese E, et al (2013). Molecular modeling and simulation of membrane lipid-mediated effects on GPCRS. *Curr Med Chem*, **20**, 22-38.
- Sahoo GC, Dikhit MR, Rani M, et al (2013). Analysis of sequence, structure of GAPDH of *Leishmania donovani* and its interactions. *J Biomol Struct Dyn*, **31**, 258-75.
- Shailza S, Sharma DK (2011). *In silico* modeling in conjunction with natural products: paving the way for rational drug-design. *Biotechnology and Molecular Biology Review*, **6**, 88-93.
- Shen MY, Sali A (2006). Statistical potential for assessment and prediction of protein structures. *Protein Science*, **15**, 2507-24.
- Shimamura T, Hiraki K, Takahashi N, et al (2008). Crystal structure of squid rhodopsin with intracellularly extended cytoplasmic region. *J Biol Chem*, **283**, 17753-6.
- Tusnády GE, Simon I (1998). Principles governing amino acid composition of integral membrane proteins - application to topology prediction. *J Mol Biol*, **283**, 489-506.
- Varsha G (2008). Metallo beta lactamases in *Pseudomonas aeruginosa* and *Acinetobacter* species. *Expert Opin Invest Drugs*, **17**, 131-43.
- Venkatachalam CM, Jiang X, Oldfield T et al (2003). LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J Mol Graphics*, **21**, 289-307.
- Warne A, Serrano-Vega MJ, Baker JG, et al (2008). Structure of the Beta1-adrenergic G Protein-Coupled Receptor. *Nature*, **454**, 454-86.
- Wiederstein M, Sippl MJ (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, **35**, 407-10.
- Wittenberger T, Schaller HC, Hellebrand S (2001). An expressed sequence tag (EST) data mining strategy succeeding in the discovery of new G-protein coupled receptors. *J Mol Biol*, **307**, 799-813.
- Zhang Y (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.