

Structural analysis and molecular modeling of Human dopamine receptor 5 (DRD5)

Raja Chakraborty*, Sayak Ganguli, Hirak Jyoti Chakraborty and Abhijit Datta

DBT Centre for Bioinformatics, Presidency College Kolkata, India

*Department of Oral Biology, University of Manitoba, Manitoba, Canada R3E0W2

Abstract- Dopamine has been shown to play a key role in the Central Nervous System. Any imbalance in dopamine production and dopamine receptor activity has been implicated widely in the pharmacology, of a number of neurobehavioral disorders existing today. The dopamine receptors interact with G-proteins to transduce dopamine stimulation into intracellular responses. In this work a 3D structure of DRD5 based on the template of High resolution crystal structure of human α -2 adrenergic receptor (PDB code 2RH1A) was generated. The model was assessed using MOLPROBITY. The results revealed that 94% of the residues were found in the favoured region of Ramachandran's plot. Active site analysis showed that predicted binding sites included all the seven transmembrane helices. Our work demonstrates that in silico modeling of proteins has been shown to play an important role in determining protein structure and provides a promising area in possible drug discovery.

Introduction

Parkinson disease and Dopaminergic receptor

Parkinson's disease (PD) is a neurodegenerative disease which affects large number of people worldwide. The symptom of Parkinson's disease (PD) comprises conditions of motor system disorders resulting in tremor, stiffness of limbs and trunk; bradykinesia, and postural instability [1](Civelli *et al.*, 1993). Evidence suggests that these conditions, are due to loss of dopamine-producing brain cells caused by degeneration of dopaminergic neurons in the substantia nigra of the mid brain region and other monoaminergic neurons in the brain stem (Forno., *et al* 1996).

Dopamine has been shown to play a key role in the central nervous system, contributing to important functions of attention, movement of different parts, perception, motivation and emotion. Imbalance in dopamine production and dopamine receptor activity has been implicated widely in the pharmacology, of a number of neurobehavioral disorders existing today (Zhang *et al.*, 2006).

Elaborate understanding of dopaminergic receptor has revealed that they follow three neuronal pathways viz the nigrostriatal pathway, mesocorticolimbic pathway and the tuberoinfundibular pathway. Of the three pathways, degeneration of nigrostriatal pathway which comprises of the neurons of the substantia nigra, and synthesizes dopamine that responds to it leads to Parkinson's disease (Civelli *et al.*, 1993).

Dopaminergic signaling

Dopamine an important neurotransmitter plays critical role in transmission of stimuli.

Dopaminergic signaling in brain has been shown to be mediated by five receptor subtypes through a cascade of intermolecular reactions and can be organized into D1-like (D1 and D5) and D2-like (D2, D3, and D4) classes based on pharmacological and physiological evidences (Missale *et al.*, 1998). The dopamine receptors interact with G-proteins to transduce dopamine stimulation into intracellular responses. The relative homology of D1 and D5 receptors suggests that receptors of the same subfamily might couple to the same set of G proteins for signal transduction. This observation has been shown for D1 like receptors where these receptors interact with Gs complexes resulting in the activation of adenylyl cyclase increasing the intracellular CAMP levels whereas the D2 receptor interacts with Gi complexes to inhibit CAMP production. These biological activities placed the two dopamine receptors in the super family of G protein-coupled receptors (Civelli *et al.*, 1993).

G-Protein coupled Receptors

GPCR form the largest group of membrane receptors having seven trans membrane α helices which share similar sequence stretches of about 25-35 consecutive residues. Crucial to the mechanism of the receptors they must be able to interact with G-Protein to mediate signal transduction in response to extra cellular stimuli. Current estimation shows that almost 50% of the modern drugs sold throughout the world are targeted to these receptors (Wolf *et al.*, 2008, Chelikani *et al.*, 2004).

GPCRs super family comprises of five structurally distinct families which include the Rhodopsin like

receptors (family A) which are shown to contain several highly conserved >90% amino acid sequences in their transmembrane helices, secretin like receptors (family B), glutamate receptors (family C), Adhesion and Frizzled (Taste 2) receptors. The availability of high resolution structure of Rhodopsin and that of β 2 - Adrenoreceptor very recently will allow us to find insights on structure function relationship in other G-protein coupled receptors (*Ballesteros et al., 2001*).

Dopamine receptor 5 (DRD5)

The *DRD5* gene is located in chromosome 4 of human and the protein encoded has similar structure and function as DRD1. The receptor protein is considered to be a direct target for therapeutic applications of small molecules (*Arbouh et al., 2007*). Analogous to the mode of ligand binding by rhodopsin and adrenergic receptors the dopaminergic receptor's ligands must involve the hydrophobic sites because of high sequence similarity of around 75% in their transmembrane domain for D1 like receptors. Identification of potential drug targets and the ability to predict whether a protein will bind with high affinity and specificity to small drug like compounds using 3-D structures has emerged as key area of research.

Methods

Sequence alignment

The binary alignment was used as the basis of modeling. Binary alignment of human dopamine receptor 5 (uniprot accession number P21918) was subjected to binary alignment with human beta 2 adrenergic receptor (uniprot accession number P07550) using BLAST program (Blosum 62 matrix) (*Altschul et al., 1997*). Additional multiple sequences comparison with GPCR sequences P25115, P21728, P02699 was done using ClustalW (*Higgins et al., 1994*). The binary alignment was used as the basis for modeling whereas the multiple alignment was used to cross check whether regions found conserved in the binary alignment could be regarded as being modular (*Wolf et al., 2008*). The location of the transmembrane domains in GPCR was predicted using HMMTOP (*Tusnady et al., 2001*), and DAS (*Cserzo et al., 1997*).

Comparative Homology Modeling

The query protein sequence of Dopamine receptor 5 (DRD5) was searched to find out related protein structure to be used as template using BLAST (Basic local Alignment Search Tool) (*Altschul et al., 1997*) against PDB (Protein Data Bank). The sequence with maximum identity and less e-value was chosen and used as the reference structure for homology modeling. Chain A of β -2 adrenergic receptor (PDB code 2RH1A)

(*Cherezov et al., 2007*) was used as the template structure for modeling of proteins.

Homology model of the seven transmembrane 7TM α helices of the dopamine receptor 5 (DRD5) was constructed using the internet based protein modeling software server, SWISS-MODEL developed by (GlaxoSmithKline) (*Arnold et al., 2006*). The homology model was generated using the 2.4Å resolution of x-ray crystallographic structure of β -2 adrenergic receptor as template.

A major problem in homology modeling is in having a sequence similarity of less than ~ 20% between the target sequence and the template sequence. Homology model generated using such low similarity as reported by (*Wilson et al., 2000*) is considered to be unreliable and less accurate. In this case a sequence similarity of more than ~35% between DRD5 and template human β -2 adrenergic receptor was obtained.

Structure Validation

The GPCR structure obtained from Swiss Model was analysed for Ramachandran's Plot using tools like MOLPROBITY (*Lovell, Davis, et al., 2003*) and RAMPAGE (*Bakker et al., 2002*) to check whether the model structure was stereo-chemically stable and the number of outliers. The model was used further to identify active sites for Ligand binding as well as for docking studies of agonists and antagonists.

Identification of active sites for Ligand binding

The active sites or the rough binding sites of ligands on both human DRD5 and DRD1 (GPCR) were identified using Q-SITE FINDER and CASTp (Computed Atlas of surface Topography of proteins) (*Dundas et al., 2006*). The above programmes are not only used for identification of active sites, binding sites located on the protein surface but are also used to identify void regions buried in the interior of proteins by measuring concave surfaces on the 3-D structure of the proteins.

Accessible solvent area (ASA view)

According to (*Ahmad et al., 2004*) functional properties of proteins strongly correlate with accessible solvent area of amino acids. Therefore a comparative study of the outer surface of GPCR (DRD5) and human β -2 adrenergic receptor was done to have a simple evaluation of the homology model generated using the SWISS-MODEL work bench and also to find out the functional properties of the protein.

Identification of interacting partners:

The protein - protein interaction partners were identified using the STRINGS database.

Results and discussion

Homology modeling of Human GPCR [Dopamine 5 Receptor (DRD5)]

The target protein sequence of human GPCR (DRD5) was compared for selection of homologues with known structures from the PDB using BLAST that searches sequence databases for optimal local alignment to the query sequence. From the BLAST results a hit of 30 reference proteins was found. High sequence similarity and with less E-value was found with Turkey β 1 adrenergic receptor with stabilizing mutations and bound cyanopindolol [PDB code 2vt4B], Cholesterol bound form of human β 2 adrenergic receptor [PDB code 3d4sA], Crystal structure of the human β 2 adrenoreceptor [PDB code 2r4Ra], High resolution crystal structure of human β 2 adrenergic G protein coupled receptor [PDB code 2rh1A], and Crystal structure of Bovine Rhodopsin [PDB code 1f88A] with sequence identity of 39%, 38%, 33%, 37% and 20% respectively.

Although a high percentage of sequence identity were obtained with [PDB code 2vt4B], [PDB code 2r4Ra], and [PDB code 2r4Ra] no significant theoretical models of the GPCR were obtained using the online modeling server SWISS-MODEL. The related High resolution crystal structure of human β 2 adrenergic G protein coupled receptor [PDB code 2rh1A] when used as template resulted in generating a model which was then used for subsequent studies. A major problem in homology modelling is in having a sequence similarity of less than ~ 20% between the target sequence and the template sequence. Homology model generated using such low similarity as reported by (Wilson *et al.*, 2000) is considered to be unreliable and less accurate.

A phylogenic analysis between Human DRD5, Rat DRD5, Human DRD1, Human β 2 adrenergic receptor and Bovine Rhodopsin was also done. The results showed that human β 2 adrenergic receptor was closely related to Human Dopamine receptor 5 (DRD5) than that of Bovine Rhodopsin. Similarly Human DRD5 and Rat DRD5 were also very closely related with almost 100% sequence similarity whereas Human DRD1 and Human DRD5 also shared a high sequence similarity.

Validation of the modeled GPCR

Peptides are built from a chain of amino acids held together by covalent bonds. The conformation of the protein backbone in space can be explained by the rotation of each amino acid around two such bonds, one is the N-C α (called the Phi ϕ) and the other around the C α -C bond (called the Psi ψ). Interestingly most combinations of the ϕ and ψ angles are not allowed due to steric hindrances, yet the

molecules try to orient themselves in such a manner to reach a state of optimal conformational stability. The distribution of these two angles for the amino acids of a particular protein is plotted in a graph also known as the Ramachandran's Plot (Ramachandran *et al.*, 1963).

The stereo chemical property of the model obtained was evaluated with Ramachandran's Plot calculations using the MOLPROBITY server. The results obtained showed that the dihedral angles ϕ and ψ of the theoretical model GPCR had 94% and 97.6% of all the residues were in favored and allowed regions respectively. In comparison, the stereo chemical quality of the model template GPCR (2rh1a) had 98.9% and 100% of all residues were in favored and allowed regions and there were no outliers. Plot result also revealed that there were 8 outliers for the theoretical model GPCR, which include ASN74, MET75, TRP203, GLU205, CYS278, ALA279, PRO280 and PRO329. Since the model GPCR had almost 98% of the all the residues (324/329) in allowed and 94% of all the residues in favored regions the model was stereo chemically stable and reliable for subsequent studies. The higher stereo chemical Quality of the model might be due to high sequence similarity of ~35% between the template and the model sequence.

Active Site Identification

The active sites of the theoretical model GPCR was built using the online server Q-SITE FINDER and CASTp (Computed Atlas of surface Topography of proteins) (Dundas *et al.*, 2006). The results revealed a total number of 42 binding sites and are shown in Figure 4a. When a comparative study with the template structure of 2rh1A was done with that of the model GPCR it was found the binding pocket (Green) in the model which had 72 residues; 99 LYS, 102 ALA, 103 GLU, 106 GLY, 107 TYR, 108 TRP, 112 ALA, 113 PHE, 115 ASP, 117 TRP, 118 VAL, 121 ASP, 122 ILE, 125 SER, 126 THR, 129 ILE, 172 ILE, 173 SER, 177 VAL, 180 ASN, 182 HIS, 183 ARG, 184 ASP, 185 GLN, 186 ALA, 187 ALA, 189 TRP, 202 THR, 205 GLU, 206 GLU, 207 ASP, 208 PHE, 211 PRO, 212 ASP, 214 ASN, 216 GLU, 217 ASN, 218 CYS, 219 ASP, 220 SER, 221 SER, 222 LEU, 223 ASN, 224 ARG, 226 TYR, 227 ALA, 230 SER, 231 SER, 234 SER, 235 PHE, 306 PHE, 310 TRP, 313 PHE, 314 PHE, 316 LEU, 317 ASN, 319 MET, 320 VAL, 321 PRO, 323 CYS, 324 SER, 325 GLY, 326 HIS, 328 GLU, 333 PHE, 334 PRO, 335 CYS, 336 VAL, 338 GLU, 341 PHE, 342 ASP, 346 TRP, 350 ALA was conserved. Therefore the site (Green) as shown in Figure 4b

can be said to be more favourable in binding agonists and antagonists.

Location of Transmembrane (TM) Helices

The dopamine receptor subtypes are shown to be highly conserved family of GPCR. Evidence suggest, the dopamine receptors show around 40%-50% of sequence identity overall and almost 87%-90% similarity in their transmembrane domains, which is significant from the point of view that these regions are important sites for ligand binding (*Livingstone et al., 1992*). The transmembrane domains of Human GPCR (DRD5), rat GPCR (DRD5) and $\beta 2$ Human adrenergic receptor as shown in Figure 5, were found using online server tools HMMTOP and DAS. The results obtained from online server (HMMTOP in predicting transmembrane helices) showed that TM-I of human GPCR (DRD5) is present between amino acid residues 50 and 65, TM-II is present between 78 and 96, TM-III present between 117 and 136, TM-IV present between 156 and 176, TM-V present between 225 and 244, TM-VI present between 299 and 320 and TM-VII is present between 336 and 356. Similar results were also obtained from DAS [supplementary data]. From the transmembrane domains it can be seen that PRO 96 and PRO 356 is present at the end of TM-II and TM-VII respectively, and is breaking the helix and providing a definite bend in the helical structure. Aspartic acid residue ASN 118 in TM-III may be responsible for having electrostatic interactions with cationic ligands. The serine residues SER 228, SER 229, SER 230 and SER 233 in TM-V may be responsible for hydrogen bonding with hydroxyl group present on the catechol ring of the ligands. Similarly aromatic interactions between agonists or antagonists with aromatic residues 298 PHE, 306 PHE and 307 PHE in TM-VI of the GPCR may help in ligand interactions with the receptor. The aromatic residues in TM-VII, PHE 341, PHE 344, PHE 347 and TRP 346 may be important residues for ligand binding. PRO 171 in TM-IV and PRO 237 in TM-V may be causing definite bends in the helical structure of the protein. These predictions are made on the basis of mutagenesis studies on the GPCR, $\beta 2$ adrenergic receptor which has been shown to interact with catecholamine ligands and show primary contact points of receptor-ligand interactions (*Livingstone et al., 1992*).

Comparison of outer surface

Since the GPCR (DRD5) and the template β -2 adrenergic receptor are membrane proteins, there outer surface is exposed to a lipid environment *in vitro*. So a comparative study of the outer surface of GPCR (DRD5) and human β -2 adrenergic receptor may provide a simple evaluation of the homology model generated

using the SWISS-MODEL work bench. Comparison of the outer surface was done using accessible solvent area (ASA) view. The 2D spiral plot of ASA view provides a graphical accessible representation of the amino acids residues in terms of their exposed surface area (*Ahmad et al., 2004*). The amino acids residues with smallest accessible area are placed at the origin of the spiral where as the residues with a larger accessible area are successively placed. Evidence suggests that functional property of a protein and the active binding sites correlate with the accessible surface area. Possible active sites are potentially found in the higher accessible region of a protein for binding ligands (*Ahmad et al., 2004*). The results obtained could be useful in a crude estimation of the active binding sites for ligand docking in a protein. The active sites obtained from CASTp and Q site finder results were in conformation with the ASA view results obtained.

References

- [1] Ao Zhang, Ying Kan and Fuying Li Behavioral Profile Pages 587-630.
- [2] Olivier Civelli, James R. Bunzow, and David K. Grandy (1993) *Annu. Rev. Pharmacol. Toxicol.*, 281-307.
- [3] Sunahara R. K., Guan H. C., O'Dowd, B. F., Seeman P., Laurier L. G., et al. (1991) *Nature*, Pages 614-19.
- [4] Grandy D. K., Zhang Y., Bouvier C., Zhou Q. Y., Johnson R. A., et al. (1991) *Proc. Natl. Acad. Sci. USA*, Pages 175-79.
- [5] Juana. Ballesteros, Lei Shi, and Jonathana. Javitch (2001) *Mol Pharmacol*, Pages 1-19.
- [6] Meena K. Sakharkar and Kishore R. Sakharkar (2007) *Current Drug Discovery Technologies*, Pages, 48-58.
- [7] Maurits EL Arbouw, Jeroen PP van Vugt, Toine CG Egberts and Henk-Jan Guchelaar (2007) *Pharmacogenomics*, Pages 159-176
- [8] Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W. and Lipman D. J. (1997) *Nucleic Acids Res*, Pages 3389-3402.
- [9] Lovell S.C., Davis I.W., Arendall III W.B., Bakker P.I.W. de, Word J.M., Prisant M.G., Richardson J.S. and Richardson D.C. (2002) *Proteins: Structure, Function & Genetics*, Pages 437-450.
- [10] Laurie A.T., Jackson R.M. (2005) *Bioinformatics*, 1908-1916.

- [11] Prashen Chelikani, Viktor Hornak, Markus Eilers, Phillip J. Reeves, Steven O. Smith, Uttam L. RajBhandary, and H. Gobind Khorana (2007) *PNAS*, Pages 7027–7032.
- [12] Wolf S., Bockmann M., Howeler U., Schlittera J., Gerwert K. (2008) *FEBS LETTER*, 8.
- [13] Jeremiah Stitham, Scott R. Gleim, Karen Douville, Eric Arehart, and John Hwa (2006) *The Journal Of Biological Chemistry*, Pages 37227–37236.
- [14] Zhexin Xiang (2006) *Curr Protein Pept Sci*, Pages 217–227.
- [15] Thompson JD, Higgins DG, and Gibson TJ. (1994) *Nucleic Acids Res*, Pages 4673-80.
- [16] Davis I.W., Leaver-Fay A., Chen V.B., Block J.N., Kapral G.J., Wang X., Murray L.W., Arendall W.B. 3rd, Snoeyink J., Richardson J.S., Richardson D.C. (2007) *Nucleic Acids Res*, Pages 375-83.
- [17] Bordoli A. K., Kopp L., and Schwede T. (2006) *Bioinformatics*, Pages 195-201.
- [18] Joe Dundas, Zheng Ouyang, Jeffery Tseng, Andrew Binkowski, Yaron Turpaz, and Jie Liang (2006) *Nucleic Acids Res*, W316-W118.
- [19] Tusnady G.E. and Simon I. (1998) *J. Mol. Biol*, Pages 489-506.
- [20] Cserzo M., Wallin E., Simon I., von Heijne G. and Elofsson A. (1997) *Prot. Eng*, Pages 673-676.
- [21] Cherezov V., Rosenbaum D.M., Hanson M.A., Rasmussen S.G., Thian F.S., Kobilka T.S., Choi H.J., Kuhn P., Weis W.I., Kobilka B.K., Stevens R.C. (2007), *Science*, Pages 1258-1265.
- [22] Ritchie D.W. and Kemp G.J.L. (2000) *PROTEINS: Struct. Funct. Genet*, Pages 178-194.
- [23] Laurie A.T., Jackson R.M. (2005) *Bioinformatics*, Pages 1908-1916.
- [24] Cristina Missale S., Russel Nash, Susan W. Robinson, Mohamed Jaber, and Marc G. Caron (1998) *Physiol Rev*, Pages 189-225.
- [25] Shandar Ahmad, M. Michael Gromiha, Hamed Fawareh and Akinori Sarai (2004) *BMC Bioinformatics*
- [26] Lipinski C.A., Lombardo F., B.W. Dominy, P.J. Feeney, (1997) *Adv.Drug.Delivery Rev*, 4-25.
- [27] Ramachandran G.N., Ramakrishnan C. and Sasisekharan V. (1963) *J. Mol. Biol.* 95-99.

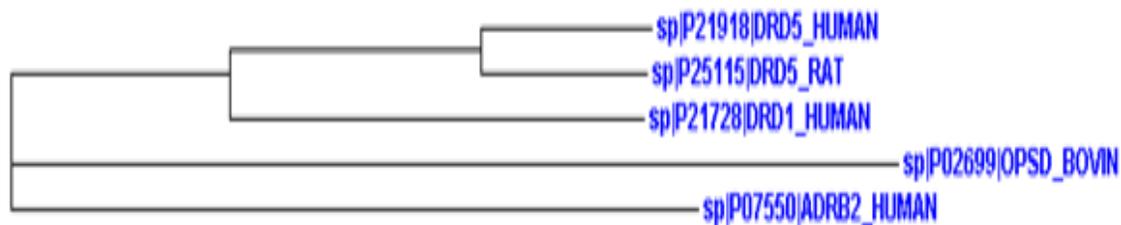


Figure 1: Phylogenetic analysis of the sequences of Human DRD5, Rat DRD5, Human DRD1, Human Beta 2 adrenergic receptor, and Bovine Rhodopsin.

ClustalW was used to align each of the query template sequence. Gap penalties (P0) of 7, 8,9,10,11,12,15 and 20 were used along with gap extension (Pc) values of 0.1 or 1. The results obtained showed no significant changes to the alignment so the default values of gap extension 10 and gap penalty 0.1 were used.

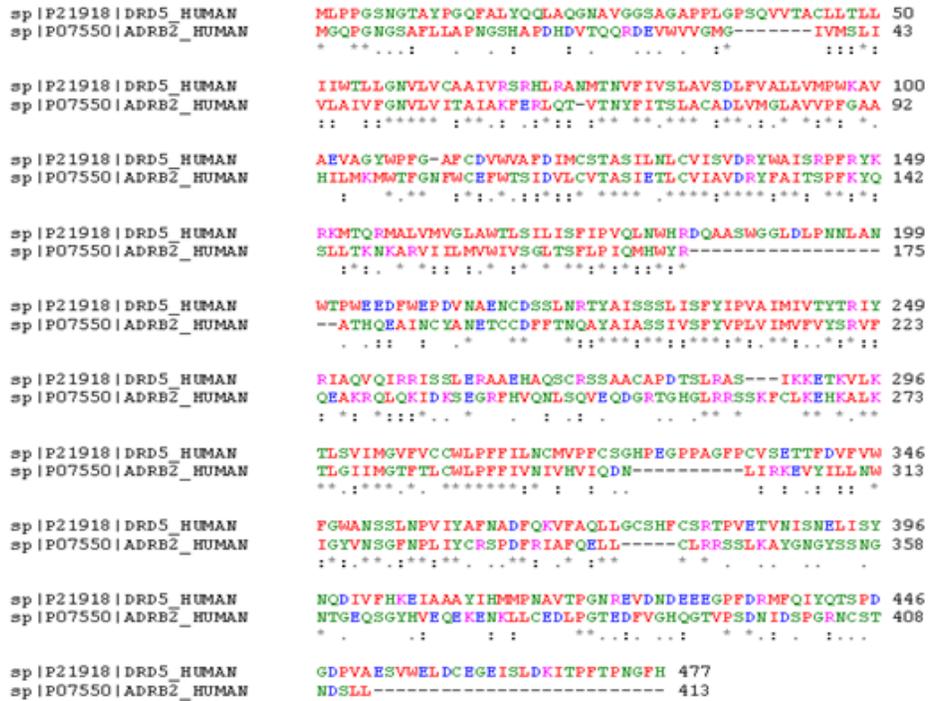


Figure 2: Sequence alignment of Human Dopamine Receptor 5 with Beta 2 Andrenergic Receptor. The default values of Gap penalty (Po) of 10 and Gap extension (Pc) of 0.1 were used for aligning the sequences.

Ramachandran Plot Analysis

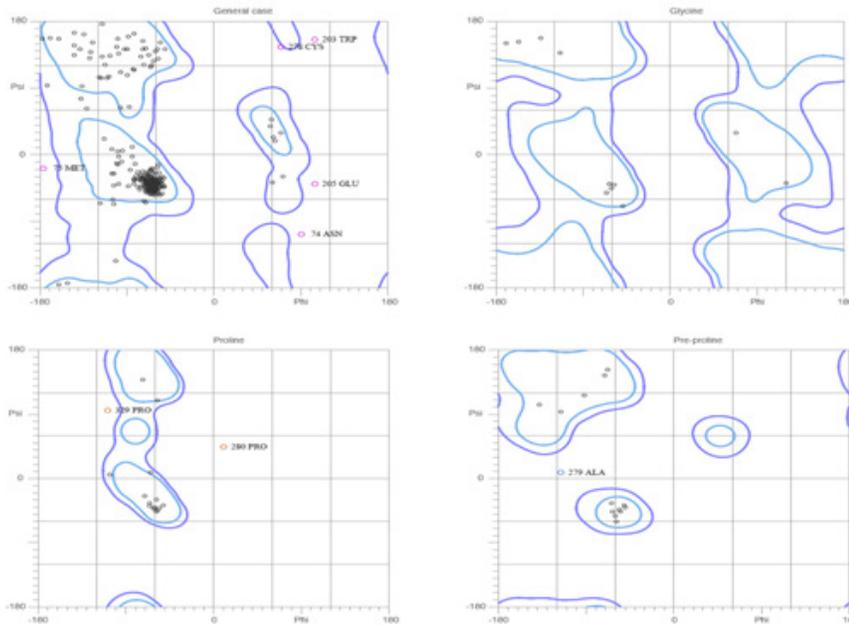


Figure 3: Ramachandran's Plot analysis was carried out using MOLPROBITY. The figure shows almost all the residues are in the energetically favored regions except 8 residues which can be viewed as red dots.

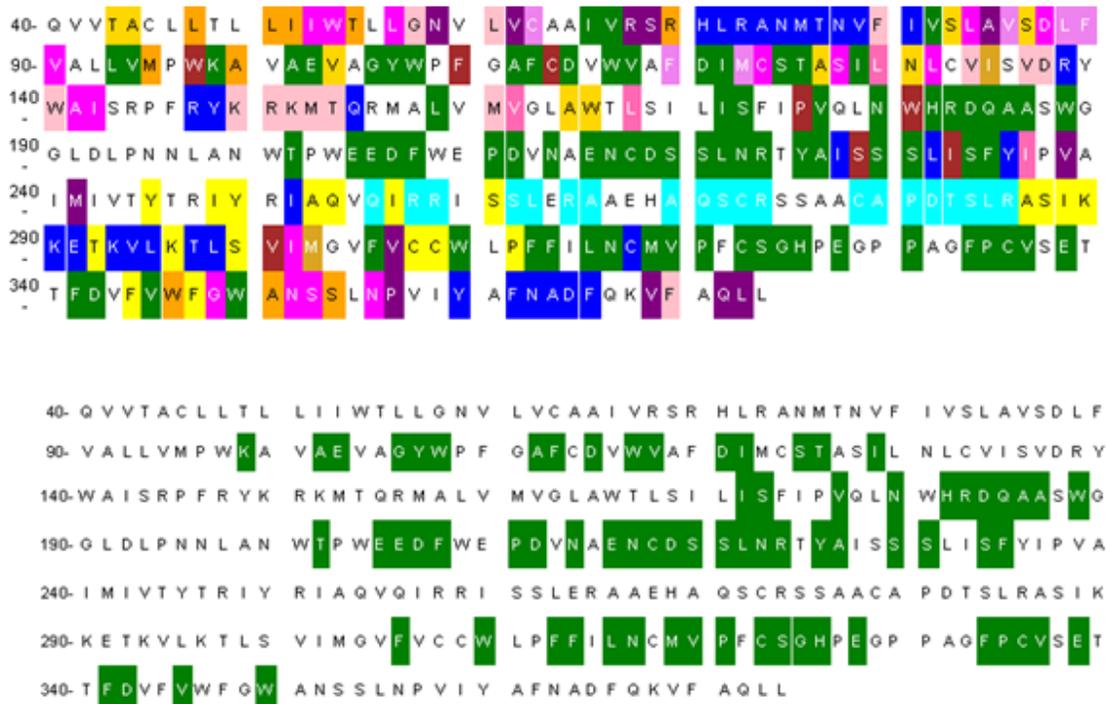


Figure 4: a) The figure shows the active binding sites of the model GPCR. A total of 42 binding sites were obtained represented in different colors in the figure. b) The figure shows the favored binding site of agonists and antagonists in the modeled GPCR.

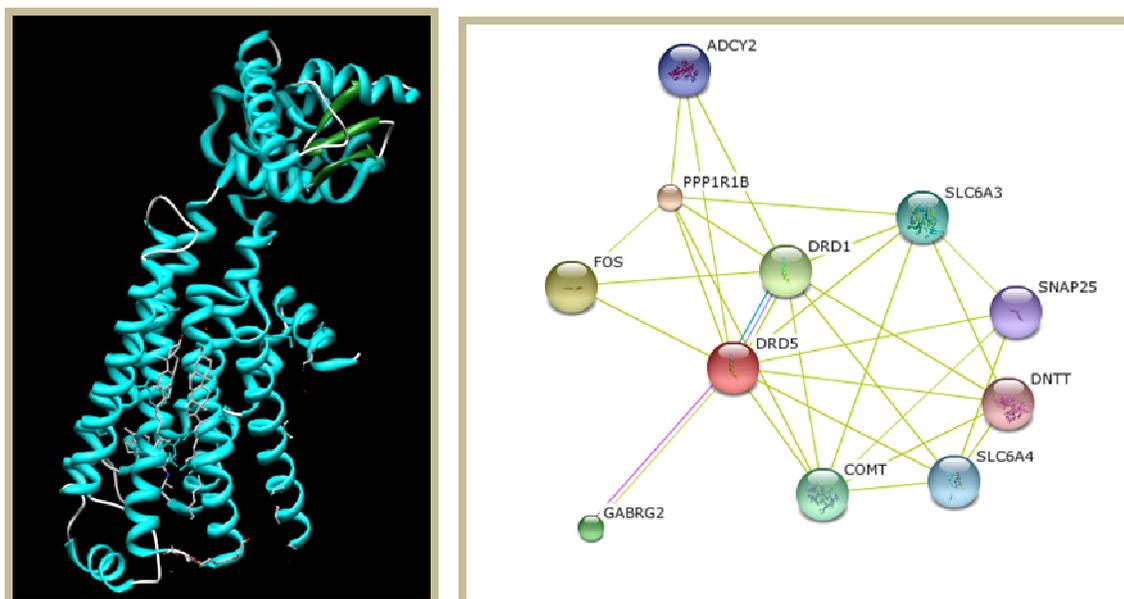


Figure 5-Structure of the modeled protein and its interacting partner.