

# Ligand Docking Methods to Recognize Allosteric Inhibitors for G-Protein-Coupled Receptors

K Harini<sup>1</sup>, S Jayashree<sup>2,3</sup>, Vikas Tiwari<sup>1</sup>, Sneha Vishwanath<sup>4,5</sup> and Ramanathan Sowdhamini<sup>1</sup> 

<sup>1</sup>Department of Bioinformatics, National Centre for Biological Sciences, Bangalore, India. <sup>2</sup>Department of Biotechnology, Vellore Institute of Technology, Vellore, India. <sup>3</sup>Royal Melbourne Institute of Technology (RMIT) University, Melbourne, VIC, Australia. <sup>4</sup>Department of Biophysics, Indian Institute of Science, Bangalore, India. <sup>5</sup>Department of Zoology, University of Cambridge, Cambridge, UK.

Bioinformatics and Biology Insights  
Volume 15: 1–7  
© The Author(s) 2021  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/11779322211037769



**ABSTRACT:** G-protein-coupled receptors (GPCRs) are membrane proteins which play an important role in many cellular processes and are excellent drug targets. Despite the existence of several US Food and Drug Administration (FDA)-approved GPCR-targeting drugs, there is a continuing challenge of side effects owing to the nonspecific nature of drug binding. We have investigated the diversity of the ligand binding site for this class of proteins against their cognate ligands using computational docking, even if their structures are known already in the ligand-complexed form. The cognate ligand of some of these receptors dock at allosteric binding site with better score than the binding at the conservative site. Interestingly, amino acid residues at such allosteric binding site are not conserved across GPCR subfamilies. Such a computational approach can assist in the prediction of specific allosteric binders for GPCRs.

**KEYWORDS:** G-protein-coupled receptors, allosteric ligands, AutoDock, cognate ligands, Tanimoto co-efficient

**RECEIVED:** March 11, 2021. **ACCEPTED:** July 20, 2021.

**TYPE:** Original Research

**FUNDING:** The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: R.S. would like to acknowledge her JC Bose Fellowship (JC Bose fellowship (SB/S2/JC-071/2015) from Science and Engineering Research Board, Bioinformatics Centre Grant funded by Department of Biotechnology, India (BT/PR40187/BTIS/137/9/2021 and NCBS (TIFR) for infrastructural facilities.

**DECLARATION OF CONFLICTING INTERESTS:** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**CORRESPONDING AUTHORS:** K Harini, Department of Bioinformatics, National Centre for Biological Sciences, GKVK Campus, Bellary Road, Bangalore 560065, India. Email: harinik@ncbs.res.in  
Ramanathan Sowdhamini, National Centre for Biological Sciences, GKVK Campus, Bellary Road, Bangalore 560065, India. Email: mini@ncbs.res.in

## Introduction

Membrane proteins constitute nearly 40% of the human genome.<sup>1,2</sup> G-protein-coupled receptors (GPCRs), being important drug targets, are one of the most well studied<sup>3–5</sup> among the membrane proteins. There are close to 800 known GPCRs in the human genome, and they are grouped into 5 families depending on their substrates (such as peptides and amines<sup>6</sup>). GPCRs act as an important environmental sensor and play a vital role in diverse signalling processes. As a result, there is enormous diversity in their position in different biological pathways.

The hallmark of GPCRs is the presence of 7 transmembrane helices (TMHs), where extracellular loop regions and parts of TMHs recognize the cognate substrates and indentations in the intracellular loop regions provide the capacity to participate in diverse biological pathways.<sup>7</sup> In the past 2 decades, there have been substantial insights into the structural features of few important GPCRs, despite the inherent challenges in the structure determination of membrane proteins. Consequently, we now have a great deal of information about the conformational changes that occur subsequent to ligand binding, distinct sites for agonist and antagonist binding, and so on. Characteristic functional motifs in the intracellular regions, the presence of conserved charged residues in the intracellular face and the presence of prolyl residues are known to contribute to signalling and conformational changes.<sup>8</sup> We now have close to 389 Protein Data Bank (PDB) entries (also organized in specialized GPCR-EXP database <https://zhanglab.dcm.bmed.umich.edu/GPCR-EXP/>)<sup>9</sup> which pertain

to GPCRs, either in the apo-form or ligand-bound form (agonist or antagonist-bound states)). There has been successful design of GPCR drugs, such as haloperidol,<sup>10</sup> over the years. As of November 2017, the US Food and Drug Administration (FDA) has approved drugs against 134 GPCRs.<sup>11</sup>

Despite the structural insights available, drug design for GPCRs remains highly challenging due to the inherent characteristics of substrate promiscuity, structural similarities of drug molecules and sequence similarity within subfamilies of GPCRs. There have been continuing demands to provide drug solutions to address this prevailing feature of GPCRs. Mutations in either GPCRs or their interacting proteins have been widely implicated in neurodegenerative and other diseases.<sup>12,13</sup> Some of the important GPCR drug targets are in the area of neurodegenerative diseases, where there is appalling amounts of side effects observed in patients who are treated with GPCR drugs. Recent efforts have, therefore, focused on the design of allosteric inhibitors.<sup>14–16</sup>

In this article, we report systematic computational ligand docking experiments using selected GPCRs with known information on ligand binding, to present which factors enable best capture of near-native ligand binding and how this can be employed subsequently to identify allosteric inhibitors. Our choice of ligand docking algorithm is AutoDock which is a well-known software that has gone through rigorous analyses by other groups as well.<sup>17</sup> We first describe the analysis of binding poses, where different levels of known information on structural or evolutionary conservation can guide the docking process. Objective measures, such as Tanimoto co-efficient,



have been employed to assess the comparison between docking poses and the structural data for ligand binding. This approach enables the identification of structurally reasonable docking poses which are not close to the native pose, namely allosteric binding. Finally, to test this approach, we have performed blind docking on one of the GPCRs, chemokine receptors, using specialized set of ligands known to bind GPCRs.<sup>18</sup> Few allosteric binding sites could indeed be recognized using this novel computational approach. Interactions between ligand and chemokine receptor at the allosteric site were found to be stable, when subject to molecular dynamics (MD) simulations.

## Methods

### *Selection of GPCR structures and ligands for analysis*

All the available GPCR structures were downloaded from PDB (April 2017).<sup>19</sup> Among 71 such structures, several identical GPCR structures with different resolution and agonist/antagonist bound forms were observed. Thus, for every GPCR with bound ligand, one best structure with highest resolution was selected to avoid redundancy. A set of 27 GPCRs were selected for final analysis (Table 1). G-protein-coupled receptor alignments were obtained from GPCRdb for the subfamily of GPCR receptors to study evolutionary conservation.<sup>20</sup> Ligands were retrieved from PDB, PubChem,<sup>21</sup> or GLASS database.<sup>18</sup> GLASS database is a resource to retrieve ligands that are known to bind to GPCRs and were used to explore allosteric binding. Out of 873 ligands (Supplementary Table 1), 46 were chosen based on Lipinski's drug likeness and XlogP (less than 2). The 46 ligands were clustered using ChemMine tool<sup>22</sup> and finally 29 compounds were used for docking (Supplementary File). These compounds were converted to PDB format using Open babel tool.<sup>23</sup>

### *Grid setting for GPCR-ligand docking*

The protein and ligand coordinates were separated from PDB file and used for docking. In 2 out of 3 experiments, ligand coordinates were derived from PubChem (please see below). Both the agonist and the antagonist bind in the extracellular domains for most of the GPCRs in distinct sites. Hence, the grid region of choice for docking was chosen in the extracellular region. However, different grid settings were considered to identify the parameters in grid setting that will help in correct prediction of ligand binding to GPCRs. These are as follows:

Level 1: Ligand coordinates were taken as such from PDB, but ligand coordinates were kept flexible for docking. Only the extracellular half of the receptor region was selected as grid box (semiblind docking) (Figure 1A).

Level 2: Ligand coordinates were taken from PubChem and ligand was kept flexible. A smaller grid that covers just the known ligand binding site was selected as grid box (guided docking) (Figure 1B).

**Table 1.** List of GPCRs used in the study.

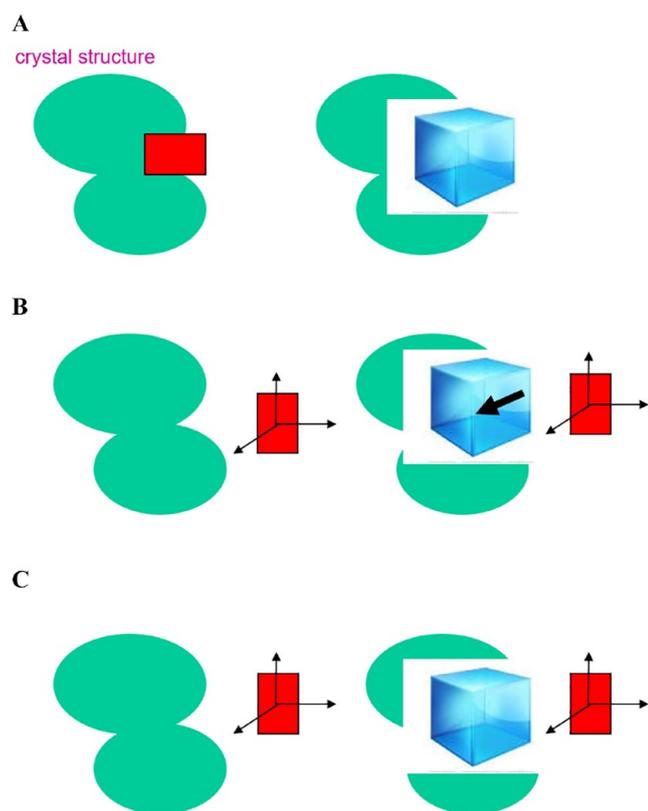
SERIAL NO.	GPCR NAME	PDB ID
GPCR-1	5-HT1B receptor	4IAQ
GPCR-2	5-HT2B receptor	4IB4
GPCR-3	A2A receptor	2YDO
GPCR-4	M2 receptor	3UON
GPCR-5	M3 receptor	4DAJ
GPCR-6	$\beta$ 1-adrenoceptor	2VT4
GPCR-7	$\beta$ 2-adrenoceptor	2RH1
GPCR-8	AT1 receptor	4YAY
GPCR-9	CCR5	4MBS
GPCR-10	CRF1 receptor	4K5Y
GPCR-11	CXCR4	3ODU
GPCR-12	D3 receptor	3PBL
GPCR-13	FFA1 receptor	4PHU
GPCR-14	mGlu5 receptor	4OO9
GPCR-15	H1 receptor	3RZE
GPCR-16	LPA1 receptor	4Z34
GPCR-17	$\delta$ receptor	4BUO
GPCR-18	$\kappa$ receptor	4N6H
GPCR-19	$\mu$ receptor	4DJH
GPCR-20	NOP receptor	4DKL
GPCR-21	Rhodopsin	4EA3
GPCR-22	OX2 receptor	4S0V
GPCR-23	P2Y1 receptor	4XNV
GPCR-24	P2Y12 receptor	4NTJ
GPCR-25	PAR1	3VW7
GPCR-26	S1P1 receptor	3V2Y
GPCR-27	SMO	4JKV

Abbreviations: GPCR, G-protein-coupled receptors; PDB, Protein Data Bank. GPCR proteins were selected based on the highest-resolution structure available in PDB. All class-A GPCRs structures were selected from PDB and a nonredundant data set was obtained.

Level 3: Ligand coordinates were taken from PubChem and ligand was kept flexible. Only the extracellular half of the receptor region was selected as grid box (semiblind docking) (Figure 1C).

### *Molecular docking*

AutoDock 4.2.6 was used for docking studies.<sup>17</sup> Protein and ligand structures were prepared using ADT (AutoDock Tool). The hydrogen atoms were added and water molecules



**Figure 1.** Different grids used for docking of ligands to GPCRs. (A) For Level 1, extracellular half of GPCRs was considered for grid box generation and ligand coordinates were taken from PDB. (B) For Level 2, smaller grid box covering only the known ligand-binding site was defined, and ligand coordinates were retrieved from PubChem. (C) For Level 3, the ligands coordinates were retrieved from PubChem while the grid box was defined by considering the extracellular half region of GPCRs. GPCR indicates G-protein-coupled receptors.

were removed followed by addition of gasteiger charges. The grid parameter file was generated with default distance (0.375 Å) between grid points. Grid space was defined around the extracellular site of GPCRs. The search parameter was set to 100 Genetic Algorithm (GA) runs and Lamarckian genetic algorithm was used. Autogrid was run followed by AutoDock. Similar parameters were used for all ligands docking to a single protein. The docking results were analysed, and protein-ligand complex was made using ADT.

#### Calculation of Tanimoto co-efficient and identification of allosteric sites

To identify how similar the protein-ligand docking results are as observed in the X-ray crystallographic structure of complexes, we used the Tanimoto co-efficient score. Any atom within 4 Å distance was considered to be a contact with the ligand.

Let 'a' be the contacts in PDB structure and 'b' be the contacts in the AutoDock complex then the Tanimoto co-efficient (TC) is calculated as

$$TC \text{ value} = \frac{\text{No. of common contacts}(a \cap b)}{\left( \frac{\text{No. of contacts in } a + \text{No. of contacts in } b}{2} \right) - (a \cap b)}$$

Furthermore, for each ligand, all 100 poses were checked for their binding site compared to native ligand. The minimum distance between all the atoms of native ligand and all the atoms of docked ligand was calculated and if the minimum distance is more than 5 Å for any pose, then that pose was considered as allosteric site-binding pose.

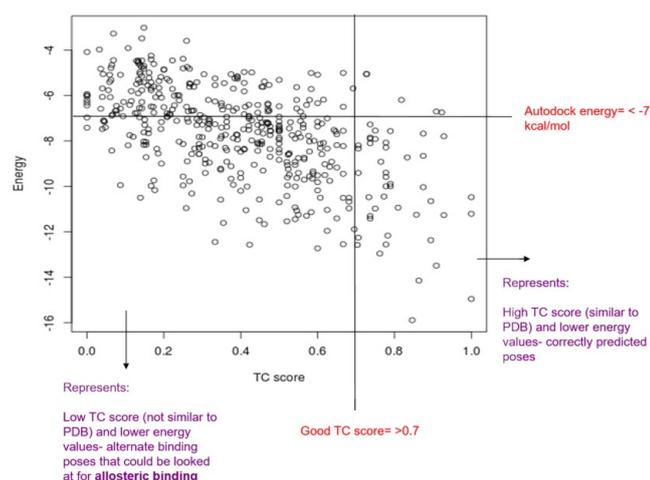
#### MD simulation

The protein-ligand complex structure was subjected to MD simulation using Desmond module of Schrodinger to assess the interaction stability of complex.<sup>24</sup> Initially, the complex structure was processed using protein preparation wizard of maestro (Schrodinger Release 2019-4: Maestro, Schrodinger, LLC, New York, NY, 2019) which assigns bond orders, sets protonation state, optimizes H-bonds and then minimizes the structure.<sup>25</sup> The structure (CXCR4-ligand complex) was protonated at pH 5.5 before minimization in protein preparation wizard. The membrane information was retrieved from OPM database.<sup>26</sup> After processing, the structure was solvated in the presence of membrane (POPC) using TIP4P water model. Orthorhombic box, with buffer distance of 10 Å, was used followed by minimization of box size. The system was neutralized and 150 mM salt (NaCl) was added. System builder was run and output of system builder (solvated system) was used for MD simulation. The system was subjected to default relaxation protocol of Desmond followed by production MD run for 100 ns. Simulation results were analysed using Simulation Interaction Diagram and Simulation Event analysis modules of Schrodinger. The analysis was performed for entire range of simulation time. Root-mean-square deviation (RMSD) is calculated for each frame by aligning the complex to protein backbone of the reference frame. Significantly higher values of "Lig fit Prot" than protein RMSD signifies the diffusion of ligand away from its initial binding site. "Lig fit lig RMSD" is calculated by aligning the ligand on the reference ligand conformation and it indicates the internal fluctuation of ligand.

## Results

### Prior knowledge of the binding mode helps to predict correct protein-ligand interaction

Twenty-seven GPCRs from PDB were selected for analysis. Each of the protein-ligand complexes was separated, and docking was performed as mentioned in Methods. The 3 different grid conditions were used as mentioned in Methods (Level 1, 2, and 3) (Figure 1A-C). We observed that in 50% of the instances, when the ligand coordinates were taken from PDB and semiguided docking was performed, we could



**Figure 2.** TC score vs docking energy score plot. High TC score indicates a high similarity between docked ligand pose and crystal pose. Lower docking score indicates the stable binding of ligand to GPCR. GPCR indicates G-protein-coupled receptors; PDB, Protein Data Bank; TC, Tanimoto co-efficient.

get the protein-ligand complex with high TC score with the original PDB structure and AutoDock energy. In Level 2 of docking (runs with use of ligand coordinates from PubChem and guided docking), complexes with high TC score and low energy could be obtained in 30% of instances. Level 3 (ligand from PubChem and semiguided docking), however, did not result in complexes similar to the original structure in most of the cases (only 3 out of 27 GPCRs had high TC scores) (Figure 2: TC score Vs Energy correlation plot for Level 1). This trend can be explained owing to the increased levels of difficulty in docking and lower clues we provide from the native pose of the ligand under question. In level 1, the position of the ligand is quite close to the native pose of the ligand (please see Figure 1).

### Identifying novel binding modes using the TC score and energy values for GPCR-ligand complexes

We selected complexes which had low TC score (that indicates a different binding mode compared to original ligand binding site [hereafter referred as 'OLBS']) and low-and-negative docking energy (that indicates stable binding) on the basis of Level 1 results. First 3 highest scoring complexes were selected for detailed analysis of the docking poses. These include beta-2-adrenergic receptor, chemokine receptor, and 5-HT<sub>2B</sub> receptor. It was observed that for beta-2-adrenergic receptor and 5-HT<sub>2B</sub> receptors, the ligand binds in the same pocket as in original complex, but is displaced along Y-axis of the protein with a large overlap with OLBS (Figure 3A and B). This results in low TC score for these complexes. For chemokine receptor, however, we could find novel binding sites for the cognate ligand with minimum overlap with OLBS (Figure 3C). Thus, we used chemokine receptor-ligand complex for further analysis on allosteric binding in GPCRs.

### Residues at the alternate novel binding sites are not conserved

Twenty-three homologous sequences of chemokine receptors were used to identify conservation of residues at the endogenous (orthosteric) binding site and allosteric binding site. It was observed that 7 out of 12 residues at OLBS were conserved, while only 3 out of 12 residues were evolutionarily conserved at the allosteric site, even within the same subfamily of receptors. This clearly suggests that allosteric binding site is novel (Figure 4).

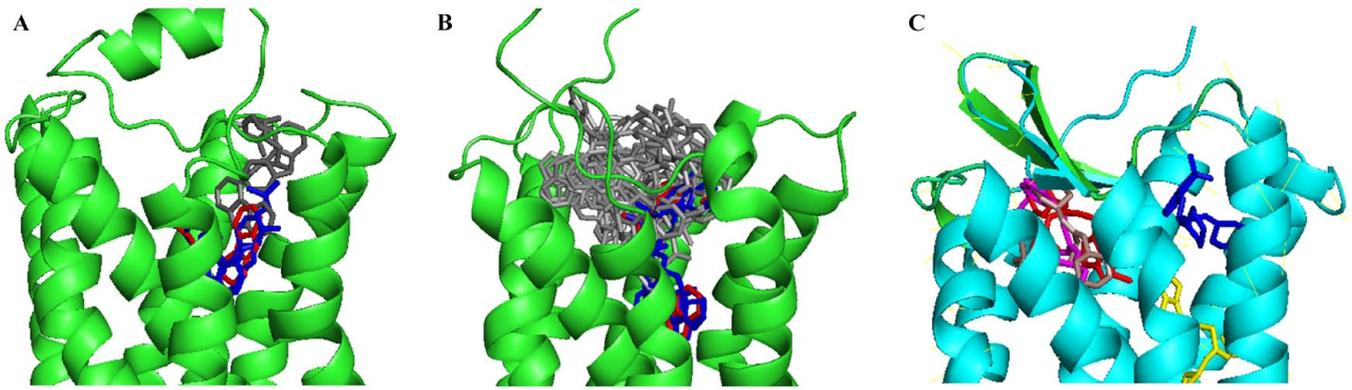
### Identifying novel allosteric binders from GLASS database for chemokine receptor

In order to explore allosteric binders for CXCR4, ligands were retrieved from GLASS database and blind docking was initiated. The docking energy of best docking pose for ligands varied from -6.25 kcal/mol (for ligand '25178561') to -11.27 kcal/mol (for ligand '483559'). Out of 29 ligands, the overall best pose (in terms of docking energy) for 6 ligands also belongs to the largest cluster. Similar analysis was performed for all the ligands. Interestingly, it was observed that all 100 poses are allosteric poses for ligand '76381' (Figure 5). The best pose of ligand '76381' in terms of best docking energy is also the best pose among the largest cluster. Therefore, this best pose for '76381' ligand to CXCR4 was subjected to MD simulation to assess the interaction stability of the complex.

The CXCR4-76381 complex was subjected to 100 ns simulation and the RMSD was calculated. Root-mean-square deviation plot indicates that the ligand moves from its original binding position during simulations (Figure 6A). Major interacting residues of CXCR4 are Tyr256 and Gln200 (Figure 6B). Most of the interactions involve H-bonding and atom-wise interactions of 76381 are shown (Figure 6C). All contacts with the CXCR4 are lost at around 40 ns but regains subsequently, as indicated by total contact plot (Figure 6D). H-bond analysis between 76381 and CXCR4 (including H-bond with water) was also performed, and it was found that at multiple instants during simulation, the H-bond contact was lost and re-formed (Figure 7). However, at no point during simulation, there was any overlap between the native ligand binding mode (ITD) and 76381 binding mode (Figure 8; Supplementary Video file <https://drive.google.com/file/d/1GvO7XiVRjFz6S4H6e5sQwtCxP4VV3jBO/view?usp=sharing>) suggesting that these are true allosteric binding sites.

### Discussion

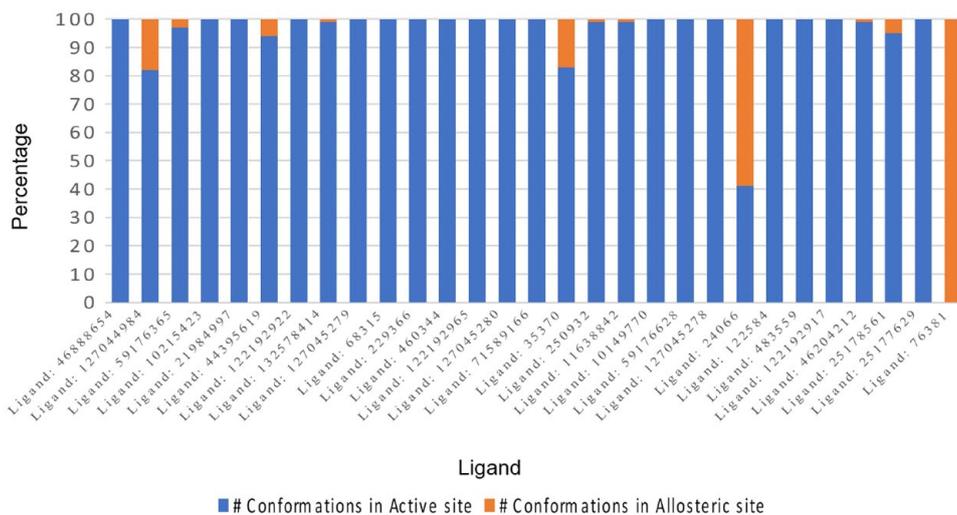
G-protein-coupled receptors are important drug targets for several approved human drugs. Since all the GPCRs have similar 7-transmembrane structure, it becomes challenging to specifically design drugs to target particular GPCRs. Thus, designing specific allosteric modulators is an important and challenging task.<sup>14</sup> In this study, we first understand how well



**Figure 3.** Novel ligand binding poses for high scoring complexes. (A) Ligand poses for beta-2 adrenergic receptor indicating the overlap between novel docked ligand pose (grey) with native ligand-binding pose (red) and best scoring ligand pose is shown in blue. (B) Ligand poses for 5HT2b receptor. The novel binding sites (grey) are close to the PDB-complex ligand binding site (red) and the best scoring pose (blue). (C) Ligand poses for chemokine receptor CXCR4. Different structures of CXCR4 are bound to different ligands in PDB. Native ligand-binding pose (red) and best scoring ligand pose is shown in blue. Other binding modes observed in PDB are represented by yellow, grey, and magenta colour ligand poses. PDB indicates Protein Data Bank.

	3.33x33	3.37x37	5.43-44x44	5.44-45x45	5.47x47	5.48x48	6.49x49	6.52x52	6.53x53	6.56x56	6.57x57	AA-ECL2		1.27x27	2.61x61	2.64x64	3.29x29	3.33x33	7.40-x39	ECL1	ECL2	ECL2	ECL2	ECL2	
CCR1	Y	L	L	N	G	L	W	Y	N	I	L	L	P	CCR1	-	W	Y	L	Y	E	W	H	T	C	S
CCR2	Y	Y	R	N	G	L	W	Y	N	I	L	L	P	CCR2	-	W	S	F	Y	E	W	V	V	C	G
CCR3	Y	L	M	T	C	L	W	Y	N	I	L	L	A	CCR3	-	W	Y	L	Y	E	W	E	L	C	S
CCR4	Y	F	I	N	G	L	W	Y	N	L	L	F	T	CCR4	-	W	Y	I	Y	E	W	H	Y	C	K
CCR5	Y	F	I	V	G	L	W	Y	N	L	L	S		CCR5	-	W	Y	L	Y	E	W	H	T	C	S
CCR6	Y	F	E	L	G	F	Q	H	N	L	L	P		CCR6	-	W	S	L	Y	E	W	S	V	C	E
CCR7	Y	F	Q	M	G	F	Q	Y	N	V	L	L		CCR7	-	W	S	I	Y	Y	W	E	A	C	S
CCR8	Y	F	M	N	G	L	W	F	N	L	L	S		CCR8	-	Q	Y	V	Y	E	W	V	Q	C	Y
CCR9	Y	F	K	V	G	F	Q	Y	N	L	L	M		CCR9	-	W	A	V	Y	Q	W	I	C	T	
CXR1	Y	F	Q	V	G	F	Q	Y	N	L	L	L		CXR1	-	A	G	I	Y	S	W	Q	R	C	
CXR2	K	F	P	H	G	F	W	Y	N	L	L	E		CXR2	-	W	S	V	K	E	W	S	P	C	
CXR3	F	F	Q	L	G	F	W	Y	H	V	L	Y		CXR3	F	W	D	A	F	S	W	L	A	C	
CXR4	Y	L	H	I	G	L	W	Y	I	S	R		CXR4	E	W	D	V	Y	E	W	R	I	C	D	
CXR5	H	F	Y	H	G	F	W	Y	H	I	F	F		CXR5	L	A	E	V	H	E	W	N	S	C	
CXR6	Y	F	Q	M	G	F	Q	Y	N	K	F	Y		CXR6	L	W	A	L	Y	E	W	L	-	C	
CX3CR1	F	F	T	N	G	F	W	Y	N	I	F	G		CX3CR1	-	W	Y	T	F	E	-	-	-	C	
XCR1	F	L	H	N	-	F	W	Y	N	L	F	Y		XCR1	-	W	P	L	F	R	W	-	-	C	
ACKR1	W	A	T	V	-	L	W	H	G	L	G	L		ACKR1	-	L	G	G	W	E	-	-	-	C	
ACKR2	Y	F	Q	N	G	F	W	Y	N	L	F	A		ACKR2	-	W	S	V	Y	E	-	G	W	C	
ACKR3	F	L	S	V	G	F	W	Y	H	V	L	S		ACKR3	-	W	S	T	F	Q	-	N	E	C	
ACKR4	Y	F	E	I	G	F	Q	Y	N	K	F	P		ACKR4	-	W	N	T	F	E	-	-	-	C	
CCRL2	Y	L	M	N	V	L	W	Y	N	F	F	F		CCRL2	-	W	A	L	L	Y	-	-	K	K	
	65% Y	65% F	22% Q	35% N	83% G	61% F	74% W	83% W	74% N	48% L	52% L	17% L			83% W	30% S	V/L 30%	65% Y	70% E	74% W	13% H	19% X	100% +17% S	17% L	

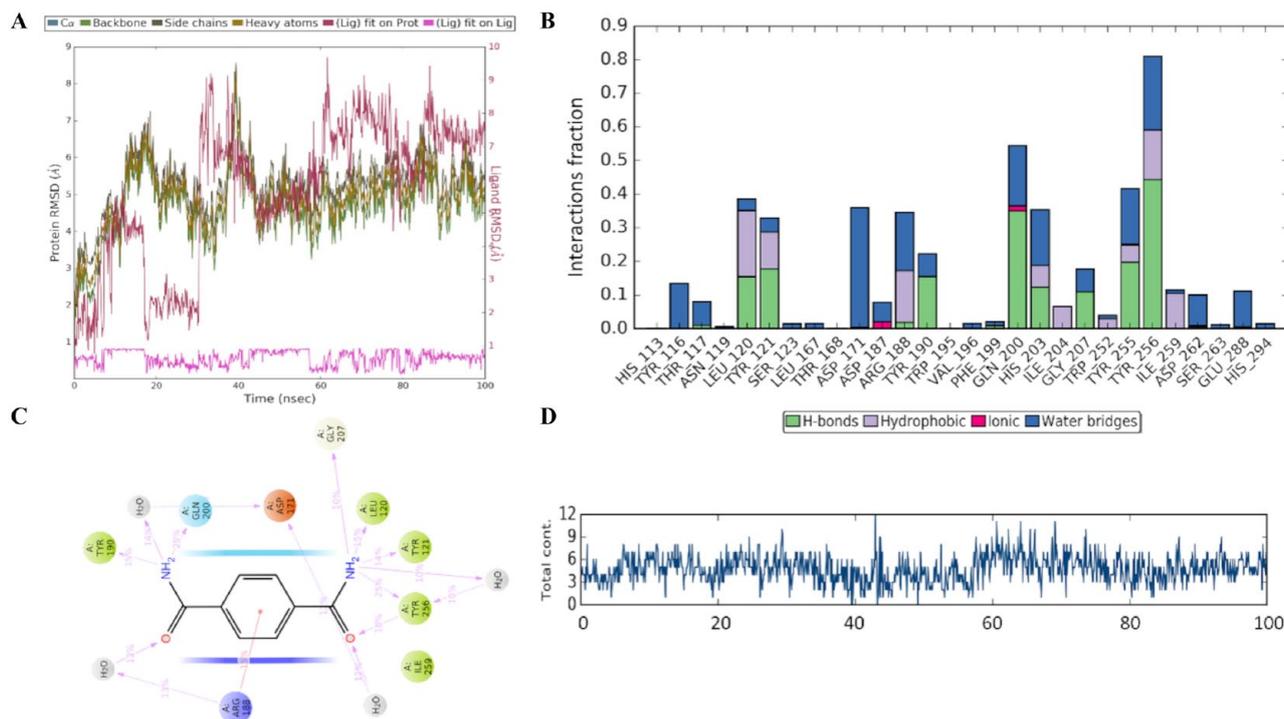
**Figure 4.** Sequence alignment of chemokine receptor subfamily. Yellow indicates the residues involve in binding with novel ligand pose. Grey residues interact with native ligand pose. Green indicates the common residues that are involved in interaction with novel pose as well as native pose.



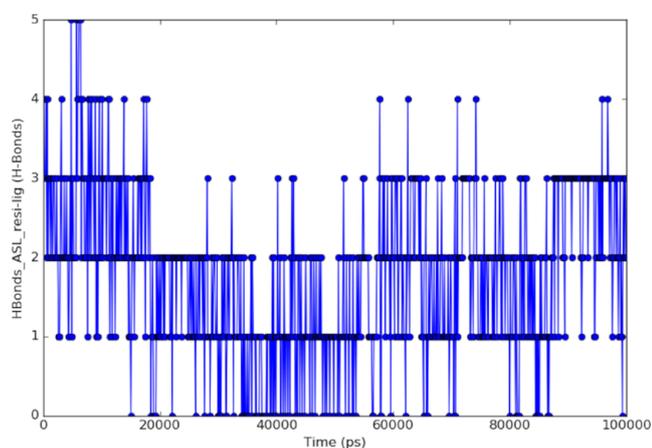
**Figure 5.** Bar graph indicating the percentage of each ligand conformation in either native site or allosteric site. MD indicates molecular dynamics; RMSD, root-mean-square deviation.

different protocols (Levels 1-3) play a role in modelling GPCR-ligand interaction. We found that accurate guidance of docking (smaller grid) is a crucial parameter and any

biochemical knowledge on ligand binding sites (as in Level 1 and 2) greatly improves the chances of reproducing native ligand pose using computational docking methods. We selected

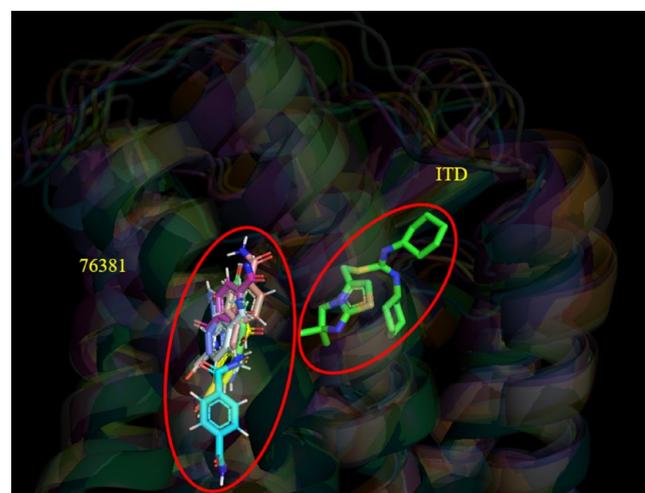


**Figure 6.** MD simulation results of CXCR4-76381 complex. (A) RMSD plot of CXCR4 and 76381. (B) Interacting residues and interaction types of CXCR4 with 76381 over the course of simulation time. Normalized stacked bars indicate the fraction of simulation time for which a particular type of interaction was maintained. Values more than 1.0 suggests that the residue forms multiple interactions of same subtype with ligand (C) interactions of 76381 atoms with residues of CXCR4 along with type and duration of interactions. Interactions that persist for more than 10% of simulation time have been shown. If some residues form multiple interaction of same type with the same atom of ligand, then interaction value can be more than 100% (D) total number of contacts (H-bonds, Water bridges, Hydrophobic, Ionic) between CXCR4 and ligand 76381 throughout the simulation. MD indicates molecular dynamics; RMSD, root-mean-square deviation.



**Figure 7.** H-bond interactions between 76381 ligand and CXCR4 (and water) throughout the simulation period. Simulation time is in picoseconds.

structural entries of GPCRs with alternate binding sites and a specialized database of ligands known to bind GPCRs, as documented in GLASS database, for an in-depth study. Structural analysis with smaller set of ligands against different proteins has been reported in the literature as well.<sup>27-32</sup> We performed docking to check the binding of ligands from GLASS database with GPCRs. We were able to obtain ligands that are reported as targets for particular GPCRs that bind better to the allosteric site as compared to the orthosteric binding site.



**Figure 8.** 76381 poses with CXCR4 compared to native ligand (ITD) of CXCR4 at different time points (in nanoseconds) during simulation. Native ligand pose (ITD) is shown as green-blue sticks. Cyan: 0 ns; Magenta: 33 ns; Yellow: 55 ns; Pink: 61 ns; White: 77 ns; Violet: 85 ns; Orange: 90 ns; Green: 100 ns. ITD, isothiouria derivative.

## Conclusion

Identification or computational design of highly specific drugs for GPCRs has been a problem for many decades. This is due to similar overall shape of ligand binding site, existence of sub-families, and broad sequence variation among GPCRs. Nonavailability of drugs, which is highly specific for one

GPCR subfamily member, gives rise to serious consequences and side effects to patients. Thus, design strategies to invest on allosteric sites acquire prime importance. Our computational approach employs reliable ligand binding software to recognize nonorthosteric ligand-binding sites which appear energetically favourable. Furthermore, our sequence analysis of such GPCRs, showed that residues at the allosteric site are not as conserved as in the orthosteric binding site, which provides hope for achieving the specificity of a particular GPCR. Furthermore, for certain GPCRs, some of their known cognate ligands were predicted to have better binding preference towards the allosteric site than orthosteric site. We conclude that few of the known GPCR-specific ligands may bind to the allosteric site and alter the function of the GPCRs. Therefore, this computational approach for ligand binding can be used for prediction of allosteric binders for GPCRs in general.

### Acknowledgements

The authors would like to acknowledge Prof. N. Srinivasan of Indian Institute of Science for useful discussions.

### Author Contributions

K.H., S.J., and V.T. carried out all the work and analyses. R.S. conceptualized the work and both R.S. and S.V. participated in discussions. K.H. and V.T. wrote first draft of the manuscript and S.J., S.V., and R.S. improved it.

### ORCID iD

Ramanathan Sowdhamini  <https://orcid.org/0000-0002-6642-2367>

### SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

### REFERENCES

- Ahram M, Litou ZI, Fang R, Al-Tawallbeh G. Estimation of membrane proteins in the human proteome. *In Silico Biol.* 2006;6:379-386.
- Almén MS, Nordström KJV, Fredriksson R, Schiöth HB. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol.* 2009;7:50. doi:10.1186/1741-7007-7-50.
- Rosenbaum DM, Rasmussen SGF, Kobilka BK. The structure and function of G-protein-coupled receptors. *Nature.* 2009;459:356-363. doi:10.1038/nature08144.
- Nagarathnam B, Kannan S, Dharnidharka V, Balakrishnan V, Archunan G, Sowdhamini R. Insights from the analysis of conserved motifs and permitted amino acid exchanges in the human, the fly and the worm GPCR clusters. *Bioinform.* 2011;7:15-20. doi:10.6026/97320630007015.
- Venkatakrishnan AJ, Flock T, Prado DE, Oates ME, Gough J, Madan Babu M. Structured and disordered facets of the GPCR fold. *Curr Opin Struct Biol.* 2014;27:129-137. doi:10.1016/j.sbi.2014.08.002.
- Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol.* 2003;63:1256-1272. doi:10.1124/mol.63.6.1256.
- Zhang D, Zhao Q, Wu B. Structural studies of G protein-coupled receptors. *Mol Cells.* 2015;38:836-842. doi:10.14348/molcells.2015.0263.
- Schwartz TW, Frimurer TM, Holst B, Rosenkilde MM, Elling CE. Molecular mechanism of 7TM receptor activation – a global toggle switch model. *Annu Rev Pharmacol Toxicol.* 2006;46:481-519. doi:10.1146/annurev.pharmtox.46.120604.141218.
- Chan W, Zhang Y. Virtual screening of human Class-A GPCRs using ligand profiles built on multiple ligand-receptor interactions. *J Mol Biol.* 2020;432:4872-4890.
- Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2018;46:D1074-D1082. doi:10.1093/nar/gkx1037.
- Sriram K, Insel PA. G protein-coupled receptors as targets for approved drugs: how many targets and how many drugs? *Mol Pharmacol.* 2018;93:251-258. doi:10.1124/mol.117.111062.
- Heng BC, Aubel D, Fussenegger M. An overview of the diverse roles of G-protein coupled receptors (GPCRs) in the pathophysiology of various human diseases. *Biotechnol Adv.* 2013;31:1676-1694. doi:10.1016/j.biotechadv.2013.08.017.
- Stoy H, Gurevich V. How genetic errors in GPCRs affect their function: possible therapeutic strategies. *Genes Dis.* 2015;2:108-132. doi:10.1016/j.gendis.2015.02.005.
- Conn PJ, Christopoulos A, Lindsley CW. Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. *Nat Rev Drug Discov.* 2009;8:41-54. doi:10.1038/nrd2760.
- Lindsley CW, Emmitte KA, Hopkins CR, et al. Practical strategies and concepts in GPCR allosteric modulator discovery: recent advances with metabotropic glutamate receptors. *Chem Rev.* 2016;116:6707-6741. doi:10.1021/acs.chemrev.5b00656.
- Azam S, Haque ME, Jakaria M, Jo S-H, Kim I-S, Choi D-K. G-protein-coupled receptors in CNS: a potential therapeutic target for intervention in neurodegenerative disorders and associated cognitive deficits. *Cells.* 2020;9:506. doi:10.3390/cells9020506.
- Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem.* 2009;30:2785-2791. doi:10.1002/jcc.21256.
- Chan WKB, Zhang H, Yang J, et al. GLASS: a comprehensive database for experimentally validated GPCR-ligand associations. *Bioinformatics.* 2015;31:3035-3042. doi:10.1093/bioinformatics/btv302.
- Berman H, Henrick K, Nakamura H. Announcing the worldwide Protein Data Bank. *Nat Struct Biol.* 2003;10:980. doi:10.1038/nsb1203-980.
- Pándy-Szekeres G, Munk C, Tsonkov TM, et al. GPCRdb in 2018: adding GPCR structure models and ligands. *Nucleic Acids Res.* 2018;46:D440-D446. doi:10.1093/nar/gkx1109.
- Kim S, Chen J, Cheng T, et al. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* 2019;47:D1102-D1109. doi:10.1093/nar/gky1033.
- Backman TWH, Cao Y, Girke T. ChemMine tools: an online service for analyzing and clustering small molecules. *Nucleic Acids Res.* 2011;39:W486-W491. doi:10.1093/nar/gkr320.
- O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: an open chemical toolbox. *J Cheminform.* 2011;3:33. doi:10.1186/1758-2946-3-33.
- Bowers KJ, Chow E, Xu H, et al. Scalable algorithms for molecular dynamics simulations on commodity clusters. Paper presented at: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, SC'06; November 11-17, 2006; Tampa, FL. doi:10.1145/1188455.1188544.
- Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des.* 2013;27:221-234. doi:10.1007/s10822-013-9644-8.
- Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* 2012;40:D370-D376. doi:10.1093/nar/gkr703.
- Mothay D, Ramesh KV. Binding site analysis of potential protease inhibitors of COVID-19 using AutoDock. *Virusdisse.* 2020;31:1-6. doi:10.1007/s13337-020-00585-z.
- Sultan S, Singh GKS, Ashraf K, Ashraf M. Molecular docking studies of enzyme inhibitors and cytotoxic chemical entities. In: Vlachakis D, ed. *Molecular Docking.* London, England: IntechOpen; 2018:13-30. doi:10.5772/intechopen.76891.
- Samant M, Chadha N, Tiwari AK, Hasija Y. In silico designing and analysis of inhibitors against target protein identified through host-pathogen protein interactions in malaria. *Int J Med Chem.* 2016;2016:2741038. doi:10.1155/2016/2741038.
- Gagic Z, Ruzic D, Djokovic N, Djikic T, Nikolic K. In silico methods for design of kinase inhibitors as anticancer drugs. *Front Chem.* 2020;7:873. doi:10.3389/fchem.2019.00873.
- Azam SS, Abbasi SW. Molecular docking studies for the identification of novel melatonergic inhibitors for acetylserotonin-O-methyltransferase using differential docking routines. *Theor Biol Med Model.* 2013;10:63. doi:10.1186/1742-4682-10-63.
- Haouz A, Vanheusden V, Munier-Lehmann H, et al. Enzymatic and structural analysis of inhibitors designed against Mycobacterium tuberculosis thymidylate kinase: new insights into the phosphoryl transfer mechanism. *J Biol Chem.* 2003;278:4963-4971. doi:10.1074/jbc.M209630200.