Homology Modeling and Docking Evaluation of Aminergic G Protein-Coupled Receptors

Fiona M. McRobb, Ben Capuano, Ian T. Crosby, David K. Chalmers,* and Elizabeth Yuriev*

Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Monash University (Parkville Campus), 381 Royal Parade, Parkville, VIC 3052 Australia

Received November 15, 2009

We report the development of homology models of dopamine (D_2, D_3, and D_4), serotonin (5-HT_{1B}, 5-HT_{2A}, 5-HT_{2B}, and 5-HT_{2C}), histamine (H_1), and muscarinic (M_1) receptors, based on the high-resolution structure of the β_2-adrenergic receptor. The homology models were built and refined using Prime. We have addressed the required modeling of extracellular loop 2, which is often implicated in ligand binding. The orthosteric sites of the models were optimized using induced fit docking, to allow for side-chain flexibility, and the resulting receptor models have been evaluated using protein validation tools. Of the nine homology models developed, six models showed moderate to good enrichment in virtual screening experiments (5-HT_{2A}, 5-HT_{1B}, D_2, 5-HT_{2C}, D_3, and M_1). The 5-HT_{2A} receptor displayed the highest enrichment in virtual screening experiments with enrichment factors of 6.1, 6.9, and 5.9 at 2, 5, and 10%, respectively, of the screened database. However, three of the models require further refinement (5-HT_{2B}, D_4, and H_1), due to difficulties in modeling some of the binding site residues as well as the extracellular loop 2. Our effort also aims to supplement the limited number of tested G protein-coupled receptor homology models based on the β_2 crystal structure that are freely available to the research community.

INTRODUCTION

There is an ongoing need for improved treatments for the disease states of schizophrenia, Parkinson's disease, and obesity, as many of the existing treatments cause undesirable side effects. For example, existing antipsychotic agents treat symptoms by acting at the dopamine D_2 and the serotonin 5-HT_{2A} receptors, however, undesired affinity at the serotonin 5-HT_{2C} and the histamine H_1 receptors has been implicated in weight gain.[1] We aim to use a structure-based approach to assist in the design of compounds that are potent at particular receptor subtypes, with reduced affinity for receptors believed to be responsible for undesirable effects. To aid our design process, we require receptor models (crystal structures or homology models), which can be used to predict the bound conformation and affinity of G protein-coupled receptor (GPCR) ligands.

Until late 2007, the only high-resolution structures of GPCRs available to inform the design of new drugs were of bovine rhodopsin,[2−4] and GPCR homology models were necessarily based on this template.[5−13] Recently, there has been a steady stream of GPCR crystal structures reported (which we hope will continue and grow). Structures of the human β_2-adrenergic receptor (β_2),[14−17] squid rhodopsin,[18,19] turkey β_1-adrenergic receptor (β_1),[20] bovine opsin,[21] bovine opsin bound to the carboxyl terminus of the Gα-subunit,[22] and human adenosine A_{2A} receptor (A_{2A})[23] have been solved, giving a wealth of structural information and revealing important structural differences between rhodopsin and other class A GPCRs, particularly in the orientation and positions of the transmembrane helices (TM) and in the structure of the loop regions.[16,24,25] The key features of these crystal structures are detailed in the Supporting Information (Table S1). The liganded structures have confirmed that the orthosteric binding site is in a similar position to that of the *cis*-retinal site in rhodopsin,[16,25] but compared to rhodopsin, the binding sites of β_1, β_2, and A_{2A} are large and open to the extracellular space. They also reveal a significantly different conformation of extracellular loop (ECL) 2, including a previously unseen helix and, in the β_1 and β_2 receptors, a second disulfide bond in ECL2, in addition to the previously identified conserved disulfide bond between ECL2 and TM3.[16,20] The recently crystallized receptors have higher homology with most other class A GPCRs than with rhodopsin and are better templates for the development of homology models (Table S2, Supporting Information).[26,27]

Since the publication of the β_2 receptor structure, a number of new GPCR homology models based on this template have been reported. A summary of the models listing their templates, modeling programs, and evaluation methods is given in Table 1. A number of automatically generated models are also present in online databases including the Protein Model Portal (http://proteinmodelportal.org)[45] and ModBase (http://modbase.compbio.ucsf.edu).[46] In developing their models, researchers have used a wide variety of alignment tools and homology modeling software including: MODELLER,[47] Sybyl,[48] Prime,[49] and ICM.[50] In most cases, these models have been built based on the 2RH1 crystal structure, but some used a combination of templates[35] (e.g., 2RH1 and 3EML). Accurate prediction of the loops, particularly those surrounding the orthosteric binding site, remains one of the more difficult aspects in GPCR homology modeling, as highlighted in the recent GPCR Dock 2008 experiment.[24] It is important to model these loops, particularly ECL2, as they are often implicated directly or indirectly

* Corresponding authors. Telephone: +61 3 9903 9110 (D.K.C.) and +61 3 9903 9611 (E.Y.). Fax: +61 3 9903 9582 (D.K.C.) and +61 3 9903 9582 (E.Y.). E-mail: David.Chalmers@pharm.monash.edu.au. (D.K.C.) and Elizabeth.Yuriev@pharm.monash.edu.au (E.Y.).

10.1021/ci900444q
Published on Web 03/01/2010

Table 1. Summary of Recently Reported GPCR Homology Models Based on β_2 Receptor Structure

receptor(s)	template	modeling programs	ECL2 modeling method	evaluation	docking program	ref
A_1 and A_{2A}	2RH1	CAChe	disulfide constraint	docking	CAChe	28
5-HT_{1A}, 5-HT_{2A}, 5-HT_{2B}, 5-HT_{2C}, 5-HT_6, 5-HT_7, α_1, α_2, D_2, D_3, D_4, H_1, M_1, M_4	2RH1	MODELLER	disulfide constraint	docking	GOLD	29
5-HT_4	2RH1	–	–	–	–	30
MC_4	2RH1	QUANTA	–	–	QUANTA	31
H_4	2RH1	MOE	disulfide and other residue constraints	–	–	32
H_1	2RH1	SYBYL	derived from 2RH1 template (partial)	–	Manual	33
NK_1	2RH1	MODELLER	derived from 2RH1 template	virtual screening	Glide	34
CCK_1	2RH1, 3EML	ICM	–	docking	ICM	35
A_{2A}	2RH1	MOE	derived from 2RH1 template	docking	IFD, Glide MOE	36
CCR_5	2RH1	InsightII	derived from 2RH1 template	docking	GOLD	37
secretin	2RH1	ICM	–	docking	ICM	38
D_1	2R4R	Swiss-model/Deep view	–	docking	GRAMM	39
M_2	2RH1	ICM	derived from 2RH1 template (partial)	docking	ICM	40
H_4	2RH1	MODELLER	–	docking	–	41
$P2Y_{14}$	3EML	Prime	derived from 3EML template (partial)	docking	Glide	42
D_2, D_3, D_4	2RH1	MODELLER	derived from 2RH1 template (partial)	docking	AUTODOCK	43
5-HT_{2A}, H_1	2RH1	MODELLER	derived from 2RH1 template	docking	GOLD	44

in ligand binding.[51] However, as loop length increases, current methods struggle, and loops over six amino acids are generally not modeled well.[52] If the reported models contained loops, then they were generated using either template-based or de novo modeling approaches (Table 1). Often a combination of these two methods was used, with the crystal structure as a template for all loops except ECL2, due to its significance in relation to the orthosteric site. A number of different techniques have been used to generate the conformation of ECL2 including modeling completely[28,34,37] or partially[33,36,40,42] based on the template or building with constraints on certain residues in the template[29,32] (usually the conserved cysteine in ECL2). Several models were further optimized using molecular dynamics, generally in a lipid bilayer.[28−30,33,37,41]

GPCR homology modeling is essentially an uncertain process, and it is important that the quality of models is evaluated. Potential causes of error, in addition to fundamental differences in structure between the template and target, include incorrect alignment and difficulties in modeling the variable loop regions as well as the refinement procedures utilized. Often the capacity to evaluate models is limited, but several approaches can be used to assess the quality of homology models. The stereochemical quality of models can be evaluated by assessment tools, such as PROCHECK,[53] WHAT_CHECK,[54] and MolProbity.[55] Site-directed mutagenesis data indicates residues that are implicated in ligand binding. Thus, by relating models and docking results to site-directed mutagenesis data, an inference can be made whether the model is a realistic representation of ligand binding. Finally, the ability of the binding site to correctly dock ligands can be assessed by docking known active compounds using programs, such as GOLD,[56] Glide,[57] or ICM.[50]

It is a significant problem for the scientific community that many reported modeling studies do not evaluate (or do not report evaluation) of published homology models. Of the studies reported in Table 1, a significant fraction do not report any evaluation. Furthermore, only a limited number of these models are freely available for use and comparison, and in many cases, the modeling procedures are described without sufficient detail to be reproducible. Our effort also aims to supplement the limited number of evaluated GPCR homology models, based on the β_2 crystal structure, that are freely available to the research community.

Specifically, to further investigate the receptors implicated in the disease states of interest to us and those implicated in drug side effects, we have built a series of nine homology models using the recent crystal structure of the β_2-adrenergic receptor (2RH1).[16] These models have been refined using experimental data and evaluated by docking and virtual screening studies. Additionally, to validate our homology modeling procedure, a β_2 model was built using the A_{2A} crystal structure as a template, and an A_{2A} model was built using the β_2 template.

EXPERIMENTAL SECTION

GPCR residues are identified using the Ballesteros−Weinstein nomenclature,[58] except for loop regions where crystal structure numbering is used. Molecular modeling was performed principally using the Schrödinger software suite. Homology models were built with Prime 1.6[49] and manually refined in Maestro 8.0.[59] Ligand molecules were prepared using LigPrep,[60] and docking was carried out with Glide 5.0.[57] Default settings were used, unless stated otherwise.

Cognate Ligand Docking. Ligands, extracted from the crystal structures of respective complexes and processed using LigPrep (in order to change their conformation), were redocked into β_2 and the A_{2A} crystal structures using the Glide program in the standard precision (SP) and extra precision (XP) modes and the induced fit docking (IFD) protocol.[61] The docking site was defined as a box 28 × 28 × 28 Å, identified by the centroid of the cocrystallized ligand. When docking into the A_{2A} crystal structure in the presence of crystallized water molecules, those close to the ligand were retained (residue numbers 505, 519, 550, 559, 565, 567, 573, and 576).

Homology Modeling. Homology models were built using the β_2 (2RH1) crystal structure[14,16] as the template. The T4-lysozyme inserted into the β_2 receptor between Gln 231 and Ser 262 to assist crystallization was removed. As the T4-lysozyme replaced ICL3, which is distant from the binding site, ICL3 was not modeled in the receptors. When using the A_{2A} crystal structure as a template, the T4-lysozyme between residues Leu 208 and Arg 222 was similarly removed.

The sequences of the human dopamine, serotonin, α- and β-adrenergic, adenosine, histamine, muscarinic, and bovine rhodopsin receptors were obtained from the Universal Protein

Resource (http://www.uniprot.org/) and aligned using ClustalW.[62] The structure alignment was manually adjusted to remove gaps in helices. Highly conserved residues in each TM were anchored, and the models were generated in Prime. The loops were refined using the refine loops tool with the extended high-loop refinement procedure. Since ECL2 is a long loop (between 13 and 33 residues), it was refined separately, in four steps, refining 6−8 residues in each step (Table S3, Supporting Information). Homology models were inspected to ensure that the side chains of the conserved residues were aligned to the template. Side chains were manually adjusted using the rotamers tool, if required. MolProbity[55] and PROCHECK[53] were used to assess the quality of the models as well as the protein report tool. MolProbity was also used to assign protonation states for the receptors and to optimize the placement of hydrogen atoms.

Comparison of Models and Crystal Structures. The A_{2A} and β_2 homology models were compared to their corresponding crystal structures by calculating the root-mean-square deviation (rmsd) of the α-carbons in Sybyl.[48] Only the residues present in the homology models were compared to the crystal structures. Rmsd values were also calculated for heavy atoms within 6 Å of the cocrystallized ligand (β_2 residues: 82, 86, 109, 110, 113, 114, 115, 117, 118, 193, 195, 199, 200, 203, 204, 207, 208, 286, 289, 290, 293, 308, 312, and 316 and A_{2A} residues: 66, 84, 85, 88, 167, 168, 169, 174, 177, 181, 246, 249, 250, 253, 264, 265, 267, 270, 271, and 274, crystal structure numbering).

Binding Site Refinement. The side-chain positions of the ligand binding site residues in each model were refined by docking an appropriate ligand into the site using IFD, which allows receptor flexibility. The docking site was defined as a box 28 × 28 × 28 Å centered on Asp 3.32, Trp 6.48, Phe 6.52, and Tyr 7.43 for all aminergic GPCRs. Up to 50 poses per ligand were collected. For initial docking, the van der Waals (vdW) radii and the partial atomic charges of both the ligand and the receptor were scaled to 0.5 to reduce the effect of steric clashes and to "soften" the surfaces of both the protein and the ligand so that a wider variety of poses could be generated. Prime was then used to optimize residues within 5 Å of the ligand poses. Finally, ligands were redocked back into all new receptor conformations, using the default vdW radii and the charge scaling (1.0 receptor, 0.8 ligand). The final models were selected after multiple iterations of model construction and refinement.

Enrichment Studies. A set of reported antagonists active at each receptor was obtained from the GLIDA database (http://pharminfo.pharm.kyoto-u.ac.jp/services/glida/index.php).[63] Active compounds (Table S4, Supporting Information) were built in ChemDraw and converted to three-dimensional (3D) using LigPrep, which also assigned formal charges according to physiological pH (pH 7.4). A set of 1000 drug-like decoy compounds with an average molecular weight of 360 g/mol was obtained from Schrödinger (http://www.schrodinger.com).[64] The properties of the ligands were assessed using QikProp[65] (Table S5, Supporting Information). Using the previously prepared 3D ligand structures, the average shape Tanimoto score was measured using ROCS,[66] taking the ligand from flexible receptor docking as a reference. The 2D Tanimoto score was calculated using UNITY in Sybyl. The decoy set, enriched with respective active compounds, was docked into each receptor structure using Glide XP and ranked by GlideScore. The docking site was defined as a box 28 × 28 × 28 Å, and the center of the binding site was identified using the coordinates of the center of carazolol in β_2 (or ZM-241,385 for docking into the A_{2A} receptor). One pose per ligand was collected. Enrichment factors were calculated at 2, 5, and 10% of the total database screened, using the following equation:[67]

$$EF^{x\%} = (Hits_{sampled}/N_{sampled}) \div (Hits_{total}/N_{total})$$

Table 2. Rmsd Values for Cognate Ligand Docking into the β_2 and A_{2A} Crystal Structures

	rmsd to crystal structure (Å)		
docking method	β_2	A_{2A}[a]	A_{2A}[b]
SP	0.85	9.83	0.79
XP	0.54	8.43	1.09
IFD	1.13	2.67	3.01

[a] Without crystallographic water. [b] With eight crystallographic water molecules.

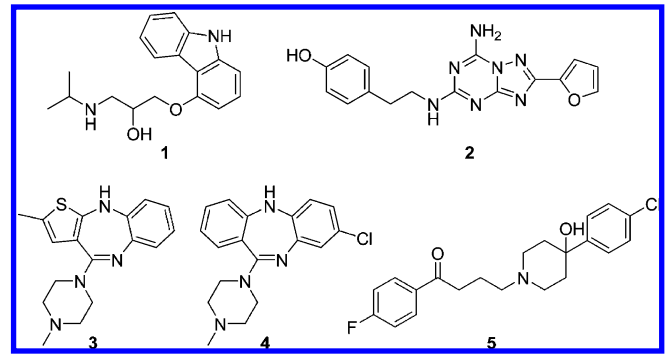

Figure 1. Structures of carazolol (**1**), ZM-241,385 (**2**), olanzapine (**3**), clozapine (**4**), and haloperidol (**5**).

RESULTS AND DISCUSSION

Cognate Ligand Docking. In a preliminary step, we investigated the ability of Glide to reproduce the bound poses of carazolol (**1**) in the β_2 and ZM-241,385 (**2**) in the A_{2A} crystal structures. We used the Glide SP and Glide XP methods, both of which use a rigid receptor, and IFD, which allows conformational changes within the receptor. The results are shown in Table 2.

Docking into the β_2 structure gave an rmsd of less than 2 Å in each case. In cognate ligand docking, an rmsd between a crystal structure and a docked pose of less than 2 Å is generally considered as a good result.[68] The A_{2A} crystal structure resolution is lower than that of the β_2 crystal structure, and the A_{2A} receptor does not make a strong ionic interaction with the ligand, such as the salt bridge observed in the aminergic GPCRs. Thus, docking into the A_{2A} crystal structure proved more challenging. Cognate docking into the A_{2A} receptor without crystallographic water molecules being present failed to reproduce the binding mode accurately with rmsd values all above 2 Å (Table 2). Only IFD was able to place the ligand in a similar conformation to the crystal structure, however, this caused a small conformational change to the binding site (rmsd 1.09 Å, residues within 6 Å of ligand). Due to the presence of water molecules interacting with the ligand in the crystal structure, we also docked into the A_{2A} receptor in the presence of eight

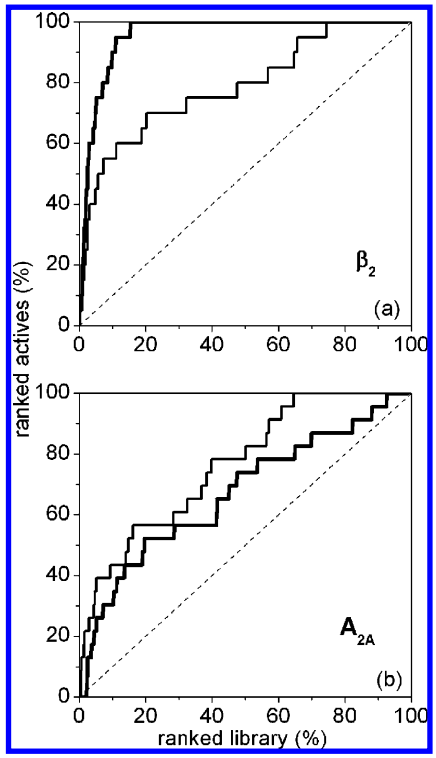

Figure 2. Enrichment plots for the (a) β_2 and (b) A_{2A} crystal structures. Bold line indicates virtual screening using Glide XP; Fine line indicates virtual screening using Glide SP; Dotted line indicates random.

crystallographic water molecules. This docking gave substantially better results for the SP and XP docking modes, with rmsd values less than 2 Å (Table 2). The furan and adenine rings superimposed well with the crystal structure, but as the substituted phenol was at the entrance of the binding site, its position proved more difficult to predict (Figure S1, Supporting Information). The results obtained with the crystallographic water molecules present indicate their importance in mediating protein−ligand interactions in the A_{2A} receptor, and their importance in being maintained for further docking studies.[69]

Virtual Screening Using the β_2 and A_{2A} Crystal Structures. To assess the ability of Glide to identify known GPCR antagonists in a library of drug-like decoys, we performed virtual screening experiments using the β_2 and A_{2A} crystal structures. To be useful in the prediction of active compounds, a docking program must be able to rank the known compounds higher than the decoy compounds. Accordingly, we docked sets of 20 (β_2) and 23 (A_{2A}) antagonists (Table S4, Supporting Information) and 1000 decoy compounds into each crystal structure. Figure 2 shows cumulative plots of the active compounds recovered against the library ranked by Glide Score.

Table 3 lists the enrichment factors and the maximum possible enrichment factor in each case. The maximum values are dependent on the total number of active compounds in each library. At the β_2 receptor, SP docking performed well, with enrichment factors of 12.1, 9.0, and 5.5 at 2, 5, and 10%, respectively, but the more computationally expensive XP docking was clearly superior with enrichment factors of 19.4, 14.0, and 9.0. XP docking also identified 95% of the active compounds in the decoy library within the top 20% of the library and placed all the ligands,

Table 3. Enrichment Factors for Virtual Screening with the β_2 and A_{2A} Crystal Structures (at *x*% of the Ranked Database Screened)

enrichment factor, %	β_2[a]		A_{2A}[b]	
	SP	XP	SP	XP
2	12.1	19.4	10.6	0.0
5	9.0	14.0	6.8	4.3
10	5.5	9.0	4.3	3.0

[a] Maximum enrichment factors for β_2: $EF^{2\%} = 48.6$, $EF^{5\%} = 20$, and $EF^{10\%} = 10$. [b] Maximum enrichment factors for A_{2A}: $EF^{2\%} = 44.5$, $EF^{5\%} = 19.7$, and $EF^{10\%} = 9.9$.

with the exception of the lowest ranked one, in similar conformations to the crystal structure of carazolol. This is a good result because not only did the Glide XP identify active compounds over the decoy compounds but also the docked poses conformed to experimental data. In one case, we were able to compare timolol in the 3D4S crystal structure to the docked pose of timolol obtained during virtual screening. An rmsd of 0.61 Å was obtained, indicating that our docking method was able to predict the binding mode of timolol very well.

In virtual screening, it is often found that large ligands rank higher due to a larger number of interactions with a target. This presents a challenge for the scoring function in order to identify lower molecular weight actives. In screening against the β_2 crystal structure, the active compounds had a lower molecular weight (300 g/mol) compared to that of the decoy ligands (360 g/mol) (Table S5, Supporting Information). Therefore, the enrichment factors obtained in this exercise are very encouraging and increase our confidence in the employed virtual screening procedure.

In the virtual screen against the A_{2A} receptor, the SP docking method performed moderately better than the XP throughout the virtual screening, with enrichment factors of 10.6, 6.8, and 4.3 at 2, 5, and 10%, respectively (Table 3). While enrichment was observed, overall the results were worse than those observed with the β_2 crystal structure. This is likely to be a result of both the different binding mode and site of the A_{2A} receptor in comparison to the β_2 receptor, as the adenosine ligand binds closer to the extracellular side of the receptor and does not make a strong ionic interaction with the receptor that is commonly observed in aminergic GPCRs. From the diverse range of ligand poses obtained from virtual screening, it is evident that the docking method in this instance may not be suitable for adenosine ligands or that, at least, it requires further optimization.

While we cannot directly compare our virtual screening results to others in the literature, due to different docking methods and decoy libraries, our virtual screening experiments with β_2 gave similar results to those obtained by de Graff et al. using Surflex and are substantially better than their GOLD results.[70] The results obtained were also comparable to the average enrichment factors obtained by Halgren et al. using Glide for virtual screening and using a number of different crystal structures.[71]

Development of GPCR Homology Models. We have used the 2RH1 crystal structure[14,16] of the β_2 GPCR as the template for homology modeling of the pharmaceutically significant D_2, D_3, D_4, 5-HT_{1B}, 5-HT_{2A}, 5-HT_{2B}, 5-HT_{2C}, M_1, H_1, and A_{2A} receptors. Of the currently available GPCR

crystal structures, 2RH1 has been obtained at the highest resolution (2.4 Å) and contains structural information for the extracellular side of the receptor, near the binding site. The β_1 structure would be a viable template, but due to the high similarity of the β_1 and β_2 crystal structures and the β_2 structure having the highest resolution, we did not use the β_1 structure as a template. We did not use the A_{2A} structure as a template because purinergic ligands bind to the A_{2A} receptor in a different orientation in the orthosteric site, compared to the adrenergic ligands, and adenosine receptors do not have the conserved Asp 3.32 residue, which is present in aminergic GPCRs. The A_{2A} structure was used for validation of the homology modeling procedure by building a model of β_2 based on this template.

Generating models for proteins with less than 30% overall homology to the template often means that the alignment can be unreliable.[72] However, GPCRs are a unique case, as the low sequence identity is compensated for by a high structural similarity, namely the 7 TM helices. Therefore, the key GPCR residues can be aligned to generate good quality homology models, particularly within the TM region.[73] The β_2 template shares a higher sequence identity with the receptors of interest, particularly the aminergic GPCRs, indicating that it is a better template than rhodopsin (Table S2, Supporting Information). The complete alignment for all homology models developed is provided in the Supporting Information.

To validate our homology modeling method, we built a model of the A_{2A} receptor based on the β_2 crystal structure and a model of the β_2 receptor based on the A_{2A} crystal structure. The rmsd (for the α-carbons) of the homology model to the crystal structure was 4.5 Å for the β_2 receptor, and the rmsd for the model and crystal structure for the A_{2A} receptor was 3.3 Å. The rmsd of binding site residues, within 6 Å of the ligand, for the β_2 model and the crystal structure was 3.0 Å and for the A_{2A} model and crystal structure was 3.1 Å.

The stereochemical quality of the models was evaluated using PROCHECK and MolProbity, with over 90% of backbone dihedral angles residing in the favored regions for all receptors. Any deviations in the Ramachandran plots[74,75] were investigated. The majority of deviations resided in the loop regions or away from the orthosteric site, and these deviations were left unchanged.

The significance of the ECLs in ligand binding, either directly or indirectly, is becoming increasingly apparent, from site-directed mutagenesis (e.g., Ile 184 in D_2),[76] docking, and most importantly, the recent crystal structures. In the β_2 receptor, the residues between the conserved Cys 191 in ECL2 and the top of TM5 play an essential role in shaping the binding site, maintaining the structure of the loop and interacting with the ligand (e.g., Phe 193).[14] Therefore, loop modeling of ECL2, in particular, is crucial.

Our homology models contain intracellular and extracellular loops based on the template structure (excluding ICL3 and termini), although we acknowledge that these regions have low reliability. In different receptors, these regions vary significantly in sequence, structure, and length (particularly for ECL2, which has between 11−33 amino acids). Therefore, we optimized the loop regions using the loop refinement tool in Prime, which is reported to work well (but becomes less reliable as the loop length increases).[52] The loop refinement procedure was able to successfully generate conformations for most loops. However, for a few models (5-HT_{1B}, 5-HT_{2C}, and D_2), we had to resort to the initial loop conformations for the longer loops, namely ECL2 and/or ECL3, as the loops obstructed the binding site.

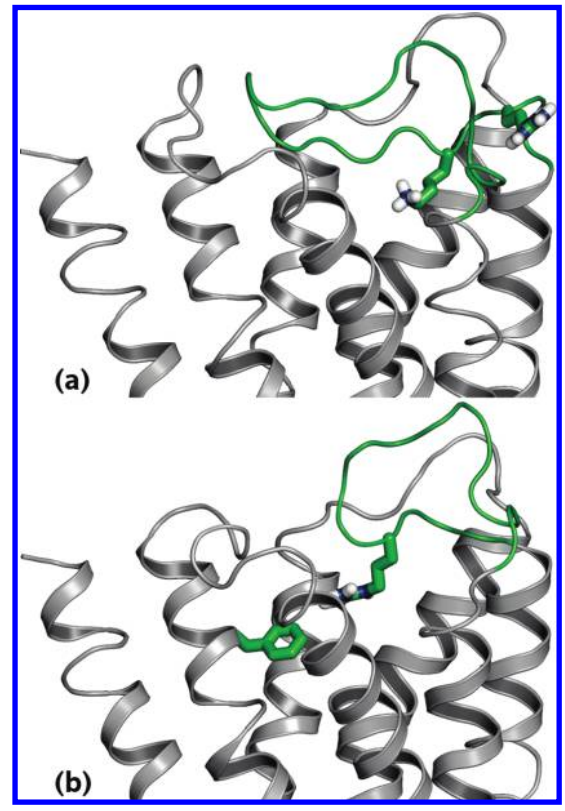

Figure 3. Problematic residues: (a) ECL2 residues Arg 211 and Lys 213 in the 5-HT_{2B} model and (b) Phe 2.61 and Arg 186 in the D_4 model (ECL2 highlighted in green). This image was created using PyMOL.[78]

The residues in ECL2 directly above the orthosteric binding site (between the cysteine in the conserved disulfide bond and the TM5 (residue 5.37)) play a significant role in defining the shape of the binding site. In the H_1, M_1, and 5-HT_{2B} receptors, this portion of the loop is longer than the template by two residues. In these receptors, it is likely that the architecture of ECL2 is different to that of the β_2 template, possibly with a helix directly above the binding site (between the disulfide bond and the TM5), or that TM5 is actually longer. We came to this conclusion because, in the case of 5-HT_{2B}, the extra residues include Arg 211 and Lys 213 (Figure 3a), which commonly make significant interactions with solvent and indicate the end of a helix in GPCRs.[77] In our models, we were unable to resolve this matter and have retained the length of TM5 from the template. In the H_1 and M_1 receptors, no such residues are present to indicate the potential start of TM5, so we again retained the length of TM5 from the β_2 crystal structure.

The conserved disulfide linkage between TM3 and ECL2 (e.g., from Cys 3.25 and Cys 191 in β_2) greatly restricts the position of ECL2. Difficulties were experienced when modeling ECL2 for the D_4 receptor, as the residue adjacent to the conserved Cys 185 was Arg 186 (Figure 3b), which protruded into the binding site and could interact with ligands in docking studies, altering the binding modes. Other problems encountered when modeling D_4 were with the

residue at position 2.61, which may be involved in subtype selectivity.[43] In the template, position 2.61 is a glycine, whereas it is a phenylalanine in D_4, which partially obstructs the entrance to the binding site. This restricts the size of the binding site, and there is no template to aid in the placement of this side chain. Similar situations were encountered with the H_1 and M_1 receptors, with large bulky residues in the models replacing smaller amino acids in the template. Further work is required to optimize these models.

Optimization of Ligand Binding Sites. The shapes of the binding sites in the homology models are likely to be influenced by the presence of carazolol in the β_2 crystal structure used as a template. We have used flexible receptor docking to moderate this bias and to also increase the volume of the binding pocket. In each receptor model, a known antagonist (Table S2, Supporting Information) was docked by IFD, generating multiple ligand−receptor structures. In each case, the shape and size of the binding site was visually inspected using surface rendering, ensuring that binding site residues were not obstructing the binding site. A single model complex was selected from the IFD structures on the basis of both the position of the ligand in the binding site and contacts observed between the ligand and the receptor and the comparison with site-directed mutagenesis. The models were also assessed using MolProbity to ensure that the flexible receptor docking did not introduce structural problems, such as steric clashes.

Our flexible receptor docking approach can be exemplified using the D_2 and 5-HT_{2A} models; upon docking olanzapine (**3**) into the D_2 homology model, several key interactions were observed that were consistent with site-directed mutagenesis data.[76,79−89] The ionic interaction between the protonated nitrogen of the ligand and the conserved Asp 3.32 (2.8 Å) and the hydrogen bonding between the secondary amine on olanzapine and Ser 5.42 (3.5 Å) were observed (Figure 4a). Many site-directed mutagenesis experiments have indicated the significance of the Asp 3.32 residue in aminergic GPCRs, due to its crucial interaction with the protonated nitrogen of the ligand.[85,90,91] TM5 residues in positions 5.42, 5.43, and 5.46, while not conserved throughout the aminergic GPCRs, are often involved in aminergic ligand−receptor interactions.[90] Olanzapine also made a substantial number of vdW contacts, particularly with the aromatic network (involving residues Trp 6.48, Phe 6.51, and Phe 6.52), and hydrophobic interactions with Ile 183 in ECL2, adjacent to the conserved Cys 182. The importance of the aromatic network has been identified by site-directed mutagenesis.[84,88,90,92]

Docking clozapine (**4**) into the 5-HT_{2A} receptor also displayed interactions consistent with site-directed mutagenesis.[91,92,94−99] Clozapine made an ionic interaction with Asp 3.32 (3.0 Å) (Figure 4b). In this instance, and unlike olanzapine in the D_2 receptor, clozapine failed to make a hydrogen bond with serine residues in TM5. Strong interactions with the aromatic network were observed, as was an interaction with Ala 230 in ECL2, adjacent to the conserved Cys 227. Similar results were obtained for the remaining homology models (Figure S2, Supporting Information).

Ligand Properties. Having built models of the D_2, D_3, D_4, 5-HT_{1B}, 5-HT_{2A}, 5-HT_{2B}, 5-HT_{2C}, M_1, and H_1 receptors, we utilized a virtual screening set of 20−57 known active compounds, specific for each receptor and subtype (Table S4, Supporting Information), mixed with 1000 drug-like decoys obtained from Schrödinger,[64] to assess model quality in the region of the ligand binding site. To reduce the potential structural bias, the ligand used for binding site optimization (see above) was excluded from virtual screening experiments. The properties of the Schrödinger decoy library have been assessed by Friesner et al. and are believed to be representative of ligands of a pharmaceutical compound library that would be capable of competing with known actives in virtual screening.[64]

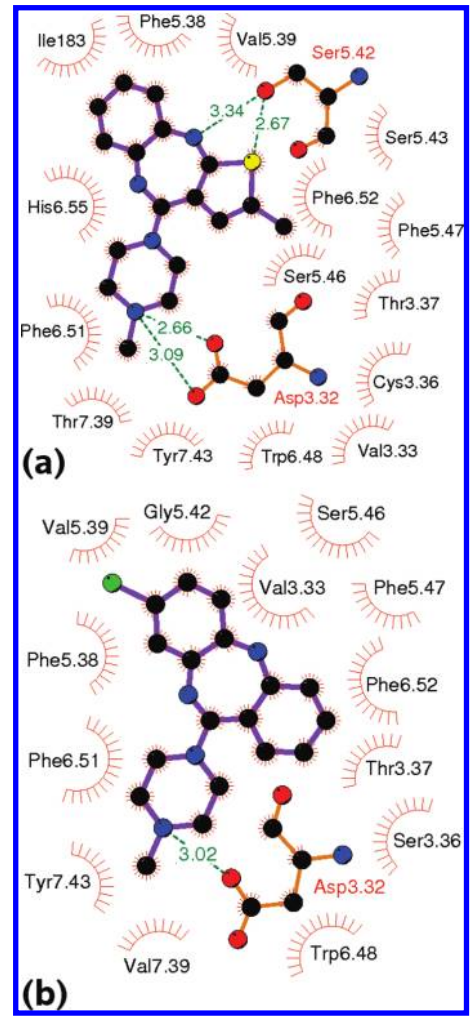

Figure 4. Schematic 2D plots of intermolecular interactions observed for flexible receptor docking: (a) olanzapine docked into the D_2 receptor and (b) clozapine docked into the 5-HT_{2A} receptor. Hydrogen atoms omitted for clarity. Nonbonded interactions are vdW, red spokes, and hydrogen bonds, dashed green lines. The plots were created with the program LIGPLOT.[93]

In this work, the properties of the decoy ligands and the active library compounds were predicted using QikProp, ROCS, and UNITY (Table S5, Supporting Information). Both the 2D and 3D Tanimoto calculations utilized the ligand from binding site refinement as the reference compound. The ROCS shape Tanimoto score was used to determine, on average, how similar the ligands were within each set and was found to be approximately 0.5 for each set, indicating that the sets are of similar diversity. An average 2D Tanimoto score was calculated using UNITY; the majority of the decoy ligands showed, on average, high diversity (0.2−0.3), with only the β_2 ligands displaying an average 2D diversity of 0.5. The properties investigated using QikProp included:

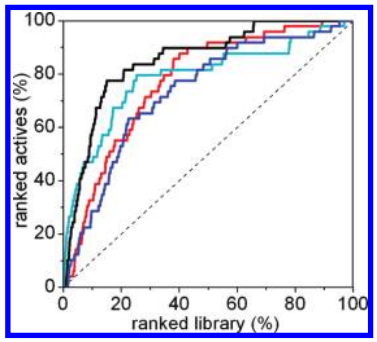

Figure 5. Enrichment plots for 5-HT_{2A} homology models: initial, red (model 1); after loop refinement, blue (model 2); after flexible receptor docking, cyan (model 3); and final, black (model 4).

molecular weight, number of rotatable bonds, polar surface area, calculated log *P*, number of hydrogen-bond donors and acceptors, and the solvent accessible volume. Generally, the properties of the active compounds were similar to those of the decoy library. The average molecular weight of the decoys was 360 g/mol, and the average molecular weight of all the active compounds was 354 g/mol. However, the average molecular weights for the β_2, H_1, and M_1 active compounds were smaller, on average, than the decoys (301, 328, and 320 g/mol, respectively). The average polar surface area of the active compounds was substantially lower than the decoys at 49 and 86.7 $Å^2$, respectively. The average number of hydrogen-bond donors in the active compounds was also lower than the decoys at 0.9 and 1.9, respectively. Of the decoy compounds, 49% contained positively charged amines, which is a characteristic known to be essential for activity for aminergic GPCRs.[90] Variations in the properties of the active compounds compared to those of the decoys were unavoidable, as the number of active compounds was limited.

Assessment of Homology Modeling Procedure by Virtual Screening. The 5-HT_{2A} homology model was used to assess our modeling procedure at different stages throughout the refinement process. Virtual screening was employed to dock a library of 1000 decoys and 49 active compounds into the 5-HT_{2A} model at four key stages during the homology modeling and the refinement process (Figure 5). First, homology model 1 (prior to any refinement) was used to screen the database. This model displayed relatively poor enrichment at 2 and 5% of the ranked database, however, 16 active compounds were recovered in the top 10% of the ranked database, showing modest enrichment of 3.3. Model 2 (after loop refinement) behaved similarly to model 1, with a slight improvement at the early stages of the screen. Model 2 also had a moderate enrichment of 2.9 at 10% of the ranked database, however only 14 actives were recovered at 10%. Model 3 (after flexible receptor docking) showed high early enrichment at 2% of the virtual screen, with an enrichment factor of 13.3. This model also displayed good enrichment at 5 and 10% of the ranked database screened (7.7 and 4.7, respectively), with 23 active compounds recovered in the top 10% of the ranked database. Model 4 (the final model), while marginally losing some very early enrichment, compared to model 3 showed the highest enrichment of 5.9 at 10% of the ranked database, recovering 29 out of the 49 active compounds. The area under the curve (AUC) was also calculated for the four enrichment plots to compare the four graphs (Table 4). Model 4 has the highest AUC value of 8463, indicating that this model displayed the best recovery of active compounds overall.

Table 4. AUC for the Enrichment Plots for the Initial, Intermediate, and Final Homology Models for the 5-HT_{2A} Receptor

model	AUC
initial (model 1)	7665.0
after loop refinement (model 2)	7279.0
after flexible receptor docking (model 3)	7837.6
final (model 4)	8463.1

Since we are interested in utilizing GPCR homology models for virtual screening purposes, we are aiming to achieve high enrichment within the top 10% of a virtual screen, as this is a common cutoff used to differentiate potential leads from inactive compounds. In this respect, the final model of the 5-HT_{2A} receptor shows a considerable improvement over the initial and intermediate models. This demonstrates the importance of such modeling steps as loop refinement and, particularly, binding site optimization.

Homology Model Evaluation by Virtual Screening. Figure 6 shows the virtual screening results as cumulative plots of the active compounds identified. Enrichment factors were calculated at 2, 5, and 10% of the ranked database screened (Table 5). The maximum possible enrichment factors for each screening run varied due to differences in the number of active compounds available for each receptor, reducing the utility of enrichment factors for direct comparison between models, particularly at 2% of the ranked database screened. However, we do report enrichment factors at 2% of the screen, as this describes early enrichment and is indicative of the success of the virtual screen. Enrichment factors at 5 and 10% of the screened database could be more easily used for more direct comparisons of the homology models. At 5% of the ranked database screened, we deemed that enrichment factors above 10 indicated high enrichment, above 5 indicated good enrichment, above 3 indicated moderate enrichment, and any enrichment factors less than or equal to 3 indicated poor enrichment. At 5% of the screen, three models showed good enrichment (5-HT_{2A}, 5-HT_{1B}, and D_2) and four models displayed moderate enrichment (M_1, 5-HT_{2C}, D_3, and 5-HT_{2B}), while two models showed little, if any, enrichment (D_4 and H_1). Similar results were also obtained at 10% of the ranked database screened. Sets of active compounds for receptors which contained more hydrogen-bond acceptors, on average, displayed slightly higher enrichment factors (e.g., 5-HT_{2A} and M_1) (Table S5, Supporting Information).

Of all homology models developed, 5-HT_{2A} performed the best with good enrichment factors obtained at 2, 5, and 10% of the screened database (Table 5). The curve tapers off after approximately 75% of the active compounds had been recovered (Figure 6b). Allowing receptor flexibility in virtual screening may improve the identification of diverse classes of active compounds, since subtle conformational changes in the protein may be required in order to accommodate different ligand types. However, this would significantly increase the computational cost. Binding modes were concordant with experimental data, with the earlier hits making more of the expected contacts, but generally, the majority of the key interactions were observed, as exemplified in Figure 7.

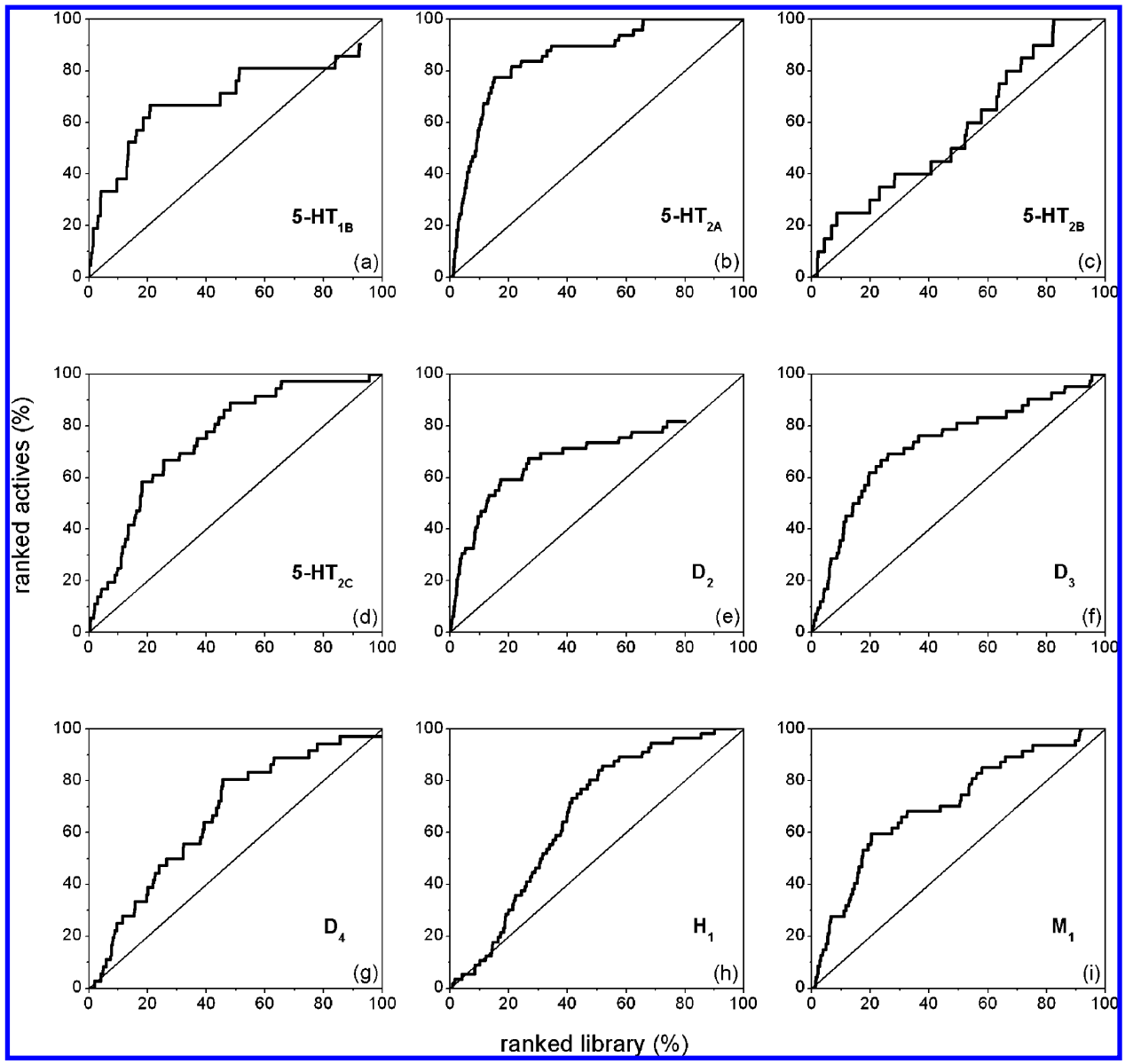

Figure 6. Enrichment plots for homology models (bold) with the percentage of the ranked library (*x*-axis) vs the ranked actives (*y*-axis). Fine line indicates random.

Table 5. Enrichment Factors for Homology Models and Maximum Enrichment Factors at *x*% of the Ranked Database Screened

	enrichment factor (at *x*% of the ranked database screened)					
receptor	2%	max 2%	5%	max 5%	10%	max 10%
5-HT_{2A}	6.1	21.4	6.9	19.8	5.9	10.0
5-HT_{1B}	9.3	48.6	6.5	19.6	3.8	9.9
D_2	7.1	21.4	6.1	19.8	4.5	10.0
M_1	3.2	22.3	3.4	19.8	2.8	10.0
5-HT_{2C}	5.3	28.0	3.2	19.9	2.4	10.0
D_3	4.6	24.3	3.2	19.7	3.5	9.9
5-HT_{2B}	4.9	48.6	3.0	20.0	2.5	10.0
D_4	1.4	28.8	1.7	19.9	2.5	10.0
H_1	1.7	18.9	1.1	18.9	1.1	10.0

The 5-HT_{1B} and D_2 models were similarly predictive. Both displayed high early enrichment at 5% of the ranked database (Table 5). The graphs tapered off after approximately 60−65% of the active compounds were retrieved (Figure 6a and e), with approximately 80% of the active compounds recovered in these screens. The compounds that were not identified in the virtual screen were often larger, bulkier compounds with a substituted tricyclic nucleus and an extended chain, indicating that the receptor conformation may need optimization in order to accommodate these ligands.

The 5-HT_{2B}, 5-HT_{2C}, D_3, and M_1 models all displayed moderate enrichment at 5 and 10% of the ranked database screened (Table 5). For the 5-HT_{2C}, D_3, and M_1 models, the majority of active compounds were identified in the virtual screen, but the number of active compounds identified tapers off after approximately 60% of actives were recovered. The docked poses generated for the active compounds retrieved later in the screen often lacked key interactions.

However, the enrichment plot for the 5-HT_{2B} model (Figure 6c) tapered off after approximately 25% of the active compounds were recovered, following which the identification of active compounds was sporadic and the curve showed little enrichment, as the curve tapered off outside the top 10% of the ranked library. This is not evident in the enrichment factors, as they were not calculated after 10% of the ranked database. The moderate virtual screening results for the 5-HT_{2B} receptor could be attributed to difficulties in

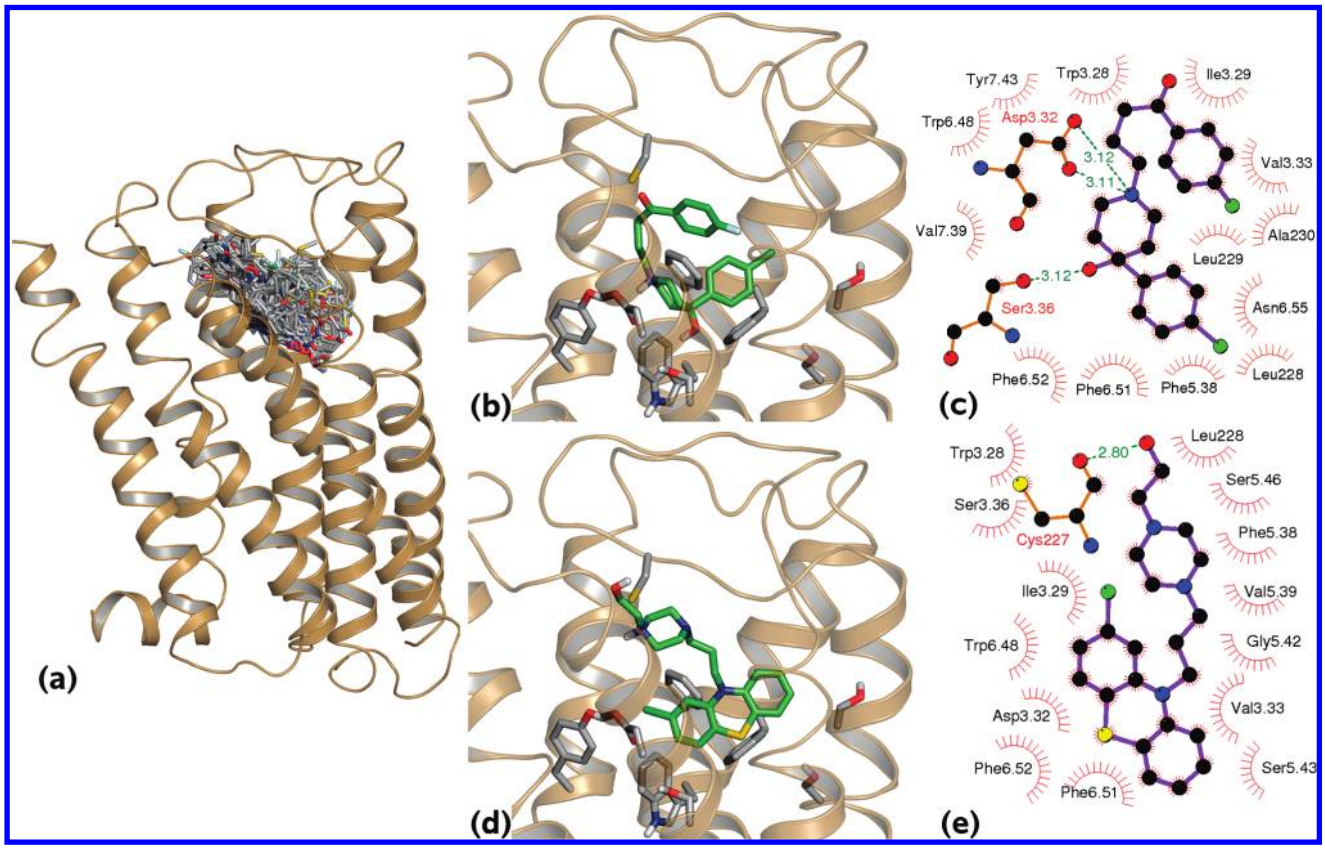

Figure 7. (a) All active compounds docked into the 5-HT_{2A} receptor during virtual screening. (b) A 3D image of an early hit, haloperidol, and (c) its corresponding schematic 2D plot showing intermolecular interactions. (d) A 3D image of a late hit, perphenazine, and (e) its corresponding schematic 2D plot showing intermolecular interactions. The 2D schematic plots were created with the program LIGPLOT.[93] Hydrogen atoms were omitted for clarity (nonbonded interactions; vdW, red spokes; hydrogen bonds, dashed greenlines). The 3D images were created using PyMOL.[78]

modeling ECL2, such as extra residues in the loop above the orthosteric site as well as bulky residues, such as Arg 211 and Lys 213 that protruded into the binding site from the loop (refer to Figure 3a). This model, particularly the loop region, will need further refinement.

The D_4 and H_1 homology models performed poorly in the virtual screens, showing little, if any, enrichment (Figure 6g and h). This is a result of difficulties experienced modeling ECL2 as well as optimizing the position of bulky residues in the binding site. The poses of the active compounds greatly varied, indicating additional refinement of the binding site, and ECL2 is required for these models.

GPCRs bind a diverse range of ligands, even at single receptors: for example, the D_2 receptor binds dibenzodiazepines (e.g., clozapine, **4**) and butyrophenones (e.g., haloperidol, **5**). Each ligand type is known to act at D_2 receptors, however, due the diversity in ligand shapes (as indicated by the average Tanimoto score in Table S5 of the Supporting Information), it is likely that the receptor undergoes conformational changes to accommodate these ligands. One receptor conformation may not, therefore, identify all the active compounds from a diverse set of ligands. During virtual screening, it was commonly observed that one class of ligand was favored over others, indicating that different receptor conformations may be required for extensive virtual screening studies.

To the best of our knowledge, the only study where a GPCR model (based, at least, partially on β_2) was evaluated by Glide-driven virtual screening is that of Kneissl et al.[34] These authors obtained an enrichment factor of 2.6 at 10% of the screened database for their model of the NK_1 receptor. The direct comparison of enrichment factors is not possible due to differences in the virtual libraries used. However, the combination of the enrichment factor values and the enrichment plots, presented by Kneissl et al., allows a cautious optimism with respect to our models. Five of our models delivered enrichment factors at 10% of the screened database greater than 2.6 and three, only marginally smaller. Furthermore, the enrichment curves presented here (except for 5-HT_{2B}, D_4 and H_1) have demonstrated better enrichment than the enrichment curves generated for NK_1.

CONCLUSIONS

We have developed a straightforward approach for the construction of aminergic GPCR homology models that utilizes Prime for initial model construction, induced fit docking for binding site optimization, and Glide for the evaluation by virtual screening of known ligands and decoys. The 5-HT_{2A} homology model was tested at four key stages during the refinement process to assess the developed protocol. Overall, each step in the refinement process advanced the models. Loop refinement improved the initial model slightly, whereas binding site optimization (by flexible receptor docking) dramatically increased the early enrichment. Using this protocol, we have developed a set of homology models (5-HT_{2A}, 5-HT_{1B}, D_2, 5-HT_{2C}, D_3, and M_1) which are able to distinguish active compounds from a set of drug-like decoys, with the 5-HT_{2A} receptor showing the

highest enrichment. A minority of the models developed using this method (5-HT_{2B}, D_4, and H_1) provided poorer results, which we believe is due to deficiencies in modeling extracellular loop 2. The homology models are available as Supporting Information so that researchers can use these structures and compare them to their own results. Further work is underway in our laboratory to investigate the refinement of these homology models using molecular dynamics in a solvated lipid bilayer.

ACKNOWLEDGMENT

F.M.M. is the recipient of an Australian Postgraduate Award (APA) scholarship. This work was supported by the Victorian Partnership for Advanced Computing (VPAC) and by the National Computational Infrastructure (NCI), which is supported by the Australian Commonwealth Government.

Supporting Information Available: A summary of recent GPCR crystal structures, sequence identity, the ligands used in flexible receptor docking and in virtual screening, length of ECL2 for each receptor and the ECL2 sections used in the loop refinement protocol, list of the active compounds docked during virtual screening, average ligand properties, cognate ligand docking results, schematic 2D plots of intermolecular interactions from flexible receptor docking, multiple sequence alignment file, and PDB files of homology models. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Kirk, S.; Glazebrook, J.; Grayson, B.; Neill, J.; Reynolds, G. Olanzapine-induced weight gain in the rat: role of 5-HT2C and histamine H1 receptors. *Psychopharmacology (Heidelberg, Ger.)* **2009**, *207*, 119–125.

(2) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Trong, I. L.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **2000**, *289*, 739–745.

(3) Okada, T.; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E. M.; Shichida, Y. Functional role of internal water molecules in rhodopsin revealed by x-ray crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5982–5987.

(4) Okada, T.; Sugihara, M.; Bondar, A.-N.; Elstner, M.; Entel, P.; Buss, V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **2004**, *342*, 571–583.

(5) Kalani, M. Y. S.; Vaidehi, N.; Hall, S. E.; Trabanino, R. J.; Freddolino, P. L.; Kalani, M. A.; Floriano, W. B.; Kam, V. W. T.; Goddard, W. A. III. The predicted 3D structure of the human D_2 dopamine receptor and the binding site and binding affinities for agonists and antagonists. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3815–3820.

(6) Tehan, B. G.; Lloyd, E. J.; Wong, M. G.; Chalmers, D. K. Analysis of agonism by dopamine at the dopaminergic D_2 G-protein coupled receptor based on comparative modelling of rhodopsin. *Mol. Simulat.* **2002**, *28*, 865–888.

(7) Boeckler, F.; Lanig, H.; Gmeiner, P. Modeling the similarity and divergence of dopamine D_2-like receptors and identification of validated ligand-receptor complexes. *J. Med. Chem.* **2005**, *48*, 694–709.

(8) Ortore, G.; Tuccinardi, T.; Bertini, S.; Martinelli, A. A theoretical study to investigate D2DAR/D4DAR selectivity: receptor modeling and molecular docking of dopaminergic ligands. *J. Med. Chem.* **2006**, *49*, 1397–1407.

(9) Bissantz, C.; Schalon, C.; Guba, W.; Stahl, M. Focused library design in GPCR projects on the example of 5-HT_{2C} agonists: comparison of structure-based virtual screening with ligand-based search methods. *Proteins* **2005**, *61*, 938–952.

(10) Evers, A.; Klabunde, T. Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the alpha1A adrenergic receptor. *J. Med. Chem.* **2005**, *48*, 1088–1097.

(11) Xhaard, H.; Rantanen, V. V.; Nyronen, T.; Johnson, M. S. Molecular evolution of adrenoceptors and dopamine receptors: implications for the binding of catecholamines. *J. Med. Chem.* **2006**, *49*, 1706–1719.

(12) Kiss, R.; Noszál, B.; Rácz, Á.; Falus, A.; Erõs, D.; Keserú, G. M. Binding mode analysis and enrichment studies on homology models of the human histamine H4 receptor. *Eur. J. Med. Chem.* **2008**, *43*, 1059–1070.

(13) Evers, A.; Hessler, G.; Matter, H.; Klabunde, T. Virtual screening of biogenic amine-binding G-protein coupled receptors: comparative evaluation of protein- and ligand-based virtual screening protocols. *J. Med. Chem.* **2005**, *48*, 5448–5465.

(14) Rosenbaum, D. M.; Cherezov, V.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Yao, X.-J.; Weis, W. I.; Stevens, R. C.; Kobilka, B. K. GPCR engineering yields high-resolution structural insights into β_2-adrenergic receptor function. *Science* **2007**, *318*, 1266–1273.

(15) Rasmussen, S. G. F.; Choi, H.-J.; Rosenbaum, D. M.; Kobilka, T. S.; Thian, F. S.; Edwards, P. C.; Burghammer, M.; Ratnala, V. R. P.; Sanishvili, R.; Fischetti, R. F.; Schertler, G. F. X.; Weis, W. I.; Kobilka, B. K. Crystal structure of the human β2 adrenergic G-protein-coupled receptor. *Nature* **2007**, *450*, 383–387.

(16) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-resolution crystal structure of an engineered human β_2-adrenergic G protein coupled receptor. *Science* **2007**, *318*, 1258–1265.

(17) Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V.-P.; Chien, E. Y. T.; Velasquez, J.; Kuhn, P.; Stevens, R. C. A specific cholesterol binding site Is established by the 2.8 Å structure of the human β_2-adrenergic receptor. *Structure* **2008**, *16*, 897–905.

(18) Murakami, M.; Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **2008**, *453*, 363–367.

(19) Shimamura, T.; Hiraki, K.; Takahashi, N.; Hori, T.; Ago, H.; Masuda, K.; Takio, K.; Ishiguro, M.; Miyano, M. Crystal structure of squid rhodopsin with intracellularly extended cytoplasmic region. *J. Biol. Chem.* **2008**, *283*, 17753–17756.

(20) Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X. Structure of a β_1-adrenergic G-protein-coupled receptor. *Nature* **2008**, *454*, 486–491.

(21) Park, J. H.; Scheerer, P.; Hofmann, K. P.; Choe, H.-W.; Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **2008**, *454*, 183–187.

(22) Scheerer, P.; Park, J. H.; Hildebrand, P. W.; Kim, Y. J.; Krausz, N.; Choe, H.-W.; Hofmann, K. P.; Ernst, O. P. Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **2008**, *455*, 497–502.

(23) Jaakola, V.-P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **2008**, *322*, 1211–1217.

(24) Michino, M.; Abola, E.; Brooks, C. L.; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat. Rev. Drug Discov.* **2009**, *8*, 455–463.

(25) Topiol, S.; Sabio, M. X-ray structure breakthroughs in the GPCR transmembrane region. *Biochem. Pharmacol.* **2009**, *78*, 11–20.

(26) Kobilka, B.; Schertler, G. F. X. New G-protein-coupled receptor crystal structures: insights and limitations. *Trends Pharmacol. Sci.* **2008**, *29*, 79–83.

(27) Mobarec, J. C.; Sanchez, R.; Filizola, M. Modern homology modeling of G-protein coupled receptors: which structural template to use. *J. Med. Chem.* **2009**, *52*, 5207–5216.

(28) Yuzlenko, O.; Kieæ-Kononowicz, K. Molecular modeling of A_1 and A_2 adenosine receptors: comparison of rhodopsin- and β_2-adrenergic-based homology models through the docking studies. *J. Comput. Chem.* **2009**, *30*, 14–32.

(29) Selent, J.; López, L.; Sanz, F.; Pastor, M. Multi-receptor binding profile of clozapine and olanzapine: a structural study based on the new β_2 adrenergic receptor template. *ChemMedChem* **2008**, *3*, 1194–1198.

(30) Pellissier, L. P.; Sallander, J.; Campillo, M.; Gaven, F.; Queffeulou, E.; Pillot, M.; Dumuis, A.; Claeysen, S.; Bockaert, J.; Pardo, L. Conformational toggle switches implicated in basal constitutive and agonist-induced activated states of 5-hydroxytryptamine-4 receptors. *Mol. Pharmacol.* **2009**, *75*, 982–990.

(31) Tan, K.; Pogozheva, I. D.; Yeo, G. S. H.; Hadaschik, D.; Keogh, J. M.; Haskell-Leuvano, C.; O'Rahilly, S.; Mosberg, H. I.; Farooqi, I. S. Functional characterization and structural modeling of obesity associated mutations in the melanocortin 4 receptor. *Endocrinology* **2009**, *150*, 114–125.

(32) Lim, H. D.; Jongejan, A.; Bakker, R. A.; Haaksma, E.; de Esch, I. J. P.; Leurs, R. Phenylalanine 169 in the second extracellular loop of the human histamine H_4 receptor is responsible for the difference in agonist binding between human and mouse H_4 receptors. *J. Pharmacol. Exp. Ther.* **2008**, *327*, 88–96.

(33) Straβer, A.; Wittmann, H.-J.; Seifert, R. Ligand-specific contribution of the N terminus and E2-loop to pharmacological properties of the histamine H_1-receptor. *J. Pharmacol. Exp. Ther.* **2008**, *326*, 783–791.
(34) Kneissl, B.; Leonhardt, B.; Hildebrandt, A.; Tautermann, C. S. Revisiting automated G-protein coupled receptor modeling: the benefit of additional template structures for a neurokinin-1 receptor model. *J. Med. Chem.* **2009**, *52*, 3166–3173.
(35) Dong, M.; Lam, P. C. H.; Pinon, D. I.; Abagyan, R.; Miller, L. J. Elucidation of the molecular basis of cholecystokinin peptide docking to its receptor using site-specific intrinsic photoaffinity labeling and molecular modeling. *Biochemistry* **2009**, *48*, 5303–5312.
(36) Ivanov, A. A.; Barak, D.; Jacobson, K. A. Evaluation of homology modeling of G-protein-coupled receptors in light of the A_{2A} adenosine receptor crystallographic structure. *J. Med. Chem.* **2009**, *52*, 3284–3292.
(37) Li, G.; Haney, K. M.; Kellogg, G. E.; Zhang, Y. Comparative docking study of anibamine as the first natural product CCR5 antagonist in CCR5 homology models. *J. Chem. Inf. Model.* **2009**, *49*, 120–132.
(38) Dong, M.; Lam, P. C.-H.; Pinon, D. I.; Sexton, P. M.; Abagyan, R.; Miller, L. J. Spatial approximation between secretin residue five and the third extracellular loop of its receptor provides new insight into the molecular basis of natural agonist binding. *Mol. Pharmacol.* **2008**, *74*, 413–422.
(39) Sudandiradoss, C.; Priya Doss, C. G.; Rajasekaran, R.; Ramanathan, K.; Purohit, R.; Sethumadhavan, R. Investigations on the interactions of scorpion neurotoxins with the predicted structure of D_1 dopamine receptor by protein-protein docking method. A bioinformatics approach. *C. R. Biol.* **2008**, *331*, 489–499.
(40) Valant, C.; Gregory, K. J.; Hall, N. E.; Scammells, P. J.; Lew, M. J.; Sexton, P. M.; Christopoulos, A. A novel mechanism of G protein-coupled receptor functional selectivity. *J. Biol. Chem.* **2008**, *283*, 29312–29321.
(41) Tanrikulu, Y.; Proschak, E.; Werner, T.; Geppert, T.; Todoroff, N.; Klenner, A.; Kottke, T.; Sander, K.; Schneider, E.; Seifert, R.; Stark, H.; Clark, T.; Schneider, G. Homology model adjustment and ligand screening with a pseudoreceptor of the human histamine H_4 receptor. *ChemMedChem* **2009**, *4*, 820–827.
(42) Ko, H.; Das, A.; Carter, R. L.; Fricks, I. P.; Zhou, Y.; Ivanov, A. A.; Melman, A.; Joshi, B. V.; Kováè, P.; Hajduch, J.; Kirk, K. L.; Harden, T. K.; Jacobson, K. A. Molecular recognition in the $P2Y_{14}$ receptor: Probing the structurally permissive terminal sugar moiety of uridine-5′-diphosphoglucose. *Bioorg. Med. Chem.* **2009**, *17*, 5298–5311.
(43) Ehrlich, K.; Götz, A.; Bollinger, S.; Tschammer, N.; Bettinetti, L.; Häbner, H.; Lanig, H.; Gmeiner, P. Dopamine D_2, D_3, and D_4 selective phenylpiperazines as molecular probes to explore the origins of subtype specific receptor binding. *J. Med. Chem.* **2009**, *52*, 4923–4935.
(44) Shah, J. R.; Mosier, P. D.; Roth, B. L.; Kellogg, G. E.; Westkaemper, R. B. Synthesis, structure-affinity relationships, and modeling of AMDA analogs at 5-HT_{2A} and H_1 receptors: structural factors contributing to selectivity. *Bioorg. Med. Chem.* **2009**, *17*, 6496–6504.
(45) Arnold, K.; Kiefer, F.; Kopp, J.; Battey, J.; Podvinec, M.; Westbrook, J.; Berman, H.; Bordoli, L.; Schwede, T. The protein model portal. *J. Struct. Funct. Genomics* **2009**, *10*, 1–8.
(46) Pieper, U.; Eswar, N.; Braberg, H.; Madhusudhan, M. S.; Davis, F. P.; Stuart, A. C.; Mirkovic, N.; Rossi, A.; Marti-Renom, M. A.; Fiser, A.; Webb, B.; Greenblatt, D.; Huang, C. C.; Ferrin, T. E.; Sali, A. MODBASE, a database of annotated comparative protein structure models, and associated resources. *Nucleic Acids Res.* **2004**, *32*, D217-D222.
(47) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
(48) *Sybyl-X*, version 1.0; Tripos: St. Louis, MO, 2009.
(49) *Prime*, version 1.6; Schrödinger, LLC: New York, NY, 2007.
(50) Abagyan, R.; Totrov, M.; Kuznetsov, D. ICM - a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **1994**, *15*, 488–506.
(51) Nygaard, R.; Frimurer, T. M.; Holst, B.; Rosenkilde, M. M.; Schwartz, T. W. Ligand binding and micro-switches in 7TM receptor structures. *Trends Pharmacol. Sci.* **2009**, *30*, 249–259.
(52) Dalton, J. A. R.; Jackson, R. M. An evaluation of automated homology modelling methods at low target template sequence similarity. *Bioinformatics* **2007**, *23*, 1901–1908.
(53) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
(54) Hooft, R. W. W.; Vriend, G.; Sander, C.; Abola, E. E. Errors in protein structures. *Nature* **1996**, *381*, 272.
(55) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., III; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **2007**, *35*, W375–383.
(56) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
(57) *Glide*, version 5.0; Schrödinger, LLC: New York, NY, 2008.
(58) Ballesteros, J. A.; Weinstein, H.; Stuart, C. S. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In. *Methods Neurosci., Academic Press:* **1995**, *25*, 366–428.
(59) *Maestro*, version 8.0; Schrödinger, LLC: New York, NY, 2007.
(60) *LigPrep*, version 2.2; Schrödinger, LLC: New York, NY, 2005.
(61) *Schrödinger Suite 2008 Induced Fit Docking protocol*; Schrödinger, LLC: New York, NY, 2005.
(62) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
(63) Okuno, Y.; Tamon, A.; Yabuuchi, H.; Niijima, S.; Minowa, Y.; Tonomura, K.; Kunimoto, R.; Feng, C. GLIDA: GPCR ligand database for chemical genomics drug discovery database and tools update. *Nucleic Acids Res.* **2008**, *36*, D907–912.
(64) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
(65) *QikProp*; version 3.1; Schrödinger, LLC: New York, NY, 2008.
(66) *ROCS*, version 2.3.1; OpenEye Scientific Software Inc.: Santa Fe, New Mexico, 2007.
(67) Pearlman, D. A.; Charifson, P. S. Improved scoring of ligand-protein interactions using OWFEG free energy grids. *J. Med. Chem.* **2001**, *44*, 502–511.
(68) Jain, A. Bias, reporting, and sharing: computational evaluations of docking methods. *J. Comput. Aided Mol. Des.* **2008**, *22*, 201–212.
(69) Huang, N.; Shoichet, B. K. Exploiting ordered waters in molecular docking. *J. Med. Chem.* **2008**, *51*, 4862–4865.
(70) de Graaf, C.; Rognan, D. Selective structure-based virtual screening for full and partial agonists of the β2 adrenergic receptor. *J. Med. Chem.* **2008**, *51*, 4978–4985.
(71) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
(72) Forster, M. J. Molecular modelling in structural biology. *Micron* **2002**, *33*, 365–384.
(73) Cavasotto, C. N.; Phatak, S. S. Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today* **2009**, *14*, 676–683.
(74) Lovell, S. C.; Davis, I. W.; Arendall III, W. B.; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. Structure validation by Cα geometry: ϕ, Ψ and Cβ deviation. *Proteins* **2003**, *50*, 437–450.
(75) Ho, B. K.; Thomas, A.; Brasseur, R. Revisiting the Ramachandran plot: hard-sphere repulsion, electrostatics, and H-bonding in the α-helix. *Protein Sci.* **2003**, *12*, 2508–2522.
(76) Shi, L.; Javitch, J. A. The second extracellular loop of the dopamine D_2 receptor lines the binding-site crevice. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 440–445.
(77) Zhang, D.; Weinstein, H. Polarity conserved positions in transmembrane domains of G-protein coupled receptors and bacteriorhodopsin. *FEBS Lett.* **1994**, *337*, 207–212.
(78) DeLano, W. L. *The PyMOL molecular graphics system*, DeLano Scientific: Palo Alto, CA, 2002.
(79) Shi, L.; Simpson, M. M.; Ballesteros, J. A.; Javitch, J. A. The first transmembrane segment of the dopamine D2 receptor: accessibility in the binding-site crevice and position in the transmembrane bundle. *Biochemistry* **2001**, *40*, 12339–12348.
(80) Javitch, J. A.; Shi, L.; Simpson, M. M.; Chen, J.; Chiappa, V.; Visiers, I.; Weinstein, H.; Ballesteros, J. A. The fourth transmembrane segment of the dopamine D2 receptor: accessibility in the binding-site crevice and position in the transmembrane bundle. *Biochemistry* **2000**, *39*, 12190–12199.
(81) Javitch, J. A.; Ballesteros, J. A.; Chen, J.; Chiappa, V.; Simpson, M. M. Electrostatic and aromatic microdomains within the binding-site crevice of the D2 receptor: contributions of the second membrane-spanning segment. *Biochemistry* **1999**, *38*, 7961–7968.
(82) Fu, D.; Ballesteros, J. A.; Weinstein, H.; Chen, J.; Javitch, J. A. Residues in the seventh membrane-spanning segment of the dopamine D2 receptor accessible in the binding-site crevice. *Biochemistry* **1996**, *35*, 11278–11285.
(83) Javitch, J. A.; Fu, D.; Chen, J. Residues in the fifth membrane-spanning segment of the dopamine D2 receptor exposed in the binding-site crevice. *Biochemistry* **1995**, *34*, 16433–16439.

(84) Javitch, J. A.; Ballesteros, J. A.; Weinstein, H.; Chen, J. A cluster of aromatic residues in the sixth membrane-spanning segment of the dopamine D2 receptor is accessible in the binding-site crevice. *Biochemistry* **1998**, *37*, 998–1006.
(85) Javitch, J. A.; Fu, D.; Chen, J.; Karlin, A. Mapping the binding site crevice of the dopamine D2 receptor by the substituted-cysteine accessibility method. *Neuron* **1995**, *14*, 825–831.
(86) Mansour, A.; Meng, F.; Meador-Woodruff, J. H.; Taylor, L. P.; Civelli, O.; Akil, H. Site-directed mutagenesis of the human dopamine D2 receptor. *Eur. J. Pharm., Mol. Pharmacol. Sect.* **1992**, *227*, 205–214.
(87) Cox, B. A.; Henningsen, R. A.; Spanoyannis, A.; Neve, R. L.; Neve, K. A. Contributions of conserved serine residues to the interactions of ligands with dopamine D2 receptors. *J. Neurochem.* **1992**, *59*, 627–635.
(88) Taylor, L. P.; Mansour, A.; Akil, H. Hydrophobic residues of the D2 dopamine receptor are important for binding and signal transduction. *J. Neurochem.* **1995**, *65*, 2105–2115.
(89) Coley, C.; Woodward, R.; Johansson, A. M.; Strange, P. G.; Naylor, L. H. Effect of multiple serine/alanine mutations in the transmembrane spanning region V of the D2 dopamine receptor on ligand binding. *J. Neurochem.* **2000**, *74*, 358–366.
(90) Shi, L.; Javitch, J. A. The binding site of aminergic G-protein coupled receptors: the transmembrane segments and second extracellular loop. *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 437–467.
(91) Wang, C.; Gallaher, T.; Shih, J. Site-directed mutagenesis of the serotonin 5-hydroxytrypamine2 receptor: identification of amino acids necessary for ligand binding and receptor activation. *Mol. Pharmacol.* **1993**, *43*, 931–940.
(92) Roth, B. L.; Shoham, M.; Choudhary, M. S.; Khan, N. Identification of conserved aromatic residues essential for agonist binding and second messenger production at 5-hydroxytryptamine2A receptors. *Mol. Pharmacol.* **1997**, *52*, 259–266.
(93) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **1995**, *8*, 127–134.
(94) Roth, B. L.; Willins, D. L.; Kristiansen, K.; Kroeze, W. K. 5-Hydroxytryptamine2-family receptors (5-hydroxytryptamine2A, 5-hydroxytryptamine2B, 5-hydroxytryptamine2C): where structure meets function. *Pharmacol. Ther.* **1998**, *79*, 231–257.
(95) Shapiro, D. A.; Kristiansen, K.; Kroeze, W. K.; Roth, B. L. Differential modes of agonist binding to 5-hydroxytryptamine2A serotonin receptors revealed by mutation and molecular modeling of conserved residues in transmembrane region 5. *Mol. Pharmacol.* **2000**, *58*, 877–886.
(96) Runyon, S. P.; Mosier, P. D.; Roth, B. L.; Glennon, R. A.; Westkaemper, R. B. Potential modes of interaction of 9-aminomethyl-9,10-dihydroanthracene (AMDA) derivatives with the 5-HT2A receptor: a ligand structure-affinity relationship, receptor mutagenesis and receptor modeling investigation. *J. Med. Chem.* **2008**, *51*, 6808–6828.
(97) Braden, M. R.; Nichols, D. E. Assessment of the roles of serines 5.43(239) and 5.46(242) for binding and potency of agonist ligands at the human serotonin 5-HT2A receptor. *Mol. Pharmacol.* **2007**, *72*, 1200–1209.
(98) Braden, M. R.; Parrish, J. C.; Naylor, J. C.; Nichols, D. E. Molecular interaction of serotonin 5-HT2A receptor residues Phe339(6.51) and Phe340(6.52) with superpotent *N*-benzyl phenethylamine agonists. *Mol. Pharmacol.* **2006**, *70*, 1956–1964.
(99) Choudhary, M.; Sachs, N.; Uluer, A.; Glennon, R.; Westkaemper, R.; Roth, B. Differential ergoline and ergopeptine binding to 5-hydroxytryptamine2A receptors: ergolines require an aromatic residue at position 340 for high affinity binding. *Mol. Pharmacol.* **1995**, *47*, 450–457.

CI900444Q