Ligand-Steered Modeling and Docking: A Benchmarking Study in Class A G-Protein-Coupled Receptors

Sharangdhar S. Phatak, Edgar A. Gatica, and Claudio N. Cavasotto*

School of Biomedical Informatics, The University of Texas Health Science Center at Houston, 7000 Fannin, Suite 690, Houston, Texas 77030, United States

Received May 11, 2010

Class A G-protein-coupled receptors (GPCRs) are among the most important targets for drug discovery. However, a large set of experimental structures, essential for a structure-based approach, will likely remain unavailable in the near future. Thus, there is an actual need for modeling tools to characterize satisfactorily at least the binding site of these receptors. Using experimentally solved GPCRs, we have enhanced and validated the ligand-steered homology method through cross-modeling and investigated the performance of the thus generated models in docking-based screening. The ligand-steered modeling method uses information about existing ligands to optimize the binding site by accounting for protein flexibility. We found that our method is able to generate quality models of GPCRs by using one structural template. These models perform better than templates, crude homology models, and random selection in small-scale high-throughput docking. Better quality models typically exhibit higher enrichment in docking exercises. Moreover, they were found to be reliable for selectivity prediction. Our results support the fact that the ligand-steered homology modeling method can successfully characterize pharmacologically relevant sites through a full flexible ligand−flexible receptor procedure.

INTRODUCTION

G-protein-coupled receptors (GPCRs) constitute the largest family of membrane receptors (∼1000 members) and mediate nearly 80% of signal transduction across cellular boundaries.[1,2] The ability of GPCRs to bind a variety of ligands, influence downstream effectors (e.g., enzymes),[3] and regulate physiological functions and processes (e.g., cardiovascular, nervous, endocrine and immune)[4] has been exploited by nearly 50% of existing drug discovery programs.[5] These programs target a mere 3% of GPCR members, highlighting the untapped potential of this structurally conserved seven-transmembrane (7TM) class of drug targets.[6] Though experimental methods have successfully identified small-molecular therapeutics, there is significant interest in using structural information to understand ligand−protein interactions, conformational changes, and rationalize the drug discovery process for GPCRs.[7,8] However, this important drug target proved difficult to express, purify, and crystallize,[9] which explains the paucity of high-quality structures across GPCR families.

The appearance of a high-quality bovine rhodopsin (bRho) crystal structure[10] in 1999 confirmed structural features like the 7TM architecture and putative side chain orientations and explained intramolecular interactions.[8] This structure paved the way for homology modeling of other members of the class A GPCRs and subsequent structure-based drug discovery efforts for the past decade.[11−22] On the basis of a succession of promising technological advancements, e.g., microfocus beamlines, sensitive detectors, advanced software, solution and solid-state NMR for crystallizing membrane proteins,[23,24] the last 3 years have seen a parade of class A GPCR structures: β_2-adrenergic receptor[25,26] (β_2), β_1-adrenergic receptor,[27] adenosine A_{2A} receptor (A2A),[28] and ligand-free opsin.[29] These structures have added to the repository of structural knowledge concerning GPCRs and prompted their use both in concrete drug discovery projects[30−32] and in methodological studies.[33−41] It is clearly evident that despite sharing the overall 7TM shape, GPCRs have unique structural signatures, e.g., the plasticity of helical membranes, binding site residues, conformational variations in the extracellular loop 2 regions, and the diverse position of ligand-binding pockets.[42] These structural differences can be smartly exploited for designing selective compounds across the GPCR family, a long-standing problem in GPCRs.[7] In order to do so, researchers would need a larger set of crystal structures and the development of advanced homology modeling methods for the vast majority of uncrystallized GPCRs. Moreover, it has been reported that docking to homology models yields in some cases enrichments comparable to those of crystal structures.[43−45] This highlights the need for accurate modeling techniques, since limitations of current crystallography methods for GPCRs may be insurmountable for the near future.[2]

The quality of models obtained from homology modeling is dependent on the sequence identity between the template protein and the target, the choice of template, the sequence alignment, and subsequent refinement methods of the crude model.[46] The low sequence identity among class A GPCRs is offset by the common 7TM structure, thus allowing modelers to build overall crude models. However, uncertainties at the side chain and backbone level, mainly due to low

* To whom correspondence should be addressed. E-mail: Claudio.N.Cavasotto@uth.tmc.edu. Phone: +1 (713) 500-3934. Fax: +1 (713) 500-3929.

10.1021/ci100285f
Published on Web 11/16/2010

sequence similarity, may introduce errors that can void the applicability of those models in drug discovery projects.

Proteins often bind to diverse chemotype ligands, thus underlining the importance of incorporating receptor flexibility in the optimization of binding sites.[47] Thus, it is reasonable to include in the modeling process information about known ligands, whenever available.

Klebe and co-workers incorporated ligand−protein restraints derived from manual or rigid-receptor docking in the modeling process;[48] this method was successfully applied to identify novel antagonists for neurokinin-1 class A GPCR.[49] Moro et al. used a rigid-receptor soft docking protocol to generate an ensemble of ligand poses in a crude model; the spatial information of each ligand pose is used to regenerate homology models with diverse side-chain orientations. Then, the side chains and ligand are locally minimized, and the ligand is then redocked to the single lowest energy model.[50] Costanzi's approach amalgamates expert experimental knowledge to optimize binding site residue side chains, with computational modeling to mimic the induced fit effects,[51] thus generating reliable β_2 adrenergic models.[33] Cavasotto et al. developed the ligand-steered homology modeling method, in which known ligands are used to shape and optimize the binding site through flexible ligand−flexible receptor stochastic docking.[22,46,52] This constituted an extension to comparative modeling of the method developed to optimize binding sites of proteins complexed with non-native ligands by accounting for protein flexibility.[53−57] Models of the melanin concentrating hormone receptor 1 (MCH-R1), an antiobesity class A GPCR target, were built through a ligand-steered (LS) procedure using bRho as template and later used in high-throughput docking (HTD). Experimental evaluation of top-ranking compounds confirmed the discovery of six low-micromolar inhibitors.[22] This LS method was also used to rationalize SAR data of ligand binding to the cannabinoid CB2 receptor.[58,59] Using this idea, recently several GPCR models generated using induced fit docking[51] were used to assess their performance in high-throughput docking.[60] Other optimization studies for GPCRs used molecular dynamics[5,61−63] and multiple template based methods.[35]

Originally, the quality of the MCH models generated through ligand-steered homology modeling (LSHM) was indirectly assessed using a small-scale HTD procedure (since by then bRho was the only crystal structure available) and by its performance in an actual lead discovery project. In this work, we present an enhanced version of the LSHM method, validating it through cross-modeling current crystallized GPCRs and by assessing the performance of the models in HTD. We aim to answer the following questions: (i) Is the LSHM method able to generate a small set of near-native models of experimentally crystallized GPCRs, by using one structural template and accounting for protein flexibility? (ii) Can ligand-steered models perform well in HTD when compared to crystal structures, crude homology models prior to any refinement, and to random selection? (iii) What is the relationship between top-ranking models, best models in terms of ligand root-mean-square deviation (rmsd), and the top performing models in HTD? (iv) Are ligand-steered models able not only to perform well in HTD but to be reliable in selectivity prediction?

Table 1. Target−Template[a] Cases Studied

case	target	template (pdb code)
1	β_2	bRho (1gzm)
2	β_2	A2A (3eml)
3	A2A	bRho (1gzm)
4	A2A	β_2 (2rh1)

[a] Abbreviations: β_2, human β_2 adrenergic receptor; bRho, bovine rhodopsin; A2A, human A_{2A} adenosine receptor.

In an attempt to answer these questions, we cross-modeled experimentally solved GPCRs with the LSHM method. The cases studied were (template → target): bRho → β_2, A2A → β_2, bRho → A2A, and β_2 → A2A (Table 1). We assessed the model quality through ligand- and binding site-rmsd. For each case, the top-ranking models were used in small-scale HTD experiments, measuring their performance through the enrichment factor (EF) at 2% of the screened database and comparing it with the corresponding crystal structures, crude homology models, and random selection. We investigated the relationship between model quality and docking enrichment and studied the performance of models in selectivity prediction.

In the Methods section, we describe the LSHM method, the docking protocol, the preparation of chemical libraries, and how docking performance was assessed. We then describe the LS cross-modeling, the HTD performance on crystal structures, crude models and LS models, and selectivity prediction (Results and Discussion). Comparison with the work of other research groups is also discussed. Finally, on the basis of our results on modeling, docking, and selectivity, we go back to answer the above questions (Conclusions).

METHODS

Ligand-Steered Modeling of the Binding Site: General Description. An outline of the LSHM method[22,46] used is shown in Figure 1. In the preparation stage, a crude model of the target is built from an appropriate structural template and the target−template sequence alignment. Information about validated disulfide bonds is included at this stage. Structural strain from the crude homology model is relieved through a restrained local energy minimization. A known ligand is selected, energy minimized, and placed within the binding site. Usually, if many ligands were available, those representing the most diverse chemotypes would be selected, thus ensuring the most structural diversity in the binding site, as was done in the modeling of MCH-R1;[22] however, in the present work, since the objective is to assess model quality, cocrystallized ligands were selected to allow comparison with crystal structures. After seeding the ligand within the binding site, an ensemble of structures is generated by randomizing the positional and rotational coordinates of the ligand, followed by a multistep energy minimization, where the van der Waals interaction is gradually switched from soft to full as in other cases.[22,64] The ligand and receptor are held flexible in this stage.[19] Validated ligand−target interactions from experimental sources are also incorporated at this state in the form of quadratic distance restraints. The energy of the complexes is calculated using crude binding energy estimation where the van der Waals, electrostatic, hydrogen-bonding, and torsional energy terms were considered. The

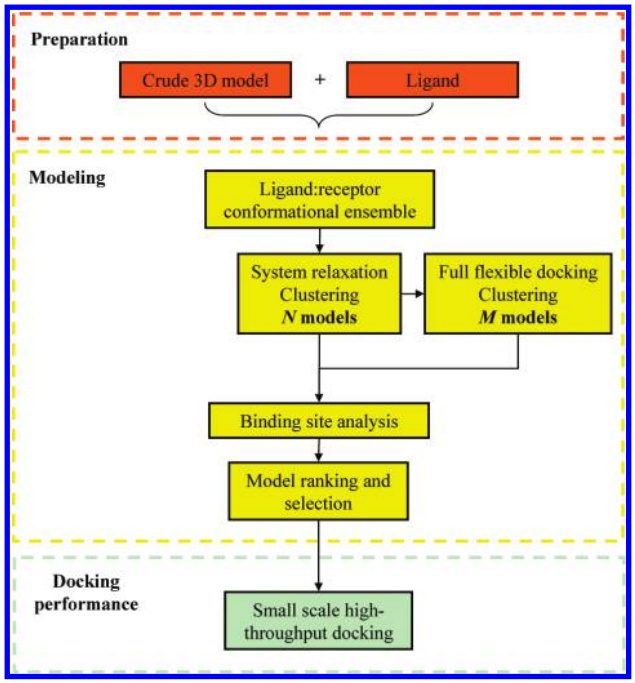

Figure 1. Flowchart of the ligand-steered modeling method. From a crude 3D model of the template and a known ligand (preparation stage, red block), an ensemble of complexes is generated by seeding and randomization the ligand within the binding site (modeling stage, yellow block). This conformational set undergoes relaxation, ligand−target binding energy estimation, and clustering, from which *N* models are selected. Side chain optimization through full flexible ligand−flexible receptor docking follows on these models, yielding *M* models. The two sets of models are merged, binding site analysis is performed, and the models are ranked. The performance of these models is evaluated via small-scale high-throughput docking (docking performance, light green box).

solvation energy in the exposed binding site is accounted for by a simple model of atomic solvation parameters.[65] Structures within the top 10 kcal window are clustered, and the *N* cluster-centers are then subjected to a flexible ligand−flexible receptor docking using a Monte Carlo-based global energy minimization, as has been already described.[22,53,56,57] Side chains within 6 Å of the ligand and facing the binding pocket are considered free, while the backbone is kept fixed. The six rigid coordinates and dihedral angles of the ligand are also considered free. Structures within the top 10 kcal energy window are clustered, and the resulting (*M*) cluster-centers are merged with those (*N*) previously obtained. The resulting set of models is analyzed for key ligand−receptor contacts (from mutagenesis data), visually inspected, and ranked.

System Representation and Energy Calculation. The molecular system was described using the ECEPP/3 force field,[66] within the ICM software platform.[67] Ligand atomic charges were taken from the MMFF force field.[68] The global energy minimization was performed using a Monte Carlo procedure in the torsional space with local minimization[69] and as implemented in ICM. A conformational ensemble of nonredundant structures was generated and stored during the Monte Carlo procedure.[70] Ligand−receptor binding energy was estimated as the interaction energy.

Crude Homology Modeling. The target−template sequences for each of the four cases modeled (Table 1) were aligned on the basis of conserved residue patterns within class A GPCRs[71] (see Supporting Information). The pairwise sequence identity in the helical transmembrane region was case 1, 22%; cases 2 and 4, 31%; case 3, 20%. The extracellular loop 2 (ECL2) of bRho was ignored in the modeling process. The T4-lysozyme residues connecting helixes V and VI of β_2-adrenergic receptor and A2A receptor (PDB 3eml) were omitted. In all cases, the disulfide bond between the cysteines in TM3 and ECL2 was included in our homology model. For each target−template pair, one homology model was generated using Modeler-9v6.[72] The residues were renumbered using the nomenclature of Ballesteros and Weinstein, whereby the most conserved residue in helix X is labeled as X.50.[73]

Clustering of Complexes. The number of clusters was determined using the data mining toolkit Rattle,[74] within the statistical package R (http://www.r-project.org/). The binding energies of complexes from the top 10 kcal window after the ensemble generation stage and the full flexible docking stage were used as input for Rattle. The program in its default mode partitions the data set into a training and a test data set, iterating through multiple clusters and calculating the within-cluster sum of squares (WCSS). A plot was generated, and the number of clusters was chosen where the minimum WCSS was observed. This number was used to cluster the complexes using K-means clustering within the ICM package.[67] The cluster center was chosen as the representative complex for any subsequent calculations.

Preparation of Compound Libraries. Chemical libraries of 29 β_2 antagonists and 51 A2A antagonists were taken from Costanzi's earlier work (Table S1, ref 39) and from the GLIDA[75] database, respectively, including the cocrystallized carazolol (β_2) and ZM241385 (A2A). They were sketched using ChemDraw and converted into three-dimensional structures using the LigPrep module within the Schrödinger software suite.[76] The β_2 compound library had the following physicochemical properties: molecular weight (MW), 318 ± 63; rotatable bonds (RB), 7.5 ± 2.1; log *P*, 2.4 ± 1.2; hydrogen-bond acceptors (HBacc), 3.2 ± 1.1; hydrogen-bond donors (HBdonor) 3.9 ± 0.9. The corresponding values for the A2A chemical library were MW, 300 ± 52; RB, 4.0 ± 1.9; log *P*, 1.5 ± 1.2; HBacc, 4.3 ± 1.2; HBdonor 1.8 ± 1.1. Given that all β_2 antagonist have +1 charge, two decoy libraries were built by random selection from the ZINC[77] database: The β_2 decoy library containing only +1 charge compounds and the A2A decoy library with average zero charge. These libraries were prepared with the LigPrep module, generating protonation states and tautomers compatible with pH 7.0. The physicochemical values corresponding to the β_2 decoy library were MW, 312 ± 27; RB, 5.1 ± 1.8; log *P*, 2.4 ± 1.2; HBacc, 2.7 ± 1.2; HBdonor, 2.6 ± 1.0. Those of the A2A decoy library were MW, 303 ± 32; RB, 4.5 ± 1.7; log *P*, 1.7 ± 0.6; HBacc, 3.9 ± 1.3; HBdonor, 1.3 ± 1.2. The two final screening libraries were built by merging each decoy library with their corresponding ligand libraries, with a total number of compounds of 3500 each.

High-Throughput Docking. Crystal and modeled structures were prepared for docking using the protein preparation workflow within the Schrödinger's Maestro suite,[78] using default settings. Water molecules were removed for the docking process. For virtual screening, the Glide[79,80] software using standard precision (SP) mode with default settings was used.[81] Grids of 20 × 20 × 20 Å^3 were centered between W6.48 and V3.33 for the β_2 models and between W6.48 and

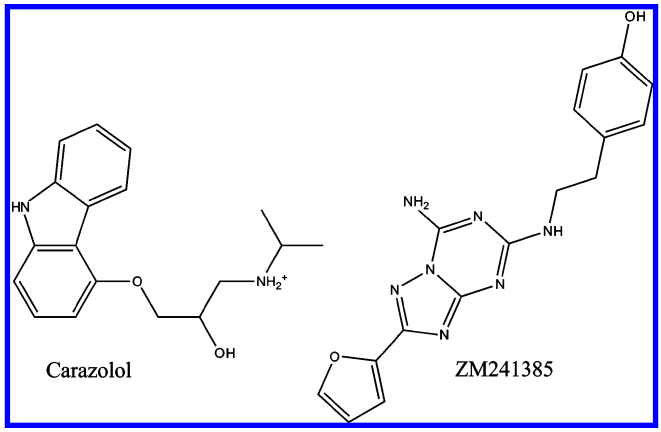

Figure 2. Compounds used in ligand-steered modeling: Carazolol for cases 1 and 2; ZM241385 for cases 3 and 4.

T3.36 for the A2A models, thus covering the overall GPCR pocket bounded by TMs 1−7. No constraints were included in the grid calculations. The enrichment factor was calculated at 2% [EF(2)] of the total database screened according to

$$EF(2) = (hits_{sampled}/N_{sample})/(hits_{total}/N_{total})$$

where $hits_{sampled}$ is the number of active compounds in the top 2% of the compound database, N_{sample} is the subset of 2% of the compound database, $hits_{total}$ is the total number of active compounds in the database, and N_{total} is the total number of compounds in the database.

RESULTS AND DISCUSSION

Ligand-Steered Modeling Identifies Near-Native Models of Experimentally Characterized GPCRs. Structural models for cases 1−4 (Table 1) were generated using the LSHM method (Figure 1). For each target, the corresponding native ligand (carazolol for β_2, ZM241385 for A2A) was used in the modeling stage. For β_2, bRho and A2A were used as templates (cases 1 and 2, respectively). For A2A, bRho and β_2 were used (cases 3 and 4, respectively).

In case 1 (bRho → β_2), the well-documented disulfide bond between C3.25 and C191[25] in the ECL2 was used as a distance restraint in the protein modeling. The remaining ECL2 was built de novo using Modeler-9v6 without any alignment between the target and the template sequences. The cocrystallized ligand of β_2 (carazolol, Figure 2) was used to shape and optimize the binding site of the crude β_2 model, as described in the LSHM flowchart (Figure 1). A distance restraint between carazolol's charged amine and D3.32 was used, which corresponds to an important interaction between known inhibitors and D3.32 in aminergic GPCRs, the lack of which may adversely affect ligand binding.[31] An ensemble of 1000 structures was generated, and the 265 complexes in the top 10 kcal energy window were clustered in five clusters. The five cluster centers (*N* models from Figure 1) were subjected to the full flexible docking step. Of the resulting 314 complexes in the top 10 kcal energy window, seven cluster centers (*M* models from Figure 1) were selected. The binding site of each model was analyzed, and models were discarded whenever the following happens: (i) there was no interaction between the charged amine of carazolol and D3.32 and (ii) the ligand was displaced from the conventional binding pocket surrounding TMs 3, 5, 6, 7 of aminergic GPCRs.[2] The remaining nine models were ranked according to their binding energies and the top five retained. Carazolol exhibited a suitable binding pose (defined as rmsd <2.5 Å) in four out of the five models (Table 2), the best being 1.3 Å (rmsd values were in the range 1.3−3.2 Å). All the five models showed reasonable binding site rmsd values (<3.1 Å), similar to other published results.[33,35,60] In all studied cases, the binding site is defined as those residues having at least one heavy atom within 5 Å of the ligand in the corresponding crystal structure. The top three LS models (which also exhibited the lowest ligand rmsd) had correct rotameric states for all the TM binding site residues, with the exception of F6.52 and N7.39.

Table 2. Model Quality of Ligand-Steered Cross-Modeling

	case 1: bRho → β_2		case 2: A2A → β_2		case 3: bRho → A2A		case 4: β_2 → A2A	
model	ligand[a]	binding site[b]	ligand[a]	binding site[b]	ligand[a,c]	binding site[b,d]	ligand[a,c]	binding site[b]
1	1.3	3.1	1.4	2.9	5.9/4.0	3.8	3.0/2.9	4.1
2	1.4	3.1	1.3	2.9	6.0/4.6	4.1	2.9/2.8	4.1
3	1.6	2.8	3.1	3.4	8.7/5.6	4.0	4.7/4.2	4.6
4	3.2	3.2	1.0	2.9	3.1/3.0	4.2	2.9/2.9	4.2
5	2.1	2.9	0.9	3.0	3.7/3.4	4.3	7.2/5.4	4.1

[a] Heavy-atom rmsd (Å) between the modeled and cocrystallized ligand (carazolol for β_2, ZM241385 for A2A). [b] Binding site heavy-atom rmsd (Å) between the model and the corresponding crystallographic structure (PDB 2rh1 for β_2, 3eml for A2A). [c] The first value corresponds to the whole ligand, the second excluding the high B-factor phenoxy moiety of ZM241385. [d] Residues from the ECL2 were excluded in this calculation.

In the top-ranked model of case 1, the ligand is positioned in the binding site surrounded by TMs 3, 5, 6, and 7. Carazolol's carbazol moiety is oriented toward the serine-dominated region of TM 5 (S5.42, S5.43, and S5.46). The positively charged amine in the tail region of carazolol interacts with D3.32. However, compared to the crystal structure, residue F6.52 has different orientation (Figure 3). Specifically, the trans conformation of F6.52 allows the carbazol moiety to move slightly toward TM 6. The $\chi 1$ angle of F6.52 is difficult to model by default homology modeling due to the corresponding alanine residue in the template (bRho). Our flexible ligand−flexible receptor Monte Carlo based approach was successful in modeling the correct orientation of F6.52 in several complexes in the modeling stage, which also had low ligand rmsd. However, they were not selected, either because they had poor energies due to force field limitations and/or due to the clustering method.

Other $\beta 2$ modeling studies have also placed carazolol in the binding pocket surrounded by TMs 3−7, albeit with some differences. Using bRho as a single template, Mobarec et al.[35] obtained a ligand rmsd of 2.55 Å. Another detailed study using bRho as the only template highlights the importance of correctly modeling the ECL2 and F6.52 rotamer state in the β_2 adrenergic receptor.[33] Carazolol docking was significantly improved (1.7 Å) with the manual changes introduced in F6.52 as compared to the model obtained directly from docking to a homology model.

Similar to case 1, carazolol was used to develop LS models for case 2 (A2A → β_2). The ECL2 loop was included in the construction of the crude model. Only the cysteine residues involved in the disulfide bond, C3.25 and C191, were used to align the ECL2 region. The remaining ECL2 residues

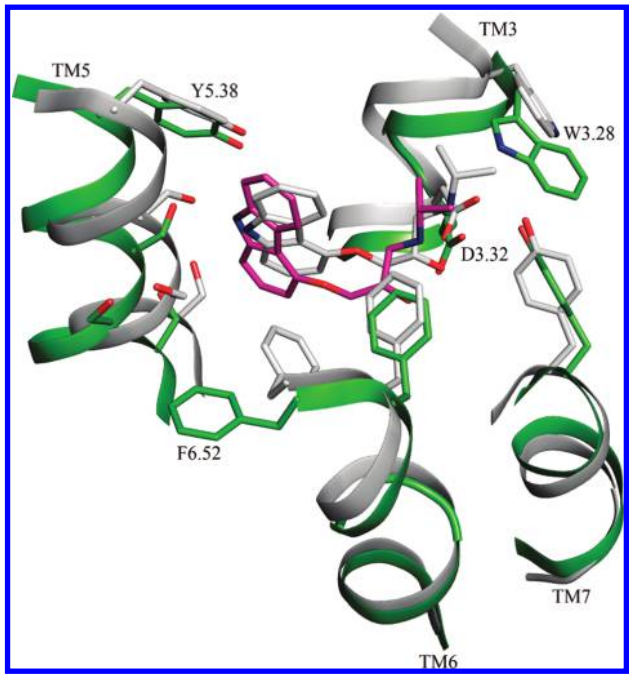

Figure 3. Top-ranked ligand-steered model of β_2 using bRho as template (case 1), superimposed with the β_2 adrenergic crystal structure (PDB 2rh1). Color code: green TMs and carbon atoms, ligand-steered model; magenta carbon atoms, modeled carazolol; white TMs and carbon atoms, crystal structure; white carbon atoms, cocrystallized carazolol. Figure prepared with Pymol (www.pymol.org).

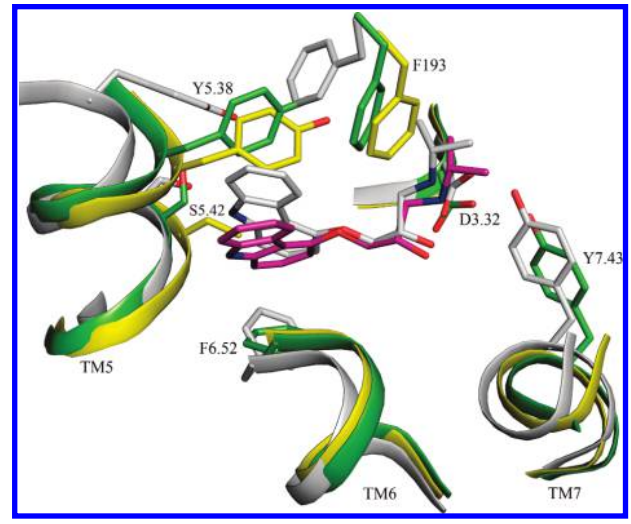

Figure 4. Top-ranked ligand-steered model of β_2 using A2A as template (case 2), superimposed with the β_2 adrenergic crystal structure (PDB 2rh1) and the crude homology model. Color code: Yellow TMs and carbon atoms, crude homology mode; green TMs and carbon atoms, ligand-steered model; magenta carbon atoms, modeled carazolol; white TMs and carbon atoms, crystal structure; white carbon atoms, cocrystallized carazolol. Figure prepared with Pymol (www.pymol.org).

inherited their corresponding conformation from the template structure. From an initial ensemble of 1000 structures, 170 complexes from the top 10 kcal window were clustered, which resulted in nine clusters (*N* models). After the Monte Carlo side-chain optimization step, 169 complexes were clustered into nine more (*M* models), for a total of 18 LS models. Inspection of the binding site, model selection, and ranking were performed as described above for case 1. Four LS models possessed ligand rmsd <1.5 Å (Table 2), and the remaining model showed an acceptable ligand rmsd of 3 Å. With the exception of W3.28 and Y5.38, and similar to what others reported,[33] the three best models in terms of ligand rmsd possessed correct rotameric states for all residues in the binding site, including the gauche+ conformation of F6.52.

The need to shape and optimize the binding site is evident in this case (Figure 4). The crude model residues in the binding pocket clash with the native pose of carazolol. The orientations of Y5.38 (due to the corresponding M177 in A2A template), S5.42, and F193 from ECL2 in the crude model results in steric clashes with the carbazol moiety of carazolol and its flexible moiety, respectively. This reduces the size of the binding site. The LSHM introduces flexibility in those side chains, such that the binding site can now accommodate carazolol. The contributions of the correct gauche+ conformation of F6.52 and modeling of the ECL2 loop also aided in the correct orientation of carazolol. A similar conclusion was drawn by Costanzi.[33]

In case 3 (bRho → A2A), one potential binding site was identified toward TMs 6 and 7. In case 4 (β_2 → A2A), two potential binding sites were identified, one toward TMs 4−6, and the other toward TMs 3, 6, and 7. Both binding sites were oriented quasiparallel to the helices, while this orientation was unlikely to change during the optimization stage. In the preparation stage, the fairly rigid ZM241385 ligand (Figure 2) could not be accommodated in either binding site perpendicular to the TM domains. Hence, in both cases 3 and 4 we seeded ZM241385 in parallel to the membranes in the region surrounded by TMs 3−7. To avoid any bias in the orientation of the ligand, two starting conformations of ZM241385, with the phenoxy group oriented toward the ECLs and toward the intracellular loops (ICLs), were used. A documented hydrogen bond between the exocyclic N15 atom of the ligand and side chain carbonyl oxygen atom of N6.55 was used as a distance restraint.[82] Since two starting positions were used, we generated twice as many (2000) complexes in the ensemble-generation stage (both in cases 3 and in 4).

In case 3, the ECL2 loop was de novo modeled, except for the disulfide bond between C3.25 and Cys166, which was conserved. The ensemble generation stage resulted in 272 complexes in the top 10 kcal window. They were clustered in seven (*N*) models, and after the full flexible docking stage, six (*M*) models were merged with those of the first stage, resulting in 13 ligand-steered models. Previous studies suggest that mostly residues in TMs 3, 6, and 7 are implicated in adenosine 2A antagonist binding.[83,84] This experimental knowledge, along with mutagenesis data on the hydrogen-bond interaction with N6.55, was used to discard eight models during the binding site analysis, in an analogous way as described in case 1. The remaining five models were ranked according to their binding energies. The best model had a ligand rmsd of 5.9 Å (4.0 Å excluding the methoxy group) (Figure 5), and two models had reasonable ligand rmsd values of 3.1 and 3.7 Å (ranked fourth and fifth, respectively). However, it is quite likely that bRho may not be a suitable template structure to model A2A. It is evident from the crystal structure that the ECL2 plays an important role in the accurate positioning of ZM241385. Moreover, modeling A2A signatures like the rotation of TM5 and

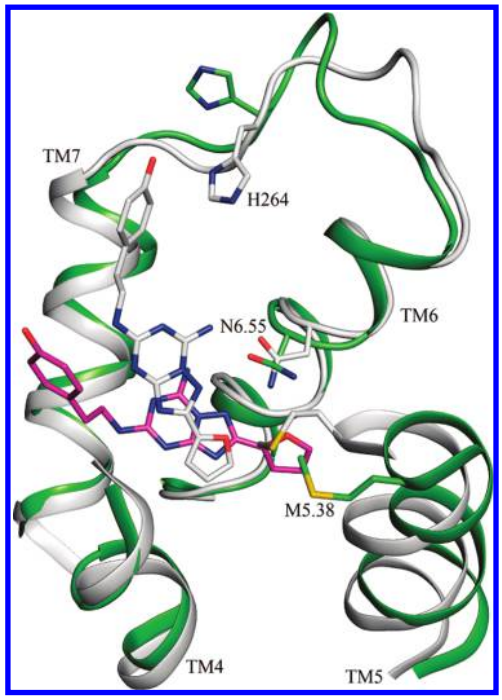

Figure 5. Top-ranked ligand-steered model of A2A using bRho as template (case 3), superimposed with the A2A adrenergic crystal structure (PDB 3eml). Color code: green TMs and carbon atoms, ligand-steered model; magenta carbon atoms, modeled ZM241385; white TMs and carbon atoms, crystal structure; white carbon atoms, cocrystallized ZM241385. Figure is prepared with Pymol (www.pymol.org).

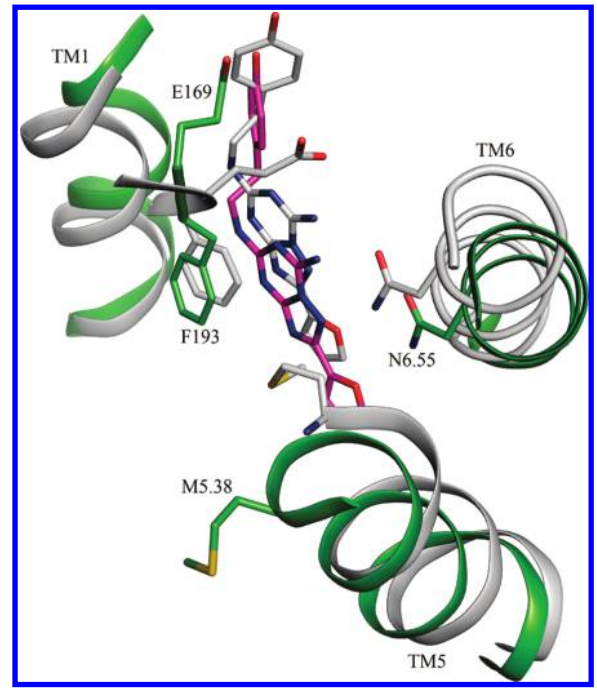

Figure 6. Top-ranked ligand-steered model of A2A using β_2 as template (case 4), superimposed with the A2A adrenergic crystal structure (PDB 3eml). Color code: green TMs and carbon atoms, ligand-steered model; magenta carbon atoms, modeled ZM241385; white TMs and carbon atoms, crystal structure; white carbon atoms, cocrystallized ZM241385. Figure is prepared with Pymol (www.pymol.org).

variability in ECL3, e.g., the orientation of H264, has been a problem for other modelers.[85]

The sequence alignment for case 4 was done using the standard conserved residues as in the other three cases. The cysteine residues in the ECL2 were used in the alignment. Gaps were introduced in the ECL2 in a way that did not affect the alignment for TMs 3 and 4. Out of 168 complexes within the 10 kcal window, seven cluster centers (*N* models) were generated, and nine cluster centers (*M* models), from the 127 complexes obtained via the side-chain optimization step, formed the final set of 16 A2A ligand-steered models. The corresponding binding sites were analyzed and ranked, as described for case 3. The two best-ranked LS models have low ligand rmsd values <3.0 Å (Table 2), comparable to what was obtained recently by Costanzi.[85] Model 1 (Figure 6) reproduces well the distance between the side chain carbonyl oxygen atom in N6.55 and the exocyclic N15 atom of the ligand ZM241385 at 3.7 Å, comparable to other models.[82,85] The F168 and the bicyclic triazolotriazine core of ZM241385 show aromatic stacking interaction, as in the crystal structure. Similar to other models,[85] ours did not predict the inward orientation of M5.38, due to the unique extended bulge of A2A in this region. Overall, our top-ranking model was able to predict the near perpendicular orientation of the ligand.

On an eight-core CPU, and depending on the type of system, generation of final LS models can take approximately 1−2 weeks.

Ligand-Steered Models Exhibit Improved Performance in High-Throughput Docking. The five top-ranking LS models (for each case, see Table 2), the crude homology models, and the crystal structures were subjected to a small scale HTD. The EF(2) was used to evaluate the HTD performance (Table 3).

Table 3. High-Throughput Docking Enrichment for Crystal Structures, Crude Homology Models, and Ligand-Steered Models

	EF(2)[a]			
	case 1: bRho → β_2	case 2: A2A → β_2	case 3: bRho → A2A	case 4: β_2 → A2A
crystal structure	24.1	24.1	8.8	8.8
crude model	17.2	1.7	2.0	4.9
model 1	17.2	15.5	0.0	7.8
model 2	13.8	17.2	0.9	7.8
model 3	8.6	15.5	3.9	0.9
model 4	13.8	6.9	0.0	6.8
model 5	13.8	8.6	0.9	2.9

[a] Enrichment factor at 2% of the compound database.

Glide SP Docking to Crystal Structures. Docking with Glide SP[81] onto the β_2 crystal structure resulted in an EF(2) of 24.1 (Table 3), with the top-ranked pose of carazolol having 0.6 Å rmsd as compared to the cocrystallized structure. Docking onto the A2A crystal structure gave an EF(2) of 8.8, with the top-ranked pose of ZM241385 showing an rmsd of 3.5 Å as compared to the crystal structure, reduced to 1.6 Å if ignoring the high b-factor phenoxy group. Thus, there was an excellent performance for both receptors in terms of docking accuracy and EF. In spite of excluding water molecules in the docking, our EFs are comparable to what has been recently published.[60]

High-Throughput Docking onto Ligand-Steered Models. In case 1, the LS models showed similar or marginally inferior early enrichment rates than the crude model (Table 3). The top-ranked LS models had an EF(2) of 17.2, like the crude model. These results were somewhat unexpected, but other studies on GPCRs also showed inferior—or not significantly better—performance of optimized models as compared to crude models.[38]

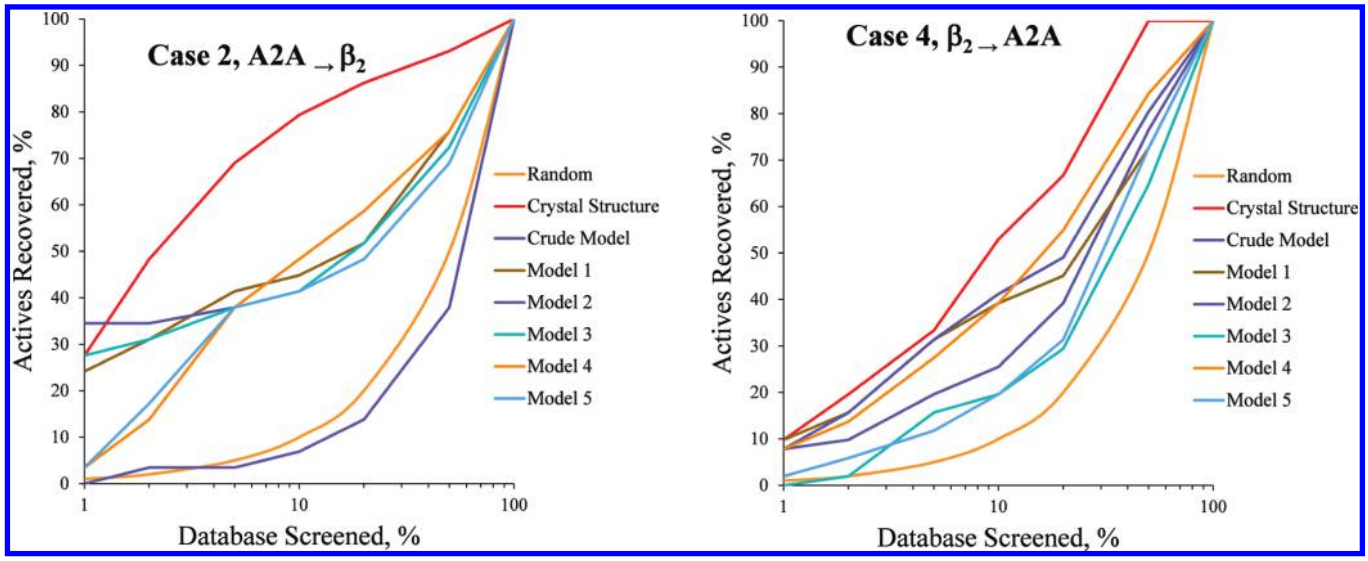

Figure 7. Enrichment curves for cases 2 and 4 for the five top-ranking ligand-steered models. Enrichment for the corresponding crystal structure, crude model, and random selection is displayed as well (β_2, human β_2 adrenergic receptor; A2A, human A_{2A} adenosine receptor).

In case 2, the crude β_2 model fared extremely poorly, with an EF(2) of 1.7 (Table 3). It is clear that refinement is essential for any meaningful application of this model. In the small-scale HTD, the LS models performed significantly better than the crude model (Table 3 and Figure 7). As in case 1, the top-ranked models had excellent ligand rmsd at the modeling stage and the highest enrichment.

The HTD performance of the crude [EF(2) = 2.0] and ligand-steered models of case 3 was rather poor. Except for one model [EF(2) = 3.9], the rest showed very low initial enrichment rates. Our results corroborate other studies that suggest the limitations of using bRho as a template to model A2A,[37,85] and usually β_2 is the preferred template to model A2A.[85] All ligand-steered models failed to identify low-rmsd ligand poses at the HTD performance stage.

Using β_2 as template (case 4), three of the five LS A2A models showed improved or equal performance in HTD as compared to the crude A2A model (Table 3 and Figure 7). The two top ranked models possessed the best enrichment, also correlating with the best ligand rmsd in the modeling stage.

The EFs of our top models are in the same range of those obtained by McRobb et al. in other GPCRs.[60] Moreover, they perform comparably to the crystal structure, what has been also found in other receptors.[43,86]

Use of Ligand-Steered Models To Assess Receptor Selectivity. We analyzed the ability of the LS models to selectively identify target-specific active compounds. For this study, the β_2 and A2A active compounds were merged to the decoy database, and the fraction of β_2 and A2A inhibitors recovered at 5% of the screening database was calculated (Figure 8). In case 1, although the LS model (41%) performed fractionally inferior to the crude model in selectively identifying β_2 inhibitors (45%), it identified fewer A2A actives than both the β_2 crystal structure and crude model: 2.0% for model 1, 12.0% for the crude model, and 4.0% for the crystal structure. In case 2, the top ranked LS model showed a significant improvement compared to the crude model in identifying β_2 antagonists, while displaying a similar trend for selectivity as the crystal structure. Model 1 identified slightly lesser percentage of A2A inhibitors (13%) as compared to the β_2 crystal structure (15%). Neither of the top ranked β_2 models from cases 1 and 2 could identify as many β_2 inhibitors as the crystal structure.

In view of the poor results for case 3 models on HTD (Table 3), we did not test its performance in selectivity prediction. In case 4, although the enrichment of the crude A2A model could be considered acceptable compared to that of the crystal structure, its selectivity profile is inverted, recovering at almost twice the rate β_2 ligands than A2A ones. This is clearly corrected in the optimized model (model 1), although its performance in selectivity prediction is inferior compared to that of the crystal.

As can be seen from Figure 8, docking to templates (rather than models) yielded poor results, contrary to what has been observed for other receptors.[43,44] Our results may hint that quality models may exhibit a better combined performance in HTD and selectivity prediction than crude models or templates. However, a more extensive and general validation in this regard would be needed to prove this statement.

Current Limitations. Despite these promising results, the modeling protocol remains imperfect. Modeling the major structural deviations in GPCRs (e.g., the ECL2 regions or the differences in the TM rearrangement for different ligands) is still challenging, and it would be perhaps desirable to assess the correct fold of TM helices before building the crude model. On the basis of our results, and in agreement with Filizola and co-workers,[35] the authors believe that experimental information from additional template structures may provide logical hints for modeling this challenging class of receptors. However, at least modeling the binding site correctly provides a good starting point for structure-based drug discovery projects.

Incorporating ligand and receptor flexibility in the LSHM method allows a convenient exploration of the ligand and binding site conformational space. However, our modeling is dependent on the force field of choice. For example, several better models (e.g., those with the correct gauche+ conformation of F6.52 in case 1, models with low ligand rmsd values) in the ensemble generation stage were assigned poor energies and hence discarded from further calculations.

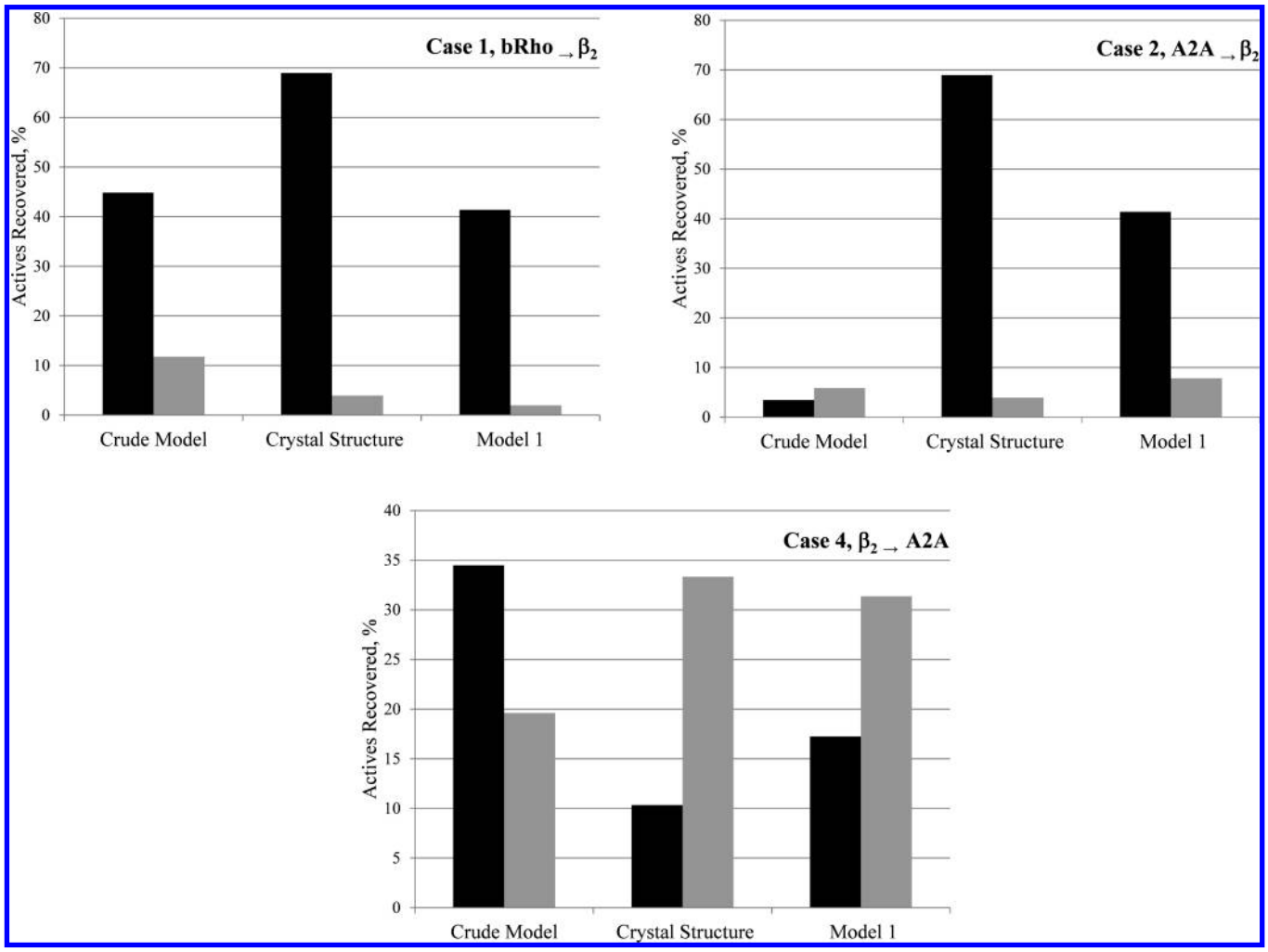

Figure 8. Selectivity prediction using crystal structures, crude models, and best ligand-steered models for cases 1, 2, and 4. The crude model refers to that obtained from the preparation stage of the flowchart. Model 1 refers to the top-ranking ligand-steered model as illustrated in Table 2. The black and gray bars indicate the percentage of β_2 and A2A inhibitors recovered at 5% of the screening database, respectively.

CONCLUSIONS

We conclude by answering the questions we posed in the Introduction about the quality of ligand-steered models, their performance in docking, and model selectivity.

(i) Is the LSHM method able to generate a small set of near-native models of experimentally crystallized GPCRs, by using one structural template and accounting for protein flexibility? Yes, in all the four cases the LSHM was able to generate a small set of near-native models with very good (cases 1 and 2) or acceptable (cases 3 and 4) ligand binding modes.

(ii) Can ligand-steered models perform well in HTD when compared to crystal structures, crude homology models prior to any refinement, and to random selection? In cases 2 and 4, the LSHM performed better than the crude homology models and random selection. In case 1, the performance of the LSHM was similar to that obtained by the crude model, which may be attributed to the lack of ECL2 in the models. In case 3, only one model performed barely acceptably in terms of enrichment. The most likely cause is the significant difference between the binding site profiles of bRho and A2A. None of the models outperformed their respective crystal structures, though the performance of the top-ranked model in cases 1, 2, and 4 was very satisfactory in that regard.

(iii) What is the relationship between top-ranking models, best models in terms of ligand rmsd, and the top performing models in HTD? Typically, models with low ligand rmsd performed well in the HTD exercise. However, models with higher values could also be competitive in terms of enrichment factor. In cases 1, 2, and 4, the top-ranked models had a low ligand rmsd (from the modeling stage) and the highest enrichment.

(iv) Are ligand-steered models able not only to perform well in HTD but to be reliable in selectivity prediction? In cases 1 and 2, the performance of top models in selectivity prediction was comparable to the crystal structure. A similar trend, though less marked, was observed in case 4, correcting the inverted profile of the crude model.

Our results suggest that in spite of the fact that in some cases docking to templates or crude homology models yields comparable results to refined models, or even crystal structures, quality models may exhibit a better balance of good performance in HTD and in selectivity prediction.

An important feature of the LSHM is the concurrent optimization of the ligand position, orientation, and conformation with that of the binding site side chains. This differs from other approaches, which introduce receptor flexibility a posteriori[33] using induced-fit docking[51] or rely on an initial

placement of the ligand either manually[48] or through rigid receptor docking.[48,50] In the second case, it might be difficult in some cases to know in advance how the ligand should be placed within the binding site. In the third case, initial misplacement of the side chains in the crude homology model could preclude the generation of a correct ensemble of ligand conformations through rigid docking. Either option could have a negative impact on the quality of the final model.

Class A GPCRs are one of the most important targets for drug discovery. It is evident that a large set of experimental structures, essential for structure-based drug discovery, will remain unavailable in the near future.[87] Thus, there is a need to develop methods that can optimally model at least the binding site for these receptors. Our results support the fact that the LSHM method addresses this problem by optimizing the pharmacologically relevant sites of GPCRs with a full flexible ligand−flexible receptor approach.

Supporting Information Available: Target−template sequence alignments for the four cases studied. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Millar, R. P.; Newton, C. L. The year in G protein-coupled receptor research. *Mol. Endocrinol.* **2010**, *24*, 261–274.

(2) Rosenbaum, D. M.; Rasmussen, S. G.; Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **2009**, *459*, 356–363.

(3) Jalink, K.; Moolenaar, W. H. G protein-coupled receptors: The inside story. *Bioessays* **2010**, *32*, 13–16.

(4) Lundstrom, K. An overview on GPCRs and drug discovery: Structure-based drug design and structural biology on GPCRs. *Methods Mol. Biol.* **2009**, *552*, 51–66.

(5) Wolf, S.; Bockmann, M.; Howeler, U.; Schlitter, J.; Gerwert, K. Simulations of a G protein-coupled receptor homology model predict dynamic features and a ligand binding site. *FEBS Lett.* **2008**, *582*, 3335–3342.

(6) Schlyer, S.; Horuk, R. I want a new drug: G-protein-coupled receptors in drug development. *Drug Discovery Today* **2006**, *11*, 481–493.

(7) Congreve, M.; Marshall, F. The impact of GPCR structures on pharmacology and structure-based drug design. *Br. J. Pharmacol.* **2009**, *159*, 986–996.

(8) Costanzi, S.; Siegel, J.; Tikhonova, I. G.; Jacobson, K. A. Rhodopsin and the others: A historical perspective on structural studies of G protein-coupled receptors. *Curr. Pharm. Des.* **2009**, *15*, 3994–4002.

(9) Alkhalfioui, F.; Magnin, T.; Wagner, R. From purified GPCRs to drug discovery: The promise of protein-based methodologies. *Curr. Opin. Pharmacol.* **2009**, *9*, 629–635.

(10) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; 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.

(11) Becker, O. M.; Shacham, S.; Marantz, Y.; Noiman, S. Modeling the 3D structure of GPCRs: Advances and application to drug discovery. *Curr. Opin. Drug. Discovery Dev.* **2003**, *6*, 353–361.

(12) Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. Protein-based virtual screening of chemical databases. II. Are homology models of G-protein coupled receptors suitable targets? *Proteins* **2003**, *50*, 5–25.

(13) Bissantz, C.; Logean, A.; Rognan, D. High-throughput modeling of human G-protein coupled receptors: Amino acid sequence alignment, three-dimensional model building, and receptor library screening. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1162–1176.

(14) Freddolino, P. L.; Kalani, M. Y.; Vaidehi, N.; Floriano, W. B.; Hall, S. E.; Trabanino, R. J.; Kam, V. W.; Goddard, W. A. 3rd. Predicted 3D structure for the human beta 2 adrenergic receptor and its binding site for agonists and antagonists. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2736–2741.

(15) Kalani, M. Y.; Vaidehi, N.; Hall, S. E.; Trabanino, R. J.; Freddolino, P. L.; Kalani, M. A.; Floriano, W. B.; Kam, V. W.; Goddard, W. A., 3rd. The predicted 3D structure of the human D2 dopamine receptor and the binding site and binding affinities for agonists and antagonists. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3815–3820.

(16) Costanzi, S.; Joshi, B. V.; Maddileti, S.; Mamedova, L.; Gonzalez-Moa, M. J.; Marquez, V. E.; Harden, T. K.; Jacobson, K. A. Human P2Y(6) receptor: Molecular modeling leads to the rational design of a novel agonist based on a unique conformational preference. *J. Med. Chem.* **2005**, *48*, 8108–8111.

(17) Moro, S.; Jacobson, K. A. Molecular modeling as a tool to investigate molecular recognition in P2Y receptors. *Curr. Pharm. Des.* **2002**, *8*, 2401–2413.

(18) Moro, S.; Spalluto, G.; Jacobson, K. A. Techniques: Recent developments in computer-aided engineering of GPCR ligands using the human adenosine A3 receptor as an example. *Trends Pharmacol. Sci.* **2005**, *26*, 44–51.

(19) Cavasotto, C. N.; Orry, A. J.; Abagyan, R. A. Structure-based identification of binding sites, native ligands and potential inhibitors for G-protein coupled receptors. *Proteins* **2003**, *51*, 423–433.

(20) Martinelli, A.; Tuccinardi, T. Molecular modeling of adenosine receptors: New results and trends. *Med. Res. Rev.* **2008**, *28*, 247–277.

(21) Patny, A.; Desai, P. V.; Avery, M. A. Homology modeling of G-protein-coupled receptors and implications in drug design. *Curr. Med. Chem.* **2006**, *13*, 1667–1691.

(22) Cavasotto, C. N.; Orry, A. J.; Murgolo, N. J.; Czarniecki, M. F.; Kocsi, S. A.; Hawes, B. E.; O'Neill, K. A.; Hine, H.; Burton, M. S.; Voigt, J. H.; Abagyan, R. A.; Bayne, M. L.; Monsma, F. J., Jr. Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. *J. Med. Chem.* **2008**, *51*, 581–588.

(23) Blois, T. M.; Bowie, J. U. G-protein-coupled receptor structures were not built in a day. *Protein Sci.* **2009**, *18*, 1335–1342.

(24) Tate, C. G.; Schertler, G. F. Engineering G protein-coupled receptors to facilitate their structure determination. *Curr. Opin. Struct. Biol.* **2009**, *19*, 386–395.

(25) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **2007**, *318*, 1258–1265.

(26) Rosenbaum, D. M.; Cherezov, V.; Hanson, M. A.; Rasmussen, S. G.; 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 beta2-adrenergic receptor function. *Science* **2007**, *318*, 1266–1273.

(27) Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G.; Tate, C. G.; Schertler, G. F. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* **2008**, *454*, 486–491.

(28) Jaakola, V. P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y.; 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.

(29) 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.

(30) Carlsson, J.; Yoo, L.; Gao, Z. G.; Irwin, J. J.; Shoichet, B. K.; Jacobson, K. A. Structure-based discovery of A(2A) adenosine receptor ligands. *J. Med. Chem.* **2010**, *53*, 3748–3755.

(31) Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K. Structure-based discovery of beta2-adrenergic receptor ligands. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6843–6848.

(32) Sabio, M.; Jones, K.; Topiol, S. Use of the X-ray structure of the beta2-adrenergic receptor for drug discovery. Part 2: Identification of active compounds. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5391–5395.

(33) Costanzi, S. On the applicability of GPCR homology models to computer-aided drug discovery: A comparison between in silico and crystal structures of the beta2-adrenergic receptor. *J. Med. Chem.* **2008**, *51*, 2907–2914.

(34) Dror, R. O.; Arlow, D. H.; Borhani, D. W.; Jensen, M. O.; Piana, S.; Shaw, D. E. Identification of two distinct inactive conformations of the beta2-adrenergic receptor reconciles structural and biochemical observations. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 4689–4694.

(35) 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.

(36) Sherbiny, F. F.; Schiedel, A. C.; Maass, A.; Muller, C. E. Homology modelling of the human adenosine A2B receptor based on X-ray structures of bovine rhodopsin, the beta2-adrenergic receptor and the human adenosine A2A receptor. *J. Comput.-Aided Mol. Des.* **2009**, *23*, 807–828.

(37) Yuzlenko, O.; Kiec-Kononowicz, K. Molecular modeling of A1 and A2A adenosine receptors: Comparison of rhodopsin- and beta2-adrenergic-based homology models through the docking studies. *J. Comput. Chem.* **2009**, *30*, 14–32.

(38) 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.
(39) Vilar, S.; Karpiak, J.; Costanzi, S. Ligand and structure-based models for the prediction of ligand-receptor affinities and virtual screenings: Development and application to the beta(2)-adrenergic receptor. *J. Comput. Chem.* **2010**, *31*, 707–720.
(40) Morizzo, E.; Federico, S.; Spalluto, G.; Moro, S. Human A3 adenosine receptor as versatile G protein-coupled receptor example to validate the receptor homology modeling technology. *Curr. Pharm. Des.* **2009**, *15*, 4069–4084.
(41) Topiol, S.; Sabio, M. Use of the X-ray structure of the Beta2-adrenergic receptor for drug discovery. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1598–1602.
(42) Mustafi, D.; Palczewski, K. Topology of class A G protein-coupled receptors: Insights gained from crystal structures of rhodopsins, adrenergic and adenosine receptors. *Mol. Pharmacol.* **2009**, *75*, 1–12.
(43) Kairys, V.; Fernandes, M. X.; Gilson, M. K. Screening drug-like compounds by docking to homology models: A systematic study. *J. Chem. Inf. Model.* **2006**, *46*, 365–379.
(44) Fan, H.; Irwin, J. J.; Webb, B. M.; Klebe, G.; Shoichet, B. K.; Sali, A. Molecular docking screens using comparative models of proteins. *J. Chem. Inf. Model.* **2009**, *49*, 2512–2527.
(45) Villoutreix, B. O.; Eudes, R.; Miteva, M. A. Structure-based virtual ligand screening: Recent success stories. *Comb. Chem. High Throughput Screen.* **2009**, *12*, 1000–1016.
(46) Cavasotto, C. N.; Phatak, S. S. Homology modeling in drug discovery: Current trends and applications. *Drug Discovery Today* **2009**, *14*, 676–683.
(47) Cavasotto, C. N.; Singh, N. Docking and high throughput docking: Successes and the challenge of protein flexibility. *Curr. Comput.-Aided Drug Design* **2008**, *4*, 221–234.
(48) Evers, A.; Klebe, G. Ligand-supported homology modeling of G-protein-coupled receptor sites: Models sufficient for successful virtual screening. *Angew. Chem., Int. Ed.* **2004**, *43*, 248–251.
(49) Evers, A.; Klebe, G. Successful virtual screening for a submicromolar antagonist of the neurokinin-1 receptor based on a ligand-supported homology model. *J. Med. Chem.* **2004**, *47*, 5381–5392.
(50) Moro, S.; Deflorian, F.; Bacilieri, M.; Spalluto, G. Ligand-based homology modeling as attractive tool to inspect GPCR structural plasticity. *Curr. Pharm. Des.* **2006**, *12*, 2175–2185.
(51) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534–553.
(52) Orry, A. J. W.; Cavasotto, C. N. Ligand-docking-based homology model of the melanin-concentrating hormone 1 receptor. 231st Meeting of the American Chemical Society, Atlanta, GA, 2006.
(53) Cavasotto, C. N.; Abagyan, R. A. Protein flexibility in ligand docking and virtual screening to protein kinases. *J. Mol. Biol.* **2004**, *337*, 209–225.
(54) Cavasotto, C. N.; Kovacs, J. A.; Abagyan, R. A. Representing receptor flexibility in ligand docking through relevant normal modes. *J. Am. Chem. Soc.* **2005**, *127*, 9632–9640.
(55) Kovacs, J. A.; Cavasotto, C. N.; Abagyan, R. A. Conformational sampling of protein flexibility in generalized coordinates: Application to ligand docking. *J. Comp. Theor. Nanosci.* **2005**, *2*, 354–361.
(56) Monti, M. C.; Casapullo, A.; Cavasotto, C. N.; Napolitano, A.; Riccio, R. Scalaradial, a dialdehyde-containing marine metabolite that causes an unexpected noncovalent PLA(2) inactivation. *ChemBioChem* **2007**, *8*, 1585–1591.
(57) Monti, M. C.; Casapullo, A.; Cavasotto, C. N.; Tosco, A.; Dal Piaz, F.; Ziemys, A.; Margarucci, L.; Riccio, R. The binding mode of petrosaspongiolide M to the human group IIA phospholipase A(2): Exploring the role of covalent and noncovalent interactions in the inhibition process. *Chem.—Eur. J.* **2009**, *15*, 1155–1163.
(58) Diaz, P.; Phatak, S. S.; Xu, J.; Astruc-Diaz, F.; Cavasotto, C. N.; Naguib, M. 6-Methoxy-*N*-alkyl isatin acylhydrazone derivatives as a novel series of potent selective cannabinoid receptor 2 inverse agonists: Design, synthesis and binding mode prediction. *J. Med. Chem.* **2009**, *52*, 433–444.
(59) Diaz, P.; Phatak, S. S.; Xu, J.; Fronczek, F. R.; Astruc-Diaz, F.; Thompson, C. M.; Cavasotto, C. N.; Naguib, M. 2,3-Dihydro-1-benzofuran derivatives as a series of potent selective cannabinoid receptor 2 agonists: Design, synthesis, and binding mode prediction through ligand-steered modeling. *ChemMedChem* **2009**, *4*, 1615–1629.
(60) McRobb, F. M.; Capuano, B.; Crosby, I. T.; Chalmers, D. K.; Yuriev, E. Homology modeling and docking evaluation of aminergic G protein-coupled receptors. *J. Chem. Inf. Model.* **2010**, *50*, 626–637.
(61) Kimura, S. R.; Tebben, A. J.; Langley, D. R. Expanding GPCR homology model binding sites via a balloon potential: A molecular dynamics refinement approach. *Proteins* **2008**, *71*, 1919–1929.
(62) Varady, J.; Wu, X.; Fang, X.; Min, J.; Hu, Z.; Levant, B.; Wang, S. Molecular modeling of the three-dimensional structure of dopamine 3 (D3) subtype receptor: Discovery of novel and potent D3 ligands through a hybrid pharmacophore- and structure-based database searching approach. *J. Med. Chem.* **2003**, *46*, 4377–4392.
(63) Durdagi, S.; Papadopoulos, M. G.; Zoumpoulakis, P. G.; Koukoulitsa, C.; Mavromoustakos, T. A computational study on cannabinoid receptors and potent bioactive cannabinoid ligands: Homology modeling, docking, de novo drug design and molecular dynamics analysis. *Mol. Diversity* **2010**, *14*, 257–276.
(64) Apostolakis, J.; Pluckthun, A.; Caflisch, A. Docking small ligands in flexible binding sites. *J. Comput. Chem.* **1998**, *19*, 21–37.
(65) Wesson, L.; Eisenberg, D. Atomic solvation parameters applied to molecular dynamics of proteins in solution. *Protein Sci.* **1992**, *1*, 227–235.
(66) Nemethy, G.; Gibson, K. D.; Palmer, K. A.; Yoon, C. N.; Paterlini, M. G.; Zagari, A.; Rumsey, S.; Scheraga, H. A. Energy parameters in polypeptides. 10. Improved geometrical parameters and nonbonded interactions for use in the ECEPP/3 algorithm, with application to proline-containing peptides. *J. Phys. Chem.* **1992**, *96*, 6472–6484.
(67) *ICM. Version 3.6*; MolSoft, LLC: La Jolla, CA, 2009.
(68) Halgren, T. Merck molecular force field I-V. *J. Comput. Chem.* **1995**, *17*, 490–641.
(69) Li, Z.; Scheraga, H. A. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6611–6615.
(70) Abagyan, R.; Argos, P. Optimal protocol and trajectory visualization for conformational searches of peptides and proteins. *J. Mol. Biol.* **1992**, *225*, 519–532.
(71) Mirzadegan, T.; Benko, G.; Filipek, S.; Palczewski, K. Sequence analyses of G-protein-coupled receptors: Similarities to rhodopsin. *Biochemistry* **2003**, *42*, 2759–2767.
(72) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
(73) Ballesteros, J.; Weinstein, H. Integrated methods for the construction of three-dimensional models of structure−function relations in G protein-coupled receptors. *Methods Neurosci.* **1995**, *25*, 366–428.
(74) Williams, G. J. Rattle: A data mining GUI for R. *The R Journal* **2009**, *1*, 45–55.
(75) 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–D912.
(76) *LigPrep. Version 2.3*; Schrödinger, LLC: New York, 2009.
(77) Irwin, J. J.; Shoichet, B. K. ZINC−A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182.
(78) *Maestro. Version 9.0*; Schrödinger, LLC: New York, 2009.
(79) 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.
(80) 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.
(81) *Glide. Version 5.5*; Schrödinger, LLC: New York, NY, 2009.
(82) 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.
(83) Kim, S. K.; Gao, Z. G.; Van Rompaey, P.; Gross, A. S.; Chen, A.; Van Calenbergh, S.; Jacobson, K. A. Modeling the adenosine receptors: Comparison of the binding domains of A2A agonists and antagonists. *J. Med. Chem.* **2003**, *46*, 4847–4859.
(84) Kim, J.; Wess, J.; van Rhee, A. M.; Schoneberg, T.; Jacobson, K. A. Site-directed mutagenesis identifies residues involved in ligand recognition in the human a(2a) adenosine receptor. *J. Biol. Chem.* **1995**, *270*, 13987–13997.
(85) Michino, M.; Abola, E.; Brooks, C. L., 3rd; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat. Rev. Drug Discovery* **2009**, *8*, 455–463.
(86) Novoa, E. M.; Ribas de Pouplana, L.; Barril, X.; Orozco, M. Ensemble docking from homology models. *J. Chem. Theory Comput.* **2010**, *6*, 2547–2557.
(87) Zhang, Y.; Devries, M. E.; Skolnick, J. Structure modeling of all identified G protein-coupled receptors in the human genome. *PLoS Comput. Biol.* **2006**, *2*, e13.

CI100285F