JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

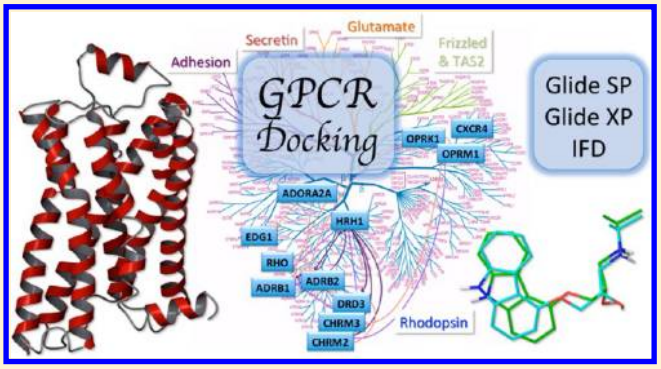
Current Assessment of Docking into GPCR Crystal Structures and Homology Models: Successes, Challenges, and Guidelines

Thijs Beuming[†,*] and Woody Sherman[†]

[†]Schrödinger, Inc., 120 West 45th Street, New York, New York, United States

Supporting Information

ABSTRACT: The growing availability of novel structures for several G protein-coupled receptors (GPCRs) has provided new opportunities for structure-based drug design of ligands against this important class of targets. Here, we report a systematic analysis of the accuracy of docking small molecules into GPCR structures and homology models using both rigid receptor (Glide SP and Glide XP) and flexible receptor (Induced Fit Docking; IFD) methods. The ability to dock ligands into different structures of the same target (cross-docking) is evaluated for both agonist and inverse agonist structures of the A2A receptor and the β1- and β2-adrenergic receptors. In addition, we have produced homology models for the β1-adrenergic, β2-adrenergic, D3 dopamine, H1 histamine, M2 muscarine, M3 muscarine, A2A adenosine, S1P1, κ-opioid, and C-X-C chemokine 4 receptors using multiple templates and investigated the ability of docking to predict the binding mode of ligands in these models. Clear correlations are observed between the docking accuracy and the similarity of the sequence of interest to the template, suggesting regimes in which docking can correctly identify ligand binding modes.



INTRODUCTION

Major advances in GPCR structural biology have led to the elucidation of several unique structures of this important family of proteins. At the time of writing, inactive state structures were available for the β1-adrenergic receptor (β1AR),[1] the β2-adrenergic receptor (β2AR),[2] the dopamine D3 receptor (D3R),[3] the histamine H1 receptor (H1R),[4] the muscarinic M2 receptor (M2R),[5] the muscarinic M3 receptor (M3R),[6] the adenosine 2A receptor (A2AR),[7] the sphingosine-1-phosphate receptor 1 (S1P1R),[8] the μ-opioid receptor (MOR),[9] the κ-opioid receptor (KOR),[10] the C-X-C chemokine receptor type 4 (CXCR4),[11] and rhodopsin.[12] In addition, for several receptor types (β1AR,[1,13,14] β2AR,[4,15−17] and A2AR[18,19]), structures with a variety of ligands have been solved, showing the extent to which GPCRs undergo induced-fit effects upon binding. Some of these structures are agonist-bound and have been solved in the active state (β2AR,[4,17] A2AR,[19,20] and rhodopsin[21,22]), showing subtle differences in binding site shape and size between agonist and antagonist/inverse agonist bound states. Thus, these structures have opened new avenues for structure-based drug design, not only for receptors where the structure is known but also for GPCRs that can be reliably modeled using available structures as templates.

Several studies have addressed the efficiency of GPCR high-throughput virtual screening exercises,[23−33] showing that both homology models and crystal structures can be effectively used. However, for drug design it is important to understand the extent to which crystal structures and homology models can be used to predict the binding mode of compounds. A further issue to address is the relationship between sequence similarity and the quality of a homology model when no crystal structure is available. Recent results of two blind GPCR modeling competitions[34,35] showed a wide range of accuracies in both ligand and protein structure predictions for several diverse GPCRs (D3R, A2AR, and CXCR4), with the accuracy depending on the availability of similar templates. For example, the first competition in 2008 with A2AR as the target and the 2010 competition with CXCR4 as the target yielded few models with RMSD under 3.0 Å, with most entries greater than 4.0 Å. On the other hand, the 2010 competition with D3 as the target resulted in more than 20 models with a RMSD of the ligand less than 2.5 Å from the crystal structure. Nonetheless, there were still many entries with RMSDs greater than 4.0 Å. While the D3 competition appeared to be easier based on the higher sequence similarity to a template, with only three data points it is not possible to determine a meaningful relationship between similarity to a template and docking accuracy. The ability to produce binding modes for agonist bound receptors from inactive state structures has been explored for A2AR[36] and β2AR.[37] Other studies have systematically studied the relationship between the choice of template and the accuracy of the homology model for GPCRs,[38] with limited investigation on

Received: August 30, 2012
Published: November 3, 2012

   dx.doi.org/10.1021/ci300411b | J. Chem. Inf. Model. 2012, 52, 3263−3277

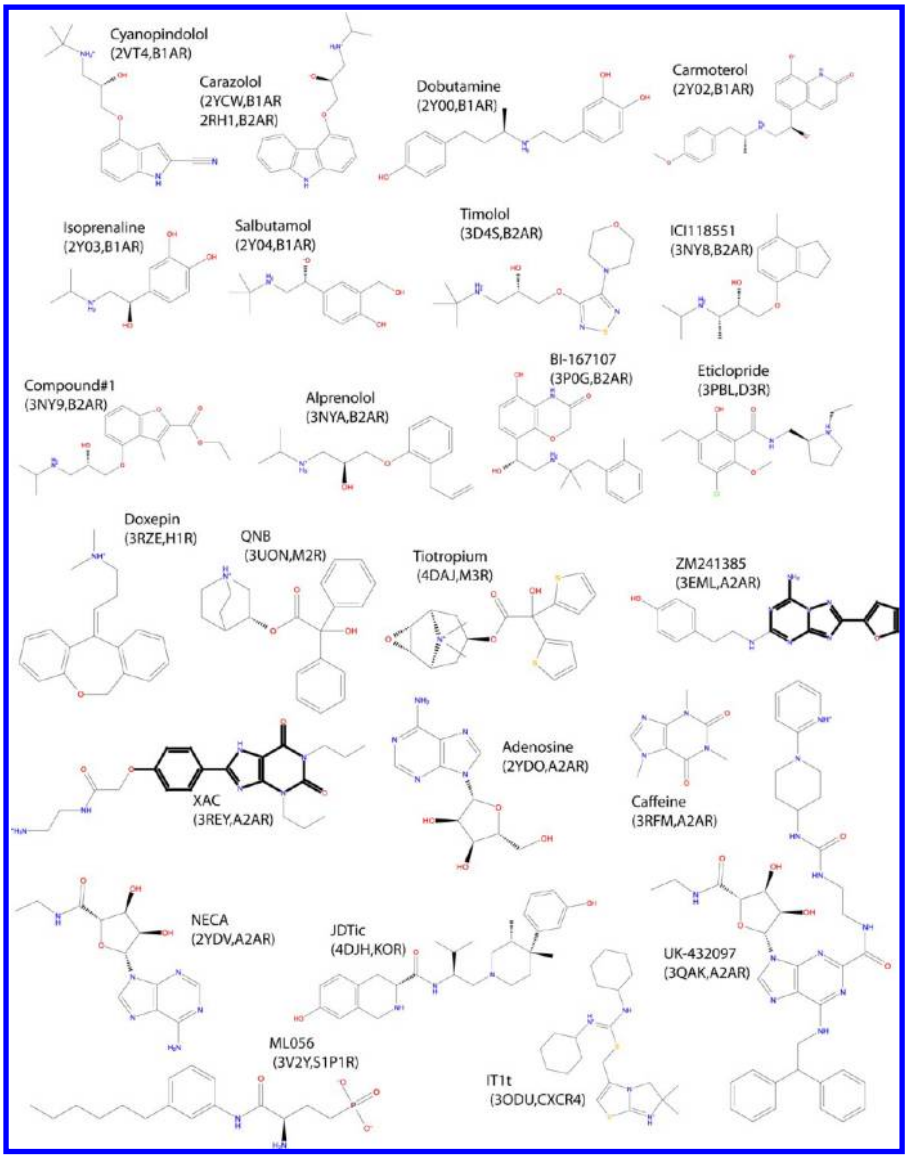


Figure 1. Ligands used in this study. The corresponding PDB code and receptor are shown in parentheses. The cores of A2AR ligands XAC and ZM241385 used for RMSD analysis are shown as bold lines (see text for details).

the impact on pose prediction accuracy for a single type of receptor (β2AR).

Here, we have expanded upon these previous studies by systematically modeling a large number of noncovalently bound GPCRs using each of the available GPCR crystal structures (at the time of writing) as a template. We assess the extent to which these models allow accurate prediction of ligand binding modes using both rigid receptor (Glide[39−41]) and flexible receptor (IFD[42]) docking strategies. These docking methods have been successfully evaluated for use with GPCRs[43,44] as well as many other targets.[42,45] These docking algorithms are applied by cognate redocking and cross-docking into the GPCR crystal structures. Then, we evaluate the ability of these programs to predict the binding mode of compounds using homology models of GPCRs. A strong correlation between template/target similarity and docking success is found, and representation of the results on a phylogenetic tree suggests regimes of applicability for model-based structure prediction given the currently available set of templates.

RESULTS

The ability of docking programs Glide[46] and IFD[42] to predict binding modes of GPCR ligands was investigated at three separate levels of increasing difficulty. We first redock the representative set of compounds (Figure 1) into their cognate receptor structures. Then, for GPCR targets with multiple available crystal structures with different ligands (i.e., β1AR, β2AR, and A2AR) we performed cross-docking (i.e., non-cognate docking). The cross-docking work included docking inverse agonists into active state structures and (partial) agonists into inactive state structures. Finally, we generated homology models of the majority of GPCRs for which the structure is known, using each of them in turn as templates. Docking accuracy was assessed for these models.

Cognate Redocking. The results for the cognate redocking of compounds into GPCR crystal structures using Glide SP, Glide XP, and IFD calculations are shown in the heat map in Figure 2 (RMSD values are reported in Supporting Information Table S1). Cells are split into RMSD for the top ranking pose (top left) and RMSD for the best pose from the maximum of

dx.doi.org/10.1021/ci300411b | J. Chem. Inf. Model. 2012, 52, 3263−3277

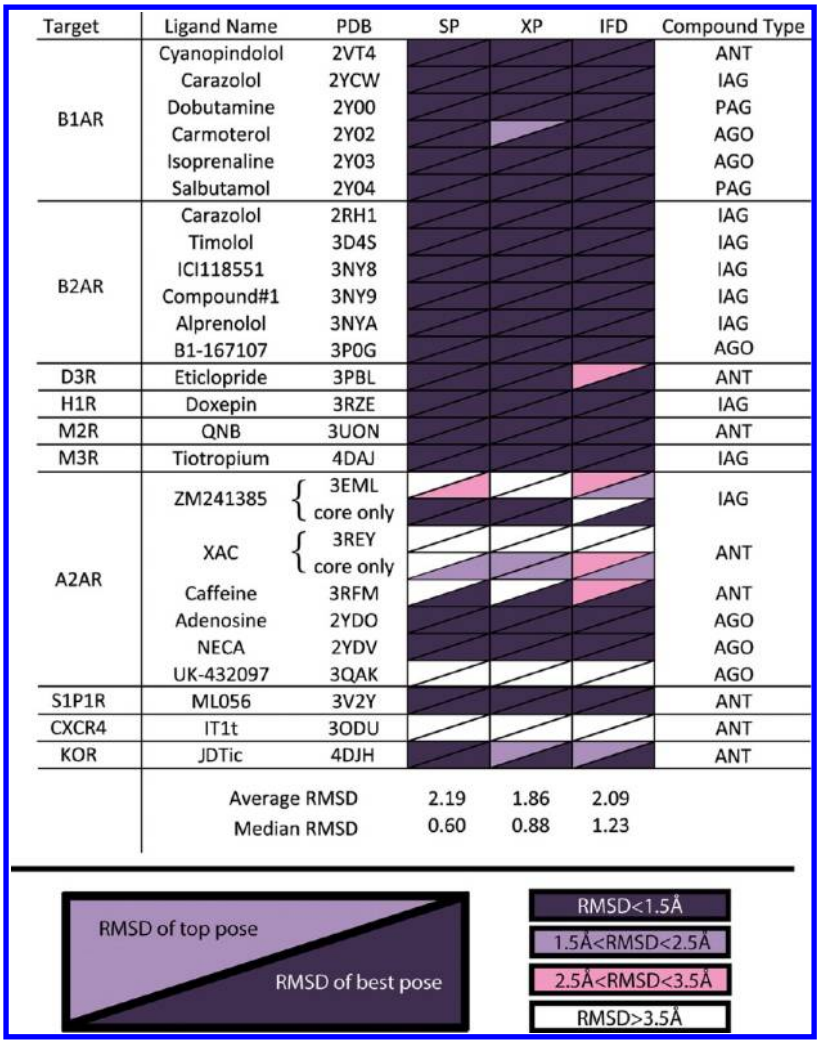


Figure 2. Cognate redocking into GPCR crystal structures. Each cell contains two heavy-atom RMSD values: the RMSD of the top pose and the RMSD of the best pose (top-left and bottom-right halves, respectively). Color-coding is described in the legend. For the A2AR ligands ZM241385 and XAC, the RMSD of the rigid core (indicated in bold in Figure 1) is reported as well. Compound type abbreviations are IAG, inverse agonist; ANT, antagonist; PAG, partial agonist; and AGO, agonist. RMSD values are shown in Supporting Information Table S1.

20 poses analyzed (bottom right). The cells are colored on the basis of the quality of the docking results, with high accuracy predictions (<1.5 Å) colored dark purple, medium accuracy predictions (1.5−2.5 Å) colored light purple, and low accuracy predictions (2.5−3.5 Å) colored pink. Cells where there is no reasonable prediction (RMSD > 3.5 Å) are not colored. On average, Glide SP, Glide XP, and IFD calculations perform with similar accuracy for cognate redocking, with Glide XP having the lowest average RMSD of top poses (1.9 Å) and Glide SP having the lowest median RMSD (0.5 Å).

For most of the complexes, highly accurate predictions are possible with all three methods. Among the exceptions are A2AR blockers ZM241385 and XAC, which have flexible solvent-exposed groups with high B-factors that are difficult to predict with high accuracy. The cores of the molecules (indicated in bold in Figure 1) are significantly more accurate, with the best results obtained using Glide XP (0.24 Å for ZM241385 and 2.01 Å for XAC). Another issue that possibly complicates the prediction of A2AR complexes is the abundance of water molecules in the binding sites of several of the structures (e.g., 3EML and 2YDO). In these structures, the water molecules, which were removed prior to docking in this study, form direct interactions with the ligands in a nonconserved way that can impact pose-prediction accuracy in the case of ZM241385, as has been noted and discussed previously in detail by others.[29,44] It is likely that for other compounds the results presented here could be improved with the selective inclusion of certain water molecules, although that was not considered in this work and will be the focus of a future publication. In another challenging case, no good poses could be found for UK-432097 bound to A2AR, likely due to the large size of the molecule (57 heavy atoms) with a significant amount of flexibility (18 rotatable bonds and two saturated rings), although the lack of inclusion of explicit water molecules could also play a role here. In the case of IT1t bound to CXCR4, several problems contribute to the poor docking accuracy. First, the CXCR4 binding site is the largest of all receptors studied here, as revealed by SiteMap[47−49] calculations, which complicates the sampling problem significantly. In addition, the ligand has an intramolecular stacking arrangement resulting in a U-shaped ligand conformation, which is not identified in the ConfGen[50] conformational search algorithm that is used in the Glide docking protocol, thereby eliminating the possibility of finding good poses in the docking

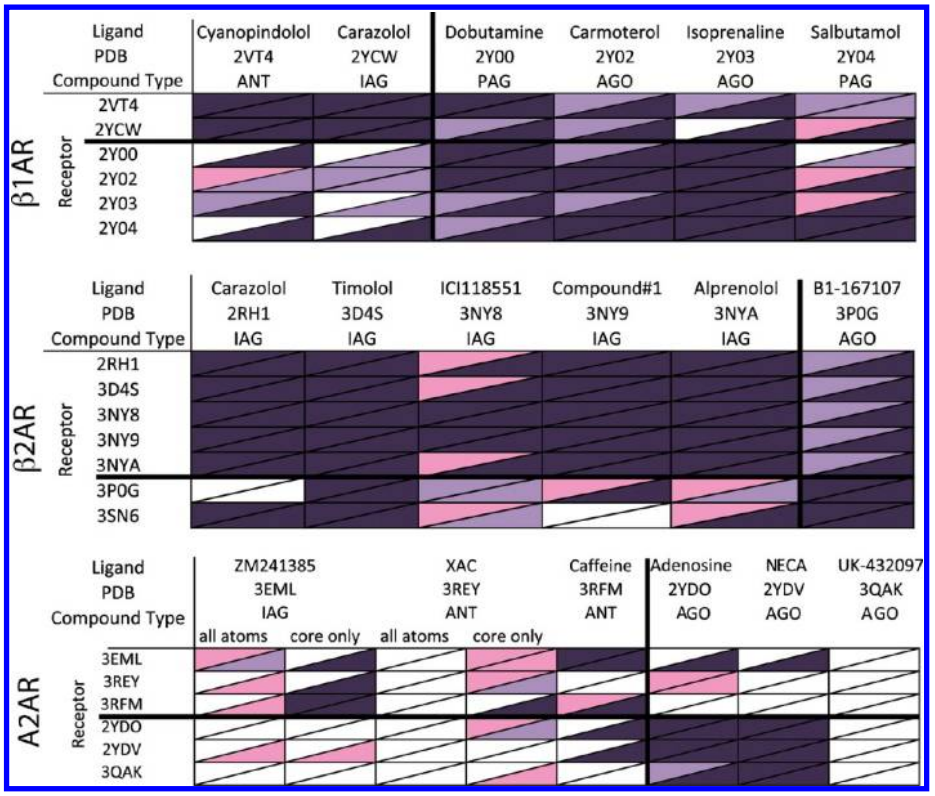


Figure 3. Cross-docking into β1AR (top), β2AR (middle), and A2AR (bottom) structures. Ligands from structures on the *x*-axis are docked into the receptor structures on the *y*-axis. Color-coding is described in Figure 2. For each of the three targets, the thick horizontal and vertical lines separate inverse agonists and antagonists (left, top) from partial and full agonists (right, bottom). Actual RMSD values are shown in Supporting Information Tables S2, S3, and S4.

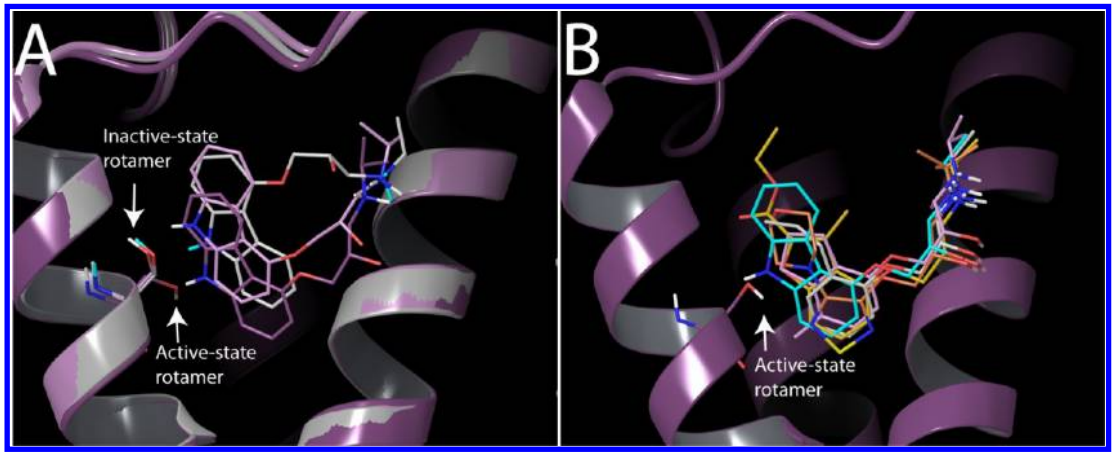


Figure 4. Incorrect cross-docking of inverse agonists into agonist–bound adrenergic GPCR structures. (A) In β1AR, different rotameric states of Ser211 (5.42) between agonist and inverse agonist bound structures lead to an incorrect top-pose prediction for carazolol in 2Y00 (white) and a medium-RMSD prediction in 2Y02 (light purple), while the redocked pose in 2YCW is highly accurate (dark purple). (B) In β2AR, a similar effect is observed for carazolol, but not the other inverse agonists studied here. Superposition of all inverse agonist structures on the 3P0G active state structure shows that this is due to the large overlap of carazolol (cyan) compared to other blockers with the active state rotamer of Ser203 (5.42). TMs 1, 2, 6, and 7 have been omitted for clarity.

stage. Indeed, when the molecule is docked rigidly into the receptor, it is possible to identify a highly accurate pose (see Supporting Information Figure S1). Finally, the prediction of the D3R/eticlopride complex using IFD ranks an incorrect pose first, while the best scoring pose (0.29 Å) is ranked third (Supporting Information Table S1). This problem with scoring can be resolved by an MM-GBSA postprocessing calculation on the IFD poses (data not shown), suggesting potential improvements to the IFD score that will be explored in future work.

Cross-Docking. A more difficult problem in docking is the prediction of binding modes of compounds using crystal structures derived in the presence of another compound or without any ligand present (cross-docking). This becomes increasingly difficult when large structural changes of the receptor occur upon binding. Here, we have performed cross-docking calculations to targets for which multiple complexes

dx.doi.org/10.1021/ci300411b | J. Chem. Inf. Model. 2012, 52, 3263–3277

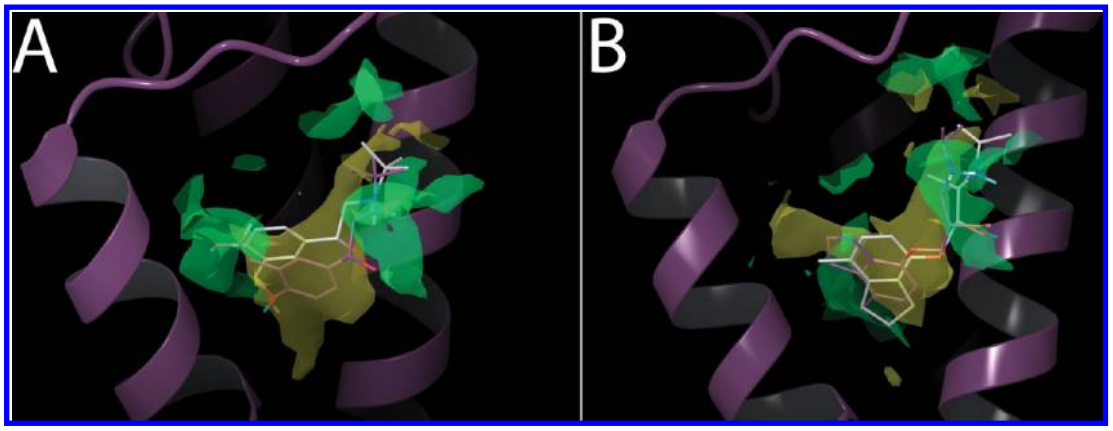


Figure 5. Flipped binding modes of partial agonists in β1AR and β2AR structures. Hydrophobic (yellow) and hydrophilic (green) binding site volumes were calculated with SiteMap.[47−49] (A) Docking of β1AR partial agonist salbutamol from 2Y04 into 2Y02 leads to a flipped pose (white) compared to the crystal structure (2Y04; purple). (B) Docking of β2AR partial agonist ICI118551 from 3NY8 into 3NYA leads to a flipped pose (white) compared to the crystal structure (3NY8; purple). TMs 1, 2, 6, and 7 have been omitted for clarity.

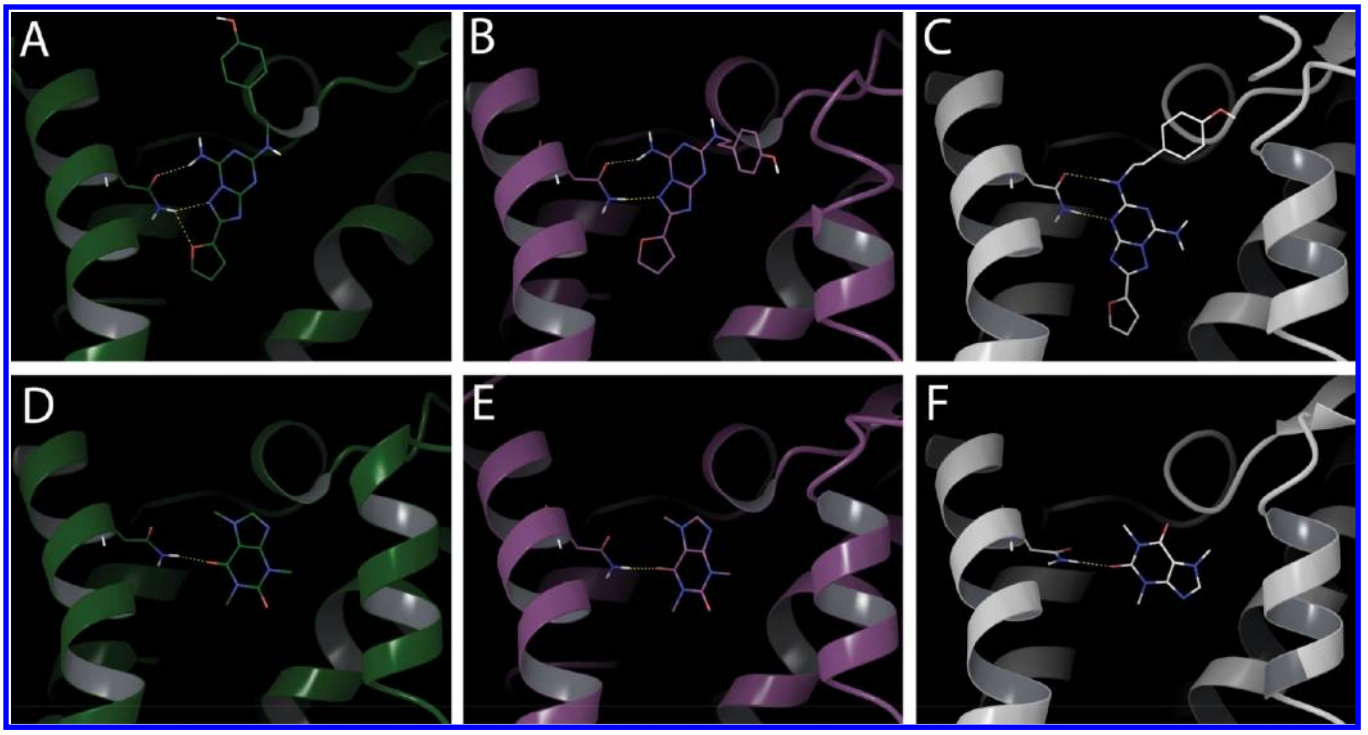


Figure 6. Prediction of ZM241385 (top) and caffeine (bottom) binding modes. Crystal structures are shown in green (A and D). For ZM241385, docking into the 3RFM structure correctly positions the core of the compound (B) while docking into 2YDV favors an incorrect hydrogen bonding pattern between Asn253 (6.55) and the ligand (C). For caffeine, docking into 3EML produced an accurate pose (E), while docking into 2YDO predicts a pose that is rotated by 90° in the binding site (F).

are available, namely β1AR, β2AR, and A2AR. As is the case for cognate redocking, results for all three methods are similar, with an average RMSD of 2.64 Å for Glide SP, 3.00 Å for Glide XP, and 2.55 Å for IFD (docking results for 3QAK, which cannot be docked with any of the methods, are not included in these averages). For a total of 108 docking calculations, Glide SP predicts 59% of all cases within 1.5 Å, 65% of all cases within 2.5 Å, and 71% of all cases within 3.5 Å. For Glide XP, these values are 50%, 63%, and 69%, while for IFD they are 44%, 64%, and 73%. In addition, in many cases where the top pose is not within 3.5 Å, the full set of poses contains at least one pose <3.5 Å, especially in the case of Glide SP (86%) and IFD (86%). Henceforth, only results using IFD (which produced the lowest average RMSD and has the most predictions <3.5 Å for cross-docking) are discussed and shown as heat maps in Figure 3 for β1AR, β2AR, and A2AR. RMSD values for all IFD, SP, and XP calculations are shown in Supporting Information Tables S2, S3, and S4.

For both β1AR and β2AR, cross-docking of antagonists and inverse agonists into structures solved with blockers (top left quadrant) is always highly accurate, while cross-docking into structures solved with (partial) agonists (bottom left quadrant) is more difficult. Indeed, the average RMSD values for the binding site residues when comparing inactive or active structures among themselves are significantly lower (0.29 Å ± 0.12) than when comparing active with inactive structures (0.63 Å ± 0.15). As has been observed previously,[37,51] this is due to the contracted size of the active-state binding site

dx.doi.org/10.1021/ci300411b | J. Chem. Inf. Model. 2012, 52, 3263−3277

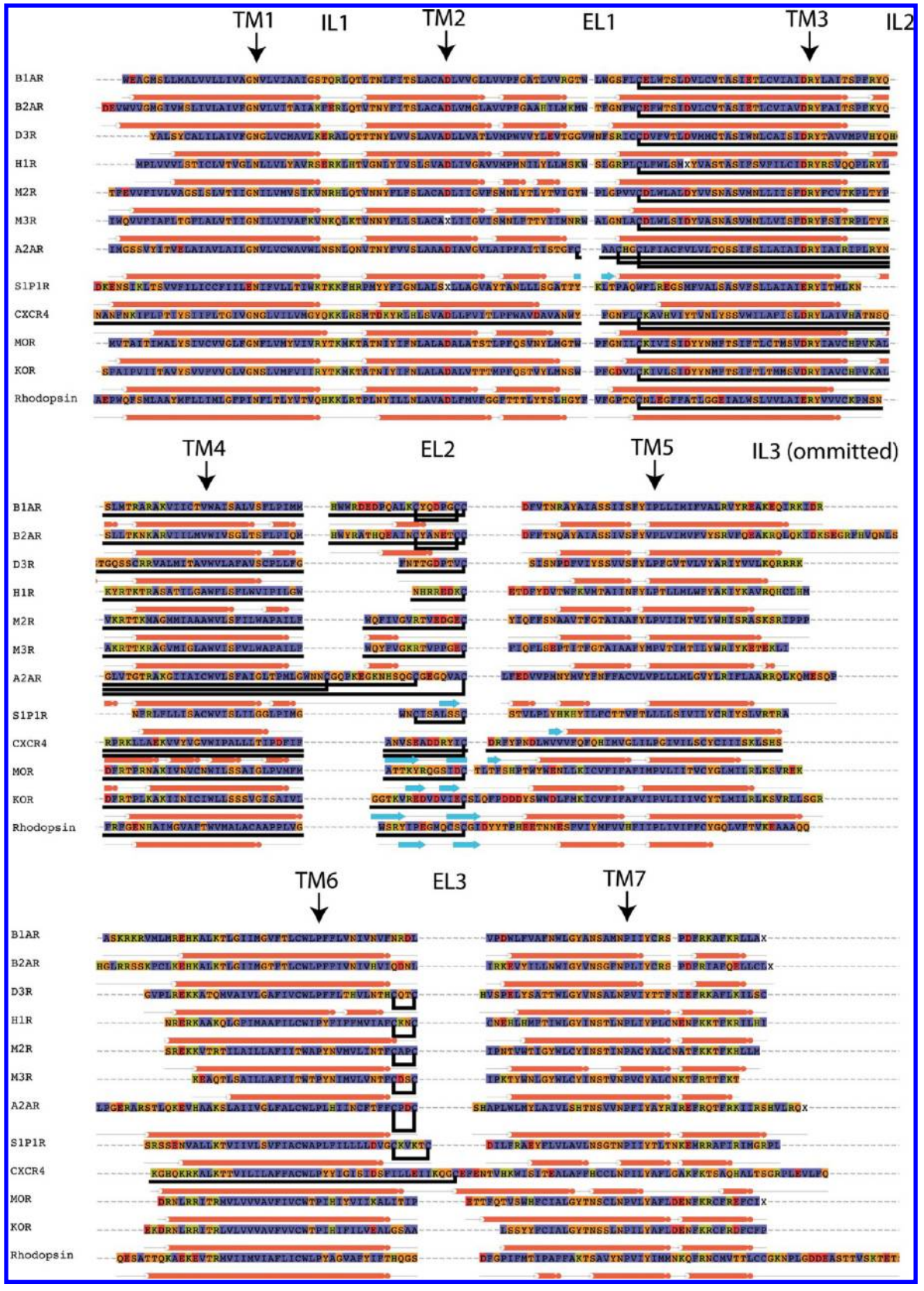


Figure 7. Sequence alignment of the available GPCR crystal structures. Secondary structure is represented by orange bars (helices) and cyan arrows (strands). Cysteine residues involved in disulfide bonds are connected by black lines. Transmembrane (TM), extra-cellular loops (EL), and intracellular loops (IL) are indicated above the alignment. For each TM, the conserved Ballesteros-Weinstein x.50 residues are indicated with an arrow. These sequences represent the residues present in the construct used in structure determination (with the T4 lysozyme removed) and are not equivalent to the biologically relevant sequence. Differences include truncated N- and C-termini and a deletion in IL3. Figure produced using the Multiple Sequence Viewer in Maestro 9.3.

dx.doi.org/10.1021/ci300411b | J. Chem. Inf. Model. 2012, 52, 3263−3277

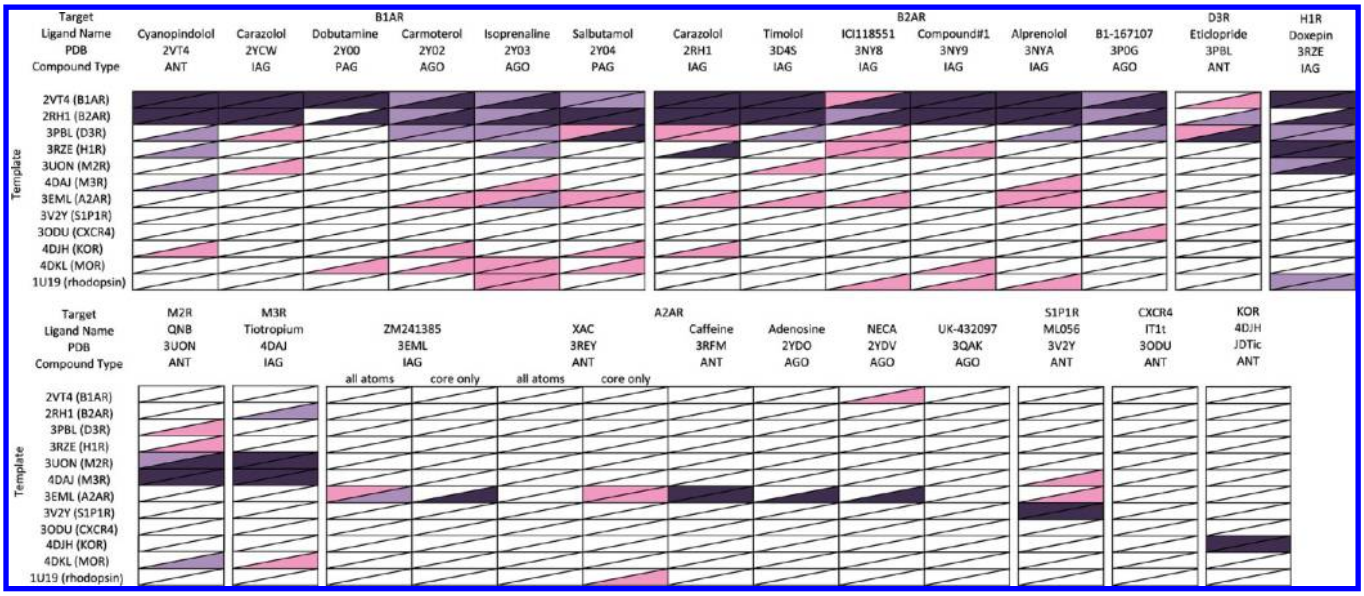


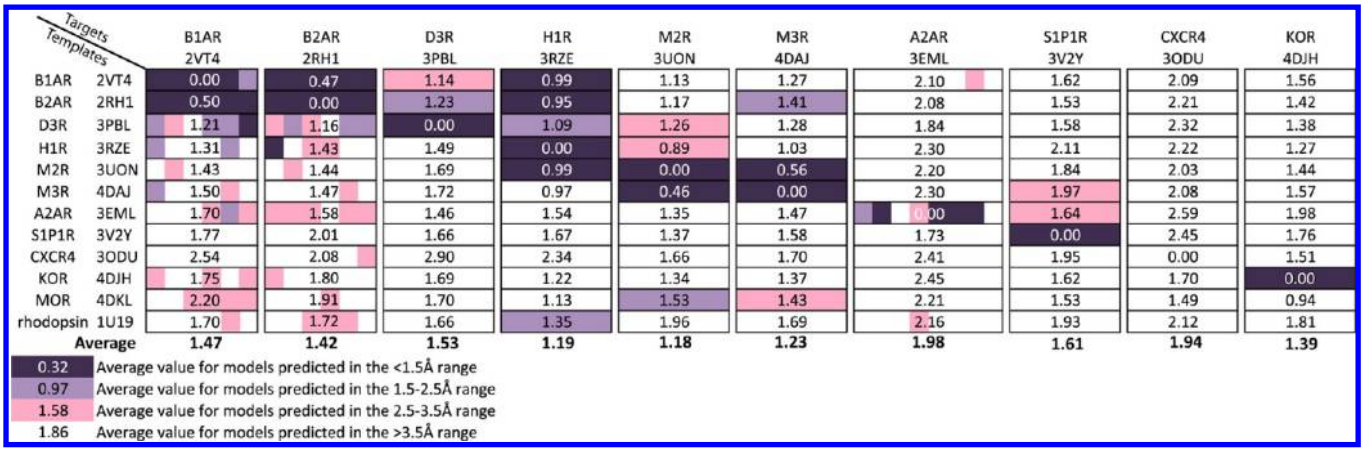
Figure 8. Docking of GPCR ligands into homology models of β1AR, β2AR, D3R, H1R, M2R, M3R, A2AR, S1P1R, CXCR4, and KOR using IFD. Color coding is described in Figure 2. The template structures used are indicated on the *y*-axis and include covalently bound structures for MOR (4DKL) and rhodopsin (1U19).

compared to the inactive state. The difference is most pronounced for Ser5.42 (Ballesteros and Weinstein numbering[52]), which adopts a different rotamer in the active state, and this can strongly affect the prediction of binding modes of inverse agonists. For example, in the case of carazolol docked into active β1AR structures, the active-state Ser rotamer stabilizes two distinct conformations, one high RMSD conformation where the entire carazolol molecule is flipped with respect to the membrane (Figure 4A, white model) and a medium RMSD pose where the ligand is shifted ~3 Å toward the extracellular side of the pocket (Figure 4A, light purple model). Docking of carazolol into the β2AR active state structure 3P0G is not possible for the same reason, but interestingly, all other β2AR inverse agonists studied here can be docked with medium to high accuracy into 3P0G. Indeed, superimposing structures 2RH1 and 3P0G shows a sterically incompatible distance <1 Å between carazolol and Ser203 (5.42; Figure 4B, carazolol shown in cyan), while other inverse agonists have significantly less overlap. Although the IFD method is designed to overcome these types of induced-fit effects by sampling of side-chain rotamers, in the case of carazolol docked into 3P0G the magnitude of the clash prevents an approximate solution to be found among the initial poses (see Experimental Section), thereby preventing correct sampling of Ser203 (5.42) during the second stage of IFD. Interestingly, we observe that an accurate pose of carazolol is identified when the 3SN6 active-state structure is used for docking. This structure differs from 3P0G in that several of the side chains in EL2 are not resolved and are represented only by CB atoms. This leverages an option within the IFD protocol to mutate key amino acids to alanine in the initial docking stage, and reinserting them automatically in the optimization stage of the IFD protocol. The absence of residues in EL2 allows the initial poses to come sufficiently close to TM5 to force Ser203 (5.42) to sample a less energetically favorable rotamer during the optimization stage, which in turn allows for accurate binding mode prediction of carazolol.

Docking of partial and full agonists into all available structures is possible with medium to high accuracy, although salbutamol docked to β1AR and ICI118551 docked to β2AR are more difficult. This is the result of a high degree of the internal symmetry in these molecules, together with a rather flat shape of the ligand binding site that can lead to stabilized "flipped" binding modes, where the ring systems are rotated 180° with respect to the cognate structure (see Figure 5).

As was the case for the cognate redocking experiment, prediction of A2A poses in cross-docking is of significantly lower accuracy than for the two beta receptors. However, similarly to cross-docking in β1AR and β2AR, it is easier to cross-dock agonists into active structures and antagonists into inactive structures. This is consistent with the lower binding-site RMSD values within either inactive or active structures (0.57 ± 0.37) than between active and inactive A2A structures (1.11 ± 0.24). The primary difficulty is in the correct prediction of the Asn253 (6.55) hydrogen bond with the ligand, which is complicated by the availability of several additional donor and acceptor atoms in most of the compounds (see Figure 6A–C for an example of a correctly and incorrectly predicted hydrogen bonding pattern for ZM241385), and the flexibility that Asn253 displays among the different crystal structures (as reflected by the high B-factors for the noncore part of the ligand). Finally, the A2A binding site is larger than that of the beta-receptors, offering more accessible space for the ligands. While water molecules in most crystal structures occupy this space, the nonconserved nature of the water molecules makes it hard to use them in cross-docking experiments. However, including explicit water molecules[29] and/or a constraint with Asn253 (6.55)[44] has been shown to dramatically improve prediction for some of the A2A ligands. Here, we do not explore the treatment of explicit water molecules during docking, although other docking programs have an option to treat the presence/absence of water molecules explicitly during docking.[53] We will explore the explicit treatment of water molecules more exhaustively in a future publication. In this study, docking of ZM241385 is possible with all inactive state structures (3EML, 3REY, and 3RFM) to the extent that the core of the molecule is predicted accurately for at least one pose in the ensemble. 3REY docking is even more problematic, with

dx.doi.org/10.1021/ci300411b | J. Chem. Inf. Model. 2012, 52, 3263−3277



Templates \ Targets		B1AR 2VT4	B2AR 2RH1	D3R 3PBL	H1R 3RZE	M2R 3UON	M3R 4DAJ	A2AR 3EML	S1P1R 3V2Y	CXCR4 3ODU	KOR 4DJH
B1AR	2VT4	0.00	0.47	1.14	0.99	1.13	1.27	2.10	1.62	2.09	1.56
B2AR	2RH1	0.50	0.00	1.23	0.95	1.17	1.41	2.08	1.53	2.21	1.42
D3R	3PBL	1.21	1.16	0.00	1.09	1.26	1.28	1.84	1.58	2.32	1.38
H1R	3RZE	1.31	1.43	1.49	0.00	0.89	1.03	2.30	2.11	2.22	1.27
M2R	3UON	1.43	1.44	1.69	0.99	0.00	0.56	2.20	1.84	2.03	1.44
M3R	4DAJ	1.50	1.47	1.72	0.97	0.46	0.00	2.30	1.97	2.08	1.57
A2AR	3EML	1.70	1.58	1.46	1.54	1.35	1.47	0.00	1.64	2.59	1.98
S1P1R	3V2Y	1.77	2.01	1.66	1.67	1.37	1.58	1.73	0.00	2.45	1.76
CXCR4	3ODU	2.54	2.08	2.90	2.34	1.66	1.70	2.41	1.95	0.00	1.51
KOR	4DJH	1.75	1.80	1.69	1.22	1.34	1.37	2.45	1.62	1.70	0.00
MOR	4DKL	2.20	1.91	1.70	1.13	1.53	1.43	2.21	1.53	1.49	0.94
rhodopsin	1U19	1.70	1.72	1.66	1.35	1.96	1.69	2.16	1.93	2.12	1.81
Average		1.47	1.42	1.53	1.19	1.18	1.23	1.98	1.61	1.94	1.39

Figure 9. Cα-RMSD values for superposition of binding sites. Representative target structures (top) were superimposed on all the templates that were used (left). For β1AR, β2AR, and A2AR, the multiple cells represent the results for the multiple available ligands used in Figure 8. Color coding is based on the best-pose RMSD (bottom right cell in Figure 8). For most targets, especially the aminergic GPCRs, there is a clear correlation between low binding-site RMSD and high docking accuracy.

accurate poses for the core obtained only using the 3RFM (caffeine-bound) structure.

Caffeine is the smallest molecule studied here, and correct poses can be identified using structures 3EML, 3RFM, 2YDO, and 2YDV. Alternative binding modes that are related by a 90° rotation of the molecule are predicted as top poses in the majority of structures (see Figure 6F). For adenosine and NECA, docking is accurate in all agonist-bound structures (2YDO, 2YDV, and 3QAK) and to some extent 3EML. The main difference between the agonist-bound structures and the inappropriate models (3REY and 3RFM) is found in TM3, where a significant shift of the helix moves residues Val84 (3.32), Ile85 (3.33), and Thr88 (3.36) closer to the binding site. Interestingly, 3EML has an intermediate shift of TM3, which allows for poses to be predicted with medium accuracy. Finally, as discussed above, the 3QAK ligand cannot be docked accurately, presumably due to its high degree of conformational flexibility (18 rotatable bonds and two saturated rings).

Finally, and most relevant to projects when a crystal structure of the GPCR of interest is not available, we evaluated the ability of Glide and IFD to predict the binding mode of compounds to homology models. Models of 10 proteins for which a noncovalently ligand bound structure was solved (β1AR, β2AR, D3R, H1R, M2R, M3R, A2AR, S1P1R, KOR, and CXCR4) were predicted using these 10 proteins plus covalently bound GPCR structures for rhodopsin and MOR as templates. The pairwise alignments between template and targets were taken from the global alignment of solved GPCR structures shown in Figure 7. These target−template pairs cover a wide range of pairwise sequence identities (see Supporting Information Table S11), ranging from highly similar pairs (e.g., M2R/M3R or KOR/MOR, both 68%) to very remotely related receptors (e.g., the KOR/S1P1R pair has the lowest value at 15% identity).

A total of 120 homology models were produced; whole-protein Cα RMSD values ranged from 1.1 Å (β1AR modeled using β2AR) to 3.5 Å (CXCR4 modeled using H1R). This range reduces to 0.6−3.1 Å when only the TM domains are considered. All ligands listed in Figure 1 were docked into the homology models of the target with which they were crystallized, producing a total of 300 complexes, of which 48% are dockings for the β1AR and β2AR. Docking was performed using IFD, Glide SP, and Glide XP. The results for IFD are shown in the heat map in Figure 8. Full numeric results for all three methods are reported in Supporting Information Tables S6, S7, and S8. Loop structures were obtained directly from the initial Prime model building step, and no effort was made to perform loop refinement, making it likely that loops with incorrect structures exist that would prevent successful docking. More sophisticated loop sampling has been successful at predicting GPCR loops in difficult cases;[54] however, applying computationally intensive loop sampling (often requiring many CPU-days of computer time) to this number of homology models is beyond the scope of this work. In order to explore the possible negative implications of a loop being placed into the binding site and precluding the correct ligand binding mode, a second set of models was generated by deleting all loops. Compounds were then docked into these TM-only models using IFD (see Supporting Information Table S9). For models that included the loops, results obtained using IFD were slightly better than those using SP, which in turn were more accurate than XP. In general, Glide XP is more sensitive to the accuracy of the receptor structure and requires a structural ensemble to achieve the best results in docking to noncognate structures.[55] For IFD, 14.5% of docking calculations have a best pose under 1.5 Å, 20.7% under 2.5 Å, and 34.1% under 3.5 Å. For SP, the comparable numbers are 12.3%, 18.5%, and 33.7%, and for XP they are 9.8%, 12.7%, and 19.6% (see Supporting Information Table S10 for a summary of method performance). Finally, removing the loops from the models did not result in an overall improvement for the best performing method (IFD) when averaging over the entire data set; however, a number of cases got better while others got worse. A closer inspection of the data revealed that most of the cases where removing loops improved results were for non-beta adrenergic targets. Indeed, comparing the effect of removing loops for the subset of the data that excluded those two targets showed a significant improvement in terms of prediction accuracy (Supporting Information Table S10B). These improvements were offset by the beta-adrenergic receptor results, which comprise 48% of the cross-docking cases, and show a degradation in the results when the loops are removed. This is likely a consequence of good loop placement in the case of close homologues and suggests that retaining loops is beneficial for docking when they can be accurately modeled.

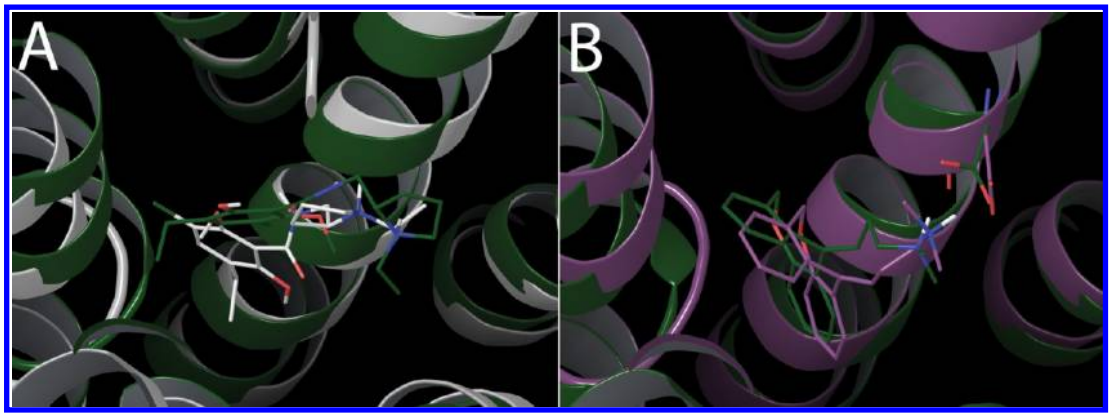


Figure 10. (A) Unsuccessful prediction (RMSD 5.0 Å) of the eticlopride binding mode in a D3R model based on the H1R template. Model in white, crystal structure in green. (B) Successful prediction (RMSD 1.5 Å) of the doxepin binding mode in a H1R model based on the D3R template. Model in purple, crystal structure in green.

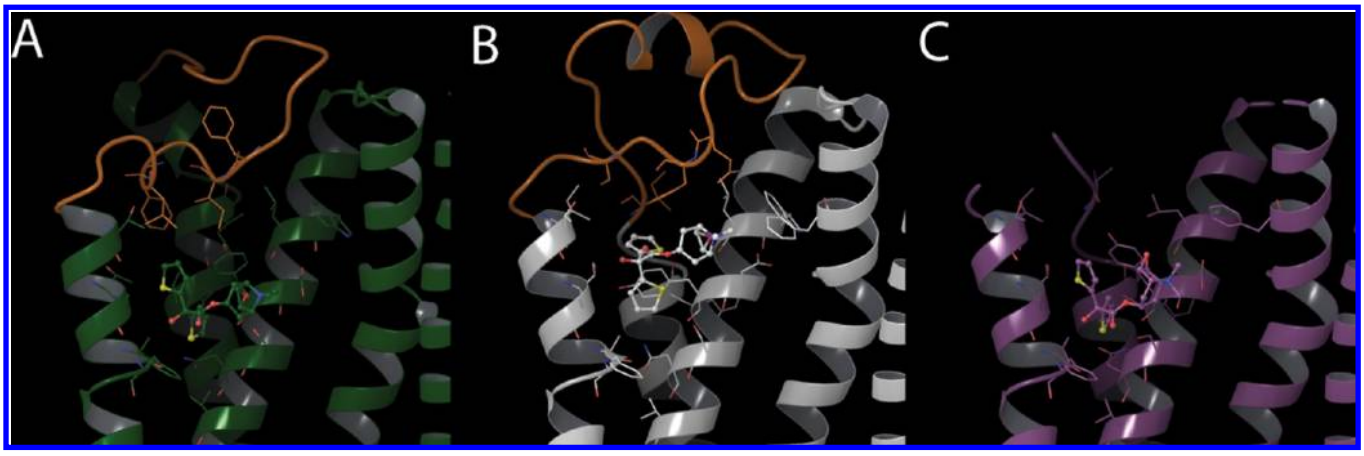


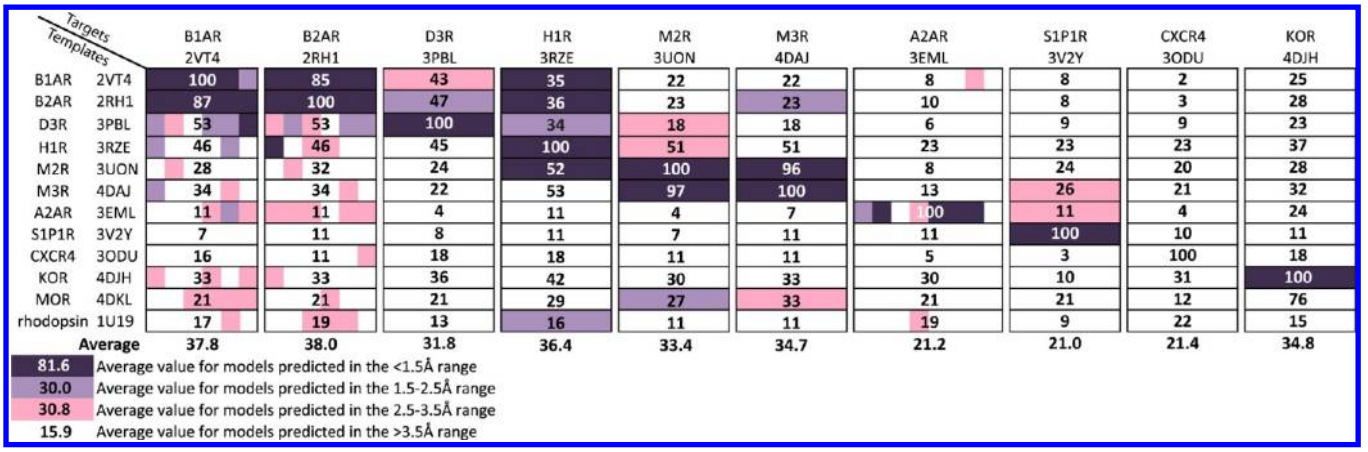
Figure 11. Effects of incorrectly predicted loops on pose accuracy. (A) crystal structure of tiotropium bound to M3R (4DAJ) interacting with residues in EL2 (orange). (B) Top scoring pose (RMSD 4.82 Å) in a model of M3R based on 2RH1 with an incorrect EL2 loop (orange); note the introduction of incorrect interactions due to a shift of the ligand toward the loop. (C) Top scoring pose (RMSD 1.85 Å) in a model of M3R based on 2RH1 with loops removed. Only TM residues are available for interaction, and most of the native interactions are correctly predicted.

To facilitate the structural interpretation of the results, structures were compared in terms of binding site RMSD (see Figure 9). Template/target pairs with small RMSD values should result in accurate homology models, and so docking into these structures should be relatively easier than cases with less accurate homology models. Indeed, for cases with a binding site Cα RMSD less than 1.0 Å, 87% of cases have a best pose under 2.5 Å. Increasing the binding site RMSD cutoff to 1.5 Å results in 45% of dockings under 2.5 Å (compared to 20.7% for the entire data set), thus confirming the relationship between accurate homology modeling and the ability to correctly predict binding modes via docking.

Docking into models of the β1AR and β2AR based on β2AR and β1AR templates, respectively, is almost as accurate as cognate redocking and cross-docking into the crystal structures of these receptors, which is encouraging for cases where a closely related GPCR structure to the target of interest is available. In addition, the D3R template is good for both beta-receptors, as there are a number of ligands with good poses within the docked ensemble, and even a few where the top scoring pose is accurate. For a small number of ligands (e.g., cyanopindolol and isoprenaline in β1AR or carazolol in β2AR), the H1 receptor is a good template, while the M3 receptor is an appropriate template for docking of cyanopindolol into β1AR. Many borderline accurate results can be found for other compounds using these aminergic structures as templates. In addition, for the more remotely related templates A2AR, the opioid receptors KOR and MOR, and rhodopsin, several borderline accurate poses for some of the compounds can be found in the ensemble.

The D3R/eticlopride complex is a harder target for docking than the beta-receptors, as indicated by the difficulty in obtaining a good top pose in the cognate redocking experiment. Indeed, the only templates that allow poses within 3.5 Å of the crystal structure to be identified are the two beta-receptors. This is in contrast to the H1R/doxepin complex, which can be modeled with medium to high accuracy using β1AR, β2AR, D3R, M2R, and even rhodopsin (see Figure 10 for docking results of D3R and H1R using these structures as templates). Comparing the pairwise binding site RMSD values for these four aminergic targets, it is clear that D3R is significantly less similar to the other GPCR templates than are the beta receptors and H1R. So, in addition to the fact that eticlopride itself is a harder molecule to dock than the beta receptor ligands and doxepin, the lower accuracy for the D3R/eticlopride models results in part from a substantial structural difference with the templates.

dx.doi.org/10.1021/ci300411b | J. Chem. Inf. Model. 2012, 52, 3263−3277



Targets / Templates		B1AR 2VT4	B2AR 2RH1	D3R 3PBL	H1R 3RZE	M2R 3UON	M3R 4DAJ	A2AR 3EML	S1P1R 3V2Y	CXCR4 3ODU	KOR 4DJH
B1AR	2VT4	100	85	43	35	22	22	8	8	2	25
B2AR	2RH1	87	100	47	36	23	23	10	8	3	28
D3R	3PBL	53	53	100	34	18	18	6	9	9	23
H1R	3RZE	46	46	45	100	51	51	23	23	23	37
M2R	3UON	28	32	24	52	100	96	8	24	20	28
M3R	4DAJ	34	34	22	53	97	100	13	26	21	32
A2AR	3EML	11	11	4	11	4	7	100	11	4	24
S1P1R	3V2Y	7	11	8	11	7	11	11	100	10	11
CXCR4	3ODU	16	11	18	18	11	11	5	3	100	18
KOR	4DJH	33	33	36	42	30	33	30	10	31	100
MOR	4DKL	21	21	21	29	27	33	21	21	12	76
rhodopsin	1U19	17	19	13	16	11	11	19	9	22	15
Average		37.8	38.0	31.8	36.4	33.4	34.7	21.2	21.0	21.4	34.8

81.6 Average value for models predicted in the <1.5Å range
30.0 Average value for models predicted in the 1.5-2.5Å range
30.8 Average value for models predicted in the 2.5-3.5Å range
15.9 Average value for models predicted in the >3.5Å range

Figure 12. Sequence identity of binding site residues between targets and templates, based on the alignment in Figure 7. All residues within 4 Å of the template ligand are included in the sequence identity calculation. For β1AR, β2AR, and A2AR, the multiple cells represent the results for each available ligand used in Figure 8. Color coding is based on best-pose RMSD (bottom right cell in Figure 8). For most targets, especially the aminergic GPCRs, there is a correlation between high binding-site sequence identity and high docking accuracy.

The final two aminergic receptors in the data set (M2R and M3R) proved to be difficult even for cognate redocking (see Figure 2). While it is possible to build a suitable model of M2R using M3R and vice versa, only a few other templates produced sufficiently accurate models. For M2R, the templates that produced low to medium accurate poses were MOR, H1R, and D3R, and for M3R, the best templates were β2AR and MOR. In contrast to D3R, where the challenges lie with the structural differences between the templates and D3R, the lower accuracy for M2R and M3R is a consequence of the inherent difficulty of predicting the M2R/QNB and M3R/tiotropium complexes, as the binding sites are relatively similar to other aminergic GPCRs (Figure 9). This difficulty is due to the formation of a bifurcated H-bond with Asn507 (6.52), which is highly sensitive to the rotameric state of the side chain and the internal conformation of the ligand. Indeed, an IFD calculation with a rigid docking Glide stage using the most similar template (H1R) leads to highly accurate predictions (see Supporting Information Figure S2). Interestingly, removing the loops from the muscarinergic models significantly improves the prediction accuracy (compare Supporting Information Tables S6 and S9). In cases where the presence of loops negatively affects pose prediction, this is not necessarily due to steric hindrance of the loops but rather due to stabilization of incorrect poses away from the binding site due to interactions with incorrectly placed loop elements (for example see Figure 11).

Given the difficulty in predicting the binding mode of A2AR ligands into A2AR crystal structures, it is not surprising that when using homology models, virtually no successful predictions of A2A ligands are possible. The only exceptions where low-accuracy results were found in the IFD ensemble are docking of XAC into a model based on rhodopsin and NECA into a model based on β1AR. In addition, the S1P1R can only be modeled with low accuracy using A2AR and M3R, while the absence of any good results for CXCR4 is not surprising, as the docking methodology used here is unable to access the internal conformation of the ligand (see cognate redocking section above). Finally, while KOR and MOR receptors are reasonable templates for the beta and muscarinergic receptors studied here, no good templates are available for the docking of JDTic into models of KOR (MOR was not studied here as it is solved with a covalently bound ligand).

When comparing the docking RMSD values as shown in Figure 8 with the measures of structural similarity (binding site RMSD, Figure 9) between templates and targets, it is clear that there is a strong correlation between these measures and the probability of a successful docking experiment (R^2 = 0.41, see Supporting Information Figure S3). Thus, most successful pairs have binding site RMSD values smaller than 1.0 Å. Medium and low accuracy pairs typically have binding site RMSD values between 1.0 and 1.5 Å, and pairs where docking was not successful have RMSD values >1.5 Å. For example, the GPCR complex that is easiest to model is H1R/doxepin, which has relatively low binding site RMSD values with multiple templates. On the other hand, D3R has much higher RMSD values with most templates, and homology model docking results are on average much worse than H1R even though cognate redocking works well. The exceptions to this intuitive finding, as mentioned before, are the muscarinergic receptors, which have low RMSDs to multiple templates, but low docking accuracy due to the complicated nature of the complex.

Sequence Analysis. The docking benchmark and structural comparison of targets and templates described above indicates that a high degree of structural similarity is necessary for reliable unguided homology-model based docking experiments within the GPCR family. To assess whether suitable target/template pairs could be identified using sequence alone, we measured the pairwise sequence identity of the binding site residues of all GPCRs studied here. To allow for a prospective comparison, we identified the residues within 4 Å of the template ligand and determined the identity with their counterparts in the aligned target. Results are shown in Figure 12.

The data in Figure 12 show that most successfully predicted complexes were modeled on templates with binding site sequence identities of 30% and higher. While there are exceptions to this rule (for example, the use of opioid receptor templates for beta and muscarinic receptors), the results of blind, unconstrained docking using templates with lower similarities than 30% should be evaluated with caution. On the other hand, the prediction of target/template residue identity can be calculated in the absence of a structure for the target, and hence the criteria for successful prediction identified here can be used in prospective modeling projects.

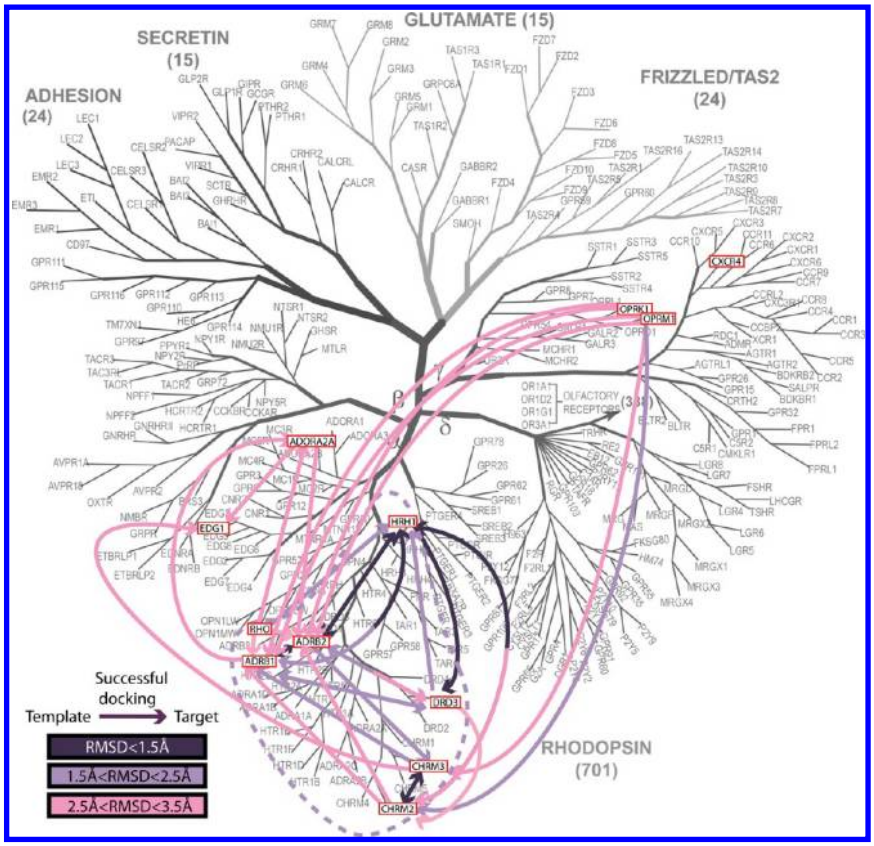


Figure 13. Phylogenetic tree (based on layout in ref 57) of the GPCR family indicating successful docking predictions carried out in this study. Gene names used are ADRB1 (β1AR), ADRB2 (β2AR), DRD3 (D3R), HRH1 (H1R), CHRM2 (M2R), CHRM3 (M3R), ADORA2A (A2AR), EDG1 (S1P1R), CXCR4 (CXCR4), OPRK1 (KOR), OPRM1 (MOR), and RHO (rhodopsin). Arrows are from template to target, with coloring based on the best-pose RMSD (lower-left cells) from Figure 8. For example, A2AR can be used as a low-accuracy template for S1P1R, but not *vice versa*. H1R is a medium-accurate template for β1AR (light purple half a rrow), while β1AR is a high-accuracy template for H1R (dark purple half arrow). In cases where multiple ligands were available, the best result among all compounds was used to color the arrow. Note the high density of the light and dark purple arrow (many good docking results) among the members of central section of the α-branch.

In addition to this pairwise comparison, we also indicated the relationships between targets and templates on a phylogenetic tree[56] of the GPCR family (see Figure 13). Here, arrows indicate the successful use of a template to model a target complex, with colors corresponding to those in the previous figures. For targets within the "middle" section of the α-branch (indicated with a dashed purple line), sufficient suitable templates exist for reliable prediction using homology modeling and IFD.

■ DISCUSSION

This study aimed to provide benchmarks for the prediction of rigid and flexible receptor docking into both crystal structures and homology models of GPCRs using a comprehensive set of GPCR crystal structures. Calculations were performed using Glide SP, Glide XP, and IFD with default settings and no manual intervention. As such, the results presented here represent a worst-case scenario for real projects, as the inclusion of experimental information tends to either improve or in the worst case generally does not alter the predictions of an automated protocol. Indeed, none of the top scoring groups in the GPCR Dock competitions[34,35] (including ours) used a fully automated protocol. Nonetheless, the results presented here represent an objective measure of the accuracy of these calculations and can be used to estimate the accuracy of real-world docking applications to crystal structure and/or homology model-based GPCR projects when no experimental information is known about the target of interest. The strength of a retrospective study using the entire available data set of GPCR structures as compared with true prospective applications of docking programs to novel targets in the GPCR Dock competitions is the larger number of examples and therefore a broader range of applicability to new projects. In addition, for the majority of models, the accuracy of docking has been estimated on the basis of a single compound, and for targets for which multiple ligands are available, the variation of accuracies suggests that these results will change when additional compounds are considered. Consequently, the relationships between targets highlighted in the phylogenetic tree in Figure 13 are incomplete and will have to be re-evaluated as new GPCR structures become available.

To test docking accuracy, we first evaluated the ability to repredict the binding mode of ligands to their cognate crystal structures. Accurate cognate redocking represents a necessary but insufficient condition for real projects, as the structure of the receptor already conforms to the bound ligand. While still a challenging problem, the docking methods explored here (Glide SP, Glide XP, and IFD) produced excellent cognate redocking results (~80% top scoring poses under 1.5 Å), with the exception of a few cases where the ligand is either very flexible or the internal geometry is challenging for the conformational sampling algorithm to identify.

dx.doi.org/10.1021/ci300411b | J. Chem. Inf. Model. 2012, 52, 3263−3277

A more challenging and practical test is the cross-docking of compounds into crystal structures solved with other ligands. Here, we only studied the receptors with several cocrystallized ligands. For beta receptors, a high level of accuracy was obtained, on par with the cognate redocking results, with the noted exception of docking antagonists into the smaller agonist binding site. The difficulty of such calculations has been noted previously and can be addressed to some extent with the use of structural interaction fingerprints[51,58] or increasing local flexibility.[37] A more complicated case is the set of A2AR compounds, which are challenging even for cognate redocking, and where docking calculations have been shown to depend heavily on the presence of explicit water molecules in the binding site.[29] The accurate prediction of water positions, orientations, and energetics during the docking calculations would likely improve the A2AR results but is beyond the scope of this work. However, previous work suggests that explicit solvent simulations that compute water molecule locations and energetics can explain complex SAR in A2AR.[59]

The most challenging and relevant task in GPCR docking is the application to homology models. Here, as expected, the results are significantly worse than docking into crystal structures. However, consistent successes can be identified between closely related subtypes of GPCRs (β1AR/β2AR or M2R/M3R) and among the larger aminergic class of GPCRs in general. On the basis of three prospective GPCR docking results (GPCR Dock competition) for targets with close (D3R), intermediate (A2AR), and distant (CXCR4) available templates, sequence identity boundaries of 35−40% for reliable homology based pose prediction have been previously suggested.[34] The values obtained in this study further define those boundaries and show that binding site sequence identity as low as 30% can be sufficient for reliable predictions, especially if nontemplate related issues (for example, insufficient ligand conformational sampling) are addressed. On the basis of this criterion and from an inspection of the docking relationship between targets plotted on a phylogenetic tree of the GPCR family, the majority of receptors in the rhodopsin (class A) α-branch now have an appropriate template in the Protein Data Bank.

All models generated here were based on a single template. The use of multiple templates to generate homology models is a common approach, and this has been systematically evaluated for GPCRs as well.[60] It was found that for situations where only distant templates are available, the use of multiple templates can improve the accuracy of the model, but when a similar template was available, no improvement was observed. Given the difficulty with docking into models based on distant templates observed here, it is not likely that multiple template modeling will lead to a significant improvement for most cases.

An important part of homology modeling is the refinement of loops. In the case of GPCRs, the second extra-cellular loop (EL2) often forms an important part of the binding site, and the structural difference among known structures is large.[61] Thus, the simple template-derived loop structures are inaccurate for many of the models produced here, and it is likely that this has adversely affected the docking results presented here. While prediction of long GPCR loops in crystal structures was shown to be highly successful in previous studies,[62] the application to homology models is not straightforward, and exploring this direction is beyond the scope of this docking benchmark. However, we did attempt to investigate the effects of the error in the loops by docking into models with the loops entirely removed, and we found an improvement for 46% of the models. However, in the cases of close homologues (e.g., the beta-receptors) where the modeled loops were more accurate, the removal of the loops degraded docking results. This suggests that proper loop placement, especially for EL2, is important for docking to GPCR homology models for cases where the loop position cannot be determined from the template and that some form of improved loop prediction will be needed in order to increase the general accuracy of docking into GPCR homology models. Given the structural differences between most of the templates, it is likely that *de novo* loop predictions will continue to present challenges, although experimental constraints can often be used to guide the prediction of loops and improve the accuracy. Indeed, experimental data indicating which residues were part of the binding site in the D2/D3 receptor family[63] were used successfully in the prospective modeling of EL2 in a D3 receptor model based on the β2AR template (2010 GPCR Dock competition,[34] Beuming et al., entry 3041).

In general, the best results for cross-docking or docking to homology models comes when incorporating other information, either from experiments (e.g., mutagenesis) or additional computational approaches (e.g., pharmacophore modeling). Indeed, the literature contains examples of successful docking studies using distant templates where the use of a non-automated protocol with additional constraints was needed to obtain good results. For example, the modeling procedure used here was unable to identify a good pose for the CXCR4/It1t complex, but successful predictions for this target were reported in the 2010 GPCR Dock competition. Interestingly, the most accurate predictions from that study (Vaidehi et al., entry 2560; Roumen et al., entry 1006) both involved a modification of the template-derived model to optimize the orientation of TM2, bringing residue D97 (2.63) into the binding site. This residue forms a hydrogen bond with the ligand in the crystal structure but points away from the binding site in most of the models built using the automated protocol presented here. In cases where an appropriate pose could be identified in the IFD ensemble, but where this pose was not the highest ranked, other methods or knowledge-based criteria might help identify the correct solution. In the case of the model of D3 based on the β2AR template (Beuming et al., entry 3041), the closest model to the crystal structure was correctly identified by comparing the IFD ensemble with results from WaterMap analysis[64] and pharmacophore modeling using Phase.[65] Similarly, in the case of the optimal submission from the 2008 GPCR Dock competition,[35] the correct solution for the A2AR-ZM241385 model was found by performing iterative IFD calculations using a constraint with residue Asn253 (6.55), which was correctly predicted to form a hydrogen bond with the ligand (Costanzi, personal communication). In all of these cases, accurate modeling of the ligand−receptor complex was facilitated by the availability of experimental data, in particular from mutagenesis experiments. Thus, while the regime of applicability of model-based predictions will increase with every new GPCR structure that is solved, and with increasing accuracy of the available modeling and docking algorithms, indirect structural approaches will remain important for understanding this important class of proteins.

EXPERIMENTAL SECTION

Structures. GPCR structures for β1AR (2VT4, 2YCW, 2Y00, 2Y02, 2Y03, 2Y04), β2AR (2RH1, 3D4S, 3NY8, 3NY9,

3NYA, 3P0G), D3R (3PBL), H1R (3RZE), M2R (3UON), M3R (4DAJ), A2AR (3EML, 3REY, 3RFM, 2YDO, 2YDV, 3QAK), S1P1R (3V2Y), CXCR4 (3ODU), KOR (4DJH), MOR (4DKL), and rhodopsin (1U19) were retrieved from the Protein Data Bank (PDB).[66] Structures were prepared with the Protein Preparation Wizard in Maestro 9.3.[67] When present, the T4-lysozyme (T4L) insertion in intracellular loop 3 (IL3) was removed from the structure. All ligands were isolated from the structure and assigned bond orders and protonation states as shown in Figure 1.

Sequence Alignment. GPCR sequences were extracted from the PDB files using the Multiple Sequence Viewer in Maestro 9.3. Sequences were aligned using ClustalW[68] and manually refined to ensure correct alignment of the loops. Insertions and deletions were placed at the central residues of the loops, with the exception of EL2, where the conserved cysteine involved in a disulfide bridge with Cys 3.25 in TM3 was aligned, when present. Pairwise sequence identities for the binding site region were defined using all residues within 4 Å of the ligand in the template. All possible pairs of alignments between targets (10) and templates (12) were extracted from the multiple sequence alignment. Gap-only positions were removed to yield the final pairwise alignments.

Homology Modeling. Homology models were generated using Prime[69] using the prepared template structures and the pairwise alignments described above. For cases where multiple structures existed for a template, the first published representative structure was chosen, namely 2VT4, 2RH1, and 3EML for β1AR, β2AR, and A2AR, respectively. Models were built using default Prime settings, including the template ligand in the initial model building. A separate set of models was prepared from these initial models that had all loops removed. The definition of loop boundaries was based on the secondary structure definitions in the PDB files.

Docking. Docking calculations were performed using Glide[46] SP,[40,41] Glide XP,[39] and Induced-Fit Docking (IFD).[42] For all methods, the docking region (grid) was centered on the template ligand in the model with default box sizes. Glide docks flexible ligands into a rigid receptor structure by sampling of the conformational, orientational, and positional degrees of freedom of the ligand. The SP and XP modes of Glide differ in how ligand degrees of freedom are sampled and in the scoring function employed. Both modes generate a large number of conformations for a ligand followed by a series of hierarchical filters to enable rapid evaluation of ligand poses. Glide XP begins with Glide SP poses and then performs further refinement using an anchor-and-grow algorithm to more thoroughly sample ligand degrees of freedom. In addition, Glide XP has a more sophisticated scoring function that includes a number of energetic penalty terms that are not in the SP scoring function. The IFD protocol is a multistep workflow that consist of (1) an initial Glide SP docking using a softened potential (scaling of van der Waals radii to 0.7 and 0.5 for receptor and ligand heavy atoms, respectively), (2) Prime refinement (side-chain prediction and minimization) for each protein−ligand complex of all residues within 5 Å of any ligand pose, (3) Glide SP redocking into each refined receptor structure using the default potential (van der Waals radii scaling of 1.0 and 0.8 for receptor and ligand, respectively), and (4) scoring of each pose with a linear combination of 0.05 × Prime energy plus the GlideScore (IFDScore).

RMSD Calculations. RMSD values compared to the X-ray ligand were calculated for all non-hydrogen atoms using a Python script that utilizes the Schrödinger Python API for handling atom matching and symmetry. RMSD values were based on all non-hydrogen atoms of the ligands. In the case of the A2AR ligands ZM241385 and XAC, which have a highly flexible solvent-exposed tails, a separate RMSD value was calculated that only took into account the non-hydrogen atoms of the heterocyclic core (shown as bold lines in Figure 1). RMSD values were binned into high accuracy (<1.5 Å), medium accuracy (1.5−2.5 Å), and low accuracy (2.5−3.5 Å). These bins have been colored dark purple, light purple, and pink throughout the figures in the manuscript. RMSD values are reported for both the lowest energy pose (*top pose*), as well as the most accurate pose (*best pose*) in the ensemble. The Cα RMSD of the binding site residues shown in Figure 9 was calculated using the Align Binding Sites tool in Maestro 9.3.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Full RMSD results for all methods and benchmarks are available as Tables S1−S9. A comparison of the performance of all methods with and without loops is shown in Table S10. Pairwise sequence identities for the full sequences are shown in Table S11. Rigid docking example poses are provided in Figures S1 and S2. A correlation plot between docking RMSD and binding site RMSD is shown in Figure S3. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

PDB ID Codes: 2VT4, 2YCW, 2Y00, 2Y02, 2Y03, 2Y04, 2RH1, 3D4S, 3NY8, 3NY9, 3NYA, 3P0G, 3PBL, 3RZE, 3UON, 4DAJ, 3EML, 3REY, 3RFM, 2YDO, 2YDV, 3QAK, 3V2Y, 3ODU, 4DHJ, 4DKL

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: thijs.beuming@schrodinger.com. Phone: 212 295 5800. Fax: 212 295 5801.

Present Address

Schrödinger, Inc., 120 West Forty-Fifth Street, New York, NY, 10036

Author Contributions

T.B. carried out the research. T.B. and W.S wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Stefano Costanzi for helpful discussions, Ramy Farid for reviewing the manuscript, and Ray Stevens and Angela Walker for providing the template of the phylogenetic tree used in Figure 13.

■ ABBREVIATIONS USED

β1AR, β1-adrenergic receptor; β2AR, β2-adrenergic receptor; D3R, dopamine D3 receptor; H1R, histamine H1 receptor; M2R, muscarinic M2 receptor; M3R, muscarinic M3 receptor; A2AR, adenosine 2A receptor; S1P1R, sphingosine-1-phosphate receptor 1; MOR, μ-opioid receptor; KOR, κ-opioid receptor; CXCR, C-X-C chemokine receptor type 4; SP, Standard Precision; XP, Extra Precision; IFD, Induced Fit Docking; PDB, Protein Data Bank; RMSD, root-mean-square deviation; PAG, partial agonists; AGO, agonist; ANT, antagonist; IAG, inverse agonist

REFERENCES

(1) 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.

(2) 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.

(3) Chien, E. Y.; Liu, W.; Zhao, Q.; Katritch, V.; Han, G. W.; Hanson, M. A.; Shi, L.; Newman, A. H.; Javitch, J. A.; Cherezov, V.; Stevens, R. C. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **2010**, *330*, 1091.

(4) Rasmussen, S. G.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **2011**, *477*, 549.

(5) Haga, K.; Kruse, A. C.; Asada, H.; Yurugi-Kobayashi, T.; Shiroishi, M.; Zhang, C.; Weis, W. I.; Okada, T.; Kobilka, B. K.; Haga, T.; Kobayashi, T. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* **2012**, *482*, 547.

(6) Kruse, A. C.; Hu, J.; Pan, A. C.; Arlow, D. H.; Rosenbaum, D. M.; Rosemond, E.; Green, H. F.; Liu, T.; Chae, P. S.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Wess, J.; Kobilka, B. K. Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* **2012**, *482*, 552.

(7) 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.

(8) Hanson, M. A.; Roth, C. B.; Jo, E.; Griffith, M. T.; Scott, F. L.; Reinhart, G.; Desale, H.; Clemons, B.; Cahalan, S. M.; Schuerer, S. C.; Sanna, M. G.; Han, G. W.; Kuhn, P.; Rosen, H.; Stevens, R. C. Crystal structure of a lipid G protein-coupled receptor. *Science* **2012**, *335*, 851.

(9) Manglik, A.; Kruse, A. C.; Kobilka, T. S.; Thian, F. S.; Mathiesen, J. M.; Sunahara, R. K.; Pardo, L.; Weis, W. I.; Kobilka, B. K.; Granier, S. Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature* **2012**.

(10) Wu, H.; Wacker, D.; Mileni, M.; Katritch, V.; Han, G. W.; Vardy, E.; Liu, W.; Thompson, A. A.; Huang, X. P.; Carroll, F. I.; Mascarella, S. W.; Westkaemper, R. B.; Mosier, P. D.; Roth, B. L.; Cherezov, V.; Stevens, R. C. Structure of the human kappa-opioid receptor in complex with JDTic. *Nature* **2012**.

(11) Wu, B.; Chien, E. Y.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **2010**, *330*, 1066.

(12) 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.

(13) Moukhametzianov, R.; Warne, T.; Edwards, P. C.; Serrano-Vega, M. J.; Leslie, A. G.; Tate, C. G.; Schertler, G. F. Two distinct conformations of helix 6 observed in antagonist-bound structures of a beta1-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 8228.

(14) Warne, T.; Moukhametzianov, R.; Baker, J. G.; Nehme, R.; Edwards, P. C.; Leslie, A. G.; Schertler, G. F.; Tate, C. G. The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor. *Nature* **2011**, *469*, 241.

(15) Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V. P.; Chien, E. Y.; Velasquez, J.; Kuhn, P.; Stevens, R. C. A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor. *Structure* **2008**, *16*, 897.

(16) Wacker, D.; Fenalti, G.; Brown, M. A.; Katritch, V.; Abagyan, R.; Cherezov, V.; Stevens, R. C. Conserved binding mode of human beta2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J. Am. Chem. Soc.* **2010**, *132*, 11443.

(17) Rasmussen, S. G.; Choi, H. J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; Devree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Schnapp, A.; Konetzki, I.; Sunahara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K. Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **2011**, *469*, 175.

(18) Dore, A. S.; Robertson, N.; Errey, J. C.; Ng, I.; Hollenstein, K.; Tehan, B.; Hurrell, E.; Bennett, K.; Congreve, M.; Magnani, F.; Tate, C. G.; Weir, M.; Marshall, F. H. Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure* **2011**, *19*, 1283.

(19) Xu, F.; Wu, H.; Katritch, V.; Han, G. W.; Jacobson, K. A.; Gao, Z. G.; Cherezov, V.; Stevens, R. C. Structure of an agonist-bound human A2A adenosine receptor. *Science* **2011**, *332*, 322.

(20) Lebon, G.; Warne, T.; Edwards, P. C.; Bennett, K.; Langmead, C. J.; Leslie, A. G.; Tate, C. G. Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* **2011**, *474*, 521.

(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.

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

(23) 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.

(24) 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.

(25) Costanzi, S.; Vilar, S. In Silico screening for agonists and blockers of the beta(2) adrenergic receptor: implications of inactive and activated state structures. *J. Comput. Chem.* **2012**, *33*, 561.

(26) de Graaf, C.; Kooistra, A. J.; Vischer, H. F.; Katritch, V.; Kuijer, M.; Shiroishi, M.; Iwata, S.; Shimamura, T.; Stevens, R. C.; de Esch, I. J.; Leurs, R. Crystal structure-based virtual screening for fragment-like ligands of the human histamine H(1) receptor. *J. Med. Chem.* **2011**, *54*, 8195.

(27) Vilar, S.; Ferino, G.; Phatak, S. S.; Berk, B.; Cavasotto, C. N.; Costanzi, S. Docking-based virtual screening for ligands of G protein-coupled receptors: not only crystal structures but also in silico models. *J. Mol. Graphics Modell.* **2011**, *29*, 614.

(28) Carlsson, J.; Coleman, R. G.; Setola, V.; Irwin, J. J.; Fan, H.; Schlessinger, A.; Sali, A.; Roth, B. L.; Shoichet, B. K. Ligand discovery from a dopamine D3 receptor homology model and crystal structure. *Nat. Chem. Biol.* **2011**, *7*, 769.

(29) Katritch, V.; Jaakola, V. P.; Lane, J. R.; Lin, J.; Ijzerman, A. P.; Yeager, M.; Kufareva, I.; Stevens, R. C.; Abagyan, R. Structure-based discovery of novel chemotypes for adenosine A(2A) receptor antagonists. *J. Med. Chem.* **2010**, *53*, 1799.

(30) Shoichet, B. K.; Kobilka, B. K. Structure-based drug screening for G-protein-coupled receptors. *Trends Pharmacol. Sci.* **2012**.

(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.

(32) Carlsson, J.; Yoo, L.; Gao, Z. G.; Irwin, J. J.; Shoichet, B. K.; Jacobson, K. A. Structure-based discovery of A2A adenosine receptor ligands. *J. Med. Chem.* **2010**, *53*, 3748.

(33) 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.

(34) Kufareva, I.; Rueda, M.; Katritch, V.; Stevens, R. C.; Abagyan, R. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **2011**, *19*, 1108.

(35) 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.
(36) Katritch, V.; Abagyan, R. GPCR agonist binding revealed by modeling and crystallography. *Trends Pharmacol. Sci.* **2011**, *32*, 637.
(37) Vilar, S.; Karpiak, J.; Berk, B.; Costanzi, S. In silico analysis of the binding of agonists and blockers to the beta2-adrenergic receptor. *J. Mol. Graphics Modell.* **2011**, *29*, 809.
(38) Filizola, M.; Weinstein, H. The structure and dynamics of GPCR oligomers: a new focus in models of cell-signaling mechanisms and drug design. *Curr. Opin. Drug. Discovery Dev.* **2005**, *8*, 577.
(39) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177.
(40) 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.
(41) 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.
(42) 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.
(43) 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.
(44) 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.
(45) Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **2004**, *56*, 235.
(46) *Glide*, 5.8 ed.; Schrödinger, LLC: New York, NY, 2012.
(47) *SiteMap*, 2.6 ed.; Schrödinger, LLC: New York, NY, 2012.
(48) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377.
(49) Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des.* **2007**, *69*, 146.
(50) Watts, K. S.; Dalal, P.; Murphy, R. B.; Sherman, W.; Friesner, R. A.; Shelley, J. C. ConfGen: a conformational search method for efficient generation of bioactive conformers. *J. Chem. Inf. Model.* **2010**, *50*, 534.
(51) de Graaf, C.; Rognan, D. Selective structure-based virtual screening for full and partial agonists of the beta2 adrenergic receptor. *J. Med. Chem.* **2008**, *51*, 4978.
(52) Ballesteros, J. A.; Weinstein, H. In *Methods in Neurosciences*; Stuart, C. S., Ed.; Academic Press: 1995; Vol. *25*, p 366.
(53) Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W.; Taylor, R. D.; Taylor, R. Modeling water molecules in protein-ligand docking using GOLD. *J. Med. Chem.* **2005**, *48*, 6504.
(54) Goldfeld, D. A.; Zhu, K.; Beuming, T.; Friesner, R. A. Successful prediction of the intra- and extracellular loops of four G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 8275.
(55) Repasky, M.; Murphy, R.; Banks, J.; Greenwood, J.; Tubert-Brohman, I.; Bhat, S.; Friesner, R. Docking performance of the glide program as evaluated on the Astex and DUD datasets: a complete set of glide SP results and selected results for a new scoring function integrating WaterMap and glide. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 787.
(56) Fredriksson, R.; Lagerström, M. C.; Lundin, L.-G.; Schiöth, H. B. The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Mol. Pharmacol.* **2003**, *63*, 1256.
(57) Katritch, V.; Cherezov, V.; Stevens, R. C. Diversity and modularity of G protein-coupled receptor structures. *Trends Pharmacol. Sci.* **2012**, *33*, 17.
(58) Deng, Z.; Chuaqui, C.; Singh, J. Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein-ligand binding interactions. *J. Med. Chem.* **2004**, *47*, 337.
(59) Higgs, C.; Beuming, T.; Sherman, W. Hydration Site Thermodynamics Explain SARs for Triazolylpurines Analogues Binding to the A2A Receptor. *ACS Med. Chem. Lett.* **2010**, *1*, 160.
(60) 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.
(61) Peeters, M. C.; van Westen, G. J. P.; Li, Q.; Ijzerman, A. P. Importance of the extracellular loops in G protein-coupled receptors for ligand recognition and receptor activation. *Trends Pharmacol. Sci.* **2011**, *32*, 35.
(62) Goldfeld, D. A.; Zhu, K.; Beuming, T.; Friesner, R. A. Successful prediction of the intra- and extracellular loops of four G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 8275.
(63) Shi, L.; Javitch, J. A. The second extracellular loop of the dopamine D2 receptor lines the binding-site crevice. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 440.
(64) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817.
(65) *Phase*, 3.4 ed.; Schrödinger, LLC: New York, NY, 2012.
(66) Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* **2003**, *10*, 980.
(67) *Maestro*, 9.3 ed.; Schrödinger, LLC: New York, NY, 2012.
(68) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947.
(69) *Prime*, 3.1 ed.; Schrödinger, LLC: New York, NY, 2012.