Advances in GPCR modeling evaluated by the GPCR Dock 2013 assessment: meeting new challenges

Irina Kufareva[1], **Vsevolod Katritch**[2], **participants of GPCR Dock 2013 (Box 1)**, **Raymond C. Stevens**[2,*], and **Ruben Abagyan**[1,*]

[1]UCSD Skaggs School of Pharmacy and Pharmaceutical Sciences, La Jolla, CA, 92039, USA

[2]Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA, 92037, USA

Abstract

Despite tremendous successes of GPCR crystallography, the receptors with available structures represent only a small fraction of human GPCRs. An important role of the modeling community is to maximize structural insights for the remaining receptors and complexes. The community-wide GPCR Dock assessment was established to stimulate and monitor the progress in molecular modeling and ligand docking for GPCRs. The four targets in the present third assessment round presented new and diverse challenges for modelers, including prediction of allosteric ligand interaction and activation states in 5-hydroxytryptamine receptors 1B and 2B, and modeling by extremely distant homology for smoothened receptor. Forty-four modeling groups participated in the assessment. State-of-the-art modeling approaches achieved close-to-experimental accuracy for small rigid orthosteric ligands and models built by close homology, and correctly predicted protein fold for distant homology targets. Predictions of long loops and GPCR activation states remain unsolved problems.

Keywords

G protein-coupled receptor; 5-hydroxytryptamine receptor 1B; 5-hydroxytryptamine receptor 2B; smoothened homolog receptor; homology modeling; docking

Introduction

Recent years have been marked by a rapid increase in the number of experimentally solved structures of G protein-coupled receptors (GPCRs) (Katritch et al., 2013; Venkatakrishnan et al., 2013), mostly due to technological breakthroughs in membrane protein crystallization

(Bill et al., 2011; Cherezov, 2011; Chun et al., 2012). The 24 receptors with publicly available structures to-date remain, however, only a small fraction of more than 350 non-olfactory/taste GPCRs in the human genome (Fredriksson et al., 2003; Lagerstrom and Schioth, 2008; Ono et al., 2005). It is an important role of the computational modeling community to help elucidate the structural details of ligand interactions, specificity and function for as many of the remaining receptors as possible.

The community-wide GPCR Dock assessment was established to evaluate the progress of the molecular modeling and ligand docking field in the context of GPCR targets. While of similar philosophy and organization to CASP and CAPRI assessments (Kryshtafovych et al., 2014; Lensink and Wodak, 2010; Moult et al., 2014), GPCR Dock focuses on specifics of GPCR-ligand complexes as opposed to the wider structural proteome dominated by soluble and globular proteins. It thus addresses a different set of challenges including distinct intramolecular packing principles influenced by the native lipid environment of GPCRs, a dynamic conformational equilibrium, evolutionally and conformationally variable loops, all of which ultimately affect ligand binding and function.

During the first and second assessments conducted in 2008 and 2010, respectively, research groups from all over the world were challenged to predict the details of ligand interactions in the A_{2A} adenosine receptor ($A_{2A}AR$), dopamine D3 receptor, and chemokine receptor CXCR4, each time using all available experimental information to date (Kufareva et al., 2011; Michino et al., 2009). These assessments identified 35–40% sequence identity between target and template as an empirical cutoff for reliable homology-based prediction of ligand-receptor interactions and highlighted the modeling techniques leading to best prediction accuracy.

In the year of 2013, the third round of the assessment was organized after the crystal structures of two human serotonin receptors ($5HT_{1B}$ and $5HT_{2B}$) (Wacker et al., 2013; Wang et al., 2013a) and a human smoothened homolog receptor (SMO) (Wang et al., 2014; Wang et al., 2013b) were solved in complex with small molecule modulators. Modelers were invited to submit their predictions of complex structures prior to release of the experimental coordinates. The assessment targets spanned a wide range of prediction difficulties, flanked by the $5HT_{1B}$ receptor on one side (with transmembrane (TM) domain sequence identity of as high as 48% to the existing structures) and by the SMO receptor on the other side (with TM sequence identity as low as 14% and with previously unseen conformations of large extracellular loops (ECLs)). Consequently, the modelers faced a new set of challenges (compared to previous assessment rounds) including building a reliable alignment between the target sequence and the available structural templates, prediction of long loop structures, and prediction of diverse binding pocket locations. The analysis of the submitted models highlighted the new successful computational methodologies and areas still in need of development.

Results

GPCR Dock 2013: description and submission statistics

The GPCR modeling and docking assessment 2013 was performed for four ligand-receptor complexes: human 5-hydroxytryptamine (5HT or serotonin) receptors 1B and 2B, both in complex with an agonist ergotamine (Wacker et al., 2013; Wang et al., 2013a) and the TM domain of the human SMO receptor in complex with two distinct small molecule antagonists, LY-2940680 (Wang et al., 2013b) and SANT-1 (Wang et al., 2014). For these four targets, 181, 171, 88 and 88 unique models were submitted by 40, 39, 20, and 20 groups, respectively. The list of participating groups, names, and affiliations is given in Box 1.

The models were assessed by several independent criteria evaluating: (i) the accuracy of alignment of the target sequence to the available structural templates, (ii) the TM bundle structure, (iii) the structure of the extracellular domains and loops, (iv–v) definition and geometry of the binding pocket, (vi) ligand position, and (vii) the atomic contacts between the ligand and the receptor. The last two criteria constituted the primary focus of the assessment; they were converted into a single score representing the likelihood of observing similar deviations among experimentally solved co-crystal structures (the so-called model "correctness"), and used for final model ranking. All assessment criteria for all models are reported in Tables S1 and S2. In the assessment report below, we report the best model by each assessment criterion as well as models within 0.1 Z-score unit (by that criterion) from it.

Target challenges

TM homology to crystallographic templates—Targets in the GPCR Dock 2013 assessment represented two extremes on the modeling difficulty scale. The level of sequence identity of the TM domain of $5HT_{1B}$ to its highest homology templates, turkey β_1 adrenoreceptor (β_1AR) and human dopamine receptor D3, exceeds 45%, and the corresponding TM domain backbone RMSD values approach 1 Å (Figure 1), qualifying it as one of the easier homology modeling cases, approximately the same difficulty as D3 in GPCR Dock 2010. The $5HT_{2B}$ receptor appears slightly more challenging with the closest structural template being ~41% identical in the TM domain, and the lowest backbone RMSD to the templates being 1.7 Å. The third assessment target, however, represents a challenge previously unseen in the GPCR Dock assessments with TM domain sequence identity of less than 15% and significant structural deviations exceeding, even in the TM domain, 2.7–3 Å from any of the structures solved to date.

Extracellular loop challenge—In contrast to the conserved seven-helical fold of the TM domains, the structural repertoire of the ECLs and termini in GPCRs is highly diverse and thus more challenging from the modeling perspective. Except for highly homologous receptors, GPCR extracellular domains often have different folds (Figure 2). In the serotonin receptor structures, the observed conformations of ECL2, the longest and the most important of the ECLs due to its involvement in ligand binding, differ significantly from any GPCRs solved to date; moreover, there are notable differences between the two receptor subtypes.

While the dimensions of ECL2 in the SMO receptor structure coincide exactly with those of 5HT$_{1B}$, the 3D conformation of this loop shares much more similarity with that of bovine rhodopsin. The SMO receptor also features an intricately folded 36-residue long fragment, the so-called *extracellular domain* (*ECD*) *linker*, which connects TM1 to the extracellular cysteine-rich domain (CRD, absent from the crystallization construct). Finally, SMO demonstrates a novel topology of ECL1 and ECL3 whose lengths (23 and 39 amino acids, respectively) far exceed those of any other known GPCR structures and whose conformations are unique in the current GPCR structural universe.

Binding pocket location—Despite the conserved receptor topology, the compounds and peptides co-crystallized with GPCRs thus far exhibit remarkable diversity in their interactions. Because determination of the binding pocket (i.e. a subset of residues interacting with the ligand) is a prerequisite for most docking procedures, this variability poses an additional challenge for the modelers.

For comparing ligand fingerprints in Figure 3A, residue contacts were projected through the alignment onto Ballesteros-Weinstein indices (Ballesteros and Weinstein, 1995) that were originally introduced for class A GPCRs and recently extended to Class F receptors (Wang et al., 2013b). Ergotamine interactions with the 5HT$_{1B}$ and 5HT$_{2B}$ receptors involves two distinct sites (Figure 3A,B): (i) "orthosteric" site deep in the TM domain, where ergoline core interactions closely resemble those observed in other biogenic amine receptors, and (ii) an "allosteric" site closer to the ECL region. The latter interactions of the cyclic tripeptide moiety, and especially of the benzyl substituent, have no analogs among the existing structures and also differ substantially between 5HT$_{1B}$ and 5HT$_{2B}$ complexes. Even more pronounced is the difference in the pocket location between the two SMO receptor complexes and the previously solved structures of other GPCRs. SANT-1 is mostly occupying the pocket in the TM domain making extensive contacts with residues in TMs 2, 3, 6, and 7 (Figure 3A,C) and in the ECL2 that is submerged deeply in the binding pocket similarly to rhodopsin. In contrast, LY-2940680 makes the majority of its residue contacts in ECL2 and ECL3 with little to no contacts in TM 3 or 6. The unique contact fingerprints of SANT-1, and especially LY-2940680, make it impossible to infer the pocket locations by homology.

Activation state—A part of the serotonin receptor challenge was prediction of the receptor activation states. Ergotamine is a classical full agonist of 5HT$_{1B}$ but a biased agonist of 5HT$_{2B}$, eliciting full β-arrestin-mediated response but only partial G-protein activation. Corroborating its functional selectivity profile, the crystallographic conformation of 5HT$_{1B}$ with ergotamine has signatures of a classical active state and closely resembles the active conformation of β$_2$AR (Rasmussen et al., 2011; Wacker et al., 2013), while the conformation of 5HT$_{2B}$ is more consistent with an inactive state in the TM5-TM6 region, but an active state in TM7.

Model alignment accuracy

Establishing a residue correspondence between the target sequence and the homologous structures is a prerequisite for any homology modeling procedure. Although the

complementary *ab initio* modeling approaches do not require the alignment as such, they still rely on identification of structural elements such as TM helices or loops, within the sequence. Alignment inaccuracies as small as a one-residue shift in a single helix inevitably lead to spatial misplacement of residues in the model and disruption of correct intra- and intermolecular contacts.

We constructed a structure-based alignment of the TM domains in the assessment targets to all available structural templates. The obtained residue correspondence is straightforward for the $5HT_{1B}$ and $5HT_{2B}$ receptors, but somewhat ambiguous for the SMO receptor whose TM7 can be aligned with the corresponding parts in class A receptor structures in two possible ways differing by one helical turn. Next, a structure-based alignment of each model to all available structural templates was built and compared to the "correct" alignment in the TM region. The obtained alignments coincided precisely for the majority of the $5HT_{1B}$ and $5HT_{2B}$ models, but had one or more mismatches for the majority of the SMO models (Figure 4A–C). At least one of the TM helices 1, 2, 4, 5, 6, and 7 was shifted by one to four residues in most cases (TM3 homology appeared to be stronger so most groups aligned it correctly); in some models the shift was as large as 30 residues (Table S1). The fewest alignments errors were observed in SMO receptor models by groups GaTech (1-residue shift in part of TM5), UWash (1 residue shift in TM7), INRIA (1 residue shifts in each of TMs 5, 6, and 7), UMich-Zhang (3 residue shift in TM6 and 1 residue shift in TM7), COH-Vaidehi (1 residue shifts in parts of TMs 5 and 6), and SNU (3-, 2-, 1, and 1 residue shifts in TMs 1, 2, 5, and 7, respectively).

As described in Supplemental Methods, the successful alignments were obtained with the use of profile-profile comparisons, 3D model input via threading, or their combination, as opposed to purely sequence-based alignment methods such as BLAST or Clustal W. Groups used such packages as HHSEARCH (Soding, 2005), FUGUE (Shi et al., 2001), RaptorX (Peng and Xu, 2011), SPARKS-X (Zhou and Zhou, 2004), I-TASSER (Roy et al., 2010), or the meta-threading package LOMETS (Wu and Zhang, 2007). The only accurate alignment (that by COH-Vaidehi) obtained by Clustal W (Larkin et al., 2007) was built in conjunction with secondary structure prediction via PSIPRED (McGuffin et al., 2000) and manual adjustments in order to achieve agreement with experimental residue mutagenesis (Dijkgraaf et al., 2011). In contrast, the procedure employed by UMich-Zhang was fully automatic, but still resulted in one of the most accurate target-template alignments in the SMO assessment. Threading and profile comparisons have therefore proved their usefulness in recognition of distant homology for GPCRs.

TM bundle structure prediction

The accuracy of prediction of TM bundle geometry was assessed by measuring the RMSD of the backbone atoms in the models following their optimal superposition with the corresponding atoms in the answers as described in Supplemental Analysis Methods. The conserved heptahelical topology in GPCRs facilitates the prediction of this domain; however, ambiguities may arise due to residue insertions or deletions leading to helical kinks and bulges, as well as due to large variations in the length of the loops connecting the helices, leading to displacement of the helix ends.

Among $5HT_{1B}$ predictions in GPCR Dock 2013, the median TM bundle RMSD to the answer was 1.89 Å, with the lowest values achieved by groups COH-Vaidehi (model #5, 1.52 Å), UWash (model #1), and EPFL (model #3) (Figure 4D–F, Table S1). Many predictions had were found in the classical active state (Supplemental Results; Figure S1D). $5HT_{2B}$ predictions had the median TM domain RMSD of 2.14 Å, with the best accuracy achieved by group NCATS (model #3, also 1.52 Å). The biased active state of this receptor was not predicted in any of the models (Supplemental Results; Figure S1E). The median TM bundle RMSD for the SMO receptor was 6.33 Å, with the lowest values achieved by group GaTech (SMO/SANT-1 models #1–5, from 2.78 Å to 3 Å, SMO/LY-2940680 models #1–5, from 2.97 Å to 3.04 Å) and UWash (SMO/LY-2940680 model #4, 3.29 Å). The decrease in prediction accuracy was consistent with the target difficulty. As expected, errors in the sequence alignment led to significant geometry prediction inaccuracies.

Most successful TM predictions for the serotonin receptors (those by COH-Vaidehi, EPFL, and NCATS) were obtained by homology with the existing aminergic receptor structures using Modeller (Eswar et al., 2006). Models were selected by either one of the Modeller energy functions (EPFL, NCATS) or manually (COH-Vaidehi, NCATS). UWash generated their models for all three receptors by homology with multiple templates and by fragment assembly with *de novo* sampling of unaligned regions in RosettaCM. Group GaTech did not provide the description of their modeling methods.

We compared the achieved prediction accuracy with the so called "naïve" approach based on threading the target sequence through the coordinates of a homologous experimental structure. From this perspective, the level of accuracy of the best $5HT_{2B}$ prediction (1.52 Å) is remarkable given that the TM domain of the closest experimentally solved structure, that of dopamine receptor D3 (PDB 3pbl), has a higher backbone RMSD of 1.7 Å. In the cases of $5HT_{1B}$ and SMO models, the predictions were no closer to the answers than the best homology modeling templates (β_2AR, PDB 3ny9, with the TM backbone RMSD of 1.19 Å to $5HT_{1B}$ and CXCR4, PDB 3oe8, with the TM backbone RMSD of 2.74 Å to SMO).

Extracellular domain prediction

The prediction of ECL conformations in the two serotonin receptors was facilitated by the presence of homologous similar length loops in the existing experimental GPCR structures. ECL1 that has the length of 10 and 11 residues in the two receptors, respectively, was predicted in 31 $5HT_{1B}$ models by 18 groups with backbone RMSD of 1.06 to 2 Å, and in 11 $5HT_{2B}$ models by 6 groups with backbone RMSD of 1.34 to 2 Å (Table S1). More than 50% of ECL1 atomic contacts with the environment were predicted correctly in 21 $5HT_{1B}$ models by 10 groups and in 60 $5HT_{2B}$ models by 20 groups (data not shown).

ECL2 which is partially disordered in both $5HT_{1B}$ and $5HT_{2B}$ structures was predicted with much lower accuracy: the best ECL2 backbone RMSD for the experimentally resolved part of the loop exceeded 3.7 Å for all $5HT_{1B}$ and 3.9 Å for all $5HT_{2B}$ models (Figure 4D–F, Table S1). This was despite the fact that in the majority (120 $5HT_{1B}$ and 97 $5HT_{2B}$) of models, the ECL2-TM3 disulfide bridge was captured correctly.

ECL3 was partially disordered in the 5HT$_{1B}$ structure making it impossible to compare the predictions to it; the corresponding 13 residue loop in 5HT$_{2B}$ was predicted with backbone RMSD between 2.27 and 3 Å in 11 models by 6 groups (Table S1). In some cases, the loop displacement was associated with inaccurately predicted positions of the extracellular ends of the TM helices while the backbone conformation of the loop was correct. Consistent with this observation, more than 50% of the loop contacts with its atomic environment were captured correctly in 64 models by 17 groups (data not shown). The ECL3 conformation and the orientation of the two cysteine residues in it (C340 and C344 in 5HT$_{1B}$, C350 and C353 in 5HT$_{2B}$) were favorable for the disulfide bridge formation in 67 5HT$_{1B}$ and 59 5HT$_{2B}$ models.

Unfortunately, modeling of long loops or termini with no homology to the existing structures remains an unsolved problem. None of the SMO models predicted the geometry of ECD linker (36 residues), ECL1 (23 residues), or ECL3 (39 residues) with acceptable accuracy: the lowest backbone RMSD after the optimal superposition of the TM domain exceeded 12 Å, 8 Å, and 7 Å for the ECD linker, ECL1 and ECL3, respectively (Table S1). None of the models captured the intra-ECD-linker disulfide bond (C193–C213) or the disulfide connecting the linker to ECL1 (C217–C295). However, it was an encouraging finding that the helical fold and the disulfide bond in the ECL3 (C490–C507) were correctly predicted in 12 models submitted by groups UWash (SMO/LY-2940680 models #1,4,5, SMO/SANT-1 models #4,5), Warsaw (SMO/LY-2940680 models #1,3,5 and SMO/SANT-1 models #1,3,5), and Vanderbilt (SMO/SANT-1 model #4).

ECL2 modeling in SMO was facilitated by its structural similarity to the corresponding fragment of bovine rhodopsin: when the TM domains are optimally superimposed, the backbone RMSD of the 19 ECL2 residues equals 3.21 Å. Group UWash modeled SMO ECL2 with RMSD of around 3 Å (2.92 Å for model #1 and 3.53 Å for model #4). 3.68 Å and 3.76 Å were achieved by group UMich-Zhang. For both of these groups, the initial target-template sequence alignment appears to have only minor errors, all localized to TM6 and TM7. Also, group Helsinki-Xhaard achieved the ECL2 RMSD of 3 Å in their model #5 of the SMO/SANT-1 complex despite generally inaccurate target-template sequence alignment; this was due to the fact that their TM3, TM4, and TM5 alignment was correct and the large alignment shifts were localized elsewhere. On the contrary, some models with few alignments errors failed to predict the loop geometry. Therefore, an accurate alignment in the TM3/TM4/TM5 region appeared necessary, but not sufficient for ECL2 geometry prediction. The conserved ECL2-TM3 disulfide bond (C314–C390) was correctly predicted in 70 of the 176 models of SMO complexes.

TM domain prediction accuracy did not correlate with ECL prediction accuracy, except for the fact that most accurate predictions in both categories were produced based on an alignment with the least errors. Thus, we also assessed the model population by the combined accuracy of TM domain and ECL2 prediction (ECL2 was chosen over other extracellular regions because of its importance in ligand binding). Independent Z-scores were calculated for TM and ECL2 backbone predictions after eliminating negative outliers (models deviating from the mean by more than 2 standard deviations) and averaged. The highest accuracy models differed from the respective population means by 1.12

(Copenhagen-Gloriam $5HT_{1B}$ models #1 and 2), 1.38 (UMich-Zhang $5HT_{2B}$ model #3), 1.33 (UWash SMO/LY-2940680 model #1), and 1.44 (UMich-Zhang SMO/SANT-1 model #2) standard deviations.

Group Copenhagen-Gloriam built their serotonin receptor models with Modeller using hybrid templates and selected the best solutions by Modeller DOPE score and visual inspection. For all three of the assessment target receptors, UMich-Zhang used a composite approach including profile-profile comparisons, threading, and iterative fragment reassembly with consequent model refinement at atomic level using fragment-guided molecular dynamics simulation (FG-MD (Zhang et al., 2011)). UWash generated their models by homology with multiple templates and by fragment assembly with *de novo* sampling of unaligned regions in RosettaCM. Table S1 shows models ranked by their combined protein prediction Z-scores.

Binding pocket definition and geometry

The correctness of the binding site residue selection and the local accuracy of the binding site atom geometry are parameters only weakly correlating with the overall structural accuracy of the model but much more important for ligand docking and screening applications. We therefore evaluated these parameters separately from the overall model accuracy. The correctness of the binding site residue identification was assessed by calculating the ligand contact strength "fingerprint" on the model residues and comparing it with the corresponding "fingerprint" in the target structure. The RMSD of the binding site residues was used to assess the geometrical accuracy of the pocket and calculated using the optimal superposition of the TM domains as described in Supplemental Analysis Methods.

The RMSD of the binding site residues was as low as 1.31 Å for the $5HT_{1B}$ predictions (achieved in model #3 by group Stockholm-Carlsson) (Figure 5). This model also featured as high as 81.1% accuracy in the ligand fingerprint determination. Models #2 and 4 by the same group also had very low binding site RMSD, but were less accurate in the ligand fingerprint determination. A similar situation was observed in models #1, 2 and 4 by group Copenhagen-Gloriam. The median binding site RMSD among the $5HT_{1B}$ models was equal to 2.35 Å and the median fingerprint accuracy to 57%.

Group Stockholm-Carlsson also produced the most accurate predictions of the binding site location (79.4% in model #2, with pocket residue RMSD of 1.94 Å) and geometry (1.88 Å in model #4) in the $5HT_{2B}$ assessment. The median values were 58.7% and 3.42 Å.

Both groups (Stockholm-Carlsson and Copenhagen-Gloriam) built their predictions using Modeller (Eswar et al., 2006) using hybrid templates (β_1AR, β_2AR, D3, H1 receptors). A critical step in the Stockholm-Carlsson modeling procedure was assessment of the models by VLS (virtual ligand screening) enrichment, in an iterative manner. Alternatively, Copenhagen-Gloriam prioritized their predicted compound poses by agreement with other aminergic GPCRs, ligand SAR, mutation and other species data. Both groups manually inspected their models at the final selection step.

Far less accurate predictions were made for the SMO receptor complexes. The best pocket RMSD among the SMO/LY-2940680 predictions was equal to 9.06 Å (model #5 by group GaTech), median 13.28 Å. The highest pocket fingerprint accuracy was only 25% (model #1 by group SNU), median 10.7%. However, the binding site predictions were considerably more accurate for the SMO/SANT-1 complexes with the lowest pocket RMSD of 4.63 Å (model #2 by group UMich-Zhang), and the best binding site location accuracy of 58.9% (model #5 by group GaTech) (Table 1). The binding mode of the SANT-1 molecule involves fewer ECL contacts and more closely resembles the previously observed binding models of agonists and antagonists in other GPCRs, possibly contributing to better predictions of both pocket location and geometry in the SMO/SANT-1 assessment in selected models. The median values were 10.66 Å and 15.8%.

Understandably, top ranking predictions for SMO were obtained in the models with fewest alignment errors. Group SNU docked the ligands with GalaxyDock (Shin et al., 2013) and performed manual mutagenesis-guided ligand pose selection. Group UMich-Zhang identified the binding sites by a combination of binding-specific substructure comparisons and sequence profile alignments (the so-called COACH (Yang et al., 2013b)). To generate ligand-GPCR complexes, they used a low-resolution coarse grain docking method, BSP-SLIM (Lee and Zhang, 2012), that recognizes local shape and feature complementarities between the ligand and the binding pocket.

Ligand RMSD and ligand-pocket contacts

The main focus of the assessment for the serotonin receptor targets was prediction of the binding pose of the ligand and its contacts with the surrounding residues. For the $5HT_{1B}$/ ergotamine complex, the median ligand RMSD was 5.44 Å, with the best result as low as 1.51 Å (Figure 6A, Table S2). The median contact prediction accuracy was equal to 15.1%, the best being 51.1%. The best RMSD and the best contact prediction were achieved in models #2 and #3, respectively, by group Stockholm-Carlsson. The corresponding statistics for the $5HT_{2B}$ receptor were (Figure 6B, Table S2): the median ligand RMSD of 5.54 Å, best 1.05 Å (achieved in model #3 by group Warsaw); the median contact accuracy of 17.8%, best 53.1% (achieved in model #4 by group Stockholm-Carlsson). The most accurate models were generally accurate by both criteria, RMSD and contacts.

Because the goal of molecular modeling is to ultimately complement the experimental structure determination efforts, it was interesting to compare the achieved levels of accuracy to the variation observed in the crystallographic experiments and resulting from the natural protein flexibility and/or limited experimental resolution. Using a large set of pairs of PDB complexes of identical protein/ligand composition, we derived statistics for answering this question (Supplemental Methods; Figure S2). According to our calculations, the accuracy of the best models corresponded to only 3.35% of the experimental distribution for $5HT_{1B}$ and to 5.11% for $5HT_{2B}$ (Table S2). We further refer to these percentiles as the measure of "correctness" of the models.

The above results show that the 2013 serotonin receptor ligand prediction accuracy was slightly below than that achieved in 2010 for the D3/eticlopride complex (0.96 Å RMSD with 58% of contacts, correctness of 8.09%). However, it would be incorrect to conclude

that the computational modeling field has not progressed, because the complexity of the ligand docking problem in the present assessment was much higher. As described above, ergotamine simultaneously occupies orthosteric and allosteric pockets in the target receptors. Extensive and diverse contacts in the ECL regions create an additional modeling challenge that was not a part of 2010 D3 assessment. To quantitatively estimate the progress, we evaluated the quality of predictions the orthosteric part of the ergotamine molecule, i.e. its ergoline core (Table S2). For $5HT_{1B}$, the median ergoline core RMSD was 3.94 Å, and the best results was as low as 0.72 Å (achieved in model #1 by group Schrödinger). The median contact prediction accuracy was 22.9%, and the best as high as 65.3% (model #3 by group Copenhagen-Norn). For $5HT_{2B}$, the median ergoline core RMSD was equal to 3.69 Å, the best being 0.64 Å (model #2 by group NCATS). The median contact accuracy equaled 23.9%, with the best achieving 72.6% (model #2 by MerzPharma). This level of accuracy places the ergoline core predictions in the shaded area of the ligand RMSD/contact plot in Figure 6, which represents the experimental variation of the corresponding parameters. The correctness of the best ergoline core predictions corresponds to 13.01% and 19.55% of the experimental distribution in the $5HT_{1B}$ and $5HT_{2B}$ assessments, respectively.

The most accurate $5HT_{1B}$ ligand predictions also featured low RMSD of the binding pocket residues (Figure 6E). However, the correlation was not as straightforward for $5HT_{2B}$ complexes, as some accurate ligand predictions were achieved in relatively inaccurate pockets. A closer inspection of the models showed that the pocket inaccuracies in these models were of local nature and observed either in the ECL2 residues or in a single incorrectly aligned TM helix (for the models by MerzPharma).

Far less accurate predictions were made for the SMO receptor complexes (Figure 6C,D, Table S2). The lowest ligand RMSD for SMO/LY-2940680 complex models was as high as 4.42 Å (model #3 by group Vanderbilt), and the highest achieved contact ratio was only 8.8% (model #1 by the same group). Due to the more "traditional" location of the binding site in the SMO/SANT-1 complex, the accuracy was marginally better for this target: the lowest RMSD of 4.31 Å (model #5 by group GaTech) and the best contact ratio of 12.3% (model #5 by group SNU). The median ligand RMSD exceeded 10 Å and median fraction of contacts was less than 1% for both SMO receptor complexes. In the context of experimental variation, these values correspond to approximately 0.1% percentile of the distribution (Table S2). Unlike in the serotonin receptor models, no correlation was observed between RMSD and contacts in the SMO predictions.

By the rules of the assessment, the participants submitted at most five models for each target and ranked these models according to the degree of their confidence in the correctness of the prediction. As in the 2010 assessment, no correlation was observed between the rank and the accuracy of the models (Figure 7).

Analysis of best models

Two $5HT_{1B}$ and seven $5HT_{2B}$ models stood out as their ligand prediction accuracy brushed the experimentally observed variation within an arbitrarily chosen cutoff of 2% (the "correctness" column in Table 1, Figure 6). Among $5HT_{1B}$ models, both first and second top ranking predictions were submitted by group Stockholm-Carlsson as their models #2 (Figure

8A) and 3, respectively. The models contained no alignment errors. For model #2, the TM backbone RMSD and ECL2 RMSD were 1.82 Å and 4.34 Å from the answer, whereas model #2 was marginally closer to the answer in the TM domain (TM backbone RMSD of 1.66 Å). The receptor prediction accuracy was at 0.93 standard deviations above the mean of the model population. The geometry and the composition of the binding pocket are among the best in the assessment (RMSD of 1.41 Å and 1.31 Å over the heavy atoms of the 19 pocket residues, 1.2 Å and 1.07 Å for the 14 residues in the TM domain and less than 1.5 Å for the 3 ECL2 residues, 77% and 81% accuracy in the pocket fingerprint prediction). The docked ligand featured non-hydrogen atom RMSD of 1.51 Å and 2.09 Å for the two models, and reproduced 47% and 51% of the experimental ligand-pocket atomic contacts. To derive their complex models, the authors used the approach that has proved successful in the previous assessment and in multiple GPCR modeling applications (Carlsson et al., 2011; Katritch and Abagyan, 2011; Katritch et al., 2011; Katritch et al., 2012; Mysinger et al., 2012; Weiss et al., 2013); namely, they carried the refinement of their models in an iterative manner evaluating, at each iteration, the enrichment of known ligands in virtual screening.

The prediction of the ergoline core was only moderate in these two top $5HT_{1B}$ models (Table S2). The best ergoline core accuracy was achieved in the models #1 by Schrödinger (0.72Å RMSD, 65% of correct contacts) and #3 by Copenhagen-Norn (0.86Å RMSD, 65% of correct contacts). Both demonstrate highly accurate prediction for the pocket residues. In the former model, the correct ergoline core pose was ranked high by taking into account conservation of pocket residues between $5HT_{1B}$ and other ergoline-binding aminergic receptors (see Supplemental Methods).

The most accurate $5HT_{2B}$ model, Warsaw #3 (Figure 8B), contained no detectable alignment errors and had 2.12 Å backbone RMSD over the 192 TM domain residues and 2.69 Å RMSD over all non-hydrogen atoms of the 23 binding site residues. 76% of the pocket fingerprint was identified correctly. The ligand deviated by RMSD of only 1.05 Å from its crystallographic pose and correctly reproduced 50% of experimental ligand-pocket contacts. The model was obtained using the procedure implemented in the GPCRM web-site (Modeller, loop refinement by Rosetta, model selection by DOPE and Rosetta total score) with subsequent two-step ligand docking, GlideScore scoring, and human expert analysis of the complexes.

In contrast to the top scoring model, $5HT_{2B}$ models ranking second, third, and fourth (MerzPharma #3, 2, 1, respectively) featured one or more residue alignment shift errors in the top parts of TM helices 4, and 7, resulting in the TM backbone RMSD of 2.68 Å, ECL2 RMSD exceeding 9 Å, and the protein prediction Z-score of below −1 (i.e. one or more standard deviation *below* the mean of the model population). However, the alignment and geometry of TMs 3, 5, and 6 containing most of the ergotamine molecule interactions were predicted correctly. Consequently, in model #3, 68% of the pocket fingerprint was correct; the ligand RMSD from the target structure was as low as 1.15 Å, and ligand-pocket contacts were reproduced with the accuracy of 45%. Moreover, these models feature some of the most accurate predictions for the ergoline core with RMSD of 0.66 Å and 73% of correct contacts (Table S2). These predictions are closely followed by three models from the

Scandinavian teams (Stockholm-Carlsson #2, Copenhagen-Norn #1, and Stockholm-Carlsson #4) with ligand RMSD below 2 Å and contact prediction accuracy exceeding 50%.

Along with, and possibly enabling the ligand prediction, Stockholm-Carlsson models also feature high precision in prediction of binding pocket residues not only in the TM domain (RMSD of 1.53–1.68 Å), but also in ECL2 (RMSD of ~2.3 Å for the 4 ECL2 residues in direct contact with the ligand).

The top scoring model of the SMO/LY-2940680 complex (Vanderbilt #3, Table 1, Figure 8C) only weakly resembles the experimental structure with the ligand deviating from the answer by as much as 4.42 Å and no strong correct contacts (contact prediction accuracy of 2%). Marginally better contact prediction accuracy was achieved in the top-ranking models of SMO/SANT-1 complex (SMO/SANT-1: GaTech #5, UMich-Zhang #3, SNU #5, Table 1, Figure 8D), probably owing to a more traditional TM location of the SANT-1 binding pocket (7%, 11%, and 12% of contacts captured, respectively). As with other assessment criteria, the accuracy of target-template alignment appeared a prerequisite for ligand pose prediction in the SMO receptor. Proper placement of some pocket residues allowed the authors to follow their docking procedure with mutagenesis-guided manual ligand pose selection. However, even when alignment in the pocket-defining helices is largely "correct", partial offsets and shifts of the helical backbone between modeling templates and the SMO receptor structure impact positioning of the contact residues and ultimately lower contact recall.

Discussion

The targets in the present assessment presented a new set of challenges for the GPCR modelers. Despite the high homology to the previously determined structures, the complexes of the $5HT_{1B}$ and $5HT_{2B}$ with ergotamine were hard to predict because the latter makes extensive (and different) contacts with the flexible ECLs. Additionally, the modelers had to predict the canonical active state of $5HT_{1B}$ and the biased non-canonical state of $5HT_{2B}$. The ECL interactions of the ligands could only be predicted with moderate precision: the binding poses of ergotamine were predicted with RMSD of 1–1.5 Å and about 50% of correct contacts. Only 3.3–5% of experimentally solved structures pairs with identical molecular composition deviate from each other in their ligand RMSD and contacts as much as the best $5HT_{1B}$ and $5HT_{2B}$ models deviate from the target receptor structures. This moderate success correlates with the accuracy in the ECL predictions. In contrast, the accuracy of best predictions for the small and rigid ergoline core that predominantly interacts with the TM domain residues, is comparable to ~13–20 percentile of the experimental variation, with RMSD as low as 0.64–0.67 Å and up to 73% of correct contacts.

Most successful serotonin receptor predictions were derived using homology modeling, often using the Modeller software and multiple templates in the form of the available aminergic receptor structures. As in the previous assessment, the strategy of selecting the models by their ability to recognize known ligands among similar inactive molecules played out well. Some of the top-ranking complexes were refined by molecular dynamics

simulations. Visual inspection and manual or semi-automatic model selection using subfamily-specific residue contacts, ligand SAR, and mutagenesis data have also proved productive.

Many of the submitted models successfully captured the activation state of $5HT_{1B}$, but not the biased state of $5HT_{2B}$ (see Supplemental Results). Availability of structures in various functional states is a contributing factor for this partial success; however, the lack of successful predictions for $5HT_{2B}$ highlights the need for either experimental determination of multiple functional states of GPCRs or for the development of reliable computational methods for their prediction. The SMO structure presented a level of challenge previously unseen in the GPCR Dock assessments, not only because of its extremely low homology to the existing structures, but also because of its long ECLs whose fold is unique among GPCR structures solved to-date. The present assessment highlighted uncertainties in the target-template sequence alignment as the critical bottleneck in the task of modeling by distant homology. It demonstrated the advantage of composite approaches that incorporate threading, fragment assembly, and energy-based refinement (e.g. iterative threading assembly refinement, I-TASSER (Roy et al., 2010; Wu et al., 2007)) in finding the proper residue correspondence. The growing number of GPCR structures increase our understanding of the principles of dynamic intra-molecular packing in these proteins; such knowledge may promote better methods for residue contact prediction (e.g. MemBrain (Yang et al., 2013a)) and help in recognizing distant homology.

SMO assessment also demonstrated that while the alignment ambiguities can be tackled by modern algorithms, the structural accuracy of the distant homology models is still in need of improvement, especially in loop regions. Wise selection of targets for experimental structure determination can help advance the structural coverage of the GPCR superfamily and boost elucidation of structure and function of the closer homologs by computational methods.

Materials and Methods

Data collection and models analysis methods, as well as methods, techniques, and approaches used by the participants of GPCR Dock 2013 for complex model generation are described in detail in Supplemental Materials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Bibliography

Ballesteros, JA.; Weinstein, H. Methods in Neurosciences. Academic Press; 1995. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors; p. 366-428.

Bill RM, Henderson PJF, Iwata S, Kunji ERS, Michel H, Neutze R, Newstead S, Poolman B, Tate CG, Vogel H. Overcoming barriers to membrane protein structure determination. Nat Biotech. 2011; 29:335–340.

Carlsson J, Coleman RG, Setola V, Irwin JJ, Fan H, Schlessinger A, Sali A, Roth BL, Shoichet BK. Ligand discovery from a dopamine D3 receptor homology model and crystal structure. Nat Chem Biol. 2011; 7:769–778. [PubMed: 21926995]

Cherezov V. Lipidic cubic phase technologies for membrane protein structural studies. Current Opinion in Structural Biology. 2011; 21:559–566. [PubMed: 21775127]

Chun E, Thompson Aaron A, Liu W, Roth Christopher B, Griffith Mark T, Katritch V, Kunken J, Xu F, Cherezov V, Hanson Michael A, et al. Fusion Partner Toolchest for the Stabilization and Crystallization of G Protein-Coupled Receptors. Structure. 2012; 20:967–976. [PubMed: 22681902]

Dijkgraaf GJP, Alicke B, Weinmann L, Januario T, West K, Modrusan Z, Burdick D, Goldsmith R, Robarge K, Sutherlin D, et al. Small Molecule Inhibition of GDC-0449 Refractory Smoothened Mutants and Downstream Mechanisms of Drug Resistance. Cancer Research. 2011; 71:435–444. [PubMed: 21123452]

Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics. 2006; Chapter 5(Unit 5):6. [PubMed: 18428767]

Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol. 2003; 63:1256–1272. [PubMed: 12761335]

Katritch V, Abagyan R. GPCR agonist binding revealed by modeling and crystallography. Trends Pharmacol Sci. 2011; 32:637–643. [PubMed: 21903279]

Katritch V, Cherezov V, Stevens RC. Structure-Function of the G Protein-Coupled Receptor Superfamily. Annual Review of Pharmacology and Toxicology. 2013; 53:531–556.

Katritch V, Kufareva I, Abagyan R. Structure based prediction of subtype-selectivity for adenosine receptor antagonists. Neuropharmacology. 2011; 60:108–115. [PubMed: 20637786]

Katritch V, Rueda M, Abagyan R. Ligand-guided receptor optimization. MethodsMol Biol. 2012; 857:189–205.

Kryshtafovych A, Moult J, Bales P, Bazan JF, Biasini M, Burgin A, Chen C, Cochran FV, Craig TK, Das R, et al. Challenging the state of the art in protein structure prediction: Highlights of experimental target structures for the 10th Critical Assessment of Techniques for Protein Structure Prediction Experiment CASP10. Proteins: Structure, Function, and Bioinformatics. 2014; 82:26–42.

Kufareva I, Rueda M, Katritch V, Stevens RC, Abagyan R. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. Structure. 2011; 19:1108–1126. [PubMed: 21827947]

Lagerstrom MC, Schioth HB. Structural diversity of G protein-coupled receptors and significance for drug discovery. Nat Rev Drug Discov. 2008; 7:339–357. [PubMed: 18382464]

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948. [PubMed: 17846036]

Lee HS, Zhang Y. BSP-SLIM: A blind low-resolution ligand-protein docking approach using predicted protein structures. Proteins: Structure, Function, and Bioinformatics. 2012; 80:93–110.

Lensink MF, Wodak SJ. Docking and scoring protein interactions: CAPRI 2009. Proteins: Structure, Function, and Bioinformatics. 2010; 78:3073–3084.

McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics. 2000; 16:404–405. [PubMed: 10869041]

Michino M, Abola E, Brooks CL, Dixon JS, Moult J, Stevens RC. participants, G.D. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. Nat Rev Drug Discov. 2009; 8:455–463. [PubMed: 19461661]

Moult J, Fidelis K, Kryshtafovych A, Schwede T, Tramontano A. Critical assessment of methods of protein structure prediction (CASP) — round x. Proteins: Structure, Function, and Bioinformatics. 2014; 82:1–6.

Mysinger MM, Weiss DR, Ziarek JJ, Gravel S, Doak AK, Karpiak J, Heveker N, Shoichet BK, Volkman BF. Structure-based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. Proceedings of the National Academy of Sciences. 2012; 109:5517–5522.

Ono Y, Fujibuchi W, Suwa M. Automatic gene collection system for genome-scale overview of G-protein coupled receptors in eukaryotes. Gene. 2005; 364:63–73. [PubMed: 16126348]

Peng J, Xu J. Raptorx: Exploiting structure information for protein alignment by statistical inference. Proteins: Structure, Function, and Bioinformatics. 2011; 79:161–171.

Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, et al. Crystal structure of the b2 adrenergic receptor-Gs protein complex. Nature. 2011; 477:549–555. [PubMed: 21772288]

Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protocols. 2010; 5:725–738.

Shi J, Blundell TL, Mizuguchi K. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. Journal of Molecular Biology. 2001; 310:243–257. [PubMed: 11419950]

Shin WH, Kim JK, Kim DS, Seok C. GalaxyDock2: Protein–ligand docking using beta-complex and global optimization. Journal of Computational Chemistry. 2013; 34:2647–2656. [PubMed: 24108416]

Soding J. Protein homology detection by HMM-HMM comparison. Bioinformatics. 2005; 21:951–960. [PubMed: 15531603]

Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. Nature. 2013; 494:185–194. [PubMed: 23407534]

Wacker D, Wang C, Katritch V, Han GW, Huang XP, Vardy E, McCorvy JD, Jiang Y, Chu M, Siu FY, et al. Structural Features for Functional Selectivity at Serotonin Receptors. Science. 2013; 340:615–619. [PubMed: 23519215]

Wang C, Jiang Y, Ma J, Wu H, Wacker D, Katritch V, Han GW, Liu W, Huang XP, Vardy E, et al. Structural Basis for Molecular Recognition at Serotonin Receptors. Science. 2013a; 340:610–614. [PubMed: 23519210]

Wang C, Wu H, Evron T, Vardy E, Han G, Huang X, Hufeisen S, Mangano T, Urban D, Katritch V, et al. Structural basis for Smoothened receptor modulation and chemoresistance. Nature Communications. 2014 accepted.

Wang C, Wu H, Katritch V, Han GW, Huang XP, Liu W, Siu FY, Roth BL, Cherezov V, Stevens RC. Structure of the human smoothened receptor bound to an antitumour agent. Nature. 2013b; 497:338–343. [PubMed: 23636324]

Weiss DR, Ahn S, Sassano MF, Kleist A, Zhu X, Strachan R, Roth BL, Lefkowitz RJ, Shoichet BK. Conformation Guides Molecular Efficacy in Docking Screens of Activated b2 Adrenergic G Protein Coupled Receptor. ACS Chemical Biology. 2013; 8:1018–1026. [PubMed: 23485065]

Wu S, Skolnick J, Zhang Y. Ab initio modeling of small proteins by iterative TASSER simulations. BMC Biol. 2007; 5:17–17. [PubMed: 17488521]

Wu S, Zhang Y. LOMETS: A local meta-threading-server for protein structure prediction. Nucleic Acids Research. 2007; 35:3375–3382. [PubMed: 17478507]

Yang J, Jang R, Zhang Y, Shen H-B. High-accuracy prediction of transmembrane inter-helix contacts and application to GPCR 3D structure modeling. Bioinformatics. 2013a

Yang J, Roy A, Zhang Y. Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics. 2013b; 29:2588–2595. [PubMed: 23975762]

Zhang J, Liang Y, Zhang Y. Atomic-Level Protein Structure Refinement Using Fragment-Guided Molecular Dynamics Conformation Sampling. Structure. 2011; 19:1784–1795. [PubMed: 22153501]

Zhou H, Zhou Y. Single-body residue-level knowledge-based energy score combined with sequence-profile and secondary structure information for fold recognition. Proteins: Structure, Function, and Bioinformatics. 2004; 55:1005–1013.

Highlights

- The four targets in the 2013 assessment presented diverse challenges for modelers

- Highly accurate predictions for close homology models and rigid orthosteric ligands

- Alignment inaccuracies are the major hurdle in modeling by distant homology

- Predictions of loops and receptor activation states remain unsolved problems
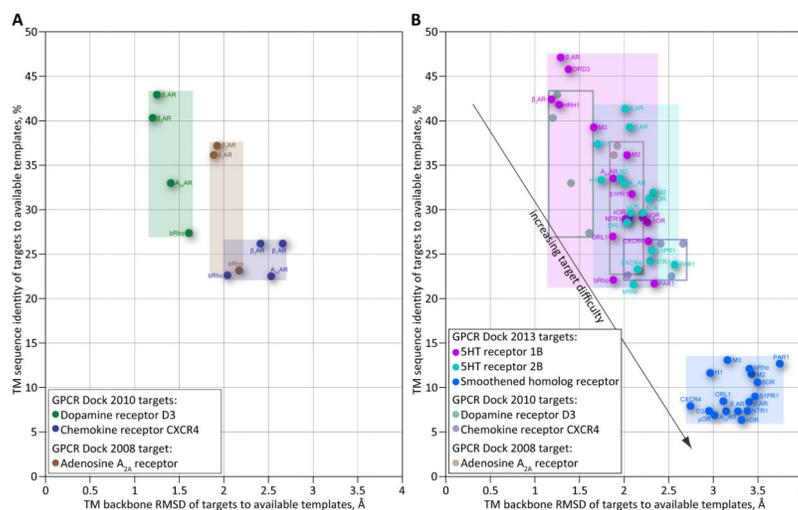
Figure 1.
Difficulty of targets in the GPCR Dock 2008–10 (A) and 2013 (B) assessments, evaluated as their sequence and structural similarity to GPCR structures available at the time of the assessment. Each dot represents a pair of a GPCR Dock target and a PDB template. Higher TM backbone RMSD (x-axis) and lower sequence identity (y-axis) indicate a more challenging modeling case. For receptors with multiple crystal structures, the lowest RMSD to the assessment target is plotted.
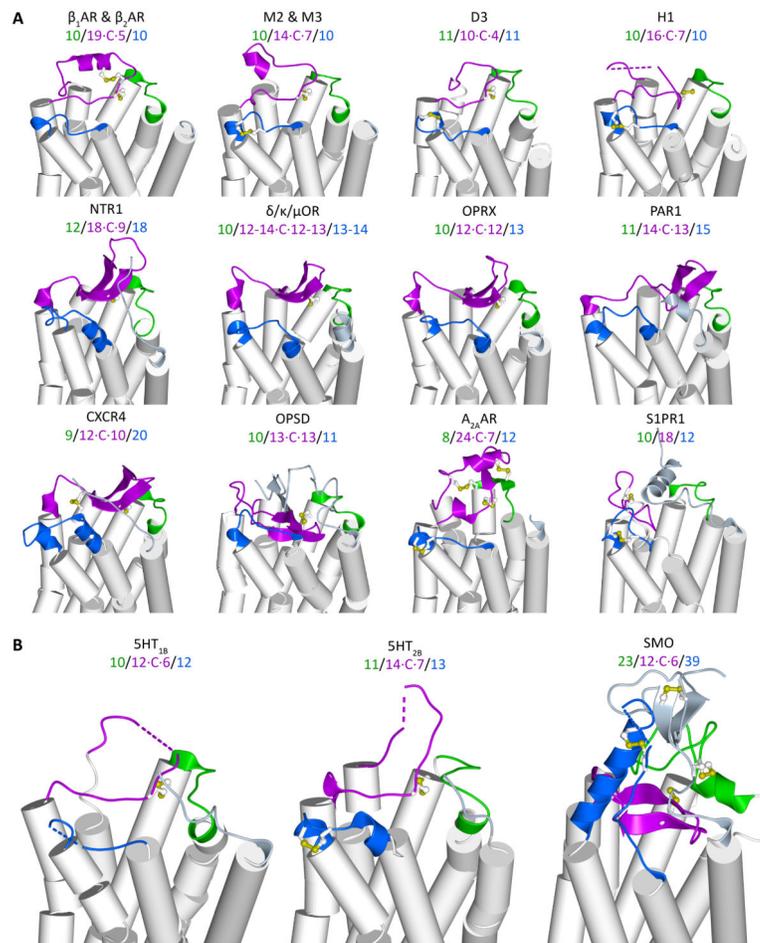
Figure 2.
Topology and conformation of the extracellular loops in previously solved structures (A) and the three assessment targets (B).
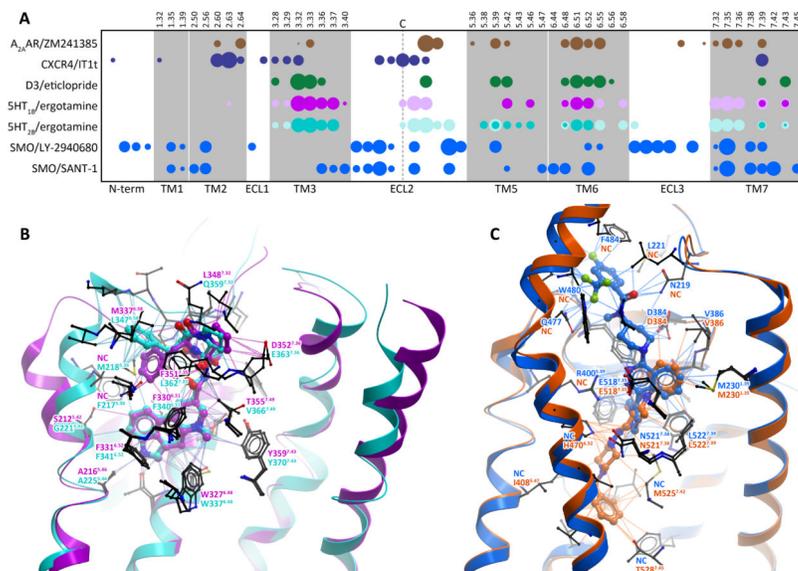
Figure 3.
(A) Contact strength fingerprints of the ligands on the receptor residues in the GPCR Dock 2008, 2010 and 2013 targets; (B–C) 2013 assessment target contacts in 3D. $5HT_{1B}$ and $5HT_{2B}$ receptor interactions with ergotamine are color-coded magenta and cyan, respectively; ribbon not displayed for TM helices 6 and 7. SMO receptor residue contacts with LY-2940680 and SANT1 are color-coded blue and orange, respectively; ribbon not displayed for TM7. Only residue contacts with cumulative strength above 1, according to Supplementary Analysis Methods, are depicted. NC – no contact. For the SMO receptor, one of two possible Ballesteros-Weinstein numberings in TM7 is arbitrarily chosen.
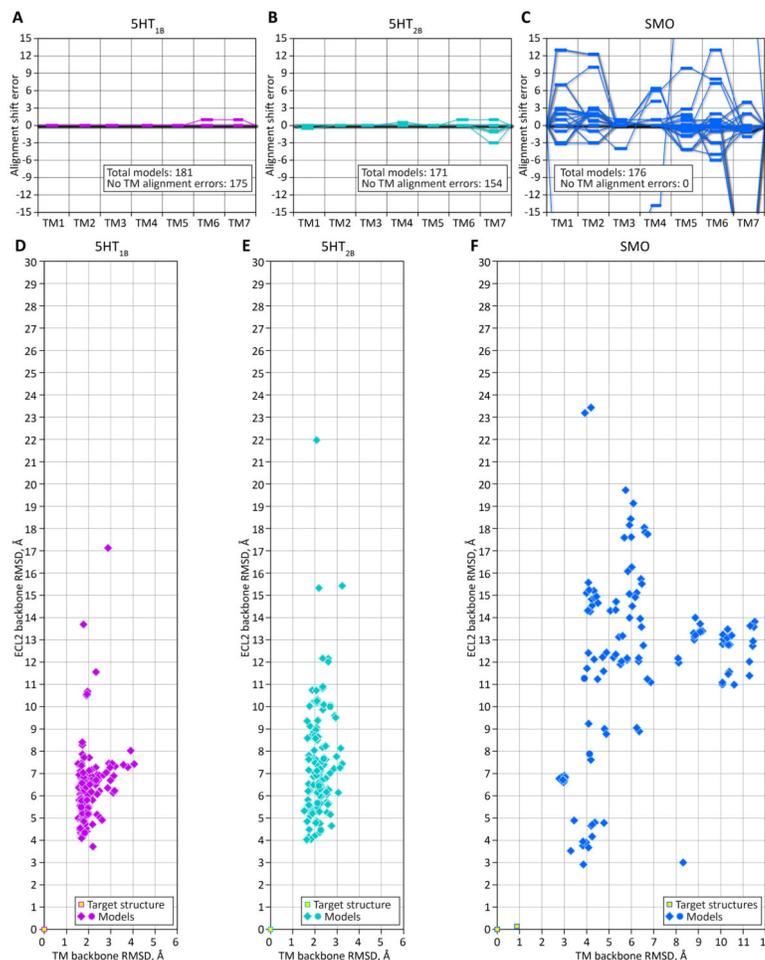
Figure 4.
Protein structure prediction in GPCR Dock 2013. (A–C) Sequence alignment errors observed in the submitted models of $5HT_{1B}$, $5HT_{2B}$ and SMO receptors. Each model is represented by a continuous line. TM helices 1 through 7 are shown on the horizontal axis; the vertical shift of a line with respect to the axis indicated indicates the average error in the alignment of the corresponding TM helix. (D–F) Scatter plot of TM domain backbone RMSD (horizontal) vs ECL2 backbone RMSD (vertical) for the predictions of three targets in GPCR Dock 2013. See also Supplemental Results, Table S1 and Figure S1.
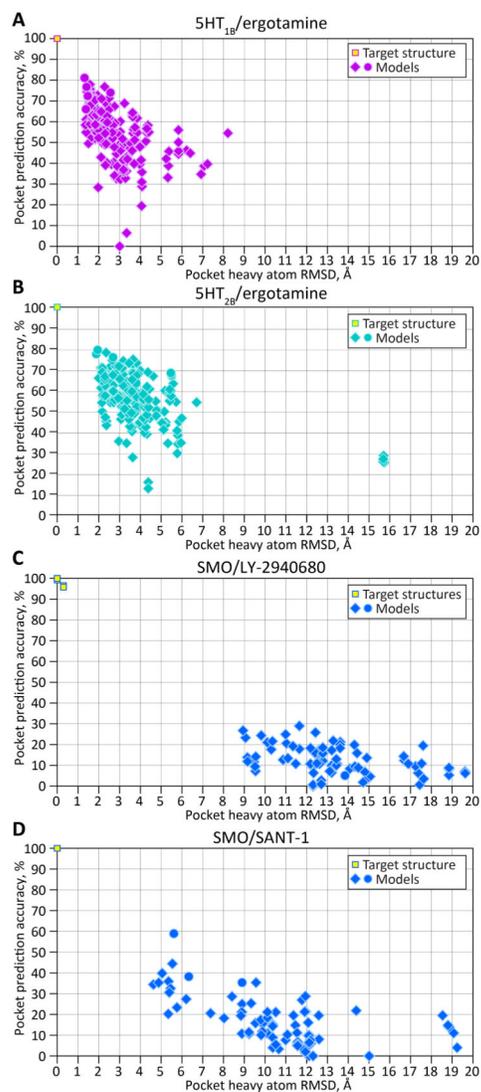
Figure 5.
Scatter plot of binding pocket heavy atom RMSD (horizontal) vs the accuracy of prediction of ligand pocket fingerprint (vertical) in the models of the four target complexes in GPCR Dock 2013. See also Table S2.
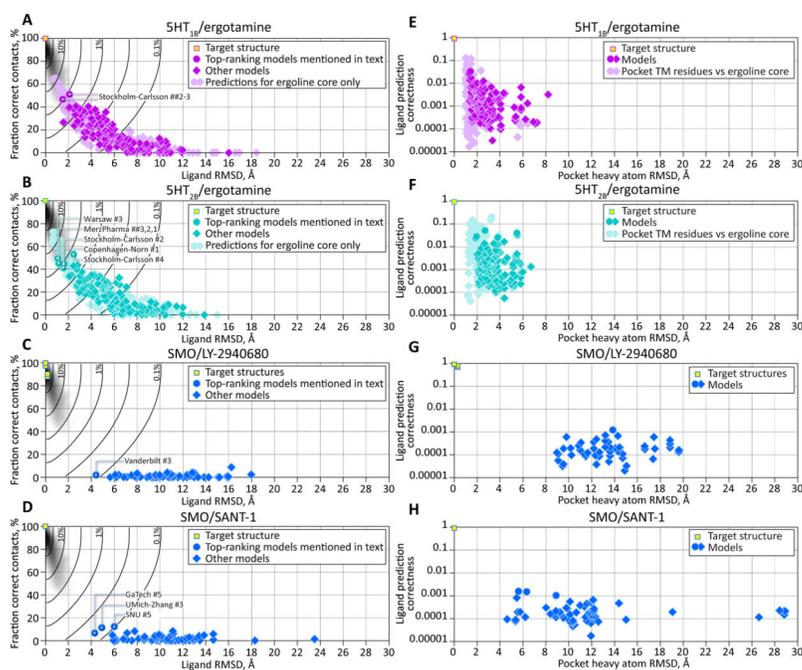
Figure 6.
Ligand prediction for the targets in the GPCR Dock 2013 assessment. (A–D) Prediction of ligand pose and atomic contacts with the pocket residues. The shaded background represents the experimentally observed distribution of the parameters on the plot (see also Figure S2). Solid black curves represent isolines of the model "correctness" used for model scoring and ranking. (E–H) Scatter plot of ligand correctness plotted against the RMSD of binding pocket residues. In (A), (B), (E) and (F), predictions for the ergoline core and the TM part of the binding pocket are shown in light pink and light cyan for $5HT_{1B}$ and $5HT_{2B}$, respectively. See also Table S2.
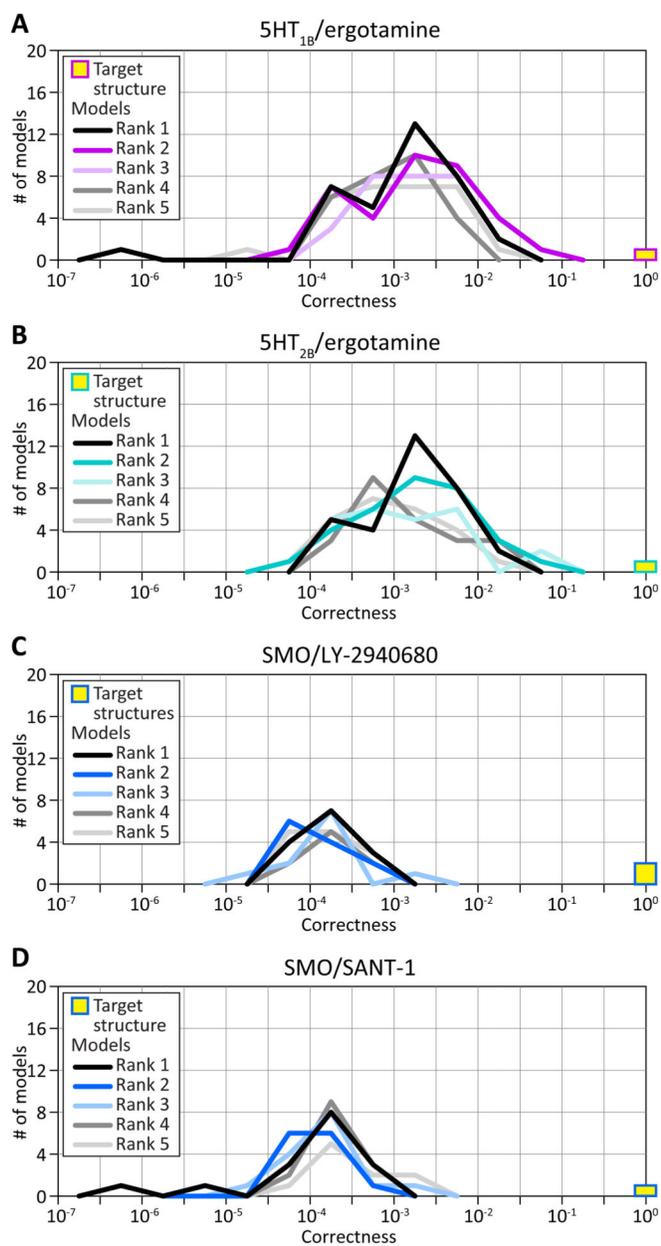
Figure 7.
Distribution of model correctness by confidence ranks assigned to the models by the authors. See also Table S2.
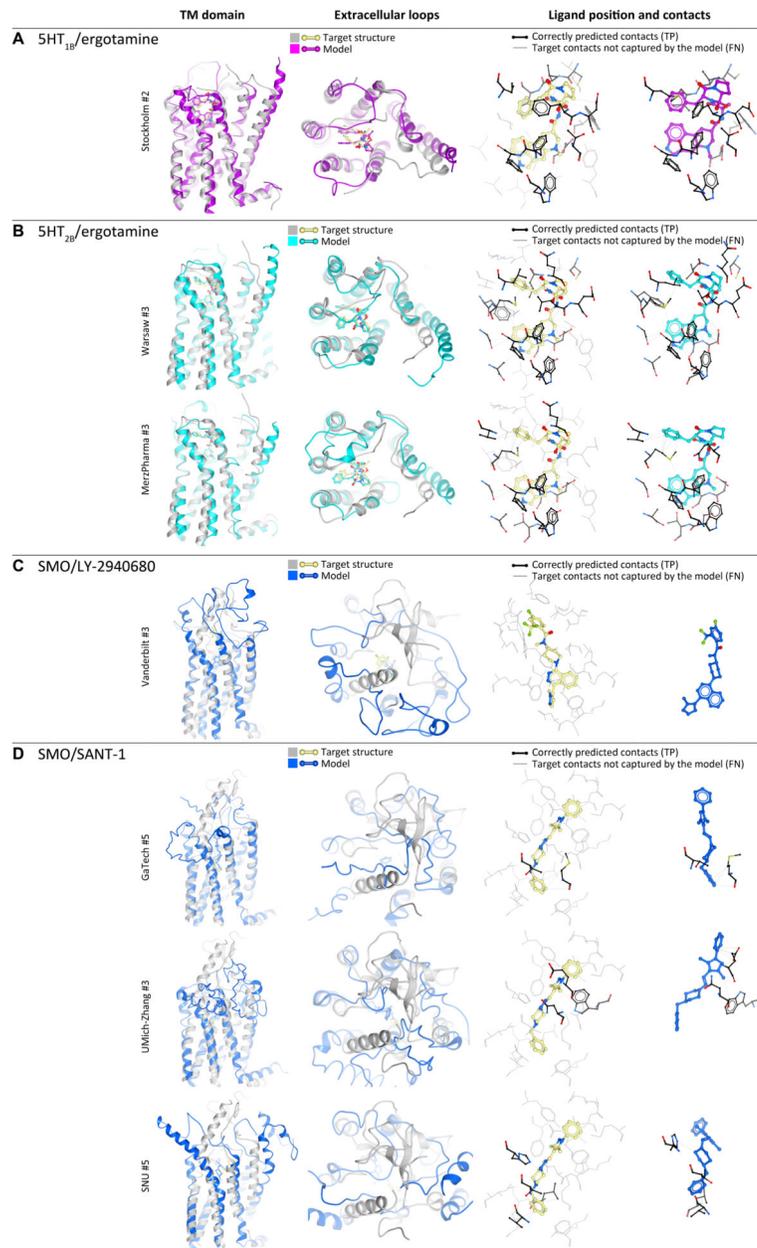
Figure 8.
Highest ranking models of the 4 targets in GPCR Dock 2013.

Table 1

Top predictions in the GPCR Dock 2013 assessment. The listed predictions feature correctness above 1% (for the 5HT$_{1B}$ and 5HT$_{2B}$ complexes) or above 0.1% (for SMO complexes), the most accurate target-template alignment, TM/ECL2 conformation, or pocket fingerprint.

Group	Model	Mean TM alignment error, # residues	TM RMSD, Å (# residues)	TM Q$_{super}$	ECL2 RMSD, Å (# residues)	Protein prediction Z-score	Pocket RMSD, Å (# residues)	Pocket prediction accuracy	Ligand RMSD[a], Å	Contact strength prediction accuracy[a]	Correctness[a,b]
5HT$_{1B}$/ergotamine											
Stockholm-Carlsson	2	0	1.82(190)	65%	4.34(13)	0.86	1.41(19)	77%	1.51(1.9)	47%(44%)	3.35%(2.29%)
Stockholm-Carlsson	3[f]	0	1.66(190)	69%	5.49(13)	0.7	1.31(19)	**81%**	2.09(1.22)	51%(55%)	2.77%(5.69%)
Warsaw	3	0	1.78(190)	66%	4.85(13)	0.74	1.47(19)	72%	2.03(2.25)	39%(35%)	1.76%(1.31%)
Warsaw	1	0	1.78(190)	66%	4.85(13)	0.74	1.47(19)	72%	2.1(2.4)	40%(38%)	1.75%(1.4%)
Copenhagen-Gloriam	2[e]	0	1.68(190)	70%	4.09(13)	**1.12**	1.38(19)	66%	2.45(2.07)	39%(39%)	1.37%(1.74%)
UCBPharma-Pitt	2	0	1.78(190)	66%	6.36(13)	0.27	1.51(19)	73%	2.67(1.56)	41%(43%)	1.33%(2.73%)
COH-Vaidehi	5	0	1.52(190)	77%	7.44(13)	0.26	2.56(19)	74%	1.57(1.94)	27%(29%)	1.25%(1.19%)
LMC-Monash	2	0	2.18(190)	64%	4.71(13)	0.27	2.49(19)	71%	2.97(3.4)	39%(35%)	1.06%(0.74%)
PompeuFabra	2	0	1.72(190)	67%	5.2(13)	0.71	2.48(19)	66%	2.86(1.34)	36%(57%)	1.02%(5.85%)
PompeuFabra	1	0	1.73(190)	67%	4.89(13)	0.8	2.01(19)	65%	2.61(2.68)	33%(39%)	1.01%(1.23%)
Schrödinger	1[c]	0	1.98(190)	68%	5.18(13)	0.38	1.56(19)	69%	5.27(0.72)	36%(65%)	0.31%(**13.01%**)
5HT$_{2B}$/ergotamine											
Warsaw	3	0	2.12(191)	61%	5.54(19)	0.46	2.69(23)	76%	1.05(1.07)	50%(64%)	5.11%(9.56%)
MerzPharma	3	0.23	2.68(191)	55%	9.79(19)	−1.04	5.46(23)	68%	1.14(0.68)	45%(68%)	3.92%(15.57%)
MerzPharma	2[c]	0.21	2.68(191)	55%	9.81(19)	−1.04	5.49(23)	68%	1.37(0.66)	45%(73%)	3.37%(**19.69%**)
MerzPharma	1	0.23	2.69(191)	54%	9.85(19)	−1.06	5.51(23)	67%	1.39(0.76)	44%(73%)	3.15%(17.73%)
Stockholm-Carlsson	2[f]	0	2.16(191)	59%	6.32(19)	0.27	1.94(23)	**79%**	1.64(1.74)	45%(44%)	2.82%(2.56%)
Copenhagen-Nom	1	0	2.38(191)	54%	7.51(19)	−0.24	3.81(23)	63%	1.58(1.72)	41%(47%)	2.5%(2.97%)
Stockholm-Carlsson	4	0	2.26(191)	61%	4.39(19)	0.5	1.88(23)	77%	2.46(1.11)	53%(66%)	2.37%(10.04%)
Padova	2	0	2.03(191)	62%	8.74(19)	−0.01	4.36(23)	69%	2.45(1.38)	43%(65%)	1.64%(7.71%)
UCBPharma-Sands	4	0	2.75(191)	63%	4.55(19)	−0.17	2.65(23)	74%	2.81(2.46)	44%(49%)	1.38%(2.08%)
PompeuFabra	4	0	2.07(191)	61%	8.83(19)	−0.07	3.46(23)	73%	2.09(1.8)	34%(47%)	1.36%(2.79%)

Group	Model	Mean TM alignment error, # residues	TM RMSD, Å (# residues)	TM Q$_{super}$	ECL2 RMSD, Å (# residues)	Protein prediction Z-score	Pocket RMSD, Å (# residues)	Pocket prediction accuracy	Ligand RMSD[a], Å	Contact strength prediction accuracy[a]	Correctness[a,b]
PompeuFabra	5	0	2.07(191)	61%	6.74(19)	0.3	3.65(23)	75%	2.29(2.49)	35%(49%)	1.27%(2.03%)
LMC-Monash	2	0	3.04(191)	47%	6.08(19)	−0.82	3.83(23)	67%	2.73(1.87)	40%(50%)	1.25%(3.04%)
UMich-Zhang	3 [e]	0	1.63(191)	65%	3.94(19)	**1.38**					
SMO/LY-2940680											
Vanderbilt	3	1.35	5.36(189)	18%	14.34(19)	0.05	13.85(21)	5%	4.42	2%	0.12%
GaTech	4 [d]	**0.13**	2.97(189)	38%	6.61(19)	1.04					
UWash	1 [e]	0.2	3.86(189)	37%	2.92(19)	**1.33**					
SNU	1 [f]	0.86	4.16(189)	21%	14.27(19)	0.16	10.99(21)	**25%**			
SMO/SANT-1											
GaTech	5 [f]	0.2	2.94(190)	38%	6.71(19)	1.15	5.61(19)	**59%**	4.31	7%	0.16%
UMich-Zhang	3	0.57	4.14(190)	38%	7.88(19)	0.9	6.33(19)	38%	4.92	11%	0.15%
SNU	5	0.82	3.9(190)	22%	11.27(19)	0.49	8.89(19)	35%	6.01	12%	0.1%
GaTech	4 [d]	**0.12**	2.99(190)	38%	6.7(19)	1.15					
UMich-Zhang	2 [e]	0.57	3.83(190)	42%	3.76(19)	**1.44**					

[a]For 5HT$_1$B and 5HT$_2$B models, values for the ergoline core are provided in parentheses;

[b]correctness assessed as the likelihood of observing the same or higher ligand position/contact variation among experimental structures;

[c]highest accuracy prediction of the ergoline core in the 5HT$_1$B and 5HT$_2$B receptors;

[d]lowest target-template alignment error in the corresponding assessment;

[e]highest protein prediction z-score in the corresponding assessment;

[f]most accurate pocket fingerprint prediction in the corresponding assessment.