



HHS Public Access

Author manuscript

J Comput Aided Mol Des. Author manuscript; available in PMC 2021 October 01.

Published in final edited form as:

J Comput Aided Mol Des. 2020 October ; 34(10): 1027–1044. doi:10.1007/s10822-020-00325-x.

Benchmarking GPCR homology model template selection in combination with *de novo* loop generation

Gregory L. Szwabowski¹, Paige N. Castleman¹, Chandler K. Sears¹, Lee H. Wink¹, Judith A. Cole², Daniel L. Baker¹, Abby L. Parrill¹

¹Department of Chemistry, The University of Memphis, Memphis, TN, 38152, U.S.A

²Department of Biological Sciences, The University of Memphis, Memphis, TN, 38152, U.S.A

Abstract

G protein-coupled receptors (GPCR) comprise the largest family of membrane proteins and are of considerable interest as targets for drug development. However, many GPCR structures remain unsolved. To address the structural ambiguity of these receptors, computational tools such as homology modeling and loop modeling are often employed to generate predictive receptor structures. Here we combined both methods to benchmark a protocol incorporating homology modeling based on a locally selected template and extracellular loop modeling that additionally evaluates the presence of template ligands during these modeling steps. Ligands were also docked using three docking methods and two pose selection methods to elucidate an optimal ligand pose selection method. Results suggest that local template-based homology models followed by loop modeling produce more accurate and predictive receptor models than models produced without loop modeling, with decreases in average receptor and ligand RMSD of 0.54 Å and 2.91 Å, respectively. Ligand docking results showcased the ability of MOE induced fit docking to produce ligand poses with atom root-mean-square deviation (RMSD) values at least 0.20 Å lower (on average) than the other two methods benchmarked in this study. In addition, pose selection methods (software-based scoring, ligand complementation) selected lower RMSD poses with MOE induced fit docking than either of the other methods (averaging at least 1.57 Å lower), indicating that MOE induced fit docking is most suited for docking into GPCR homology models in our hands. In addition, target receptor models produced with a template ligand present throughout the modeling process most often produced target ligand poses with RMSD values 4.5

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. <https://www.springer.com/aam-terms-v1>

Corresponding Author Abby L. Parrill.

Present Addresses 107 Scates Hall, University of Memphis, Memphis, TN 38152

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ASSOCIATED CONTENT

Supporting Information.

The following file containing coordinates of homology models in PDB format is available free of charge. File names are the format target_template_modelID.pdb where model ID is CE for contact energy selected homology models and 1–10 for loop modeled results ordered from lowest to highest energy.

Homology_Models.zip

Publisher's Disclaimer: This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

Å and Tanimoto coefficients > 0.6 after selection based on ligand complementation than target receptor models produced in the absence of template ligands. Overall, the findings produced by this study support the use of local template homology modeling in combination with de novo ECL2 modeling in the presence of a ligand from the template crystal structure to generate GPCR models intended to study ligand binding interactions.

Keywords

Comparative modeling; Comparative protein modeling; Homology modeling; Ligand identification; GPCR; Ligand Docking; Loop Modeling; G protein-coupled receptor

1. INTRODUCTION

G Protein Coupled Receptors (GPCR) are involved in a multitude of cellular signaling pathways. When GPCR signaling is dysregulated, diseases such as cancer, diabetes, and nervous system disorders can manifest. [1] About 34% of FDA-approved drugs target GPCR, reflecting their physiological roles in the regulation and development of disease. [2]

Structurally, GPCR consist of 7 transmembrane (TM) helical domains, 3 extracellular and 3 intracellular loops that connect the membrane spanning domains, an extracellular amino terminus and an intracellular carboxy terminus. [3] The Ballesteros-Weinstein numbering scheme is often used to relate structurally similar sites among different GPCR sequences and classes. [4] In this scheme, the most conserved residue within a transmembrane helix is denoted as the TM.50 residue and other residues within the same domain are numbered relative to this position. For example, the most conserved residue within class A GPCR transmembrane helix 3 is an arginine, thus it is identified as R^{3.50}. An alanine located 5 amino acids prior to the reference arginine (i.e. nearer the amino terminus) would be A^{3.45}.

Of the more than 800 GPCR encoded within the human genome [5], only 70 are represented in the Protein Data Bank [6] by experimentally determined, three-dimensional structures as of April 27, 2020. The lack of experimentally resolved structures for many GPCR has led to the use of computational modeling as a GPCR structure prediction tool. Despite this trend, GPCR modeling is not without challenges and critical decision points, including but not limited to, template structure selection and template-target alignment. Additional challenges in GPCR modeling include effectively sampling and selecting conformations for the extracellular loop (ECL) and intracellular loop (ICL) regions of the target receptor. Accurately modeling the second ECL region of GPCR models likely impacts applications aimed at investigating ligand binding, whereas the ICL region likely impacts applications targeting recognition of intracellular signaling partners and G protein selection and activation. Homology modeling, frequently used to model receptors with unresolved structure, is rooted in the theory that proteins with similar amino acid sequences and common function possess similar structures due to common evolutionary origins. [7] A receptor with high amino acid sequence identity (and similar function) is typically chosen as the template upon which a target receptor sequence is to be modeled. Amino acid identities higher than 30% in these applications are generally considered acceptable. [8] However, selecting a template based solely on global sequence identity may not emphasize GPCR

regions most relevant to the purpose for which the models are being generated. [9] For example, if the goal is to study GPCR interactions with G-proteins, prioritizing intracellular loop sequence homology would most likely produce better models than homology models generated using templates with high ligand binding pocket similarity. The opposite is true for ligand docking studies. Our recent study challenged the conventional use of global sequence identities for GPCR template selection. In that study CoINPocket scores developed by Ngo et al. [10] were utilized to select templates with which to build GPCR homology models. [9] The CoINPocket comparison bases its scoring on the importance and strength of individual ligand-residue interactions across a representative set of class A GPCR and was used to find closely related pharmacological receptor “neighbors” as a ligand identification strategy. [10] In the previous work, models using templates selected using CoINPocket similarities were compared to homology models constructed from templates selected using conventional global sequence identity metrics. The resulting homology models were evaluated in terms of their overall structural similarity and the similarity of docked poses to the reference crystal structures. [9] Homology modeling based on CoINPocket nearest neighbors resulted in models with greater docked ligand pose accuracy than models whose templates were selected based on global sequence similarity alone, although overall similarity of the protein models to the crystal structure did not differ substantially. [9]

Of the characteristic GPCR regions, extracellular loop 2 (ECL2) is often the longest and most variable in terms of both length and amino acid composition. [11] This length produces intrinsic flexibility. In addition, ECL2 often contributes to GPCR ligand binding and selectivity. [12] Site-directed mutagenesis within the ECL2 region can produce drastic changes in ligand binding activity (such as in the human histamine H₄ receptor (H4R)), demonstrating the role ECL2 can play in recognizing and binding ligands. [13] A recent review published by Woolley et al. discussed the various structural impacts of ECL2 on ligand binding. [14] In certain GPCR such as rhodopsin, sphingosine-1-phosphate receptor 1 (S1P1), and free fatty acid receptor 1 (FFAR1), the intrahelical space between transmembrane domains is open in absence of a bound ligand. However, upon ligand binding, ECL2 and the N-terminal domain form a “lid” that covers the binding site, forming a more stable ligand-receptor complex that results in slower ligand dissociation. In other GPCR such as peptide-binding receptors, the β -hairpin ECL2 structure remains open during a ligand binding event. [14] A disulfide bond between conserved cysteine residues 3.25 in transmembrane domain 3 (TM3) and 45.50 in ECL2 often contributes to receptor stabilization, and removal of this bond using site-directed mutagenesis has proven detrimental to GPCR ligand binding. [11] The structural variability of ECL2 is typically addressed by loop modeling, wherein loop conformations can be computationally sampled using *de novo* approaches to best describe loop structure in three dimensions. [15] In another of our previous benchmark studies, ECL2 was modeled in the context of crystallized receptor structures (using accurate loop anchor residues) as a method of testing the structural accuracy of loops generated with a variety of loop modeling methods. [16] This study identified the kinematic loop closure with fragments (KICF) algorithm within Rosetta [17] as most frequently sampling GPCR loop conformations within a 2.5 Å RMSD of the reference crystal structure. [16] While loop modeling within crystallized receptor structures is not generally needed unless the loop region has unresolved atomic coordinates, the results

of the benchmark in question suggest a preferred method for loop structure prediction within the context of a homology model.

The current benchmark study assessed a combination of previously benchmarked modeling choices with new variables on the accuracy of docking into GPCR homology models. The previously benchmarked local similarity-guided template selection and loop modeling protocols are assessed in combination with the presence or absence of the template ligand while modeling the target receptor, as well as three distinct methods of ligand docking. We speculated that inclusion of a ligand from the template receptor during the modeling process would produce better quality receptor models and docked ligand poses, as closely-related GPCR often share ligands and ligands within different GPCR often contact similar residues. [10] A set of 10 crystallized class A GPCR were subjected to this analysis (Scheme 1), with a subset of receptors being used to analyze various docking methods in addition to the protein modeling process. The 10 GPCR employed in this study were: angiotensin type II receptor (AT2R), chemokine receptor type 4 (CXCR4), free fatty acid receptor 1 (FFAR1), histamine receptor 1 (H1R), muscarinic receptor 1 (M1), muscarinic acetylcholine receptor 4 (M4R), nociceptin opioid receptor (NOP), kappa opioid receptor (OPRK), P2Y purinoceptor 12 (P2Y12), and protease-activated receptor 1 (PAR1). Performance of this GPCR modeling workflow was benchmarked using root mean square deviation (RMSD) of alpha carbon positions after superposition of each protein model on its respective reference crystal structure to assess protein model quality and two ligand pose quality metrics: 1) RMSD of ligand atomic positions after superposition of each docked complex on the crystallographic complex; and 2) Tanimoto coefficients representing the proportion of common ligand interaction sites in the modeled and crystallographic complexes.

The long-term goal of the current research is to optimize a modeling protocol to investigate GPCR complexes for any target GPCR sequence. Potential applications of this modeling protocol include prioritization of candidate ligands for experimental screening and generation of hypotheses regarding receptor sites involved in ligand binding to be tested by subsequent site-directed mutagenesis. Improved accuracy in candidate ligand prioritization will help accelerate receptor deorphanization [18] and help improve the identification of chemical tool compounds to probe receptor function or to serve as preclinical lead compounds in the drug discovery process. Overall, this work demonstrates that the integration of loop modeling with homology models constructed from locally selected template structures produces better receptor models (0.43 Å average RMSD decrease), as well as better docked ligand poses (2.91 Å average ligand RMSD decrease) than non-loop modeled local template homology models. In addition, this work exhibits that inclusion of a pharmacological neighbor receptor's ligand throughout the receptor modeling process produces a greater proportion of high quality docked complexes than receptors modeled without a ligand present (30% of best docked poses exhibited RMSD < 4.5Å when selected via ligand complementation).

2. RESULTS AND DISCUSSION

2.1 Homology Model Template Selection

A summary of target and template GPCR, the local similarity measure (CoINPocket score), GenBank accession numbers, and PDB identification codes used in this study can be found in Table 1. Local templates used for the homology modeling of GPCR target proteins were selected from a pool of GPCR with available crystal structures using the CoINPocket local similarity measure. [6, 10] GPCR from the same subfamily as the target were excluded as templates, for example, C-C chemokine receptor type 5 (CCR5) was not selected as a template from which to model CXCR4. Four receptors were modeled using two different target-template pairings so models of the same receptor with differing template structures (and therefore different local similarity scores) could be compared.

According to the CoINPocket scores set forth by Ngo et al., any receptor compared to itself has a GPCR-CoINPocket score of 5.47. [10] This self-similarity establishes a maximal binding pocket similarity score for target sequence/template receptor pairs. It should be noted that a higher local similarity score does not always translate to a high global unweighted sequence similarity and vice versa. Local similarity scores used herein ranged from 1.23 (M₄R/NOP) to 2.58 (M₁/H₁R), indicating that the binding site residues were not perfectly conserved between receptor pairings but still shared appreciable residue conservation. Unweighted global sequence identities ranged from 7% (FFAR1/P2Y12 and P2Y12/PAR1) to 36% (M₁/H₁R). Percent identities at the low end of this range would generally be considered unacceptable for the purpose of constructing homology models unless substantial further refinement (such as loop modeling) were included in the workflow.

2.2 Protein Model Development and Analysis

Homology models developed in this study were benchmarked against crystallographic reference structures. In order to place these comparisons in the appropriate context, RMSD values between different PDB entries for the same receptor were calculated (Table 2) to set a baseline for experimental variability. On average, the experimental alpha carbon variation among multiple structures of the same receptor was 1.75 Å, which sets a range of expectation for models generated using our methodology. Our expectation is that models that differ from the target crystal structure by no more than 2 times the average experimental variation, or 3.5 Å, should be considered high quality models.

Homology models were constructed in MOE using two different model selection settings, Generalized Born Solvation (GBVI) scoring [19] and contact energy [20]. The latter produced models with binding pockets that better matched ligand locations in GPCR crystallographic complexes in terms of location and volume based on the Alpha Shapes methodology discussed in the methods section. This is illustrated using the M1 receptor in Figure 1 (ligand structure 3 shown in Figure 2).

Two approaches were used to generate homology models to be used as the starting points for loop modeling and in ligand docking experiments. Both approaches produced eleven homology models for each target:template GPCR pairing modeled using the default modeling options in MOE, with the exception of selecting effective atomic contact energy as

the scoring method. Approach A included no ligand in the binding pocket of the template receptor and was applied to all pairings in Table 1. Approach B was applied only to the target:template pairing with the higher CoINPocket score for each target receptor and retained the crystallographic ligand from the template receptor and utilized this ligand as an 'Environment for Induced Fit'. [21]

Software packages and algorithms for ECL2 loop modeling in the context of GPCR crystal structures have been previously benchmarked, with the best performance achieved by Rosetta's kinematic closure with fragments (KICF) algorithm. [16] In the current study, KICF was used to sample ECL2 conformations in the context of homology models generated by approaches A and B with no ligand present in the binding pocket, herein referred to as loop modeling approaches A1 and B1, respectively. Additionally, KICF was used to sample ECL2 conformations in the context of homology models generated by approach B with ligand present during the loop modeling process (overall process of protein model generation with ligand present through both modeling steps considered approach B2).

For each set of loop modeling results, the ten lowest-scored models with intersulfur (Cys 3.25-Cys 45.50) distances ≤ 5.1 Å were selected for further examination. Each of the ten loop modeled structures and the local template homology model was superposed on the reference crystal structure using non-loop residues, followed by calculation of alpha-carbon receptor RMSD values for the entire structure (Figures 3 and 4 and Supplementary Tables S1 and S2) and ECL2 region (Supplementary Tables S3 and S4). A representative superposition of a local template homology model on the reference crystal structure is shown in Figure 5. Superpositions of loop modeled structures on reference crystal structures are shown in Figure 6. Note that RMSD values for local template homology models are different from the values discussed in the prior benchmark [9], as the homology models discussed therein were chosen based on GBVI scoring rather than effective atomic contact energies. Alpha carbon RMSD values for models generated using approach A1 based on the highest-similarity templates ranged from 2.76 (M4R) to 6.32 Å (FFAR1) with an average of 4.34 Å (Table S1 and Figure 3A). Initial homology models for three target:template pairings achieved our high quality metric of 3.5 Å, H1R (3.15 Å), M1 (2.93 Å), and M4R (2.76 Å). While these receptor RMSD values are not sub-angstrom (<1.0 Å) or near-atomic (<2.5 Å), comparison of the generated RMSD values to structural variation within crystallographic structures for each receptor allows for a better examination of model quality (Table 2). For example, the best scoring initial homology model for P2Y12 had an alpha carbon RMSD of 4.07 Å, which falls within 2.32 Å of the average variation present in crystallized structures and within 1.4 Å of the observed 2.67 Å variation between the most diverse pair of P2Y12 PDB entries, indicating that models being generated with the methodology discussed thus far are adequate representations of the receptors being modeled. Loop modeling led to substantial improvements in model quality in two cases, FFAR1 and OPRK. In these cases, a loop-modeled structure was 2.27 Å and 1.24 Å lower in RMSD relative to the crystallographic reference structure. There were no cases in which loop modeling resulted in a substantial (>0.25 Å) loss of protein model quality.

Four receptors were modeled using two different templates (AT2R, H1R, M4R, P2Y12, Table 1). Initial homology model alpha carbon RMSD values were lower for 3 of 4 receptors

(H1R, M4R, P2Y12) when modeled with a more similar template (Figure 4, Table S2), indicating that the use of a template with a higher local similarity score leads to better homology models. The largest difference in initial homology model quality due to template choice was for the P2Y12 receptor, for which the distribution of alpha carbon RMSD values for loop optimized models based on the low and high similarity templates are significantly different based on the Kolmogorov-Smirnov test at 95% confidence. The lower similarity template, FFAR1 (PDB entry 5TZR) has an unusual ligand binding mode that involves ligand insertion between TM segments 3 and 4, with a resulting offset of TM3 relative to other known GPCR structures (Figure 7). Thus, the dramatic difference in P2Y12 model quality reflected in the over 4 Å RMSD difference is likely less a consequence of similarity differences than in the truly unusual structure features of the FFAR1 crystal structure relative to all other currently known GPCR structures. Loop modeling provided substantial improvements in three of four receptors modeled based on lower similarity templates (H1R, M4R, P2Y12) without detrimental impact in other cases, indicating the value of loop modeling for its potential to produce improved receptor models. This is further supported by the observation that distributions of loop optimized model RMSD values between low and high similarity templates are not significantly different for H1R or M4R based on the Kolmogorov-Smirnov test at 95% confidence. This suggests that loop modeling can compensate in some cases for differences in initial model quality.

The effect of including the template ligand only during homology modeling (approach B1) or during both homology modeling and loop modeling (approach B2) was also assessed (Figure 3 and Table S1). Initial homology models generated using the 'Environment for Induced Fit' option via approach B1 possessed similar average receptor RMSD values to models created by approach A1 (4.24 Å vs. 4.34 Å, respectively), indicating that inclusion of a binding pocket ligand during the homology modeling process did not substantially impact protein model quality. Loop modeling produced improved models in a similar number of cases by all approaches.

2.4 Ligand Docking and Analysis

Three docking methods were compared in this study: MOE induced fit, MOE rigid receptor, and Rosetta docking. These methods were first assessed for their performance in redocking ligands (Figure 2) into six reference crystal structures. In order to compare the docked ligand poses generated by each method to the crystallographic ligand positions, ligand RMSD (LRMSD) values were calculated using MOE. Though LRMSD values are an output of Rosetta's ligand docking process, LRMSD values reported here were calculated in MOE to ensure that a consistent superposition process was used prior to LRMSD calculation. Rosetta performed worst of the three methods when docking into crystallized receptor structures. No poses produced possessed LRMSD values under 3 Å (Table S5), in contrast to poses with LRMSD under 2 and 3 Å in 5 of 6 cases produced by the MOE rigid and induced fit docking, respectively, in a previous benchmark. [9]

These docking algorithms were further assessed by docking ligands (Figure 2) into models from approach A1. Examples of poses produced by each docking method can be found in Figure 8. The ability of each method to sample docked poses similar to the crystal structure

was assessed using the pose with the lowest LRMSD value resulting from each method (Figure 9). In addition, LRMSD averages without FFAR1 were calculated as all ligand poses for this receptor had LRMSD above 7 Å (Table S5), which can most likely be attributed to the unusual ligand binding mode in the crystallized reference structure (Figure 7). The distribution of LRMSD values was not significantly different at the 95% confidence level between methods based on a Kolmogorov-Smirnov test. However, guidance on docking protocol selection can still be drawn on the basis of the proportion of results meeting a performance target. In this case, a docking performance target LRMSD of 4.5 Å was set as a reasonable increase of 1.5 Å higher than the majority of re-docking results for MOE induced fit. The best ligand poses sampled with MOE induced fit docking met our performance target in 5 of 6 cases while both MOE rigid and Rosetta docking sampled best poses < 4.5 Å in only 2 of 6 and 4 of 6 cases, respectively, (Figure 9, Plot A), indicating that MOE Induced Fit docking sampled quality poses most often. This holds true when considering average LRMSD values without FFAR1 as well: the average MOE induced fit best pose LRMSD averaged 3.73 Å, lower than both MOE rigid (4.64 Å) and Rosetta (4.29 Å) docking (Table S5). The best poses produced by MOE rigid receptor docking typically had the highest LRMSD values (all receptors except NOP, Figure 9), illustrating the importance of flexible residue side chains in sampling ligand poses representative of the crystallographic ligand pose. This can likely be attributed to the differences between the homology modeled structures and the crystallographic reference structures, as MOE rigid docking performed well at re-docking ligands into crystal structures in a previous benchmark study. [9]

When a crystallographic reference is not available, ligand poses must be selected based on available information from the docked pose alone, rather than determination of LRMSD using a reference structure. Ligand poses are typically selected in such cases based on pose scores. In order to measure scoring performance of the docking methods, the lowest LRMSD among the top 10 scoring ligand poses using either the scoring function associated with the method (T10) or a complementation scoring method (T10 Comp) for each receptor was determined (Figure 9 panels B and C, Table S5). Excluding the comparison of MOE induced fit and MOE rigid T10 Comp LRMSD values ($p = 0.03$), the distribution of LRMSD values across the six receptors was not significantly different at the 95% confidence level between pose selection methods based on a Kolmogorov-Smirnov test. However, guidance on pose selection method can still be drawn based on differences in average results. Average T10 Comp values for MOE induced fit, MOE rigid, and Rosetta docking across the subset of six receptors were 5.58, 8.02, and 6.47 Å, respectively. Average T10 values for the same methods were 5.90, 8.13, and 7.47 Å, respectively. The T10 Comp values were lower by 0.32, 0.11, and 1.00 Å, respectively. These data support two conclusions: 1) that pose selection using complementation scoring provided a slight decrease in LRMSD regardless of the docking method used to generate the poses, although the differences were not significant based on the Mann Whitney U test and 2) selected poses from MOE induced fit docking had lower LRMSD than those selected from the other docking methods (significant difference at the 95% confidence level achieved only for the induced fit versus rigid comparison based on the Mann Whitney U test).

Comparisons of docked poses for receptors shared between the current and previous benchmark studies [9] demonstrate the value of including loop modeling in the protein

modeling protocol. Induced fit poses (derived from approach B2) selected by complementation score in the current study (Table S6) had an average LRMSD of 6.13 Å and an average Tanimoto coefficient of 0.51 in contrast to the prior study with an average LRMSD of 10.44 Å and an average Tanimoto coefficient of 0.31. Thus docking results are improved by sampling ECL2 loop conformations after homology modeling.

Ligands were docked into four sets of receptor models generated using modeling approach A1 based on two different templates (Figure 10 and Table S7). In three out of four cases, a lower LRMSD value and higher Tanimoto coefficient was obtained for the best pose sampled when docking into the models based on higher similarity templates. In every case pose selection based on complementation score (as evidenced by T10 Comp LRMSD values) selected lower RMSD poses (three of the four also with higher Tanimoto coefficients) from docking into models based on the higher similarity template. Though Tanimoto coefficients were not found to be significantly different at the 95% confidence level based on the Kolmogorov-Smirnov test, distributions of selected pose RMSD values between the high and low similarity templates were significantly different at the 95% confidence level ($p = 0.03$).

MOE induced fit docking was used to dock ligands into models based on the highest CoINPocket scored templates generated using modeling approaches A1, B1 and B2 (Figures 11 and 12, Table S6). These results were assessed using two metrics, LRMSD (Figure 11) and Tanimoto coefficients (Figure 12). The distributions of selected pose RMSD value or Tanimoto coefficients between the methods were not significantly different at the 95% confidence level based on the Kolmogorov-Smirnov test. However, comparison of results guided by a target LRMSD threshold of 4.5 Å or lower for high-quality poses, coupled with a Tanimoto coefficient of 0.6 or greater (at least 60% of ligand contact residues shared) does provide some guidance for protocol selection. Based on these targets, approach B2 coupled with complementation scoring for pose selection can be clearly identified as the best protocol for obtaining high quality ligand poses when using homology models in docking studies. In particular, docking into five of the targets sampled a pose that met both of these thresholds, and the complementation scoring included a pose that met both thresholds in three of the five cases. Approaches A1 and B1 yielded zero cases in which both thresholds were met after pose selection. Overall, homology modeling and loop modeling with a template protein ligand present produces target protein ligand poses that meet performance goals for a greater proportion of docking targets than in the absence of the template protein ligand. Thus, we recommend the use of receptor modeling approach B2, wherein a template ligand is present throughout both homology modeling and loop modeling of the target receptor.

3. CONCLUSIONS

The overall goal of the work described here was to assess a combination of previously benchmarked modeling choices with new variables on the accuracy of docking into GPCR homology models. The previously benchmarked local similarity-guided template selection and loop modeling protocols are assessed in combination with the presence or absence of the

template ligand while modeling the target receptor as well as three distinct methods of ligand docking.

Loop modeling led to substantial improvements (>1 Å decreases in alpha carbon RMSD compared to the crystallographic reference structure) in protein model quality in two cases by all three protein modeling approaches, FFAR1 and OPRK (Figure 3, Table S1). There was only one case in which loop modeling caused a substantial reduction in protein model quality (PAR1 modeled with a P2Y12 ligand as environment for induced fit). Loop modeling also provided substantial improvements in three of four receptors modeled based on lower similarity templates without detrimental impact in other cases (Figure 4), indicating the value of loop modeling for its potential to produce improved receptor models. The refined ECL2 regions of each receptor also played a role in producing more accurate ligand poses, as evident by the average 2.59 Å decrease in LRMSD (from 7.67 Å to 4.76 Å) and 0.11 increase (from 0.45 to 0.56) in Tanimoto coefficients between contact residues compared to ligand poses docked into initial homology models (Table S6).

When docking native ligands into homology models generated using the protocols discussed herein, ligand poses with LRMSD values within 4.5 Å of the crystallized reference structure ligand pose were most often sampled using MOE induced fit docking (Figure 9). When looking at methods of pose selection (pose scoring and ligand complementation) across all 3 docking methods (Figure 9), MOE induced fit docking poses selected via T10 or T10 Comp scoring were far better than MOE rigid receptor or Rosetta docking. NOP docking results illustrate this point, as the best ligand pose selected via complementation scoring from the MOE induced fit docking was 3.45 Å and 5.84 Å lower than from Rosetta and MOE rigid receptor docking, respectively (Supplementary Table S5). Though the need to further validate docked ligand models via methods such as site-directed mutagenesis is clear, these results remain promising in terms of producing ligand poses resembling those of crystallized ligands.

While MOE induced fit docking often produces the best ligand poses, nuances in the other two docking methods must be considered. The Site Finder function within MOE was used to provide user-identified docking sites for MOE induced fit and rigid receptor docking. Rosetta docking requires a user-defined XYZ coordinate binding pocket centroid, which was defined in this work as the centroid of the site identified by the MOE Site Finder function. Rosetta also uses 'movers', which change the conformation of the ligand-receptor complex during the docking process. [22] Arguments for these movers include parameters such as 'box_size', a maximum translation of a ligand from its starting point, and scoring grid width, which defines the cubical space around which the ligand will be scored. [23] Since all methods utilize user-defined parameters to guide the process, docking results can vary depending on the values used for these parameters. Efforts were undertaken to match parameters between methods as much as possible in order to provide a fair comparison.

Though homology modeling receptors using the 'Environment for Induced Fit' option in MOE produced protein models of relatively similar quality as those produced using the default homology modeling protocol in MOE (4.24 Å vs 4.34 Å, respectively), complementation score pose selection on docking results from receptors modeled with a

template ligand present throughout both homology and loop modeling is the method that most often selected high quality poses for any target (LRMSD = 4.5 Å and Tanimoto coefficient = 0.6) (Figures 11 and 12, Table S6), suggesting the use of receptor modeling approach B2 in future efforts.

These results provide further evidence that GPCR homology model construction from templates selected on the basis of similarity scores weighted toward sites involved in strong ligand binding interactions (CoINPocket scores) improves docking pose accuracy (Figure 10). Among 4 receptors modeled using templates of differing local similarity (average similarity score 1.58 versus 2.26), average LRMSD after pose selection by complementation was below 6 Å for the models constructed based on higher similarity templates, but over 8 Å for the models constructed based on lower similarity templates (Supplementary Table S7).

A suggested workflow to generate GPCR models to be used to study ligand interactions can be extracted based on these comparative performance results. First, homology models should be constructed based on templates with the highest local similarity scores and with template ligand included as the 'Environment for Induced Fit' in MOE. ECL2 conformations should be sampled with the template ligand present using the KICF algorithm in Rosetta constraining formation of the C3.25-C45.50 disulfide bond. Ligand docking into the top 10 scored resulting models using induced fit docking in MOE followed by pose selection via ligand complementation will serve to select high quality poses from the set of sampled poses.

4. METHODS

4.1 Target/Template Selection and Preparation

Template sequences for homology modeling of targeted, previously crystallized receptors used in this study (Table 1) were selected using the contact-informed neighboring pocket (CoINPocket) scores developed by Ngo et al. to emphasize similarities at sites that make important and strong ligand interactions in a set of 27 unique class A GPCR crystal structures. [10] In addition to calculating CoINPocket scores, Ngo et al. calculated unweighted global similarity values for each possible sequence pairing. Global similarity values for the receptors used in this benchmark can be found in Table 1. For the initial subset of receptors, a template for each target GPCR was selected that exhibited the highest CoINPocket local similarity score and possessed a previously solved and deposited crystal structure. Each template selected was not 1) the target GPCR or 2) a closely related GPCR that binds the same endogenous ligand. The CoINPocket score-based models were termed "local template" models. A subset of four target receptors were additionally modeled on the basis of a lower similarity template for comparison.

4.2 Homology Model Construction and Analysis

Homology modeling began with the deletion of non-GPCR sequence segments from template and crystallographic reference structures including fusion partners such as T4 lysozyme or thermostabilized cytochrome b₅₆₂RIL from the selected template, as these are non-native segments used to stabilize a single receptor conformation for crystallization. [24]

Each target sequence was aligned to the selected template sequence using a two-step procedure in MOE 2018.01. [21] First, the sequences were aligned using MOE's "sequence only" method of automatic alignment. After the initial alignment, any gaps in helical segments were manually shifted into the structurally variable intracellular and extracellular loop regions while ensuring that conserved TM.50 residues remained aligned. [9] This structure-independent alignment was performed to account for the variability in sequence length and composition of loop regions and to avoid distortions within the more structurally conserved helical transmembrane domains. Homology models were then generated using two approaches. Models created using approach A utilized the default settings in the MOE homology modeling interface, with the exception of scoring models based on effective atomic contact energy. The second approach (approach B) utilized the same settings for homology model generation as approach A, with the addition of retaining the crystallographic ligand from the template structure as the 'Environment for Induced Fit'. Approach B was applied only to the target:template pairing with the higher CoINPocket score for each target receptor. The resulting homology models were then superposed onto the crystallized reference structure based on non-loop residues prior to calculation of receptor alpha-carbon RMSD values, both for the entire sequence and for loop segments, as metrics of structural similarity.

4.3 De Novo Extracellular Loop 2 (ECL2) Modeling

Prior to ECL2 modeling, loop 'anchor' residues were selected. For each receptor, the final helical residue of TM4 and first helical residue of TM5 of the lowest contact energy local template homology model were used as anchor points (Table 3), with loop modeling then sampling conformations of all residues between the anchor points. Fragment libraries required by the KICF algorithm [17, 25] were then generated using the *Robetta* server. [26] To generate these fragments, a FASTA formatted sequence containing the nine residues prior to the first loop anchor, the ECL2 sequence and the nine residues after the second loop anchor was submitted to the server. An atomic disulfide constraint that restricts the distance of sulfur atoms in critical cysteine residues 3.25 of TM3 and 45.50 of ECL2 to 5.1 Å (representative of the known disulfide bond in many GPCR structures) was incorporated into the loop modeling protocol. [27] This constraint is meant to emulate the filtering of models with unrealistic disulfide distances done in the previous benchmark, as filtering ECL2 models based on disulfide distance > 5.1 Å often produced models with better loop RMSD values. [16] When the constraint was applied to the Rosetta loop modeling protocol, far fewer models exhibiting disulfide distances uncharacteristic of GPCR resulted. Examples of models with unrealistic disulfide distance include models with steric clashes due to sub-angstrom disulfide distances or models with extremely large inter-sulfur distances. This atomic constraint also reduces the loop model conformational space. ECL2 models were produced using three different approaches, each utilizing a different combination of homology and loop modeling methods:

- A1.** Homology and loop modeling without a template ligand present.
- B1.** Homology models created using template ligand as 'Environment for Induced Fit', loop modeled without template ligand present in the binding pocket.

- B2.** Homology models created using template ligand as ‘Environment for Induced Fit’, loop modeled with template ligand present in the binding pocket.

Each approach generated a total of 250 disulfide-constrained ECL2 models for each of the target:template pairings in this benchmark. Greater loop sampling was achieved for loops meeting the 5.1 Å threshold than in the previous benchmark [16] because all 250 of the constructed loops met the disulfide distance filter. This number was ten-fold higher than the number of models typically meeting the 5.1 Å disulfide distance filter out of the 1000 generated models for each target in the prior benchmark study. [16] The ECL2-TM3 disulfide bond was formed in the top 10 lowest scoring models followed by geometry optimization of the ECL2 segment in MOE. The resulting local template derived and ECL2 optimized models were used for subsequent ligand docking. Receptor alpha-carbon RMSD values were calculated for ECL2 optimized models as described in the prior section.

4.4 Ligand Docking

Ligand docking was initially performed with a subset of 6 targets generated via modeling approach A1 using both the MOE and Rosetta software packages. Three distinct methods were examined in this study: MOE induced fit docking, MOE rigid receptor docking, and RosettaScripts ligand docking (herein referred to as Rosetta docking). MOE induced fit docking places the active ligand into a user-defined binding site inside a target receptor whose residue side chains are allowed to move freely during the refinement stage. MOE rigid receptor docking places the active ligand into a user-defined binding site inside a target receptor whose side chains are held static during both the placement and refinement stage. The docking methods employed by MOE continuously sample ligand conformations as the docking proceeds, allowing for a best fit of the ligand within a potential binding pocket. In contrast, Rosetta docking differs from MOE in that ligand conformations are generated prior to the docking process, rather than actively sampling ligand conformations within the binding pocket during the docking analysis. Ligand conformations and a user-defined binding pocket of a target receptor with flexible residue side chains are required inputs for Rosetta docking. [23] Conformations for each ligand docked using Rosetta were generated using MOE’s Conformational Search tool, which outputs a database of energetically reasonable ligand conformations. In addition to homology models, reference crystal structures were used as docking targets for Rosetta in order to compare docking performance to the previous benchmark. [9] The remainder of ligand docking was performed using only MOE induced fit docking based on the results of the docking method comparisons.

Each local template model and the top ten sampled ECL2 optimized models were utilized as docking targets. Each protein and ligand structure was prepared at pH 7.4 using the “QuickPrep” function in MOE to 1) ensure proper protonation and charge at the desired pH and 2) minimize the structure using the AMBER10:EHT forcefield. [28] Once each receptor model was prepared, the Site Finder function in MOE was used to define a binding pocket within the receptor model. This function organizes potential binding sites by the volume of alpha spheres within a potential binding pocket, based on the Alpha Shapes methodology of Edelsbrunner et al. [29] Both forms of MOE ligand docking used in this study utilized the MOE Site Finder function to define the docking site, though it is not the sole method of binding pocket selection available within MOE. Rosetta, on the other hand, uses XYZ

coordinates to define a theoretical binding site within a receptor that restricts ligand movement within that defined site. The XYZ coordinates of the center of the binding site defined as the docking site for MOE docking were used to define the binding site during Rosetta docking.

Ligands docked into each receptor can be found in Figure 2. Both MOE induced fit and rigid receptor docking protocols generated 1000 initial ligand placement poses for each of the top ten lowest scoring receptor models, from which the top 400 poses based on the London dG scoring function were passed on to the refinement stage. [21]. Refinement used the Generalized-Born Volume Integral/Weighted Surface area (GBVI/WSA) scoring function. [19] For each of the 10 receptor models with different ECL2 conformations, the top 5 refined ligand poses were retained as final complexes after the refinement stage to provide 50 poses overall for each target modeled. In order for the Rosetta ligand docking to adequately match the sampling of MOE docking, 1,000 ligand poses were generated for each of the top ten scoring models. All 1,000 poses were retained for each run as Rosetta docking does not remove poses through the workflow.

Once docked, an alpha carbon superposition of each receptor-ligand complex onto a crystallized reference structure was constructed and a heavy atom ligand RMSD (LRMSD) was calculated between the two ligand poses. Tanimoto coefficients were calculated to compare first neighbor residues to the ligand in the residue interaction network calculated using the RING 2.0 server between docked poses and crystallographic reference structures. [30] Two sets of poses were selected using different criteria. The first pose set included the ten lowest scoring models based on the scoring function for each respective docking method. The second pose set included the ten poses with the top ten ligand complementation scores, which reflect the ratio of hydrogen bonds made by the ligand when docked into a receptor to the number of ligand hydrogen bonding sites. This ligand complementation score reflects the proportion of polar functional groups that are involved in hydrogen bonding interactions. Polar functional groups within a ligand are able to participate in a maximal number of hydrogen bonds while free in solution. Docked ligand poses with polar functional groups not involved in hydrogen bonds are less energetically favorable in a bound environment than in water for both entropic and enthalpic reasons, which this score attempts to capture.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number R15MH109034. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

MOE 2018.01 was used throughout this work and was provided as a courtesy of the Chemical Computing Group. In addition, the authors would like to thank the Meiler Lab at Vanderbilt University for their outstanding discussion and support of the Rosetta software suite. The University of Memphis High-Performance Computing (HPC) facilities were used constantly throughout this work and the authors wish to express our gratitude.

Funding Sources

This work was supported by the National Institutes of Health

[grant R15 MH109034].

ABBREVIATIONS

AT2R	angiotensin II type 2 receptor
CCR5	C-C chemokine receptor type 5
CXCR4	C-X-C chemokine receptor 4
DP2	Prostaglandin D2 receptor 2
ECL	extracellular loop
FFAR1	free fatty acid receptor 1
GPCR	G protein-couple receptors
H1R	histamine receptor H1
ICL	intracellular loop
KICF	kinematic closure with fragments
LRMSD	ligand RMSD
M1	muscarinic acetylcholine receptor 1
M4R	muscarinic acetylcholine receptor 4
MOE	Molecular Operating Environment
NOP	nociceptin opioid receptor
OPRK	κ -opioid receptor
P2Y12	P2Y purinoceptor 12
PAR1	protease-activated receptor 1
RMSD	root-mean squared deviation
TM	transmembrane

References

1. Hu G, Mai T, Chen C. Visualizing the GPCR Network: Classification and Evolution. *Sci. Rep.* 2017;7(1):15495; 15495–15495. [PubMed: 29138525]
2. Hauser AS, Chavali S, Masuho I, Jahn LJ, Martemyanov KA, Gloriam DE, et al. Pharmacogenomics of GPCR Drug Targets. *Cell.* 2018;172(1–2):41,54.e19. [PubMed: 29249361]
3. Gacasan SB, Baker DL, Parrill AL. G protein-coupled receptors: the evolution of structural insight. *AIMS Biophys.* 2017;4(3):491–527. [PubMed: 29951585]

4. Ballesteros JA, Weinstein H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors In: *Methods in neurosciences*. Elsevier; 1995 p. 366–428.
5. Fredriksson R, Lagerström MC, Lundin L, Schiöth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol*. 2003;63(6):1256–72. [PubMed: 12761335]
6. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Res*. 2000;28(1):235–42. [PubMed: 10592235]
7. Cavasotto CN, Phatak SS. Homology modeling in drug discovery: current trends and applications. *Drug Discov Today*. 2009;14(13–14):676–83. [PubMed: 19422931]
8. Xiang Z Advances in homology protein structure modeling. *Curr Protein Peptide Sci*. 2006;7(3):217–27. [PubMed: 16787261]
9. Castleman PN, Sears CK, Cole JA, Baker DL, Parrill AL. GPCR homology model template selection benchmarking: Global versus local similarity measures. *J. Mol. Graph. Model*. 2019 1;86:235–46. [PubMed: 30390544]
10. Ngo T, Ilatovskiy AV, Stewart AG, Coleman JLJ, McRobb FM, Riek RP, et al. Orphan receptor ligand discovery by pickpocketing pharmacological neighbors. *Nat. Chem. Biol*. 2017;13(2):235–42. [PubMed: 27992882]
11. Wheatley M, Wooten D, Conner MT, Simms J, Kendrick R, Logan RT, et al. Lifting the lid on GPCRs: the role of extracellular loops. *Br J Pharmacol*. 2012;165(6):1688–703. [PubMed: 21864311]
12. Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, et al. The dynamic process of β 2-adrenergic receptor activation. *Cell*. 2013;152(3):532–42. [PubMed: 23374348]
13. Wifling D, Bernhardt G, Dove S, Buschauer A. The extracellular loop 2 (ECL2) of the human histamine H4 receptor substantially contributes to ligand binding and constitutive activity. *PLoS One*. 2015;10(1):e0117185. [PubMed: 25629160]
14. Woolley MJ, Conner AC. Understanding the common themes and diverse roles of the second extracellular loop (ECL2) of the GPCR super-family. *Mol Cell Endocrinol*. 2017;449:3–11. [PubMed: 27899324]
15. Yarnitzky T, Levit A, Niv MY. Homology modeling of G-protein-coupled receptors with X-ray structures on the rise. *Curr Opin Drug Discov Devel*. 2010;13(3):317–25.
16. Wink LH, Baker DL, Cole JA, Parrill-Baker AL. A Benchmark Study of Loop Modeling Methods Applied to G Protein-Coupled Receptors. *J. Comput. Aided Mol. Des*. 2019;33(6):573–95. [PubMed: 31123958]
17. KIC with fragments [Internet]; 2014 [updated July 27; cited October 15 2019]. Available from: https://www.rosettacommons.org/docs/latest/application_documentation/structure_prediction/loop_modeling/KIC_with_fragments.
18. Chung S, Funakoshi T, Civelli O. Orphan GPCR research. *Br J Pharmacol*. 2008;153(S1):S339–46. [PubMed: 18071299]
19. Wojciechowski M, Lesyng B. Generalized Born model: Analysis, refinement, and applications to proteins. *J. Phys. Chem. B*. 2004;108(47):18368–76.
20. Zhang C, Vasmatazis G, Cornette JL, DeLisi C. Determination of atomic desolvation energies from the structures of crystallized proteins. *J Mol Biol*. 1997;267(3):707–26. [PubMed: 9126848]
21. Chemical Computing Group U. Molecular Operating Environment (MOE). 2018.
22. Fleishman SJ, Leaver-Fay A, Corn JE, Strauch E, Khare SD, Koga N, et al. RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS one*. 2011;6(6):e20161. [PubMed: 21731610]
23. Lemmon G, Meiler J. Rosetta Ligand docking with flexible XML protocols In: *Computational Drug Discovery and Design*. Springer; 2012 p. 143–55.
24. Heydenreich FM, Vuckovic Z, Matkovic M, Veprintsev DB. Stabilization of G protein-coupled receptors by point mutations. *Front Pharmacol*. 2015;6:82. [PubMed: 25941489]
25. Stein A, Kortemme T. Improvements to robotics-inspired conformational sampling in rosetta. *PLoS one*. 2013;8(5):e63090. [PubMed: 23704889]

26. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 2004;32(suppl_2):W526–31. [PubMed: 15215442]
27. Constraint File [Internet]; 2015 [updated Aug 30,; cited October 15, 2019]. Available from: https://www.rosettacommons.org/docs/latest/rosetta_basics/file_types/constraint-file.
28. Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, Duke RE, et al. Amber 10. University of California; 2008.
29. Edelsbrunner H, Facello M, Fu P, Liang J. Measuring proteins and voids in proteins. ; 1995.
30. Piovesan D, Minervini G, Tosatto SC. The RING 2.0 web server for high quality residue interaction networks. *Nucleic Acids Res.* 2016;44(W1):W367–74. [PubMed: 27198219]
31. Zhang H, Han GW, Batyuk A, Ishchenko A, White KL, Patel N, et al. Structural basis for selectivity and diversity in angiotensin II receptors. *Nature.* 2017;544(7650):327. [PubMed: 28379944]
32. Wang L, Yao D, Deepak RK, Liu H, Xiao Q, Fan H, et al. Structures of the human PGD2 receptor CRTH2 reveal novel mechanisms for ligand recognition. *Mol Cell.* 2018;72(1):48,59. e4. [PubMed: 30220562]
33. Wu B, Chien EYT, Mol CD, Fenalti G, Liu W, Katritch V, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science (New York, N.Y.).* 2010;330(6007):1066–71.
34. Lu J, Byrne N, Wang J, Bricogne G, Brown FK, Chobanian HR, et al. Structural basis for the cooperative allosteric activation of the free fatty acid receptor GPR40. *Nat. Struct. Mol. Biol.* 2017;24(7):570. [PubMed: 28581512]
35. Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, et al. Structure of the human histamine H1 receptor complex with doxepin. *Nature.* 2011;475(7354):65. [PubMed: 21697825]
36. Thal DM, Sun B, Feng D, Nawaratne V, Leach K, Felder CC, et al. Crystal structures of the M1 and M4 muscarinic acetylcholine receptors. *Nature.* 2016;531(7594):335. [PubMed: 26958838]
37. Thompson AA, Liu W, Chun E, Katritch V, Wu H, Vardy E, et al. Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature.* 2012;485(7398):395. [PubMed: 22596163]
38. Wu H, Wacker D, Mileni M, Katritch V, Han GW, Vardy E, et al. Structure of the human κ -opioid receptor in complex with JDTic. *Nature.* 2012;485(7398):327. [PubMed: 22437504]
39. Zhang C, Srinivasan Y, Arlow DH, Fung JJ, Palmer D, Zheng Y, et al. High-resolution crystal structure of human protease-activated receptor 1. *Nature.* 2012;492(7429):387. [PubMed: 23222541]
40. Zhang J, Zhang K, Gao Z, Paoletta S, Zhang D, Han GW, et al. Agonist-bound structure of the human P2Y12 receptor. *Nature.* 2014;509(7498):119. [PubMed: 24784220]

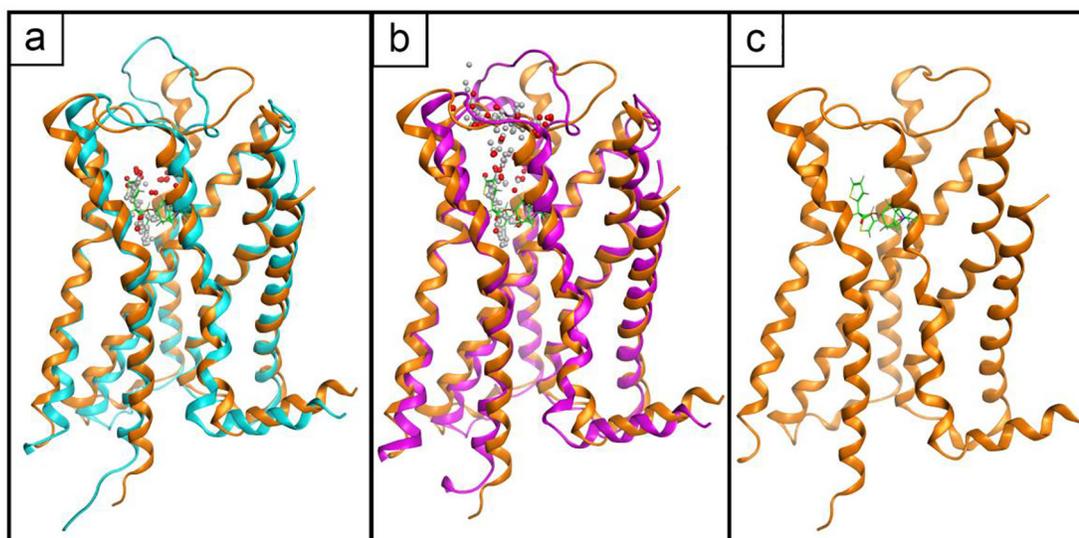
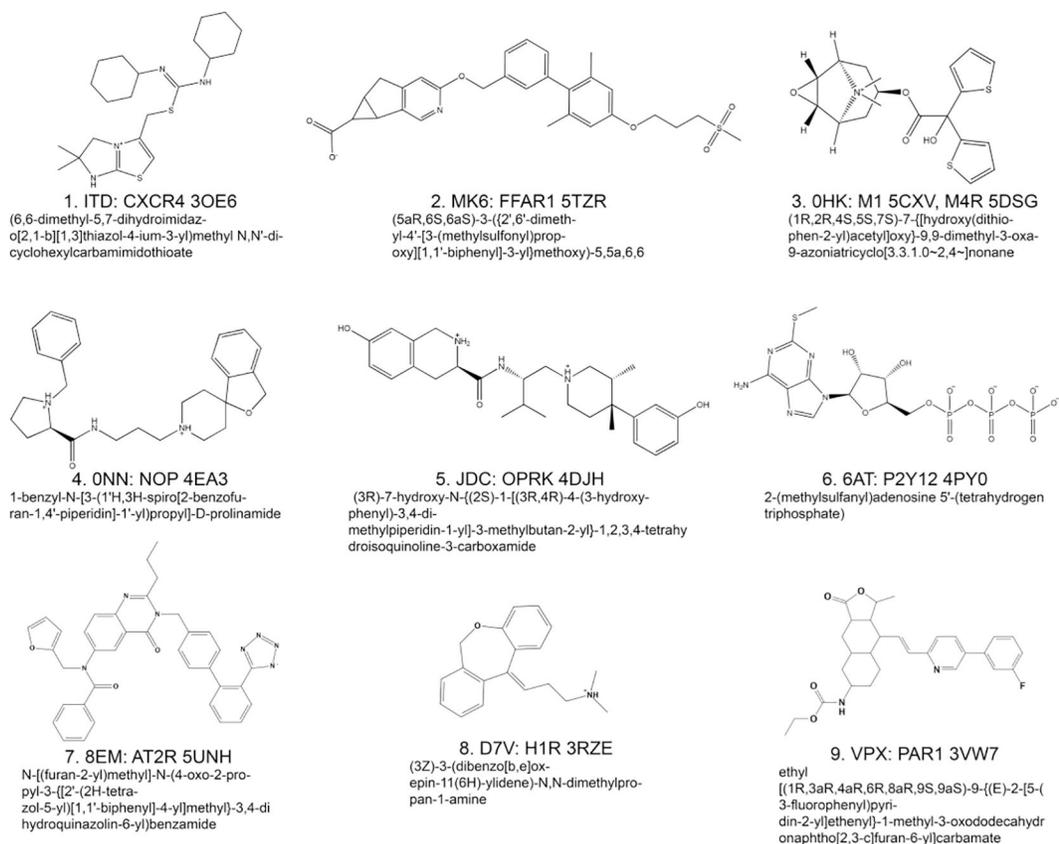


Fig. 1. Comparison of binding sites (spheres) identified in M1 homology models constructed from PDB 3RZE selected using A) Contact energy (cyan ribbon) and B) GBVI score (magenta ribbon) using the MOE site finder application with each respective structure superposed onto the crystal structure (PDB 5CXV orange) with ligand (OHK) visible (green). C) Crystal structure (PDB 5CXV orange) with ligand **3** (OHK green, see Figure 2 for structure).

**Fig. 2.**

Names and structures of ligands docked into protein models. The first line of text represents an abbreviated description of the ligand, name of the receptor, and PDB entry code in the PDB. The subsequent lines provide the IUPAC name of the ligand.

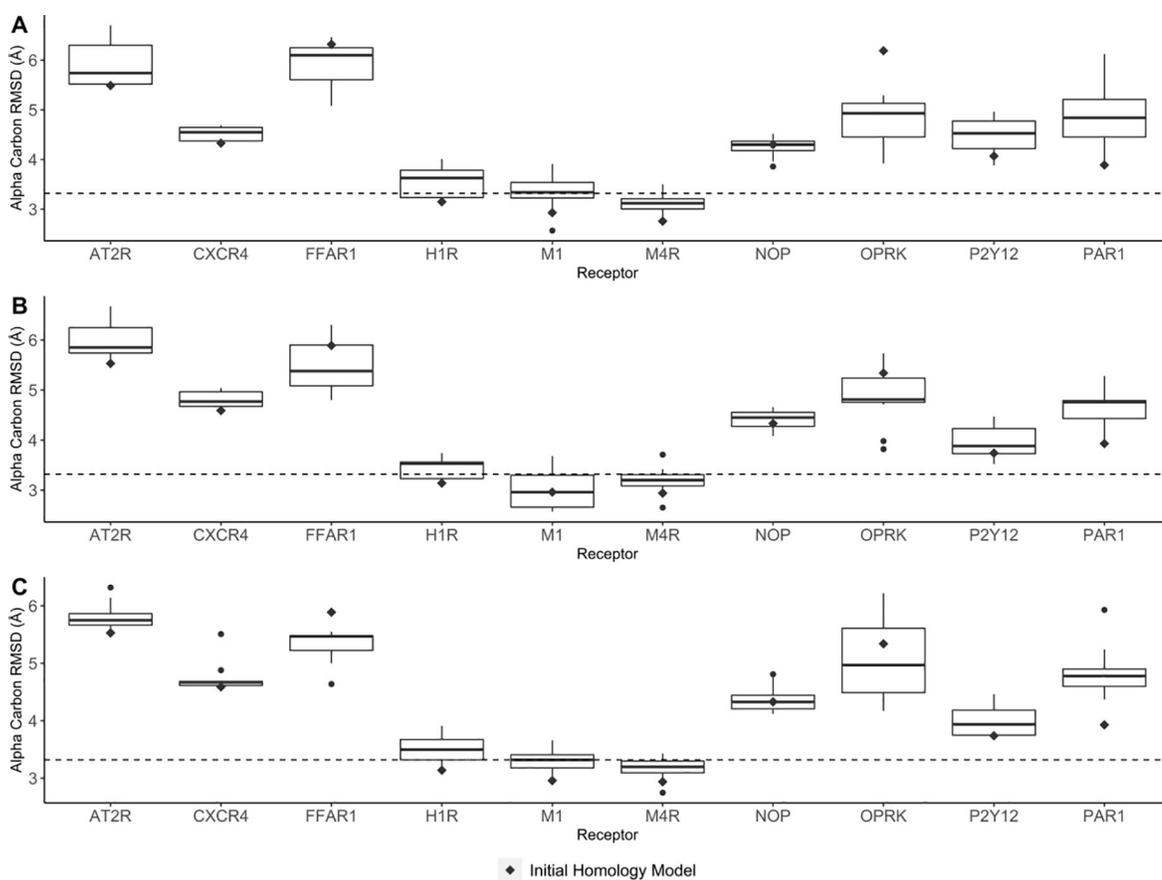
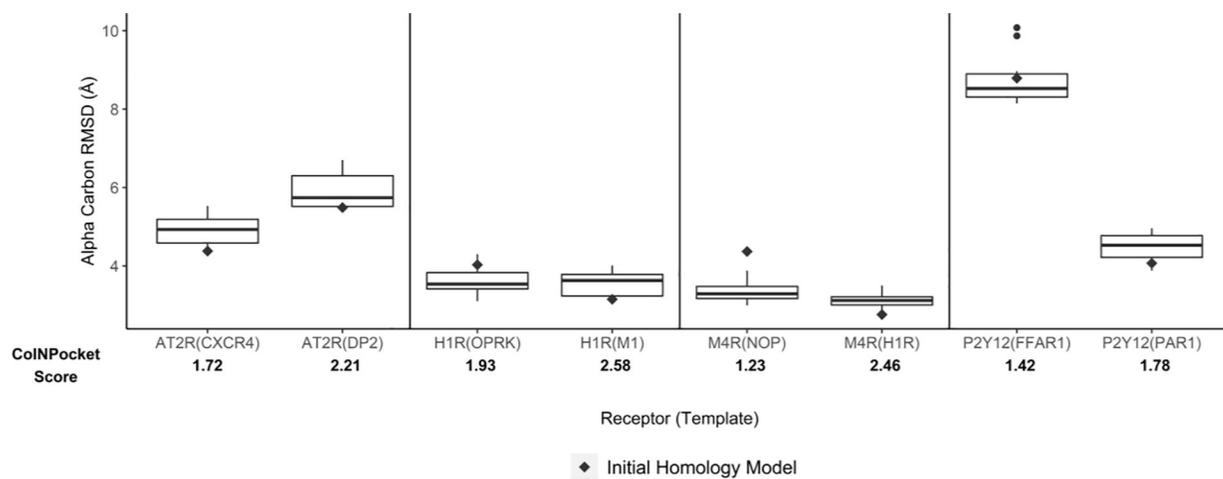


Fig. 3. Alpha carbon receptor RMSD values for the homology models generated with and without loop modeling for three different modeling approaches: (A) Approach A1, (B) Approach B1, (C) Approach B2. The dashed line appearing in each plot represents our receptor model quality metric of 3.5 Å.

**Fig. 4.**

Alpha carbon receptor RMSD values relative to crystallographic reference structures for receptor models generated by approach A1 with and without loop modeling for receptors modeled using two templates of varying local similarity score.

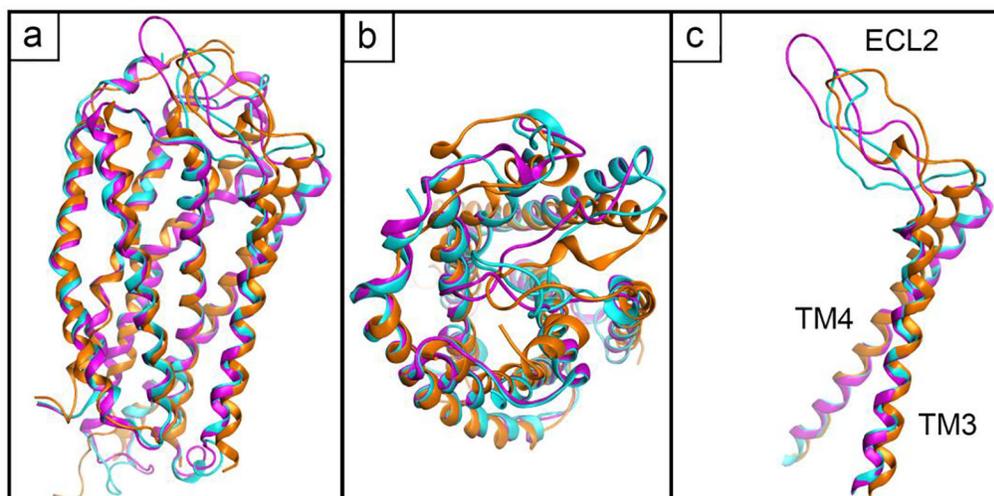


Fig. 5. Lowest RMSD P2Y12 homology model constructed from PDB 3VW7 superposed on reference crystallographic structure (PDB 4PY0). A) View from within membrane plane of P2Y12 local template homology model (magenta) and lowest RMSD loop modeled local template homology model (cyan) superposed over the crystallized reference structure (orange). B) Extracellular view of the same superposition. C) Ribbons for TM4-ECL2-TM5 segments only shown from same viewpoint used in panel A.

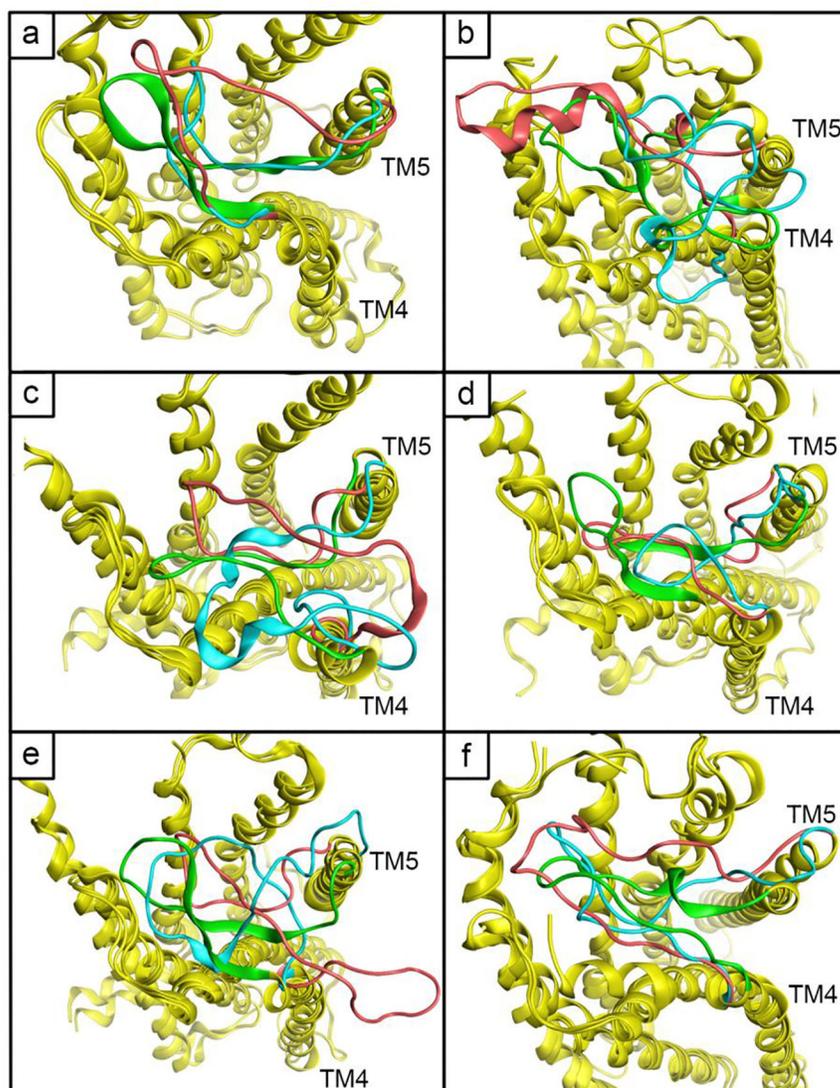


Fig. 6. The lowest RMSD model of the top 10 scoring ECL2 models (cyan) and local template homology model (salmon) was superposed onto the crystallized reference structure (green). Loop RMSD values can be found in Tables 5 and 6. (A) CXCR4 based on PDB 5UNH, (B) FFAR1 based on PDB 4PY0, (C) M1 based on PDB 3RZE, (D) NOP based on PDB 5DSG, (E) OPRK based on PDB 3RZE, (F) P2Y12 based on PDB 3VW7

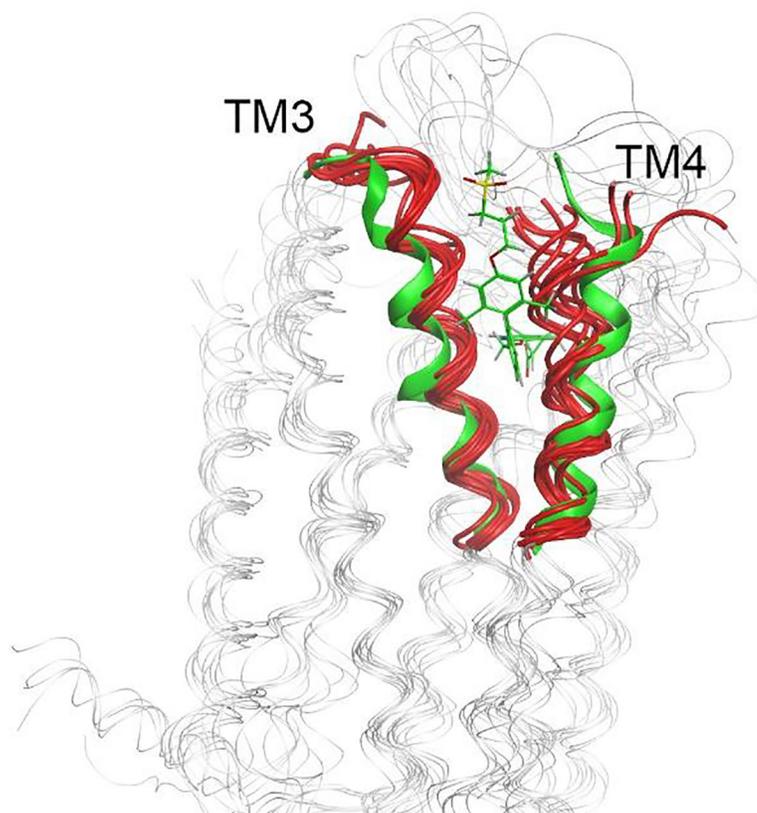


Fig. 7. All atom superposition of crystal structures used in this study with segments of TM3 and TM4 highlighted (green: FFAR1, red: all other receptors) to showcase the unusual binding mode of FFAR1. The bound conformation of ligand MK6 within FFAR1 is highlighted in green as well

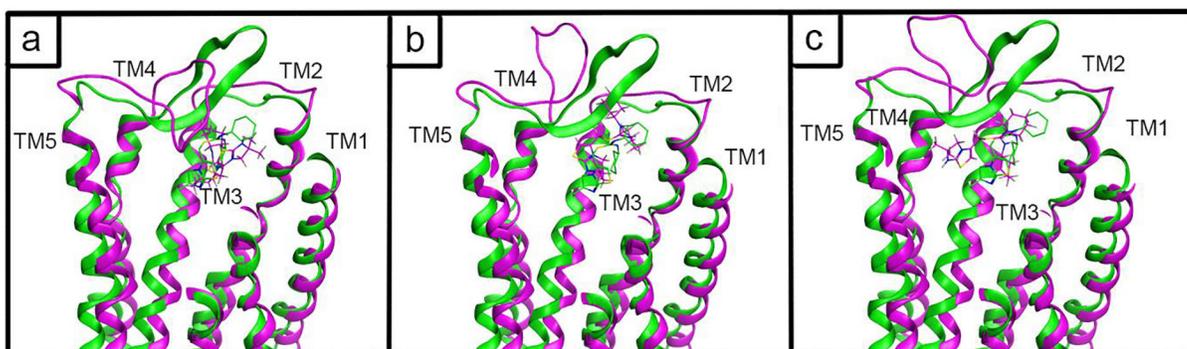
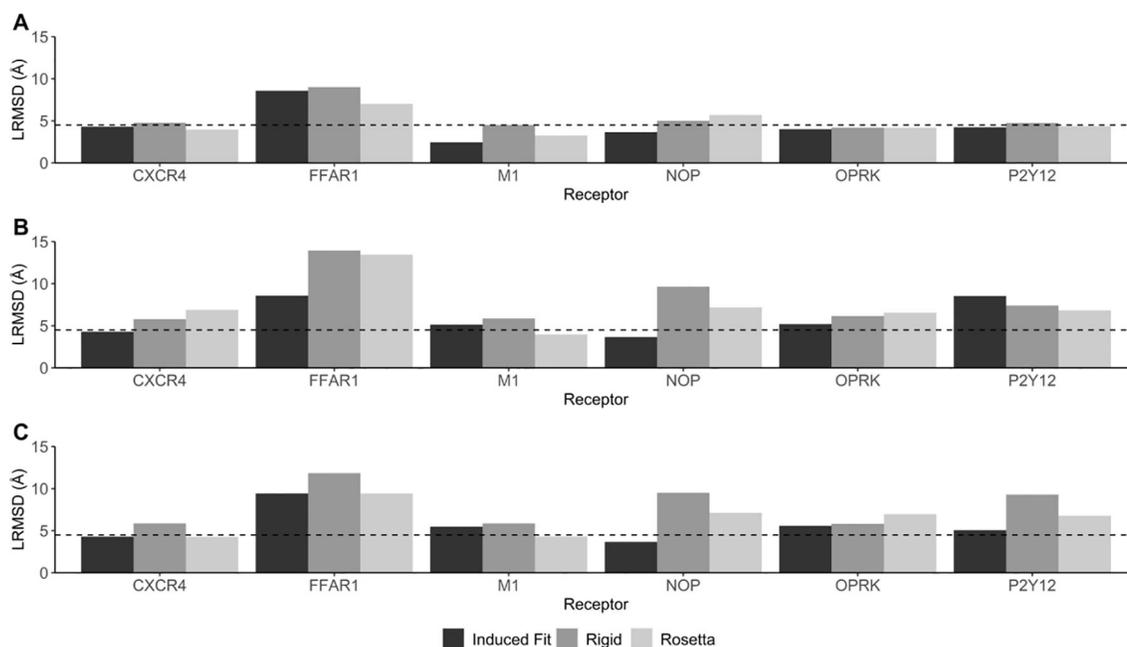
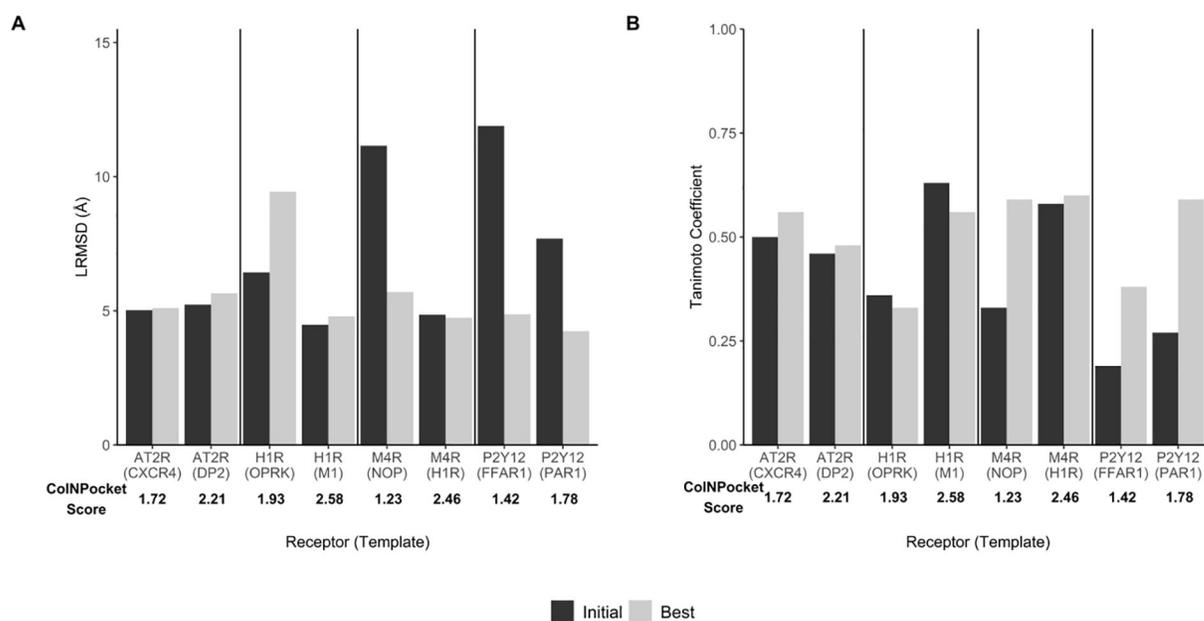


Fig. 8.

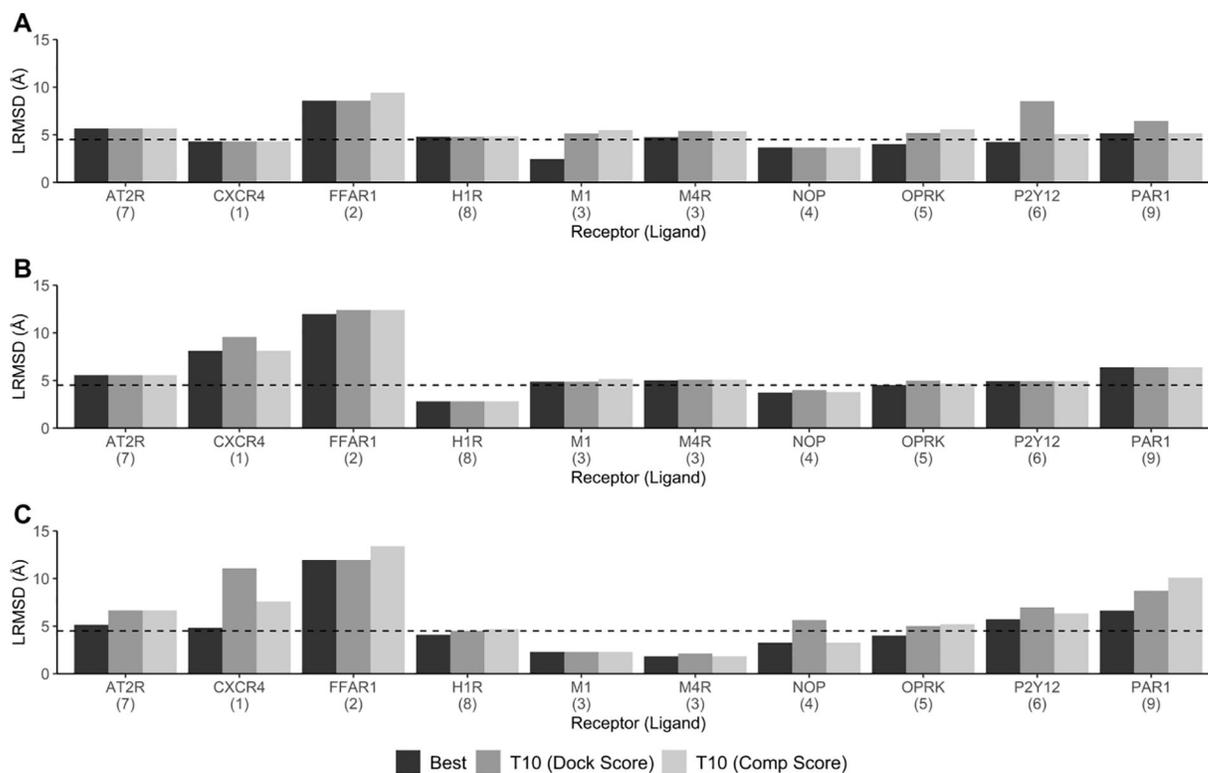
An example of CXCR4 ligand 1 docked using three different methods with the lowest LRMSD pose shown. Ligand superpositions of poses docked into CXCR4 models based on PDB 5UNH (magenta) and crystallographic reference (PDB 3OE6, green) are shown for three docking methods: MOE induced fit (panel A), MOE rigid receptor (panel B), and Rosetta (panel C)

**Fig. 9.**

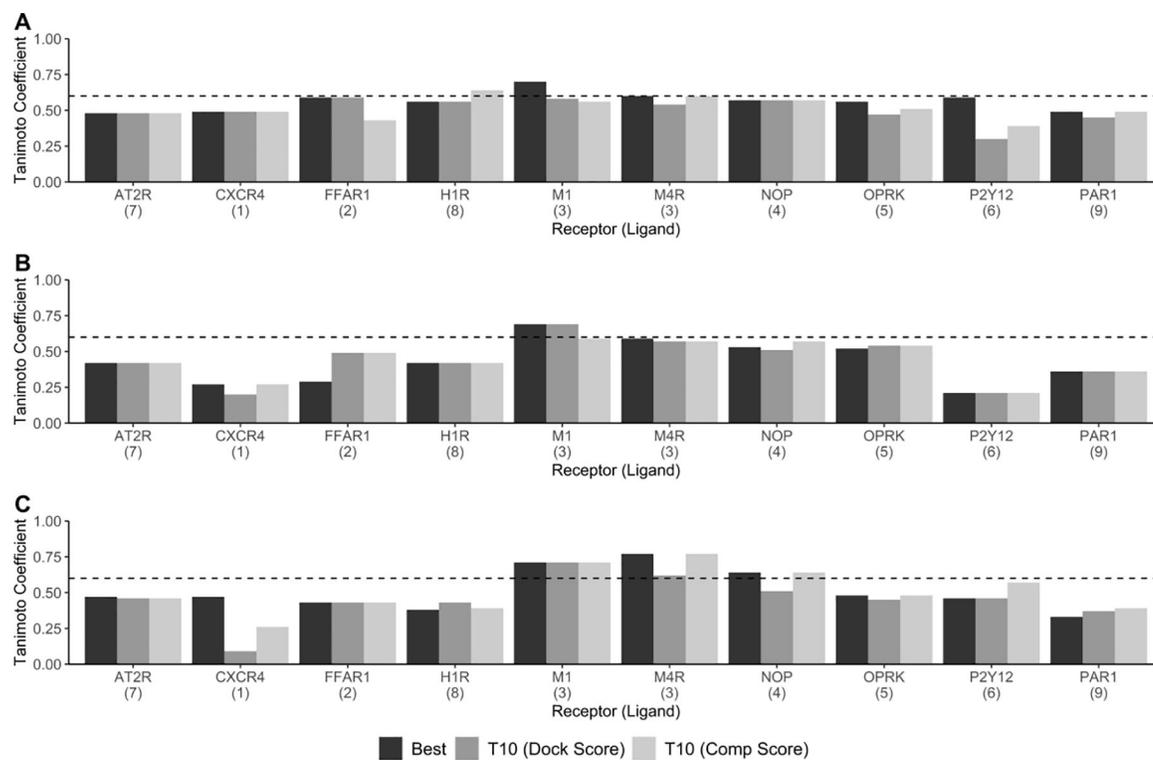
Ligand RMSD (LRMSD) values calculated in comparison to the crystallized reference structure for three different docking methods employed in the context of approach A1 models. The dashed line appearing in each plot represents our docking performance target of 4.5 Å. (A) Lowest RMSD value found within the retained ligand poses for each method. All methods sampled 10,000 ligand poses per receptor (1000 per model). Both MOE Induced Fit and MOE Rigid retained 50 ligand poses per receptor (5 per model) and Rosetta retained all ligand poses. (B) Lowest LRMSD value within the top 10 scoring ligand poses. (C) Lowest LRMSD value within the top 10 poses based on adjusted percent complementation score.

**Fig. 10.**

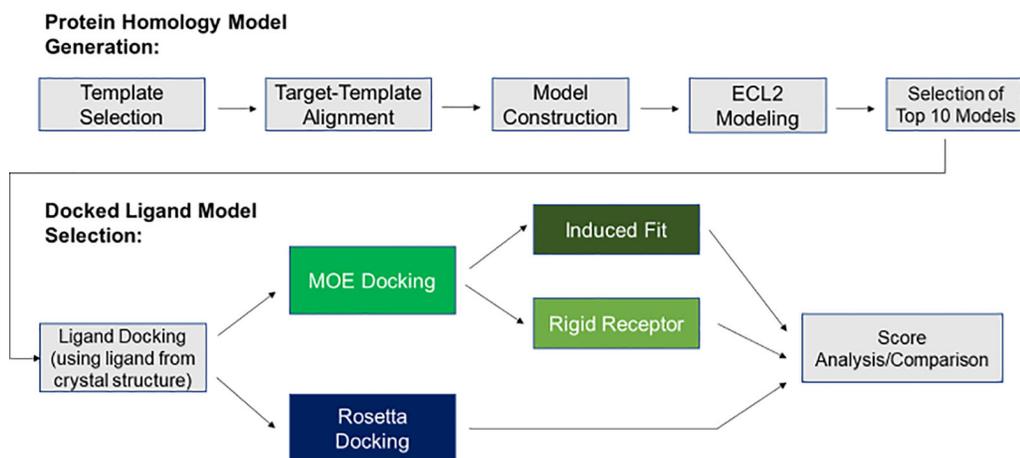
Comparison of MOE induced fit docking poses and crystallographic ligand poses for receptors modeled using two templates when docked into approach A1 receptor models. (A) Lowest LRMSD docked pose obtained from docking into the top 10 scoring loop refined homology models (black) and best T10 comp LRMSD (grey) for each receptor. (B) Calculated Tanimoto coefficients corresponding to the aforementioned docked poses.

**Fig. 11.**

Ligand RMSD values for ligand poses docked into receptor models generated using three different receptor modeling approaches. The dashed line appearing in each plot represents our pose quality metric of 4.5 Å. Ligands docked into each receptor are represented below each receptor name and follow the numbering scheme found in Figure 2. (A) Approach A1, (B) Approach B1, (C) Approach B2.

**Fig. 12.**

Tanimoto coefficients for ligand poses docked into receptor models generated using three different receptor modeling approaches. The dashed line appearing in each plot represents our Tanimoto coefficient target of 0.6. Ligands docked into each receptor are represented below each receptor name and follow the numbering scheme found in Figure 2. (A) Approach A1, (B) Approach B1, (C) Approach B2



Scheme 1.
Homology modeling/loop modeling protocol.

Table 1.

GenBank accession numbers and PDB ID numbers for GPCR used in this study.

Receptor	Local Template	Local Similarity Measure ^a	Unweighted Global Similarity (%)	GenBank Accession Number	Target PDB ID	Template PDB ID
AT2R	CXCR4	1.72	31.43	P50052.1	5UNH [31]	3OE6
AT2R	DP2	2.21	33.99	P50052.1	5UNH	6D26 [32]
CXCR4	AT2R	1.72	31.43	CAA12166.1	3OE6 [33]	5UNH
FFAR1	P2Y12	1.42	7.09	AAI20945.1	5TZR [34]	4PY0
H1R	OPRK	1.93	20.31	P35367.1	3RZE [35]	4DJH
H1R	M1	2.58	35.98	P35367.1	3RZE	5CXV [36]
M1	H1R	2.58	35.98	CAA68560.1	5CXV	3RZE
M4R	NOP	1.23	14.41	P08173.2	5DSG [36]	4EA3
M4R	H1R	2.46	34.37	P08173.2	5DSG	3RZE
NOP	M4R	1.23	14.41	NP_872588.1	4EA3 [37]	5DSG
OPRK	H1R	1.93	20.31	AAC50158.1	4DJH [38]	3RZE
PAR1	P2Y12	1.78	16.26	N/A	3VW7 [39]	4PY0
P2Y12	FFAR1	1.42	16.26	Q9H244.1	4PY0 [40]	5TZR
P2Y12	PAR1	1.78	7.09	Q9H244.1	4PY0	3VW7

^aCompared to the maximal self-similarity measure of 5.47. A pairing of two receptors with a local similarity score of 5 would indicate a near 100% ligand binding pocket similarity, while a receptor pairing with a local similarity score of 1 or less would indicate low ligand binding pocket similarity.

Table 2.

Variation among experimental structures for each receptor used as a template and/or target in this study.

Receptor	Entries in PDB	Variation (Å) ^a
AT2R	5UNH, 5UNG, 5UNF, 6DO1, 5XJM	1.78
CXCR4	3ODU, 3OE0, 3OE6, 3OE9, 4RWS	1.35
DP2	6D26, 6D27	0.49
FFAR1	4PHU, 5TZR, 5TZY	1.34
H1R	3RZE	N/A
M1	5CXV, 6OIJ	2.35
M4R	5DSG	N/A
NOP	4EA3, 5DHG, 5DHH	0.70
OPRK	4DJH, 6B73	3.29
P2Y12	4NTJ, 4PXZ, 4PY0	2.67
PAR1	3VW7	N/A
Average		1.75

^aHighest alpha carbon RMSD between any two structures for each receptor. RMSD was calculated using an alpha carbon superposition between residues present within all PDB entries for each receptor. Receptors with a value of "N/A" had only one crystal structure available at the time of this research.

Table 3.

ECL2 loop start/end residues for each receptor's crystal structure and lowest RMSD homology model.

Receptor	Crystal Structure			Homology Model		Length Difference ^a	Avg. Anchor Residue RMSD (Å) ^b
	Local Template	ECL2 Start Residue	ECL2 End Residue	Anchor Residue 1	Anchor Residue 2		
AT2R	CXCR4	F181	P201	F181	Q206	+5	1.69
AT2R	DP2	F181	P201	F181	E202	-1	3.89
CXCR4	AT2R	F174	L194	F174	D193	-1	2.86
FFAR1	P2Y12	E145	P176	F142	P176	+3	6.65
H1R	OPRK	L163	T188	L163	V187	-1	3.15
M1	H1R	V168	P186	W164	I187	+12	2.97
M4R	NOP	V175	P193	F170	N192	+4	2.24
M4R	H1R	V175	P193	I168	P193	+7	2.42
NOP	M4R	M188	Q208	Q192	V214	+2	9.67
OPRK	H1R	L196	Y219	L192	D223	-3	2.74
P2Y12	FFAR1	I161	E181	T163	I193	+10	5.70
P2Y12	PAR1	I161	E181	M160	V185	+5	4.72
PAR1	PAR1	L238	G265	K240	E264	-3	5.02

Structures of both the crystal structure and lowest alpha carbon RMSD homology model for each were aligned and superposed in MOE, then renumbered from 1 starting at the beginning of TM1.

^aDifference in sequence length between segment bookended by anchor points of the homology model and ECL2 of crystal structure. For example, the segment loop modeled for the CXCR4 homology model had one more residue than the actual ECL2 of the crystal structure of CXCR4, etc.

^bRMSD of loop anchor residue positions in the lowest RMSD loop model from the corresponding residues in the crystal structure once superposed.