JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

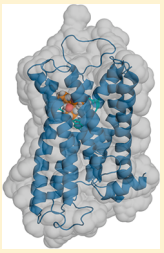
Homology Modeling of Human Muscarinic Acetylcholine Receptors

Trayder Thomas,[†] Kimberley C. McLean,[†] Fiona M. McRobb,[§] David T. Manallack, David K. Chalmers,* and Elizabeth Yuriev*

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

Supporting Information

ABSTRACT: We have developed homology models of the acetylcholine muscarinic receptors M_1R-M_5R, based on the β_2-adrenergic receptor crystal as the template. This is the first report of homology modeling of all five subtypes of acetylcholine muscarinic receptors with binding sites optimized for ligand binding. The models were evaluated for their ability to discriminate between muscarinic antagonists and decoy compounds using virtual screening using enrichment factors, area under the ROC curve (AUC), and an early enrichment measure, LogAUC. The models produce rational binding modes of docked ligands as well as good enrichment capacity when tested against property-matched decoy libraries, which demonstrates their unbiased predictive ability. To test the relative effects of homology model template selection and the binding site optimization procedure, we generated and evaluated a naïve M_2R model, using the M_3R crystal structure as a template. Our results confirm previous findings that binding site optimization using ligand(s) active at a particular receptor, i.e. including functional knowledge into the model building process, has a more pronounced effect on model quality than target–template sequence similarity. The optimized M_1R-M_5R homology models are made available as part of the Supporting Information to allow researchers to use these structures, compare them to their own results, and thus advance the development of better modeling approaches.



INTRODUCTION

The use of structure-based design methods for G protein-coupled receptors (GPCRs) is an active area of research.[1−4] It commenced in the early 2000s after the landmark report of the crystal structure of bovine rhodopsin[5] and accelerated after 2007, when the first crystal structures of ligand-infusible GPCR complexes were solved.[6−8] Technological advances have greatly improved the success of GPCR crystallization[7,9,10] and, at the time of writing, over 30 crystal structures of GPCRs have been solved.[11] However, GPCR crystallization is still an area of highly specialized expertise with most structures coming from a limited number of research groups. As a result, the number of available structures is still very small given the ∼800 GPCRs present in the human genome, including 342 nonolfactory receptors.[12] Many GPCR families are still not covered by the currently available high resolution structural information, and it is accepted that, at present, solving structures for all members of the GPCR superfamily is not a realistic goal.[1,4] Consequently, in the absence of experimental structural data, researchers who wish to use structure-based methods to target GPCRs turn to homology models for docking and virtual screening (VS).[13] In several of these drug discovery campaigns, GPCR homology models have proven useful for discovering agents for a range of GPCR targets (Table 1).

While generally an established technique, generation of GPCR homology models for virtual screening can be a speculative exercise, relying on many assumptions and suppositions. Therefore, careful consideration of several related aspects is required when such an exercise is undertaken. (i) Robustness of the computational protocol. This aspect comprises quality of both homology modeling and docking algorithms and should always be evaluated against relevant targets for which experimental data is available: structural data for validating homology modeling and activity data for validating VS. The ultimate question that must be answered is whether the combination of the evaluative model and the protocol used can distinguish between known actives and drug-like decoy molecules. (ii) Quality and appropriateness of the input structural data; specifically the choice of template. Choosing a template for GPCR homology modeling has been previously evaluated;[14−17] however, with the ever-increasing number of available templates, this question cannot be resolved once and for all and requires regular re-evaluation. (iii) Predictive quality of the generated homology models. To address this final issue, homology models should be evaluated in a virtual screening scenario with a particular focus on decoy selection. Because of the importance of these issues, there is currently a considerable interest in evaluating homology modeling and VS protocols as applied to GPCRs.[18−20]

In this study, we have addressed all of the above issues by modeling the five subtypes of muscarinic acetylcholine receptors (mAChRs) and evaluating them using virtual screening. The mAChRs receptors (M_1R-M_5R) can be subdivided into two functional classes based on their G protein coupling preference.[21] The M_1R, M_3R, and M_5R selectively couple to G proteins of the G_q/G_{11} family while the M_2R and M_4R preferentially activate G_i/G_o-type G proteins. Activation of mAChRs leads to a wide range of biochemical and physiological effects, primarily depending on the mAChR location and

Received: September 2, 2013
Published: December 15, 2013

Table 1. Prospective Virtual Screening Campaigns against GPCR Homology Models

target[a]	template[a]	homology modeling program	docking/ screening program	screening library	hit rate[b]	affinity (number of compounds)[c]	ref
D_3R	β_2AR, β_1AR	MODELLER	DOCK3.6	prefiltered ZINC[28] (3000K+)	23% (20%)	K_i = 0.2 − 3.1 μM (6) optimized K_i = 81 nM (1)	29
CXCR4	rhodopsin, β_2AR, β_1AR, $A_{2A}R$	MODELLER	DOCK3.6	lead-like subset of ZINC (3300K)	4% (17%)	IC_{50} = 107 μM (1)	30
CXCR7	rhodopsin, β_2AR, β_1AR, $A_{2A}R$, CXCR4	MOE	CONSENSUS-DOCK	3 proprietary collections (187K, 402K, 196K)	3.3%	IC_{50} = 1.29−11.4 μM (21)	31
S1PR1[d]	rhodopsin	GPCRgen	Snooker	diverse subset of MSD/ Organon library (50K)	NR	pK_i = 4.3−4.7 (3)	32, 33
$A_{2A}R$	β_1AR	MODELLER, MOE	Glide	CAP, BioFocus SoftFocus (545K)	9%	pK_i = 7.5−9.0 (6) (13 to >100-fold selective vs A_1R)	34
5-HT_7R[d]	rhodopsin	MODELLER	Glide	Enamine Screening Collection (730K)	NR	K_i = 0.197 and 0.265 μM (2)	35, 36
5-HT_2R	β_2AR	MODELLER	DOCK3.5, MM-GBSA	FDA drug library, filtered by MW (1430)	NR	K_i = 1.959 mM (1)	37
MCH-1R	β_2AR	MOE	GOLD	commercial vendor catalogues (45K)	14%	IC_{50} = 131 and 213 nM (2 most potent out of 10 novel chemotypes)	38
CB_2R	β_2AR	CHARMM for "activation"	GOLD	filtered subset of ZINC (273K)	12%	K_i = 2.3 nM−71.43 μM (13)	39

[a]Receptor abbreviations: adenosine A_X receptor, A_XR; β_X-adrenergic receptor, β_XAR; cannabinoid receptor 2, CB_2R; C−X−C chemokine receptor 4, CXCR4; dopamine D_3 receptor, D_3R; melanin-concentrating hormone-1 receptor, MCH-1R; serotonin 5-HT_X receptor, 5-HT_XR; sphingosine 1-phosphate receptor, S1PR1. [b]Hit rates are estimated differently in various studies. Where available, we quote hit rates for VS against crystal structures for comparison, in parentheses. NR = not reported. [c]Reported as per original papers. [d]A combination of ligand-based and structure-based approaches were used in this campaign.

subtype. The M_1R, M_4R, and M_5R subtypes are mainly expressed in the central nervous system (CNS); whereas, the M_2R and M_3R subtypes are widely distributed both in the CNS and in peripheral tissues. Specifically, we have generated homology models of mAChRs M_1−M_5, using the β_2-adrenergic receptor (β_2AR) crystal structure (PDB ID: 2RH1)[6] as the template and have optimized their orthosteric binding sites using the induced fit docking (IFD) procedure.[22] (i) We have demonstrated the robustness of our homology modeling/VS protocol using the β_2AR crystal structure in complex with the inverse agonist carazolol and β_2AR antagonist and inverse agonist activity data. We have further verified the protocol by a validation against the β_2AR crystal structure in complex with alprenolol.[23] (ii) To assess the predictive quality of the M_1R−M_5R models, we have carried out virtual screening investigations of all five homology models. The models have been tested against property-matched decoy libraries to demonstrate their unbiased predictive capacity. (iii) Furthermore, after the M_2R (human)[24] and M_3R (rat)[25] crystal structures became available, a naïve (i.e., nonoptimized) M_2R model was generated using the M_3R crystal structure as a template. Evaluating VS performance allowed comparison between the models: naïve but based on a close-sequence template and optimized but based on a more remote-sequence template. Our results support previous findings that binding site optimization using ligand(s) active at a particular receptor, i.e. including functional knowledge into the model building process,[26] has a pronounced effect on model quality for virtual screening. It is clear from our results that carefully designed and knowledge-based homology structures, built with templates with greater than 35% overall similarity in the trans-membrane region,[1] are at least as useful in VS as crystal structures. Finally, similar to our previous work,[27] we have released the coordinates of the five optimized muscarinic receptor structures in the spirit of open science research.

EXPERIMENTAL SECTION

Software. Molecular modeling was performed with the Schrödinger software suite.[40,41] Homology models of the five muscarinic M_1−M_5 acetylcholine receptors were built in Prime[42] (v 3.0 and 3.1) from a multiple sequence alignment generated in ClustalW[43] using the Maestro interface (v 9.2 and 9.3). Ligand molecules were prepared using LigPrep[44] (v 2.5), and the binding site was optimized using the IFD protocol[22] following the previously developed procedure.[27] Ligands were docked into the homology models using Glide[45,46] (v 5.7 and 5.8). Default settings were used, unless otherwise stated. Physical descriptors evaluated for comparison of the decoy sets with the active compounds included molecular weight (MW), number of rotatable bonds, number of hydrogen bond donor and acceptor atoms, and calculated logP (ClogP). These physical properties, along with polar surface area (PSA) and vdW volume, were computed using the ChemAxon Marvin Calculator (cxcalc) (http://www.chemaxon.com). The 2D Tanimoto score (calculated using fragment sizes of 1−7 atoms, ignoring hydrogens) was measured to demonstrate the diversity of the structures within the ligand sets.[47] The workflow followed in this study is shown in Figure 1 and described in detail in the following sections.

Homology Modeling. The sequences of the human dopamine, serotonin, α- and β-adrenergic, adenosine, histamine, muscarinic, and bovine rhodopsin receptors were obtained from the Universal Protein Resource (http://www.uniprot.org/) and aligned using ClustalW. The multiple sequence alignment generated was manually edited to remove gaps in helices and to anchor highly conserved residues in each transmembrane (TM) helix. Naïve homology models for the five human mAChRs were built in Prime v 3.0 from the multiple sequence alignment, using the β_2-adrenergic receptor (PDB ID: 2RH1) crystal structure[6] as the template. The human muscarinic M_2 acetylcholine receptor was also built in Prime v 3.1, using the rat muscarinic M_3 acetylcholine receptor (PDB

dx.doi.org/10.1021/ci400502u | J. Chem. Inf. Model. 2014, 54, 243−253

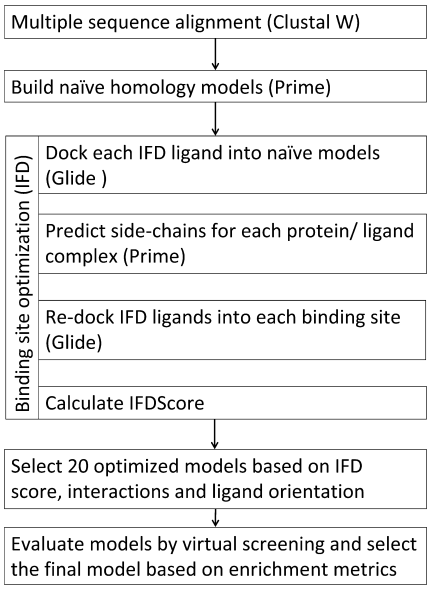


Figure 1. Flowchart of homology modeling and model evaluation.

ID: 4DAJ) crystal structure[25] as the template. Further details are described in our previous work.[27]

Binding Site Optimization. The β_2AR crystal structure (PDB ID: 2RH1) and the mAChRs homology models were treated by the Protein Preparation Wizard workflow,[44] prior to docking. Hydrogen atoms were added and minimized using the OPLS_2005 force field. The side chain conformations of the residues within the ligand binding site were refined by docking an appropriate antagonist or inverse agonist into each of the built muscarinic receptor homology models and the β_2AR crystal structure using the IFD protocol. The docking site was centered upon the residues Asp 3.32, Trp 6.48, Phe 6.52, and Tyr 7.43 (Ballesteros–Weinstein numbering[48]) and was defined by a box of dimensions 28 × Å 28 Å × 28 Å. Up to 50 poses per ligand were collected in the initial Glide docking step, with both the van der Waals (vdW) radii and the partial atomic charges scaled to 0.5 in order to collect a more extensive range of poses.

Prime was used to optimize residues within 5 Å of ligand atoms, excluding Asp 3.32 and Trp 6.48, which play a critical role in correctly orienting ligand molecules. Trp 6.48 is a key residue of the aromatic cluster of TM5 and TM6, believed to act as a "micro-switch", important for receptor activation and inactivation.[49] The IFD protocol was found to consistently cause Trp 6.48 to undergo a conformational "flip" during the Prime step, forcing the bulky indole side chain down and away from the binding pocket. For M_1R–M_5R models, when Trp 6.48 and Asp 3.32 were omitted from binding site optimization, more credible ligand poses were obtained, which led to better enrichment. Pala et al.[16] report a similar observation for VS-evaluated homology models of the MT_2 melatonin receptor, namely that residues known to form critical ligand contacts tended to adopt a conformation not favorable to forming such contacts. They have taken this observation as another reason for "calibrating" models (e.g., by VS evaluation) to determine the domain of their applicability.

Following optimization with Prime, the ligand was redocked into the optimized receptor conformations with Glide, using default vdW and charge scaling parameters. Multiple ligand–receptor poses were generated for each model. Successful poses were chosen on the basis of the position and orientation of the ligand within the binding pocket, key hydrogen bonding and vdW interactions, and the relative energy of interaction (a composite of the protein and ligand energy scores: IFDScore = GlideScore + 0.05 × PrimeEnergy). A maximum of 20 poses were collected. During the IFD optimization of the binding sites, we monitored the distance (ndist) between the ionizable or quaternary nitrogen of the ligand (for simplicity we will just refer to this atom as the "ionizable nitrogen") and the closest carboxylate oxygen of the conserved Asp 3.32 residue. This residue has been determined by site-directed mutagenesis to be crucial in the ligand-binding mode of all aminergic GPCRs.[50] The term ndist is a quantitative measure of this important ionic interaction, and receptors with ndist > 3.0 Å were excluded from further analysis.

Virtual Screening Libraries. Active compounds known to act at the β_2-adrenergic and muscarinic receptors were used to enrich the decoy compound databases (20 actives for β_2AR and 48 actives for mAChRs; see Table S1 for the lists of actives and ref 27 for chemical structures). The active compounds were downloaded from the GLIDA database[51] (http://pharminfo.pharm.kyoto-u.ac.jp/services/glida/). Protonation states and formal charges at physiological pH (pH 7.4 ± 2.0) for each active ligand and decoy compound were assigned in LigPrep. One structure per compound was selected for screening.

Three sets of decoy compounds were used in this study. Set 1, containing 1000 drug-like decoy compounds, was obtained from Schrödinger (http://www.schrodinger.com). This set had been randomly selected from a library of one million compounds having properties characteristic of drug molecules.[45,46] We have analyzed the properties of the decoy ligands and active compounds: molecular weight (g/mol), number of rotatable bonds, polar surface area (Å^2), calculated logP, number of hydrogen bond donors and acceptors, solvent accessible volume (Å^3), and 2D Tanimoto score. Generally, the properties of the active compounds were found to be similar to those of the decoy library (Table 2). The molecular weights varied from 151 to 645 g/mol, with an average of 360 g/mol. These decoys were not specifically chosen to mimic muscarinic antagonist compounds, as we first wanted to ascertain whether our models were capable of identifying active ligands from within a broad representation of drug-like compound space.

Decoy set 2 was derived from the ZINC database[28] (7 233 297 compounds, database version 7) by a process of successive

Table 2. Average Ligand Properties

ligand set	MW (g/mol)	rotatable bonds	PSA (Å^2)	ClogP	H-bond donor	H-bond acceptor	vdW volume (Å^3)	2D Tanimoto score
M_1R actives	324	5.1	31	3.03	1.4	1.6	318	0.233
				decoy sets				
1: Schrödinger	360	5.0	84	2.90	2.0	4.2	316	0.125
2: ZINC	320	4.3	38	3.43	1.4	1.7	302	0.185
3: refined Schrödinger	343	4.8	79	2.59	2.4	3.3	312	0.143

dx.doi.org/10.1021/ci400502u | J. Chem. Inf. Model. 2014, 54, 243–253

eliminations, creating a subset of molecules that closely adhered to the physical properties of the actives (Table 2). Specifically, decoys were required to fall within a similar normal distribution as the active compounds (265−434 g/mol; mean 322 g/mol; standard deviation 40 g/mol). Decoys were also required to contain an ionizable nitrogen and not to contain more than three hydrogen bond donors or four hydrogen bond acceptors. Finally, each decoy was required to have a Tanimoto score of less than 0.8 with respect to all other molecules within the set to ensure topological diversity. 1000 molecules were randomly selected from a larger subset satisfying the applied criteria, so that direct comparisons could be made between the screening results using the Schrödinger and ZINC libraries, in terms of enrichment factors and early hits. A carefully selected set of 1000 molecules seems to be sufficient to detect enrichment trends. Huang et al. found that there was little size-dependent behavior detected when screening with their entire Directory of Useful Decoys (DUD)[52] of 98 266 molecules compared to a randomly selected subset of 1000 molecules.

Decoy set 3 (refined Schrödinger) was a subset of the decoy set 1, with molecular weight limited to be consistent with that of the active compounds (260−410 g/mol). All compounds from the Schrödinger decoy library with a molecular weight which fell outside the range of the active compounds were removed. Furthermore, all decoy compounds which did not contain an ionizable nitrogen were similarly removed to create a more challenging decoy set of 261 compounds.

Enrichment Studies. Molecular docking studies were performed using Glide, which flexibly docks ligands into a rigid receptor model. The docking site was centered upon the coordinates of carazolol (the inverse agonist present in the β_2AR crystal structure) and limited to accommodate ligands up to 18 Å in length. The midpoint of each ligand was bound to an inner box of 10 Å^3. Postdocking minimization retained a single pose per ligand. Both the Standard Precision (SP) and the Extra Precision (XP) scoring functions were evaluated, and XP gave marginally better results, which are presented here. Poses were ranked using GlideScore. Following docking, models were visually inspected to ensure that the ligands were well oriented within the defined binding pocket and to ensure that important expected interactions, based on mutagenesis studies,[53] were found between ligand and receptor molecules. Enrichment factors (EF) were calculated at 2, 5, and 10% of the total number of compounds (N_{total}) screened, according to $EF^{x\%} = (Hits_{sampled}/N_{sampled}) \div (Hits_{total}/N_{total})$.

RESULTS

Method Evaluation: β_2 Adrenergic Receptor Ligand Docking. Our modeling protocol encompasses generating multiple IFD complex structures and selecting final receptor models. To evaluate the protocol, we used the β_2AR as a test case. Thirty-three structures of the β_2AR/carazolol complex were generated using IFD. To test the ability of our modeling and VS evaluation workflow to preferentially retrieve known actives, we docked 20 known β_2AR antagonists (see Table S1 in the Supporting Information) and the library of Schrödinger decoys into all 33 receptor models. Higher enrichment factors and area under the enrichment curve (AUC) values and lower average distance between the ligand ionizable nitrogen and Asp 3.32 (ndist) correlated with greater model efficiency in selecting active molecules early in the screen. The properties of the top 5 highest ranked models are shown in Table 3, and a complete list is provided in Supporting Information Table S2. A detailed comparison between β_2AR/carazolol IFD complexes and carazalol- or alprenolol-bound crystal structures is presented in the Supporting Information. This test case shows that our protocol can retrieve correct binding modes for β_2AR/carazolol complexes (i.e., consistent with crystal structures).

Table 3. Five Top Ranked Models from Virtual Screening of the β_2AR Structures Generated by IFD Using Carazolol

	enrichment factor						
ranking	2%	5%	10%	AUC	mean ndist (Å)	Carazolol RMSD (Å)	Alprenolol RMSD (Å)
1	21.8	13.7	8.9	0.96	2.09	0.75	0.65
2	21.8	13.7	9.4	0.93	2.30	1.50	1.95
3	14.5	12.7	8.9	0.95	2.51	1.21	1.62
4	12.1	11.7	9.4	0.95	2.36	4.37	1.95
5	12.1	12.7	8.9	0.96	2.79	1.63	1.09

Homology Modeling of Muscarinic Receptors. *Binding Site Optimization by IFD.* Clozapine and atropine were chosen as the optimizing ligands for IFD since they have high affinity for the M_1−M_5 receptors; reported clozapine K_i values vary from 1.4−5.0 nM and atropine K_i values range between 0.2 and 1.5 nM.[54] Following the VS procedure, described below, we found that the atropine-optimized model for the M_1R gave the best enrichment, while the best M_2R−M_5R models were optimized using clozapine.

Model Quality Evaluation by VS. We evaluated the ability of the receptor models to prioritize active compounds over decoy molecules. The decoy libraries, enriched with the respective active compounds (see Table S1 in the Supporting Information), were docked into the receptor models. The IFD ligands, used for binding site optimization, were excluded from virtual screening to remove any potential structural bias. While enrichment plots and enrichment factors (EFs) are still routinely used for evaluating VS performance (e.g., ref 18), they are not ideal and do not account for several aspects of virtual screening. ROC curves are superior to enrichment plots in that they not only reflect the selection of actives, but also the nonselection of decoys.[55,56] The metric afforded by a ROC curve is the area under the receiver operating characteristic curve (ROC AUC), which gives an indication of the total number of compounds successfully docked into the model and is interpreted as the probability that a randomly chosen active has a higher score than a randomly chosen inactive. Several metrics, such as NSQ_AUC[57] and LogAUC,[58] have also been developed to focus on early, rather than overall, enrichment.

ROC curves for the M_1R−M_5R models are shown in Figure 2. Enrichment plots and semilogarithmic ROC curves are provided in the Supporting Information (Figures S2 and S3). We also report the ROC AUC and LogAUC metrics, as enrichment measures (Table 4). The LogAUC preferentially weighs early enrichment by computing the percentage of the ideal area under the semilog ROC curve.The results reveal excellent enrichment capacity for M_2R, M_4R, and M_5R models with the latter having particularly good early enrichment. Although, the M_1R and, particularly, the M_3R gave lower enrichments using all three decoy sets, their enrichment metrics are comparable to and sometimes better than those obtained in recent reports. For example, homology models of the MT_2 melatonin receptor,[16] based on the β_2AR and optimized for antagonists gave $EF_{2\%}$ = 3.1−18.7 and of antagonists against

dx.doi.org/10.1021/ci400502u | J. Chem. Inf. Model. 2014, 54, 243−253

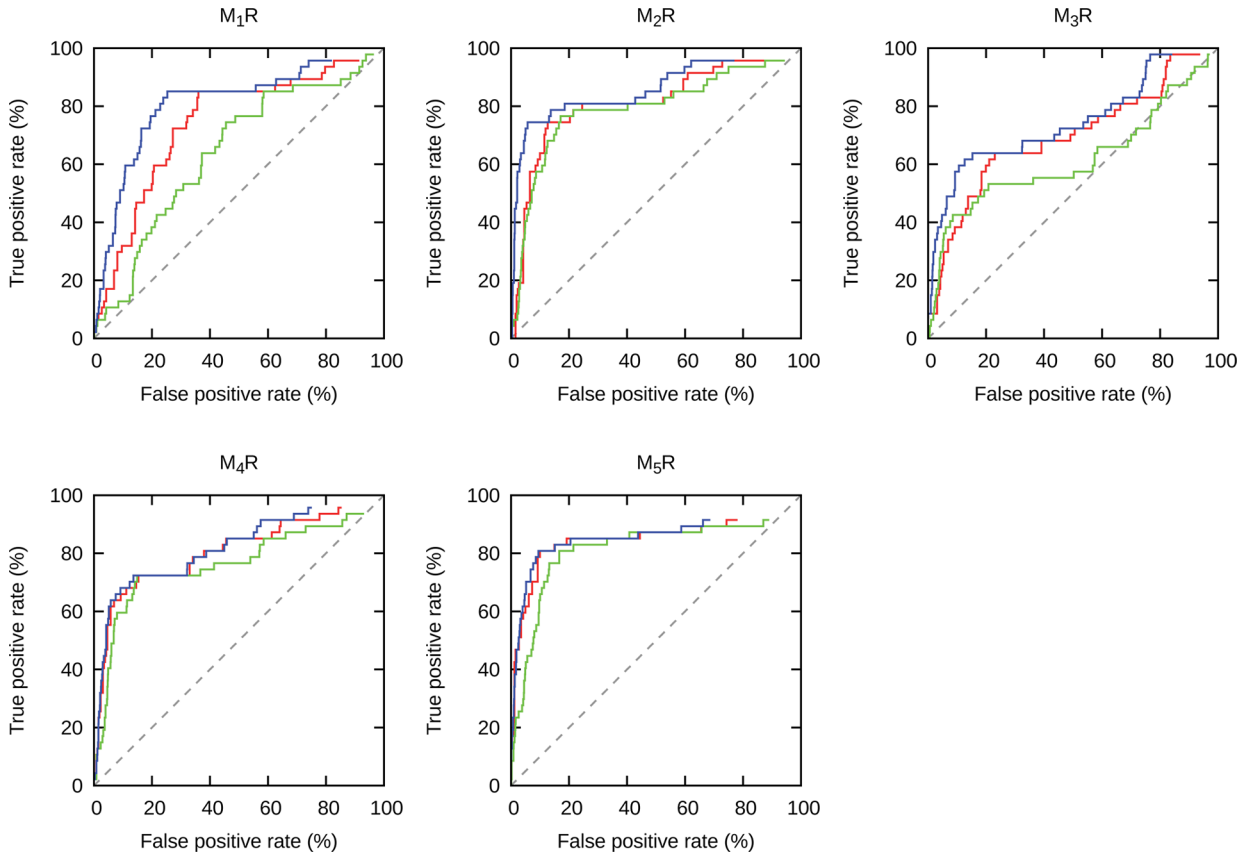


Figure 2. ROC curves for M_1R–M_5R models: (blue) set 1, Schrödinger; (green) set 2, ZINC; (red) set 3, refined Schrödinger. The dotted line indicates random choice (no enrichment).

Table 4. Virtual Screening Evaluation of Muscarinic Receptors

receptor	ROC AUC	$LogAUC_{0.001}$	mean ndist (Å)	EF (at *X* % of ranked database) 2	5	10
set 1 (Schrödinger decoy set)						
M_1R	0.81	0.35	3.89	5.3	5.5	4.7
M_2R	0.86	0.50	3.72	11.7	11.4	7.4
M_3R	0.74	0.38	4.11	8.5	7.6	4.9
M_4R	0.82	0.41	5.07	7.4	8.4	6.4
M_5R	0.85	0.53	4.00	12.7	10.1	7.4
set 2 (ZINC decoy set)						
M_1R	0.64	0.22	3.92	3.2	2.1	1.3
M_2R	0.79	0.36	3.57	3.2	5.9	5.3
M_3R	0.62	0.26	4.81	3.2	5.5	4.2
M_4R	0.76	0.35	5.10	6.4	5.5	5.7
M_5R	0.81	0.40	4.02	8.5	5.5	5.3
set 3 (refined Schrödinger decoy set)						
M_1R	0.74	0.28	3.89	2.8	2.5	2.3
M_2R	0.81	0.36	3.72	2.8	3.7	4.0
M_3R	0.69	0.30	4.11	3.7	2.9	3.0
M_4R	0.81	0.40	5.07	4.7	4.5	4.4
M_5R	0.84	0.51	4.00	5.6	5.3	5.1

multiple β_2AR crystal structures gave $EF_{2\%}$ = 0.3–11.7 and $EF_{10\%}$ = 1.5–3.9.[59] While our results compare favorably with the cited work, such comparisons should not be overinterpreted given the studies used different actives, decoy sets, and receptor types.

The main deficiencies of the models are the failure to dock some of the actives, shown as a gap at the end of the ROC curves, and in the inability of the M_3R model to identify a substantial fraction of actives, shown by the M_3R plots dropping down to the "random" line at approximately 60% of the false positive rate when using the ZINC decoy set. The properties of actives that either did not dock or produced docked poses with a scoring energy greater than the set acceptable cutoff are reported in the Supporting Information (Table S3). This data suggests that the most likely reason for docking failure is the large size of these compounds; thus a better M_3R model might be developed by using an alternative bulkier IFD ligand.

As a simple evaluation of the binding geometries, we calculated the distance between the ionizable nitrogen of the actives and Asp 3.32 (ndist). Mean values for each set are reported in Table 4. In the majority of cases, ndist fell within the range exhibited by ligands in 22 GPCR crystal structures (2.52 Å (PDB ID: 2Y01)–4.02 Å (PDB ID: 4DAJ); mean 2.92 Å). This salt bridge and other key receptor–ligand hydrogen bonding and ionic interactions were observed among many of the top-ranked poses of active compounds. This confirmed that not only were the models capable of producing high enrichment, they were also generating the expected contacts. Figure 3 illustrates binding modes of three active ligands, demonstrating interactions with binding site residues. It could be therefore suggested that a requirement for ndist to be less than 4 Å may serve as a useful pharmacophore filter in prospective virtual screening against aminergic GPCRs. However, it should be noted that recent work by Lin et al.

dx.doi.org/10.1021/ci400502u | J. Chem. Inf. Model. 2014, 54, 243–253

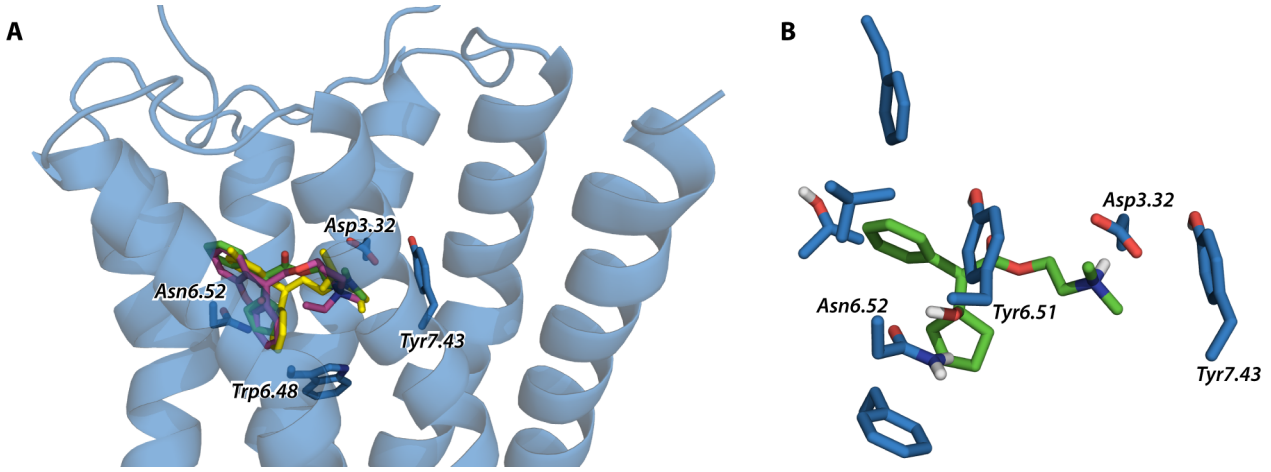


Figure 3. Cartoon representation of the M_2R model, showing the docked poses of the three highest ranked actives (A) and a close-up of cyclopentolate surrounded by its interacting residues (B). Color coding: cyclopentolate (green), tolterodine (yellow), and methantheline (pink). Binding site residues are blue.

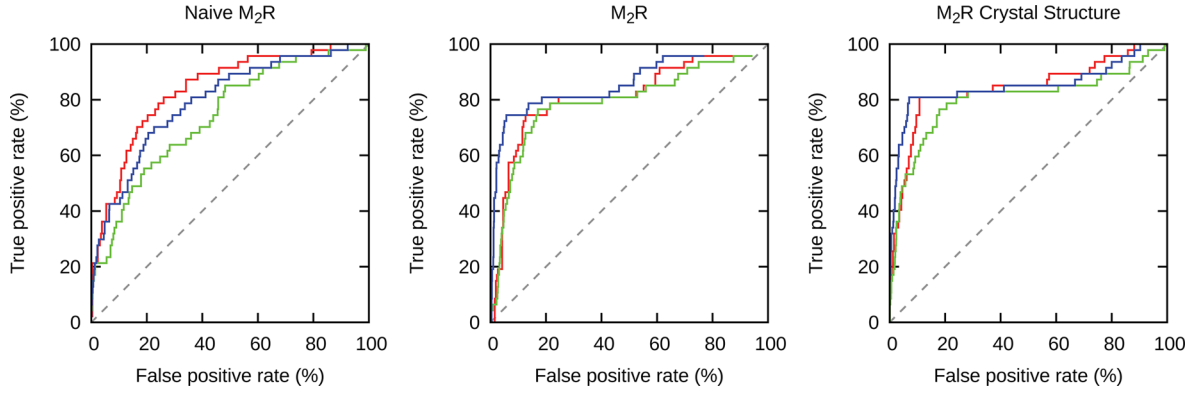


Figure 4. ROC curves for the M_2R naïve model (left), the M_2R optimized model (middle), and the M_2R crystal structure (right): (blue) set 1, Schrödinger; (green) set 2, ZINC; (red) set 3, refined Schrödinger. The dotted line indicates random choice (no enrichment).

has demonstrated that activity may be achieved without making this contact.[37]

Due to the high similarity of the five subtypes M_1R-M_5R, compounds which act at the M_1R usually also have some affinity for the other subtypes.[21] A rigorous test of model quality would be to dock compounds with a high level of specificity for individual subtypes into all subtypes so that an assessment of the selectivity of the homology models could be made. However, a significant difficulty encountered in this project has been to identify a sufficient number of compounds that are generally agreed to be selective for one receptor over the other four subtypes.

Comparison of Decoy Sets. The analysis of the VS data obtained using the Schrödinger decoy set (set 1) revealed that the results are biased toward low molecular weight compounds (across both active and decoy sets), which is reasonable given the characteristic small binding pocket of mAChRs.[24,25] Recent publications have given considerable attention to the development of receptor-appropriate decoy libraries.[15,59] Decoy sets, where the physical properties of compounds differ substantially from the corresponding active ligands, have been shown to lead to biased virtual screening results and often artificially good enrichment.[52]

Both the ZINC and refined Schrödinger decoy sets (sets 2 and 3) were matched to actives in terms of their physical properties, including the requirement to contain only compounds with an ionizable nitrogen at physiological pH. This filter was based on one of the benchmarks for model success, specifically their ability to generate the salt bridge between the ionizable nitrogen of a ligand and the Asp 3.32 residue of the receptor. Thus, these challenging sets of decoys were designed to investigate whether the docking and scoring process could select for this interaction in actives ahead of decoys that also contained an ionizable nitrogen.

The enrichment metrics (Table 4) and ROC and enrichment curves (Figures 2, S2, and S3 (Supporting Information)) demonstrate that indeed these sets of decoys are more challenging (particularly, for M_1R and M_3R). But encouragingly, the models produced enrichment and early enrichment values similar to that of nonproperty matched decoys (particularly, for M_4R and M_5R). These results indicate that our models are indeed capable of preferentially identifying active compounds among property-matched decoys.

Template Selection vs Binding Site Optimization. The choice of an appropriate template for GPCR homology modeling is an area of long-standing debate.[14−16] It has recently been demonstrated that, while important, the choice of template should be made while also considering issues such as binding site optimization and knowledge-enhancement of homology models. Specifically, Tropsha and co-workers[26] have compared the VS effectiveness of β_2AR crystal structures with a range of historical β_2AR models that were built before

dx.doi.org/10.1021/ci400502u | J. Chem. Inf. Model. 2014, 54, 243−253

the crystal structures became available. They demonstrated that several models produced VS enrichment comparable to and even exceeding that of crystal structures.

Here we investigated the proposal that an optimized homology model may approach the quality of a crystal structure, even though it is based on a remote template. Using the recently solved structure of the rat M_3R[25] as a template, we built a naïve human M_2R homology model, i.e. a homology model that has not been optimized by IFD. This naïve model had an RMSD of 1.64 Å to the human M_2R crystal structure. It can be seen from the results of VS (Figures 4, S4, and S5 (Supporting Information) and Table 5) that the optimized model, based on the β_2AR template, significantly outperforms the naïve M_3R-based variant and produces results close to those for the M_2R crystal structure. These results mirror those obtained in VS against the homology models of the MT_2 melatonin receptor where the $EF_{2\%}$ increased from 0 to 5.2 for a crude model to 3.1−18.7 for an optimized model.[16] Similar to the observations for other muscarinic models (Table 4), decoy sets 2 and 3 (ZINC and refined Schrödinger) make discrimination of decoys and actives more difficult. However, even with these demanding decoys, the optimized model still outperforms the naïve model in terms of enrichment, if not early enrichment.

Table 5. Virtual Screening Evaluation of M_2 Muscarinic Receptors

receptor	ROC AUC	$LogAUC_{0.001}$	ndist	EF (at X % of ranked database) 2	5	10
set 1 (Schrödinger decoy set)						
optimized model	0.86	0.47	3.72	11.7	11.4	7.4
naïve model	0.80	0.38	6.28	9.6	5.9	4.2
crystal structure	0.85	0.55	4.82	15.9	10.9	7.9
set 2 (ZINC decoy set)						
optimized model	0.79	0.36	3.57	3.2	5.9	5.3
naïve model	0.74	0.33	6.02	8.5	4.2	3.4
crystal structure	0.80	0.42	5.16	8.5	8.0	5.3
set 3 (refined Schrödinger decoy set)						
optimized model	0.81	0.36	3.72	2.8	3.7	4.0
naïve model	0.84	0.40	6.28	5.6	4.1	3.6
crystal structure	0.84	0.47	4.82	6.6	4.9	4.2

Binding site optimization takes into account the structural plasticity of a binding site and its adjustment to the structural demands of an active ligand. Our results demonstrate that the optimized M_2R model, based on the remote sequence template, was better at distinguishing actives from decoys than the naïve M_2R model, based on the close sequence template. Thus, it is clear that the choice of IFD ligand and the robustness of the IFD protocol could be as important for the production of a useful receptor model as the extent of target−template sequence similarity.

DISCUSSION

Several muscarinic receptor models have been generated over the past few years (summarized in Table 6), with the majority being of the M_1R.[60−69] Two models each of the M_3R[70,71] and the M_2R[72−74] and one model of the M_5R[75] have been also reported. These models were generally constructed in the course of molecular pharmacology studies to address issues of receptor activation and selectivity, allosterism, and bitopic binding, or receptor dimerization, although some groups have used predominantly modeling approaches to investigate the structural mechanisms of antagonist binding, receptor activa-

Table 6. Muscarinic Receptor Modeling Studies

receptor	template	homology modeling program	purpose	additional techniques used	ref
M_1R	rhodopsin	MODELLER	molecular pharmacology of allosteric modulation by a peptide ligand	loop modeling, MD, protein−protein docking	65
		Prime	molecular pharmacology of allosteric potentiation		62
		VEGA	modeling study to investigate receptor activation	MD in hydrated lipid bilayer	60
	β_2AR	MOE	modeling study to investigate allosteric modulation by a peptide ligand	MD in hydrated lipid bilayer, protein−protein docking	67
		QUANTA, MODELLER	molecular pharmacology of activation and selectivity	loop modeling	61, 76
			molecular pharmacology of allosterism and bitopic binding	loop modeling	63, 76
			molecular pharmacology of activation	loop modeling	64
	D_3R	MOE	molecular pharmacology of allosterism and bitopic binding	loop modeling	66
	M_3R	Prime	modeling study to investigate receptor activation	binding site refinement	77
	M_3R	MOE	molecular pharmacology of allosterism and bitopic binding		68
	M_2R	MOE	homology modeling		69
M_2R	M_3R	Prime, MODELLER, YASARA	modeling study to investigate the effect of template choice	IFD	72
	β_2AR	ICM	molecular pharmacology of allosterism and bitopic binding	flexible receptor docking of two agonists using BDMC algorithm	73, 74
M_3R	rhodopsin	MODELLER	modeling study to investigate structural mechanism of antagonist binding	MD in hydrated lipid bilayer	71
	β_1AR	Prime	molecular pharmacology of dimerization		70
M_5R	β_1AR	MODELLER	modeling study to investigate structural mechanism of antagonist binding	MD in hydrated lipid bilayer	75

dx.doi.org/10.1021/ci400502u | J. Chem. Inf. Model. 2014, 54, 243−253

tion, and allosteric modulation. A range of templates were used in these studies: rhodopsin, β_1AR and β_2AR, as well as the more recently solved D_3R, M_2R, and M_3R. Models were constructed using QUANTA/MODELLER, Prime, MOE, ICM, VEGA, and YASARA. Several approaches to additional model refinement were also implemented including MD in a hydrated lipid bilayer, loop modeling, and protein−protein docking. Significantly, the majority of the reported mAChR models were not optimized to generate knowledge-based models. In this study, we have developed such knowledge-based homology models of the muscarinic acetylcholine receptors M_1R−M_5R.

Binding site optimization has gained significant traction in the GPCR modeling field as an important way of using experimental knowledge (such as SAR and/or site-directed mutagenesis) to improve the quality and predictive power of naïve, or crude, homology models. Using 5-$HT_{2A}R$ as a test-case,[27] we have previously demonstrated the importance of loop refinement and, particularly, binding site optimization for improving model quality and VS performance. Such improvements have been also achieved for GPCR models in a number of studies focused on what has been termed ligand-steered,[78] ligand-guided,[79,80] ligand-adapted,[16] or ligand-optimized[15] homology modeling (Table 7). Binding site optimization via

Table 7. Binding Site Optimization Methods

method	receptor	ref
randomizing and clustering receptor complex structures	β_2AR and $A_{2A}R$	78
side chain conformation sampling in the presence of docked ligands	5-HT_XRs	37
	A_XRs	80
backbone perturbation and binding site reshaping with elastic normal-mode analysis	CXCR4	30, 79
IFD	MT_2 melatonin receptor	16
	D_1R and D_2R	15

a variety of methods—particularly those utilizing available experimental data about a target and its ligands—have been commonly used and shown to be successful in GPCR Dock assessments.[19,20]

Ideally, an optimized model, based on a close sequence template, would be the best choice for virtual screening.[69,72] However, close sequence templates are not always available. In such cases, knowledge-based optimization, e.g. by using established actives, can improve a model (Table 7). Using the M_2R as a case study, we compared a naïve model, based on a close sequence template (M_3R), and an optimized model, based on a more remote template (β_2AR). The IFD optimized model outperformed the naïve model in virtual screening. This observation parallels that of Kolaczkowski et al.[15] who generated ligand-optimized homology models of the D_1 and D_2 dopamine receptors using IFD and tested them in VS against ZINC- and Schrödinger-based decoy libraries spiked with ligands specific for dopamine receptors. They found that binding site optimization significantly improved VS performance, while observing no advantage in using a D_3R-based D_2R model compared to a model based on the more evolutionary distant β_2AR. Our findings are also in agreement with those of Tropsha and co-workers,[26] who suggest that such knowledge-based models "may be even more useful for practical structure-based drug discovery than X-ray structures".[26] Thus, our results and those of others[15,17,26] provide evidence that binding site optimization greatly improves homology models for VS. Future work is required to evaluate homology models in a flexible receptor scenario: by on-the-fly receptor flexibility,[81,82] molecular dynamics,[83] or using receptor ensembles.[84,85]

Finally, we tested the M_1R−M_5R models against increasingly demanding decoy sets. Specifically, to avoid artificial enrichment due to active-favoring biases, we have matched physicochemical decoy properties to those of ligands active at muscarinic receptors. The results showed that indeed these sets of decoys were more challenging. However, even using our matched decoy sets, the models produced enrichment (including early enrichment), similar to that obtained using nonproperty matched decoys. Recently, Gatica and Cavasotto have published a GPCR decoy database, where 39 decoy molecules were selected for each GPCR ligand.[59] Similar to our findings, they observed a marked decrease in enrichment for matched decoys compared to bias-uncorrected decoys.

CONCLUSIONS

In this work, we have developed homology models of the muscarinic acetylcholine receptors M_1R−M_5R and evaluated them in VS for the identification of antagonists. The models were generated by Prime and optimized using IFD (Glide + Prime). Model refinement was guided by experimental knowledge of active compounds and critical binding site residues. The refinement resulted in ligand-induced adaptation of the receptor binding sites, which optimized them for antagonist recognition. The homology models were evaluated in retrospective VS using Glide and were capable of distinguishing known antagonists from matched decoy compounds. These results bolster confidence for prospective virtual screening using these receptor models. Even more significantly, our results support the following suppositions about homology modeling of GPCRs: (i) binding site optimization is a crucial step in model generation, (ii) knowledge-based homology models of GPCRs are appropriate for prospective VS, and (iii) property-matched decoys should be used in VS evaluation of homology models. In line with our past practice, we make the optimized M_1R−M_5R homology models freely available as part of the Supporting Information. We consider such open access as crucial in our field since it allows researchers to use these structures, compare them to their own results,[37,69] and thus advance the development of better modeling methods.

ASSOCIATED CONTENT

Supporting Information

List of actives used in virtual screening enrichment studies; properties of actives that either did not dock into M_1R−M_5R models or produced docked poses with a scoring energy greater than the set acceptable cutoff; enrichment plots and semilog ROC curves for M_1R−M_5R models and for the M_2R optimized model, compared to the M_2R naïve model and the M_2R crystal structure; and PDB files of homology models. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*Phone: +61 3 9903 9611. E-mail: Elizabeth.Yuriev@monash.edu.

*Phone: +61 3 9903 9110. E-mail: David.Chalmers@monash.edu.

Present Address

[§]F.M.M.: Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, 9500 Gilman Drive, MC 0747, La Jolla, CA, 92093-0747.

Author Contributions

[†]T.T. and K.C.M.: These authors have contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

T.T. is a recipient of an Australian Postgraduate Award (APA) scholarship. K.C.M. was a recipient of a UROP scholarship supported by BIO21, The University of Melbourne. F.M.M. was a recipient of an Australian Postgraduate Award (APA) scholarship. E.Y. was a recipient of the travel award provided by the Australian Academy of Science (COST Action CM1103). This work was supported by the Victorian Life Sciences Computation Initiative (VLSCI, grant number VR0004) and by the National Computational Infrastructure (grant number: y96), which is supported by the Australian Commonwealth Government.

■ REFERENCES

(1) Shoichet, B. K.; Kobilka, B. K. Structure-based drug screening for G-protein-coupled receptors. *Trends Pharmacol. Sci.* **2012**, *33*, 268−272.

(2) Mason, J. S.; Bortolato, A.; Congreve, M.; Marshall, F. H. New insights from structural biology into the druggability of G protein-coupled receptors. *Trends Pharmacol. Sci.* **2012**, *33*, 249−260.

(3) Granier, S.; Kobilka, B. A new era of GPCR structural and chemical biology. *Nat. Chem. Biol.* **2012**, *8*, 670−673.

(4) Stevens, R. C.; Cherezov, V.; Katritch, V.; Abagyan, R.; Kuhn, P.; Rosen, H.; Wuthrich, K. The GPCR Network: a large-scale collaboration to determine human GPCR structure and function. *Nat. Rev. Drug. Discov.* **2013**, *12*, 25−34.

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

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

(7) Rasmussen, S. G.; Choi, H. J.; Rosenbaum, D. M.; Kobilka, T. S.; Thian, F. S.; Edwards, P. C.; Burghammer, M.; Ratnala, V. R.; Sanishvili, R.; Fischetti, R. F.; Schertler, G. F.; Weis, W. I.; Kobilka, B. K. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* **2007**, *450*, 383−387.

(8) Rosenbaum, D. M.; Cherezov, V.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Yao, X. J.; Weis, W. I.; Stevens, R. C.; Kobilka, B. K. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* **2007**, *318*, 1266−1273.

(9) Serrano-Vega, M. J.; Magnani, F.; Shibata, Y.; Tate, C. G. Conformational thermostabilization of the β1-adrenergic receptor in a detergent-resistant form. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 877−882.

(10) Caffrey, M.; Li, D.; Dukkipati, A. Membrane protein structure determination using crystallography and lipidic mesophases: recent advances and successes. *Biochemistry* **2012**, *51*, 6266−6288.

(11) Topiol, S. X-ray structural information of GPCRs in drug design: what are the limitations and where do we go? *Expert Opin. Drug Discov.* **2013**, *8*, 607−620.

(12) Fredriksson, R.; Lagerstrom, M. C.; Lundin, L. G.; Schioth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **2003**, *63*, 1256−1272.

(13) Kooistra, A. J.; Roumen, L.; Leurs, R.; de Esch, I. J.; de Graaf, C. From heptahelical bundle to hits from the haystack: structure-based virtual screening for GPCR ligands. *Methods Enzymol.* **2013**, *522*, 279−336.

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

(15) Kolaczkowski, M.; Bucki, A.; Feder, M.; Pawlowski, M. Ligand-optimized homology models of D_1 and D_2 Dopamine receptors: Application for virtual screening. *J. Chem. Inf. Model.* **2013**, *53*, 638−648.

(16) Pala, D.; Beuming, T.; Sherman, W.; Lodola, A.; Rivara, S.; Mor, M. Structure-based virtual screening of MT_2 Melatonin receptor: Influence of template choice and structural refinement. *J. Chem. Inf. Model.* **2013**, *53*, 821−835.

(17) Beuming, T.; Sherman, W. Current assessment of docking into GPCR crystal structures and homology models: successes, challenges, and guidelines. *J. Chem. Inf. Model.* **2012**, *52*, 3263−3277.

(18) Anighoro, A.; Rastelli, G. Enrichment factor analyses on G-protein coupled receptors with known crystal structure. *J. Chem. Inf. Model.* **2013**, *53*, 739−743.

(19) Kufareva, I.; Rueda, M.; Katritch, V.; Stevens, R. C.; Abagyan, R. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **2011**, *19*, 1108−1126.

(20) Michino, M.; Abola, E.; Brooks, C. L., 3rd; Dixon, J. S.; Moult, J.; Stevens, R. C. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat. Rev. Drug Discov.* **2009**, *8*, 455−463.

(21) Wess, J.; Eglen, R. M.; Gautam, D. Muscarinic acetylcholine receptors: mutant mice provide new insights for drug development. *Nat. Rev. Drug. Discov.* **2007**, *6*, 721−733.

(22) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534−553.

(23) Wacker, D.; Fenalti, G.; Brown, M. A.; Katritch, V.; Abagyan, R.; Cherezov, V.; Stevens, R. C. Conserved binding mode of human beta2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J. Am. Chem. Soc.* **2010**, *132*, 11443−11445.

(24) Haga, K.; Kruse, A. C.; Asada, H.; Yurugi-Kobayashi, T.; Shiroishi, M.; Zhang, C.; Weis, W. I.; Okada, T.; Kobilka, B. K.; Haga, T.; Kobayashi, T. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* **2012**, *482*, 547−551.

(25) Kruse, A. C.; Hu, J.; Pan, A. C.; Arlow, D. H.; Rosenbaum, D. M.; Rosemond, E.; Green, H. F.; Liu, T.; Chae, P. S.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Wess, J.; Kobilka, B. K. Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* **2012**, *482*, 552−556.

(26) Tang, H.; Wang, X. S.; Hsieh, J. H.; Tropsha, A. Do crystal structures obviate the need for theoretical models of GPCRs for structure-based virtual screening? *Proteins* **2012**, *80*, 1503−1521.

(27) McRobb, F. M.; Capuano, B.; Crosby, I. T.; Chalmers, D.; Yuriev, E. Homology modeling and docking evaluation of aminergic G protein-coupled receptors. *J. Chem. Inf. Model.* **2010**, *50*, 626−637.

(28) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **2012**, *52*, 1757−1768.

(29) Carlsson, J.; Coleman, R. G.; Setola, V.; Irwin, J. J.; Fan, H.; Schlessinger, A.; Sali, A.; Roth, B. L.; Shoichet, B. K. Ligand discovery from a dopamine D_3 receptor homology model and crystal structure. *Nat. Chem. Biol.* **2011**, *7*, 769−778.

(30) Mysinger, M. M.; Weiss, D. R.; Ziarek, J. J.; Gravel, S.; Doak, A. K.; Karpiak, J.; Heveker, N.; Shoichet, B. K.; Volkman, B. F. Structure-based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5517−5522.

(31) Yoshikawa, Y.; Oishi, S.; Kubo, T.; Tanahara, N.; Fujii, N.; Furuya, T. Optimized method of G-protein-coupled receptor homology modeling: Its application to the discovery of novel CXCR7 ligands. *J. Med. Chem.* **2013**, *56*, 4236−4251.
(32) van Loenen, P. B.; de Graaf, C.; Verzijl, D.; Leurs, R.; Rognan, D.; Peters, S. L. M.; Alewijnse, A. E. Agonist-dependent effects of mutations in the sphingosine-1-phosphate type 1 receptor. *Eur. J. Pharmacol.* **2011**, *667*, 105−112.
(33) Sanders, M. P.; Roumen, L.; van der Horst, E.; Lane, J. R.; Vischer, H. F.; van Offenbeek, J.; de Vries, H.; Verhoeven, S.; Chow, K. Y.; Verkaar, F.; Beukers, M. W.; McGuire, R.; Leurs, R.; Ijzerman, A. P.; de Vlieg, J.; de Esch, I. J.; Zaman, G. J.; Klomp, J. P.; Bender, A.; de Graaf, C. A prospective cross-screening study on G-protein-coupled receptors: Lessons learned in virtual compound library design. *J. Med. Chem.* **2012**, *55*, 5311−5325.
(34) Langmead, C. J.; Andrews, S. P.; Congreve, M.; Errey, J. C.; Hurrell, E.; Marshall, F. H.; Mason, J. S.; Richardson, C. M.; Robertson, N.; Zhukov, A.; Weir, M. Identification of novel adenosine A_{2A} receptor antagonists by virtual screening. *J. Med. Chem.* **2012**, *55*, 1904−1909.
(35) Kołaczkowski, M.; Nowak, M.; Pawłowski, M.; Bojarski, A. J. Receptor-based pharmacophores for serotonin 5-HT_7R antagonists: Implications to selectivity. *J. Med. Chem.* **2006**, *49*, 6732−6741.
(36) Kurczab, R.; Nowak, M.; Chilmonczyk, Z.; Sylte, I.; Bojarski, A. J. The development and validation of a novel virtual screening cascade protocol to identify potential serotonin 5-HT(7)R antagonists. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2465−2468.
(37) Lin, X.; Huang, X. P.; Chen, G.; Whaley, R.; Peng, S.; Wang, Y.; Zhang, G.; Wang, S. X.; Wang, S.; Roth, B. L.; Huang, N. Life beyond kinases: structure-based discovery of sorafenib as nanomolar antagonist of 5-HT receptors. *J. Med. Chem.* **2012**, *55*, 5749−5759.
(38) Heifetz, A.; Barker, O.; Verquin, G.; Wimmer, N.; Meutermans, W.; Pal, S.; Law, R. J.; Whittaker, M. Fighting obesity with a sugar-based library: Discovery of novel MCH-1R antagonists by a new computational-VAST approach for exploration of GPCR binding sites. *J. Chem. Inf. Model.* **2013**, *53*, 1084−1099.
(39) Renault, N.; Laurent, X.; Farce, A.; El Bakali, J.; Mansouri, R.; Gervois, P.; Millet, R.; Desreumaux, P.; Furman, C.; Chavatte, P. Virtual screening of CB_2 receptor agonists from bayesian network and high-throughput docking: structural insights into agonist-modulated GPCR features. *Chem. Biol. Drug. Des.* **2013**, *81*, 442−454.
(40) Suite 2012: Maestro, version 9.3; LigPrep, version 2.5; Schrödinger Suite 2012 Protein Preparation Wizard; Schrödinger Suite 2012 Induced Fit Docking protocol; Glide version 5.8; Prime version 3.1; Schrödinger, LLC: New York, NY, 2012.
(41) Suite 2011: Maestro, version 9.2; LigPrep, version 2.5; Schrödinger Suite 2011 Protein Preparation Wizard; Schrödinger Suite 2011 Induced Fit Docking protocol; Glide version 5.7; Prime version 3.0; Schrödinger, LLC: New York, NY, 2011.
(42) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J.; Honig, B.; Shaw, D. E.; Friesner, R. A. A hierarchical approach to all-atom protein loop prediction. *Proteins* **2004**, *55*, 351−367.
(43) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673−4680.
(44) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221−234.
(45) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739−1749.
(46) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.
(47) Chalmers, D. K.; Roberts, B. P. *Silico—A Perl Molecular Modelling Toolkit*; Monash University: Melbourne, 2011.
(48) Ballesteros, J. A.; Weinstein, H.; Stuart, C. S. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **1995**, *25*, 366−428.
(49) Holst, B.; Nygaard, R.; Valentin-Hansen, L.; Bach, A.; Engelstoft, M. S.; Petersen, P. S.; Frimurer, T. M.; Schwartz, T. W. A conserved aromatic lock for the tryptophan rotameric switch in TM-VI of seven-transmembrane receptors. *J. Biol. Chem.* **2010**, *285*, 3973−3985.
(50) Spalding, T. A.; Birdsall, N. J.; Curtis, C. A.; Hulme, E. C. Acetylcholine mustard labels the binding site aspartate in muscarinic acetylcholine receptors. *J. Biol. Chem.* **1994**, *269*, 4092−4097.
(51) Okuno, Y.; Tamon, A.; Yabuuchi, H.; Niijima, S.; Minowa, Y.; Tonomura, K.; Kunimoto, R.; Feng, C. GLIDA: GPCR−ligand database for chemical genomics drug discovery−database and tools update. *Nucleic Acids Res.* **2008**, *36*, D907−D912.
(52) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789−6801.
(53) Shi, L.; Javitch, J. A. The binding site of aminergic G protein-coupled receptors: the transmembrane segments and second extracellular loop. *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 437−467.
(54) Bymaster, F. P.; Felder, C. C.; Tzavara, E.; Nomikos, G. G.; Calligaro, D. O.; McKinzie, D. L. Muscarinic mechanisms of antipsychotic atypicality. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2003**, *27*, 1125−1143.
(55) Hawkins, P. C. D.; Warren, G. L.; Skillman, A. G.; Nicholls, A. How to do an evaluation: pitfalls and traps. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 179−190.
(56) Nicholls, A. What do we know and when do we know it? *J. Comput.-Aided Mol. Des.* **2008**, *22*, 239−255.
(57) Katritch, V.; Rueda, M.; Lam, P. C.; Yeager, M.; Abagyan, R. GPCR 3D homology models for ligand screening: lessons learned from blind predictions of adenosine A2a receptor complex. *Proteins* **2010**, *78*, 197−211.
(58) Mysinger, M. M.; Shoichet, B. K. Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model.* **2010**, *50*, 1561−1573.
(59) Gatica, E. A.; Cavasotto, C. N. Ligand and decoy sets for docking to G protein-coupled receptors. *J. Chem. Inf. Model.* **2012**, *52*, 1−6.
(60) Espinoza-Fonseca, L. M.; Pedretti, A.; Vistoli, G. Structure and dynamics of the full-length M1 muscarinic acetylcholine receptor studied by molecular dynamics simulations. *Arch. Biochem. Biophys.* **2008**, *469*, 142−150.
(61) Lebon, G.; Langmead, C. J.; Tehan, B. G.; Hulme, E. C. Mutagenic mapping suggests a novel binding mode for selective agonists of M1 muscarinic acetylcholine receptors. *Mol. Pharmacol.* **2009**, *75*, 331−341.
(62) Ma, L.; Seager, M. A.; Wittmann, M.; Jacobson, M.; Bickel, D.; Burno, M.; Jones, K.; Graufelds, V. K.; Xu, G.; Pearson, M.; McCampbell, A.; Gaspar, R.; Shughrue, P.; Danziger, A.; Regan, C.; Flick, R.; Pascarella, D.; Garson, S.; Doran, S.; Kreatsoulas, C.; Veng, L.; Lindsley, C. W.; Shipe, W.; Kuduk, S.; Sur, C.; Kinney, G.; Seabrook, G. R.; Ray, W. J. Selective activation of the M1 muscarinic acetylcholine receptor achieved by allosteric potentiation. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15950−15955.
(63) Avlani, V. A.; Langmead, C. J.; Guida, E.; Wood, M. D.; Tehan, B. G.; Herdon, H. J.; Watson, J. M.; Sexton, P. M.; Christopoulos, A. Orthosteric and allosteric modes of interaction of novel selective agonists of the M1 muscarinic acetylcholine receptor. *Mol. Pharmacol.* **2010**, *78*, 94−104.
(64) Kaye, R. G.; Saldanha, J. W.; Lu, Z. L.; Hulme, E. C. Helix 8 of the M1 muscarinic acetylcholine receptor: scanning mutagenesis delineates a G protein recognition site. *Mol. Pharmacol.* **2011**, *79*, 701−709.
(65) Marquer, C.; Fruchart-Gaillard, C.; Letellier, G.; Marcon, E.; Mourier, G.; Zinn-Justin, S.; Menez, A.; Servent, D.; Gilquin, B.

Structural model of ligand-G protein-coupled receptor (GPCR) complex based on experimental double mutant cycle data: MT7 snake toxin bound to dimeric hM1 muscarinic receptor. *J. Biol. Chem.* **2011**, *286*, 31661−31675.
(66) Daval, S. B.; Valant, C.; Bonnet, D.; Kellenberger, E.; Hibert, M.; Galzi, J. L.; Ilien, B. Fluorescent derivatives of AC-42 to probe bitopic orthosteric/allosteric binding mechanisms on muscarinic M1 receptors. *J. Med. Chem.* **2012**, *55*, 2125−2143.
(67) Xu, J.; Chen, H. Interpreting the structural mechanism of action for MT7 and human muscarinic acetylcholine receptor 1 complex by modeling protein-protein interaction. *J. Biomol. Struct. Dyn.* **2012**, *30*, 30−44.
(68) Daval, S. B.; Kellenberger, E.; Bonnet, D.; Utard, V.; Galzi, J. L.; Ilien, B. Exploration of the orthosteric/allosteric interface in human M1 muscarinic receptors by bitopic fluorescent ligands. *Mol. Pharmacol.* **2013**, *84*, 71−85.
(69) Jójárt, B.; Balint, A. M.; Balint, S.; Viskolcz, B. Homology modeling and validation of the human M1 muscarinic acetylcholine receptor. *Mol. Inf.* **2012**, *31*, 635−638.
(70) McMillin, S. M.; Heusel, M.; Liu, T.; Costanzi, S.; Wess, J. Structural basis of M3 muscarinic receptor dimer/oligomer formation. *J. Biol. Chem.* **2011**, *286*, 28584−28598.
(71) Martinez-Archundia, M.; Cordomi, A.; Garriga, P.; Perez, J. J. Molecular modeling of the M3 acetylcholine muscarinic receptor and its binding site. *J. Biomed. Biotechnol.* **2012**, *2012*, 789741.
(72) Jakubik, J.; Randakova, A.; Dolezal, V. On homology modeling of the M2 muscarinic acetylcholine receptor subtype. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 525−538.
(73) Valant, C.; Gregory, K. J.; Hall, N. E.; Scammells, P. J.; Lew, M. J.; Sexton, P. M.; Christopoulos, A. A novel mechanism of G protein-coupled receptor functional selectivity. Muscarinic partial agonist McN-A-343 as a bitopic orthosteric/allosteric ligand. *J. Biol. Chem.* **2008**, *283*, 29312−29321.
(74) Gregory, K. J.; Hall, N. E.; Tobin, A. B.; Sexton, P. M.; Christopoulos, A. Identification of orthosteric and allosteric site mutations in M2 muscarinic acetylcholine receptors that contribute to ligand-selective signaling bias. *J. Biol. Chem.* **2010**, *285*, 7459−7474.
(75) Huang, X.; Zheng, G.; Zhan, C. G. Microscopic binding of M5 muscarinic acetylcholine receptor with antagonists by homology modeling, molecular docking, and molecular dynamics simulation. *J. Phys. Chem. B* **2012**, *116*, 532−541.
(76) Blaney, F. E.; Raveglia, L. F.; Artico, M.; Cavagnera, S.; Dartois, C.; Farina, C.; Grugni, M.; Gagliardi, S.; Luttmann, M. A.; Martinelli, M.; Nadler, G. M.; Parini, C.; Petrillo, P.; Sarau, H. M.; Scheideler, M. A.; Hay, D. W.; Giardina, G. A. Stepwise modulation of neurokinin-3 and neurokinin-2 receptor affinity and selectivity in quinoline tachykinin receptor antagonists. *J. Med. Chem.* **2001**, *44*, 1675−1689.
(77) Chin, S. P.; Buckle, M. J. C.; Chalmers, D. K.; Yuriev, E.; Doughty, S. W. Towards activated homology models of the human M_1 muscarinic acetylcholine receptor. *J. Mol. Graph. Model.* **2014**, submitted.
(78) Phatak, S. S.; Gatica, E. A.; Cavasotto, C. N. Ligand-steered modeling and docking: a benchmarking study in class a g-protein-coupled receptors. *J. Chem. Inf. Model.* **2010**, *50*, 2119−2128.
(79) Neves, M. A.; Simoes, S.; Sa e Melo, M. L. Ligand-guided optimization of CXCR4 homology models for virtual screening using a multiple chemotype approach. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 1023−1033.
(80) Katritch, V.; Kufareva, I.; Abagyan, R. Structure based prediction of subtype-selectivity for adenosine receptor antagonists. *Neuropharmacology* **2011**, *60*, 108−115.
(81) Yuriev, E.; Agostino, M.; Ramsland, P. A. Challenges and advances in computational docking: 2009 in review. *J. Mol. Recognit.* **2011**, *24*, 149−164.
(82) Yuriev, E.; Ramsland, P. A. Latest developments in molecular docking: 2010−2011 in review. *J. Mol. Recognit.* **2013**, *26*, 215−239.
(83) Miao, Y.; Nichols, S. E.; Gasper, P. M.; Metzger, V. T.; McCammon, J. A. Activation and dynamic network of the M2 muscarinic receptor. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 10982−10987.
(84) Bottegoni, G.; Rocchia, W.; Rueda, M.; Abagyan, R.; Cavalli, A. Systematic exploitation of multiple receptor conformations for virtual ligand screening. *PLoS ONE* **2011**, *6*, e18845.
(85) Rueda, M.; Bottegoni, G.; Abagyan, R. Recipes for the selection of experimental protein conformations for virtual screening. *J. Chem. Inf. Model.* **2010**, *50*, 186−193.