Chapter 2

Homology Modeling and Docking Evaluation of Human Muscarinic Acetylcholine Receptors

Trayder Thomas, David K. Chalmers, and Elizabeth Yuriev

Abstract

The development of GPCR homology models for virtual screening is an active area of research. Here we describe methods for homology modeling of the acetylcholine muscarinic receptors M_1R–M_5R. The models are based on the β_2-adrenergic receptor crystal structure as the template and binding sites are optimized for ligand binding. An important aspect of homology modeling is the evaluation of the models for their ability to discriminate between active compounds and (presumed) inactive decoy compounds by virtual screening. The predictive ability is quantified using enrichment factors, area under the ROC curve (AUC), and an early enrichment measure, LogAUC. The models produce good enrichment capacity, which demonstrates their unbiased predictive ability. The optimized M_1R–M_5R homology models have been made freely available to the scientific community to allow researchers to use these structures, compare them to their results, and thus advance the development of better modeling approaches.

Key words Acetylcholine muscarinic receptor, Binding site optimization, Decoy, Docking, GPCR, Homology modeling, Virtual screening

1 Introduction

The use of structure-based design methods for G protein-coupled receptors (GPCRs) commenced in the early 2000s with the landmark report of the structure of bovine rhodopsin [1]. The first crystal structures of ligand-infusible GPCRs became available in 2007 [2–4], and, at the time of writing, the number of available structures has grown to a total of 119 crystal structures for 22 receptor subtypes [5]. Despite the considerable technical advances in the field, GPCR crystallization remains an area of highly specialized expertise, and the solved structures make up only a small fraction of the ~800 GPCRs present in the human genome (including 342 nonolfactory receptors) [6]. It is accepted that the prospect of solving the structures of all members of the GPCR superfamily is not realistic in the foreseeable future [7, 8]. Therefore, when receptor models are required for structure-based investigations and experimental data is

Jaromir Myslivecek and Jan Jakubik (eds.), *Muscarinic Receptor: From Structure to Animal Models*, Neuromethods, vol. 107,
DOI 10.1007/978-1-4939-2858-3_2,

lacking, researchers turn to homology models. A homology model of a protein (also known as a comparative model) is an atomic-resolution model of a protein (the "target") built based on its amino acid sequence and experimental three-dimensional structure of a related homologous protein (the "template").

GPCR homology models are important tools for understanding GPCR function and for structure-based drug design [7–11]. Virtual screening campaigns against GPCR homology models have identified novel active agents for a range of GPCR targets [12] in a prospective manner (i.e., where compounds initially identified through a virtual screen have been sourced and experimentally validated). A wider overview of GPCR modeling is provided by reference [13]; for brief summaries of GPCR docking studies (as well as other docking-related surveys), see [14, 15].

This chapter describes the procedural steps involved in building, optimizing, and evaluating models of muscarinic acetylcholine receptors (mAChRs). We start with the overview of muscarinic receptor modeling (Section 2) and then discuss the approaches we have used to develop refined GPCR homology models which are able to identify active compounds through virtual screening (Sections 3–4) [12, 16].

There are five subtypes of muscarinic acetylcholine receptors, denoted M_1R–M_5R [17]. Development of mAChR ligands (particularly, subtype-selective ligands) holds potential for the treatment of many diseases such as Alzheimer's, schizophrenia, drug addiction, type 2 diabetes, and cancer [18].

2 Overview of Muscarinic Receptor Modeling

Several muscarinic receptor models have been generated over the past few years. They have employed a variety of different templates as the basis for homology model construction. The template proteins used to generate the homology models in each case are listed in Table 1.

3 Homology Modeling of Muscarinic Acetylcholine Receptors

The modeling workflow is shown in Fig. 1 and described in detail in the following sections. We have built homology models of mAChRs M_1–M_5 [12], using the β_2-adrenergic receptor (β_2AR) crystal structure (PDB ID: 2RH1) [2] as the template and employing the induced fit docking (IFD) procedure [19] to optimize the models to improve their identification of compounds which bind to their orthosteric binding sites. The predictive quality of all five models was assessed through retrospective virtual screening investigations. The results obtained using property-matched decoy libraries demonstrated the unbiased predictive capacity of these models.

Table 1
Templates used for modeling muscarinic receptors

Template	Receptor modeled	Template PDB ID	References
Rhodopsin	M_1R	1U19	[22]
		1F88	[35, 36, 39]
	M_2R	1U19	[26]
	M_3R	1GZM	[23]
β_2AR	M_1R	2RH1	[16, 28–31]
	M_2R	3D4S	[26]
	M_2R	2RH1	[20, 21, 26]
	M_1R–M_5R	2RH1	[12]
M_2R	M_1R	3UON	[34]
M_3R	M_1R	4DAJ	[33, 37]
	M_2R	4DAJ	[26]
	M_5R	4DAJ	[25]
D_3R[a]	M_1R	3PBL	[32]
β_1AR	M_2R	2VT4	[26]
	M_3R	2VT4	[38]
	M_5R	2VT4	[24]

[a]The original D_3R-based model was edited to replace the extracellular loop 2 by fragments extracted from the structures of the human β_2AR (PDB ID: 2RH1) and $A_{2A}AR$ (PDB ID: 3EML) receptors

3.1 Software

Many software packages have been used for modeling of mAChRs including ICM [20, 21], MODELLER [22–30], MOE [31–35], Prime [12, 26, 35–38], YASARA [26], QUANTA [27–30], and VEGA [39]. Molecular modeling steps and options described in this protocol refer to the Schrödinger software suite [40], as used by us [12, 16]. Default settings were used, unless otherwise stated.

The following Schrödinger modules and programs were used for specific tasks (**Note 1**):

1. Homology modeling—Prime [41].
2. Multiple sequence alignment—ClustalW [42].
3. Ligand preparation—LigPrep [43].
4. Ligand docking—Glide [44, 45].
5. Binding site optimization—IFD [19].
6. Computation of physical descriptors for comparison of the decoy sets with the active compounds—ChemAxon Marvin Calculator (cxcalc) (http://www.chemaxon.com): The descriptors include molecular weight (MW), number of rotatable bonds, number of hydrogen bond donor and acceptor atoms, calculated logP (ClogP), polar surface area (PSA), and vdW volume.

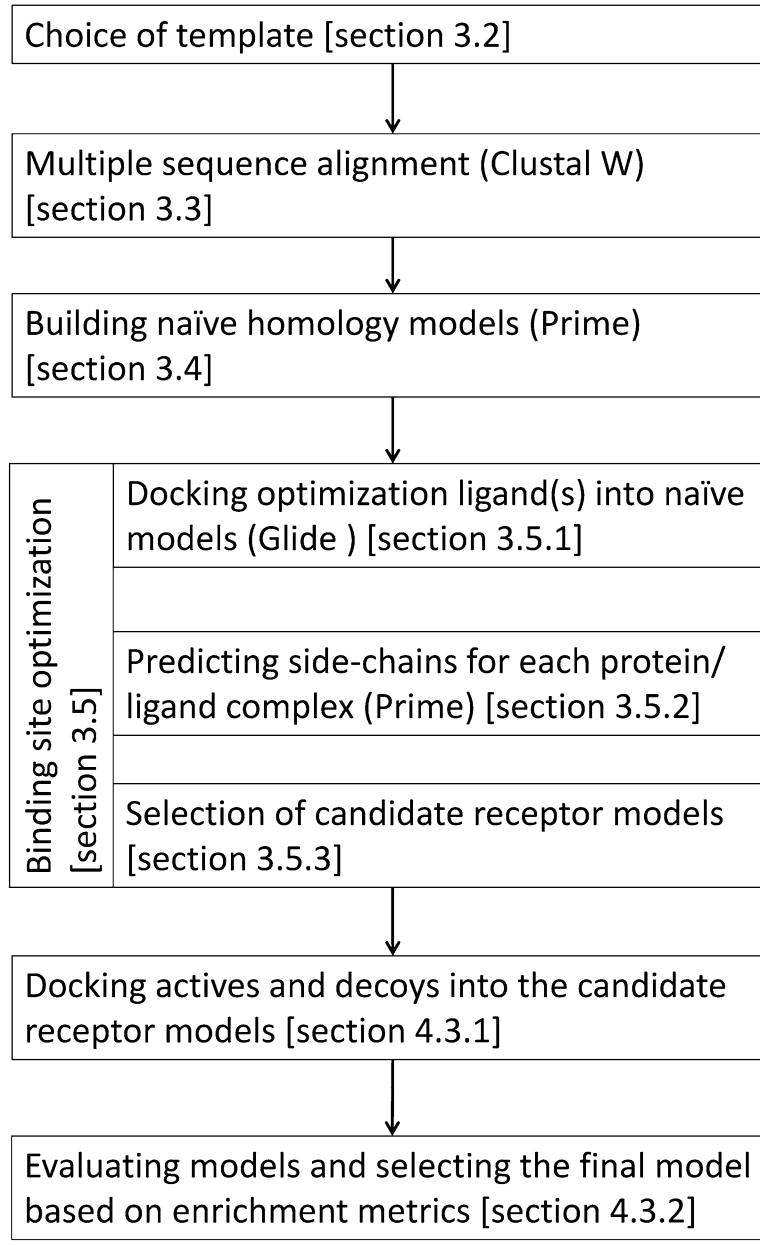

Fig. 1 Flow chart of homology modeling and model evaluation

7. Computation of the 2D Tanimoto score (using fragment sizes of 1–7 atoms, ignoring hydrogens) to demonstrate the diversity of the structures within the ligand sets—Silico [46].
8. Sorting of docked poses based on rank and calculating enrichment metrics—Silico [46].

3.2 Choice of Template

The choice of an appropriate template for GPCR homology modeling is an area of an ongoing debate [47–50]. While close sequence similarity is a very important factor in template selection, it has been shown that model refinement approaches such as binding site optimization and model enhancement based on experimental knowledge (discussed in more detail in Section 3.5 below) are capable of compensating for more distant sequence relationships [51]. Additionally, the ever-increasing number of available structures and improvements in methods for model refinement means that the most optimal choice of a template for a particular target cannot be determined once and for all and requires regular re-evaluation. As can be seen in Table 1, a variety of templates have been used as the basis for modeling of muscarinic acetylcholine receptors. Furthermore, it has also been shown that model quality can be improved by combining multiple templates [32].

A receptor model, built on a close sequence template, is usually considered to be preferable for structure-based drug design [26, 34]. However, when close sequence templates are not available (as still is the case for many GPCRs), knowledge-based optimization can be used to improve a model that is based on a more remote sequence template. We have previously demonstrated that an optimized model of the M_2R, based on a more remote template (β_2AR), outperformed a naïve (i.e., generated without binding site optimization) M_2R model, based on a close sequence template (M_3R), in virtual screening [12]. These results agree with similar observations for modeling the dopamine D_1 and D_2 receptors [48], the β_2AR [51], as well as a diverse panel of receptors (β_1A and β_2A, dopamine D_3, histamine H_1, muscarine M_2 and M_3, A_2A adenosine, S1P1, kappa-opioid, and C-X-C chemokine 4 receptors) [52].

3.3 Multiple Sequence Alignment

Due to a significant level of sequence conservation within the transmembrane regions (helices) of GPCRs, it is possible to align sequences by making use of highly conserved residues to identify the positions of gaps and inserts. GPCR sequences are available for download from the Universal Protein Resource (http://www.uniprot.org/). For effective receptor space coverage, representatives from different relevant GPCR subfamilies should be used for multiple sequence alignment. To establish the alignment between the mAChRs and the template (β_2AR) structure, we used ClustalW [12] (Fig. 1) to create a multiple sequence alignment of the muscarinic receptors with the human dopamine, serotonin, α- and β-adrenergic, adenosine, histamine, and bovine rhodopsin receptors. The ClustalW multiple sequence alignment required manual editing to remove gaps in helices and to ensure that highly conserved residues in each transmembrane helix were properly aligned.

3.4 Generation of a Starting (Naïve) Model

Using a well-aligned sequence of an appropriate template, a user can generate a starting (naïve) 3D receptor model. In our work,

naïve homology models of the five human mAChRs were built in Prime (Fig. 1) from the multiple sequence alignment, using the β_2-adrenergic receptor (PDB ID: 2RH1) crystal structure [2] as the template.

3.5 Binding Site Optimization

A naïve model is not necessarily very good at identifying active ligands and can be improved by binding site optimization. Binding site optimization takes into account the structural plasticity of a binding site and its ability to adjust to the structural demands of the ligand. Binding site optimization via a variety of approaches has been widely used to improve model quality. These have been variously described as ligand-steered [53], ligand-guided [54, 55], ligand-adapted [49], or ligand-optimized [48] homology modeling. Binding site optimization via a variety of approaches—mainly those employing accessible experimental data relating to a target and its ligand knowledge (such as structure-activity relationships and/or site-directed mutagenesis)—has been commonly used and shown to improve model quality in GPCR Dock assessments [56–58].

The following steps describe binding site optimization via induced fit docking, as implemented by us for muscarinic receptors using the IFD module of the Schrödinger software suite [12] (Fig. 1).

3.5.1 Docking Optimization Ligands into Naïve Models

The mAChR homology models are first treated by the Maestro Protein Preparation Wizard workflow [43] to add and minimize hydrogen atoms using the OPLS_2005 force field. Following model preparation, the ligand-binding site is refined by docking an appropriate ligand (here referred to as the "optimization ligand") into each of the homology models using Glide within the IFD protocol. The optimization ligand should ideally be representative of the hits a user intends to find, e.g., by having a similar scaffold. In order to identify a larger range of hits, it is important to consider how adjusting the binding site around the optimization ligand will limit the ability of other ligands to dock. For example, optimization of the residues around a smaller ligand can reduce the volume of the binding site and preclude the ability of larger ligands to dock successfully.

Glide docks ligands within a predefined, cuboid region. This site should be roughly centered on the binding site, and be large enough, so as to allow the binding of large ligands. We chose to center the docking site upon residues Asp 3.32, Trp 6.48, Phe 6.52, and Tyr 7.43 (Ballesteros-Weinstein nomenclature [59]) (**Note 2**). Both the van der Waals (vdW) radii and the partial atomic charges were scaled by 0.5 in order to collect a more diverse range of poses. In the initial Glide docking step, up to 50 poses per ligand were collected.

Clozapine and atropine have been demonstrated as useful optimizing ligands for IFD [12] since they have high affinity for the

M_1–M_5 receptors; reported clozapine K_i values vary from 1.4 to 5.0 nM and atropine K_i values range between 0.2 and 1.5 nM [60]. Following the virtual screening evaluation procedure (described below), the atropine-optimized model for the M_1R gave the best enrichment, while the best M_2R–M_5R models were optimized using clozapine [12].

3.5.2 Predicting Conformations of Binding Site Side Chains

The user should select specific residues to include into the binding site refinement. We recommend selecting the side-chain conformations of the residues within 5 Å of ligand atoms, excluding Asp 3.32 and Trp 6.48, for optimization with Prime. Asp 3.32 and Trp 6.48 play a critical role in correctly orienting ligand molecules within receptor binding sites (**Note 3**). 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 in virtual screening [12, 49].

3.5.3 Selection of Candidate Receptor Models

Following binding site optimization with Prime, candidate receptor models are selected by evaluating how the optimization ligand fits within the binding site. The optimization ligand should be re-docked into the optimized receptors with Glide using default vdW and charge scaling parameters. Multiple ligand-receptor poses for each model should be generated and inspected. At this stage receptor models are chosen on the basis of the position and orientation of the ligand within the binding site, key hydrogen bonding and vdW interactions, and the relative energy of interaction, which is a composite of the protein and ligand energy scores (Eq. 1). 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 should also be taken into account (**Note 4** and **5**):

$$\text{IFDScore} = \text{GlideScore} + 0.05 \times \text{Prime\_Energy} \quad (1)$$

A maximum of 20 poses are required to be collected for further evaluation of receptor models by retrospective virtual screening.

4 Evaluation of Muscarinic Acetylcholine Receptor Models by Virtual Screening

The predictive quality of the candidate receptor models is established by measuring the ability of the candidate models to distinguish between known active compounds and chemically similar decoy compounds with physicochemical properties that closely match those of the actives. Methods for evaluating homology modeling and virtual screening protocols as applied to GPCRs are a focus of active research [56–58, 63, 64].

4.1 Actives

Known antagonists of muscarinic receptors (actives) can be obtained from the GLIDA database [65] (http://pharminfo.pharm.kyoto-u.ac.jp/services/glida/). We have used a set of 48 actives (Note 6) in our retrospective virtual screening studies [12]. Reference [16] contains the chemical structures of these actives and Table 2 lists their average physicochemical properties. We used the Maestro module LigPrep to generate compound 3D structures and to assign tautomeric states and formal charges at physiological pH (pH 7.4 ± 2.0) of active and decoy compounds with a single, likely structure per compound being retained for screening.

4.2 Decoys

Although there are many large decoy libraries available for retrospective virtual screening studies, it has been demonstrated that a set of approximately 1000 molecules is sufficient to detect enrichment trends. For example, it has been shown that little library size-dependent behavior is detected when screening with the entire Directory of Useful Decoys (DUD) set of approximately 100,000 molecules compared to a randomly selected subset of 1000 DUD molecules [66]. It is important that the decoys must not be readily distinguishable from the active compounds. Decoy sets where the physicochemical properties of the decoys differ substantially from those of the active ligands have been shown to lead to biased virtual screening results, and often artificially good enrichment [66].

In our studies we have evaluated three sets of decoys. Table 2 lists the properties of the decoy and active compounds: molecular weight MW (g/mol), number of rotatable bonds (NRB), polar surface area PSA (Å^2), calculated logP, number of hydrogen bond donors and acceptors (HBD and HBA, respectively), solvent accessible volume (Å^3), and 2D Tanimoto score. It can be seen that, generally, the properties of the active compounds are similar to those of the decoy libraries.

4.2.1 Set 1

The Schrödinger decoy set (http://www.schrodinger.com), previously used by us [12, 16], contains 1000 drug-like decoys, randomly selected from a library of one million compounds having

Table 2
Average ligand properties [12]

Ligand set	MW (g/mol)	NRB	PSA (Å^2)	ClogP	HBD	HBA	vdW volume (Å^3)	2D Tanimoto score
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

properties characteristic of drug molecules [44, 45]. The molecular weights of this set vary from 151 to 645 g/mol, with an average of 360 g/mol. These decoys were not specifically selected to mimic muscarinic antagonist compounds.

4.2.2 Set 2

We derived the ZINC decoy set from the ZINC database (http://zinc.docking.org/) [67] (7.2 million compounds, database version 7) by a process of successive eliminations, generating a subset of molecules closely matching the physicochemical properties of the actives (Table 2). 1000 molecules were randomly selected satisfying the physicochemical properties criteria (Note 7). This set is more challenging in terms of distinguishing between decoys and active compounds.

4.2.3 Set 3

The refined Schrödinger decoy set is a subset of the decoy Set 1, with molecular weight limited to be consistent with that of the active compounds (260–410 g/mol). To generate this set, all compounds from the Schrödinger decoy library with a molecular weight outside the range of the active compounds and without ionizable nitrogen were removed. This set contains 261 molecules and is more challenging than Set 1 from which it was derived.

4.3 Enrichment Studies

4.3.1 Docking

To establish the ability of the receptor to identify muscarinic antagonists, the decoys and actives are docked into the candidate receptor models (Note 8), generated and selected at Sections 3.5.1–3.5.3 above (Fig. 1). The top pose for each ligand (determined by GlideScore) is retained following post-docking minimization.

Following docking, models should always be visually inspected to ensure that the ligands bind within the defined binding pocket. Furthermore, the intermolecular interactions in these poses should be examined to ensure that important expected interactions, based on mutagenesis studies [68], are observed between ligand and receptor molecules.

4.3.2 Numeric and Graphic Assessment of Models

Both enrichment plots and receiver operating characteristic curve (ROC) plots have been used to establish the performance of homology models and crystal structures. It has been argued that ROC curves are superior to enrichment plots: not only do they reflect the selection of actives, but also the non-selection of decoys [69, 70]. One metric that can be derived from ROC plots is the area under the curve (AUC). The AUC, which has an ideal value of 1, gives an indication of the general ability of the model to distinguish active compounds (true positives) from decoy compounds (true negatives). It should be noted that the AUC does not specifically focus on the best docking scores being allocated to active compounds (early enrichment). In order to weight the AUC towards early enrichment, the ROC curve can instead be plotted with a logarithmic *x*-axis. The resulting LogAUC [71] is then

calculated by computing the fraction of the ideal area under the semilog ROC curve. Another metric, NSQ_AUC [72], has also been developed to probe early, rather than overall, enrichment.

Calculating Curves

Due to the large scale of most virtual screening efforts there is significant advantage in automating the process of calculating ROC curves. There are many programs already designed to calculate ROC curves for popular docking packages. Below we detail the general procedure available in our scripts implemented in Silico [46].

The set of docked actives and decoys is ranked by GlideScore. The ROC curve is plotted by stepping through the list and plotting the cumulative fraction of actives encountered (true positives, *y*-value) against the cumulative fraction of decoys (true negatives, *x*-value).

An enrichment curve is generated similarly, differing in that the *x*-axis reflects the percentage of compounds encountered rather than decoys. The enrichment factors (EF) are the *y*-values corresponding to each % (*x*-value). Enrichment factors are usually calculated at 2, 5, and 10 % of the total number of compounds (N_{total}) screened, according to Eq. (2):

$$\text{EF} = \left(\frac{\text{Hits}_{\text{sampled}}}{N_{\text{sampled}}}\right) \div \left(\frac{\text{Hits}_{\text{total}}}{N_{\text{total}}}\right) \tag{2}$$

Area Under the Curve

The AUC is calculated by integrating the area under the ROC curve (Note 9) according to Eq. (3):

$$\text{AUC} = \sum_{1}^{d_{\text{T}}} \frac{1}{d_{\text{T}}} \times \frac{a_{\text{F}}}{a_{\text{T}}} \tag{3}$$

where d_{T} is the total number of decoys, a_{F} is the number of actives found, and a_{T} is the total number of actives.

LogROC curves follow a similar principle but weight the ROC curve to favor early enrichment. Because logarithmic curves have an asymptote at zero it is necessary to introduce a lower limit. This lower limit (λ) functions to restrict the length of the *x*-axis to a finite value. It is important to realize that LogAUC is only comparable for identical λ values (usually 0.001). There are also additional considerations: any coordinates smaller than λ should not be included in the calculations and unless one of the *x*-coordinates coincides with λ, an additional term needs to be added to calculate the first partial step.

In its reduced form, the formula for LogAUC is Eq. (4):

$$\text{for } \frac{d_{\text{F}}}{d_{\text{T}}} > \lambda, \text{LogAUC}_{\lambda} = \sum_{d=1}^{d_{\text{T}}} \frac{\log_{10}\left(\frac{d_{\text{F}}}{d_{\text{F}}-1}\right)}{-\log_{10}(\lambda)} \times \frac{a_{\text{F}}}{a_{\text{T}}} \tag{4}$$

However, this formula can be understood more intuitively when written as Eq. (5):

$$\text{for } \frac{d_F}{d_T} > \lambda, \text{LogAUC}_\lambda = \sum_{d=1}^{d_T} \frac{\log_{10}\left(\frac{d_F}{d_T}\right) - \log_{10}\left(\frac{d_F - 1}{d_T}\right)}{\log_{10}(1) - \log_{10}(\lambda)} \times \frac{a_F}{a_T} \quad (5)$$

To account for the first partial step, an additional term (Eq. 6) may need to be added to both of these previous formulas:

$$\text{if } \frac{d_F}{d_T} > \lambda \text{ and } \frac{d_F - 1}{d_T} < \lambda, \frac{\log_{10}\left(\frac{d_F}{d_T}\right) - \log_{10}(\lambda)}{\log_{10}(1) - \log_{10}(\lambda)} \times \frac{a_F}{a_T} \quad (6)$$

where d_F is the number of decoys found, d_T is the total number of decoys, a_F is the number of actives found, and a_T is the total number of actives.

The intuitive formula above can be visualized (Fig. 2) as a ratio of two distances on the x-axis multiplied by the height of the curve (the fraction of actives found). The numerator is the distance

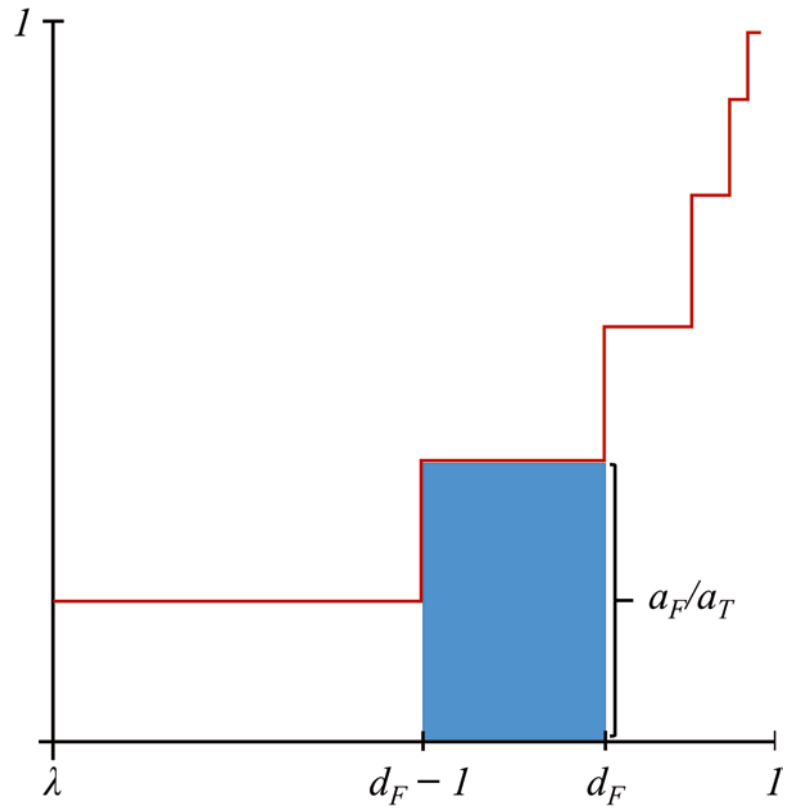

Fig. 2 Construction of a rectangle under a LogROC curve

between the decoy just found and the previous decoy (one horizontal step). The denominator is the length of the x-axis (λ to 1). The additional term is necessary because if none of the coordinates coincide with λ the first full horizontal step will precede λ.

Practically, this means that the LogAUC is being constructed in steps by a series of rectangles. Each step is calculated as a fraction of the Cartesian length of the logarithmic x-axis and multiplied by the height of the ROC curve at this point. In this way, the LogAUC of the perfect enrichment curve will be 1.

5 Representative Results

The receptor homology models are evaluated by testing their ability to rank active compounds above decoy molecules. In such evaluation, the decoy libraries (see Section 4.2 above) and active compounds (see Section 4.1 above) are docked into the receptor models. The optimization ligands, used for binding site optimization (see Section 3.5 above), should be excluded from virtual screening to remove any potential structural bias.

Representative ROC curves, enrichment plots, and semilogarithmic ROC curves for the M_5R model are shown in Fig. 3 and the corresponding metrics are listed in Table 3. The M_5R model shows excellent enrichment capacity and has particularly good early enrichment.

The main deficiency of the models is the failure to dock some of the actives, shown as a gap at the end of the ROC curves (Fig. 3). The properties of actives that did not dock or produced docked poses with a scoring energy greater than the set acceptable cutoff (a GlideScore of 0 kcal/mol) suggest that the most likely reason for this docking failure is the large size of these compounds. Therefore, a better model might be developed by using an alternative bulkier optimization ligand. Work is currently in progress in our laboratory which demonstrates that this is indeed the case and muscarinic receptor models generated using alternative optimization ligands are able to dock a wider range of compounds.

The metrics shown in Table 3 compare favorably with other similar docking studies (although such comparisons should not be over-interpreted given different actives, decoy sets, and receptor types used). For example, the MT_2 melatonin receptor models [49] which were based on the β_2AR and optimized for antagonists produced $EF_{2\%} = 3.1–18.7$; a range of antagonist-bound GPCR crystal structures gave $EF_{2\%} = 0.3–11.7$ and $EF_{10\%} = 1.5–3.9$ [64].

Both decoy sets 2 and 3 (ZINC and refined Schrödinger) were property-matched to actives. In addition, the decoy selection criteria for these sets included the requirement to contain only compounds with an ionizable nitrogen at physiological pH, based on the hypothesis that a ligand ionizable nitrogen should be able to

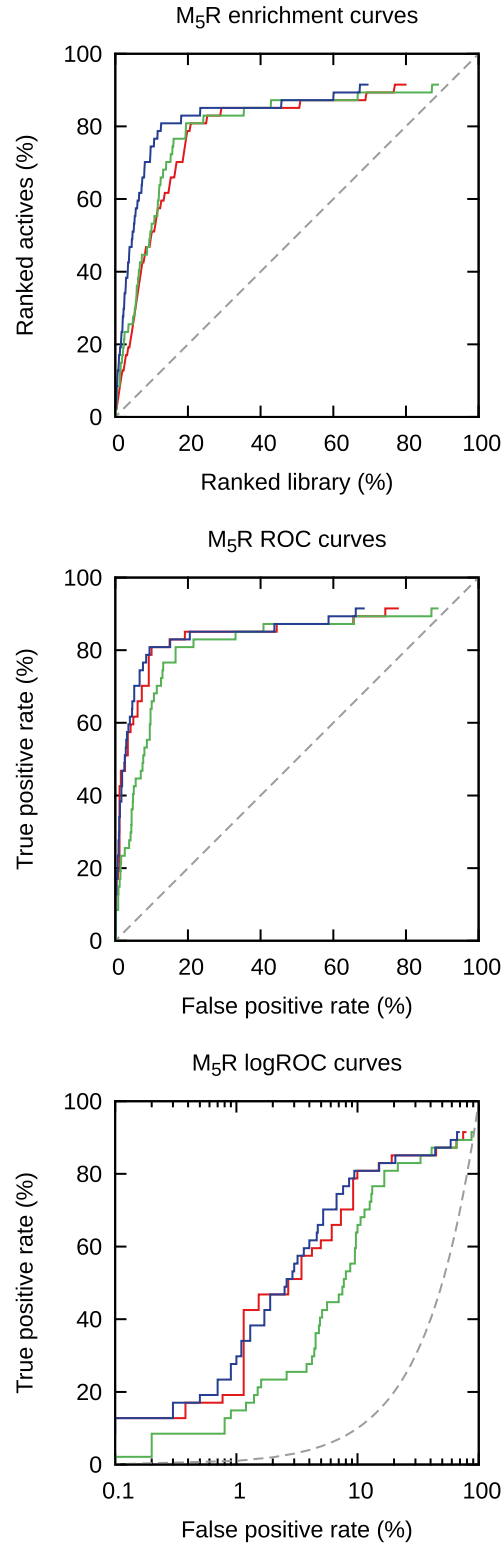

Fig. 3 Curves for the M_5R model. *Blue*: Set 1, Schrödinger; *green*: Set 2, ZINC; *red*: Set 3, refined Schrödinger. *Dotted line* indicates random choice (no enrichment) (Color figure online)

Table 3
Virtual screening metrics of the M_5R

Decoy set	AUC	$LogAUC_{0.001}$	EF (at x% of ranked database)		
			2	5	10
Set 1 (Schrödinger decoy set)	0.85	0.53	12.7	10.1	7.4
Set 2 (ZINC decoy set)	0.81	0.40	8.5	5.5	5.3
Set 3 (refined Schrödinger decoy set)	0.84	0.51	5.6	5.3	5.1

form the salt bridge with Asp 3.32. Thus, these decoys sets were designed to be challenging with respect to selecting for this interaction in actives ahead of decoys that also contained an ionizable nitrogen.

The enrichment metrics (Table 3) and ROC and enrichment curves (Fig. 3) demonstrate that indeed these sets of decoys are more challenging. However, they also show that the enrichment and early enrichment values are similar to values of non-property-matched decoys, indicating the model's capability to preferentially identify active compounds even amongst property-matched decoys.

The five subtypes M_1R–M_5R have pairwise sequence identities in the range of 50–70 %. For residues within 6 Å of the ligand (from the M_2R crystal structure) the sequence identity increases to 90–100 %. Due to this high similarity, compounds that act at one receptor subtype usually also have some affinity for the other subtypes [17]. A rigorous test of model quality would be to dock compounds with a high level of specificity for individual subtypes into all subtypes. However, a significant challenge encountered in homology modeling and evaluation of muscarinic acetylcholine receptors is to identify a sufficient number of compounds that are selective for one receptor over the other four subtypes.

6 Availability of Models

It is a significant problem for researchers interested in GPCR structure-based design that only a limited number of GPCR homology models (including models of the muscarinic acetylcholine receptors) are freely available for use and comparison. We aim to supplement the limited number of evaluated homology models that are available to the research community [12, 16]. The optimized M_1R–M_5R homology models, built and evaluated as described above, are freely available as part of the Supporting Information for

Refs. [12, 16]. ZINC-derived decoy sets (Sections 4.2.2 and 4.2.3) are available for users upon request. We consider such open access as crucial in our field since it allows researchers to use these structures, compare them to their results [34, 73], and thus advance the development of better modeling methods.

7 Conclusions

This chapter describes development of homology models of the muscarinic acetylcholine receptors M_1R–M_5R and their evaluation through retrospective virtual screening for the identification of antagonists. Model refinement, guided by experimental knowledge of active compounds and critical binding site residues, results in ligand-induced adaptation of the receptor binding sites and their optimization for antagonist recognition. Specifically, the generated homology models are capable of distinguishing known antagonists from matched decoy compounds. The confirmed predictive power of these models gives greater confidence in the use of these models for prospective virtual screening. The following aspects of the modeling procedure are particularly important: (1) binding site optimization is a critical step in model generation; (2) knowledge-based homology models of GPCRs are appropriate for prospective virtual screening, once confirmed in retrospective tests; and (3) property-matched decoys should be used in virtual screening evaluation of homology models. Future work is required to evaluate homology models in a flexible receptor scenario: by on-the-fly receptor flexibility [14, 15], molecular dynamics [74], or using receptor ensembles [75–78].

8 Notes

1. For specific version numbers of programs and modules, readers should refer to original publications. For example, for homology modeling of the five muscarinic M_1–M_5 acetylcholine receptors we employed Prime (versions 3.0 and 3.1) using the Maestro interface (versions 9.2 and 9.3). Ligand molecules were prepared with LigPrep (version 2.5) and docked into the homology models using Glide (versions 5.7 and 5.8).
2. In the Ballesteros-Weinstein residue numbering system, the first number corresponds to the helix number and the second number represents the position relative to the most conserved residue in that helix (assigned the arbitrary number "50"). This nomenclature is not used for the variable loop regions, where receptor sequence numbering is used, usually following the crystal structure (PDB) numbering.

3. Trp 6.48 is a key residue of the aromatic cluster of transmembrane helices 5 and 6. It has been suggested that it acts as a "micro-switch" for receptor activation and inactivation [61]. The IFD protocol consistently caused Trp 6.48 to undergo a conformational "flip" during the Prime step [12], which forced the bulky indole side chain down and away from the binding pocket.
4. During the IFD optimization of the binding sites, monitoring the distance (ndist) between the ionizable or quaternary nitrogen of the ligand and the closest carboxylate oxygen of the conserved Asp 3.32 residue is advisable. The Asp 3.32 residue crucial for ligand binding of all aminergic GPCRs has been determined by site-directed mutagenesis [62]. The distance ndist is a quantitative measure of this important ionic interaction and receptors with ndist >3.0 Å should be excluded from further analysis.
5. In the new IFD protocol (2013-1 release: label "IFD"; 2013-3 release: label "Extended Sampling"), the re-optimized IFDScore is
   $$\text{IFDScore} = 1.0 \times \text{Prime\_Energy} + 9.057 \times \text{GlideScore} + 1.428 \times \text{Glide\_Ecoul}$$
   Source: http://www.schrodinger.com/kb/307.
6. Actives used in model evaluation via retrospective virtual screening: atropine, benzquinamide, benztropine, biperiden, buclizine, carbinoxamine, chlorpromazine, chlorprothixene, clidinium, clozapine, cyclizine, cyclopentolate, cycrimine, desipramine, dicyclomine, diphenidol, dosulepin, doxepin, doxylamine, ethopropazine, flavoxate, glycopyrrolate, homatropine methyl bromide, hyoscyamine, methantheline, methotrimeprazine, metixene, metoclopramide, olanzapine, orphenadrine, oxybutynin, oxyphencyclimine, oxyphenonium, pirenzepine, procyclidine, promazine, promethazine, propantheline, propiomazine, quinacrine, scopolamine, solifenacin, thiethylperazine, tolterodine, tridihexethyl, trifluprom azine, trihexyphenidyl, and trospium. Where atropine was used as the induced fit ligand, it was excluded from the virtual screen, and likewise for clozapine.
7. To generate the ZINC-based decoy library, molecules were required to fall within a similar normal distribution as the active compounds (MW = 265–434 g/mol; mean 322 g/mol; standard deviation 40 g/mol). They were also required to contain an ionizable nitrogen and not to contain more than three hydrogen bond donors or four hydrogen bond acceptors. Finally, to ensure topological diversity, each decoy was required to have a Tanimoto score <0.8 with respect to all other molecules within the set.

8. The docking site is centered as per Section 3.5.1. If using Glide, both the Standard Precision (SP) and the Extra Precision (XP) scoring functions could be used. XP gives marginally better results [12].
9. Calculating the AUC can be done during, or separately from, the curve plotting process. In general the *x*, *y*-coordinates are used to construct a series of geometric shapes that fit the curve and sum their areas. For an ROC curve the geometric shapes will always be rectangles but if one wants to calculate the area under an enrichment curve (note that this is generally less meaningful or comparable as the ideal area can change with the size of the library) then trapezoids will be necessary.

Acknowledgments

T.T. is a recipient of an Australian Postgraduate Award (APA) scholarship. 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. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Crystal structure of rhodopsin: a G protein-coupled receptor. Science 289(5480):739–745
2. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318(5854):1258–1265
3. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature 450(7168):383–387
4. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC, Kobilka BK (2007) GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. Science 318(5854):1266–1273
5. Yang J, Zhang Y (2014) GPCRS-EXP: a database for experimentally solved GPCR structures. http://zhanglab.ccmb.med.umich.edu/GPCR-EXP/. Accessed 3 Dec 2014
6. Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol 63(6):1256–1272. doi:10.1124/mol.63.6.1256
7. Shoichet BK, Kobilka BK (2012) Structure-based drug screening for G-protein-coupled receptors. Trends Pharmacol Sci 33(5):268–272. doi:10.1016/j.tips.2012.03.007
8. Stevens RC, Cherezov V, Katritch V, Abagyan R, Kuhn P, Rosen H, Wuthrich K (2013) The GPCR Network: a large-scale collaboration to determine human GPCR structure and function. Nat Rev Drug Discov 12(1):25–34. doi:10.1038/nrd3859
9. Mason JS, Bortolato A, Congreve M, Marshall FH (2012) New insights from structural biology into the druggability of G protein-coupled receptors. Trends Pharmacol Sci 33(5):249–260. doi:10.1016/j.tips.2012.02.005
10. Granier S, Kobilka B (2012) A new era of GPCR structural and chemical biology. Nat Chem Biol 8(8):670–673. doi:10.1038/nchembio.1025
11. Kooistra AJ, Roumen L, Leurs R, de Esch IJ, de Graaf C (2013) From heptahelical bundle to hits from the haystack: structure-based virtual screening for GPCR ligands. Methods Enzymol 522:279–336. doi:10.1016/B978-0-12-407865-9.00015-7

12. Thomas T, McLean KC, McRobb FM, Manallack DT, Chalmers DK, Yuriev E (2014) Homology modeling of human muscarinic acetylcholine receptors. J Chem Inf Model 54(1):243–253. doi:10.1021/ci400502u
13. Costanzi S (2013) Modeling G protein-coupled receptors and their interactions with ligands. Curr Opin Struct Biol 23(2):185–190. doi:10.1016/j.sbi.2013.01.008
14. Yuriev E, Agostino M, Ramsland PA (2011) Challenges and advances in computational docking: 2009 in review. J Mol Recognit 24:149–164
15. Yuriev E, Ramsland PA (2013) Latest developments in molecular docking: 2010–2011 in review. J Mol Recognit 26(5):215–239. doi:10.1002/jmr.2266
16. McRobb FM, Capuano B, Crosby IT, Chalmers D, Yuriev E (2010) Homology modeling and docking evaluation of aminergic G protein-coupled receptors. J Chem Inf Model 50:626–637
17. Wess J, Eglen RM, Gautam D (2007) Muscarinic acetylcholine receptors: mutant mice provide new insights for drug development. Nat Rev Drug Discov 6(9):721–733. doi:10.1038/nrd2379
18. Kruse AC, Weiss DR, Rossi M, Hu J, Hu K, Eitel K, Gmeiner P, Wess J, Kobilka BK, Shoichet BK (2013) Muscarinic receptors as model targets and antitargets for structure-based ligand discovery. Mol Pharmacol 84(4):528–540. doi:10.1124/mol.113.087551
19. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R (2006) Novel procedure for modeling ligand/receptor induced fit effects. J Med Chem 49(2):534–553
20. Gregory KJ, Hall NE, Tobin AB, Sexton PM, Christopoulos A (2010) Identification of orthosteric and allosteric site mutations in M2 muscarinic acetylcholine receptors that contribute to ligand-selective signaling bias. J Biol Chem 285(10):7459–7474
21. Valant C, Gregory KJ, Hall NE, Scammells PJ, Lew MJ, Sexton PM, Christopoulos A (2008) 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 283(43):29312–29321. doi:10.1074/jbc.M803801200
22. Marquer C, Fruchart-Gaillard C, Letellier G, Marcon E, Mourier G, Zinn-Justin S, Menez A, Servent D, Gilquin B (2011) 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 286(36):31661–31675. doi:10.1074/jbc.M111.261404
23. Martinez-Archundia M, Cordomi A, Garriga P, Perez JJ (2012) Molecular modeling of the M3 acetylcholine muscarinic receptor and its binding site. J Biomed Biotechnol 2012:789741. doi:10.1155/2012/789741
24. Huang X, Zheng G, Zhan CG (2012) Microscopic binding of M5 muscarinic acetylcholine receptor with antagonists by homology modeling, molecular docking, and molecular dynamics simulation. J Phys Chem B 116(1):532–541. doi:10.1021/jp210579b
25. Zheng G, Smith AM, Huang X, Subramanian KL, Siripurapu KB, Deaciuc A, Zhan CG, Dwoskin LP (2013) Structural modifications to tetrahydropyridine-3-carboxylate esters en route to the discovery of M5-preferring muscarinic receptor orthosteric antagonists. J Med Chem 56(4):1693–1703. doi:10.1021/jm301774u
26. Jakubik J, Randakova A, Dolezal V (2013) On homology modeling of the M2 muscarinic acetylcholine receptor subtype. J Comput Aided Mol Des 27(6):525–538. doi:10.1007/s10822-013-9660-8
27. Blaney FE, Raveglia LF, Artico M, Cavagnera S, Dartois C, Farina C, Grugni M, Gagliardi S, Luttmann MA, Martinelli M, Nadler GM, Parini C, Petrillo P, Sarau HM, Scheideler MA, Hay DW, Giardina GA (2001) Stepwise modulation of neurokinin-3 and neurokinin-2 receptor affinity and selectivity in quinoline tachykinin receptor antagonists. J Med Chem 44(11):1675–1689
28. Lebon G, Langmead CJ, Tehan BG, Hulme EC (2009) Mutagenic mapping suggests a novel binding mode for selective agonists of M1 muscarinic acetylcholine receptors. Mol Pharmacol 75(2):331–341. doi:10.1124/mol.108.050963
29. Avlani VA, Langmead CJ, Guida E, Wood MD, Tehan BG, Herdon HJ, Watson JM, Sexton PM, Christopoulos A (2010) Orthosteric and allosteric modes of interaction of novel selective agonists of the M1 muscarinic acetylcholine receptor. Mol Pharmacol 78(1):94–104. doi:10.1124/mol.110.064345
30. Kaye RG, Saldanha JW, Lu ZL, Hulme EC (2011) Helix 8 of the M1 muscarinic acetylcholine receptor: scanning mutagenesis delineates a G protein recognition site. Mol Pharmacol 79(4):701–709. doi:10.1124/mol.110.070177
31. Xu J, Chen H (2012) Interpreting the structural mechanism of action for MT7 and human muscarinic acetylcholine receptor 1 complex by modeling protein-protein interaction. J Biomol Struct Dyn 30(1):30–44. doi:10.1080/07391102.2012.674188
32. Daval SB, Valant C, Bonnet D, Kellenberger E, Hibert M, Galzi JL, Ilien B (2012) Fluorescent derivatives of AC-42 to probe

bitopic orthosteric/allosteric binding mechanisms on muscarinic M1 receptors. J Med Chem 55(5):2125–2143. doi:10.1021/jm201348t

33. Daval SB, Kellenberger E, Bonnet D, Utard V, Galzi JL, Ilien B (2013) Exploration of the orthosteric/allosteric interface in human M1 muscarinic receptors by bitopic fluorescent ligands. Mol Pharmacol 84:71–85. doi:10.1124/mol.113.085670
34. Jójárt B, Balint AM, Balint S, Viskolcz B (2012) Homology modeling and validation of the human M1 muscarinic acetylcholine receptor. Mol Inf 31(9):635–638. doi:10.1002/minf.201200062
35. Jacobson MA, Kreatsoulas C, Pascarella DM, O'Brien JA, Sur C (2010) The M1 muscarinic receptor allosteric agonists AC-42 and 1-[1′-(2-methylbenzyl)-1,4′-bipiperidin-4-yl]-1,3-dihydro-2H-benzimidazol-2-one bind to a unique site distinct from the acetylcholine orthosteric site. Mol Pharmacol 78(4):648–657. doi:10.1124/mol.110.065771
36. Ma L, Seager MA, Wittmann M, Jacobson M, Bickel D, Burno M, Jones K, Graufelds VK, 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 CW, Shipe W, Kuduk S, Sur C, Kinney G, Seabrook GR, Ray WJ (2009) Selective activation of the M1 muscarinic acetylcholine receptor achieved by allosteric potentiation. Proc Natl Acad Sci U S A 106(37):15950–15955. doi:10.1073/pnas.0900903106
37. Chin SP, Buckle MJC, Chalmers DK, Yuriev E, Doughty SW (2014) Towards activated homology models of the human M_1 muscarinic acetylcholine receptor. J Mol Graph Model 49:91–98
38. McMillin SM, Heusel M, Liu T, Costanzi S, Wess J (2011) Structural basis of M3 muscarinic receptor dimer/oligomer formation. J Biol Chem 286(32):28584–28598. doi:10.1074/jbc.M111.259788
39. Espinoza-Fonseca LM, Pedretti A, Vistoli G (2008) Structure and dynamics of the full-length M1 muscarinic acetylcholine receptor studied by molecular dynamics simulations. Arch Biochem Biophys 469(1):142–150. doi:10.1016/j.abb.2007.09.002
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 (2012). New York, NY
41. Jacobson MP, Pincus DL, Rapp CS, Day TJ, Honig B, Shaw DE, Friesner RA (2004) A hierarchical approach to all-atom protein loop prediction. Proteins 55(2):351–367
42. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22(22):4673–4680
43. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des 27(3):221–234. doi:10.1007/s10822-013-9644-8
44. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47(7):1739–1749
45. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem 47(7):1750–1759
46. Chalmers DK, Roberts BP (2011) Silico—a Perl Molecular Modelling Toolkit, Monash University: Melbourne
47. Mobarec JC, Sanchez R, Filizola M (2009) Modern homology modeling of G-protein coupled receptors: which structural template to use? J Med Chem 52(16):5207–5216. doi:10.1021/jm9005252
48. Kolaczkowski M, Bucki A, Feder M, Pawlowski M (2013) Ligand-optimized homology models of D_1 and D_2 dopamine receptors: application for virtual screening. J Chem Inf Model 53:638–648. doi:10.1021/ci300413h
49. Pala D, Beuming T, Sherman W, Lodola A, Rivara S, Mor M (2013) Structure-based virtual screening of MT_2 melatonin receptor: influence of template choice and structural refinement. J Chem Inf Model 53(4):821–835. doi:10.1021/ci4000147
50. Rataj K, Witek J, Mordalski S, Kosciolek T, Bojarski AJ (2014) Impact of template choice on homology model efficiency in virtual screening. J Chem Inf Model 54(6):1661–1668. doi:10.1021/ci500001f
51. Tang H, Wang XS, Hsieh JH, Tropsha A (2012) Do crystal structures obviate the need for theoretical models of GPCRs for structure-based virtual screening? Proteins 80(6):1503–1521. doi:10.1002/prot.24035
52. Beuming T, Sherman W (2012) Current assessment of docking into GPCR crystal structures and homology models: successes,

challenges, and guidelines. J Chem Inf Model 52(12):3263–3277. doi:10.1021/ci300411b

53. Phatak SS, Gatica EA, Cavasotto CN (2010) Ligand-steered modeling and docking: a benchmarking study in class a g-protein-coupled receptors. J Chem Inf Model 50(12): 2119–2128
54. Neves MA, Simoes S, Sáe Melo ML (2010) Ligand-guided optimization of CXCR4 homology models for virtual screening using a multiple chemotype approach. J Comput Aided Mol Des 24(12):1023–1033
55. Katritch V, Kufareva I, Abagyan R (2011) Structure based prediction of subtype-selectivity for adenosine receptor antagonists. Neuropharmacology 60(1):108–115. doi:10.1016/j.neuropharm.2010.07.009
56. Kufareva I, Rueda M, Katritch V, Stevens RC, Abagyan R (2011) Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. Structure 19(8): 1108–1126. doi:10.1016/j.str.2011.05.012
57. Michino M, Abola E, Brooks CL III, Dixon JS, Moult J, Stevens RC (2009) Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. Nat Rev Drug Discov 8(6):455–463
58. Kufareva I, Katritch V, Participants of GPCR Dock 2013, Stevens RC, Abagyan R (2014) Advances in GPCR modeling evaluated by the GPCR Dock 2013 assessment: meeting new challenges. Structure 22(8):1120–1139. doi:10.1016/j.str.2014.06.012
59. Ballesteros JA, Weinstein H, Stuart CS (1995) Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. Methods Neurosci 25:366–428
60. Bymaster FP, Felder CC, Tzavara E, Nomikos GG, Calligaro DO, McKinzie DL (2003) Muscarinic mechanisms of antipsychotic atypicality. Prog Neuro Psychopharmacol Biol Psychiatry 27(7):1125–1143. doi:10.1016/j.pnpbp.2003.09.008
61. Holst B, Nygaard R, Valentin-Hansen L, Bach A, Engelstoft MS, Petersen PS, Frimurer TM, Schwartz TW (2010) A conserved aromatic lock for the tryptophan rotameric switch in TM-VI of seven-transmembrane receptors. J Biol Chem 285(6):3973–3985. doi:10.1074/jbc.M109.064725
62. Spalding TA, Birdsall NJ, Curtis CA, Hulme EC (1994) Acetylcholine mustard labels the binding site aspartate in muscarinic acetylcholine receptors. J Biol Chem 269(6):4092–4097
63. Anighoro A, Rastelli G (2013) Enrichment factor analyses on G-protein coupled receptors with known crystal structure. J Chem Inf Model 53(4):739–743. doi:10.1021/ci4000745
64. Gatica EA, Cavasotto CN (2012) Ligand and decoy sets for docking to G protein-coupled receptors. J Chem Inf Model 52(1):1–6. doi:10.1021/Ci200412p
65. Okuno Y, Tamon A, Yabuuchi H, Niijima S, Minowa Y, Tonomura K, Kunimoto R, Feng C (2008) GLIDA: GPCR—ligand database for chemical genomics drug discovery—database and tools update. Nucleic Acids Res 36(Suppl 1):D907–D912. doi:10.1093/nar/gkm948
66. Huang N, Shoichet BK, Irwin JJ (2006) Benchmarking sets for molecular docking. J Med Chem 49(23):6789–6801
67. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG (2012) ZINC: a free tool to discover chemistry for biology. J Chem Inf Model 52:1757–1768. doi:10.1021/ci3001277
68. Shi L, Javitch JA (2002) The binding site of aminergic G protein-coupled receptors: the transmembrane segments and second extracellular loop. Annu Rev Pharmacol Toxicol 42:437–467. doi:10.1146/annurev.pharmtox.42.091101.144224
69. Hawkins PCD, Warren GL, Skillman AG, Nicholls A (2008) How to do an evaluation: pitfalls and traps. J Comput Aided Mol Des 22(3–4):179–190
70. Nicholls A (2008) What do we know and when do we know it? J Comput Aided Mol Des 22(3–4):239–255
71. Mysinger MM, Shoichet BK (2010) Rapid context-dependent ligand desolvation in molecular docking. J Chem Inf Model 50(9): 1561–1573
72. Katritch V, Rueda M, Lam PC, Yeager M, Abagyan R (2010) GPCR 3D homology models for ligand screening: lessons learned from blind predictions of adenosine A2a receptor complex. Proteins 78(1):197–211
73. Lin X, Huang XP, Chen G, Whaley R, Peng S, Wang Y, Zhang G, Wang SX, Wang S, Roth BL, Huang N (2012) Life beyond kinases: structure-based discovery of sorafenib as nanomolar antagonist of 5-HT receptors. J Med Chem 55(12):5749–5759. doi:10.1021/jm300338m
74. Miao Y, Nichols SE, Gasper PM, Metzger VT, McCammon JA (2013) Activation and dynamic network of the M2 muscarinic receptor. Proc Natl Acad Sci U S A 110(27):10982–10987. doi:10.1073/pnas.1309755110
75. Bottegoni G, Rocchia W, Rueda M, Abagyan R, Cavalli A (2011) Systematic exploitation of multiple receptor conformations for virtual ligand screening. PLoS One 6(5), e18845. doi:10.1371/journal.pone.0018845

76. Korb O, Olsson TS, Bowden SJ, Hall RJ, Verdonk ML, Liebeschuetz JW, Cole JC (2012) Potential and limitations of ensemble docking. J Chem Inf Model 52(5):1262–1274. doi:10.1021/ci2005934
77. Rueda M, Totrov M, Abagyan R (2012) ALiBERO: evolving a team of complementary pocket conformations rather than a single leader. J Chem Inf Model 52(10):2705–2714. doi:10.1021/ci3001088
78. Xu M, Lill MA (2012) Utilizing experimental data for reducing ensemble size in flexible-protein docking. J Chem Inf Model 52(1):187–198. doi:10.1021/ci200428t