



Universiteit
Leiden
The Netherlands

Clavis Aurea? Structure-enabled approaches of identifying and optimizing GPCR ligands

Lenselink, E.B.

Citation

Lenselink, E. B. (2017, June 7). *Clavis Aurea? Structure-enabled approaches of identifying and optimizing GPCR ligands*. Retrieved from <https://hdl.handle.net/1887/49402>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/49402>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/49402> holds various files of this Leiden University dissertation.

Author: Lenselink, E.B.

Title: Clavis Aurea? Structure-enabled approaches of identifying and optimizing GPCR ligands

Issue Date: 2017-06-07

Chapter 2

Docking and virtual screening strategies for GPCR drug discovery



Adapted from:

"Docking and virtual screening strategies for GPCR drug discovery"

Thijs Beuming, Eelke B. LenseLink, Daniele Pala, Fiona McRobb, Matt Repasky, and Woody Sherman.

G Protein-Coupled Receptors in Drug Discovery: Methods and Protocols, Volume 1335 of the series Methods in Molecular Biology (2015): pages 251-276.

Chapter 2

Docking and virtual screening strategies for GPCR drug discovery

Contents

About this chapter	page 30
Introduction	page 30
Methods	page 31
General docking methodology	page 31
Accuracy of docking in GPCRs: pose prediction	page 33
Virtual Screening against GPCRs	page 36
Incorporating experimental information into docking	page 39
Combining docking with additional computational tools	page 41
Treatment of water molecules	page 42
Data fusion	page 44
Optimizing structures and models for docking	page 45
Screening for agonists and inverse agonists	page 47
A workflow for modeling of GPCR-ligand interactions	page 49
Conclusions and Future Developments	page 51
References	page 52

About this chapter

As briefly touched upon in **chapter 1**, progress in structure determination of G protein-coupled receptors (GPCRs) has made it possible to apply structure-based drug design (SBDD) methods to this pharmaceutically important target class. The quality of GPCR structures available for SBDD projects fall on a spectrum ranging from high resolution crystal structures (<2 Å), where all water molecules in the binding pocket are resolved, to lower resolution (>3 Å) where some protein residues are not resolved, and finally to homology models that are built using distantly related templates. Each GPCR project involves a distinct set of opportunities and challenges, and requires different approaches to model the interaction between the receptor and the ligands. In this review we will discuss docking and virtual screening to GPCRs, and highlight several refinement and post-processing steps that can be used to improve the accuracy of these calculations. Several examples are discussed that illustrate specific steps that can be taken to improve upon the docking and virtual screening accuracy. While GPCRs are a unique target class, many of the methods and strategies outlined in this review are general and therefore applicable to other protein families.

Introduction

The recent increase in the number of solved G protein-coupled receptor (GPCR) crystal structures was enabled by several key experimental breakthroughs, including stabilization of GPCR structures using fusion proteins¹ and site-directed mutagenesis²⁻⁵ coupled with the discovery of ligands that have favorable kinetic properties for crystallization. In addition, novel crystallization techniques, such as use of the lipidic cubic phase and serial femtosecond crystallography,^{6,7} have made it possible to get good diffraction from crystals unsuitable for ordinary structure determination. Since the emergence of the b₂AR structures in 2007¹, a steady stream of novel structures have expanded coverage to many branches of the GPCR phylogenetic tree, with a diverse coverage of the class A GPCRs and more recent structures solved for classes B, C and F⁸ (http://gpcr.scripps.edu/gpcr_targets.htm). Nonetheless, most targets for which a structure has been solved have only a small number of ligands co-crystallized (typically only one), making it necessary to predict the binding mode of other compounds using computational techniques such as docking. In some cases, the ligands of interest will be similar to the co-crystallized ligand, requiring little structural rearrangement (induced fit) of the receptor, whereas in other cases substantial movements of the receptor may need to be taken into account in order to achieve accurate docking results. In addition to the need for accurate docking of

diverse ligands to crystal structures, docking to homology models is also of critical importance — for each solved GPCR structure one or more close homologues exist that can be modeled with relatively high accuracy. Nonetheless, a homology model built based on a closely related template does not ensure that the binding site will be accommodating to the ligands of interest, thus induced fit must again be considered. Finally, there exists a large number of GPCRs for which no closely related homology modeling templates are available, adding to the challenge of docking to these targets. Fortunately, strategies can be employed to improve the accuracy of docking to homology models of more distant targets, although these strategies depend highly on available experimental data and can be quite different to the approaches employed for crystal structures and high-quality models. In this work we will discuss the various approaches to GPCR docking, beginning with a general overview of docking methodologies, as performed using tools in the Schrödinger small-molecule drug discovery suite. For reference, the various tools described throughout the manuscript are summarized in Table 2.1. We will describe their application to high-resolution crystal structures, but also explore more challenging cases involving induced fit and homology models.

Finally, we discuss how docking and complementary techniques can be used to move beyond simple models of binding to explore questions related to activation (agonists, antagonists, etc.) and dynamics.

Methods

General docking methodology

Docking of a small molecule into a protein involves sampling a large number (typically in the millions) of possible conformations and orientations of the molecule in the protein binding site with an objective to identify the ligand poses (conformation and orientation) with the most favorable binding free energies. Once one or more viable poses are found, the binding energy can be estimated using a scoring function that describes the physics of binding using an explicit physics-based approach,^{9,10} an empirical scoring function that relies on some physics,^{11,12} or a pure knowledge-based scoring function based on fitting to experimental data.¹³⁻¹⁵ Docking scoring functions can be used for i) predicting the most likely binding pose of a ligand, and ii) differentiating between ligands that interact with a protein with a favorable binding energy (actives) and those with an unfavorable binding energy (inactives).

Table 2.1 Compendium of tools in the Schrödinger suite used for docking

Category	Name	Function
Data preparation	<i>Protein Preparation Wizard</i>	Assignment of protein and complex bond orders, tautomeric and ionization states.
	<i>LigPrep</i>	Enumeration of ligand protonation and tautomeric states, chirality.
Target preparation and analysis	<i>Prime</i>	Homology modeling, loop modeling and side chain prediction.
	<i>IFD</i>	Binding site adaptation.
Docking	<i>Glide</i>	Rigid receptor docking.
	<i>IFD</i>	Flexible receptor docking.
Scoring	<i>GlideScore</i>	Virtual screening.
	<i>MM/GBSA</i>	Implicit solvation model for accurate scoring of congeneric and diverse compounds.
	<i>WaterMap</i>	Scoring using explicit water, comparing congeneric
	<i>WM/MM</i>	compounds only. Hybrid of MM/GBSA and WaterMap.
Data fusion	<i>Shape</i>	3D-based comparison of ligands.
	<i>Canvas</i>	2D-fingerprint determination, similarity comparison and
		Clustering.
Experimental data	<i>Prime</i>	Local protein structure optimization.
	<i>Structural interaction fingerprints</i>	Calculating similarity and clustering of binding modes
Auxiliary tools	<i>MacroModel</i>	MM-based ligand conformations and strain.
	<i>Jaguar</i>	QM-based ligand conformations and strain, protonation states

Different scoring functions are appropriate for these two fundamental tasks. While the ligand is almost always treated flexibly, as described below, the protein is often treated rigidly or with minimal flexibility (i.e. a few key side chains). The rigid receptor treatment greatly reduces the conformational search space of the system, thereby speeding the calculations considerably and making the problem more tractable, but neglects potential induced-fit effects.¹⁶ However, docking algorithms that account for protein flexibility exist and will be discussed later in this section. The docking program Glide¹⁷⁻¹⁹ is amongst the most widely used and best performing docking algorithms.^{20,21} The underlying methodology has been described in detail elsewhere, but in short, it involves initial sampling and filtering of millions of possible poses using crude geometric and energetic criteria, and the evaluation of much smaller sets of poses using scoring functions of increasing complexity and accuracy. The protein is kept fixed during the entire protocol, and only ligand orientational and torsional space is sampled. Glide uses the so-called “Emodel” scoring function to compare poses for the same ligand, while the GlideScore is used to compare different compounds based on a rough predicted affinity. The Standard Precision (SP) GlideScore has been optimized using virtual screening data, and is also used in the High Throughput Virtual Screening (HTVS) mode of Glide, while the Extra Precision (XP) GlideScore includes a number of penalties and rewards that attempt to capture specific protein-ligand recognition motifs. The inclusion of penalty terms in the XP scoring function, which is important for eliminating false positives (inactives), requires additional sampling, thereby increasing the computational costs. While the HTVS mode typically takes 1-2 seconds per ligand and SP takes 10-20 seconds per ligand, XP takes 5-10 minutes per ligand, with the range depending primarily on the number of rotatable bonds in the ligand.

When there are protein conformational changes associated with ligand binding (side-chain rotamer changes and backbone relaxation effects), protein flexibility can be taken into account using an approach such as Induced Fit Docking (IFD).²² This protocol consists of an initial docking stage with a softened van der Waals potential, followed by receptor optimization of an ensemble of initial ligand/protein complexes using the Prime protein structure prediction program,²³ and finally another round of docking using standard Glide settings into these optimized protein binding sites. Select side chains can also be removed in the initial docking stage to allow for greater exploration of poses within the receptor binding site. The method has been shown to correctly model induced-fit effects in a large number of cases where rigid receptor docking fails,²² and has also been applied to other areas of research, such as modeling off-target binding²⁴ and generating alternate protein structures for virtual screening.²⁵ The IFD approach is significantly

slower than rigid receptor Glide docking, and is typically used for pose prediction and generating induced fit receptor structures, not for high-throughput applications such as virtual screening.

Accurate prediction of ligand binding modes requires accurate representation of both the ligand and the protein. For the former, LigPrep²⁶ is used to generate initial 3D coordinates and enumerate the possible stereoisomers of a ligand. Epik is then used to predict accessible tautomers and ionization states.^{27,28} Importantly, accessibility of the possible tautomeric and ionization states are captured in a so-called “state penalty”, which is used to augment the raw GlideScore to enhance the performance of the algorithm in virtual screening.²⁹ Larger state penalties mean that the state is less populated in solution and therefore must pay an energetic penalty upon binding to account for the reduced apparent concentration of that state of the ligand. The state penalty is computed in units of kcal/mol and therefore can be added directly to the GlideScore.

Preparation of the protein structure is handled using the Protein Preparation Wizard in Maestro.^{29,30} A crucial step of this preparation is the assignment of rotameric states of residue with polar hydrogen atoms, such as Ser, Thr, Tyr, Asn, and Gln residues, as well as those of water molecules, to optimize the hydrogen bonding network in the protein structure. In addition, tautomeric/ionization states of His residues are evaluated using the same criteria, and the protonation state of ionizable residues is predicted using the PROPKA algorithm.³¹ Extensive benchmarking studies have shown the importance of adequate preparation of the protein structure for accurate virtual screening results.^{21,29}

Accuracy of pose prediction is typically measured in terms of pose RMSD compared to a crystal structure. However, other metrics exist to assess pose accuracy and each has its own merits. For example, the Real-Space R (RSR) metric takes into account electron density and weights inaccuracies in the ligand pose more heavily in regions where the electron density is stronger.³² Another approach, Generally Applicable Replacement for rmsD (GARD), computes a score for the accuracy of a pose using weights for each atom determined from their relative importance to binding.³³ Torsion Fingerprint Deviation (TFD) is another metric to compare ligand conformations, which compares torsional fingerprints between a query molecule and a set of conformations, and has the advantage of being scaled by the number of torsions in a molecule and can be used to compare conformational ensembles rather than just pairs of conformations.³⁴ Recent GPCR docking benchmarks³⁵⁻³⁷ have also tried to assess docking accuracy using other parameters such as ligand/receptor contact maps.

The ability of docking algorithms to distinguish between active compounds (binding typical affinity $< 10 \mu\text{M}$) and inactive compounds (typical affinity $> 100 \mu\text{M}$) in a virtual screen is measured in terms of enrichment. Many metrics exist, but the most widely used are EF(n%) (the fraction of actives in the top n% of the data), area under the receiver-operating characteristic curve (AUC), and Boltzmann-enhanced discrimination of receiver-operating characteristic (BEDROC).³⁸ More discussion on this topic is presented in the virtual screening section below. In GPCRs, the situation is further complicated by the need to consider the functional characteristics of a binding small molecule, which can be either a partial or full agonist, an inverse agonist, a neutral antagonist, an allosteric modulator, or a functionally selective (i.e. biased) agonist.³⁹

Accuracy of docking in GPCRs: pose prediction

The ability of docking algorithms to predict the structure of GPCR-ligand complexes has been assessed in several retrospective and prospective benchmarks. In general, Glide is able to reproduce the binding mode of poses in cognate GPCR crystal structures (so called redocking or self docking) within 1.5 \AA in $\sim 80\%$ of cases,^{17,40} which is consistent with the docking accuracy of Glide to other target classes. A more relevant and practical test is cross docking of non-cognate ligands into crystal structures. In our comprehensive benchmark, this was shown to be as accurate as redocking for the b_1 -adrenoceptor ($\beta_1\text{AR}$) and $\beta_2\text{AR}$, but highly problematic for a third target (the adenosine $\text{A}_{2\text{A}}$ receptor ($\text{A}_{2\text{A}}\text{R}$)). In the former case, the limited degree of induced fit observed upon binding and a highly conserved binding mode among all studied ligands probably contributed to the success. The $\text{A}_{2\text{A}}\text{R}$ ligands were much more diverse, and the ligand-receptor interactions are mediated by several non-conserved water molecules (Figure 2.1B), which were excluded from the analysis. The role of including these waters in virtual screening against the $\text{A}_{2\text{A}}\text{R}$ is discussed in detail later in this chapter. The use of receptor ensembles and physics-based post processing has been shown to improve cross-docking accuracy for non-GPCR targets⁴²⁻⁴⁴ and likely would help with GPCRs as well, although demonstrations of this have not been published. In general, GPCR structures have shown a moderate degree of induced fit upon binding, suggesting that Glide or IFD are suitable tools for docking diverse ligands in a majority of cases. Nonetheless, more recently published structures have shown dramatic conformational changes for different ligands, and in those cases docking will likely fail (Figure 2.1C).

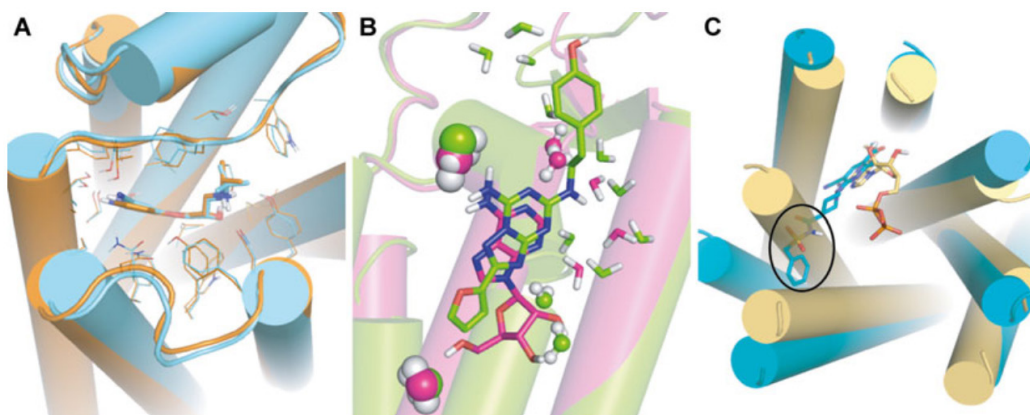


Figure 2.1 Structural comparisons between related GPCR structures and their impact on docking. A) Superimposed β -adrenoceptors β_1 AR bound to pindolol (orange, PDB ID: 2VT4) and β_2 AR bound to carazolol (cyan, PDB ID: 2RH1) illustrate the case where structures are sufficiently similar that cross docking within the same target class is relatively straightforward, and docking into homology models based on highly similar templates is also possible.⁴⁰ B) Docking into A_{2A} R structures is complicated due to the presence of non-conserved water molecules. Superimposed structures of A_{2A} R bound to inverse agonist ZM241385 (green, PDB ID: 4E1Y) and the endogenous agonist adenosine (magenta, PDB ID: 2YDO) show clusters of conserved (large CPK on the left) and non-conserved (small ball and stick on the right) waters interacting with ligands and/or receptor. A protocol to determine which waters to include in docking has been described in detail.⁴¹ C) Superimposed structures of the $P2Y_{12}$ receptor bound to an agonist 2MeSATP (yellow, PDB ID: 4PXZ) and inverse agonist AZD1283 (cyan, PDB ID: 4NTJ) show a major induced fit effect, involving a large rearrangement of a TM domain, that cannot be addressed with conventional docking techniques.

Docking into homology models is an even more difficult challenge, but represents the most relevant scenario for drug discovery projects, since most GPCRs still lack crystal structures. Methods for generating GPCR homology models have been reviewed in this series previously.⁴⁵ The GPCR Dock assessments³⁵⁻³⁷ are prospective homology modeling/docking exercises on a small number of targets, including the A_{2A} R, dopamine D_3 receptor (D_3 R), serotonin 5-HT_{1B} receptor (5-HT_{1B}R) and serotonin 5-HT_{2B} receptor (5-HT_{2B}R), chemokine receptor CXCR4 (CXCR4), and the smoothened (SMO) receptor. Success in predicting poses for these targets has been dependent on the availability of similar templates, i.e. largely successful for D_3 R and 5-HT receptors, while more difficult for the A_{2A} R and especially CXCR4 and the SMO receptor. Dependence of pose prediction on the availability of similar templates was also identified in a study where we generated a matrix of homology models for all available GPCR structures and tested the ability of Glide and IFD to predict binding modes of ligands.

Highly similar templates such as β_1 AR/ β_2 AR and muscarinic M_2 and M_3 receptors (M_2 R and M_3 R, respectively) generated highly accurate homology models. RMSDs for dock-

ing ligands into these models are, on average, equal to those redocked into the crystal structures. Sporadic success was seen docking into models from template/target pairs with intermediate similarity such as the histamine H_1 receptor (H_1R) models built using other aminergic GPCR templates. However, for the majority of models built, blind application of the docking algorithm did not produce viable poses. In general, this analysis suggested a sequence identity threshold of 30% or higher for a homology model to have a reasonable chance of producing accurate poses. Results from the GPCR Dock assessments have suggested a threshold of 35-40%,³⁶ while a more recent publication using the Rosetta suite of programs suggested a threshold of 50%.⁴⁶ These average estimates can be misleading, as each case is different and all it can take is a single side chain being in the wrong position for a homology model to be essentially useless for docking. Nevertheless, accurate pose prediction using models based on distant templates has been possible by incorporating information from experiments (e.g., mutagenesis⁴⁷) or other computational approaches (e.g., pharmacophore modeling⁴⁸). Indeed, the literature contains several examples, discussed below, of successful docking studies using distant templates where the use of a non-automated protocol with additional constraints was sufficient to obtain poses with reasonable similarity to the crystal structure.

Virtual Screening against GPCRs

A primary objective in virtual screening is to separate active from inactive compounds (i.e. “enrich” a database). There are several metrics to assess the performance of virtual screening tools in retrospective validation studies.⁴⁹⁻⁵¹ The metrics primarily focus on the rank of active compounds relative to inactive compounds; however, the most relevant metric depends on the objectives of the study. In most cases of pharmaceutical interest, the objective is to reduce a screening library from millions of compounds to a manageable number that can be run through a medium throughput assay (typically hundreds or thousands). As such, equally weighting the full receiver-operating characteristic (ROC) curve, as is done when considering the area under the curve (AUC), has little relationship to the goal of the virtual screen (i.e. to get a large number of actives in a tiny fraction of assayed database molecules). For example, a method that finds 50% of the active compounds at the top of the list but misses the other 50% of the compounds (perhaps because there are multiple receptor conformations that are not being considered) would get an AUC of 0.5. This is the same AUC as one gets from a random distribution, or worse, for not finding any active compounds until 50% of the database has been screened, at which point all of the active compounds are found (AUC = 0.5). As such, we focus primarily on early enrichment metrics, such as BEDROC³⁸

with a high value of α (typically 160.9, which corresponds to 80% of the score coming from the top 1% of the list) or enrichment in the top $n\%$, such as EF(1%). Focusing in even earlier parts of the database screen (such as EF(0.1%) or EF(0.01%)) requires larger decoy databases than are typically used in retrospective virtual screening studies.

Since the structure of β_2 AR receptor appeared in 2007¹, many examples of prospective virtual screening using these structures have been published including $A_{2A}R$ ⁵²⁻⁵⁴, CXCR4,⁵⁵ M_2R and M_3R ,⁵⁶ H_1R ,⁵⁷ κ opioid receptor (KOR),⁵⁸ D_3R ,⁵⁹⁻⁶¹ and β_2AR .⁶² In the absence of crystal structures, accurate homology models have proven useful for virtual screening.^{59,60,63-65} These studies have been reviewed in detail elsewhere.^{66,67} The number of hits varies dramatically, depending on the choice of target, docking algorithm, docking post-processing, screening library, and number of tested compounds. Nonetheless, some general conclusions can be drawn about the accuracy of these experiments. The success rates for prospective screens on X-ray structures have averaged around 40%, while for homology models the average value drops to around 15%, which is still significantly higher than one would obtain from an unbiased experimental library screen. This suggests a potentially prominent role for structure-based virtual screening in GPCR lead identification, not only using X-ray structures but also homology models.⁶³

Compound databases used in prospective and retrospective virtual screening calculations strongly differ in terms of size and distribution of physicochemical properties. Indeed, retrospective screening campaigns are often based on compound databases that are one or two orders of magnitude smaller than prospective screens, typically containing 10^4 - 10^5 molecules.^{52,63,68-71} These datasets usually consist of a small fraction of known active compounds combined with a large set of commercially available decoy molecules with physicochemical properties matching those of active derivatives. As a result, the range of properties observed in these databases is strongly influenced by the choice of the initial set of active compounds and is often narrower compared to that observed in libraries used in prospective studies. Since retrospective screening experiments are usually performed to evaluate the ability of a receptor structure to discriminate between active and decoy compounds, striking differences between the physicochemical properties of actives and decoys could favor the identification of the former compounds, leading to artificial enrichments.^{69,71,72} Thus, to ensure a more rigorous evaluation of the discriminating ability of GPCR model structures, a novel compound library has been recently created by combining an initial set of active GPCR ligands with a large dataset of decoy molecules, which were selected to enforce ligand-decoy similarity in terms of physicochemical properties whilst retaining structural dissim-

ilarity.⁷² Indeed, this novel GPCR-tailored dataset proved to be effective in reducing artificial screening performances obtained when using decoy molecules characterized by physicochemical properties significantly different from those of active compounds.

As opposed to retrospective screening experiments, prospective virtual screens often start from libraries of millions of in-house or commercially available compounds (e.g., the “lead-like” or “drug-like” subsets of ZINC or eMolecules)^{54-56,58,59,62,64,73} to ensure an exhaustive exploration of the chemical space and to favor the identification of many structurally diverse novel ligands. These large datasets are often pre-filtered on the basis of the physicochemical properties of known drugs^{74,75} or of lead molecules identified in previous high-throughput screening studies,⁷⁶ with the aim of retaining a manageable subset of ligands endowed with appropriate characteristics for further optimization and development. In some cases, the initial database is further filtered applying stricter criteria, aimed at ensuring the presence of target-specific pharmacophoric features (e.g., a basic nitrogen required to bind to aminergic GPCRs⁵⁷) or at favoring the selection of ligands characterized by specific physicochemical properties (e.g. to favor the crossing of the blood brain barrier⁷⁷). Finally, the screening library may be further reduced in size by selecting compounds with similarity to known binders, either using shape-based, or 2D fingerprint-based comparisons.

Incorporating experimental information into docking

In many instances the direct, application of a docking program with default settings to predict the pose of a ligand to a GPCR will lead to inaccurate pose prediction or an inability to separate active and non-active ligands, especially when using homology models. Combining docking algorithms with experimental information can be helpful to improve accuracy, both in general and when docking to GPCRs. For GPCRs this information typically comes from site-directed mutagenesis experiments, where a residue mutation that affects ligand binding is taken as evidence that there is direct interaction between the residue and the ligand. While such information can be very helpful in reducing the docking search space and improving overall accuracy, mutagenesis data can be misleading as well. For example, a residue determined to be critical for ligand binding might play a role in the dynamics or conformational flexibility of the receptor and might not be in direct contact with the ligand, making the use of such information in docking challenging. Nonetheless, many residues found to be important in mutagenesis studies are in contact with ligands, as has been determined in crystal structures, and thus can be used as powerful constraints in docking.

Glide offers the ability to specify constraints in the docking calculation, such as requiring certain hydrogen bonds to be made, fixing the ligand core position based on another molecule, requiring an atom of certain chemistry to be in a particular position, and defining the coordination complex geometry to metals. Glide also allows for control over atomic van der Waals radii for the ligand and receptor, which can be used to soften the van der Waals potential and simulate minor induced-fit effects. More extensive protein conformational rearrangements can be accounted for with IFD. Additional flexibility can be attained through sampling of protein hydroxyls to improve hydrogen bonding. Docking constraints are an easy way to ensure that poses are consistent with known or suspected binding patterns. In Glide and IFD the different constraint types are defined prior to docking. Alternatively, experimental information can be used to post process large numbers of predicted poses for a given ligand/receptor complex, filtering by specific interaction criteria (i.e. H-bonds, aromatic interactions, etc.) using the “Pose Filter” script in the Maestro interface. Finally, the Structural Interactions Fingerprint method (SIFt) provides an intuitive way to ensure that predicted poses match an expected binding mode.^{57,78,79} SIFt analysis works by representing each protein/ligand interaction as a bit, and concatenates all protein-ligand interactions into a bit string. Computational comparison of bit strings can be done very efficiently, allowing large numbers of poses to be compared and clustered. Practically speaking, it can be used to ensure that predicted poses have similar binding modes to ligands in a crystal structure or a well-validated homology model. From the perspective of the receptor, homology modeling can involve alignment modification to ensure experimentally observed orientation of known interacting residues towards the binding site. Finally, in many cases co-crystallized ligand structures deposited in the Protein Data Bank⁸⁰ show a strong similarity with small molecule crystal structures in the Cambridge Structural Database.⁸¹ Especially in the case of small, rigid ligands, it can be advantageous to only dock the crystallographic conformation of the ligand. When docking larger, more flexible ligands, it can be useful in Glide to define torsional constraints to rigidly treat substructures within molecules to be docked while allowing the remainder of each molecule to be treated flexibly.

Prospective examples where these strategies improve pose predictions exist in the literature. In the GPCR Dock 2008 assessment³⁵ the most accurate solution for the A_{2A}R-ZM241385 model was found by performing IFD calculations using a constraint with residue Asn253 (6.55), which was correctly predicted to form a hydrogen bond with the ligand (residues in parentheses correspond to the Ballesteros-Weinstein numbering system⁸²). The most accurate predictions for the CXCR4 model in the 2010 assessment both

involved a modification of the template-derived model to optimize the orientation of TM2, bringing residue D97 (2.63) into the binding site³⁷. In the same assessment, experimental data indicating which residues were part of the binding site in the D₂R/D₃R family⁸³ were used successfully in the prospective modeling of extracellular loop 2 (EL2) in a D₃R model.

Combining docking with additional computational tools

Docking can often be improved by combining results with other computational approaches. In the case of pose prediction, differentiating between multiple poses with similar scores can be challenging, even when experimental data exists. In that case, additional computationally-derived lines of evidence can help with pose selection. In the case of virtual screening, post processing using additional scoring functions (i.e. MM-GBSA or WaterMap-based scoring⁸⁴) can substantially increase enrichment.⁸⁵ Alternatively, data fusion by combining docking scores with other metrics, including shape-based, and 2D fingerprint based comparisons, can lead to significantly improved results (see below).⁸⁶

The ability of Glide to predict binding modes and separate active and non-active ligands derives in part from an accurate set of force field parameters that describe the ligand internal geometry and its interactions with the protein. The standard version of Glide currently uses an approximate method for determining the energetic penalty that a ligand needs to pay to adopt the bioactive conformation, which includes only torsions and 1,4 van der Waals interactions. More accurate strain energies can be computed with the “Strain Energy Calculation and Re-scoring” script in Maestro. In that protocol, for each ligand pose in the input file, a tightly constrained minimization and an unconstrained minimization are performed with MacroModel.⁸⁷ Resulting strain energies can be used to recalculate Glide docking scores. Alternatively, the determination of ligand strain can be determined at the QM level, using a program such as Jaguar.⁸⁸ This requires the determination of the global minimum energy conformation in solution phase and its associated gas-phase energy.

However, such calculations have inherent problems, since modeling discrete conformations at a temperature of 0 K in implicit solvent can be a poor surrogate for a full free energy calculation in explicit solvent.

Treatment of water molecules

The ability of small molecules to differentially displace and retain specific waters in protein binding sites is a large contributor to their affinity. The WaterMap algorithm^{89,90} computes the locations and thermodynamic properties of water molecules in protein binding sites, and has been used to reveal important aspects of binding sites relating to their ability to bind small molecules, both in globular proteins⁹¹ and GPCRs.⁹² The most typical application of the methodology is to understand structure activity relationships (SAR) for congeneric series where small structural modifications can result in substantial changes in activity. For example, the method was able to describe the activity profile of a series of triazolympurine A_{2A}R ligands, which were otherwise difficult to explain using metrics such as GlideScore or MM-GBSA.⁹³ As the size of the 2-substituted aliphatic group was initially increased to methyl and isopropyl, there was a decrease in potency. However, extending the substituent to n-butyl and n-pentyl results in a significant gain in potency. This trend could not be readily explained by ligand-receptor interactions, steric effects, or differences in ligand desolvation. WaterMap correctly predicts the trend in binding affinity for this series based on the differential water displacement patterns. In brief, small unfavorable substituents occupy a region in the adenosine A_{2A}R binding site predicted to contain stable waters, while the longer favorable substituents extend to a region that contains several unstable waters. The predicted binding energies associated with displacing water within these hydration sites correlates well with the experimental activities. In that case, it was necessary to first predict the orientation of the ligand substituents in the binding site using Glide, and then re-score resulting poses using WaterMap.

In addition to rationalizing SAR, inclusion of explicit water molecules can help with pose prediction. For instance, a relatively large improvement in terms of RMSD has been observed for the A_{2A}R when docking the small molecule ZM241385 with several waters present, compared to redocking in the absence of these waters,⁹⁴ improving the top pose from 10.3 Å to 0.9 Å. However, most crystal structures of GPCRs have relatively poor resolution, and hence the binding site water molecules are usually not resolved. Until now only two GPCR structures with resolution under 2 Å have been published in which the vast majority of water molecules in the binding site are presumably visible.^{95,96} Thus, the prediction of the location and orientation of water molecules is important especially for lower resolution crystal structures, which constitute the vast majority of currently available GPCR structures. An exhaustive retrospective study on the high resolution A_{2A}R structure has shown that the positions of these water molecules can effectively be predicted with the WaterMap algorithm. In that study, 19 out of 22 water molecules present

in the crystal structure were correctly predicted by WaterMap.⁴¹ The comparison of the high resolution structure of A_{2A}R (PDB ID: 4EIY) with an earlier structure at lower resolution (PDB ID: 3EML) shows that the main difference is in the number of resolved water molecules, not the structure of the protein itself, which is largely similar in the high and low resolution structures. This suggests that WaterMap can be used to accurately predict water positions in low-resolution structures as well.

When comparing crystal structures of GPCRs solved with multiple ligands, it is clear that the waters that interact with the ligand and/or the protein are not necessarily conserved. Certain waters interact only with certain scaffolds and not all waters directly contribute to protein-ligand binding. The effect of individual contributions of waters on docking was analyzed in detail for the A_{2A}R.⁴¹ Hundreds of orientations for both predicted and crystal structure waters were generated and the effect of including them in docking was assessed using a set of 299 ligands and 17337 decoys. Based on enrichment, five waters were found to be especially important, and, encouragingly, WaterMap predicted waters performed equally well or better than crystal structure-derived waters. To address the issue of different waters mediating interactions for different scaffolds, an ensemble of structures was selected using a machine learning approach (Decision Trees). This Decision Tree derived a structural ensemble that not only performed better in terms of enrichment but also in terms of the diversity of actives retrieved.

Additionally, WaterMap hydration site analysis can also be used in pose prediction. For example, in the case of the modeling of the D₃R/eticlopride complex two poses were generated using IFD that could not be distinguished based on docking scores or compatibility with experimental data.³⁷ WaterMap was run on the D₃ model and poses were scored by summing the free energy of hydration over waters that were displaced. Two very high-energy hydration sites were detected at the TM5 and TM6 interface. The pose ultimately reported in one of the models, but not the other, had a very favorable WaterMap score due to an effective displacement of these two high-energy waters by the ethyl- and chloro-substituents on the aromatic ring of eticlopride. Comparison of the predicted pose with the crystal structure showed that the prediction of the displacement of these high-energy sites by the ligand was correct.

Finally, WaterMap analysis can be used to determine regions of the binding site where ligand binding is expected to greatly improve ligand affinity. These regions can then be used to define positional constraints to be used in virtual screening.

Data fusion

While the previous sections in this work have focused on structure-based docking for virtual screening, many other virtual screening approaches exist, such as ligand-based pharmacophores,⁹⁷⁻⁹⁹ shape-based screening,^{100,101} fingerprint similarity,^{102,103} and other methods. We do not intend to cover those methods in this work, but it is noteworthy to mention that they can be combined with docking to improve the overall database enrichment. Strategies to combine data from multiple methods, typically called “data fusion” or “consensus scoring”, have been shown to improve the quality of virtual screening methods relative to a single approach.^{86,104-106} Data fusion works by leveraging differences in the individual methods to enhance overall enrichment by having the strengths in one complement weakness in another. For example, a docking method may yield very good scores for certain compounds that fit well into the receptor conformation being used, but might miss compounds similar to known actives that do not fit due to induce-fit effects that are not accounted for in docking. However, a 2D fingerprint or 3D shape screen may identify the active as being very similar to a known active, and therefore score it highly. Conversely, a novel active compound might get a poor fingerprint or shape similarity score but could be scored very well by a docking algorithm because it is able to pick up key interactions, irrespective of the fact that the molecule is not similar to the co-crystallized ligand. The complementarity of the individual methods can lead to reduction in false negatives (active compounds that score poorly) and false positives (inactive compounds that score well). There are many ways to fuse data from multiple methods and the references provided above cover the most widely used approaches. Of course, it is possible for data fusion to produce worse results than an individual method, especially when a poor method is combined with a good method. For data fusion to be successful, one needs to combine methods that on their own have a signal and where the signal from multiple methods complement each other.

Optimizing structures and models for docking

The shape and architecture of the binding site of crude, unrefined GPCR homology models are often inadequate to accommodate known active ligands.^{35-37,40} This might be due to i) differences in the amino acid composition of the binding site between template and model sequences, which can cause binding site collapse if one substitutes large side chains with smaller side chains; ii) differences in the relative arrangement of the trans-membrane (TM) portions between template and modeled receptor; iii) the geometry of loop segments, which are often based on the template structure and have a major effect

on the shape of the binding site in many GPCRs; and vi) the inclusion of a template ligand in the model building, which could induce unusual rearrangements of the binding site region in the resulting model. Due to these effects, a structural optimization of the binding site of the receptor model is often necessary if the model is going to be used to describe the recognition of active compounds. The Prime package for homology modeling contains a diverse set of tools for optimizing side chains and loops in models. Alternatively, the IFD methodology²² is ideally suited to optimize the arrangements of side chains in the GPCR binding site around one or multiple known binders. In particular, IFD can be used in the early stages of a virtual screening workflow to reshape the binding cavity of the crude model around a potent ligand, yielding dramatic improvements in screening performance. For example, an unrefined 5HT_{2A}R model built on the β_2 AR crystal structure (PDB ID: 2RH1) showed poor enrichment, with EF(2%) and EF(5%) of the ranked database of 2 and 2.8, respectively.⁷⁰ This was probably due to the presence of a co-crystallized ligand within the template structure, which, being less bulkier compared to the 5HT_{2A}R ligands, induced a contraction of the binding region that did not allow for a proper accommodation of larger serotonergic compounds. Therefore, clozapine, a potent 5HT_{2A}R antagonist, was then docked within the crude receptor model applying the IFD protocol with the aim to reshape and enlarge the binding cavity around a rigid and bulky scaffold. The final 5HT_{2A}R-clozapine complex was selected on the basis of the consistency of modeled receptor-ligand interactions with site-directed mutagenesis data and yielded improved with EF(2%) and EF(5%) of 13.3 and 7.7, respectively. A similar improvement in screening performance was observed when comparing a crude M₃R- based homology model of the muscarinic M1 receptor (M1R) with a ligand-refined model structure.⁶⁸ The initial unrefined M1R model showed only a limited performance in virtual screening, with EF(2%) and EF(5%) values of 2.0 and 1.6, respectively. Optimization of the model by docking of acetylcholine using IFD improved with EF(2%) and EF(5%) values to 5.9 and 3.5, respectively.

The recent boost in GPCR structure determination has allowed a systematic evaluation of the effect of template choice and of ligand-induced fit on the screening performance of GPCR homology models. The general strategy is to generate multiple models of a GPCR using different templates, and/or generate a set of model variants by docking multiple ligands into the binding site. Retrospective virtual screening calculations can then be used to select the most appropriate model(s) for prospective application.

For example, homology models of the dopaminergic D₁ receptor (D₁R) and D₂R were built using the β_2 AR as template and different receptor structures were generated by

docking twenty potent dopaminergic compounds within each crude D₁R and D₂R structure using the IFD protocol.⁶⁹ Similar to what has been observed for 5HT_{2A}R and for M₁R, the ligand-refined structures were superior to the unrefined models at identifying active compounds, with EF(2%) values for the best ligand-refined D₁R and D₂R models (22 and 12, respectively) nearly two-times higher than those obtained for the best crude D₁ and D₂ homology models (6.4 and 8.2, respectively). Interestingly, twenty ligand-refined D₂R models were built using the closely related D₃R crystal structure as template to evaluate to what extent a higher sequence homology could improve screening performance. The best performing D₃R-based D₂R model structure showed enrichment values (EF(2%) = 12) close to those obtained from the β_2 AR-derived ligand-refined D₂R models, indicating that a higher degree in sequence similarity does not necessarily lead to an increase in screening enrichments.

Similar conclusions were drawn in a recent study aimed at comparing the screening performances of a crude M₃R-based M₂R homology model with that of a ligand-optimized β_2 AR M₂R model.⁷¹ Indeed, although derived from a close sequence template (96% sequence identity in the binding site), the former model showed a limited performance in virtual screening studies, with EF(2%) and EF(5%) values of 9.6 and 5.9, respectively. Conversely, the latter model, which was built basing on a more phylogenetically distant template and which binding site was reshaped around clozapine by means of IFD, showed a much better performance, with calculated EF(2%) and EF(5%) of 11.7 and 11.4, respectively. A comparative virtual screening procedure was also conducted on the crystal structure of the human M₂R. Intriguingly, the screening performance of the ligand-refined M₂R model was close to that obtained with the M₂R crystal structure (EF(2%) and EF(5%) of 15.9 and 10.9, respectively), confirming the importance of binding site refinement over a sequence-based selection of the template structure to achieve good performances in virtual screening.

We recently performed an exhaustive analysis of the effect of template choice and ligand adaptation on the improvement of virtual screening results for the melatonin MT₂ receptor (MT₂R). The MT₂R has a sequence identity lower than 30% with all crystallized GPCRs, making its structural prediction particularly challenging. Twelve unrefined MT₂R homology models were constructed starting from the crystal structures of different GPCRs.⁴⁸ None of the crude MT₂R models was able to properly accommodate active compounds within the putative binding crevice, yielding EF(2%) values equal to 0 in most cases. To optimize the binding site, twelve representative melatonergic ligands were docked within each unrefined MT₂R model using IFD. However, the default IFD pro-

cedure failed to properly accommodate the melatonergic ligands within the MT₂R, as determined by comparison of the resulting poses with experimental data. Indeed, blind docking of ligands into models based on distantly related templates is not expected to produce viable results, as shown in the pose prediction benchmark described above. To address this issue, a customized IFD protocol was set up, in which the twelve representative compounds were manually positioned within the binding pocket of each crude MT₂R model in poses consistent with mutagenesis data and structure-activity relationships. The resulting complexes were subsequently refined applying the standard IFD protocol. The final ligand-optimized receptor models exhibited significant improvements of VS performances compared to the unrefined structures, with the best performing MT₂R model showing an EF(2%) value of 41.2.

Screening for agonists and inverse agonists

The elucidation of the agonist bound structures for several GPCRs, including β_2 AR,¹⁰⁷ β_1 AR,¹⁰⁸ A_{2A}R,¹⁰⁹ P2Y₁₂ receptor (P2Y₁₂R),¹¹⁰ neurotensin receptor type 1 (NTS1R)¹¹¹ and M₂R¹¹² enables a detailed analysis of the conformational differences between GPCRs bound to various functional classes of ligands, including inverse agonists, partial and full agonists, neutral antagonists, and biased ligands. In some cases, such as the aminergic GPCRs and the A_{2A}R, conformational changes appear to be subtle, whereas in the P2Y₁₂R case very large differences are observed. An important issue to resolve is to what extent structures solved with a ligand of a certain class of functionality can be used in docking studies focused on another class. For the β -adrenoceptors, docking of inverse agonists into active state structures was found to be problematic, requiring application of IFD for necessary expansion of the binding site.⁴⁰ Vice versa, docking of the smaller partial agonists into the larger inactive state binding site was typically easier.

A recent example used an active state structure to prospectively identify partial and full agonists of β_2 AR in a virtual screen.⁷³ First, the active state receptor was shown to effectively and selectively retrieve known binders from a set of decoy ligands, and after running a screen of the ZINC library and testing 22 compounds, six compounds were found to be active. In addition, this study attempted to build an active state model of the D₂R using the active state structure of β_2 AR as a template. Virtual screening with the active D₂R model led to hit rates and potencies that were far lower when compared to a screen of the active β_2 AR structure and a similar screen of an inactive D₃R model built using the inactive β_2 AR structure.⁵⁹

The IFD protocol has also been applied to GPCR crystal structures to optimize the recognition of agonist compounds in virtual screening.¹¹³ For example, different crystal structures of the β_2 AR were evaluated for their ability to discriminate between agonist and antagonist compounds. Results clearly showed that, while inverse agonist- and antagonist-bound crystal structures (PDB IDs: 2RH1 and 3NYA) tended to favor the selection of antagonists, the active-state crystal structure (PDB ID: 3P0G) prioritized agonists over antagonists. With the aim to evaluate the effect of a ligand-induced optimization on the receptor's ability to recognize agonist or antagonist compounds, three β_2 AR models were built by refining the binding site of an inverse agonist-bound structure with three full agonists, i.e. epinephrine, isoproterenol and fenoterol, using IFD. All three agonist-induced β_2 AR models reverted their initial preference, being as effective in prioritizing agonists over antagonists as the active-state crystal structure of the β_2 AR.

When no active ligand structure is available, it is possible to introduce local conformational changes into the inactive state structure to model the interaction with agonists. One approach is to dock a single agonist into the receptor using a flexible receptor approach, and increase the efficiency of the screen by using Structural Interaction Fingerprints to ensure hits form key interactions with residues known to be important for agonist recognition.⁷⁹ Another approach is to use the transformation matrix derived by superposition of the inactive and active structures of one receptor, and apply that transformation to a different GPCR for which only an active state structure or model is available. This approach was used to generate an active state model of the D₂R, which was then used to dock a series of G-protein biased partial agonists.¹¹⁴

A final possible application of docking has been to use receptor flexibility as a tool to predict whether a compound is a blocker or a partial agonist, and what conformational changes are associated with binding of agonists vs. inverse agonists. IFD in tandem with linear discriminant analysis (LDA) was used to generate hypotheses of conformational changes induced in the β_2 AR receptor by agonist binding.¹¹⁵ This analysis suggested agonists induce subtle movements to TM5 of the receptor, in particular around Ser207 (5.46). This hypothesis was later confirmed by a crystal structure of the active state of β_2 AR, suggesting that conformational changes associated with agonists binding can be detected in other structures with a similar approach as well.

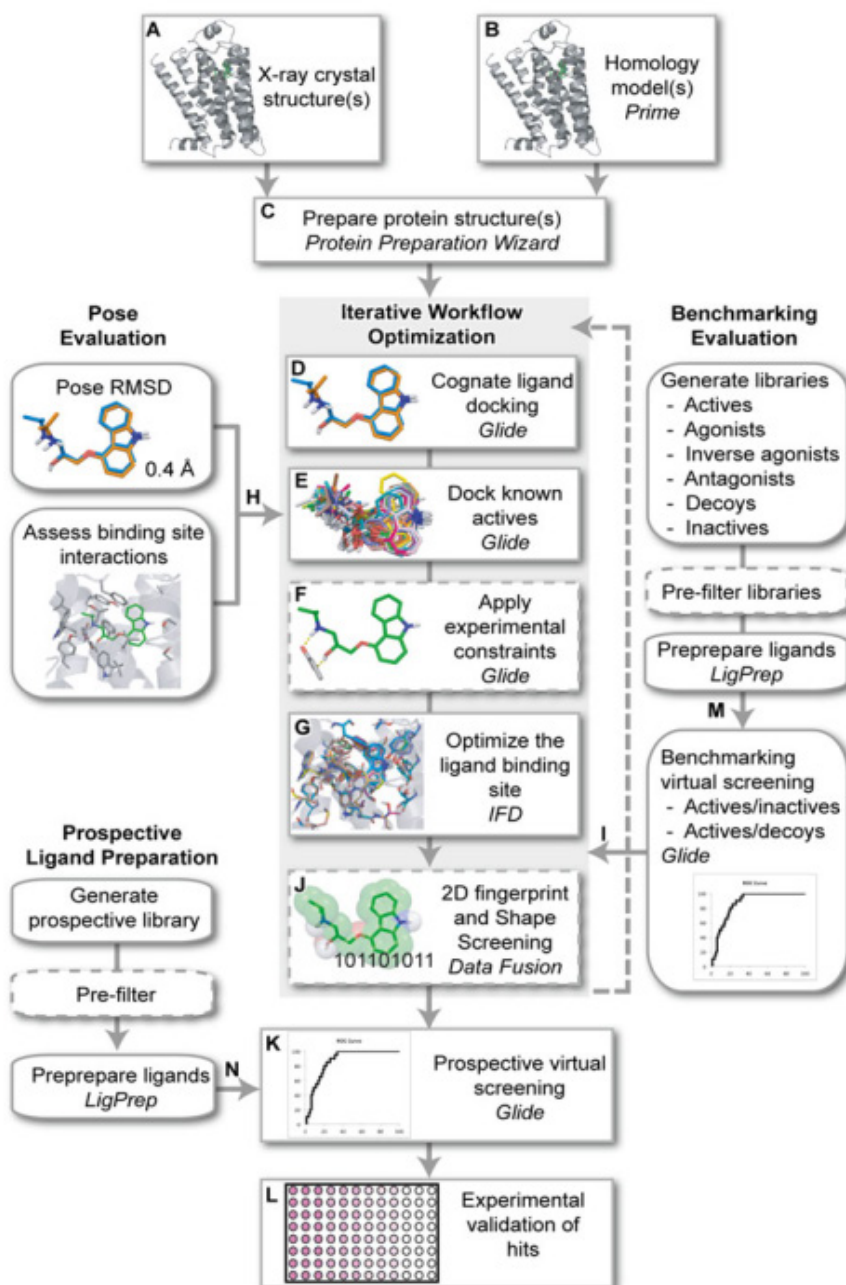


Figure 2.2 Proposed docking workflow for GPCR modeling. Customization of a docking protocol for a specific GPCR target depends on the quantity and accuracy of available data, including crystal structures, homology modeling templates, known actives, and mutagenesis data. Whenever possible, retrospective validation of the model/structure should be performed before prospective application. The programs used in the various stages are shown in *italics* and described in more detail in Table 1 and throughout the text.

A workflow for modeling of GPCR-ligand interactions

A GPCR modeling project can be approached with a number of different strategies and will involve the application of several of the tools described in this manuscript (see Table 2.1). The selection and order of the various steps will highly depend on the nature of the modeling problem, but some recommendations for a general workflow can be made (see Figure 2.2). The first step involves determining the correct structural representation of the GPCR to use. In a SBDD program, this involves selection of the correct crystal structure(s) (Figure 2.2A) or homology model(s) (Fig. 2.2B) to start from. At this stage the Protein Preparation Wizard should be used to select tautomeric and ionization states of residues, and which water molecules to include in the docking experiment, if any (Fig 2.2C). Before applying this structure to a prospective study, a number of validations and optimization steps should be carried out, often in an iterative fashion. In the case of crystal structures, a first check of the utility of the structure and chosen docking methodology is the correct redocking of the cognate ligand (Figure 2.2D). The next step is the docking of additional active molecules into a given structure or model, either using rigid receptor docking with Glide, or flexible receptor docking using IFD (Figure 2.2E). Correspondence of the resulting poses with experimental data (additional crystal structures, mutagenesis data, etc.) will need to be verified at this stage, and/or can be enforced with the use of docking constraints (Figure 2.2F). At this stage, IFD using selected active compounds can be used to modify the binding site to allow correct rigid receptor docking of Glide of a comprehensive set of active molecules (Figure 2.2G). Optimization of the structure is monitored by measuring both pose accuracy values (Figure 2.2H) and enrichment calculations using a small set of actives and decoys (Figure 2.2I) to determine the most appropriate structural variant(s) to go forward with. Data Fusion methods, bringing in results from 2D fingerprint and 3D-shape comparisons with known binders, can help increase the accuracy of the virtual screen (Figure 2J). Once this optimization of the project-specific docking workflow has been established, prospective application of the workflow, followed by experimental validation of the predictions, can proceed (Figure 2.2K and 2.2L). For all these steps, the creation of appropriate ligand libraries, including actives and decoys or tailored compounds databases to be used in prospective screening experiments is of great importance (Figure 2.2M and 2.2N).

Conclusions and Future Developments

Novel scoring functions that take into account hitherto neglected contributions to the free energy of binding, including protein desolvation and protein strain, are currently actively being developed.²¹ In addition, accurate methods for the prediction of ligand binding affinity (i.e. Free Energy Perturbation) are becoming increasingly more accurate and accessible.¹¹⁶ Together with a continued rise in the number of available structure of GPCRs, the application of structure-based approaches in GPCR drug discovery will become more frequent in the near future. In the future, as computers continue to grow in power, ab initio prediction of ligand binding through a study of the dynamical properties of proteins and ligands might become commonplace.¹¹⁷ In the meantime, the methods outlined in this review might serve as a guide to predict binding modes and perform virtual screening for this important class of targets.

References

- 1 Rosenbaum, D. M. *et al.* GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* **318**, 1266-1273, doi:10.1126/science.1150609 (2007).
- 2 Roth, C. B., Hanson, M. A. & Stevens, R. C. Stabilization of the human beta2-adrenergic receptor TM4-TM3-TM5 helix interface by mutagenesis of Glu122(3.41), a critical residue in GPCR structure. *J. Mol. Biol.* **376**, 1305-1319, doi:10.1016/j.jmb.2007.12.028 (2008).
- 3 Shibata, Y. *et al.* Thermostabilization of the neurotensin receptor NTS1. *J. Mol. Biol.* **390**, 262-277, doi:10.1016/j.jmb.2009.04.068 (2009).
- 4 Serrano-Vega, M. J., Magnani, F., Shibata, Y. & Tate, C. G. Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 877-882, doi:10.1073/pnas.0711253105 (2008).
- 5 Magnani, F., Shibata, Y., Serrano-Vega, M. J. & Tate, C. G. Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 10744-10749, doi:10.1073/pnas.0804396105 (2008).
- 6 Yin, X., Xu, H., Hanson, M. & Liu, W. GPCR crystallization using lipidic cubic phase technique. *Curr. Pharm. Biotechnol.* **15**, 971-979 (2014).
- 7 Liu, W. *et al.* Serial femtosecond crystallography of G protein-coupled receptors. *Science* **342**, 1521-1524, doi:10.1126/science.1244142 (2013).
- 8 Katritch, V., Cherezov, V. & Stevens, R. C. Diversity and modularity of G protein-coupled receptor structures. *Trends Pharmacol. Sci.* **33**, 17-27, doi:10.1016/j.tips.2011.09.003 (2012).
- 9 Guimaraes, C. R. & Cardozo, M. MM-GB/SA rescoring of docking poses in structure-based lead optimization. *J. Chem. Inf. Model* **48**, 958-970, doi:10.1021/ci800004w (2008).
- 10 Durdagi, S., Zhao, C., Cuervo, J. E. & Noskov, S. Y. Atomistic models for free energy evaluation of drug binding to membrane proteins. *Curr. Med. Chem.* **18**, 2601-2611 (2011).
- 11 Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V. & Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **11**, 425-445 (1997).
- 12 Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455-461, doi:10.1002/jcc.21334 (2010).
- 13 Gohlke, H., Hendlich, M. & Klebe, G. Knowledge-based scoring function to predict protein-ligand interactions. *J. Mol. Biol.* **295**, 337-356, doi:10.1006/jmbi.1999.3371 (2000).
- 14 Muegge, I. & Martin, Y. C. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. *J. Med. Chem.* **42**, 791-804, doi:10.1021/jm980536j (1999).
- 15 Neudert, G. & Klebe, G. DSX: a knowledge-based scoring function for the assessment of protein-ligand complexes. *J. Chem. Inf. Model* **51**, 2731-2745, doi:10.1021/ci200274q (2011).
- 16 Gschwend, D. A., Good, A. C. & Kuntz, I. D. Molecular docking towards drug discovery. *J. Mol. Recognit.* **9**, 175-186, doi:10.1002/(SICI)1099-1352(199603)9:2<175::AID-JMR260>3.0.CO;2-D (1996).
- 17 Friesner, R. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739-1749, doi:10.1021/jm0306430 (2004).
- 18 Friesner, R. A. *et al.* Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49**, 6177-6196, doi:10.1021/jm051256o (2006).
- 19 Halgren, T. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **47**, 1750-1759, doi:10.1021/jm030644s (2004).
- 20 Perola, E., Walters, W. P. & Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **56**, 235-249, doi:10.1002/prot.20088 (2004).
- 21 Repasky, M. P. *et al.* Docking performance of the glide program as evaluated on the Astex and DUD datasets: a complete set of glide SP results and selected results for a new scoring function integrating WaterMap and glide. *J. Comput. Aided Mol. Des.* **26**, 787-799 (2012).
- 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.* **49**, 534-553, doi:10.1021/jm050540c (2006).
- 23 Prime 3.7 (Schrödinger, Inc., Portland, OR, 2014).

- 24 Farid, R., Day, T., Friesner, R. A. & Pearlstein, R. A. New insights about HERG blockade obtained from protein modeling, potential energy mapping, and docking studies. *Bioorg. Med. Chem.* **14**, 3160-3173 (2006).
- 25 Sherman, W., Beard, H. S. & Farid, R. Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug Des.* **67**, 83-84 (2006).
- 26 LigPrep v3.1 (Schrödinger, Inc., Portland, OR, 2014).
- 27 Greenwood, J. R., Calkins, D., Sullivan, A. P. & Shelley, J. C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput. Aided Mol. Des.* **24**, 591-604, doi:10.1007/s10822-010-9349-1 (2010).
- 28 Shelley, J. C. *et al.* Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J. Comput. Aided Mol. Des.* **21**, 681-691, doi:10.1007/s10822-007-9133-z (2007).
- 29 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.* **27**, 221-234, doi:10.1007/s10822-013-9644-8 (2013).
- 30 Maestro 9.9 (Schrödinger, Inc., Portland, OR, 2014).
- 31 Bas, D. C., Rogers, D. M. & Jensen, J. H. Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins* **73**, 765-783, doi:10.1002/prot.22102 (2008).
- 32 Yusuf, D., Davis, A. M., Kleywegt, G. J. & Schmitt, S. An alternative method for the evaluation of docking performance: RSR vs RMSD. *J Chem Inf Model* **48**, 1411-1422, doi:10.1021/ci800084x (2008).
- 33 Baber, J. C., Thompson, D. C., Cross, J. B. & Humblet, C. GARD: a Generally Applicable Replacement for RMSD. *J Chem Inf Model* **49**, 1889-1900, doi:10.1021/ci9001074 (2009).
- 34 Schulz-Gasch, T., Scharfer, C., Guba, W. & Rarey, M. TFD: Torsion Fingerprints as a new measure to compare small molecule conformations. *J. Chem. Inf. Model* **52**, 1499-1512, doi:10.1021/ci2002318 (2012).
- 35 Michino, M. *et al.* Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat. Rev. Drug Discov.* **8**, 455-463, doi:10.1038/nrd2877 (2009).
- 36 Kufareva, I., Katritch, V., Participants of, G. D., Stevens, R. C. & Abagyan, R. Advances in GPCR modeling evaluated by the GPCR Dock 2013 assessment: meeting new challenges. *Structure* **22**, 1120-1139, doi:10.1016/j.str.2014.06.012 (2014).
- 37 Kufareva, I. *et al.* Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **19**, 1108-1126, doi:10.1016/j.str.2011.05.012 (2011).
- 38 Truchon, J. F. & Bayly, C. I. Evaluating virtual screening methods: good and bad metrics for the “early recognition” problem. *J. Chem. Inf. Model* **47**, 488-508, doi:10.1021/ci600426e (2007).
- 39 Urban, J. D. *et al.* Functional selectivity and classical concepts of quantitative pharmacology. *J. Pharmacol. Exp. Ther.* **320**, 1-13, doi:10.1124/jpet.106.104463 (2007).
- 40 Beuming, T. & Sherman, W. Current assessment of docking into GPCR crystal structures and homology models: successes, challenges, and guidelines. *J. Chem. Inf. Model* **52**, 3263-3277, doi:10.1021/ci300411b (2012).
- 41 Lenselink, E. B., Beuming, T., Sherman, W., van Vlijmen, H. W. & A.P., I. Selecting an optimal number of binding site waters to improve virtual screening enrichments against the adenosine A2A receptor. *J. Chem. Inf. Model* **54**, 1737-1746, doi:10.1021/ci5000455 (2014).
- 42 Bottegoni, G., Kufareva, I., Totrov, M. & Abagyan, R. Four-dimensional docking: a fast and accurate account of discrete receptor flexibility in ligand docking. *J. Med. Chem.* **52**, 397-406, doi:10.1021/jm8009958 (2009).
- 43 Greenidge, P. A., Kramer, C., Mozziconacci, J. C. & Sherman, W. Improving docking results via reranking of ensembles of ligand poses in multiple X-ray protein conformations with MM-GBSA. *J. Chem. Inf. Model* **54**, 2697-2717, doi:10.1021/ci5003735 (2014).
- 44 Huang, S. Y. & Zou, X. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. *Proteins* **66**, 399-421, doi:10.1002/prot.21214 (2007).
- 45 Simms, J. *et al.* Homology modeling of GPCRs. *Methods Mol. Biol.* **552**, 97-113, doi:10.1007/978-1-60327-317-6_7 (2009).

- 46 Nguyen, E. D., Norn, C., Frimurer, T. M. & Meiler, J. Assessment and challenges of ligand docking into comparative models of G-protein coupled receptors. *PloS one* **8**, e67302, doi:10.1371/journal.pone.0067302 (2013).
- 47 Costanzi, S. Modeling G protein-coupled receptors and their interactions with ligands. *Curr. Opin. Struct. Biol.* **23**, 185-190, doi:10.1016/j.sbi.2013.01.008 (2013).
- 48 Pala, D. *et al.* Structure-based virtual screening of MT2 melatonin receptor: influence of template choice and structural refinement. *J. Chem. Inf. Model* **53**, 821-835, doi:10.1021/ci4000147 (2013).
- 49 Bender, A. & Glen, R. C. A discussion of measures of enrichment in virtual screening: comparing the information content of descriptors with increasing levels of sophistication. *J. Chem. Inf. Model* **45**, 1369-1375, doi:10.1021/ci0500177 (2005).
- 50 Good, A. C. & Oprea, T. I. Optimization of CAMD techniques 3. Virtual screening enrichment studies: a help or hindrance in tool selection? *J. Comput. Aided Mol. Des.* **22**, 169-178, doi:10.1007/s10822-007-9167-2 (2008).
- 51 Kirchmair, J., Markt, P., Distinto, S., Wolber, G. & Langer, T. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection--what can we learn from earlier mistakes? *J. Comput. Aided Mol. Des.* **22**, 213-228, doi:10.1007/s10822-007-9163-6 (2008).
- 52 Katritch, V., Kufareva, I. & Abagyan, R. Structure based prediction of subtype-selectivity for adenosine receptor antagonists. *Neuropharmacology* **60**, 108-115, doi:10.1016/j.neuropharm.2010.07.009 (2011).
- 53 Katritch, V. *et al.* Structure-based discovery of novel chemotypes for adenosine A(2A) receptor antagonists. *J. Med. Chem.* **53**, 1799-1809, doi:10.1021/jm901647p (2010).
- 54 Carlsson, J. *et al.* Structure-based discovery of A2A adenosine receptor ligands. *J. Med. Chem.* **53**, 3748-3755, doi:10.1021/jm100240h (2010).
- 55 Mysinger, M. M. *et al.* Structure-based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5517-5522, doi:10.1073/pnas.1120431109 (2012).
- 56 Kruse, A. C. *et al.* Muscarinic receptors as model targets and antitargets for structure-based ligand discovery. *Mol. Pharmacol.* **84**, 528-540, doi:10.1124/mol.113.087551 (2013).
- 57 de Graaf, C. *et al.* Crystal structure-based virtual screening for fragment-like ligands of the human histamine H(1) receptor. *J. Med. Chem.* **54**, 8195-8206, doi:10.1021/jm2011589 (2011).
- 58 Negri, A. *et al.* Discovery of a novel selective kappa-opioid receptor agonist using crystal structure-based virtual screening. *J. Chem. Inf. Model* **53**, 521-526, doi:10.1021/ci400019t (2013).
- 59 Carlsson, J. *et al.* Ligand discovery from a dopamine D3 receptor homology model and crystal structure. *Nat. Chem. Biol.* **7**, 769-778, doi:10.1038/nchembio.662 (2011).
- 60 Vass, M., Schmidt, E., Horti, F. & Keseru, G. M. Virtual fragment screening on GPCRs: a case study on dopamine D3 and histamine H4 receptors. *Eur. J. Med. Chem.* **77**, 38-46, doi:10.1016/j.ejmech.2014.02.034 (2014).
- 61 Lane, J. R. *et al.* Structure-based ligand discovery targeting orthosteric and allosteric pockets of dopamine receptors. *Mol. Pharmacol.* **84**, 794-807, doi:10.1124/mol.113.088054 (2013).
- 62 Kolb, P. *et al.* Structure-based discovery of beta2-adrenergic receptor ligands. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 6843-6848, doi:10.1073/pnas.0812657106 (2009).
- 63 Vilar, S. *et al.* Docking-based virtual screening for ligands of G protein-coupled receptors: not only crystal structures but also in silico models. *J. Mol. Graph. Model* **29**, 614-623, doi:10.1016/j.jmgm.2010.11.005 (2011).
- 64 Engel, S. *et al.* A virtual screen for diverse ligands: discovery of selective G protein-coupled receptor antagonists. *J. Am. Chem. Soc.* **130**, 5115-5123, doi:10.1021/ja077620l (2008).
- 65 Costanzi, S., Santhosh Kumar, T., Balasubramanian, R., Kendall Harden, T. & Jacobson, K. A. Virtual screening leads to the discovery of novel non-nucleotide P2Y(1) receptor antagonists. *Bioorg. Med. Chem.* **20**, 5254-5261, doi:10.1016/j.bmc.2012.06.044 (2012).
- 66 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.* **522**, 279-336, doi:10.1016/B978-0-12-407865-9.00015-7 (2013).

- 67 Andrews, S. P., Brown, G. A. & Christopher, J. A. Structure-based and fragment-based GPCR drug discovery. *ChemMedChem* **9**, 256-275, doi:10.1002/cmdc.201300382 (2014).
- 68 Chin, S. P., Buckle, M. J., Chalmers, D. K., Yuriev, E. & Doughty, S. W. Toward activated homology models of the human M1 muscarinic acetylcholine receptor. *J. Mol. Graph. Model* **49**, 91-98, doi:10.1016/j.jmgm.2014.02.002 (2014).
- 69 Kolaczowski, 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* **53**, 638-648, doi:10.1021/ci300413h (2013).
- 70 McRobb, F. M., Capuano, B., Crosby, I. T., Chalmers, D. K. & Yuriev, E. Homology modeling and docking evaluation of aminergic G protein-coupled receptors. *J. Chem. Inf. Model* **50**, 626-637, doi:10.1021/ci900444q (2010).
- 71 Thomas, T. *et al.* Homology modeling of human muscarinic acetylcholine receptors. *J. Chem. Inf. Model* **54**, 243-253, doi:10.1021/ci400502u (2014).
- 72 Gatica, E. A. & Cavasotto, C. N. Ligand and decoy sets for docking to G protein-coupled receptors. *J. Chem. Inf. Model* **52**, 1-6, doi:10.1021/ci200412p (2012).
- 73 Weiss, D. R. *et al.* Conformation guides molecular efficacy in docking screens of activated beta-2 adrenergic G protein coupled receptor. *ACS Chem. Biol.* **8**, 1018-1026, doi:10.1021/cb400103f (2013).
- 74 Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **44**, 235-249 (2000).
- 75 Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Deliv. Rev.* **46**, 3-26 (2001).
- 76 Teague, S. J., Davis, A. M., Leeson, P. D. & Oprea, T. The Design of Leadlike Combinatorial Libraries. *Angew. Chem. Int Ed. Engl.* **38**, 3743-3748 (1999).
- 77 Langmead, C. J. *et al.* Identification of novel adenosine A(2A) receptor antagonists by virtual screening. *J. Med. Chem.* **55**, 1904-1909, doi:10.1021/jm201455y (2012).
- 78 Deng, Z., Chuaqui, C. & Singh, J. Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein-ligand binding interactions. *J. Med. Chem.* **47**, 337-344, doi:10.1021/jm030331x (2004).
- 79 de Graaf, C. & Rognan, D. Selective structure-based virtual screening for full and partial agonists of the beta2 adrenergic receptor. *J. Med. Chem.* **51**, 4978-4985, doi:10.1021/jm800710x (2008).
- 80 Rose, P. W. *et al.* The RCSB Protein Data Bank: views of structural biology for basic and applied research and education. *Nucleic Acids Res.* **43**, D345-356, doi:10.1093/nar/gku1214 (2015).
- 81 Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. B* **58**, 380-388 (2002).
- 82 Ballesteros, J. A. & Weinstein, H. [19] 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 (1995).
- 83 Shi, L. & Javitch, J. A. The second extracellular loop of the dopamine D2 receptor lines the binding-site crevice. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 440-445, doi:10.1073/pnas.2237265100 (2004).
- 84 Abel, R. *et al.* Contribution of explicit solvent effects to the binding affinity of small-molecule inhibitors in blood coagulation factor serine proteases. *ChemMedChem* **6**, 1049-1066, doi:10.1002/cmdc.201000533 (2011).
- 85 Rapp, C., Kalyanaraman, C., Schiffmiller, A., Schoenbrun, E. L. & Jacobson, M. P. A molecular mechanics approach to modeling protein-ligand interactions: relative binding affinities in congeneric series. *J. Chem. Inf. Model* **51**, 2082-2089, doi:10.1021/ci200033n (2011).
- 86 Sastry, G. M., Inakollu, V. S. & Sherman, W. Boosting virtual screening enrichments with data fusion: coalescing hits from two-dimensional fingerprints, shape, and docking. *J. Chem. Inf. Model* **53**, 1531-1542, doi:10.1021/ci300463g (2013).
- 87 MacroModel v10.5 (Schrödinger, Inc., Portland, OR, 2014).
- 88 Jaguar v8.5 (Schrödinger, Inc., Portland, OR, 2014).
- 89 Abel, R., Young, T., Farid, R., Berne, B. J. & Friesner, R. A. Role of the active-site solvent in the thermodynamics of factor Xa ligand binding. *J. Am. Chem. Soc.* **130**, 2817-2831, doi:10.1021/ja0771033 (2008).

- 90 Young, T., Abel, R., Kim, B., Berne, B. J. & Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 808-813, doi:10.1073/pnas.0610202104 (2007).
- 91 Beuming, T. *et al.* Thermodynamic analysis of water molecules at the surface of proteins and applications to binding site prediction and characterization. *Proteins* **80**, 871-883, doi:10.1002/prot.23244 (2012).
- 92 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.* **33**, 249-260, doi:10.1016/j.tips.2012.02.005 (2012).
- 93 Higgs, C., Beuming, T. & Sherman, W. Hydration Site Thermodynamics Explain SARs for Triazolylpurines Analogues Binding to the A2A Receptor. *ACS Med. Chem. Lett.* **1**, 160-164, doi:10.1021/ml100008s (2010).
- 94 Ivanov, A. A., Barak, D. & Jacobson, K. A. Evaluation of homology modeling of G-protein-coupled receptors in light of the A(2A) adenosine receptor crystallographic structure. *J. Med. Chem.* **52**, 3284-3292, doi:10.1021/jm801533x (2009).
- 95 Liu, W. *et al.* Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **337**, 232-236, doi:10.1126/science.1219218 (2012).
- 96 Fenalti, G. *et al.* Molecular control of delta-opioid receptor signalling. *Nature* **506**, 191-196, doi:10.1038/nature12944 (2014).
- 97 Dixon, S. L. *et al.* PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J. Comput. Aided Mol. Des.* **20**, 647-671, doi:10.1007/s10822-006-9087-6 (2006).
- 98 Salam, N. K., Nuti, R. & Sherman, W. Novel method for generating structure-based pharmacophores using energetic analysis. *J. Chem. Inf. Model* **49**, 2356-2368, doi:10.1021/ci900212v (2009).
- 99 Wolber, G. & Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model* **45**, 160-169, doi:10.1021/ci049885e (2005).
- 100 Kirchmair, J. *et al.* How to optimize shape-based virtual screening: choosing the right query and including chemical information. *J. Chem. Inf. Model* **49**, 678-692, doi:10.1021/ci8004226 (2009).
- 101 Sastry, G. M., Dixon, S. L. & Sherman, W. Rapid shape-based ligand alignment and virtual screening method based on atom/feature-pair similarities and volume overlap scoring. *J. Chem. Inf. Model* **51**, 2455-2466, doi:10.1021/ci2002704 (2011).
- 102 Eckert, H. & Bajorath, J. Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches. *Drug. Discov. Today* **12**, 225-233, doi:10.1016/j.drudis.2007.01.011 (2007).
- 103 Willett, P. Similarity-based virtual screening using 2D fingerprints. *Drug. Discov. Today* **11**, 1046-1053, doi:10.1016/j.drudis.2006.10.005 (2006).
- 104 Hert, J. *et al.* New methods for ligand-based virtual screening: use of data fusion and machine learning to enhance the effectiveness of similarity searching. *J. Chem. Inf. Model* **46**, 462-470, doi:10.1021/ci050348j (2006).
- 105 Tan, L., Geppert, H., Sisay, M. T., Gutschow, M. & Bajorath, J. Integrating structure- and ligand-based virtual screening: comparison of individual, parallel, and fused molecular docking and similarity search calculations on multiple targets. *ChemMedChem* **3**, 1566-1571, doi:10.1002/cmdc.200800129 (2008).
- 106 Zhang, Q. & Muegge, I. Scaffold hopping through virtual screening using 2D and 3D similarity descriptors: ranking, voting, and consensus scoring. *J. Med. Chem.* **49**, 1536-1548, doi:10.1021/jm050468i (2006).
- 107 Rasmussen, S. G. *et al.* Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549-555, doi:10.1038/nature10361 (2011).
- 108 Warne, T. *et al.* The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor. *Nature* **469**, 241-244, doi:10.1038/nature09746 (2011).
- 109 Lebon, G. *et al.* Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* **474**, 521-525, doi:10.1038/nature10136 (2011).
- 110 Zhang, J. *et al.* Agonist-bound structure of the human P2Y12 receptor. *Nature* **509**, 119-122, doi:10.1038/nature13288 (2014).

- 111 Egloff, P. *et al.* Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E655-662, doi:10.1073/pnas.1317903111 (2014).
- 112 Kruse, A. C. *et al.* Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**, 101-106, doi:10.1038/nature12735 (2013).
- 113 Costanzi, S. & Vilar, S. In silico screening for agonists and blockers of the beta(2) adrenergic receptor: implications of inactive and activated state structures. *J. Comput. Chem.* **33**, 561-572, doi:10.1002/jcc.22893 (2012).
- 114 Free, R. B. *et al.* Discovery and characterization of a G protein-biased agonist that inhibits beta-arrestin recruitment to the D2 dopamine receptor. *Mol. Pharmacol.* **86**, 96-105, doi:10.1124/mol.113.090563 (2014).
- 115 Vilar, S., Karpiak, J., Berk, B. & Costanzi, S. In silico analysis of the binding of agonists and blockers to the beta2-adrenergic receptor. *J. Mol. Graph. Model* **29**, 809-817, doi:10.1016/j.jmgm.2011.01.005 (2011).
- 116 Goldfeld, D. A. *et al.* Docking and free energy perturbation studies of ligand binding in the kappa opioid receptor. *J. Phys. Chem. B* **119**, 824-835, doi:10.1021/jp5053612 (2015).
- 117 Dror, R. O. *et al.* Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 13118-13123, doi:10.1073/pnas.1104614108 (2011).

