Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Title page

A molecular pharmacologist's guide to GPCR crystallography

Chayne L. Piscitelli, James Kean, Chris de Graaf, Xavier Deupi

Laboratory of Biomolecular Research, Department of Biology and Chemistry (*C.P. and X.D.*), and Condensed Matter Theory Group, Department of Research with Neutrons and Muons (*X.D.)*, Paul Scherrer Institute, CH-5232 Villigen, Switzerland; Heptares Therapeutics Ltd, BioPark, Broadwater Road, Welwyn Garden City, AL7 3AX, UK (*J.K.*); and Division of Medicinal Chemistry, Faculty of Sciences, Amsterdam Institute for Molecules, Medicines and Systems (AIMMS), VU University of Amsterdam, Amsterdam, The Netherlands (*C. dG.*).

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

CP and JK contributed equally to this work

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

MOL #99663

Running Title Page.

A molecular pharmacologist's guide to GPCR crystallography

Corresponding authors: CdG (c.de.graaf@vu.nl), XD (xavier.deupi@psi.ch)

Dr. Xavier Deupi
Paul Scherrer Institut
WHGA/114
CH-5232 Villigen PSI, Switzerland
Tel. +41 56 310 3337
xavier.deupi@psi.ch

Number of pages:	51
Number of tables:	3
Number of figures:	5
Number of references:	128
Words in the Abstract:	203
Words in the Introduction:	681
Words in the Discussion:	9461
Words in the Conclusion:	263

ABBREVIATIONS: GPCR, G protein-coupled receptor; PDB, Protein Data Bank; SPR, surface plasmon resonance; LCP, lipidic cubic phase; XFEL, X-ray free electron laser; SFX, serial femtosecond crystallography; SBDD, structure-based drug design; ER, endoplasmic reticulum; ECL, extracellular loop; ICL, intracellular loop; PDB ID, Protein Data Bank identifier; $mGlu_5R$, Metabotropic glutamate receptor 5; CRF_1R, corticotropin-releasing factor receptor 1; β_1AR, beta 1 adrenergic receptor; β_2AR, beta 2 adrenergic receptor; $P_2Y_{12}R$, P_2Y_{12} receptor; $A_{2A}R$, adenosine 2A receptor; NTS1R, neurotensin 1 receptor; CXCR4, C-X-C chemokine receptor type 4; T4L, T4 lysozyme; OX_2R, orexin 2 receptor; CCR5, C-C chemokine receptor type 5; P_2Y_1R, P_2Y_1 receptor; TM, transmembrane; Fab, fragment antigen-binding; Nb, nanobody; G protein, GTP binding protein, CX3CL1R, fraktalkine (CX3CL1) receptor; MAD, multi-wavelength anomalous diffraction; MR, molecular replacement; RMSD, root mean square deviation; ESD or ESU, estimated coordinate error value; RMSF, root mean square fluctuation; EDS, Electron Density Server; SAR, structure-activity relationship; D_3R, dopamine 3 receptor; $5HT_{1B}$, serotonin 1B receptor; PAR1, Protease-activated receptor type 1; FFA1, Free fatty acid receptor 1; M_3R, Muscarinic acetylcholine receptor M_3; H_1R, histamine 1 receptor; IFP, interaction fingerprint; S1P1, Sphingosine-1-phosphate receptor 1; M_2R, Muscarinic acetylcholine receptor M_2; $mGlu_1R$, Metabotropic glutamate receptor 1; NMR, nuclear magnetic resonance; EM, electron microscopy.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract

G protein-coupled receptor (GPCR) structural biology has progressed dramatically in the last decade. There are now over 120 GPCR crystal structures deposited in the Protein Data Bank of 32 different receptors from families scattered across the phylogenetic tree, including Class B, C, and Frizzled GPCRs. These structures have been obtained in combination with a wide variety of ligands, and captured in a range of conformational states. This surge in structural knowledge has enlightened research into the molecular recognition of biologically active molecules, the mechanisms of receptor activation, the dynamics of functional selectivity, and fuelled structure-based drug design efforts for GPCRs. Here we summarize the innovations in both protein engineering/molecular biology and crystallography techniques that have led to these advances in GPCR structural biology, and discuss how they may influence the resulting structural models. We also provide a brief molecular pharmacologist's guide to GPCR X-ray crystallography, outlining some key aspects in the process of structure determination, with the goal to encourage non-crystallographers to interrogate structures at the molecular level. Finally we show how chemogenomics approaches can be used to marry the wealth of existing receptor pharmacology data with the expanding repertoire of structures, providing a deeper understanding of the mechanistic details of GPCR function.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Introduction

The mechanisms by which drugs act on receptors involve a complex interplay of thermodynamic and kinetic parameters, dictated in large part by the structures of the molecules involved. Researchers studying molecular pharmacology use a wide range of techniques to elucidate the mechanisms of drug action. These include the measurement of direct interactions between ligands and receptors, such as labeled ligand binding studies and surface plasmon resonance (SPR), functional assays to analyze signaling pathways, and monitoring conformational changes through fluorescence labeling of receptor subdomains. Decades of elegant pharmacology research have been heavily influenced in the last years by the increasing availability of crystallographic structural data, which have provided a slew of molecular models of protein-ligand complexes, often of therapeutic interest. G protein-coupled receptors (GPCRs) have exemplified this trend, with a veritable explosion in the number of receptor structures in the last eight years (Figure 1, top panels), which have provided the opportunity to take an exquisite look into the atomic details of drug binding and receptor activation mechanisms, and propelled the application of virtual screening and structure-based drug design to this family of receptors. Interestingly, many of these structures have been obtained in complex with pharmaceutically relevant drugs (Table 1).

A prerequisite to attempt solving the structure of a protein by X-ray crystallography is the preparation of large quantities (milligrams) of purified, stable, and homogeneous protein. GPCRs have been historically hard targets for structural biology due to the difficulty to obtain samples that satisfy these ideal conditions (Kobilka, 2013). However, over the last 10 years, structural biology of GPCRs has advanced dramatically with the emergence of new technologies and techniques to tackle the problems of receptor instability and homogeneity. Perhaps the most important of such advances have been the use of protein engineering (e.g.,

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

mutagenesis, truncations, and the creation of chimeric constructs) and the generation of receptor-specific protein binding partners for co-crystallization (such as conformational antibodies (Hino et al., 2012; Rasmussen et al., 2007), camelid antibody fragments (nanobodies), directed against the receptor (Rasmussen et al., 2011a) or the G protein to stabilize the active state ternary complex (Rasmussen et al., 2011b)). Furthermore, the discovery of new detergents for protein solubilization (Chae et al., 2010; Cho et al., 2015), the development of novel crystallization techniques such as lipidic cubic phase (LCP) crystallization (Caffrey, 2015) and key advances in several aspects of crystallography, such as microcrystallography (Moukhametzianov et al., 2008), X-ray free electron lasers (XFEL) and serial femtosecond crystallography (SFX) (Liu et al., 2013) have also contributed decisively to the field of GPCR structural biology.

It is important to realize that despite the outstanding value of 3D GPCR structures to understand the mechanisms of drug action at the molecular level, they must be viewed primarily as molecular models fitted to crystallographic data. The accuracy of these models, especially at the level of individual amino acid side chains, ligands, water molecules, and polar networks are completely dependent on the quality and completeness of the measured X-ray diffraction datasets. As a general rule, 3D structures should be backed up with pharmacological data, to ratify the models and provide a broader context to the analysis of ligand-receptor interactions. Such approaches can provide new insights into the mechanisms of GPCR activation and form the basis for structure-based drug design (SBDD) (Congreve et al., 2014).

In this review, we provide an overview of the innovations in protein engineering and crystallography that have led to the recent successes in GPCR structural biology. We first outline the molecular biology techniques currently used to facilitate GPCR crystallography, discussing the rationale behind each type of receptor modification, and how these can influence the information that can be obtained from the resulting structure. We then present an overview

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

of the process of structural determination by X-ray crystallography, focusing on the concepts more relevant to molecular pharmacologists. This second section aims to provide some guidance to assess crystallographic information and on how to interpret the quality of structural models. Finally, we show an example of how structural and pharmacology data can be combined using chemogenomics techniques to gain a deeper understanding of GPCR molecular pharmacology.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Molecular biology approaches to facilitate GPCR crystallography

As in the case of soluble proteins in the early days of crystallography, the structures of the first GPCRs were solved thanks to the natural advantages that they offered to the crystallographer. Specifically, the first structures of a GPCR, bovine rhodopsin, were obtained through isolation of large quantities of the receptor from a natural source - retinal rod outer segments (Li et al., 2004; Palczewski et al., 2000). Later, crystallization of squid rhodopsin was achieved through a similar route (Murakami and Kouyama, 2008). In addition, rhodopsin is also surprisingly stable when solubilized from the membrane using short-chain detergents, making it viable for vapor diffusion crystallography (Standfuss et al., 2007). Unfortunately, other GPCRs cannot be obtained as easily as rhodopsin, and need to be overexpressed in recombinant systems, often leading to heterogeneity due to post-translational modifications. Furthermore, unlike rhodopsin, most GPCRs are highly dynamic, and unstable upon solubilization. Herein we describe the molecular biology approaches used by crystallographers to overcome these difficulties and optimize GPCR constructs for successful crystallization (Figure 1, bottom panel).

Heterogeneity from post-translational modifications

The choice of expression system has important consequences on the characteristics of the expressed protein (Tate and Grisshammer, 1996), and expression of GPCRs using heterologous systems such as bacteria, yeast, insect or mammalian cells, or cell-free systems has had varying degrees of success (Milic and Veprintsev, 2015). The folding and maturation pathways of membrane proteins differ among these expression organisms. For instance, bacterial systems do not produce N-linked glycosylation, which could be the reason why functional expression of GPCRs in bacteria has had only limited success; N-linked glycosylation is cotranslational and, therefore, can be important for the expression, folding and cell surface localization of GPCRs. In mammalian cells, the glycosylation process occurs in the endoplasmic

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

reticulum (ER) and the Golgi apparatus as the receptor is trafficked through the cell (Hossler et al., 2009). The glycosylation process is very complex, resulting in varying combinations of branched glycans attached to specific extracellular sites. Also, the more glycosylation sites a receptor has, the higher its potential heterogeneity. However, numerous studies have shown that simply eliminating receptor glycosylation either by mutagenesis or by chemical inhibition of the glycosylation machinery in eukaryotic expression systems results in poorly trafficked receptor and retardation in the ER and Golgi (Chen et al., 2010a; Lanctot et al., 2006; Norskov-Lauritsen et al., 2015). On the other hand, deglycosylation of the N-terminus of GPCRs has not been shown to affect their pharmacology (Haga et al., 2012; Kruse et al., 2012; Shimamura et al., 2011). The role of glycosylation on the extracellular loops of GPCRs may be more complicated. For instance, deglycosylation of extracellular loop (ECL) 2 in PAR1 has been shown to enhance the maximal signaling response (Soto and Trejo, 2010), possibly due to interactions between the sugar group and the ligand that influence the stability of the active receptor conformation. Unfortunately, complex sugar groups are usually not resolved in X-ray structures due to their flexibility, unless they form crystal contacts (Crispin et al., 2009; Palczewski et al., 2000).

Other post-translational modifications including phosphorylation and palmitoylation are also sometimes removed for GPCR crystallography (Warne et al., 2008). Reversible phosphorylation occurs at the C-terminus as well as the intracellular loops (ICLs) in response to receptor activation, and is important for receptor desensitization and internalization (Nobles et al., 2011; Tobin, 2008). Phosphorylation is mediated through a number of different protein kinases and different patterns of phosphorylation can occur in different tissue types. This leads to the possibility that phosphorylation influences the signaling capabilities of receptors in different tissues. Palmitoylation occurs at the C-terminus, typically at conserved cysteine residues present in helix 8, although some receptors are palmitoylated at multiple sites in the C-terminus

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

(Zuckerman et al., 2011). Palmitoylation is also reversible and serves to anchor the C-terminus of the receptor to the lipid bilayer, effectively creating a fourth, and sometimes fifth, ICL. This structural rearrangement of the C-terminus has been linked to biased signaling, as well as to the dynamics of receptor phosphorylation and desensitization (Zuckerman et al., 2011). Removal of these modifications in most GPCR structures has usually been as a consequence of truncations to reduce flexibility of unstructured regions. Helix 8 is usually retained due to its role in GPCR signaling, although some receptors, e.g. $mGlu_5R$ (Dore et al., 2014) and CRF_1R (Hollenstein et al., 2013), have been solved with helix 8 partially or completely removed.

Of the ~120 GPCR crystal structures deposited in the Protein Data Bank (Berman et al., 2000) (www.pdb.org), 54 correspond to unique sequences of 30 different non-rhodopsin receptors (for instance, the turkey β_1AR has been solved using three different constructs, see Supplementary Table 1). 51 of these sequences have some form of receptor truncation (Supplementary Table 1): 43 constructs have been crystallized with C-terminal truncations, 32 have been truncated at the N-terminus, sometimes removing whole domains (Dore et al., 2014; Hollenstein et al., 2013; Siu et al., 2013; Wu et al., 2014), and 35 were crystallized with shortened ICLs (Egloff et al., 2014; Warne et al., 2008; Zou et al., 2012).

Mutagenesis and conformational stabilization

Site-directed mutagenesis is an additional tool used to improve GPCR crystallizability. In addition to being used to remove sites of post-translational modifications (see above), mutagenesis is also used to enhance expression. For instance, mutations $C116^{3.27}L$ in the β_1AR (Warne et al., 2008), $E122^{3.41}W$ in the β_2AR (Hanson et al., 2008), $I135^{3.29}L$ in the κ opioid receptor (Wu et al., 2012), and $D294^{7.49}N$ in the $P2Y_{12}R$ (Zhang et al., 2014) resulted in phenotypes with an increased level of expression of functional receptor (Supplementary Table 1).

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

But perhaps the most creative use of site-directed mutagenesis in the field of GPCR crystallography has been its application to conformational stabilization. In order to crystallize GPCRs, the receptor molecules must first be extracted from the lipidic membrane using detergents. However, GPCRs are typically unstable in detergent solution. In addition, GPCRs are highly dynamic proteins, existing in a range of conformational states between inactive (R) and active forms (R*) (Lohse et al., 2014), which further hinders crystallization. The inactive conformation of the receptor is a low energy state, and therefore a more stable form of the receptor. This is one of the reasons why the majority of GPCR structures have been solved in the inactive conformation in the presence of a stabilizing antagonist or inverse agonist (Ghosh et al., 2015). Some receptors are inherently more stable, and the formation of a ligand complex together with favorable binding kinetics is enough to stabilize the receptor for crystallization (Cherezov et al., 2007; Shimamura et al., 2011). In other cases (Wu et al., 2012), where the low stability of the receptor precluded purification of functional protein, or sufficiently stabilizing ligands were unavailable, the receptor is stabilized by mutagenesis.

The first receptor to be stabilized using this approach was the turkey β_1AR in the inactive conformation (Warne et al., 2008). The stabilized receptor contained six mutations, and showed similar binding affinities to the wild type receptor for antagonist ligands, but a reduced affinity for agonist ligands, indicating that the receptor had been conformationally stabilized towards the R state. Comparison of the turkey β_1AR structure to that of the homologous human β_2AR inactive structures (solved without stabilizing mutations) (Cherezov et al., 2007; Rasmussen et al., 2007) showed little differences, even around the mutated positions (Warne et al., 2008). A_{2A}R is another case of a GPCR stabilized using this approach, and is the only example of a receptor solved using the same ligand with and without conformational stabilization by mutagenesis (Dore et al., 2011; Jaakola et al., 2008). Comparison of the structure around the stabilizing mutations did not show local perturbations (Dore et al., 2011), despite the receptors having

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

different overall conformations, especially in the regions of transmembrane helices (TM) 5 and 6. Consistent with the pharmacological activity of the two constructs, the stabilized mutant is apparently locked into the inactive state, whereas the non-stabilized construct appears to be more conformationally flexible (Jaakola et al., 2008).

The molecular mechanisms by which certain point mutations result in thermostabilization are not yet fully understood (Heydenreich et al., 2015; Tate, 2012; Tate and Schertler, 2009). While such mutations are, to some extent, transferrable between close orthologs (Serrano-Vega and Tate, 2009), successful thermostabilizing mutations are difficult to predict, and computational methods are being developed to address this issue (Chen et al., 2012; Lee et al., 2014).

The structures stabilized by mutagenesis mentioned above correspond to inactive states. In an effort to crystallize an active conformation of the β_2AR receptor, a mutation was introduced into the ligand binding site that enabled the covalent binding of a designed agonist, FAUC50 (Rosenbaum et al., 2011). However, the obtained structure retained an inactive conformation, indicating that agonist binding alone may not be sufficient for the stabilization of the active state. Indeed, when stabilization by mutagenesis was used to obtain the structures of agonist-bound A_{2A}R (Lebon et al., 2011) and NTS_1R (White et al., 2012), this also resulted only in partially active states (Deupi, 2014). Remarkably, through the use of the large and highly potent agonist UK-432097, the structure of A_{2A}R (fused to T4L) was solved in an active-like conformation without the requirement of additional stabilization (Xu et al., 2011). However, despite these structures reveal the interactions involved in agonist binding, they are unlikely to represent the fully activated G protein-binding conformation comparable to the structure of active rhodopsin bound to the C-terminal G protein peptide (Choe et al., 2011; Deupi et al., 2012), or the structure of active β_2AR bound to the $G\alpha_s$ protein (Rasmussen et al., 2011b).

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Finally, site-directed mutagenesis for covalent trapping of ligand-receptor complexes has also been used in the elucidation of the structures of the murine μ opioid receptor in complex with a morphinan antagonist (Manglik et al., 2012), and CXCR4 in complex with a viral chemokine (Qin et al., 2015), among others (Weichert and Gmeiner, 2015).

While the mutations reported above have no effect on the structure, it has been observed that mutagenesis may sometimes have a dramatic effect. For instance, the structures of three similar constructs of CXCR4 with and without a T240$^{6.36}$P mutation showed that it caused the disruption of a short section of helix 6, effectively uncoupling ligand binding from receptor activation (Wu et al., 2010). To date, structures of fourteen different receptors have been obtained using mutagenesis to stabilize a particular conformation of the receptor (Ghosh et al., 2015) (Supplementary Table 1).

Chimeric constructs

The most successful protein engineering technique to obtain crystalizable GPCRs has been the creation of chimeric constructs in which the receptors are genetically fused to a soluble protein. In such chimeric receptors, the fusion protein replaces an intracellular loop or is added as a tag to the N-terminus (Fenalti et al., 2015; Wang et al., 2013b; Wu et al., 2014), providing a large and stable domain that favors the formation of crystal contacts. The fusion proteins themselves are highly crystalizable and feature N- and C- termini at the right distance for their insertion into GPCR loops without resulting in a significant distortion of the transmembrane bundle (Chun et al., 2012). These fusion proteins are typically T4 lysozyme (T4L) (Cherezov et al., 2007; Jaakola et al., 2008; Wu et al., 2010) or a thermostabilized apocytochrome (b_{562}RIL) (Chun et al., 2012; Liu et al., 2012; Zhang et al., 2014), but more recently other proteins have also been used successfully as ICL3 fusions, such as the catalytic domain of *Pyrococcus abyssi* glycogen synthase in the OX_2 receptor (Yin et al., 2015), or rubredoxin in CCR5 (Tan et al., 2013) and the $P2Y_1$ receptor (Zhang et al., 2015a) (Supplementary Table 1).

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

While such dramatic protein engineering is still required in many cases to obtain crystalizable GPCR constructs, it naturally raises questions about the reliability of the obtained structures. To address this concern, many efforts have been made to determine the effect of protein fusions on receptor activity. In the β_2AR and the A_{2A}R, T4L insertion into ICL3 does not appear to constrict conformational changes associated to activation, as this fusion results in a higher affinity for agonists, a property associated with constitutive activity (Jaakola et al., 2008; Rosenbaum et al., 2007). Furthermore, while this fusion impedes coupling to the G protein, a fluorescence-based assay detected conformational changes in TM6 of the β_2AR consistent with agonist-induced movements upon activation (Rosenbaum et al., 2007). It has been suggested that, in these cases, the fusion results in changes in the cytoplasmic side of TM6 that perturb an intramolecular ionic interaction (ionic lock) that stabilizes the inactive state of some Class A GPCRs (Chien et al., 2010; Dore et al., 2011; Preininger et al., 2013). Supporting this idea, a structure of A_{2A}R solved with the same ligand but without an ICL3 fusion did indeed show the presence of the ionic interaction (Dore et al., 2011). On the other hand, replacement of T4L by b_{562}RIL in ICL3 produced a structure closer in conformation to the inactive state, although the ionic interaction was not fully formed (Liu et al., 2012). The insertion of b_{562}RIL into ICL3 of the smoothened receptor has also been proposed as a reason for the lack of structural rearrangements at the cytoplasmic surface upon agonist binding (Wang et al., 2013b). Finally, comparison of the murine δ opioid receptor structure solved with an ICL3 T4L fusion (Granier et al., 2012), and the human δ opioid receptor with an N-terminal b_{562}RIL fusion (Fenalti et al., 2014), shows a high degree of structural similarity with the main deviations occurring proximal to the sites of fusion.

In summary, creation of fusion chimeras has proven a very successful strategy, greatly accelerating our knowledge of GPCR structure across a wide range of receptors. However, it is

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

important to keep in mind that the use of this technique may introduce some artifacts in the obtained structures.

Co-crystallization tools

An additional strategy to facilitate structure determination of GPCRs is the use of crystallization chaperones to form stable complexes and/or trap the receptor in a given conformation. For instance, monoclonal antibody fragments (Fabs) have been used to determine the structure of β_2AR (Rasmussen et al., 2007) and A_{2A}R (Hino et al., 2012) in the presence of inverse agonists. Like in the fusion strategy, the Fabs create an extended hydrophilic surface area to mediate crystal contacts, and reduce the flexibility at the receptor surface. Fab5, directed against β_2AR, binds to a structural epitope on ICL3 (one of the most structurally dynamic regions in many GPCRs), but does not affect the ligand-binding properties of the receptor (Day et al., 2007). However, the crystal structure of the β_2AR-Fab5 complex bound to an inverse agonist showed an apparent intermediate conformational state that may have been influenced by Fab-mediated crystal packing constraints. For the A_{2A}R, on the other hand, Fab2838 is conformationally selective for the antagonist-bound state, abrogating agonist binding while retaining wild type antagonist pharmacology (Hino et al., 2012). Fab2838 binds to a similar pocket on the cytoplasmic side of the receptor to that used by the C-terminal α-helix of $G\alpha_s$ upon activation. However, Fab2838 binding results in an inactivated receptor by locking TM3, TM6 and TM7 together.

The development of G protein mimetics has enabled to capture the active state of certain GPCRs. Specifically, nanobodies (Nb), the recombinant antigen-binding domain of camelid heavy chain antibodies, are only a quarter of the size of conventional Fab fragments and very efficient at mimicking G proteins (Steyaert and Kobilka, 2011). For instance, immunization of a llama with purified agonist-bound β_2AR allowed to obtain a nanobody (Nb80) that recognized specifically the active state of the receptor. Interestingly, Nb80 shows similar attributes to $G\alpha_s$

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

with respect to its influence on agonist affinities and on the conformational changes stabilized in the receptor (Rasmussen et al., 2011a). Importantly, while the β_2AR-T4L fusion construct was unable to activate G protein signaling, presumably due to steric clashes between T4L and the G protein (Rasmussen et al., 2007), Nb80 is small enough to stabilize the active state of the receptor even with T4L inserted into ICL3 (Rasmussen et al., 2011a). Further engineering of Nb80 resulted in the creation of a higher affinity nanobody (Nb6B9) that allowed to solve the structure of β_2AR in the presence of a range of agonists, including some with low affinity (Ring et al., 2013). These structures revealed that different agonists can stabilize similar conformational changes during receptor activation using a different set of ligand-receptor interactions. Nanobodies have also been used to crystallize the M_2 muscarinic acetylcholine receptor in an active conformation with and without a positive allosteric modulator (Kruse et al., 2013a), and the constitutively active viral GPCR US28 in complex with the human chemokine domain CX3CL1 (Burg et al., 2015). But the most important application of nanobodies to GPCR structure determination has been their use in the elucidation of the complex between β_2AR and the $G\alpha_s$ protein (Rasmussen et al., 2011b). This impressive feat, however, was only possible through the combined use of many of the techniques discussed above (Supplementary Table 1).

Crystallization techniques

GPCRs, as integral membrane proteins, must be extracted from the lipid bilayer using detergents prior to their purification. Outside the membrane environment GPCRs are typically unstable, and tend to unfold unless they are stabilized by ligand binding, protein engineering or mutagenesis, as discussed above. Classical approaches to crystallography using the vapor diffusion technique are usually unsuitable for many GPCRs because they have a relatively small hydrophilic surface area and the long-chain detergents used for solubilization form large micelles that generally occlude the polar surfaces available to form crystal contacts. The

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

dichotomy of requiring short-chained detergents with small micelles to expose the receptor for crystallogenesis and the need to maintain folded functional receptor in such harsh conditions reduces the chances of success. Indeed, only four non-rhodopsin GPCRs have been solved using vapor diffusion crystallography, and these required either thermal stabilization (Dore et al., 2011; Egloff et al., 2014; Warne et al., 2008) or the use of co-crystallization tools such as Fabs (Hino et al., 2012) (Supplementary Table 1). However, whether this is because vapor diffusion crystallography has fallen out of fashion, or if it is truly inhibitory is still debatable. New detergents such as neopentyl glycols (Chae et al., 2010), ganglio-tripod amphiphiles (Chae et al., 2014), and steroidal amphiphiles (Lee et al., 2013), some of which have the unique property of stabilizing non-denatured GPCRs when diluted below their critical micelle concentration, may aid further crystallization efforts.

Nevertheless, the majority of GPCR structures have been solved using lipidic cubic phase (LCP) crystallization (for a comprehensive review, see (Caffrey, 2015)). In this method, the receptor is crystallized in a lipidic environment, instead of from a detergent solution. The membrane-like environment of LCP is more stabilizing than detergents, and therefore is advantageous for crystallizing unstable receptors. However due to the nature of the lipidic phase, crystal nucleation is slower compared to vapor diffusion and crystals typically take longer to grow. Consequently LCP crystals tend to be small (10-30 μm) and pose challenges to their isolation and to obtain diffraction (Liu et al., 2014). These drawbacks have been overcome by the development of microfocus beamlines and high-energy synchrotron sources, which have enabled data to be extracted from microcrystals. However, the radiation damage caused to such small crystals makes collection of high-resolution data difficult. The recent emergence of X-ray free electron lasers (XFELs) is allowing these challenges to be overcome. XFELs are capable of generating ultrafast pulses of X-rays at intensities several orders of magnitude above the brightest synchrotron sources (Chapman et al., 2011; Spence et al., 2012), enabling data

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

collection from small protein crystals before the sample is damaged (and, eventually destroyed) by the power of the beam. Hence, the crystals do not require cryoprotection and data collection can be carried out at room temperature. By supplying a continuous stream of crystals to the XFEL beam, a dataset can be compiled from hundreds of thousands of diffraction images, which forms the basis of serial femtosecond crystallography (SFX). This method was used to determine the structure of 5-HT_{2B}R using crystals formed in LCP (Liu et al., 2013). Comparison of this structure to that of the same protein determined using traditional X-ray crystallography techniques (Wacker et al., 2013) showed a remarkable agreement, with only small differences in the loops, termini, and a few side chain rotamers. As SFX allows diffraction data to be obtained at room temperature and in a lipidic environment, it can be argued that it provides a truer depiction of the native state of a receptor (Liu et al., 2013). The use of SFX in structure determination of GPCRs has since been validated in additional receptors, including smoothened (Weierstall et al., 2014), the δ opioid receptor (Fenalti et al., 2015) and the AT_1 angiotensin receptor (Zhang et al., 2015b). Further developments of this technique may enable the study of receptor kinetics within crystals and capture short-lived conformational states during activation (Barty et al., 2013; Kern et al., 2013).

Non-ligand molecules present in structures

It is sometimes overlooked that the natural environment of the receptor *in vivo*, including the lipid bilayer and its components, affects GPCR activation and signaling. Lipids can affect the function of membrane proteins in a number of ways, either through direct interaction, or by altering the physical properties of the membrane environment, such as bilayer thickness, curvature, or lateral pressure (Oates and Watts, 2011). In addition, many GPCRs have been shown to localize in certain regions of the cell membrane through association with lipid rafts, mediated by interactions with the palmitoylated cysteines of helix 8 (Chini and Parenti, 2004).

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Lipid molecules and cholesterol have been observed in the crystal structures of many GPCRs (Cherezov et al., 2007; Jaakola et al., 2008; Manglik et al., 2012; Wu et al., 2014; Zhang et al., 2014). Due to the nature of protein crystals, the presence of a lipid molecule suggests a favorable interaction between the lipid and the receptor, although it is not clear whether these interactions are benign associations, or if the interaction indicates modulation of receptor function. In a structure of inactive β_2AR, cholesterol was a prerequisite for crystallogenesis and, thus, was added in excess during crystallization (Cherezov et al., 2007). Accordingly, cholesterol was found to mediate parallel associations of receptors in the crystal lattice. A subsequent structure with a different crystal lattice showed the same cholesterol binding sites located in a shallow groove between TM1 and 4, but in this case they did not participate in crystal contacts (Hanson et al., 2008). This observation led to the discovery of a putative cholesterol binding consensus motif present in almost half of all Class A GPCRs (Hanson et al., 2008). Other cholesterol binding sites suggested from modeling studies have been later observed in the structures of A_{2A}R and μ opioid receptor (Cang et al., 2013; Jaakola et al., 2008; Manglik et al., 2012). Interestingly, an additional cholesterol binding site was predicted in the β_2AR at the top of TM1 and 7 that could potentially influence ligand binding (Cang et al., 2013). Recently, a cholesterol molecule was found in the structure of $P2Y_{12}$R in a similar position (Zhang et al., 2014). Depletion of cholesterol from lipid bilayers has been shown to change the biased signaling properties of β_2AR, from signaling through $G\alpha_s/G\alpha_i$ to predominantly through $G\alpha_s$ (Cherezov et al., 2007; Xiang et al., 2002). However, it is uncertain whether this effect stems from a direct interaction between cholesterol and receptor, or from co-localization effects of cholesterol on the receptor with specific G proteins (Pontier et al., 2008). In addition, β_2AR has been shown to form homodimers *in vivo* (Angers et al., 2000), and studies have shown that cholesterol may also affect the way β_2AR molecules assemble into dimers, leading to the intriguing idea that cholesterol could modulate receptor function through changing the way GPCRs associate in the bilayer (Prasanna et al., 2014). However, if one entertains the idea of

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

molecules such as cholesterol acting as modulators of receptor activity, one must also consider that absence of these molecules in receptor structures where their natural environment is cholesterol rich may suggest an incomplete representation of the receptor.

GPCR X-ray structure determination

The previous section outlines the recent techniques and approaches that have been used to obtain stable and purified GPCRs and grow crystals suitable to obtain X-ray diffraction data in synchrotrons or XFELs. In this section, we describe briefly how such diffraction data are translated into the final 3D structures that allow researchers to map the interactions between drugs and receptors. Comprehensive introductions to macromolecular crystallography for a general scientific audience can be found in excellent books on the subject (Rhodes, 2006; Rupp, 2009).

X-ray crystallography

In a nutshell, structural determination by X-ray crystallography involves measuring the directions and intensities of X-rays diffracted by the electron clouds of the molecules in the crystal, and using computer software to reconstruct a map of the electron density. In an iterative refinement process, the crystallographer builds a 3D model of the protein that fits the electron density while being consistent with the prior knowledge of general protein structure and on the protein that has been crystallized. It is important to keep in mind that the final 3D structure, which is all that most of non-crystallographers see, is a model representing the best fit of the protein atoms to the electron density map. Thus, in order to assess the quality of this model, it is useful for users of these models to be familiar with some basic concepts of structure determination by X-ray crystallography. This section is not meant to be a comprehensive guide to X-ray crystallography; rather we wish to provide just a brief conceptual overview, highlighting a few key points in the process along the way that are important for the critical use of crystallographic structure models.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Crystal symmetry

In a protein crystal, the molecules are arranged in an array of repeating elements called unit cells. The unit cells forming the crystal and the contents of the unit cells themselves are held together mainly by protein-protein contacts, but the molecules are loosely packed and the solvent content in the crystal is very high (around 60% in GPCR crystal structures). The contacts mediating the crystal formation are generally weak and are not always biologically relevant. Protein-protein interactions in crystal structures should always be carefully evaluated with additional experimental data before drawing conclusions about their physiological relevance, especially if the interface is small or non-conserved. As an example, in Class A GPCR structures, contacts between transmembrane regions observed in the crystal structures so far, may in some instances, resemble the biological interfaces that are hypothesized to be present in GPCR oligomers, but are not definitive proof of physiologically relevant dimeric structures (Duarte et al., 2013).

When crystals are exposed to a beam of X-rays, the lattice-like array of unit cells causes the scattered radiation to be amplified into discrete spots (reflections) at specific angles relative to the incoming X-ray beam (Bragg and Bragg, 1913) resulting in a diffraction pattern measured by a detector (Figure 2). The diffracted X-rays are waves and thus characterized by three parameters: amplitude, frequency, and phase. These parameters, together with the arrangement of the spots relative to each other on the detector, provide the details necessary to construct an “image” of the unit cell. The amplitude is recorded as intensity values on the detector pixels; the square root of the intensity value is simply the amplitude of the wave. The frequency of each reflection is related to the angle at which the reflection exits the crystal. Conceptually, the exact angle that the reflection exits the crystal is related to the resolution of the diffracted beam; higher angle reflections (recorded farther from the center of the detector) arise from diffraction of finer slices of the unit cell and thus bring higher resolution information

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

toward the reconstruction of its contents. The phase of each reflection is, quite unfortunately, not recorded by the detector and therefore must be estimated in some way. Finally, the spacing and relative arrangement of the reflections captured by the detector are related to the size and symmetry of the unit cell. The distance between reflections measured along diffraction axes are inversely proportional to the dimensions of the unit cell; e.g. closely spaced reflections measure out a large unit cell axis and *vice versa*. The symmetry of the molecular arrangements within the unit cell is also reflected in the diffraction data, which aids in the calculations, as symmetry-related reflections can be averaged together to increase the signal-to-noise ratio of the dataset.

Data Collection

From a three-dimensional crystal, reflections are recorded as the crystal is rotated in order to capture a complete dataset. The recorded reflections on individual frames are then indexed with a coordinate value *h,k,l* (the so-called Miller index), intensities are integrated, and finally the intensity values of the frames are scaled together, adjusted for systematic errors, and symmetry-related or multiply-measured spots are merged. The result is a table containing several thousand or more unique reflections, with intensity (I), standard deviation (σ), and Miller index (*h,k,l*). As stated previously, each reflection is a discrete wave resulting from the periodic scattering of the contents of the unit cell. The sum of these individual waves add together to produce a complex three-dimensional form, which is the image of the electron density of the unit cell. In practice, reflections (amplitudes plus phase estimate) are combined together by Fourier synthesis (simulating the function of the lens in a microscope) to produce the image of the unit cell. To think about it another way, the electron density distribution of the unit cell can be imagined as a complex three dimensional waveform (high density peaks where protein atoms are crystallized in place, low density where disordered solvent fills the spaces between) and the diffracted reflections are the individual waves that can be added back together to reconstruct the image of this density.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

The quality of the crystal largely determines the quality of the diffraction data and, in turn, the accuracy of the structural model. Good quality protein crystals that produce well-resolved high-resolution diffraction can be difficult to achieve, particularly for highly dynamic membrane proteins such as GPCRs. When crystals of such proteins can be coaxed into existence they are often small, resulting in faint diffraction signals, and/or are insufficiently crystalline, resulting in poor diffraction. If there is too much variation between unit cells and the protein molecules are not well ordered, the diffraction breaks down and the reflection data becomes smeared, weak, and ultimately absent. For instance, crystals may exhibit a large degree of "mosaicity" (i.e. are formed by a mosaic distribution of differently oriented blocks) or may be “twinned” (i.e. have crystalline blocks specifically oriented which give rise to overlapping diffraction). In such cases, the measured diffraction pattern is much more difficult to analyze. As such, recent technological advances in crystallography have allowed researchers to more easily overcome these hurdles. As described in previous sections, extensive genetic modification of the receptor with thermostabilizing mutations, splicing in fusion proteins such as T4L, or by co-crystallization with Fab proteins or nanobodies, has allowed GPCRs to be crystallized with sufficient quality for diffraction studies. Furthermore, developments in synchrotron microcrystallography and XFEL instrumentation have made key contributions by providing very intense focused beams of radiation that allow diffraction data to be obtained from very small crystals.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Initial phasing

As mentioned previously, in order to reconstruct the electron density map of the unit cell from diffraction data, the phases of the diffracted X-rays are required. There are several techniques to obtain this information. In the first high-resolution X-ray crystal structure of a GPCR, bovine rhodopsin, phasing information was obtained by a technique called multi-wavelength anomalous diffraction (MAD) (Hendrickson et al., 1985) using mercury-soaked crystals (Palczewski et al.,

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

2000). When heavy atoms, such as mercury are incorporated into protein crystals, at certain wavelengths the diffraction pattern changes slightly but significantly in a very precise way. This change in diffraction can be exploited to extract a limited set of phase information, allowing the reconstruction of electron density. With only one recent exception, smoothened receptor (Wang et al., 2014), which was also solved by soaking crystals with heavy atom solutions and employing anomalous diffraction measurements, all of the remaining GPCR crystal structures (over 120 to date) have been solved by a technique called molecular replacement (MR). With MR, basically, the phases are calculated from a known protein structure that is expected to be similar, and applied together with the intensities of the diffraction data to generate an initial electron density map. If the structure of the MR model is similar enough to the unknown structure, a reasonable electron density map is produced that can be further refined. With low-resolution and weak diffraction data, there is significant risk to introduce “model bias” into the calculated maps, whereby the model phase information dominates the calculation, resulting in essentially an electron density map of the MR model only. A clue that the initial MR-derived map is of sufficient quality is to see if it contains new features not present in the MR model, e.g. truncated regions or side chains. With this method, the structure of the β_2AR (Rosenbaum et al., 2007)(the second GPCR structure to be solved, after rhodopsin), was obtained using rhodopsin as a phase "template" to reconstruct the electron density. In a similar fashion, the next structure, β_1AR (Warne et al., 2008), was solved using β_2AR as a template, and so on.

Model building

The final step corresponds to building a 3D molecular model that fits in the electron density (Figure 2). As most GPCRs structures start from an MR-derived map, the MR model itself is generally a good starting point to begin building and refinement. High sequence homology within GPCR transmembrane regions provides a convenient base upon which the structure of the crystallized receptor can be modeled. After a set of changes to the model are made,

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

computer programs apply energy minimization algorithms to more finely fit the atomic coordinates and displacement factors (B-factors) to the electron density, restraining the model to known chemical properties of amino acids (e.g. length and angle of covalent bonds, Van der Waals contact distances, etc.) and to expected protein structure properties (e.g. peptide backbone dihedral restraints, secondary structure hydrogen bond restraints, etc.). The degree to which restraints are applied depends on the quality of the diffraction data. Lower resolution data demand stronger restraints, while higher resolution data can be “freed” a bit more, letting the higher certainty of the electron density guide atom placement. Thus, low-resolution structures tend to have geometrical statistics with small variance (low root mean square deviation; RMSD), closer to an average “ideal” value than higher-resolution structures.

Crystal structure quality metrics

When we visualize the structure of a protein solved by X-ray crystallography, we are looking at a molecular model that has been built to fit as well as possible into an electron density map, using advanced computational methods and, in many cases, some assumptions about missing data. In order to effectively use these models it is important to be familiar with some of the existing metrics available to validate their quality. Crystal structure validation is absolutely essential in the process of interpreting and adapting models for further research applications (Read et al., 2011). In this process, we want to understand the quality of the data that generated the model, the stereochemical quality of the model itself, and how well the model actually fits the data, both on a global basis and the local fit of residues and ligands.

Perhaps the most familiar metric to non-crystallographers is resolution, which, as stated above, essentially reports the highest angle reflections recorded in the diffraction pattern. A reported resolution value of, for example, 2.5 Å, states that the diffraction dataset contains reflections that arose from the scattering of unit cell contents in 2.5 Å increments. Importantly, resolution in this context does not refer specifically to the precision in the position of the atoms in the

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

structure, rather it can be thought of as a measure of the "fineness" of the data which gets fed into the electron density calculation and should be regarded more as a metric of the data quality, not necessarily model quality. The precision in atom positions is reported as an estimated coordinate error value (ESD or ESU) and is typically between 1/5 - 1/10 of the resolution (Brunger, 1997); i.e. a structure at a resolution of 3 Å provides a precision in the position of atoms within 0.6 - 0.3 Å, depending on the quality of the data. The average resolution of the solved GPCR structures is 3 Å, which practically speaking, allows one to visualize in the electron density the basic contours of amino acid side chains and ligands. The highest resolution obtained for a GPCR is 1.8 Å (Liu et al., 2012), which allows significantly more detail to be modeled. This structure was able to include 57 ordered water molecules and a sodium ion inside the receptor, plus two cholesterol and 23 ordered lipid molecules (Figure 3).

Another measure of crystal structure quality is the R-factor, which measures the agreement between the recorded diffraction data and the derived model built into the electron density. Thus, the R-factor quantifies how well the refined structure predicts the observed data. An R-factor of zero indicates perfect agreement, and an R-factor of approximately 0.54 indicates randomness, or essentially no agreement. Crystallographers realistically aim for R-factors of about 0.2 or less, if possible. In the available GPCR structures, this value ranges roughly from 0.2 to 0.4, with an average of 0.27. R-factors are reported as two separate values: R-work and R-free. R-work is the R-factor calculated from all of the data used in the refinement process. However, this statistic can become erroneously low if the model is "overfit" to the data. When refining against low-resolution datasets, if too few restraints are placed on the model, the refinement algorithms will, in a sense, take too many liberties in adjusting the coordinates and B-factors so that the calculated diffraction pattern matches the observed diffraction pattern. Thus, the R-work is not an independent and unbiased statistical indicator of model quality, and is always in danger of being artificially low. To guard against this, the R-free statistic was

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

introduced. R-free is calculated from a subset of data that has been randomly selected at the very start of the process and withheld from refinement throughout the entire process. Therefore, in theory it should be “free” from model bias. If the refinement algorithms are performing properly and the model truly reflects the observed data, both R-work and R-free should be in agreement. If the model is overfit, R-work and R-free will diverge. In practice though, the R-work and R-free values may differ by about 5 percentage points (e.g. 0.20 and 0.25, respectively) as it is nearly impossible to completely remove all bias from the procedure. Incidentally, however, if the values are too similar, this *may* indicate a bias in R-free, which can arise with careless application of molecular replacement phasing or if the R-free test reflections are correlated to R-work reflections by non-crystallographic symmetry elements.

While resolution, R-work, and R-free are global measures of the quality of the data and structure, temperature factors (also known as B-factors or atomic displacement factors - ADPs) and occupancies are local descriptors at the atomic level. In essence, B-factors are a measure of how smeared out the electron density is for an atom, and they provide some insights into the local disorder of the molecules. Generally, loop regions or long amino acid side chains have a higher freedom of movement, which can be thought of as a certain "blurring" of the atom in space, which translates into a higher B-factor (Figure 4, left panel). While it is tempting to equate high B-factors with highly dynamic regions, this interpretation should be done with caution, as it is not always the case. High B-factors can arise from either dynamic disorder or static disorder. Because the electron density map is an average over all unit cells contributing to the diffraction data, if atoms are locked into the crystal in slightly varying locations in each unit cell (static disorder), the electron density for that atom will be smeared out and the B-factor will be high. Alternatively, if the atom is fluctuating within the unit cell (dynamic disorder), the electron density will be similarly smeared out and the factor will also be high. The B-factor simply tells you how well the atom position is defined in the crystal. Generally, the mean B-

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

factor of a crystal structure is correlated with the resolution of the dataset. Higher resolution diffraction arises from better ordered structures and a higher degree of crystallinity across all unit cells, which means that atom positions are less variable in the crystal and give rise to well-defined electron density maps with low positional uncertainty and low B-factors. For protein modeling and interpretation purposes, B-factors can be considered as a metric of uncertainty for the coordinates. High B-factors equate with high positional uncertainty. Thus regions with high B-factors can be interpreted as being less well-defined than regions with low B-factors. In molecular dynamics studies, for example, this can be related to the degree of fluctuation observed during a simulation run, especially for regions of the model where the crystal structure is not clearly constrained by crystal contacts or other artifacts of the crystal environment (e.g. bound buffer components, etc…). As such, a low root mean square fluctuation (RMSF) value in a molecular dynamics simulation might not always recapitulate a low B-factor observed in the crystal structure due to these environmental differences.

On the other hand, occupancies are an estimate of the fraction of the diffracting molecules in which the atom occupies the position specified in the model. For moderate- to low-resolution datasets, occupancy and B-factor are nearly impossible to refine independently. Thus, fractional occupancies are calculated and refined only for very specific cases where the reduced occupancy can be clearly supported by the electron density. This generally only occurs for highly electron-dense scatters (e.g. heavy atoms such as selenium or mercury, sometimes well-ordered aromatic ligands), or for alternate conformations of amino acid side chains and loops. For instance, a side chain may exhibit two conformations, both supported by electron density. Two side chains can be built into the density, each with a fractional occupancy that adds to 1.0.

When evaluating a crystal structure, geometrical statistics should also be considered to understand how closely the model conforms to an accurate protein structure (Read et al., 2011). Covalent bond lengths and angles should be consistent with known chemical parameters, and

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Van der Waals contacts should be within allowed distances, accounting for the placement of hydrogen atoms in the structure. Of course, chirality of amino acids and ligands must be correct. The peptide backbone should generally contain planar peptide bonds, and the torsion angles phi (φ) and psi (ψ) should generally conform to expected values from updated Ramachandran plots. Side chain torsion angles chi (χ) should also be evaluated for outliers. A good quality average protein structure will not have outliers in any of these measurements, by definition. It must be stressed, however, that quite often protein structures do, in fact, contain geometrical outlier values relative to the average structure as they may contain regions wound-up energetically for some functional process. However, outliers in a good model must be supported by electron density. Going back to what was stated in the beginning of this section; this is why low-resolution structures should have better statistics than higher-resolution structures. At low-resolution, stronger geometrical restraints are needed to prevent overfitting of the model to the map. At higher resolutions, the data is stronger to more confidently fit atoms to the electron density with less restraint.

The discerning user of protein structure models should always concern themselves with these measures. While most of these values can be easily inspected in the corresponding entries in the PDB web site (e.g. the full validation report produced by the PDB), or in the PDB coordinate file itself (which is simply a text file that can be opened with any text editor) a better approach is to use an analysis program such as MolProbity (http://molprobity.biochem.duke.edu/) (Chen et al., 2010b). MolProbity can be run directly from a web interface and generates a thorough analysis of the protein structure. PDB files can be fetched directly from the PDB, or custom files can be uploaded. The program will add hydrogen atoms, identify if the side chains of Asn, Gln, or His should be flipped based on hydrogen bonding patterns (a common mistake in protein models), and then perform an all-atom contact and geometry analysis. Several clear tutorials are available at the website to assist the first-time user. A MolProbity analysis outputs a so-

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

called “multi-criterion chart” which gives a residue-by-residue list of scores on several geometrical indicators of model quality including all-atom contacts (clash score), Ramachandran score, Cβ deviations, side chain rotamer outliers, and general bond length and angle deviations. While all these scores are important in evaluating a model, particular emphasis should be placed first on the all-atom contact clash score. This is a measure of the van der Waals overlaps, which must be minimized to the greatest extent possible in a well-built protein structure no matter the resolution or quality of the data. The Ramachandran analysis provides a measure of the peptide backbone phi and psi angles and how they relate to a benchmark population of high-resolution crystal structures. Most protein structures should not contain Ramachandran outliers, but if they do, the presence of the outlier should be justified by strong electron density. Again, the lower the quality of the diffraction data (the harder the electron density maps are to interpret), the better the geometry statistics should be. Higher resolution data with strong electron density can provide evidence of deviations from an ideal protein structure, whereas poor data cannot justify outliers. Molprobity also outputs kinemage files and can be operated with the program Coot to provide a more graphical presentation of flagged outlier regions for closer inspection.

B-factors and occupancies can be visualized in a molecular graphics program (e.g. PyMOL (Schrodinger, 2010)), by coloring each residue according to these values (Figure 4). Coloring by B-factors will give an idea of which regions of the receptor are more disordered and probably have weaker electron density, while coloring by occupancy will highlight residues that were modeled with alternate conformations.

Analysis of electron density maps

For a more detailed analysis, electron density should be inspected. Maps calculated from deposited datasets can be obtained easily for most structures, in lieu of sophisticated crystallographic software packages, from the Electron Density Server (EDS,

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

http://eds.bmc.uu.se/eds/) (Kleywegt et al., 2004). This web service provides a summary analysis of the data and model, and can generate analytical plots for evaluating the electron density data. Particularly illustrative is the real-space R-factor plot, showing a residue-by-residue calculation of the fit of the model to the map. This can highlight potentially troubling areas of the model that warrant further inspection. Of course, the best way to understand how the model fits the data is to look at the map itself. For this, the EDS can calculate so-called sigma-A weighted maps, which are the most common type of maps crystallographers use for model building. Two useful flavors of these maps are the “standard” 2mFo-DFc map, and the “difference” map, mFo-DFc. The standard 2mFo-DFc map shows, essentially, the experimental electron density map into which the crystallographer has built the model. Exploring this map is useful to verify the overall quality of the map, and how different regions of the model may be built into weaker or stronger density. The difference map mFo-DFc shows the residual electron density after subtracting the calculated model density from the experimental (observed) density. This is useful to highlight errors in the model. Substantial positive electron density peaks suggest an incomplete model, e.g. missing atoms, while negative electron density peaks show areas where the model is not supported by experimental density. A useful derivative of the difference map is the “omit” map, whereby a small region of the model is purposely deleted, then the mFo-DFc map is recalculated (sometimes after performing a bit of simulating annealing dynamics to reduce model bias - this is then commonly referred to as an “SA-omit” map). Omit maps are routinely used to validate the placement of ligands in crystal structures. If the omitted region is highlighted by strong positive density peaks, this is good evidence that the model is correct. However, if the difference density for the omitted region is weak and uninterpretable, the model is probably wrong. The recent versions of most popular molecular graphics programs (e.g PyMOL (Schrodinger, 2010), Coot (Emsley et al., 2010), or CCP4MG (McNicholas et al., 2011)) provide well-documented functionality to fetch, calculate, and display electron density

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

maps directly from data deposited in the PDB or via EDS (Figure 3), though omit maps must be recalculated using a crystallographic refinement program.

With a small amount of effort, the strengths and weaknesses of crystallographic data can be assessed to determine how to process the model for downstream applications. While overall quality factors such as resolution and R-free are important metrics to consider, it is equally important, if not more so, to critically evaluate crystal structures at a finer level, down to the local environments of individual residues and ligands, in the context of the electron density maps. Not all parts of the model built by the crystallographer are equally supported by the diffraction data, thus interpretations from crystal structures *require* these density-driven inspections. This imperative is the central thesis of a recent review by Lamb, *et al.* (Lamb et al., 2015). For flexible membrane proteins such as GPCRs, map and model validation is especially relevant as most cases have only moderate overall resolution and the electron density quality can vary widely between the transmembrane domains and the solvent-exposed loops and ligand binding sites.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Complementing GPCR structural chemogenomics with molecular pharmacology data

Decades of intense research on GPCRs have produced vast amounts of molecular pharmacology data, available, for instance, at the IUPHAR/BPS Guide to Pharmacology (Pawson et al., 2014). As discussed above, the field of structural biology is catching up, and curated and up-to-date structural data can be obtained, for instance, at the GPCRDB (Isberg et al., 2014). The combination of molecular pharmacology and structural data provides a powerful lens to gain new insights into the mechanistic details of GPCR function. For instance, a systematic analysis of the functional data available for crystallized ligand-receptor complexes is crucial to elucidate the molecular determinants of ligand binding, including details of structure-activity relationships (SAR) which, in turn, can be used to extrapolate the available information to ligands and receptors that are related to known structures, but have not yet been crystallized (Kooistra et al., 2013).

In this section we present an example of how structural data can complement chemogenomics studies. Supplementary Table 2 displays ligand and mutagenesis data for 27 crystallized GPCRs. Ligand data include chemical structure, name and PDBid of co-crystallized ligands, plus the total number of small-molecule ligands (60 heavy atoms or fewer) with binding affinity (IC50/Ki) or functional potency (IC50/EC50) of at least 10 μM for each receptor, identified using the ChEMBLdb (Bento et al., 2014); below this number and in brackets, we specify how many of these ligands are similar (ECFP-4 Tanimoto similarity ≥ 0.4) to the co-crystallized compound. Mutagenesis data have been extracted from the GPCRDB and recent literature, and include the number of ligands used in mutagenesis studies to assess ligand binding and activity (and how many of those are peptide ligands), number of mutants and mutated positions in these studies,

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

and number of investigated combinations of mutants and ligands. Such mutagenesis data provides a quantitative measure of the amount of information available for each receptor that can be used in the process of drug discovery. The data in Supplementary Table 2 is summarized graphically in panel A of Figure 5.

About half of the analyzed GPCR structures have been co-crystallized with ligands that are chemically similar (red bar segments in Figure 5, panel A top) to a substantial portion (≥10%) of the known ligands for that receptor (Figure 5 panel A top, green background). Thus, for these receptors, the binding mode of a relatively large number of ligands can be confidently predicted using computational techniques such as molecular docking. Remarkably, for the $P2Y_1$ and PAR1 receptors, the binding mode of about half of the known ligands (191 out of 131, and 236 out of 574) can in principle be modeled in a relatively straightforward manner (although small differences in ligand may affect overall ligand binding mode). Still, for most of these receptors the majority of compounds have different chemotypes than the co-crystallized ligands (Figure 5, panel A top, blue bar segments). In these cases, binding modes can be inferred by analyzing mutagenesis data covering many different mutants/residue positions and different ligands (Figure 5, panel A, middle and bottom). For instance, in $A_{2A}R$, β_2AR, δ opioid receptor and CXCR4 there is a large amount of mutagenesis data (number of ligands used, number of unique mutants and mutated positions, and combinations of mutants and ligands; see Supplementary Table 2, mutagenesis data, for details) that guide the prediction of binding modes of ligands that are not yet crystallized. Moreover, community wide GPCR structure modeling assessments (GPCRDOCK) (Kufareva et al., 2014; Kufareva et al., 2011; Michino et al., 2009) to predict the coordinates of the GPCR-ligand crystal structures have indicated that the best $A_{2A}R$ (Costanzi et al., 2009), D_3R (Obiol-Pardo et al., 2011), CXCR4 (Bhattacharya et al., 2013; Roumen et al., 2011), $5HT_{1B}$ (Rodriguez et al., 2014) and $5HT_{2B}$ models were constructed by the careful consideration of receptor mutation data. Conversely, PAR1, FFA1, and $P2Y_{12}$ receptors have

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

relatively little mutation data available (Supplementary Table 2, mutation data), and therefore, additional SAR data will be required to hypothesize binding modes for ligands that are dissimilar from the co-crystallized ligands. As an example, the $P2Y_1$ crystal structures in complex with MRS2500 and BPTU illustrate that ligands can target very different binding sites (Zhang et al., 2015a)(Figure 5, panel B).

At the opposite side of the spectrum, some GPCRs (κ opioid receptor, M_3R, D_3R, CRF_{1R} and H_1R) have been co-crystallized with ligands that cover 1% or less of the chemical space of known small drugs for each of these receptors. Clearly, co-crystallization of these receptors with chemically diverse ligands would greatly benefit the drug discovery efforts in these subfamilies. Interestingly, there is a substantial amount of mutagenesis data available for these receptors (Supplementary Table 2, mutation data; and Figure 5, panel A middle and bottom), which combined with structural information from related receptors, can facilitate the generation of reasonable docking models for other ligands. This is illustrated by successful crystal structure-based virtual screening studies in which novel potent D_3R (Carlsson et al., 2011; Lane et al., 2013; Vass et al., 2014), M_3R (Kruse et al., 2013b) and H_1R (de Graaf et al., 2011) ligands were identified, in the case of H_1R by so-called interaction fingerprint (IFP) scoring to select molecules that make similar contacts with the receptor binding site as the co-crystallized ligand (de Graaf et al., 2011). It should furthermore be noted that in some receptors (Figure 5 panel A top, starred), while there is a low number of known compounds similar to co-crystallized ligands, these nevertheless share some conserved substructures and/or a conserved shape/pharmacophore. For instance, in the $5\text{-}HT_{1B}$ receptor, while the total number of known ligands similar to the co-crystallized (dihydro-)ergotamine (Wacker et al., 2013; Wang et al., 2013a) is negligible, the tryptamine substructure is present in a large portion (24%) of know $5\text{-}HT_{1B}$ ligands. Similarly, the phosphonic acid group of the co-crystallized ML056[15] is present in 11% of all $S1P_1$ ligands. Also, the co-crystallized CP-376395 antagonist[19] shares

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

perpendicularly oriented N-heterocyclic and hydrophobic aromatic rings with most CRF_1R ligands, which can in principle facilitate modeling studies of other CRF_1R ligands to the CRF_1R binding site. Finally, the co-crystallized doxepin (H_1R), (R)-3-quinuclidinylbenzilate (BZ) (M_2R), and tiotropium (M_3R) ligands share an amine, with two aromatic rings oriented in a butterfly shape (Figure 5, panel C), with many other H_1R, M_2R, and M_3R ligands (Kooistra et al., 2013). The large number of mutation data available for many of these receptors furthermore facilitates experimentally guided modeling of other ligand binding modes (Supplementary Table 2 and Figure 5, panel A middle and bottom).

As a final note of caution, structural and chemogenomics data should only be combined when they refer to similar ligand binding modes. For instance, the majority of $mGlu_1R$ and $mGlu_5R$ ligands extracted from ChEMBLdb target the (orthosteric) extracellular Venus Fly Trap domains of class C GPCRs, while the co-crystallized FITM[20] and mavoglurant[21] ligands target the (allosteric) transmembrane domain.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Conclusion

The rapid emergence of structural data for GPCRs is significantly advancing our ability to generate accurate models of ligand-receptor complexes of unknown structure, interpret ligand binding structure-activity relationships, and extrapolate these relationships to related systems. Many GPCRs have been crystallized in complex with clinically relevant drugs, or close analogs of therapeutic compounds, providing a framework to understand the molecular basis for pharmacological activities. This, in turn, is fueling renewed efforts toward SBDD and an expanding search for drugs that act through poorly understood mechanisms, such as allosteric modulation and biased signaling. Furthermore, the proliferation of structures generates new starting points for molecular dynamics simulations, providing insights into the dynamics of ligand binding and receptor activation. These breakthroughs in GPCR crystallography have required the synergistic combination of numerous technical innovations in protein engineering, detergent chemistry, crystallization, and X-ray sources, breaking the intractability of this receptor family for structural studies. However, the techniques used to enable GPCR structure determination must not be overlooked when utilizing the structures for further research, as many GPCR structures have been heavily modified are far from the wild type protein. We must also remember that deposited coordinates in a PDB file are the crystallographers' best interpretation of a dataset; the truest picture of a crystal structure only emerges when the model is validated and viewed in the context of electron density maps. With this, the structural information from crystallography and other methods, such as NMR and EM, combined with the decades of functional data on ligands, mutants, and signaling complexes is bringing forth the next chapter in molecular pharmacology.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Authorship contribution

Participated in research design: Piscitelli, Kean, de Graaf, and Deupi.

Performed data analysis: Piscitelli, Kean, de Graaf, and Deupi.

Wrote or contributed to the writing of the manuscript: Piscitelli, Kean, de Graaf, and Deupi.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

References

Angers S, Salahpour A, Joly E, Hilairet S, Chelsky D, Dennis M and Bouvier M (2000) Detection of beta 2-adrenergic receptor dimerization in living cells using bioluminescence resonance energy transfer (BRET). *Proceedings of the National Academy of Sciences of the United States of America* **97**(7): 3684-3689.

Barty A, Kupper J and Chapman HN (2013) Molecular imaging using X-ray free-electron lasers. *Annu Rev Phys Chem* **64**: 415-435.

Bento AP, Gaulton A, Hersey A, Bellis LJ, Chambers J, Davies M, Kruger FA, Light Y, Mak L, McGlinchey S, Nowotka M, Papadatos G, Santos R and Overington JP (2014) The ChEMBL bioactivity database: an update. *Nucleic Acids Res* **42** (Database issue): D1083-1090.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN and Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* **28**(1): 235-242.

Bhattacharya S, Lam AR, Li H, Balaraman G, Niesen MJ and Vaidehi N (2013) Critical analysis of the successes and failures of homology models of G protein-coupled receptors. *Proteins* **81**(5): 729-739.

Bragg WH and Bragg WL (1913) The Reflection of X-rays by Crystals. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **88**(605): 428-438.

Brunger AT (1997) X-ray crystallography and NMR reveal complementary views of structure and dynamics. *Nat Struct Biol* **4 Suppl**: 862-865.

Burg JS, Ingram JR, Venkatakrishnan AJ, Jude KM, Dukkipati A, Feinberg EN, Angelini A, Waghray D, Dror RO, Ploegh HL and Garcia KC (2015) Structural biology. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science* **347**(6226): 1113-1117.

Caffrey M (2015) A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. *Acta Crystallogr F Struct Biol Commun* **71**(Pt 1): 3-18.

Cang X, Du Y, Mao Y, Wang Y, Yang H and Jiang H (2013) Mapping the functional binding sites of cholesterol in beta2-adrenergic receptor by long-time molecular dynamics simulations. *J Phys Chem B* **117**(4): 1085-1094.

Carlsson J, Coleman RG, Setola V, Irwin JJ, Fan H, Schlessinger A, Sali A, Roth BL and Shoichet BK (2011) Ligand discovery from a dopamine D3 receptor homology model and crystal structure. *Nat Chem Biol* **7**(11): 769-778.

Chae PS, Bae HE, Ehsan M, Hussain H and Kim JW (2014) New ganglio-tripod amphiphiles (TPAs) for membrane protein solubilization and stabilization: implications for detergent structure-property relationships. *Org Biomol Chem* **12**(42): 8480-8487.

Chae PS, Rasmussen SG, Rana RR, Gotfryd K, Chandra R, Goren MA, Kruse AC, Nurva S, Loland CJ, Pierre Y, Drew D, Popot JL, Picot D, Fox BG, Guan L, Gether U, Byrne B, Kobilka B and Gellman SH (2010) Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* **7**(12): 1003-1008.

Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, Hunter MS, Schulz J, DePonte DP, Weierstall U, Doak RB, Maia FR, Martin AV, Schlichting I, Lomb L, Coppola N, Shoeman RL, Epp SW, Hartmann R, Rolles D, Rudenko A, Foucar L, Kimmel N, Weidenspointner G, Holl P, Liang M, Barthelmess M, Caleman C, Boutet S, Bogan MJ, Krzywinski J, Bostedt C, Bajt S, Gumprecht L, Rudek B, Erk B, Schmidt C, Homke A, Reich C, Pietschner D, Struder L, Hauser G, Gorke H, Ullrich J, Herrmann S, Schaller G, Schopper F, Soltau H, Kuhnel KU, Messerschmidt M, Bozek JD, Hau-Riege

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

SP, Frank M, Hampton CY, Sierra RG, Starodub D, Williams GJ, Hajdu J, Timneanu N, Seibert MM, Andreasson J, Rocker A, Jonsson O, Svenda M, Stern S, Nass K, Andritschke R, Schroter CD, Krasniqi F, Bott M, Schmidt KE, Wang X, Grotjohann I, Holton JM, Barends TR, Neutze R, Marchesini S, Fromme R, Schorb S, Rupp D, Adolph M, Gorkhover T, Andersson I, Hirsemann H, Potdevin G, Graafsma H, Nilsson B and Spence JC (2011) Femtosecond X-ray protein nanocrystallography. *Nature* **470**(7332): 73-77.

Chen KY, Zhou F, Fryszczyn BG and Barth P (2012) Naturally evolved G protein-coupled receptors adopt metastable conformations. *Proceedings of the National Academy of Sciences of the United States of America* **109**(33): 13284-13289.

Chen Q, Miller LJ and Dong M (2010a) Role of N-linked glycosylation in biosynthesis, trafficking, and function of the human glucagon-like peptide 1 receptor. *Am J Physiol Endocrinol Metab* **299**(1): E62-68.

Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS and Richardson DC (2010b) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**(Pt 1): 12-21.

Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK and Stevens RC (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **318**(5854): 1258-1265.

Chien EY, Liu W, Zhao Q, Katritch V, Han GW, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V and Stevens RC (2010) Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **330**(6007): 1091-1095.

Chini B and Parenti M (2004) G-protein coupled receptors in lipid rafts and caveolae: how, when and why do they go there? *J Mol Endocrinol* **32**(2): 325-338.

Cho KH, Husri M, Amin A, Gotfryd K, Lee HJ, Go J, Kim JW, Loland CJ, Guan L, Byrne B and Chae PS (2015) Maltose neopentyl glycol-3 (MNG-3) analogues for membrane protein study. *Analyst* **140**(9): 3157-3163.

Choe HW, Kim YJ, Park JH, Morizumi T, Pai EF, Krauss N, Hofmann KP, Scheerer P and Ernst OP (2011) Crystal structure of metarhodopsin II. *Nature* **471**(7340): 651-655.

Chun E, Thompson AA, Liu W, Roth CB, Griffith MT, Katritch V, Kunken J, Xu F, Cherezov V, Hanson MA and Stevens RC (2012) Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* **20**(6): 967-976.

Congreve M, Dias JM and Marshall FH (2014) Structure-based drug design for g protein-coupled receptors. *Progress in medicinal chemistry* **53**: 1-63.

Costanzi S, Siegel J, Tikhonova IG and Jacobson KA (2009) Rhodopsin and the others: a historical perspective on structural studies of G protein-coupled receptors. *Curr Pharm Des* **15**(35): 3994-4002.

Crispin M, Bowden TA, Coles CH, Harlos K, Aricescu AR, Harvey DJ, Stuart DI and Jones EY (2009) Carbohydrate and domain architecture of an immature antibody glycoform exhibiting enhanced effector functions. *Journal of molecular biology* **387**(5): 1061-1066.

Day PW, Rasmussen SG, Parnot C, Fung JJ, Masood A, Kobilka TS, Yao XJ, Choi HJ, Weis WI, Rohrer DK and Kobilka BK (2007) A monoclonal antibody for G protein-coupled receptor crystallography. *Nat Methods* **4**(11): 927-929.

de Graaf C, Kooistra AJ, Vischer HF, Katritch V, Kuijer M, Shiroishi M, Iwata S, Shimamura T, Stevens RC, de Esch IJ and Leurs R (2011) Crystal structure-based virtual screening for fragment-like ligands of the human histamine H(1) receptor. *J Med Chem* **54**(23): 8195-8206.

Deupi X (2014) Relevance of rhodopsin studies for GPCR activation. *Biochim Biophys Acta* **1837**(5): 674-682.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Deupi X, Edwards P, Singhal A, Nickle B, Oprian D, Schertler G and Standfuss J (2012) Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II. *Proceedings of the National Academy of Sciences of the United States of America* **109**(1): 119-124.

Dore AS, Okrasa K, Patel JC, Serrano-Vega M, Bennett K, Cooke RM, Errey JC, Jazayeri A, Khan S, Tehan B, Weir M, Wiggin GR and Marshall FH (2014) Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature* **511**(7511): 557-562.

Dore AS, Robertson N, Errey JC, Ng I, Hollenstein K, Tehan B, Hurrell E, Bennett K, Congreve M, Magnani F, Tate CG, Weir M and Marshall FH (2011) Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure* **19**(9): 1283-1293.

Duarte JM, Biyani N, Baskaran K and Capitani G (2013) An analysis of oligomerization interfaces in transmembrane proteins. *BMC Struct Biol* **13**: 21.

Egloff P, Hillenbrand M, Klenk C, Batyuk A, Heine P, Balada S, Schlinkmann KM, Scott DJ, Schutz M and Pluckthun A (2014) Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **111**(6): E655-662.

Emsley P, Lohkamp B, Scott WG and Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**(Pt 4): 486-501.

Fenalti G, Giguere PM, Katritch V, Huang XP, Thompson AA, Cherezov V, Roth BL and Stevens RC (2014) Molecular control of delta-opioid receptor signalling. *Nature* **506**(7487): 191-196.

Fenalti G, Zatsepin NA, Betti C, Giguere P, Han GW, Ishchenko A, Liu W, Guillemyn K, Zhang H, James D, Wang D, Weierstall U, Spence JC, Boutet S, Messerschmidt M, Williams GJ, Gati C, Yefanov OM, White TA, Oberthuer D, Metz M, Yoon CH, Barty A, Chapman HN, Basu S, Coe J, Conrad CE, Fromme R, Fromme P, Tourwe D, Schiller PW, Roth BL, Ballet S, Katritch V, Stevens RC and Cherezov V (2015) Structural basis for bifunctional peptide recognition at human delta-opioid receptor. *Nat Struct Mol Biol* **22**(3): 265-268.

Ghosh E, Kumari P, Jaiman D and Shukla AK (2015) Methodological advances: the unsung heroes of the GPCR structural revolution. *Nat Rev Mol Cell Biol* **16**(2): 69-81.

Granier S, Manglik A, Kruse AC, Kobilka TS, Thian FS, Weis WI and Kobilka BK (2012) Structure of the delta-opioid receptor bound to naltrindole. *Nature* **485**(7398): 400-404.

Haga K, Kruse AC, Asada H, Yurugi-Kobayashi T, Shiroishi M, Zhang C, Weis WI, Okada T, Kobilka BK, Haga T and Kobayashi T (2012) Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* **482**(7386): 547-551.

Hanson MA, Cherezov V, Griffith MT, Roth CB, Jaakola VP, Chien EY, Velasquez J, Kuhn P and Stevens RC (2008) A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor. *Structure* **16**(6): 897-905.

Hendrickson WA, Smith JL and Sheriff S (1985) Direct phase determination based on anomalous scattering. *Methods in enzymology* **115**: 41-55.

Heydenreich FM, Vuckovic Z, Matkovic M and Veprintsev DB (2015) Stabilization of G protein-coupled receptors by point mutations. *Front Pharmacol* **6**: 82.

Hino T, Arakawa T, Iwanari H, Yurugi-Kobayashi T, Ikeda-Suno C, Nakada-Nakura Y, Kusano-Arai O, Weyand S, Shimamura T, Nomura N, Cameron AD, Kobayashi T, Hamakubo T, Iwata S and Murata T (2012) G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. *Nature* **482**(7384): 237-240.

Hollenstein K, Kean J, Bortolato A, Cheng RK, Dore AS, Jazayeri A, Cooke RM, Weir M and Marshall FH (2013) Structure of class B GPCR corticotropin-releasing factor receptor 1. *Nature* **499**(7459): 438-443.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Hossler P, Khattak SF and Li ZJ (2009) Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* **19**(9): 936-949.

Isberg V, Vroling B, van der Kant R, Li K, Vriend G and Gloriam D (2014) GPCRDB: an information system for G protein-coupled receptors. *Nucleic Acids Res* **42**(Database issue): D422-425.

Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP and Stevens RC (2008) The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **322**(5905): 1211-1217.

Kern J, Alonso-Mori R, Tran R, Hattne J, Gildea RJ, Echols N, Glockner C, Hellmich J, Laksmono H, Sierra RG, Lassalle-Kaiser B, Koroidov S, Lampe A, Han G, Gul S, Difiore D, Milathianaki D, Fry AR, Miahnahri A, Schafer DW, Messerschmidt M, Seibert MM, Koglin JE, Sokaras D, Weng TC, Sellberg J, Latimer MJ, Grosse-Kunstleve RW, Zwart PH, White WE, Glatzel P, Adams PD, Bogan MJ, Williams GJ, Boutet S, Messinger J, Zouni A, Sauter NK, Yachandra VK, Bergmann U and Yano J (2013) Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. *Science* **340**(6131): 491-495.

Kleywegt GJ, Harris MR, Zou JY, Taylor TC, Wahlby A and Jones TA (2004) The Uppsala Electron-Density Server. *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1): 2240-2249.

Kobilka B (2013) The structural basis of G-protein-coupled receptor signaling (Nobel Lecture). *Angewandte Chemie* **52**(25): 6380-6388.

Kooistra AJ, Kuhne S, de Esch IJP, Leurs R and de Graaf C (2013) A structural chemogenomics analysis of aminergic GPCRs: lessons for histamine receptor ligand design. *British journal of pharmacology*.

Kruse AC, Hu J, Pan AC, Arlow DH, Rosenbaum DM, Rosemond E, Green HF, Liu T, Chae PS, Dror RO, Shaw DE, Weis WI, Wess J and Kobilka BK (2012) Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* **482**(7386): 552-556.

Kruse AC, Ring AM, Manglik A, Hu J, Hu K, Eitel K, Hubner H, Pardon E, Valant C, Sexton PM, Christopoulos A, Felder CC, Gmeiner P, Steyaert J, Weis WI, Garcia KC, Wess J and Kobilka BK (2013a) Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**(7478): 101-106.

Kruse AC, Weiss DR, Rossi M, Hu J, Hu K, Eitel K, Gmeiner P, Wess J, Kobilka BK and Shoichet BK (2013b) Muscarinic receptors as model targets and antitargets for structure-based ligand discovery. *Molecular pharmacology* **84**(4): 528-540.

Kufareva I, Katritch V, Participants of GD, Stevens RC and Abagyan R (2014) Advances in GPCR modeling evaluated by the GPCR Dock 2013 assessment: meeting new challenges. *Structure* **22**(8): 1120-1139.

Kufareva I, Rueda M, Katritch V, Stevens RC, Abagyan R and participants GD (2011) Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **19**(8): 1108-1126.

Lamb AL, Kappock TJ and Silvaggi NR (2015) You are lost without a map: Navigating the sea of protein structures. *Biochim Biophys Acta* **1854**(4): 258-268.

Lanctot PM, Leclerc PC, Escher E, Guillemette G and Leduc R (2006) Role of N-glycan-dependent quality control in the cell-surface expression of the AT1 receptor. *Biochem Biophys Res Commun* **340**(2): 395-402.

Lane JR, Chubukov P, Liu W, Canals M, Cherezov V, Abagyan R, Stevens RC and Katritch V (2013) Structure-based ligand discovery targeting orthosteric and allosteric pockets of dopamine receptors. *Molecular pharmacology* **84**(6): 794-807.

Lebon G, Warne T, Edwards PC, Bennett K, Langmead CJ, Leslie AG and Tate CG (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* **474**(7352): 521-525.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Lee S, Bhattacharya S, Grisshammer R, Tate C and Vaidehi N (2014) Dynamic behavior of the active and inactive states of the adenosine A(2A) receptor. *J Phys Chem B* **118**(12): 3355-3365.

Lee SC, Bennett BC, Hong WX, Fu Y, Baker KA, Marcoux J, Robinson CV, Ward AB, Halpert JR, Stevens RC, Stout CD, Yeager MJ and Zhang Q (2013) Steroid-based facial amphiphiles for stabilization and crystallization of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* **110**(13): E1203-1211.

Li J, Edwards PC, Burghammer M, Villa C and Schertler GF (2004) Structure of bovine rhodopsin in a trigonal crystal form. *Journal of molecular biology* **343**(5): 1409-1438.

Liu W, Chun E, Thompson AA, Chubukov P, Xu F, Katritch V, Han GW, Roth CB, Heitman LH, AP IJ, Cherezov V and Stevens RC (2012) Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **337**(6091): 232-236.

Liu W, Wacker D, Gati C, Han GW, James D, Wang D, Nelson G, Weierstall U, Katritch V, Barty A, Zatsepin NA, Li D, Messerschmidt M, Boutet S, Williams GJ, Koglin JE, Seibert MM, Wang C, Shah ST, Basu S, Fromme R, Kupitz C, Rendek KN, Grotjohann I, Fromme P, Kirian RA, Beyerlein KR, White TA, Chapman HN, Caffrey M, Spence JC, Stevens RC and Cherezov V (2013) Serial femtosecond crystallography of G protein-coupled receptors. *Science* **342**(6165): 1521-1524.

Liu W, Wacker D, Wang C, Abola E and Cherezov V (2014) Femtosecond crystallography of membrane proteins in the lipidic cubic phase. *Philos Trans R Soc Lond B Biol Sci* **369**(1647): 20130314.

Lohse MJ, Maiellaro I and Calebiro D (2014) Kinetics and mechanism of G protein-coupled receptor activation. *Curr Opin Cell Biol* **27**: 87-93.

Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK and Granier S (2012) Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature* **485**(7398): 321-326.

McNicholas S, Potterton E, Wilson KS and Noble ME (2011) Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr D Biol Crystallogr* **67**(Pt 4): 386-394.

Michino M, Abola E, participants GD, Brooks CL, 3rd, Dixon JS, Moult J and Stevens RC (2009) Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nat Rev Drug Discov* **8**(6): 455-463.

Milic D and Veprintsev DB (2015) Large-scale production and protein engineering of G protein-coupled receptors for structural studies. *Front Pharmacol* **6**: 66.

Moukhametzianov R, Burghammer M, Edwards PC, Petitdemange S, Popov D, Fransen M, McMullan G, Schertler GF and Riekel C (2008) Protein crystallography with a micrometre-sized synchrotron-radiation beam. *Acta Crystallogr D Biol Crystallogr* **64**(Pt 2): 158-166.

Murakami M and Kouyama T (2008) Crystal structure of squid rhodopsin. *Nature* **453**(7193): 363-367.

Nobles KN, Xiao K, Ahn S, Shukla AK, Lam CM, Rajagopal S, Strachan RT, Huang TY, Bressler EA, Hara MR, Shenoy SK, Gygi SP and Lefkowitz RJ (2011) Distinct phosphorylation sites on the beta(2)-adrenergic receptor establish a barcode that encodes differential functions of beta-arrestin. *Sci Signal* **4**(185): ra51.

Norskov-Lauritsen L, Jorgensen S and Brauner-Osborne H (2015) N-glycosylation and disulfide bonding affects GPRC6A receptor expression, function, and dimerization. *FEBS Lett* **589**(5): 588-597.

Oates J and Watts A (2011) Uncovering the intimate relationship between lipids, cholesterol and GPCR activation. *Current opinion in structural biology* **21**(6): 802-807.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Obiol-Pardo C, Lopez L, Pastor M and Selent J (2011) Progress in the structural prediction of G protein-coupled receptors: D3 receptor in complex with eticlopride. *Proteins* **79**(6): 1695-1703.

Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M and Miyano M (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**(5480): 739-745.

Pawson AJ, Sharman JL, Benson HE, Faccenda E, Alexander SP, Buneman OP, Davenport AP, McGrath JC, Peters JA, Southan C, Spedding M, Yu W, Harmar AJ and Nc I (2014) The IUPHAR/BPS Guide to PHARMACOLOGY: an expert-driven knowledgebase of drug targets and their ligands. *Nucleic Acids Res* **42**(Database issue): D1098-1106.

Pontier SM, Percherancier Y, Galandrin S, Breit A, Gales C and Bouvier M (2008) Cholesterol-dependent separation of the beta2-adrenergic receptor from its partners determines signaling efficacy: insight into nanoscale organization of signal transduction. *The Journal of biological chemistry* **283**(36): 24659-24672.

Prasanna X, Chattopadhyay A and Sengupta D (2014) Cholesterol modulates the dimer interface of the beta(2)-adrenergic receptor via cholesterol occupancy sites. *Biophys J* **106**(6): 1290-1300.

Preininger AM, Meiler J and Hamm HE (2013) Conformational flexibility and structural dynamics in GPCR-mediated G protein activation: a perspective. *Journal of molecular biology* **425**(13): 2288-2298.

Qin L, Kufareva I, Holden LG, Wang C, Zheng Y, Zhao C, Fenalti G, Wu H, Han GW, Cherezov V, Abagyan R, Stevens RC and Handel TM (2015) Structural biology. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science* **347**(6226): 1117-1122.

Rasmussen SG, Choi HJ, Fung JJ, Pardon E, Casarosa P, Chae PS, Devree BT, Rosenbaum DM, Thian FS, Kobilka TS, Schnapp A, Konetzki I, Sunahara RK, Gellman SH, Pautsch A, Steyaert J, Weis WI and Kobilka BK (2011a) Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **469**(7329): 175-180.

Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI and Kobilka BK (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* **450**(7168): 383-387.

Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK and Kobilka BK (2011b) Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **477**(7366): 549-555.

Read RJ, Adams PD, Arendall WB, 3rd, Brunger AT, Emsley P, Joosten RP, Kleywegt GJ, Krissinel EB, Lutteke T, Otwinowski Z, Perrakis A, Richardson JS, Sheffler WH, Smith JL, Tickle IJ, Vriend G and Zwart PH (2011) A new generation of crystallographic validation tools for the protein data bank. *Structure* **19**(10): 1395-1412.

Rhodes G (2006) Crystallography Made Crystal Clear, in *A Guide for Users of Macromolecular Models*, Academic Press, Burlington.

Ring AM, Manglik A, Kruse AC, Enos MD, Weis WI, Garcia KC and Kobilka BK (2013) Adrenaline-activated structure of beta-adrenoceptor stabilized by an engineered nanobody. *Nature*.

Rodriguez D, Ranganathan A and Carlsson J (2014) Strategies for improved modeling of GPCR-drug complexes: blind predictions of serotonin receptors bound to ergotamine. *J Chem Inf Model* **54**(7): 2004-2021.

Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Yao XJ, Weis WI, Stevens RC and Kobilka BK (2007) GPCR engineering yields high-

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

resolution structural insights into beta2-adrenergic receptor function. *Science* **318**(5854): 1266-1273.

Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SG, Choi HJ, Devree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P and Kobilka BK (2011) Structure and function of an irreversible agonist-beta(2) adrenoceptor complex. *Nature* **469**(7329): 236-240.

Roumen L, Sanders MPA, Vroling B, De Esch IJP, De Vlieg J, Leurs R, Klomp JPG, Nabuurs SB and De Graaf C (2011) In Silico Veritas: The Pitfalls and Challenges of Predicting GPCR-Ligand Interactions. *Pharmaceuticals* **4**(9): 1196-1215.

Rupp B (2009) *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. 1st ed. Garland Science, New York.

Schrodinger, LLC (2010) The PyMOL Molecular Graphics System, Version 1.3r1.

Serrano-Vega MJ and Tate CG (2009) Transferability of thermostabilizing mutations between beta-adrenergic receptors. *Mol Membr Biol* **26**(8): 385-396.

Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW, Kobayashi T, Stevens RC and Iwata S (2011) Structure of the human histamine H1 receptor complex with doxepin. *Nature* **475**(7354): 65-70.

Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, Zhou C, Xu Q, Wacker D, Joseph JS, Liu W, Lau J, Cherezov V, Katritch V, Wang MW and Stevens RC (2013) Structure of the human glucagon class B G-protein-coupled receptor. *Nature* **499**(7459): 444-449.

Soto AG and Trejo J (2010) N-linked glycosylation of protease-activated receptor-1 second extracellular loop: a critical determinant for ligand-induced receptor activation and internalization. *The Journal of biological chemistry* **285**(24): 18781-18793.

Spence JC, Weierstall U and Chapman HN (2012) X-ray lasers for structural and dynamic biology. *Rep Prog Phys* **75**(10): 102601.

Standfuss J, Xie G, Edwards PC, Burghammer M, Oprian DD and Schertler GF (2007) Crystal structure of a thermally stable rhodopsin mutant. *Journal of molecular biology* **372**(5): 1179-1188.

Steyaert J and Kobilka BK (2011) Nanobody stabilization of G protein-coupled receptor conformational states. *Current opinion in structural biology* **21**(4): 567-572.

Tan Q, Zhu Y, Li J, Chen Z, Han GW, Kufareva I, Li T, Ma L, Fenalti G, Zhang W, Xie X, Yang H, Jiang H, Cherezov V, Liu H, Stevens RC, Zhao Q and Wu B (2013) Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science* **341**(6152): 1387-1390.

Tate CG (2012) A crystal clear solution for determining G-protein-coupled receptor structures. *Trends Biochem Sci* **37**(9): 343-352.

Tate CG and Grisshammer R (1996) Heterologous expression of G-protein-coupled receptors. *Trends Biotechnol* **14**(11): 426-430.

Tate CG and Schertler GF (2009) Engineering G protein-coupled receptors to facilitate their structure determination. *Current opinion in structural biology* **19**(4): 386-395.

Tobin AB (2008) G-protein-coupled receptor phosphorylation: where, when and by whom. *British journal of pharmacology* **153 Suppl 1**: S167-176.

Vass M, Agai-Csongor E, Horti F and Keseru GM (2014) Multiple fragment docking and linking in primary and secondary pockets of dopamine receptors. *ACS medicinal chemistry letters* **5**(9): 1010-1014.

Wacker D, Wang C, Katritch V, Han GW, Huang XP, Vardy E, McCorvy JD, Jiang Y, Chu M, Siu FY, Liu W, Xu HE, Cherezov V, Roth BL and Stevens RC (2013) Structural features for functional selectivity at serotonin receptors. *Science* **340**(6132): 615-619.

Wang C, Jiang Y, Ma J, Wu H, Wacker D, Katritch V, Han GW, Liu W, Huang XP, Vardy E, McCorvy JD, Gao X, Zhou XE, Melcher K, Zhang C, Bai F, Yang H, Yang L, Jiang H,

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Roth BL, Cherezov V, Stevens RC and Xu HE (2013a) Structural basis for molecular recognition at serotonin receptors. *Science* **340**(6132): 610-614.

Wang C, Wu H, Evron T, Vardy E, Han GW, Huang XP, Hufeisen SJ, Mangano TJ, Urban DJ, Katritch V, Cherezov V, Caron MG, Roth BL and Stevens RC (2014) Structural basis for Smoothened receptor modulation and chemoresistance to anticancer drugs. *Nature communications* **5**: 4355.

Wang C, Wu H, Katritch V, Han GW, Huang XP, Liu W, Siu FY, Roth BL, Cherezov V and Stevens RC (2013b) Structure of the human smoothened receptor bound to an antitumour agent. *Nature* **497**(7449): 338-343.

Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG and Schertler GF (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* **454**(7203): 486-491.

Weichert D and Gmeiner P (2015) Covalent Molecular Probes for Class A G Protein-Coupled Receptors: Advances and Applications. *ACS chemical biology*.

Weierstall U, James D, Wang C, White TA, Wang D, Liu W, Spence JC, Bruce Doak R, Nelson G, Fromme P, Fromme R, Grotjohann I, Kupitz C, Zatsepin NA, Liu H, Basu S, Wacker D, Han GW, Katritch V, Boutet S, Messerschmidt M, Williams GJ, Koglin JE, Marvin Seibert M, Klinker M, Gati C, Shoeman RL, Barty A, Chapman HN, Kirian RA, Beyerlein KR, Stevens RC, Li D, Shah ST, Howe N, Caffrey M and Cherezov V (2014) Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nature communications* **5**: 3309.

White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, Gvozdenovic-Jeremic J, Shah P, Shiloach J, Tate CG and Grisshammer R (2012) Structure of the agonist-bound neurotensin receptor. *Nature* **490**(7421): 508-513.

Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V and Stevens RC (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**(6007): 1066-1071.

Wu H, Wacker D, Mileni M, Katritch V, Han GW, Vardy E, Liu W, Thompson AA, Huang XP, Carroll FI, Mascarella SW, Westkaemper RB, Mosier PD, Roth BL, Cherezov V and Stevens RC (2012) Structure of the human kappa-opioid receptor in complex with JDTic. *Nature* **485**(7398): 327-332.

Wu H, Wang C, Gregory KJ, Han GW, Cho HP, Xia Y, Niswender CM, Katritch V, Meiler J, Cherezov V, Conn PJ and Stevens RC (2014) Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science* **344**(6179): 58-64.

Xiang Y, Rybin VO, Steinberg SF and Kobilka B (2002) Caveolar localization dictates physiologic signaling of beta 2-adrenoceptors in neonatal cardiac myocytes. *The Journal of biological chemistry* **277**(37): 34280-34286.

Xu F, Wu H, Katritch V, Han GW, Jacobson KA, Gao ZG, Cherezov V and Stevens RC (2011) Structure of an agonist-bound human A2A adenosine receptor. *Science* **332**(6027): 322-327.

Yin J, Mobarec JC, Kolb P and Rosenbaum DM (2015) Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant. *Nature* **519**(7542): 247-250.

Zhang D, Gao ZG, Zhang K, Kiselev E, Crane S, Wang J, Paoletta S, Yi C, Ma L, Zhang W, Han GW, Liu H, Cherezov V, Katritch V, Jiang H, Stevens RC, Jacobson KA, Zhao Q and Wu B (2015a) Two disparate ligand-binding sites in the human P2Y receptor. *Nature*.

Zhang H, Unal H, Gati C, Han GW, Liu W, Zatsepin NA, James D, Wang D, Nelson G, Weierstall U, Sawaya MR, Xu Q, Messerschmidt M, Williams GJ, Boutet S, Yefanov OM, White TA, Wang C, Ishchenko A, Tirupula KC, Desnoyer R, Coe J, Conrad CE, Fromme

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

P, Stevens RC, Katritch V, Karnik SS and Cherezov V (2015b) Structure of the Angiotensin receptor revealed by serial femtosecond crystallography. *Cell* **161**(4): 833-844.

Zhang K, Zhang J, Gao ZG, Zhang D, Zhu L, Han GW, Moss SM, Paoletta S, Kiselev E, Lu W, Fenalti G, Zhang W, Muller CE, Yang H, Jiang H, Cherezov V, Katritch V, Jacobson KA, Stevens RC, Wu B and Zhao Q (2014) Structure of the human P2Y12 receptor in complex with an antithrombotic drug. *Nature* **509**(7498): 115-118.

Zou Y, Weis WI and Kobilka BK (2012) N-terminal T4 lysozyme fusion facilitates crystallization of a G protein coupled receptor. *PloS one* **7**(10): e46039.

Zuckerman DM, Hicks SW, Charron G, Hang HC and Machamer CE (2011) Differential regulation of two palmitoylation sites in the cytoplasmic tail of the beta1-adrenergic receptor. *The Journal of biological chemistry* **286**(21): 19014-19023.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Funding

This work was supported by the Swiss National Science Foundation [Grant 146520] and by COST Action GLISTEN [CM1207]; James Kean is an employee of Heptares Therapeutics Ltd, which is a wholly owned subsidiary of the Sosei Group Corporation.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Figure and Table legends

Figure 1. A. Number of GPCR structures deposited in the Protein Data Bank per year (top) and total (bottom). B. Solved GPCR structures in each main class and ligand family. C. The 2D schematic diagram shows a summary of the main receptor modifications that enable structure determination of GPCRs. Sites of N-linked glycosylation are represented by the letter 'N'; these are typically removed through mutagenesis or enzymatic digestion. Phosphorylation sites are indicated by the letter 'S', and palmitoylation sites are indicated by the letter 'C'. Mutations that improve expression or stability are represented by the letter 'X', and have been found throughout the transmembrane domains and loops. Sites of receptor truncation are shown by dotted lines, and sites of protein fusion are indicated by red scissors. The fusion proteins used, either as an N-terminal tag, ICL2 or ICL3 fusion are shown below with N- and C-termini represented as blue and red spheres respectively (T4 lysozyme, PDB ID 3NY8; b_{562}RIL, PDB ID 1M6T; rubredoxin, PDB ID 4MBS; catalytic domain of glycogen synthase, PDB ID 4S0V).

Figure 2. Main steps in the process of structure determination by X-ray crystallography.

Figure 3. Electron densities in the binding site (top panels) of the A_{2A} adenosine receptor in complex with ZM241385 (bottom panels), solved at two different resolutions. While even at 2.6 Å resolution (left) side chains and ligand can be modeled reliably, the electron density of water molecules at this contour level (2 sigma) is too weak to be visible. On the other hand, at 1.8 Å resolution (right), the electron density of most water molecules in the binding site is clearly visible, and modeling of the side chains and ligand are unambiguous. The commands to produce this type of representation in PyMOL are provided in the Supplementary Material.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Molecular Pharmacology Fast Forward. Published on July 7, 2015 as DOI: 10.1124/mol.115.099663
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 4. Structures of the A_{2A} adenosine receptor in complex with ZM241385 colored by B-factor. The structure in the right was solved at a higher resolution and also presents a better order, which translates in a better definition of, e.g., ECL2. The commands to produce this type of representation in PyMOL are provided in the Supplementary Material.

Figure 5. A. Graphical representation of the data in Supplementary Table 2 (see main text and the legend of Supplementary Table 2 for details). B. Different ligand binding modes in the P_2Y_1 receptor. C. The ligands doxepin (H_1R), (R)-3-quinuclidinylbenzilate (BZ) (M_2R), and tiotropium (M_3R) ligands share a similar butterfly shape.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Table 1. List of pharmaceutically relevant compounds co-crystallized with GPCRs.

Receptor	Class	PMID	PDBID	Drug	Efficacy	Drug ID (1)	Action
Approved drugs							
beta 1 adrenergic receptor (turkey)	A	21228877	2Y04	Salbutamol	Agonist (partial)	DB01001	Asthma and obstructive pulmonary disease.
		21228877	2Y03	Isoprenaline	Agonist	DB01064	Bronchodilator and heart stimulant.
		21228877	2Y01	Dobutamine	Agonist (partial)	DB00841	Cardiac stimulant used after heart attack.
		22579251	4AMJ	Carvedilol	Inverse agonist	DB01136	Congestive heart failure.
beta 2 adrenergic receptor (human)	A	18547522	3D4S	Timolol	Inverse agonist	DB00373	Antihypertensive, antiangina, and antiarrhythmic.
		20669948	3NYA	Alprenolol	Antagonist	DB00866	Antihypertensive, antiangina, and antiarrhythmic.
		24056936	4LDO	Adrenaline	Agonist	DB00668	Asthma and cardiac failure.
M3 muscarinic acetylcholine receptor (rat)	A	22358844	4DAJ	Tiotropium	Inverse agonist	DB01409	Bronchodilator used in the management of chronic obstructive pulmonary disease.
serotonin 1b receptor (human)	A	23519210	4IAR	Ergotamine	Agonist	DB00696	Treatment of migraine.
		23519210	4IAQ	Dihydroergotamine	Agonist	DB00320	Treatment of migraine.
serotonin 2b receptor (human)	A	23519210	4IB4	Ergotamine	Agonist	DB00696	Treatment of migraine.
H1 histamine receptor (human)	A	21697825	3RZE	Doxepin (E,Z)	Antagonist	DB01142	Sedative (2).
adenosine A2A receptor (human)	A	21593763	2YDO	Adenosine	Agonist	DB00640	Treatment for some types of tachycardia.
		21885291	3RFM	Caffeine	Antagonist	DB00201	Central nervous system stimulant, appears to be useful in the treatment of some types of headache.
CCR5 chemokine receptor (human)	A	24030490	4MBS	Maraviroc	Antagonist	DB04835	Treatment of HIV infection.
OX2 orexin receptor (human)	A	25533960	4S0V	Suvorexant	Antagonist	24965990 (3)	Treatment of insomnia.
protease-activated receptor type 1 (human)	A	23222541	3VW7	Vorapaxar	Antagonist	10077130 (3)	Treatment for acute coronary syndrome chest pain caused by coronary artery disease.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

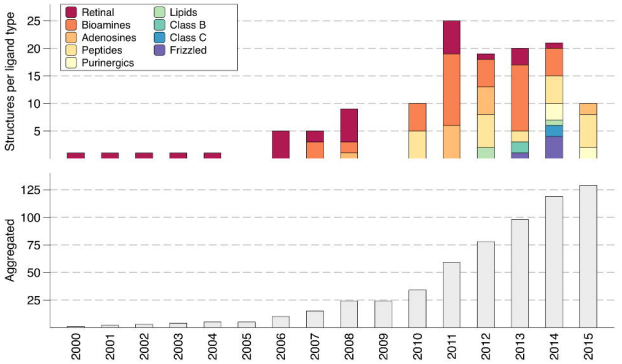
In clinical trials							
P2Y12 purinergic receptor (human)	A	24670650	4NTJ	AZD1283	Antagonist	23649325 (3)	Treatment of acute arterial thrombosis.
smoothened receptor (human)	F	23636324	4JKV	LY2940680	Antagonist	49848070 (3)	Small cell lung cancer.
Development discontinued							
kappa opioid receptor (human)	A	22437504	4DJH	JDTic	Antagonist	9956146 (3)	Treatment of cocaine abuse.
free fatty-acid receptor type 1 (human)	A	25043059	4PHU	TAK-875 (fasiglifam)	Agonist	24857286 (3)	Treatment for type 2 diabetes.
metabotropic glutamate receptor type 5 (human)	C	25042998	4OO9	Mavoglurant	Negative allosteric modulator	9926832 (3)	Treatment of fragile X syndrome.

(1) DrugBank ID (when available); (2) Doxepin has antidepressant effects, due to an overall increase in serotonergic neurotransmission; blocking of histamine H1 receptors accounts for its sedative effects. (3) PubChem CID.

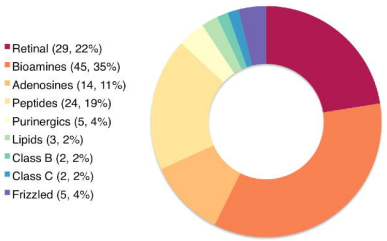
Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 20, 2024

Figure 1

A


Retinal
Bioamines
Adenosines
Peptides
Purinergics
Lipids
Class B
Class C
Frizzled
Structures per ligand type
25
20
15
10
5
Aggregated
125
100
75
50
25
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015

B


Retinal (29, 22%)
Bioamines (45, 35%)
Adenosines (14, 11%)
Peptides (24, 19%)
Purinergics (5, 4%)
Lipids (3, 2%)
Class B (2, 2%)
Class C (2, 2%)
Frizzled (5, 4%)

C

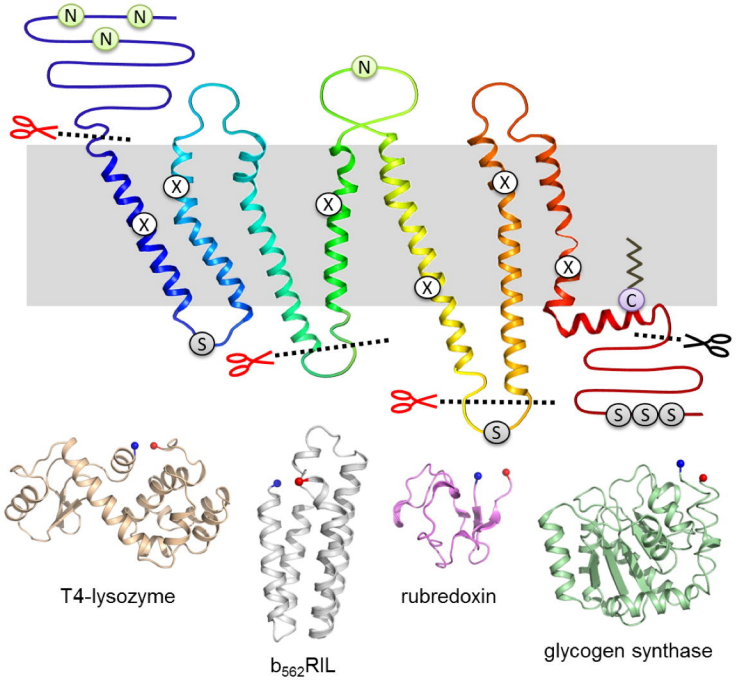

N
X
S
C
T4-lysozyme
b562RIL
rubredoxin
glycogen synthase

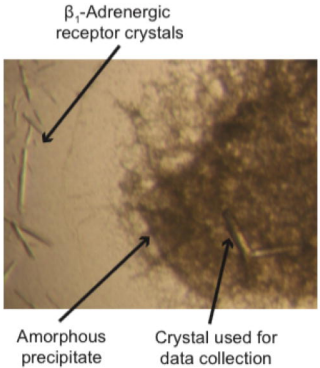
Figure 2

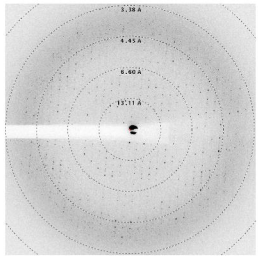
crystallization

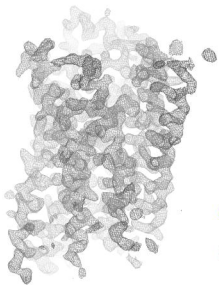
X-ray diffraction pattern

electron density

model building







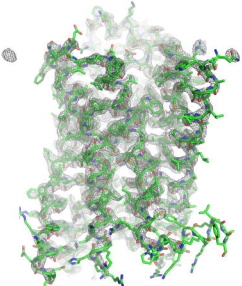


Figure 3

3EML

2.6 Å

4EIY

1.8 Å

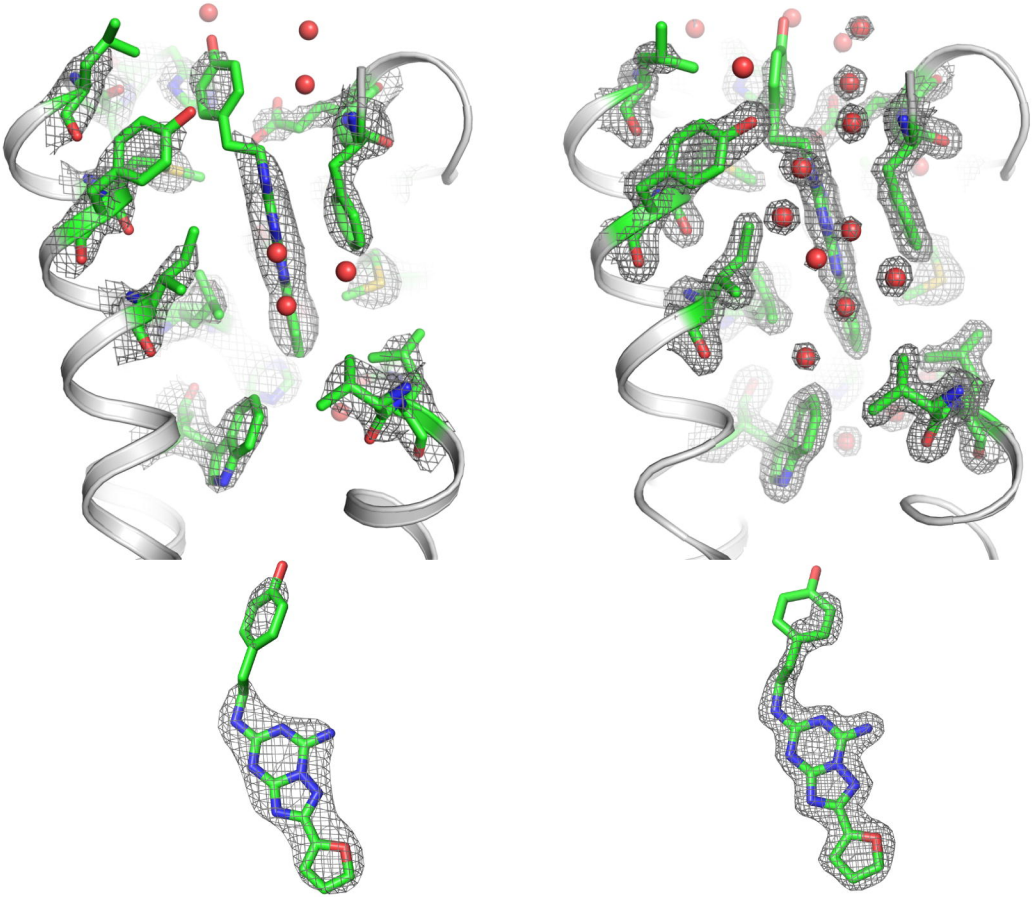


Figure 4

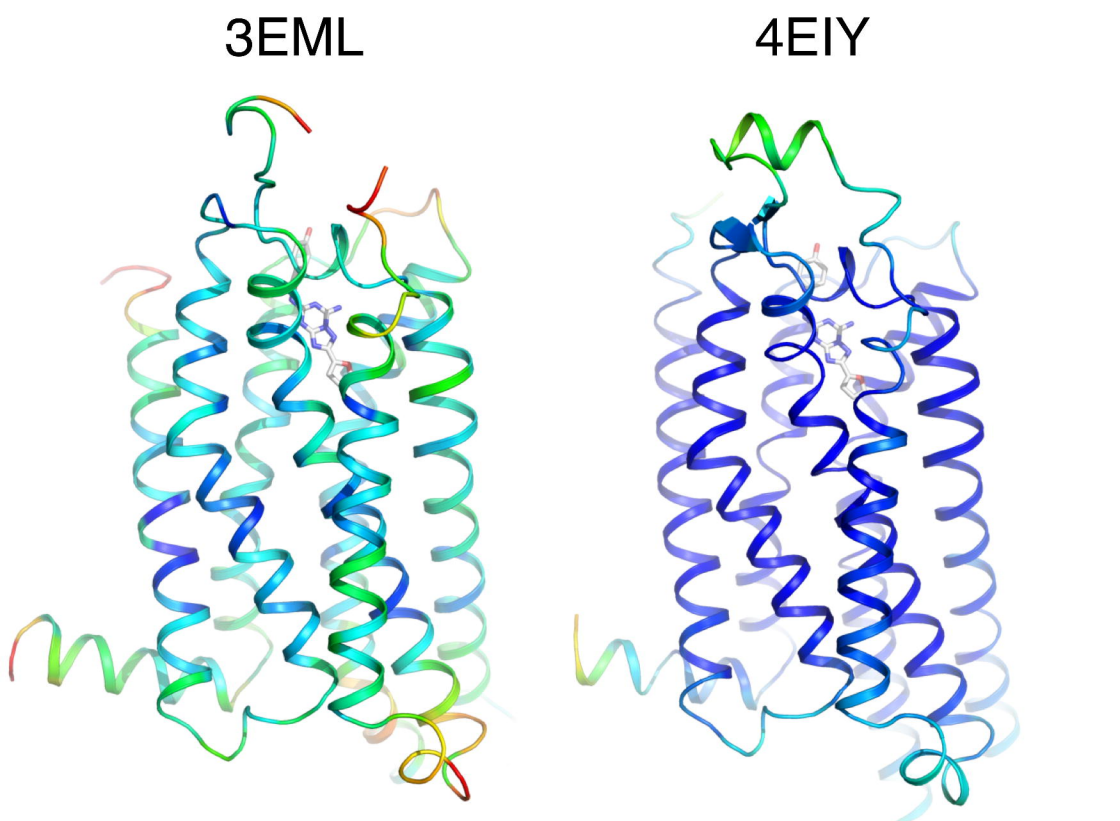
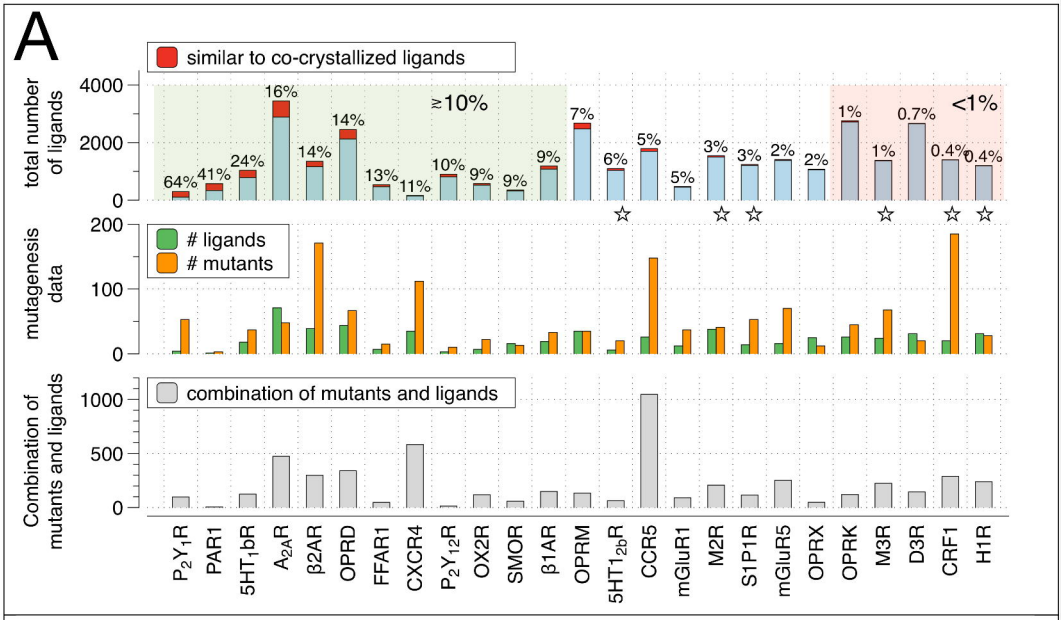
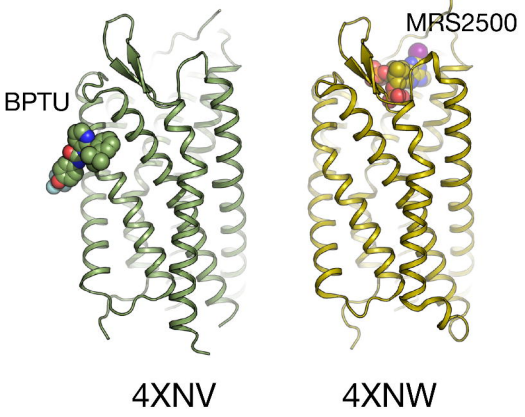
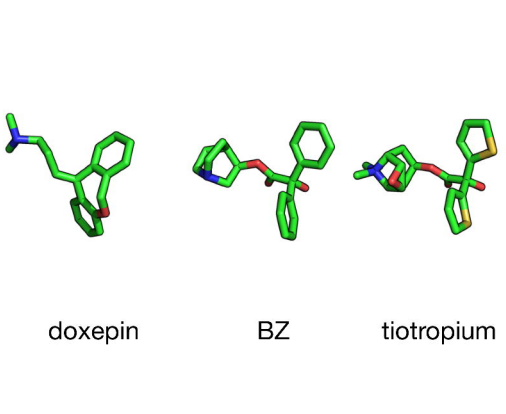


Figure 5


A
similar to co-crystallized ligands
total number of ligands
4000
2000
0
≥10%
<1%
64%
41%
24%
16%
14%
14%
13%
11%
10%
9%
9%
9%
7%
6%
5%
5%
3%
3%
2%
2%
1%
1%
0.7%
0.4%
0.4%
mutagenesis data
200
100
0
ligands
mutants
Combination of mutants and ligands
1000
500
0
combination of mutants and ligands
P2Y1R
PAR1
5HT1bR
A2AR
β2AR
OPRD
FFAR1
CXCR4
P2Y12R
OX2R
SMOR
β1AR
OPRM
5HT2bR
CCR5
mGluR1
M2R
S1P1R
mGluR5
OPRX
OPRK
M3R
D3R
CRF1
H1R


B
P2Y1 receptor
BPTU
MRS2500
4XNV
4XNW


C
doxepin
BZ
tiotropium