

OPINION

The GPCR Network: a large-scale collaboration to determine human GPCR structure and function

Raymond C. Stevens, Vadim Cherezov, Vsevolod Katritch, Ruben Abagyan, Peter Kuhn, Hugh Rosen and Kurt Wüthrich

Abstract | G protein-coupled receptors (GPCRs) are targeted by ~30–40% of marketed drugs, and their key roles in normal physiology and in disease demonstrate that an understanding of their structure and function is valuable to researchers in both basic science and drug discovery. However, until recently, detailed structural information on this protein family was limited by challenges in X-ray crystallographic analysis of such membrane proteins. The GPCR Network was created in 2010 with the goal of structurally characterizing 15–25 representative human GPCRs within 5 years, based on an active outreach programme addressing an interdisciplinary community of scientists interested in GPCR structure, chemistry and biology. Here, we provide an overview of how this collaborative effort has enabled the structural determination and characterization of eight human GPCRs so far, and discuss some of the challenges that remain in gaining more detailed insights into structure–function relationships in this receptor superfamily.

G protein-coupled receptors (GPCRs) are critical eukaryotic signal transduction gatekeepers and represent the largest protein family in the human proteome, with more than 800 members. They share a common architecture of seven transmembrane helices, and can be classified into five major classes and further divided into subfamilies based on sequence similarities (FIG. 1). Located in the plasma membrane, GPCRs recognize a variety of extracellular stimuli, including photons, ions, small molecules, peptides and proteins, and transmit the resulting extracellular signals across the membrane over a distance of ~30 Å to elicit intracellular responses. Signal transmission occurs through coupling to different intracellular proteins (for example, heterotrimeric G proteins, arrestins and kinases)¹, which then activate downstream effectors and trigger cascades of cellular and physiological responses. GPCR-mediated

signalling pathways have been related to numerous human diseases, and GPCRs are the targets of an estimated 30–40% of all drugs currently on the market².

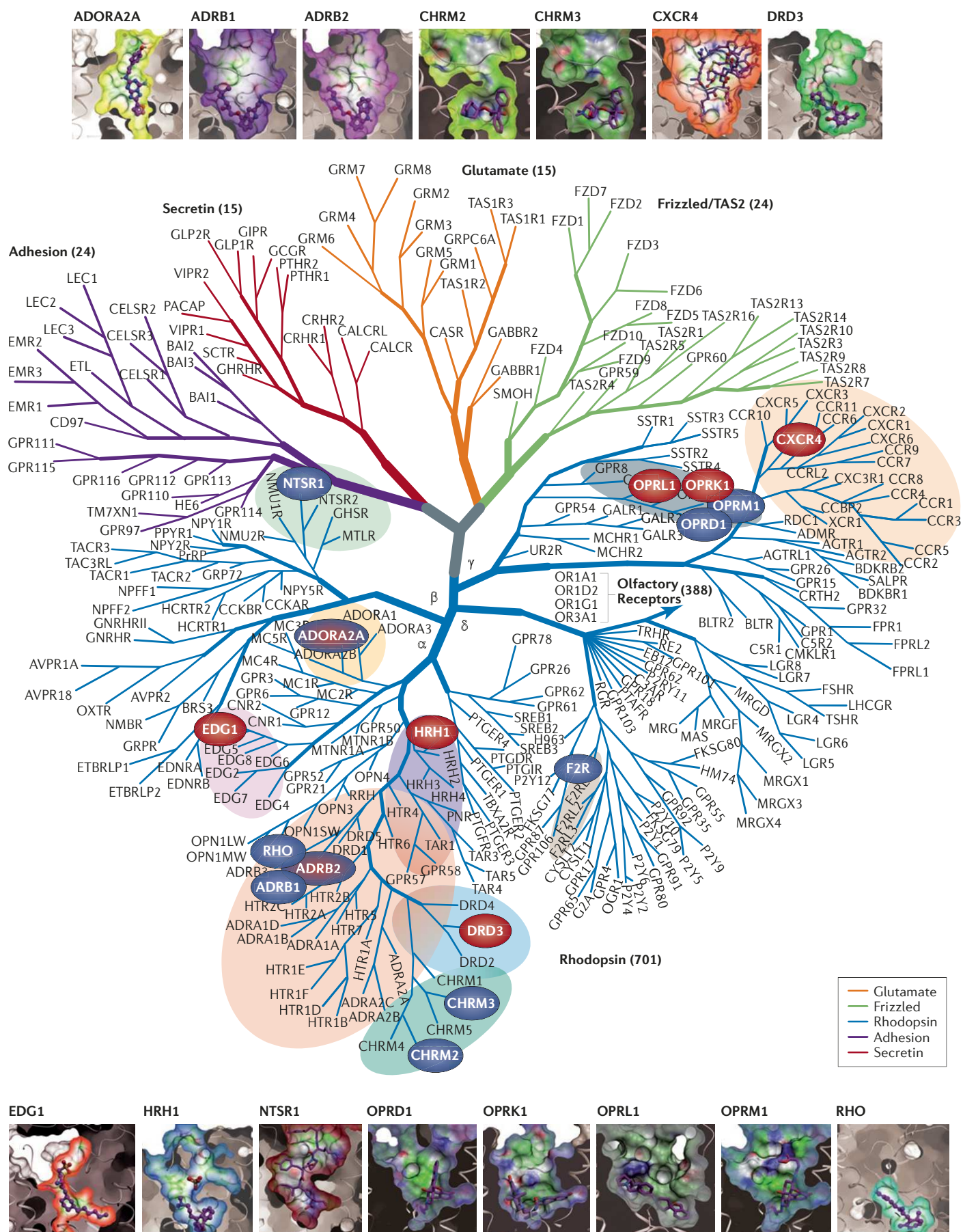
In view of their fundamental roles in health and disease, a detailed understanding of GPCR structure and function is of value to the basic science community interested in cell signalling and molecular recognition, as well as to the applied science community interested in drug discovery. However, gaining this knowledge has been challenging because of the inherent conformational flexibility of these receptors. This flexibility is essential for their activity under physiological conditions³, but has made crystallization and structure determination using traditional methods very difficult. Breakthroughs in many connected steps along the structure determination pipeline were necessary to surmount this obstacle (BOX 1). The rapid progress in the field has

brought numerous new and unexpected insights into ligand–receptor binding, selectivity, activation mechanism and allosteric interaction, and has corrected past misconceptions about how GPCRs work. These insights are only the first steps in building a comprehensive picture of GPCR family structure–function relationships and exploiting structural information in biophysical, pharmacological and computational inquiries. In addition, further progress in the field depends on our ability to identify stabilizing ligands benefiting crystallization and our understanding of how stabilizing mutations and fusions assist crystallization and how much protein engineering alters functional features of the receptors.

In this Perspective, we present a community-wide interdisciplinary infrastructure effort that was created through a funding opportunity from the US National Institute of General Medical Biosciences (NIGMS) — one of the component institutes of the US National Institutes of Health (NIH) — to achieve a thorough understanding of human GPCR structure–function relationships. The receptors and their interactions are characterized using structural biology techniques (based on X-ray crystallography and NMR spectroscopy), chemistry, biochemistry, biophysics and bioinformatics. In addition, experiments are conducted on each crystallized GPCR, such as site-specific mutagenesis of key residues and investigation of structure–activity relationships for ligand–receptor binding. An important element of the programme is the active initiation of collaborations around the globe with fellow scientists interested in specific GPCRs.

Establishment and aims of the Network GPCR Network infrastructure and pipeline.

The GPCR Network was established in 2010 as a collaborative effort funded by the NIGMS's Protein Structure Initiative: Biology programme (NIGMS PSI: Biology; specifically the U54 funding mechanism). The GPCR Network, based at The Scripps Research Institute, is part of a much larger network of research centres aimed at implementing high-throughput structure determination pipelines to study a broad range of important biological and biomedical problems.



◀ **Figure 1 | Phylogenetic tree representation of the human GPCR superfamily.** There are more than 800 human G protein-coupled receptors (GPCRs), which can be classified into five major families: Rhodopsin (class A); Secretin and Adhesion (class B); Glutamate (class C); and Frizzled/taste receptor 2 (TAS2). These families can be further divided into subfamilies on the basis of sequence similarity. This phylogenetic tree was constructed using sequence similarity within the seven-transmembrane region. GPCRs are named here according to their gene name as used by the [UniProt database](#). Family members with determined structures are highlighted within the tree, and their binding pockets with the ligand — as captured in each of the distinct structures — are shown around the tree in the same orientation for ease of comparison. ADORA2A, adenosine A_{2A} receptor (Protein Data Bank (PDB) code: 3EML); ADRB1, β_1 -adrenergic receptor (PDB code: 2VT4); ADRB2, β_2 -adrenergic receptor (PDB code: 2RH1); CHRM2, muscarinic acetylcholine receptor M_2 (PDB code: 3UON); CHRM3, muscarinic acetylcholine receptor M_3 (PDB code: 4DAJ); CXCR4, CXC chemokine receptor 4 (PDB code: 3ODU); DRD3, dopamine D_3 receptor (3PBL); EDG1, sphingosine-1-phosphate receptor 1 (PDB code: 3V2Y); HRH1, histamine H_1 receptor (PDB code: 3RZE); NTSR1, neurotensin receptor 1 (PDB code: 4GRV); OPRD1, δ -type opioid receptor (PDB code: 4EJ4); OPRK1, κ -type opioid receptor (PDB code: 4DJH); OPR1, opiate receptor-like 1 (nociceptin/orphanin FQ peptide opioid receptor) (PDB code: 4EA3); OPRM1, μ -type opioid receptor (PDB code: 4DKL); RHO, rhodopsin (PDB code: 1F88); F2R, proteinase-activated receptor 1 (PDB code: 3VW7; structure not yet available). Figure modified, with permission, from REF. 58 © (2003) American Society for Pharmacology and Experimental Therapeutics.

In addition, the technical capability of determining GPCR structures was created through a combination of technologies developed with support from the NIH Common Fund in Structural Biology to the Joint Center for Innovative Membrane Protein Technologies (JCIMPT), and by the integration and optimization of technologies developed by other research groups.

Efficient determination of GPCR structures requires the development of a robust approach to enable a high likelihood of success and to provide a platform that is amenable to optimization and cost reduction based on accumulated experience. The approach used by the GPCR Network in the initial successful structure determinations of the β_2 -adrenergic receptor (β_2 -AR)⁴ and the adenosine A_{2A} receptor⁵ led to the development of a pipeline infrastructure with feedback loops that could be used as a template for the determination of other GPCR structures (BOX 1). This pipeline is constantly being optimized using a ‘family learning approach’, which is similar to the approach used by the Structural Genomics Consortium and which has been successfully applied to other protein families, most notably protein kinases⁶.

With the family learning approach, similar protocols and reagents (for example, expression strategies, assays, inhibitors and co-factors) are re-used, and lessons learned for one family member can often be applied to other members, making it more cost-effective to study large protein families than would be possible through traditional NIH R01 funding (which is awarded to single investigators with two to three students).

In parallel with the research efforts of scientists within the GPCR Network, a constant flow of scientists from around

the world passes through the GPCR Network laboratories. These visitors are trained in the use of technologies that have been developed as part of the JCIMPT programme, and then return to their home institutions with new methods and tools, thereby encouraging an expansion of the number of active researchers studying GPCRs.

GPCR selection. A crucial aspect of the Network is that its research is focused on human GPCRs. This is because of numerous examples in which a single-amino-acid

difference between the animal and human isoforms of a protein can have important consequences for drug discovery^{7–10}, and also because of the collective experience of the Network in structure-based drug discovery.

The main goal of the Network is to structurally characterize 15–25 representative GPCRs within the 5-year funding period (2010–2015), with the vision to characterize molecular recognition and signalling mediated by this membrane protein superfamily. As well as the structures of the receptors themselves, a better understanding of their function is facilitated by studies of these receptors in complexes with a broad range of ligands, using X-ray crystallography, NMR spectroscopy and hydrogen–deuterium exchange mass spectroscopy (HDX-MS). The repertoire of biophysical methods used will be extended to include, for example, electron paramagnetic resonance (EPR) spectroscopy experiments.

In addition, computational methods of virtual ligand screening, conformational sampling of ligand-binding pockets and molecular dynamics simulations are used to explore an ever-expanding ligand universe. These structural pharmacology studies involve not only orthosteric ligands but also increasingly focus on allosteric ligands and bitopic ligands (also known as combined allosteric-orthosteric ligands) that modulate GPCR activity, which can yield beneficial

Glossary

Allosteric ligands

Ligands that bind elsewhere from the orthosteric binding site and influence the functional properties of the receptor. In some classifications, the intracellular binding partners (for example, G proteins) are considered to be allosteric molecules as they bind at a distance of 30 Å from the orthosteric ligand-binding site.

Bitopic ligands

Ligands that have both orthosteric ligand-binding properties as well as a secondary element that is able to bind to a neighbouring allosteric site on the receptor.

Ca atoms

The chiral carbon atoms to which the primary amine, the carboxylic group and the side chain are attached to in an amino acid. Comparison of three-dimensional structures of proteins is sometimes carried out by superimposing the Ca atoms of proteins, as this provides a simple estimate of the similarity of their skeleton or backbone structure.

Electron paramagnetic resonance

(EPR). Similar in concept to NMR spectroscopy, but whereas NMR examines the spins of atomic nuclei, EPR detects the spins of unpaired electrons.

Hydrogen–deuterium exchange mass spectroscopy

(HDX-MS). A technique used to probe protein conformations. The exchange rate of an amide hydrogen is substantially influenced by hydrogen bonding, and the exchange kinetics of an amide hydrogen can be highly reflective of its locations in secondary and tertiary structures.

Non-olfactory receptors

G protein-coupled receptors from the Rhodopsin family, excluding the ~388 olfactory receptors.

Orthosteric ligands

Ligands that bind to the natural ligand-binding site on the receptor and thus directly compete with this natural ligand for receptor binding. For class A G protein-coupled receptors (GPCRs), the orthosteric binding site is typically in the cavity positioned in the extracellular portion of the seven-transmembrane region. For class B and class C GPCRs, pockets in this location are considered to be allosteric because their natural ligands bind in a separate extracellular domain.

Root mean square deviation

(RMSD). A quantitative measure of the similarity between two superimposed sets of atomic coordinates. RMSD values (units of Å) can be calculated for any type and subset of atoms: for example, for chiral carbon (Ca) atoms of proteins (Ca RMSD) for all residues; for residues in the transmembrane helices or the loops; as well as for non-hydrogen atoms of small-molecule ligands (ligand RMSD).

Box 1 | The GPCR Network structure determination pipeline

To go from the nomination of a specific G protein-coupled receptor (GPCR) to its final three-dimensional structure determination and thereby develop a deeper understanding of receptor dynamics and functional behaviour, we have developed a stage-gate process (which harnesses more than 20 years of experience) that is continually being refined to increase efficiency and decrease the overall cost of this process. This structure determination approach relies on a set of family learning metrics through a pipeline with feedback loops (see the figure).

The thick black lines describe the process pathway; the thin black lines are feedback loops at which information at one stage can be used to repeat or adjust an earlier set of experiments. For example, if we find that a construct is not stable enough or does not pass one of our pre-crystallization screens, we would return (thin black lines) to construct design or search for more ligands that might help further stabilize the receptor. Once a structure is determined (steps 1–10), follow-up studies are initiated to include complementary structure–function characterization (steps 10–12).

Step 1: gene construct design and synthesis. An average of 75 constructs have been necessary for the Network to determine each of the eight human GPCR structures to date. The cDNA for each of the target's parent construct is codon-optimized and synthesized. Construct optimization includes placements of affinity tags, amino- and carboxyl termini truncations and fusion protein considerations, as well as thermostabilizing point mutations⁵⁵.

Step 2: expression using baculovirus expression system. Receptor overexpression is carried out in a high-throughput (HT) manner using a baculovirus expression system in insect cells at three different volume scales for rapid construct screening, ligand stability and biophysical characterization as well as crystallization trials.

Step 3: purification. Solubilization in detergent or cholesterol mixtures is followed by purification protocols based on immobilized metal affinity chromatography (IMAC).

Step 4: ligand selection and sample characterization. Before any crystallization trials are initiated, ligand selection is carefully evaluated either during the expression stage or at this stage.

Steps 5–6: biophysical characterization and pre-crystallization assays. Purity, monodispersity and ligand-binding affinity quality checks are followed by thermal stability analysis and by receptor mobility in lipidic cubic phase (LCP) using fluorescence recovery after photobleaching (known as LCP-FRAP)⁵⁶.

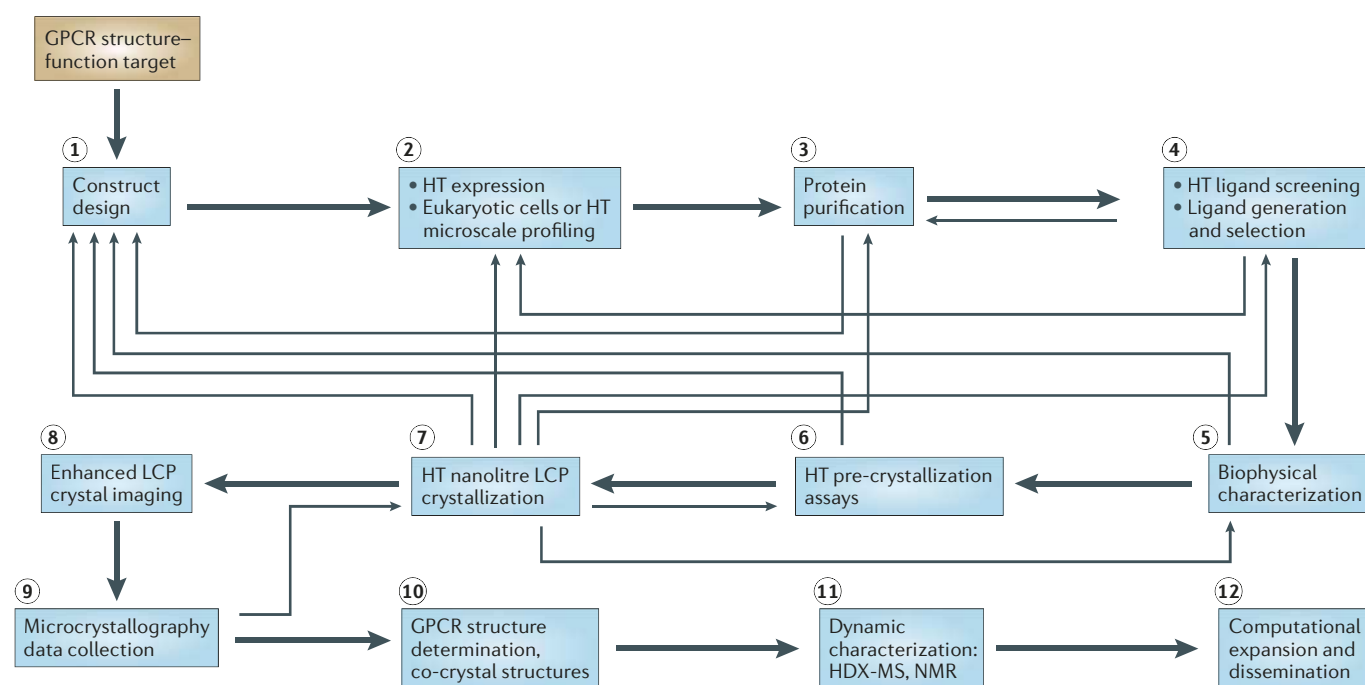
Steps 7–8: crystallization studies and crystal imaging. Crystallization trials are performed using the *in meso* (LCP)⁵⁷ approaches. This crystallization approach was selected as the media most closely resembles the cell membrane and is important in aiding the stabilization of detergent-destabilized human GPCRs. Initial coarse screening is carried out using ~400 conditions coupled with several host lipid matrices. After set-up, the plates are transferred into an automatic incubator or imager that is capable of storing up to 1,000 plates, where they are automatically imaged according to a specified schedule.

Step 9: data collection, processing and structure determination. Diffraction data collection is done at synchrotron beamlines equipped with microfocus beams (for example, the General Medical Sciences and Cancer Institutes Structural Biology Facility at the Advanced Photon Source (the GM/CA APS project), Argonne, Illinois). Data collection, processing and structure solution are optimized based on experience of working with microcrystals.

Step 10: structure determination of GPCR–ligand complexes and co-crystal structures. For each receptor, several receptor–ligand co-crystal structures are pursued to thoroughly map orthosteric and allosteric receptor binding sites.

Step 11: characterization of the dynamics of GPCR–ligand complexes. Biophysical studies using NMR and hydrogen–deuterium exchange mass spectroscopy (HDX-MS) are pursued to understand conformational equilibria and dynamic aspects of each receptor. The repertoire of techniques used will be further expanded — for example, with electron paramagnetic resonance spectroscopy.

Step 12: computational ligand screening and medicinal chemistry. Ligand docking, virtual screening and homology modelling of close receptor subtypes are conducted to further characterize and annotate each receptor structure. Medicinal chemistry studies are conducted through community outreach to yield an improved understanding of GPCR–ligand interactions.



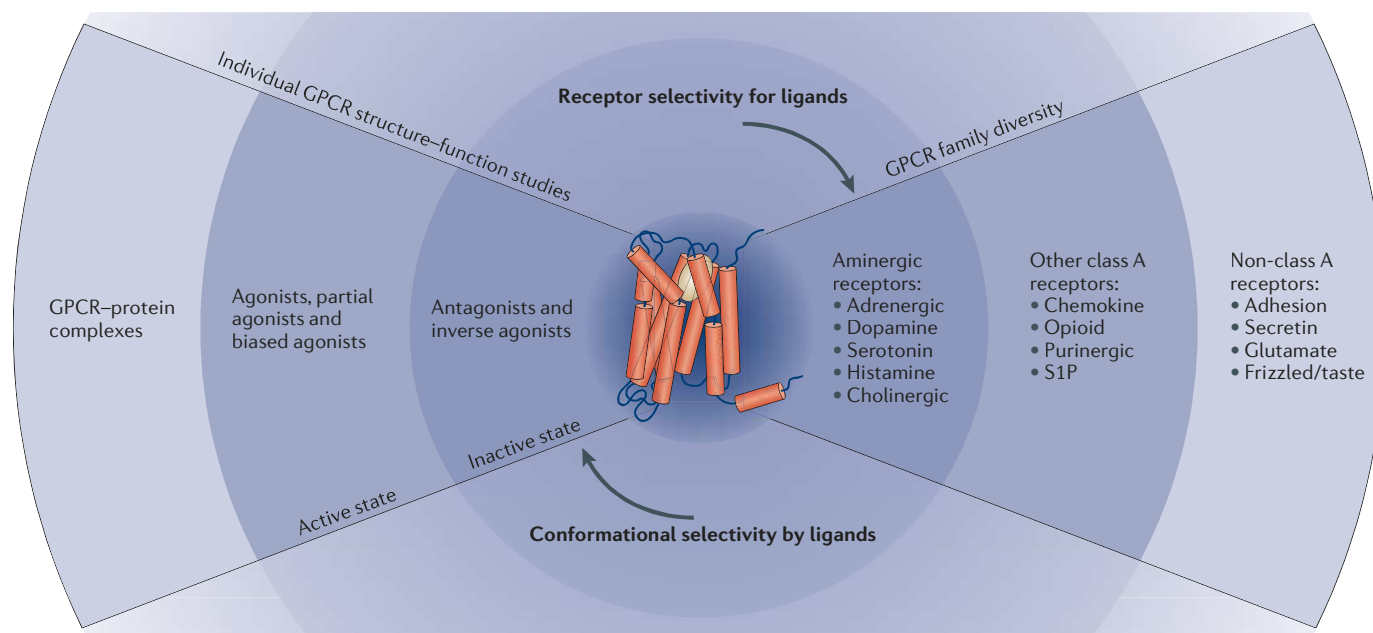


Figure 2 | The strategy of the GPCR Network. This schematic diagram demonstrates the strategy undertaken by the GPCR Network to determine receptor structure and function. It includes leveraging each experimentally determined receptor structure (for example, human β_2 -adrenergic receptor (β_2 -AR); centre) towards the understanding of G protein-coupled receptor (GPCR) family diversity and receptor selectivity for ligands (right

panel) as well as individual GPCR structure–function studies and conformational selectivity by ligands (left panel). By creating a complete data package for each receptor that includes complementary biophysical, structural, functional and ligand data, the receptors can be more thoroughly understood in contrast to just solving the structure with limited follow up. S1P, sphingosine-1-phosphate.

selectivity and functional profiles for candidate drugs¹¹. This work is then followed up by medicinal chemistry and tool compound development, whereby the biological significance of the structural information is extensively investigated by collaborating with scientists who have long-standing individual interests in particular receptor systems.

For the initial phase of the programme (through to 2015), we have selected GPCR targets from different subfamilies with distant homology in order to maximize the impact of each structure solved, as sufficiently related receptor subtypes can be analysed with homology modelling (as discussed below). The 5-year goal, based on a combination of experimentally solved structures and computationally predicted homology models of GPCRs, is to achieve 40–60% structural coverage of non-olfactory receptors (FIG. 1). The eight structures solved in the first 2 years since the GPCR Network was established are estimated to cover about 67 modelled receptors when using a 35% sequence identity threshold for homology modelling (as described below), which corresponds to more than 18% structural coverage of non-olfactory receptors. Because of the paucity of the structural information for peptide GPCRs, a few of these subfamilies — for example, opioid, chemokine and

class B (secretin) receptors — are tagged for more detailed coverage in order to experimentally characterize their structural variability and selectivity towards orthosteric and allosteric ligands.

So far, there are no published representative structures for class B and class C receptors (although solution of the structure of corticotropin-releasing factor receptor 1, a class B GPCR, has been reported by the company Heptares), and these receptors are being pursued for structural coverage by many different groups, including the GPCR Network. Limitations to the successful crystallization of these receptors include the scarcity of ligands that can stabilize the extracellular and transmembrane domains of these receptors into a unique and compact conformation. Nevertheless, published structures of class B and class C receptors are anticipated within the next year or two.

Beyond the Network

Solving the structures of all members of the GPCR superfamily in the various signalling states is not a realistic goal at present, and thus computationally predictive approaches and complementary biochemistry studies are crucial for obtaining a complete understanding of this important protein superfamily. The overall plan of the GPCR Network is

to work with the scientific community to fill in the gaps in structural knowledge and advance our understanding of GPCR structure and function.

Computational GPCR studies. An important activity of the GPCR Network is to leverage the impact of each experimental structure in order to enhance our understanding of the structures of related GPCRs, their specificity and other functional features. This goal is being achieved through comparative analyses as well as homology-based modelling of structures and their complexes^{12,13} (FIG. 2), which has previously proven to be successful in applications to other major protein families¹⁴.

To assess the state of the art in computer homology modelling for the prediction of ligand–GPCR interactions, two community-wide assessments of GPCR modelling and docking were organized: GPCR Dock 2008 and GPCR Dock 2010 (FIG. 3). In 2008, the goal was to assess the ability of existing protocols and programmes to predict the structure of the adenosine A_{2A} receptor bound to the antagonist ZM241385 (REF. 15). A few groups, which used β_2 -AR and/or β_1 -AR as the closest available template structures as well as extensive knowledge of mutagenesis and structure–activity

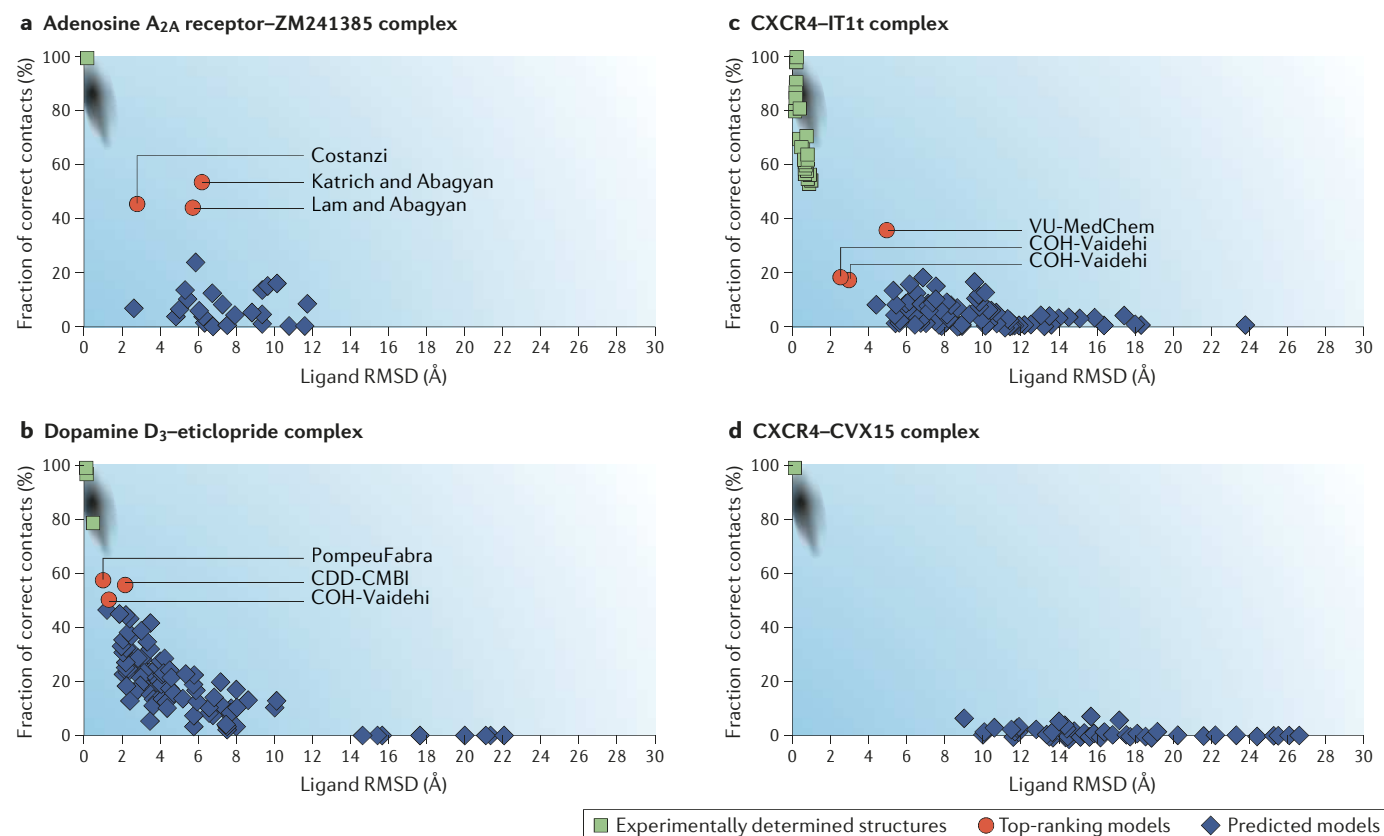


Figure 3 | Community-wide predictions of GPCR–ligand docking and receptor interactions. In 2008 (REF. 15) and in 2010 (REF. 16), two community-wide assessments were conducted whereby the participants submitted prediction models of undisclosed G protein-coupled receptor (GPCR)–ligand structures that were recently determined by the GPCR Network. The dark shaded area in the top left-hand corner of each plot represents the distribution of the corresponding parameters for pairs of symmetry-related molecules in a subset of each Protein Data Bank (PDB) structure. **a** | 2008: GPCR Dock involving the human adenosine A_{2A} receptor in complex with the antagonist ZM241385, highlighting the limited success in predicting receptor–ligand interactions. **b** | 2010: GPCR Dock involving the human dopamine D₃ receptor in complex with the antagonist eticlopride, showing significant improvements in receptor–ligand docking, which are probably due to improved experimental models for

the biogenic amine receptor subfamily. **c** | 2010: GPCR Dock involving human CXCR4 in complex with the small-molecule IT1t. **d** | 2010: GPCR Dock involving human CXCR4 in complex with the peptide molecule CVX15. Panels **a** and **b** show progress in the field based on available structural templates of related receptors (there was a sequence identity of 32% with β_2 -adrenergic receptor (β_2 -AR) in panel **a**, and a sequence identity of 41% with β_2 -AR in panel **b**), whereas panels **c** and **d** highlight the need for additional structural coverage for peptide receptors (there was a sequence identity of 25% with β_1 -AR, the most similar sequence template available, in panels **c** and **d**). The top-ranking models from each assessment are denoted in each panel as circles, with the labels representing the name of the group that submitted the model (see REFS 15, 16 for further details). RMSD, root mean square deviation.

relationships for the adenosine receptor subfamily, were able to predict the overall ligand orientation and about 40% of the ligand–receptor atomic contacts¹⁵. However, the accuracy of even the best models used in the assessment was far below that of crystal structures with regard to ligand–receptor contacts as well as ligand positioning¹⁵ (FIG. 3a).

In 2010, a second community-wide assessment included modelling of the dopamine D₃ receptor (FIG. 3b) and the CXC chemokine receptor 4 (CXCR4) (FIG. 3c,d) in complex with antagonists¹⁶. Analysis of these blind modelling results (FIG. 3) suggests that GPCR interactions with small-molecule ligands can be accurately and reliably

predicted when homology models are based on >35% sequence identity to a structural template in the transmembrane domain, as in the case of the D₃ receptor¹⁶. Consequently, we currently use 35% sequence identity as an approximate threshold for accurate homology modelling of GPCRs, although we are aware that this threshold may vary between different GPCR families. For example, in the opioid subfamily the nociceptin/orphanin FQ peptide opioid receptor (NOP) has a strikingly distinct ligand-binding pocket despite having 60% sequence identity with the classical opioid receptor subtypes δ , κ and μ ¹⁷. The accuracy of the ligand-binding predictions may also vary between different ligand classes. To engage a wider scientific

community in this discussion and encourage additional progress, the results of each assessment are freely available on the GPCR Network website.

Outreach and collaborations. Each structure determined to date through the GPCR Network programme has been studied in collaboration with scientists who have extensive knowledge of the specific receptor system from a chemical and biological perspective (TABLE 1).

For example, Brian Kobilka's expertise in the biochemistry of the β_2 -AR, including the protein engineering and understanding of receptor pharmacology, was crucial to the success in producing the first detailed

Table 1 | Crystal structures solved by the GPCR Network

GPCR	Partner (or partners)	PDB codes	Method of crystallization	Structure–function follow-up studies
β_2 -adrenergic receptor ^{4,38}	B. Kobilka	2RH1	LCP with ICL3 T4L fusion	Modelling full and partial agonists ⁵⁹
β_2 -adrenergic receptor ⁴⁰	None	3D4S	LCP with ICL3 T4L fusion	Establishing specific cholesterol binding site ⁴⁰
β_2 -adrenergic receptor ⁴¹	None	3NY8; 3NY9; 3NYA	LCP with ICL3 T4L fusion	Establishing conserved binding mode of receptor ⁴¹ ; ligand-dependent perturbation by HDX mass spectroscopy ²⁷ ; characterization of biased signalling pathways in the wild-type receptor by NMR ⁴⁶
Adenosine A _{2A} receptor ⁵	A. IJzerman and L. Heitman	3EML	LCP with ICL3 T4L fusion	GPCR Dock 2008 (REF. 15); assessment of novel agonist chemotypes ⁴⁹ ; subtype selectivity analysis by mutations ¹⁹ and modelling ⁶⁰
Adenosine A _{2A} receptor ¹⁸	K. Jacobson	3QAK	LCP with ICL3 T4L fusion	Assessment of activation-related changes in GPCRs; molecular modelling of agonists ²⁰ ; structure-based drug design ²¹
Adenosine A _{2A} receptor ⁶¹	A. IJzerman and L. Heitman	4EIY	LCP with ICL3 BRIL fusion	Allosteric regulation by sodium ions ⁶¹
CXC chemokine receptor 4 (REF. 23)	A. Brooun, P. Wells and T. Handel	3ODU; 3OE6; 3OE8; 3OE9; 3OE0	LCP with ICL3 T4L fusion	GPCR Dock 2010 (REF. 16)
Dopamine D ₃ receptor ²⁴	L. Shi, A. H. Newman and J. Javitch	3PBL	LCP with ICL3 T4L fusion	GPCR Dock 2010 (REF. 16)
Histamine H ₁ receptor ²⁵	S. Iwata and T. Kobayashi	3RZE	LCP with ICL3 T4L fusion	Virtual ligand screening ⁵²
κ -type opioid receptor ²²	B. Roth, F. I. Carroll and P. D. Mosier	4DJH	LCP with ICL3 T4L fusion	Twenty-nine mutations studied for structure–function relationships ²²
Nociceptin peptide receptor (also known as orphanin FQ receptor) ¹⁷	G. Calo and B. Roth	4EA3	LCP with amino-terminal BRIL fusion	Eight mutations studied for structure–function relationships ¹⁷
Sphingosine-1-phosphate receptor 1 (REF. 26)	M. Hanson and H. Rosen	3V2W; 3V2Y	LCP with ICL3 T4L fusion	Five mutations studied for structure–function relationships ²⁶

BRIL, thermostabilized apocytochrome *b*₅₆₂; GPCR, G protein-coupled receptor; HDX, hydrogen–deuterium exchange; ICL3, intracellular loop 3; LCP, lipidic cubic phase; T4L, T4 lysozyme.

structure of the receptor at 2.4 Å resolution⁴. Similarly, collaborations with Adriaan IJzerman and Kenneth Jacobson yielded structures of the adenosine A_{2A} receptor in complex with an antagonist at 2.6 Å resolution⁵ and with an agonist at 2.8 Å resolution¹⁸. Following the publication of the first adenosine A_{2A} receptor structure, we worked with Adriaan IJzerman, Laura Heitman and their colleagues on site-specific mutagenesis of the receptor binding site, and then with Kenneth Jacobson and colleagues at the NIH on the design and synthesis of novel ligands that bind to the receptor^{19–21}.

In the recent determination of the structure of the human κ -type opioid receptor²², Bryan Roth's laboratory was instrumental in selecting and providing the ligand that enabled the crystallization of this receptor. Upon completion of the structures of the κ -type opioid receptor and NOP, 28 and 5 mutants were made, respectively, for

the two different receptors to immediately probe their structure and function. Similar successes resulted from the following collaborations: Tracy Handel and Pfizer on the structure of CXCR4 (REF. 23); Jonathan Javitch on the structure of the dopamine D₃ receptor²⁴; So Iwata on the structure of the histamine H₁ receptor²⁵; Mike Hanson and Receptos on the structure of the sphingosine-1-phosphate receptor 1 (S1PR1)²⁶; and Girolamo Calo on the structure of NOP¹⁷. Given their long history of working with a specific receptor system, all of these colleagues were ideal partners for conducting follow-up biochemical studies, including site-specific mutagenesis and/or structure–activity relationship analyses with numerous compounds. Community collaborations have thus enabled structure determinations and added value to the structural insights once a structure is obtained, which is at the heart of the GPCR Network's mission.

To encourage and facilitate follow-up work on GPCRs with determined structures, the GPCR Network provides protein material to scientists who apply complementary biophysical methods. For example, HDX experiments in the Griffin laboratory at The Scripps Research Institute, Florida, USA, were performed with protein material produced at the GPCR Network, and this generated valuable insights into GPCR conformational equilibria in response to different pharmacological ligands²⁷. Finally, collaborative funding from the NIGMS PSI:Biological programme is proving to be encouragingly successful at partnering experts in specialized biology areas with other PSI:Biological centres. As an example, the Handel laboratory at the University of California San Diego, USA, partnered with the GPCR Network through an NIH-funded collaboration on chemokine receptors. To assist in target selection and collaborator recruitment, the PSI:Biological initiative

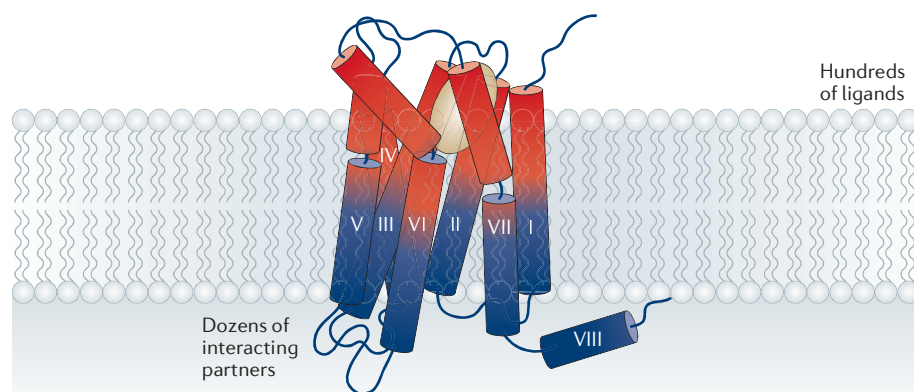


Figure 4 | Receptor diversity in class A GPCRs. The simplified diagram illustrates the diversity of class A G protein-coupled receptors (GPCRs). Comparison of experimental GPCR structures reveals a greater structural diversity in the extracellular modules (shown in red) than in the intracellular modules (shown in blue), which seems to reflect the evolutionary pressure of recognizing hundreds of endogenous ligands while transferring their signals to only dozens of interacting partners (see the main text for further details).

has also created the [Community Nominated Targets programme](#), which enables fellow scientists to suggest targets for close collaboration with the GPCR Network or another PSI:Biology centre.

Based on the experience gained by the GPCR Network so far, the most critical aspect for a collaborative effort to succeed is the active and enthusiastic participation of collaborating laboratories. This includes knowledge and access to compounds, as well as intuitive insights into specific receptor biochemistry. Of the collaborations that have not been successful to date, the primary reason has been a lack of commitment or passionate interest for the target structure. Consequently, we have adopted a policy whereby committed and active collaboration is required before we will commit our own resources, which are also very limited.

Impact of new GPCR structures

At the time of publication, the GPCR Network pipeline process has facilitated the determination of the structures of most of the available human GPCRs: β_2 -AR⁴, adenosine A_{2A} receptor⁵, CXCR4 (REF. 23), dopamine D_3 receptor²⁴, histamine H_1 receptor²⁵, S1PR1 (REF. 26), κ -type opioid receptor²² and NOP¹⁷. Beyond the GPCR Network, several groups have contributed key GPCR structures, including the structures of bovine rhodopsin²⁸, squid rhodopsin²⁹, turkey β_1 -AR³⁰, human muscarinic acetylcholine receptor M_2 (REF. 31), rat muscarinic acetylcholine receptor M_3 (REF. 32), mouse μ -type opioid receptor³³, mouse δ -type opioid receptor³⁴ and rat neurotensin receptor³⁵. Additional structure reports include the characterization

of the active and inactive states for rhodopsin³⁶ and opsin³⁷, β_2 -AR^{38–41} and adenosine A_{2A} receptor^{18,42,43}. In total, 16 unique GPCR structures (human and animal) have now been deposited into the [Protein Data Bank](#).

A key breakthrough for the field was the structure determination of a complex of β_2 -AR with the G_s protein⁴⁴, setting the stage for studies of GPCR complexes with their intracellular partner proteins. The ensemble of these structures illustrates that the seven-transmembrane protein architecture is ideally suited to recognize many different types of natural and synthetic ligands, and to transmit signals over a distance of 30 Å to the intracellular side of the membrane.

Receptor structure diversity and modularity.

The 16 distinct class A (Rhodopsin family) GPCR structures (both human and animal) that have been determined to date confirm the overall structural conservation of the seven-transmembrane fold, with a root mean squared deviation (RMSD) of C α atoms at <3.0 Å calculated for the transmembrane helices between any pair of GPCRs.

By contrast, the ligand-binding sites vary substantially among the different receptors, providing a rationale for the intricate molecular recognition potential within the GPCR family that makes each receptor unique⁴⁵.

Variations in structure are especially pronounced on the extracellular side of the receptors, reflecting a distinct evolutionary and functional modularity between the extracellular (ligand-binding) and intracellular (downstream signalling) modules of GPCRs (FIG. 4). Interestingly, although the RMSD value of the protein backbone

calculated between different GPCR pairs in the extracellular transmembrane region is almost twofold larger than the RMSD value of the intracellular transmembrane module, reflecting a greater variation in the extracellular region of the receptors, the intracellular side shows larger conformational changes in discrete molecular regions following receptor activation. A contrast in structural variability between the ligand-binding and downstream signalling modules is also evident from comparison of the primary structures, which show 6% sequence identity in the extracellular transmembrane region compared with 26% sequence identity in the intracellular transmembrane region.

For several GPCRs, structures have now been determined in different functional states, setting the stage for the formulation of hypotheses for follow-up work to the presently available data. There also seems to be modularity among the helix sub-bundles I to IV and V to VII. This observation leads to intriguing questions about splicing variants and receptor chimaeras, as well as about the role of individual helices in receptor function beyond their involvement in ligand binding or receptor oligomerization. For example, the primary role of helix I, which is structurally the most diverse among the available GPCR structures, may involve membrane insertion or oligomerization to a greater extent than previously proposed.

Receptor conformational states and dynamics.

As a complement to the crystal structure data, HDX and NMR experiments are being used to understand receptor behaviour. HDX spectroscopy is able to define the conformational flexibility of receptors at a global level, whereas the use of selective labels in NMR can enable one to ask specific questions about receptor conformational plasticity and dynamics. For example, using site-specific ¹⁹F-NMR labels to study β_2 -AR complexes with various ligands, we observed that the cytoplasmic ends of helices VI and VII adopt two major conformational states⁴⁶. Using ¹⁹F-NMR provides a novel approach for correlating chemical structures of pharmacological ligands with their signalling pathways, acting either through G proteins and/or through β -arrestin.

Drug discovery and development. Crystal structures of GPCRs also provide a robust three-dimensional structural framework for computational modelling of receptor dynamics and oligomerization state⁴⁷, as well as for ligand docking and virtual ligand screening (VLS). The growing number of

structure-based VLS studies demonstrate encouragingly high hit rates (20–70%) in the identification of new ligand chemotypes as lead compounds for adenosine A_{2A} receptor^{48,49}, CXCR4 (REF. 50), dopamine D_3 receptor⁵¹ and H_1 receptors⁵², as well as in lead optimization²¹. Following their successful applications to kinases, proteases and other protein target families, structure-based ligand screening technologies are now becoming an important part of the GPCR drug discovery process in pharmaceutical and biotechnology companies^{53,54}.

Conclusions and future directions

Twelve years after the structure of bovine rhodopsin was determined and 5 years after the first human GPCR structure with a diffusible ligand was obtained, a total of 16 unique GPCR structures have now been reported. As a consequence, leveraging through modelling has become more reliable and improvements in accuracy are anticipated as additional experimental structures are determined.

As a current working practice, the GPCR Network uses a 35% sequence identity cut-off point for homology modelling, and estimates that up to 20% of the non-olfactory GPCRs could thus be reliably modelled based on the experimental structures of the aforementioned 16 distinct receptors. For a full coverage of non-olfactory receptors at this cut-off point, our current estimate indicates that at least 100 carefully selected unique GPCR structures are needed. Beyond these considerations of 'structural coverage', the long-term focus of the GPCR Network is to obtain more detailed insights into structure–function correlations.

In this regard, additional structure determinations of GPCR complexes with different ligands and with intracellular partner proteins, combined with biophysical characterization of these receptor–ligand interactions, will provide valuable input. Of key interest is access to an ever-widening range of ligands to support more detailed characterization of the rapidly expanding group of GPCRs that have become accessible for structural biology studies, and to this end we welcome collaborations with interested parties.

The establishment of the GPCR Network was enabled by funding from the NIGMS PSI:Biology initiative, which was set up to support high-throughput structure determination pipelines to study a broad range of important biological and biomedical problems. The NIH funding mechanisms have evolved over the years to now include both highly collaborative large-scale projects

as well as the traditional single-investigator grants, and this allows for a more efficient path of scientific discovery depending on the scientific questions being proposed. Although at present it is challenging to define how much this work has cost over the years (which includes more than 20 years of investment in the field in general), it is hoped that over the coming years we will be in a better position to address this question.

Nevertheless, at the time of publication, 77 GPCR complexes with different ligands and binding partners have been deposited in the PDB. Based on the currently available structures that are providing a basis for molecular replacement, it is anticipated that this number will grow rapidly as the GPCR Network and others further pursue an improved understanding of GPCR structure and function. This ambitious project will continue to greatly benefit from close collaborations with fellow scientists who have insights into specific receptor systems and access to libraries of GPCR ligands, as well as a common vision for the continued growth of knowledge in this exciting and important area of biological and biomedical research.

With a more complete understanding of structure–function relationships within the GPCR superfamily, one can begin to investigate the next generation of questions about the basic principles of molecular recognition, variation of receptor conformational states in disease, similarities and differences in the signalling mechanisms of GPCRs, the rational design of more selective and desirable therapeutics as well as the evolution of GPCRs in the framework of the evolution of the human species.

Raymond C. Stevens, Vadim Cherezov and Vsevolod Katritch are at the Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

Ruben Abagyan is at the University of California San Diego (UCSD) Skaggs School of Pharmacy and Pharmaceutical Sciences and at the San Diego Supercomputer Center, La Jolla, California 92093, USA.

Peter Kuhn is at the Department of Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

Hugh Rosen is at the Department of Chemical Physiology and at the The Scripps Research Institute Molecular Screening Center, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

Kurt Wüthrich is at the Department of Molecular Biology and at the Skaggs Institute of Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

Correspondence to R.C.S.
e-mail: stevens@scripps.edu

doi:10.1038/nrd3859

Published online 14 December 2012

- Rajagopal, S., Rajagopal, K. & Lefkowitz, R. J. Teaching old receptors new tricks: biasing seven-transmembrane receptors. *Nature Rev. Drug Discov.* **9**, 373–386 (2010).
- Wise, A., Gearing, K. & Rees, S. Target validation of G-protein coupled receptors. *Drug Discov. Today* **7**, 235–246 (2002).
- Kobilka, B. K. & Deupi, X. Conformational complexity of G-protein-coupled receptors. *Trends Pharmacol. Sci.* **28**, 397–406 (2007).
- Cherezov, V. *et al.* High-resolution crystal structure of an engineered human β_2 -adrenergic G protein-coupled receptor. *Science* **318**, 1258–1265 (2007).
- Jaakola, V. P. *et al.* The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science* **322**, 1211–1217 (2008).
- Weigelt, J., McBroom-Cerajewski, L. D., Schapira, M., Zhao, Y. & Arrowsmith, C. H. Structural genomics and drug discovery: all in the family. *Curr. Opin. Chem. Biol.* **12**, 32–39 (2008).
- Mileni, M. *et al.* Structure-guided inhibitor design for human FAAH by interspecies active site conversion. *Proc. Natl Acad. Sci. USA* **105**, 12820–12824 (2008).
- Oksenberg, D. *et al.* A single amino-acid difference confers major pharmacological variation between human and rodent 5-HT_{1B} receptors. *Nature* **360**, 161–163 (1992).
- Tucker, A. L. *et al.* A_1 adenosine receptors. Two amino acids are responsible for species differences in ligand recognition. *J. Biol. Chem.* **269**, 27900–27906 (1994).
- Yao, B. B. *et al.* Molecular modeling and pharmacological analysis of species-related histamine H_3 receptor heterogeneity. *Neuropharmacology* **44**, 773–786 (2003).
- Valant, C., Robert Lane, J., Sexton, P. M. & Christopoulos, A. The best of both worlds? Bitopic orthosteric/allosteric ligands of G protein-coupled receptors. *Annu. Rev. Pharmacol. Toxicol.* **52**, 153–178 (2012).
- Reynolds, K., Abagyan, R. & Katritch, V. in *GPCR Molecular Pharmacology and Drug Targeting: Shifting Paradigms and New Directions* (ed. Gilchrist, A.) 385–433 (Wiley & Sons, 2010).
- Yarnitzky, T., Levit, A. & Niv, M. Y. Homology modeling of G-protein-coupled receptors with X-ray structures on the rise. *Curr. Opin. Drug Discov. Devel.* **13**, 317–325 (2010).
- Cavasotto, C. N. & Phatak, S. S. Homology modeling in drug discovery: current trends and applications. *Drug Discov. Today* **14**, 676–683 (2009).
- Michino, M. *et al.* Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nature Rev. Drug Discov.* **8**, 455–463 (2009).
- Kufareva, I., Rueda, M., Katritch, V., Stevens, R. C. & Abagyan, R. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **19**, 1108–1126 (2011).
- Thompson, A. A. *et al.* Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature* **485**, 395–399 (2012).
- Xu, F. *et al.* Structure of an agonist-bound human A_{2A} adenosine receptor. *Science* **332**, 322–327 (2011).
- Jaakola, V. P. *et al.* Ligand binding and subtype selectivity of the human A_{2A} adenosine receptor: identification and characterization of essential amino acid residues. *J. Biol. Chem.* **285**, 13032–13044 (2010).
- Deflorian, F. *et al.* Evaluation of molecular modeling of agonist binding in light of the crystallographic structure of an agonist-bound A_{2A} adenosine receptor. *J. Med. Chem.* **55**, 538–552 (2012).
- Tosh, D. K. *et al.* Optimization of adenosine 5'-carboxamide derivatives as adenosine receptor agonists using structure-based ligand design and fragment screening. *J. Med. Chem.* **55**, 4297–4308 (2012).
- Wu, H. *et al.* Structure of the human κ -opioid receptor in complex with JDTic. *Nature* **485**, 327–332 (2012).
- Wu, B. *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066–1071 (2010).
- Chien, E. Y. *et al.* Structure of the human dopamine D_3 receptor in complex with a D_3/D_2 selective antagonist. *Science* **330**, 1091–1095 (2010).
- Shimamura, T. *et al.* Structure of the human histamine H_1 receptor complex with doxepin. *Nature* **475**, 65–70 (2011).

26. Hanson, M. A. *et al.* Crystal structure of a lipid G protein-coupled receptor. *Science* **335**, 851–855 (2012).
27. West, G. M. *et al.* Ligand-dependent perturbation of the conformational ensemble for the GPCR β_2 adrenergic receptor revealed by HDX. *Structure* **19**, 1424–1432 (2011).
28. Palczewski, K. *et al.* Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289**, 739–745 (2000).
29. Murakami, M. & Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **453**, 363–367 (2008).
30. Warne, T. *et al.* Structure of a β_1 -adrenergic G-protein-coupled receptor. *Nature* **454**, 486–491 (2008).
31. Haga, K. *et al.* Structure of the human M_2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* **482**, 547–551 (2012).
32. Kruse, A. C. *et al.* Structure and dynamics of the M_3 muscarinic acetylcholine receptor. *Nature* **482**, 552–556 (2012).
33. Manglik, A. *et al.* Crystal structure of the μ -opioid receptor bound to a morphinan antagonist. *Nature* **485**, 321–326 (2012).
34. Granier, S. *et al.* Structure of the δ opioid receptor bound to naltrindole. *Nature* **485**, 400–404 (2012).
35. White, J. F. *et al.* Structure of the agonist-bound neurotensin receptor. *Nature* **490**, 508–513 (2012).
36. Standfuss, J. *et al.* The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* **471**, 656–660 (2011).
37. Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497–502 (2008).
38. Rosenbaum, D. M. *et al.* GPCR engineering yields high-resolution structural insights into β_2 -adrenergic receptor function. *Science* **318**, 1266–1273 (2007).
39. Rasmussen, S. G. *et al.* Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature* **469**, 175–180 (2011).
40. Hanson, M. A. *et al.* A specific cholesterol binding site is established by the 2.8 Å structure of the human β_2 -adrenergic receptor. *Structure* **16**, 897–905 (2008).
41. Wacker, D. *et al.* Conserved binding mode of human β_2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J. Am. Chem. Soc.* **132**, 11443–11445 (2010).
42. Lebon, G. *et al.* Agonist-bound adenosine A_{2A} receptor structures reveal common features of GPCR activation. *Nature* **474**, 521–525 (2011).
43. Dore, A. S. *et al.* Structure of the adenosine A_{2A} receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure* **19**, 1283–1293 (2011).
44. Rasmussen, S. G. *et al.* Crystal structure of the β_2 adrenergic receptor–Gs protein complex. *Nature* **477**, 549–555 (2011).
45. Katritch, V., Cherezov, V. & Stevens, R. C. Diversity and modularity of G protein-coupled receptor structures. *Trends Pharmacol. Sci.* **33**, 17–27 (2012).
46. Liu, J. J., Horst, R., Katritch, V., Stevens, R. C. & Wuthrich, K. Biased signaling pathways in β_2 -adrenergic receptor characterized by ^{19}F -NMR. *Science* **335**, 1106–1110 (2012).
47. Johnston, J. M. & Filizola, M. Showcasing modern molecular dynamics simulations of membrane proteins through G protein-coupled receptors. *Curr. Opin. Struct. Biol.* **21**, 552–558 (2011).
48. Carlsson, J. *et al.* Structure-based discovery of A_{2A} adenosine receptor ligands. *J. Med. Chem.* **53**, 3748–3755 (2010).
49. Katritch, V. *et al.* Structure-based discovery of novel chemotypes for adenosine A_{2A} receptor antagonists. *J. Med. Chem.* **53**, 1799–1809 (2010).
50. Mysinger, M. M. *et al.* Structure-based ligand discovery for the protein–protein interface of chemokine receptor CXCR4. *Proc. Natl Acad. Sci. USA* **109**, 5517–5522 (2012).
51. Carlsson, J. *et al.* Ligand discovery from a dopamine D_3 receptor homology model and crystal structure. *Nature Chem. Biol.* **7**, 769–778 (2011).
52. de Graaf, C. *et al.* Crystal structure-based virtual screening for fragment-like ligands of the human histamine H_1 receptor. *J. Med. Chem.* **54**, 8195–8206 (2011).
53. Congreve, M., Langmead, C. & Marshall, F. H. The use of GPCR structures in drug design. *Adv. Pharmacol.* **62**, 1–36 (2011).
54. Shoichet, B. K. & Kobilka, B. K. Structure-based drug screening for G-protein-coupled receptors. *Trends Pharmacol. Sci.* **33**, 268–272 (2012).
55. Roth, C. B., Hanson, M. A. & Stevens, R. C. Stabilization of the human beta2-adrenergic receptor TM4–TM3–TM5 helix interface by mutagenesis of Glu122(3.41), a critical residue in GPCR structure. *J. Mol. Biol.* **376**, 1305–1319 (2008).
56. Xu, F., Liu, W., Hanson, M. A., Stevens, R. C. & Cherezov, V. Development of an automated high throughput LCP-FRAP assay to guide membrane protein crystallization in lipid mesophases. *Cryst. Growth Des.* **11**, 1193–1201 (2011).
57. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nature Protoc.* **4**, 706–731 (2009).
58. Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. & Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **63**, 1256–1272 (2003).
59. Katritch, V. *et al.* Analysis of full and partial agonists binding to β_2 -adrenergic receptor suggests a role of transmembrane helix V in agonist-specific conformational changes. *J. Mol. Recognit.* **22**, 307–318 (2009).
60. Katritch, V., Kufareva, I. & Abagyan, R. Structure based prediction of subtype-selectivity for adenosine receptor antagonists. *Neuropharmacology* **60**, 108–115 (2011).
61. Liu, W. *et al.* Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **337**, 232–236 (2012).

Acknowledgements

The GPCR Network acknowledges support from the US National Institutes of Health (NIH)/National Institute of General Medical Sciences (NIGMS) PSI: Biology grant U54 GM094618 and the NIH Common Fund grant P50 GM073197 to the Joint Center for Innovative Membrane Protein Technologies (JCIMPT) for technology development. The authors are grateful to the members of the Scientific Advisory Board (T. W. Schwarz, B. L. Roth, R. M. Stroud, G. Wagner, S. H. White and I. A. Wilson) and community collaborators, including A. Brooun, G. Calo, R. Guerrini, T. Handel, M. Hanson, A. IJzerman, S. Iwata, K. Jacobson, J. Javitch, T. Kobayashi, B. Kobilka, P. D. Mosier, A. H. Newman, B. L. Roth, L. Shi and P. Wells. The authors thank K. Kadyshkevskaya and I. Kufareva for assistance with figure preparation, and E. Abola and A. Walker for assistance with manuscript preparation.

Competing interests statement

The authors declare [competing financial interests](#): see Web version for details.

FURTHER INFORMATION

Community Nominated Targets programme:

<http://shkb.org/cnt>

GPCR Network: <http://gpcr.scripps.edu>

IUPHAR database: <http://www.iuphar-db.org>

JCIMPT Home Page: <http://jcimpt.scripps.edu>

NIGMS PSI: Biology programme: www.nigms.nih.gov/

Research/FeaturedPrograms/PSI/psi_biology

Protein Data Bank: <http://www.rcsb.org>

UniProt database: <http://www.uniprot.org>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF