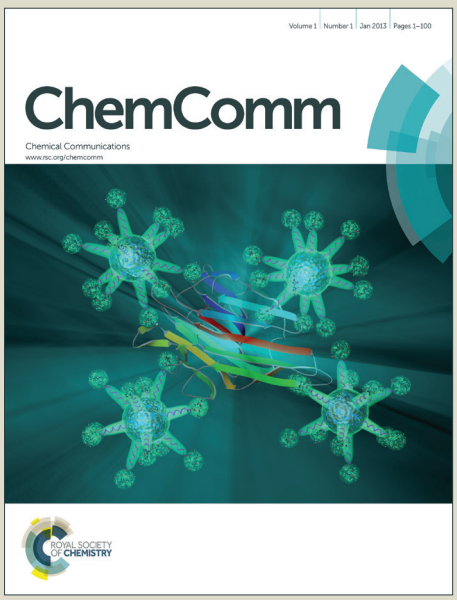
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Expanding the horizons of G protein-coupled receptor structure-based ligand discovery and optimization using homology models

Claudio N. Cavasotto[a,*] and Damián Palomba[a]

Received (in XXX, XXX) XthXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

With >800 members in humans, the G protein-coupled receptors (GPCRs) super-family is the target for more than 30% of the marketed drugs. The recent boom in GPCR crystallography has enabled the solution of ~30 different GPCR structures, what boosted the identification and optimization of novel modulators and new chemical entities through structure-based strategies. However, the number of available structures represents a small part of the human GPCR druggable target space, and its complete coverage in the near future seems unlikely. Homology modelling represents a reliable tool to fill this gap, and hence to vastly expand the horizons of structure-based drug discovery and design. In this Feature Article, we show from a wealth of retrospective and prospective studies, that in spite of the pitfalls of, and standing challenges in homology modelling, structural models have been critical for the blossoming and success of GPCR structure-based lead discovery and optimization endeavours; in addition, they have also been instrumental in characterizing receptor-ligand interaction, guiding the design of site-directed mutagenesis and SAR studies, and assessing off-target effects. Considering though its current limitations, we also discuss the most pressing issues to develop more accurate homology modelling strategies, with a special focus on the integration of computational tools with biochemical, biophysical and QSAR data, highlighting methodological aspects and recent progress. The teachings of the three GPCR Dock community-wide assessments and the fresh developments in GPCR classes B, C and F structures are commented. This is a fast growing and highly promising field of research, and in the coming years, the use of high-quality models should enable the discovery of a growing number of potent, selective and efficient GPCR drug leads with high therapeutic potential through receptor structure-based strategies.

1. Introduction: The world of GPCRs

1.1 Description and function

G protein-coupled receptors (GPCRs) are integral membrane proteins which recognize numerous messengers such as photons, odorants, neurotransmitters, fatty acids, ions, and peptides, and translate these stimuli into intracellular responses[1]. The GPCRs signalling process is linked to several physiological and pathophysiological responses affecting immune, cardiovascular and endocrine systems, among others[2-4]. Neurodegenerative, immune, metabolic, cardiovascular, psychiatric, and oncologic diseases have been tackled by a great number of drugs targeting GPCRs[5], an attractive target which currently accounts for more than 30% of the marketed drugs[6]. Considering recent efforts aimed at determining human GPCR structure and function[7], it is reasonable to expect that the number of drugs targeted to GPCRs will further increase.

With over 800 members in humans[8, 9], the GPCR super-family is usually classified into five main families[10]: class A or family 1 (rhodopsin family), which is by far the most numerous group with approximately 300 members; class B or family 2 (secretin and adhesion families); class C or family 3 (glutamate family); and the frizzled/taste2 family. GPCRs are composed of a polypeptide chain of seven α-helices crossing the cell membrane, also known as the transmembrane domains (TMs), with the N-terminus and the C-terminus located at the extracellular and intracellular side, respectively. The C-terminus possesses an α-helix (helix 8) parallel to the plasma membrane. The TMs are connected by three intracellular (ILs) and three extracellular (ELs) loops (Fig. 1).

The extracellular domains (the ELs and N-terminus) and the section of the helical-bundle facing the extracellular milieu are responsible for the binding of modulators, while the intracellular regions (the ILs and C-terminus) and the portion of the TMs domains open toward the intracellular milieu are linked to the binding of intracellular partners and the regulation of their activity[11]. Ligands can induce or stabilize different conformational states of the TMs which trigger intracellular signalling cascades controlled by heterotrimeric guanine nucleotide-binding proteins (G proteins), and whose function is related to the ability of the Gα subunit to toggle between an inactive GDP-bound conformation, and an active GTP-bound conformation that regulates the activity of downstream effector proteins[12].

In the absence of an activating ligand, GPCRs usually display basal activity that is enhanced upon binding of an agonist (full or partial), reduced by inverse agonists and unaltered by neutral antagonists[11], which block the action of both agonists and inverse agonists[13]. GPCRs can also be modulated by allosteric ligands, which bind to a site different from the orthosteric one (i.e. the natural ligand-binding site), and bitopic ligands, which have the ability to bind to both orthosteric and allosteric sites[7].

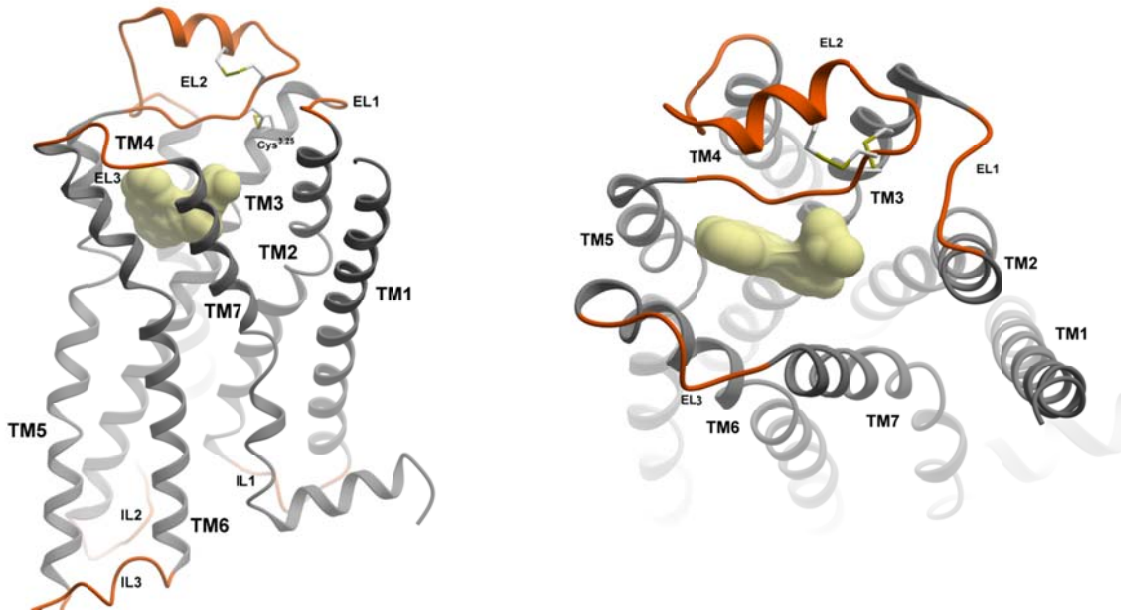


Fig. 1: Architecture of G protein-coupled receptors. Transmembrane helical regions (TM1-TM7) are shown in grey, and extracellular (EC) and intracellular (IL) loops are shown in orange. Disulfide bonds involving the EL2 are displayed in stick representation. The orthosteric binding site within TM3, 5, 6 and 7 is also displayed. Figures prepared with ICM software (Molsoft, LLC; www.molsoft.com).

1.2 The structural age: unveiling details of GPCR-ligand interaction and triggering structure-based drug design

The determination of GPCR 3D structures opened a wealth of new opportunities for structure-based virtual screening (SBVS) campaigns characterized by high hit rates and affinities, where novel ligands and new chemical entities (NCEs) were discovered[14, 15]. The advent of new structures also served as starting point for hit-to-lead optimization[6, 14, 16, 17], and has been instrumental to characterize receptor-ligand interaction, rationalize structure-activity relationships (SAR)[18], design site-directed mutagenesis (SDM) experiments, shed light into GPCR function[19], and assess off-target effects[20, 21].

The first 2D model of rhodopsin, a class A GPCR, was proposed in 1983 by Hargrave and coworkers[22]. Ten years later, a 2D projection map was calculated from two-dimensional crystals of bovine rhodopsin (bRho) by using electron cryo-microscopy[23], and based on this map, a molecular model of the receptor was built[24]. The breakthrough came in 2000 with the release of the first X-ray crystal structure of a GPCR, bRho in its inactive (dark-adapted) state covalently bound to retinal[25]. For years, bRho remained the only GPCR structure experimentally solved. It was not until 2007 that crystal structures were determined for the β_2 adrenergic receptor (β_2AR) bound to carazolol[26-29], the first druggable GPCR to be crystallized. This situation was facilitated by different crystallization strategies[16], such as the formation of fusion proteins by incorporating soluble proteins [T4 lysozyme (T4L) or apo cytochrome b_{562}RIL (BRIL)[30]] into the IL_3 or the N-terminus, the introduction of antibody fragments[31], and the insertion of mutations (thermo-stabilised receptors or StaRs[32]). These approaches were intended to decrease the flexibility of IL_3, maximize the polar surface available for crystallization, mimic part of the α subunit, and increase the conformational thermostability of GPCRs. In 2011, the experimental determination of the β_2AR bound to both an agonist and a nanobody showed the first crystal structure of a fully-activated GPCR[31]. This success was followed by a new breakthrough: the β_2AR complexed with both an agonist and the G protein, revealing for the first time the molecular details of the interaction between the latter and the intracellular surface of the receptor[33]. Thanks to these crystallization developments, several 3D structures have been solved, *e.g.*, the turkey β_1 adrenergic receptor (β_1AR)[34], the human adenosine A_{2A} (A_{2A}R)[35], histamine H_1 (H_1R)[36], dopamine D_3 (D_3R)[37], muscarinic M_2 (M_2R)[38] and M_3 (M_3R)[39], serotonin 1B ($5HT_{1B}$)[40] and 2B ($5HT_{2B}$)[41] receptors, the sphingosine 1-phosphate receptor 1 (S1PR1)[42], the chemokine receptors 4 (CXCR4)[43] and the C-C chemokine receptor 5 (CCR5)[44], the δ (OPRD)[45], μ (OPRM)[46], κ (OPRK)[47], and nociceptin (OPRX)[48] opioid receptors, the neurotensin receptor 1 (NTSR1)[49], the proteinase-activated receptor 1 (PAR1)[50], the P2Y12 ($P2Y_{12}$R)[51] and P2Y1 ($P2Y_1$R)[52] receptors, the GPR40 receptor (GPR40)[53] [also known as free fatty acid receptor 1 (FFAR1)], the orexin receptor 2 (OXR2)[54], and the angiotensin II type-1 receptor (AT_2R1)[55]. Nowadays, there are 118 class A GPCR X-ray structures, besides two class B [the corticotropin-releasing factor receptor 1 (CRF_1)[56] and the glucagon receptor (GCGR)[57]], two class C [the metabotropic glutamate receptor 1 (mGluR1)[58] and the metabotropic glutamate receptor 5 (mGluR5)[59]], and four class F [the smoothened receptor (SMO)[60, 61]].

Most GPCR structures possess a co-crystallized ligand bound, and thus three different receptor conformations can be characterized[11]: (i) an "inactive state", wherein the receptor is crystallized in complex with an antagonist or inverse agonist, (ii) an "agonist-bound state", which lacks the G protein or a substitute for it, and (iii) a "fully-active state", which consists of a trimeric complex formed by the receptor, an agonist, and the G protein (or G protein mimetic). There are also intermediate conformations among these three states enabling different structural features, which stemming from differences in the chemical structure of the

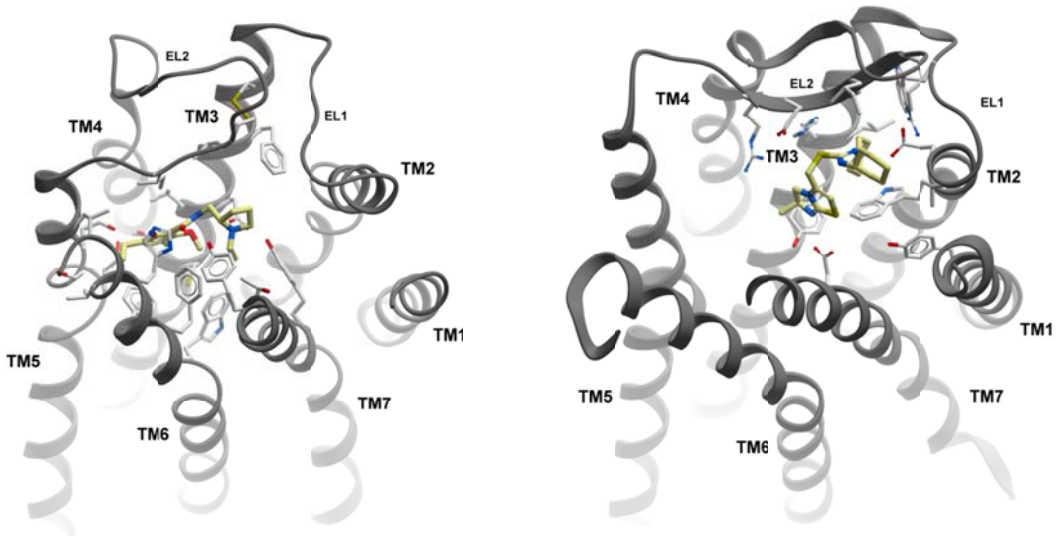


Fig. 2: Major and minor orthosteric binding sites in GPCRs. Left panel: small-molecule antagonist eticlopride bound to the major site in the dopamine D_3 receptor (PDB 3PBL), delimited by TM3, 5, 6, and 7 (in some receptors, TM4 is also involved); right panel: small-molecule antagonist isoithiourea IT1t bound to the minor site of the chemokine CXCR4 (PDB 3ODU), delimited by TM1, 2, 3, and 7. Figures prepared with ICM software (Molsoft, LLC; www.molsoft.com).

bound ligand, could be linked to partial agonism activity[62]. It has been suggested that the presence of a ligand plays an important role in GPCR stabilization. Moreover, it was even found that ligand-induced receptor conformational stabililization could help in GPCR expression[63]. Since it is believed the inactive state to be more rigid and thus more feasible for crystallization, fewer GPCRs have been crystallized in the active state[17].

Although in most of the structures co-crystallized ligands bind to the orthosteric "major" binding site delimited by TM helices 3, 4, 5, 6, and 7 (Fig. 2), isoithiourea IT1t binds to a "minor" binding site in CXCR4 delimited by TM helices 1, 2, 3, and 7[43] (Fig. 2), while peptide ligand CVX15 binds to the major binding site in the same receptor. A M_2R structure has been also co-crystallized with both an orthosteric (iperoxo) and a positive allosteric (LY2119620) ligands[64] (Fig. 3).

Different co-crystallized ligands with either the same receptor (*e.g.* $A_{2A}R$), or the same receptor subtypes (*e.g.* β_1AR and β_2AR) have provided valuable insight into ligand binding modes[65] and druggability[66]. In spite of their differences, specific GPCR-ligand interactions are necessary both for full/partial agonists and antagonists/inverse agonists[67].

The GPCR activation mechanism is a complex process in which, thanks to the structural progress over the past years, significant improvement in its understanding has been achieved[68]. Based on analyses of crystal structures of inactive and active receptors, including β_2AR, bRho, M_2R, and $A_{2A}R$, activation starts through distinct residues at the top of the different receptors, while the main processes of activation are common to all family A members[11]. The activation mechanism would affect TM3 and TM6 to generate concerted movements: the inward movement of TM5, the slight rotation and upward movement of TM3, the rotation of TM6, and the inward movements of TM7 and TM1[11]. These movements are facilitated by both the breaking of the ionic lock[11] and mainly the rearrangement of specific hydrophobic residues between TM3 and TM6.

1.3 Do we really need GPCR *in silico* models to expand ligand discovery and lead optimization?

As it has been said above, knowledge of GPCR 3D structures is a key component in structure-based drug design. To date, only ~30 different GPCR structures are available, a small fraction of the >800 GPCRs present in the human genome[8]. In spite of the recent crystallization breakthrough, a complete structural coverage of the GPCR space seems unlikely in the near future. In this scenario, accurate *in silico* modelling arises as a powerful tool approach to fill the gap. It has been shown that carefully built GPCR models accurately capture binding site structural features, and are suitable for SBDD even as x-ray structures[69-72] (see also sections 3. and 4.). Thus, modelling appears essential for SBDD on GPCRs at a genome scale

For protein modelling, two main strategies can be followed: i) Homology (or comparative) modelling, where structural models of a given protein (target) are built based on the experimentally solved structure of a homologous protein (template); ii) *De novo* modelling, whose algorithms do not rely on homologous templates and predict structures directly from sequence. In the latter category, GPCR modelling tools such as PREDICT[73] and Membstruck[74] have been developed and successfully used in SBVS[75-79]. Hybrid approaches, such as the recently developed GPCR-I-Tasser[80], should also be mentioned.

Of these approaches, homology modelling is by far the most widely used tool for GPCRs, as can be seen from the wealth of successful SBVS campaigns (see section 3.1), lead optimization endeavours (3.2), several other applications in the context of structure-based drug (3.3), and the GPCR-Dock competitions (4.), and will be the topic of this Feature Article. However, it will also be evident from this review, that in spite of truly impressive achievements, GPCR homology modelling still has limitations, and there is an actual and urgent need to address the challenges of developing more accurate modelling methods able to integrate *in silico* design with experimental knowledge[81], and benchmark

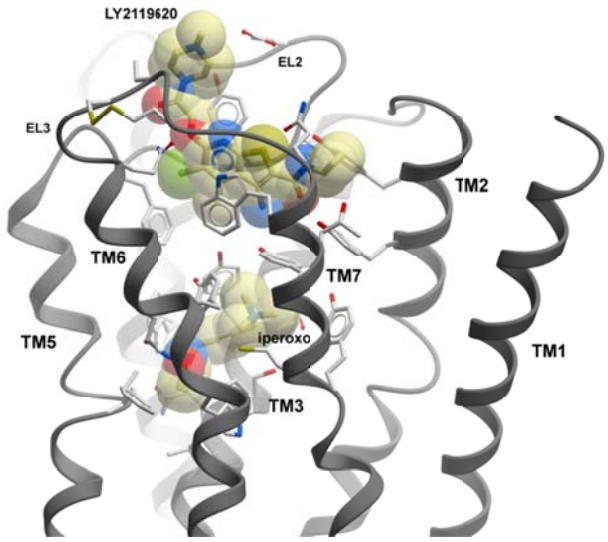


Fig. 3: Human M_2 muscarinic acetylcholine receptor bound to the small-molecule agonist iperoxo and allosteric modulator LY2119620 (PDB 4MQT). The orthosteric site is delimited by TM3, 5, 6, and 7. Figure prepared with ICM software (Molsoft, LLC; www.molsoft.com).

these methodologies in retrospective and prospective structure-based drug discovery and optimization campaigns.

2. GPCR structural homology modelling

Homology modelling aims at predicting an unknown protein structure (target or query receptor) from a related homologous protein whose 3D structure (template) has been experimentally solved[82], and consists of the following steps: i) Selection of one (or more) template(s) from a homologous protein(s); ii) target-template sequence alignment; iii) Preliminary target model (crude model) based on the template (the correspondence between amino acids in the target and template is directly taken from the alignment); iv) Refinement of the crude model (preferably in complex with a ligand), incorporating experimental data whenever available; v) Model validation.

Although model quality usually depends on the extent of target/template sequence similarity[83], in GPCR modelling, where low sequence similarity is the rule, overall structural similarity, and the presence of key conserved residues in each helix across the whole family[84] facilitate the task. A plethora of recent studies (see section 2.2), and the community-wide GPCR Dock assessments[70-72] (see section 4.) have demonstrated that crude model refinement incorporating biochemical and biophysical data yield reliable models for SBDD. Refinement is necessary especially to obtain an accurate binding site representation due to their structural differences stemming from low sequence identity, chemical diversity of GPCR ligands, and structural flexibility associated with ligand efficiency (LE)[76, 85-87].

An in-depth characterization of every step of the homology modelling process is beyond the scope of this work (the Reader may refer to Refs. [88-92] for a comprehensive description of this methodology). Instead, in this section, we will focus on four challenging and pressing issues of GPCR modelling, highlighting methodological aspects and recent advances: template selection and crude model building, loop modelling, refinement strategies, and model validation. This is followed by an up-to-date reference to available web-servers related to GPCR modelling.

2.1 Template selection and crude model building

Throughout the text we use the Ballesteros-Weinstein scheme[93], whereby residues are numbered as X.YY, wherein X represent the number of the helix in which the residue of interest is located, and YY its relative position to the most conserved amino acid in that helix, designated as number 50 (Asn in TM1, Asp in TM2, Arg in TM3, Trp in TM4, Pro in TM5, Pro in TM6, and Pro in TM7). For residues located in loops or terminal segments the sequential numbering is used.

Besides target/template TM sequence similarity, and the functional state of the receptor (active or inactive)[94], the common features in binding sites that match ligand priviledged structures[95] can be also used to select the appropriate template(s) for a given target. The recognition of specific features in the target sequence, such as amino acids responsible for helical kinks (Gly and Pro), and/or Cys residues that participate in the formation of disulfide bonds should be also accounted for in template selection[96]. In spite of the fact that the common assumption that homology model accuracy correlates with sequence similarity has been reflected in docking experiments[97, 98] and in VS studies[99-106], this has been recently challenged[107, 108], where optimal models built based on the closest related templates did not improve the VS outcome in terms of AUROC (Area Under the Receiver-Operating Characteristic curve) score[107], and BEDROC (Boltzmann-Enhanced Discrimination of ROC) and enrichment factor (EF)[108].

GPCR homology models can be built by means of either a single- or multiple-template approach[92], where the target is divided into several segments, and different templates are used to model each segment. Regardless the template approach, it is always advisable to build multiple sequence alignments across several GPCRs, from where the pairwise target/template(s) alignment (s) are to be extracted. Although the common practice is to avoid gaps in the alignment of the TMs[92], observed backbone irregularities in the TM helices of recent structures[109] should be taken into consideration at the alignment level, especially the wide π- and tight 3.10 helical turns in TM2 and TM5[110]. Even though building an initial crude model is a straightforward process[88, 111], it is advisable to account for potential structural differences such as kinks induced by Pro and/or Gly, rigid TM rotations, shifts, and tilts at this stage, since they could be critical to correctly predict GPCR-ligand interaction, and it may be difficult so solve these issues at the refinement stage[72]. The difficulty at modelling the kink induced in the CXCR4 by the $T^{2.56}XP^{2.58}$ motif[43] observed during the GPCR DOCK 2010 competition[71] clearly illustrates this point.

It should be mentioned that although the multiple-template approach often outperforms the single-template strategy in terms of structural accuracy, provided that the templates are properly selected and their sequences correctly aligned[94], it has been shown that binding site refinement using a full flexible docking approach and few geometrical constraints extracted from SDM can generate models that perform significantly better than crude models in terms of binding pose prediction, SBVS performance, and selectivity[101] (several other examples are presented in section 2.2).

2.2 Refinement strategies: Impact on retrospective docking poses and SBVS

Information inferred from SDM and SAR studies in terms of residues and ligand moieties involved in receptor-ligand interaction, respectively, and interaction patterns extracted from related GPCR-ligand crystal structures may be used to incorporate pharmacophore/geometrical constraints during the modelling process between the receptor and the ligand, or among the ligand and receptor themselves. Although it is advisable to use this information as early as possible, it has been shown to be especially valuable at the model refinement stage.

It is worth to note that analysis of SDM data has shown that it might be ligand-dependent, both in terms of ligand type (agonist or antagonist) and of different chemotypes[112]. Thus, one should be cautious when using SDM data from a given chemotype to infer interaction patterns for others[113].

Conserved interaction sites observed in the growing number of bioaminergic receptor structures, such as $Asp^{3.32}$, aromatic residues at positions 4.52, 4.56, 6.52 and 6.55, and polar amino acids at positions 5.42 and 5.46, have been successfully used to derive distance restraints to model GPCR-ligand interaction[70-72, 101]. In the $A_{2A}R$, although ligand interaction with $Asn^{6.55}$ was correctly predicted in many studies[72, 101], the lack of experimental information regarding other hydrogen-bond interactions precluded an accurate modelling of them, considering also the unpredictable fact that many of those interactions were mediated by non-conserved water molecules, thus not included in the modelling[72].

The use of distance constraints to optimize crude models represented a critical step toward high-quality homology modelling. Klebe and co-workers introduced protein-ligand restraints obtained from manual or rigid-receptor docking in the modelling procedure using MODELLER[114], and neurokinin-1 receptor (NK1R) models thus generated were successfully used in the discovery of antagonists[115].

In the ligand-steered homology modelling (LSHM) method, the binding site is co-optimized with the ligand through a flexible ligand-flexible receptor docking procedure by means of Monte Carlo sampling of the side-chain dihedral angles, and the six rigid coordinates and dihedral angles of the ligand, supplemented by receptor-ligand distance restraints whenever available from SDM or SAR data[116, 117]. The use of geometrical constraints is convenient –since it helps to decrease the number of degrees of freedom-, though not mandatory. Homology models of the melanin-concentrating hormone receptor 1 (MCH-R1) generated using the LSHM were used in a prospective SBVS campaign, where six novel low-micromolar antagonists were discovered[117]. The LSHM was further validated through cross-modelling of experimentally solved GPCR structures, observing that refined models outperformed crude models in terms of ligand pose prediction, VS performance and selectivity[101] (see also section 4.); refined models of the cannabinoid 2 receptor (CB2) using LSHM were also used for SAR data rationalization[118-120].

It should be noted that binding site optimization with non-native ligands, following the successful approach developed for protein kinases[121-123] and other receptors[124, 125], was used in crystal and modelled structures of the β_2AR for receptor ensemble docking[103], where it was observed that esemble docking outperformed the single-structure strategy.

In a method proposed by Moro at co-workers, an ensemble of ligand poses within a crude model binding site is generated using rigid receptor soft-docking followed by local energy minimization of the side chains and ligand, thus generating homology models with diverse side chain orientations[126]. The ligand is then re-docked to the best energy model. Costanzi utilized an approach wherein experimental knowledge of ligand binding is combined with *in silico* modelling of induced-fit effects[127] in order to develop β_2AR models[128].

Following this strategy, GPCR models of dopamine (D_2, D_3, and D_4), serotonin ($5\text{-}HT_{1B}$, $5\text{-}HT_{2A}$, $5\text{-}HT_{2B}$, and $5\text{-}HT_{2C}$), histamine (H_1), and muscarinic (M_1) receptors, based on the structure of the β_2AR, were created using an induced-fit docking (IFD) approach, to assess their performance in VS[129]. On models of $5\text{-}HT_{2A}$, $5\text{-}HT_{1B}$, D_2, $5\text{-}HT_{2C}$, D_3, and M_1 the Authors were able to identify active compounds from decoys, while the remaining models ($5\text{-}HT_{2B}$, D_4, and H_1) yielded poorer outcomes, probably owing to difficulties in modelling the EL_2; the same strategy was used to probe whether the availability of a novel structure of the closely related D_3 receptor would allow the construction of reliable models of D_2R and D_1R[108]; the Authors stressed that the ligand employed in the IFD procedure is a determinant factor, much more important for the performance of homology models in VS studies than the choice of template or the model preparation method. The IFD method was also used to develop optimized binding sites of the acetylcholine muscarinic receptors, where it was concluded that the optimization stage including functional knowledge has a stronger impact on model quality than target−template sequence similarity[130].

In a study of VS on a set of MT_2 melatonin receptor models[131], ligands were placed within the MT_2 modelled binding sites according to SDM data and pharmacophore modelling, and the complexes were refined using IFD. It was shown that most of the ligand-adapted MT_2 receptor models displayed important improvements in VS enrichments compared to the unrefined homology models[131].

Chin et al. developed human M_1R homology models based on the crystal structure of the rat M_3R, and then modified them by using the agonist-bound crystal structure of a β_2AR[132]. The binding sites were then refined by IFD with acetylcholine; it was observed that the models developed could be successfully used to detect agonists.

In the community-wide assessment of GPCR structure modelling and ligand docking 2008 (GPCR Dock 2008)[72] a β_2AR-based homology model combined with the ligand-guided backbone ensemble receptor optimization (LiBERO) technique was used to predict the structure of the human $A_{2A}R$ complexed with antagonist ZM241385[133]. Multiple conformations of the protein backbone were generated using heavy-atom Elastic Network Normal Mode Analysis (EN-NMA), which was followed by docking ligands into the models with flexible side chains. The models thus generated were clustered and validated through small-scale retrospective VS; the modelling of the non-conserved part of the EL_2 (residues G^{142} to A^{165}, which was the unaligned portion that was not included in the initial $A_{2A}R$ model) was performed using the ICM[134] loop modelling algorithm based on global minimization of the conformational

energy imposing disulfide bond restraints. Finally, the optimized binding site and the EL_2 conformational ensemble were ranked according to their conformational energy. The LiBERO approach was also utilized in the same assessment, although using the turkey β_1AR as template[72].

Molecular dynamics (MD) is also a useful strategy to optimize receptor-ligand interactions[135-138]. Using dynamic homology modelling[135], the activated state of β_2AR was modelled based on the "active" opsin structure, without adding any experimental information. Free MD simulations in an explicit membrane/solvent environment were conducted and representative binding modes were extracted by hierarchical clustering of interaction fingerprints (IFPs)[139]. These binding modes were assessed in VS studies in which outperformed the X-ray structure of the inactive β_2AR in prioritizing agonists over antagonists/inverse agonists[140].

Explicit-solvated MD simulations of four GPCR-ligand bound complexes ($CXCR_4$ and D_3R X-ray structures, and H_4R and 5-HT_6 homology models) were undertaken in lipid bilayers in order to develop discrete protein conformations, and thus to characterise binding site flexibility[141]. Representative structures from a RMSD-based clustering were compared to crystal structures and models, and it was observed that MD snapshots outperformed X-ray structures and homology models in terms of VS enrichment, what according to the Authors, was probably because protein conformations from MD are less biased toward a specific chemotype.

2.3 Modelling the loops

Modelling extracellular and intracellular loops in GPCRs is still a highly difficult task due to their high sequence and structural variability, as observed in the available crystal structures[142, 143]. Moreover, the substantial length of some loops, *e.g.* IL_3, hinders any attempt to successfully model them, thus being suitable to directly omit them[92].

The EL_2 links TM4 and TM5 and, in many class A GPCRs, features a highly conserved Cys residue that makes a disulfide bond with $Cys^{3.25}$ (Fig. 1). A great structural variability of EL_2, as well as a diverse array of disulfide bonds involving Cys residues of this loop have been observed among the several experimentally solved GPCR structures[144]. In an early study on bRho, Cavasotto et al. showed that omitting the EL_2 had no impact in redocking the co-crystallized ligand retinal, while it had a minor impact in retrospective VS[145]. On the same line, Nikiforovich and co-workers et al. showed that docking to loop-less crystal structures of β_1AR, β_2AR, and A_{2A}R was as good as or better than with modelled loops, in terms of binding mode prediction[146]. A study by de Graaf et al. on D_2R, A_3AR, and thromboxane A2 ($T_{A2}R$) models revealed that loop-less models of D_2R and $T_{A2}R$ were able to discriminate ligands from decoys in retrospective VS, while EL_2 modelling was only important for A_3AR[147]. This suggests that EL_2 modelling should be conducted using experimental restraints whenever available, while the impact of adding ELs should be evaluated by retrospective SBVS[65].

Recently, however, several *de novo* strategies have been introduced as alternatives for loop modelling.

By means of the Protein Local Optimization Program (PLOP), which employs a refined sampling grid, an all-atom energy function with implicit solvent, and an accurate side-chain packing algorithm, Goldfeld *et al.*[148] were able to restore the conformation of ILs and ELs of bRho, $A_{2A}R$, β_1AR, and β_2AR in their native environment. In addition, in order to deal with cases wherein loops and membrane have important interactions, they performed explicit membrane simulations where the lowest energy conformers for both short and long loops matched the corresponding crystal structures. Later, PLOP was used to predict the same ILs and ELs, both with TM domains fixed in their crystallographic positions, as well as with a homology model of β_2AR[149]. According to the Authors, this was the first successful study of an RMSD validated, physics-based loop prediction within the framework of GPCR modelling.

The EL_2 structure was predicted in 13 GPCRs by means of the CABS (C-Alpha, Beta, and Side chain) protein modelling tool[150], which is based on a coarse-grained structure representation and a Monte Carlo (MC) dynamics sampling scheme[151, 152]. The modelling approach used experimental constraints on disulfide bonds, yielding ensembles of low-energy conformers with modest computational resources. A Metropolis Monte Carlo (MMC) method has been used to model the three ELs of the transmembrane domains of the thyroid-stimulating hormone receptor (TSHR) by employing a local torsion move and a grid-based force-field method[153].

It should be mentioned that beyond the *de novo* methods, there are also computer programs and web servers which are intended to predict the loop structure. Examples include ModLoop[154], which predicts the loop conformations by satisfaction of spatial restraints, without depending on a database of known protein structures; Rosetta[155], a combined approach of fragment-based and *de novo* prediction for loop modelling; and SuperLooper[156], a knowledge-based method which predicts loop conformation from a database of known loop structures.

2.4 Structural model validation

In order to assess the actual usefulness of a homology model, validation is an essential step, regardless the target protein under study. As a basic premise, the intended application of the model should determine its desired quality[157]. Medium-quality models may be adequate for conducting mutagenesis experiments, while high-quality models are required for SBVS studies as well as mechanistic analysis. Typically, an "internal" evaluation is undertaken so as to guarantee that the model stereo-chemistry (*e.g.* bond lengths and angles, dihedral angles, and non-bonded contacts) is within acceptable limits. This can be assessed by employing computer programs such as PROCHECK[158], WHATCHECK[159], and MolProbity[160]. Despite the fact that structural properties outside the normal range could hint serious errors in the model, a successful internal consistency check in no way guarantees that the model is indeed a correct representation of the actual structure of the target.

In the context of GPCRs, retrospective docking has appeared as an efficient approach to validate homology models[117], in which a dataset of known ligands is merged with a decoy library – preferably an un-biased one[161], and docked to the models. Binding pose prediction, and/or the ability to prioritize ligands over decoys (assessed by EFs and/or area under the ROC curve), may be taken as a measure of the quality of the model[102, 103, 108, 117, 130, 131, 141, 162] (see section 2.2). Experimental knowledge

inferred from SDM, and/or quantitative SAR (QSAR) can not only be used to construct binding hypotheses to guide modelling (see section 2.2), but also to further examine and validate modelled GPCR-ligand complexes[163-165]. The successful application of modelled binding sites in prospective docking and lead optimization (sections 3.1 and 3.2, respectively) is a further step toward model validation.

For a proper interpretation of the results, it should be taken into account that the performance of homology models in VS experiments may depend on other factors not related to the modelling process itself, such as the availability of template structures, the docking program of choice, the ligand and decoy dataset[161, 166], small-molecule preparation, the specific target[167], the presence/absence of water molecules[161, 162], and whether receptor flexibility is accounted or not in the docking process[103, 121, 168, 169].

2.5 Useful web-servers in GPCR modelling

Today, many resources and tools aiding homology modelling of GPCRs are available, *e.g.* repositories of models, servers to perform homology modelling, and ligand databases, among others. Assessment meetings of protein structure prediction methods, particularly CASP[170] and GPCR-Dock[70-72], paved the way for the improvement of these services and have become the prediction of the protein structure an attainable work.

GPCRM[171] is an online platform for predicting GPCR structures, which combines several strategies for template detection, alignment generation, model building, loop refinement and model filtering based on the Z-coordinate, with the option of human intervention. Homology models are created by utilizing multiple template structures and profile-profile comparison. GPCRM provides the 10 top-scored models according to the Modeller DOPE score[177] and the Rosetta total score[178]. The URL corresponding to the GPCRM server, and the homology modelling web tool described in this section are listed on Table 1.

Table 1: On-line tools for GPCR homology modelling

Resource Name	URL	Ref.
GPCRM	gpcrm.biomodellab.eu/	171
GPCR-SSFE	www.ssfa-7tmr.de/ssfe/	172
GOMoDo	molsim.sci.univr.it/cgi-bin/cona/begin.php	173
GPCR-ModSim	gpcr-modsim.org/	174, 175
GPCRautomodel	genome.jouy.inra.fr/GPCRautomdl/cgi-bin/welcome.pl	176
GPCR-I-TASSER	zhanglab.ccmb.med.umich.edu/GPCR-I-TASSER/	80

The GPCR-Sequence-Structure-Feature-Extractor (GPCR-SSFE)[172] is a server that offers template predictions, sequence alignments, structure motifs and homology models of the transmembrane helices of 5025 class A GPCRs. The pipeline is based on a fragment approach that takes advantage of available family A crystal structures. Users are able to access the models stored either by browsing the GPCR dataset in accordance with their pharmacological classification or searching for results using a UniProt identifier.

The GPCR Online Modeling and Docking server (GOMoDo)[173] carries out automatic homology modeling, and either a blind or an information-driven ligand docking of GPCRs by combining different bioinformatic tools. It utilizes the HHsearch[179] for performing sequence alignment, MODELLER[180] for building a 3D model of a given sequence, the VADAR server[181] for verifying the obtained 3D model, AutoDock VINA[182] or HADDOCK[183] for docking small-molecules uploaded by users, Fpocket[184] for binding sites prediction, and LovoAlign[185] for conducting structural alignment of models needed for VINA docking.

GPCR-ModSim[174, 175] is a web-based service for homology modeling and all-atom MD equilibration of GPCRs. This server is intended to obtain the most accurate structural and dynamic information for a given GPCR, and it provides a stand-alone protocol for all modelling steps.

The GPCRautomodel[176] site is aimed at conducting automatic homology modeling of GPCR structures. In a first step, it uses a threading-based method to obtain a 3D model. In a second stage, it performs docking of selected small-molecules with the modelled receptor by utilizing VINA[182].

The GPCR-I-TASSER method has been already mentioned in section 1.3

These web-servers have been used in several applications. By way of illustration, to study protein-protein interaction of the human α_{2C}R with the human Filamin-2 protein[186], to rationalize SAR of A_{2A}R ligands[187], and even in the GPCR Dock 2013 assessment in the sequence alignment and template selection for the successful prediction of the 5-HT_{1B} and 5-HT_{2B} receptors in complex with ergotamine[188].

Other web tools which may aid in the homology modelling process, such as model and motif databases, chemical libraries, docking portals, among others, are listed in Table 2.

3. Structure-based drug design using GPCR homology models

3.1. Discovery of new ligands through virtual screening

Using GPCR crystal structures, ligands have been discovered for various receptors with both high hit rates (actives/tested) and structural novelty[15]. New antagonists for β_2AR[189-191], A_{2A}R[162, 192], D_3R[99], and H_1R[193] with hit rates between 20% and 73%, and at least 2 new scaffolds per receptor, were discovered (for a review of recent SBDD approaches using GPCR crystal structures cfr. Ref. [16, 194, 195]). Furthermore, in GPCR docking campaigns, hit rates and affinities in GPCRs were two to three log-orders better than those against soluble proteins[15]. It has been suggested that two main elements may be responsible for this: i) supposedly unbiased chemical libraries actually possess a large quantity of molecules with structural features in common with GPCR ligands; ii) the well-buried GPCRs orthosteric binding sites favours the identification of small molecules with high LE[15].

The use of homology GPCR models has been also instrumental for the discovery of new ligands even since bRho was the only available template. Early successfully prospective SBVS campaigns included bioaminergic receptors (α_{1A}R[196], D_3R[137], H_4R[197]), chemokine receptors (CCR4[198], CCR5[199]), peptide receptors [NK1R[115], formylpeptide receptor (FPR1R)[200],

Table 2: On-line tools useful in the GPCR homology modelling process

Resource Name	Application	URL	Ref.
GPCRpred	Server for prediction of GPCR families and subfamilies	www.imtech.res.in/raghava/gpcrpred/	203
GPCRHMM	Server for putative GPCR detection from sequence and TM segment localization prediction	gpcrhmm.sbc.su.se/	204
GPCR-HGmod	Database that contains 3D structural models of GPCRs in the human genome	zhanglab.ccmb.med.umich.edu/GPCR-HGmod/	-
GLASS	Repository for experimentally-validated GPCR-ligand interactions	zhanglab.ccmb.med.umich.edu/GLASS/	205
GPCR-exp	Database of experimentally-solved GPCR structures	zhanglab.ccmb.med.umich.edu/GPCR-EXP	-
Adenosiland	Integrated bioinformatics and chemoinformatics web-resource dedicated to adenosine receptors	mms.dsfarm.unipd.it/Adenosiland/	206, 207
GPCRserver	Server for GPCR identification and TM region prediction	genomics.fzu.edu.cn/GPCR/index.html	208
GPCR structure and VS library	GPCR modelling and virtual screening database	cssb.biology.gatech.edu/skolnick/webservice/gpcr/index.html	209
GPCRdb	Contains data, diagrams and web tools for GPCRs	gpcrdb.org/	210
GLL/GDD	Ligand libraries (GLL) and docking decoy databases (GDD) for 147 GPCRs	cavasotto-lab.net/Databases/GDD/	161
PDBTM	Protein data bank of transmembrane proteins	pdbtm.enzim.hu/	211
TinyGRAP	GPCR mutant database	www.cmbi.ru.nl/tinygrap/credits/	212
MPSTRUC	Database of membrane protein of known structure	blanco.biomol.uci.edu/mpstruc/	-
MPtopo	Database of membrane proteins with experimentally-validated TM segments	blanco.biomol.uci.edu/mptopo/	213
GPCR network	Portal of the PSI:Biology GPCR network	gpcr.scripps.edu/	-
GPCR-OKB	Information management system for GPCR oligomerization	filizolalab01.mssm.edu:8080/gpcr-okb/	214
GPCR NaVa	Database that describes sequence variants within the GPCR family	nava.liacs.nl/	215
IUPHAR GPCR database	Expert-driven knowledgebase of GPCR drug targets and their ligands	www.guidetopharmacology.org/GRAC/ReceptorFamiliesForward?type=GPCR	216
GLIDA	GPCR-ligand database	pharminfo.pharm.kyoto-u.ac.jp/services/glida/	217
GPCR-RD	Database for experimental restraints of GPCRs	zhanglab.ccmb.med.umich.edu/GPCR-RD/	218
GPCR SARfari	Integrated chemogenomics workbench focussed on GPCRs	www.ebi.ac.uk/chembl/sarfari/gpcrsarfari	-

MCH1R[116, 117]], cannabinoid receptors (CB2[201]), and purine receptors [free fatty acid receptor 1 (FFAR1)[164, 202]]. The cascade of new GPCR structures triggered by the release of the β_2AR in 2007 not only dramatically enhanced SBVS on crystal structures[16, 194, 195], but also provided structurally diverse templates for further improving GPCR models, and thus greatly expanding its use in drug design.

Inspired by the challenge of the GPCR Dock 2010 assessment[71], in which the modelling community aimed to predict the structure of the D_3R-eticlopride complex, Carlsson et al. developed a homology model of D_3R and docked more than 3.3 million molecules against it, repeating this experiment on the crystal structure of the D_3R-eticlopride complex once it had been released[99]. Concerning the model, six compounds were discovered with binding affinities in the range 0.2-3.1 μM, and one of them was subsequently optimized to 81 nM. With respect to the crystal structure, five compounds were found in the 0.3-3.0 μM range. Moreover, the hit rate for the screening on the homology model was 23% and on the crystal structure was 20%. Thus, the hit rates using the model and the crystal structure were basically equivalent. Each VS returned two novel scaffolds, different from known ligands, and among themselves. Furthermore, the active molecules found in the screening from the homology model displayed no measurable affinity for the template used in the modelling (β_2AR).

In the same context as the previous work, Mysinger et al. docked over 3 million molecules against a homology model of the CXCR4 and the crystal structure[100]. A single antagonist was found in docking against the model, which was similar to known ligands and possessed a modest specificity. The hit rate using the model was 4%, while the screening on the crystal structure yielded not only a higher hit rate (17%), but also four antagonists that were different from known scaffolds, substantially smaller than most known ligands, and specific for CXCR4. One of them had an IC_{50} value of 0.31 μM and a LE of 0.36 (placing it in the lead-like range of compounds for oral drugs), and all ligands inhibited CXCR4-mediated chemotaxis in cell culture. When comparing these two targets (D_3R and CXCR4) and these four virtual screening campaigns, the Authors drew two conclusions: first, an important factor was the ligand bias in the used database

(ZINC[219]) toward biogenic amine mimetics, rather than to CXCR4-like ligands; unlike D_3R ligands, there are relatively few molecules sharing the same size and charge properties as known CXCR4 ligands. Second, the relatively poorer result of screening against CXCR4 homology models might be related to the sequence identity with the structural templates. They suggested that accurate models may be developed for GPCRs that share ~40% or higher sequence identity, and with enough mutagenesis information (as for D_3R). On the contrary, for targets with significantly lower sequence identities, ranging from 18 to 25% (as for CXCR4), homology models suitable for drug discovery might be "out of reach".

On a homology model of $A_{2A}R$ built from the rβ_1AR, an array of agonists with diverse ligand efficiencies was discovered through SBVS, with a hit rate of 9%[220]. Hits were furthered optimized for affinity and selectivity (cf. section 3.2).

Ligand- and protein-based molecular fingerprints were applied in a virtual screening of fragment-like molecules on the H_3R[221]. The FLAP (Fingerprint of Ligands And Proteins)[222-224] method was used in a H_3R model based on the H_1R crystal structure, and refined by means of molecular docking and MD simulations with H_3R actives. The best structures for each complex were chosen on the basis of the ability to distinguish between known fragment-like H_3R actives and inactive ones in retrospective VS studies. Using a collection of 156,090 molecules filtered from the ZINC database, a prospective VS on FLAP models resulted in 18 experimentally confirmed hits, with affinities in the range of 0.5-10 μM. Moreover, these confirmed H_3R hits did not show affinity for H_4R.

Multiple homology models were developed for the A_1R, using the crystal structure of $A_{2A}R$ as template, and approximately 2.2 million lead-like compounds were docked into the models[225]. With the aim of examining the intrinsic selectivity of the models, all high-ranking molecules were tested in binding assays not only on the A_1R but also on $A_{2A}R$ and A_3AR. The screening exhibited a hit rate of 21% and the most potent compound had a K_i of 400 nM, although it yielded few selective compounds. The Authors drew three conclusions from this study: i) Even when screening is performed with the same library, distinct models of the same receptor return distinct sets of ligands; in this sense, model performance varied widely in terms of both the absolute number of actual ligands and their selectivity; ii) homology models seem to work well in GPCR docking, as evidenced by the outcomes; iii) by means of applying docking to solely one receptor subtype, obtaining selective compounds is a difficult task for targets with high degrees of similarity, e.g. the adenosine receptors.

A homology model of the D_2R in the active conformation based on the active β_2AR crystal structure was built, and a prospective VS of 2.7 million "lead-like" and 400K "fragment-like" molecules from the ZINC database was conducted against it[226]. Out of three actives found in functional assays, two were agonists and one was an inverse-agonist. However, these three hits had low affinity, the agonism was weak, and they were similar to known dopamine receptor ligands, indicating that the active β_2AR structure might not be a proper template for the active D_2R. These outcomes suggested that although the β_2AR structure possesses a high sequence identity and it was the right template for the inactive conformation[99], structural information obtained from the active β_2AR was not transferable to the active D_2R structure. The Authors argued that this fact might be either a singular case, or related to their modelling approach. Thus, the agonist state might be specific for any given GPCR−ligand pair.

A VS on CXCR7 homology models was undertaken using a dataset of commercially available compounds and a new modelling method based on multiple GPCR crystal structures[227]. The CXCR4 structure, and the structures of bRho, β_2AR, β_1AR, and $A_{2A}R$, were used as the "principal template" and "supplementary templates", respectively. Twenty-one novel hits with IC_{50} values ranging from 1.29 to 11.4 μM and a variety of scaffolds were determined. Furthermore, salt bridges between $Asp^{4.61}$ and $Asp^{6.58}$ and protonated nitrogen atoms of the ligands, as well as π–π stacking interactions between $Trp^{2.61}$ and ligands were found relevant for CXCR7 ligand binding.

Schmidt et al. docked over 2 million compounds from the ZINC database to CXCR3 homology models and to the CXCR4 crystal structure, respectively, in order to find both dual modulators and selective compounds for each target[228]. They identified selective and non-selective ligands, which were confirmed by *in vitro* assays for both receptors. Eleven novel ligands for both targets were found, with high hit rates of 57% (CXCR3-selective), 50% (CXCR4-selective), and 50% (dual binders). Most of these hits exhibited binding constants in the low-nanomolar range, and very good LE indices. It is worth noting that high hit rates were achieved in each category, even the hit rate for the CXCR3 model was higher than the one for the CXCR4 crystal structure. Moreover, the CXCR3 model did not seem to suffer template bias according to the number of potential dual modulators and the hit rate found in that category. Furthermore, all but one binder detected in this study possessed chemistry features different from known ligands of both targets from the ChEMBL database[229].

A combined ligand- and structure-based strategy for identifying H_4R antagonists was recently developed, where initially, a ligand-based VS of the ZINC database was conducted to select potential H_4R antagonists (focused library), and several H_4R homology models were built using the H_1R crystal structure as template and refined with MD in a fully atomistic lipid membrane environment[230]. Structural models were validated by their ability for discriminating active from non-active H_4R antagonists in docking using a validation set extracted from the ChEMBL database. Finally, the best model was used to screen the focused library, and thus 11 drug candidates were obtained and presented as novel lead compounds.

A hybrid strategy combining a structure- and ligand-based method was developed and used to identify novel nociceptin (NOP) ligands[231]. Homology models of the binding site of the active-state NOP receptor were built based on the opsin structure using simulated annealing, and then ranked according to the EF in retrospective docking. A structural refinement followed employing a shape-based similarity strategy along with molecular docking of known NOP agonists. Virtual screening of the CNS Permeable subset of the ZINC database was undertaken utilizing a ligand pharmacophore- and shape-based protocol, followed by a structure-based step using the refined NOP active-state conformations obtained in the enrichment calculation. Molecules containing a piperazine ring were eliminated due to off-target

effects. Small-molecules were ranked according to a consensus score, and 20 compounds were purchased and tested in binding affinity assays. From the better six compounds, four had binding affinities less than 50 μM. Further, one had a K_i of 1.5 μM and represented a NCE.

A structure-based virtual fragment screening was carried out both on the D_3R crystal structure and on a H_4R homology model (based on the H_1R crystal structure)[232]. By means of all-atom membrane-embedded MD simulations, representative receptor conformations for both targets were generated, and a library consisting of 12,905 fragments was docked on the conformational ensemble of both structures. *In vitro* assays showed hit rates in the range of 16-32%, and K_i values in the range of 0.17-2.8 μM for D_3R, and 8.4-75 μM for H_4R. Moreover, the hits possessed high LE, with values in the 0.31-0.74 range, and an admissible lipophilic efficiency. The crystal structure, homology model, and ensemble docking provided all valuable hits with little overlap. Moreover, the single homology model outperformed the single crystal structure in terms of hit rate. However, in this particular case, the ensemble docking strategy was not better than the single structure docking method, both approaches thus being complementary. The Authors thus concluded that a combined approach should be followed to maximize hit retrieval.

A structure-based VS and a functional cell-based screening were undertaken in order to identify adrenergic α_{2C}AR receptor agonists[233]. A homology model of the activated α_{2C}AR was built based on the human active-state β_2AR crystal structure, and the best conformation for VS was chosen based on retrospective docking. A library of 3071 fragments was experimentally screened, and also docked to the model, exhibiting a hit rate of 6.7% and an EF of 12. Moreover, 2 fragments out of the 16 detected hits were identified by VS at the top 1% of the screened library, and showed themselves as specific ligands of α_{2C}AR.

A structure-based virtual fragment screening along with an IFP scoring method was performed against optimized homology models of the H_4R built using the β_2AR and H_1R crystal structures as templates[234]. On the basis of the retrospective VS analysis, two β_2AR-based H_4R models and their corresponding IFP references were employed in the VS using molecules extracted from ZINC. Six compounds were confirmed as H_4R ligands, with pK_i values ranging from 5.2 to 6.8. None of the hits possessed detectable binding affinity for β_2AR, proving that the method did not suffer from template bias. Afterwards, the VS was conducted against the H_1R-based H_4R models and three hits were found. Altogether, nine compounds were confirmed as hits with binding affinities for H_4R in the range of 0.14-6.9 μM, representing five distinct scaffolds.

3.2. Getting it better: Hit-to-lead optimization

Although several hits discovered through SBVS have been optimized for affinity[117, 164, 235], there are not too many actual structure-based guided optimization studies (cf. Refs. [14, 236] for a review of early uses of GPCR models in lead optimization).

As described in section 3.1, Carlsson et al. performed a structure-based guided optimization of a D_3R SBVS hit, reaching an affinity of 81 nM[99]. Hit molecules discovered through SBVS against an $A_{2A}R$ model based on the β_1AR were optimized to selective and potent lead molecules using a structure-based design, and synthetized[220]. Substitution of the propenyl-thiophene ring[237], and replacement of the chromone ring[220] resulted in molecules with improved affinity and selectivity toward $A_{2A}R$, and selectivity toward $A_{2A}R$, respectively.

With the aim of identifying H_1-H_3 dual antagonists suitable for intranasal administration from phthalazinone analogues, a H_1R homology model based on the crystal structure of bRho was built and complexed with the second-generation of anti-histamine azelastine, what furnished evidence that the incorporation of certain fragments related to H_3R antagonism should bring about dual H_1-H_3 antagonism[238]. A series of H_1-H_3 dual antagonists were synthesized and two compounds showed a slightly lower potency toward H_1R, but a much higher potency toward H_3R than azelastine, the clinical gold-standard. Moreover, one of them exhibited improved *in vivo* pharmacokinetic properties compared to azelastine.

Novel selective CysLTR2 antagonists were discovered using a homology model of CysLTR2 built from the crystal structure of bRho as template, and refined by MD simulations[239]. Based on the proposed binding mode of the selective lead antagonist HAMI3379, a series of dicarboxylated chalcones was docked within the binding site, and six promising hits were synthesized and tested for CysLTR2 antagonism, two out of which showed potent and selective CysLTR2 antagonism with IC_{50} values of 7.5 and 0.25 μM.

Using a homology model of the CB_2 constructed using the crystal structure of β_2AR as template, and refined by MD simulations, 3D-QSAR models were generated from comparative molecular field analysis (CoMFA[240]) using 2-quinolone and 2-pyridone coumarin CB2 leads[241]. In accordance with pharmacophoric features derived from the 3D-QSAR model, a series of coumarin derivatives was subsequently designed, and SAR studies were carried out. Several compounds showed high selectivity for CB2 against CB1, among which one CB2 agonist [EC_{50} = 0.103 μM, selectivity index (SI) > 97], and one CB2 antagonist (IC_{50} = 0.019 μM, SI >500).

Homology models of the human (h) and mouse (m) A_3ARs based on a hybrid template (crystal structures of agonist-bound $hA_{2A}AR$, and active $h\beta_2AR$) were designed in order to develop sulfonated nucleoside ligands for A_3AR with affinity independent on the species[242]. Molecular docking studies of (N)-methanocarba derivatives were undertaken to model key interactions between these nucleosides series and the h- and m-A_3ARs, and thus guide substitutions at the C2 and N^6 positions for chemical synthesis. Based on this interaction analysis, the sulfonate groups on C2-phenylethynyl substituents would produce high affinity at both h- and m-A_3ARs, whereas a N^6-p-sulfophenylethyl substituent would exhibit higher hA_3AR than mA_3AR affinity. Insights gained from modelling were confirmed by pharmacological studies, wherein one agonist analogue is bound selectively to h/m A_3ARs [K_i (hA_3AR) = 1.9 nM] and the corresponding p-sulfo isomer showed mixed A_1/A_3AR agonism. Subsequently, using the same A_3AR hybrid model, the Jacobson group[243] conducted molecular docking studies of (N)-methanocarba adenosine 5′-uronamides derivatives with the aim of identifying highly selective agonists of the A_3AR, but lacking the arylethynyl group, linked to potential liver toxicity. A planar C2-triazole linker in place of an ethynyl group showed to be the best substitution which favours selective binding to the A_3AR. Several analogues

with N^6 and C2 substitution were synthesized, and pharmacologically and *in vivo* characterized. All of the derivatives exhibited K_i values ranging from 0.3 to 12 nM at the A_3AR and one of them achieved a highly prolonged and full efficacy in controlling mechano-allodynia (> 90% protection up to 4 h).

Yaziji et al.[244] synthesized two series of diaryl 2- or 4-amidopyrimidines and determined their affinities for the four human adenosine receptors (A_1R, A_{2A}R, A_{2B}R, and A_3R). Based on the results of the first series, the design of both the second set of compounds and new derivatives exploring the alkyl substituent of the exocyclic amide group was performed. This synthesis was assisted by means of an approach that combined molecular docking to a hA_3R homology model (built using the crystal structure of A_{2A}R as template) and 3D-QSAR analysis. As a result, four compounds displayed both remarkable affinities ($K_i \leq$ 6 nM) and selectivity toward the A_3R subtype. Subsequently, the same research group examined the impact of methoxyaryl substitution patterns on N-(2,6-diarylpyrimidin-4-yl)acetamides with the aim of modulating the A_3R antagonistic profile[245]. A homology model of the hA_3R was developed using as template the inactive structure of A_{2A}R and molecular docking as well as 3D-QSAR studies were carried out. Guided by the modelling results, a focused compounds library was synthesized and its pharmacological profile was studied for the four human adenosine receptor subtypes. Novel A_3R antagonists were reported, which showed excellent potency (K_i < 20 nM), wherein two ligands are highlighted with a K_i < 7 nM and highly selective profiles among ARs. The most important features of the pipelines used in the research projects aimed at targeting the A_3Rs by Sotelo and coworkers are explained in Ref. [246].

3.3 Recent applications of GPCR homology models in other structure-based drug design scenarios

Besides the use in ligand discovery through VS (section 3.1) and structure-based lead optimization (section 3.2), GPCR homology models are invaluable to study off-target effects using docking, guide the design of small-molecule and peptide ligands, rationalize SAR data, design SDM experiments, characterize receptor-ligand interaction, and complemented with MD, to understand ligand binding mechanisms and protein dynamics. In Table 3 we present recent applications of GPCR models in several structure-based drug design scenarios.

4. Modelling and docking accuracy

Although homology models are usually used when experimental structures are not available, retrosprospective modelling and comparison with crystal structures, and retrosprospective docking analysed in terms of ligand RMSD (if known) and enrichment data, furnish valuable information in terms of methodology, strategies, and further developments needed.

Using the experimentally solved structures of bRho, β_2AR, A_{2A}AR, the LSHM method (section 2.2) was validated through cross-modeling, and performance of the thus generated models investigated in docking experiments[101]. This methodology was able to generate quality models of the receptors complexed with their co-crystallized ligands (~1 Å for β_2AR modelled using bRho or A_{2A}AR as templates; 2.8 Å for A_{2A}AR using β_2AR as template). It was also observed that: i) LSHM performed better than templates, crude models, and random ligand selection in small-scale high-throughput retrospective docking; ii) higher quality models typically displayed better enrichment in docking. Interestingly, homology models were found to be reliable for selectivity prediction. Clearly, these results support the fact that the LSHM method can successfully characterize GPCR binding sites through a fully flexible ligand- receptor approach. It should be noted, however, that models underperformed with respect to crystal structures in terms of docking enrichment and selectivity prediction, likely because of inaccuracies at the backbone level.

The community-wide GPCR modelling and docking (GPCR Dock) assessment was established to monitoring and stimulate the advancement of GPCR structure prediction and ligand docking, as well as emphasizing areas for methodological improvement. The rationale and organization of GPCR Dock is analogous to the way of CASP (Critical Assessment of methods of Protein Structure)[247] and CAPRI (Critical Assessment of PRediction of Interactions)[248, 249]. In the GPCR Dock blind prediction assessment, the participants predict and submit models of a receptor-ligand complex from the sequence of the receptor and a 2D representation of the ligand prior to the public release of the 3D coordinates of the complex.

The first round of GPCR Dock was carried out in October 2008 in conjunction with the public release of the crystal structure of the human A_{2A}AR bound to the high-affinity antagonist ZM241385[35, 72], where 29 groups participated. The most successful models, which had an average heavy-atom RMSD of 2.8 Å for the ligand, and 3.4 Å for the residues of the binding site, were constructed by homology modelling taking into account the β_2AR structure as template, which shares ~35% sequence identity with A_{2A}AR, and using experimental information derived from SDM. However, they could not account for most of the receptor-ligand contacts (only ~50%) and rank models properly. In fact, most of the participants were far from accurately predicting the native ligand pose and the correct conformation of EL_2, which has a lower degree of sequence similarity and structural conservation. The EL_2 was *de novo* modelled in many predictions, although the best approach (S. Costanzi) utilized a combination of homology modelling (in a short segment around a conserved cysteine residue) along with *de novo* modelling for the remainder residues of the loop. The crystal structure also revealed four well conserved water molecules around the ligand, but none was included in the submitted predictions. Even though it can be shown that ZM241385 pose can be recovered upon docking with no waters[72, 161], waters may be necessary for a more accurate binding pose prediction and for binding free energy calculations.

The second round, GPCR Dock 2010[71], was performed in parallel with the solution of the crystal structures of the D_3R[37] and the CXCR4[43] so as to model three different classes of complexes showing three levels of difficulty: i) the small-molecule antagonist eticlopride bound to hD_3R, which has two close structural templates; ii) the small-molecule antagonist isoithiourea IT1t bound within a large peptide binding site of hCXCR4, which has more distant templates; and iii) the CVX15 peptide [RR-Nal-CTQKdPPTR-Cit-CRGdP, where Nal represents the non-natural amino acid L-3-(2-naphthyl)alanine,

Table 3: Recent applications of GPCR homology models in the context of structure-based drug design

Target	Template structure(s)[a]	Aim	Ref.
$RXFP_3$	CXCR4 [3OE0]	Recognize potential sites of interaction for binding of the native ligand human relaxin-3	250
CCR_4	CCR_5 [4MBS]	Identify the active site, and the residues involved in CCR_4–naphtelene-sulphonamide derivatives interaction using MD	251
GPR84	active β_2AR [3P0G]	Investigate GPR84-ligand molecular recognition, and comparison with other lipid receptors	252
$CXCR_7$	$CXCR_4$ [3ODU]	Assess the binding mode within $CXCR_7$ of the agonist cyclic peptide TC14012[b], and detect essential residues involved in the interaction with synthetic agonists using MD and SDM	253
H_4R	H_1R [3RZE]	Study the binding pathway of histamine from the extracellular side to the orthosteric binding site of the H_4R using unconstrained MD	254
A_3AR	A_2AR [2YDO]	Understand the positive allosterism mediated by imidazoquinoline toward A_3AR using supervised MD	255
FPR_2	μOR [4DKL], $CXCR_4$ [3ODU]	Study the binding mode of non-peptide and formyl peptide ligands	256
D_1R	β_2AR [3P0G]	Understand D_1R-agonist interaction using SDM and MD of D_1R-cathecol-amines	257
H_4R	β_2AR [2RH1]	Determine ligand binding modes to the H_4R binding site using 3D-QSAR, SDM and MD	258
β_2AR	β_2AR [2RH1] and others	Study differences at local and global conformational dynamic level of the N-terminal variants of the β_2AR using MD	259
D_4R	D_3R [3PBL], M_2R [3UON]	Analyze conformational dynamics induced upon ligand binding (dopamine and spiperone) using MD simulation in a lipid environment	260
FFA2	β_2AR [2RH1]	Determine residues involved in recognition and function of potent and selective orthosteric agonists within FFA2	261
CB1, CB2	$S1P_1$ [3V2Y], $A_{2A}AR$ [3QAK]	SAR rationalization of tricyclic ring systems binding to CB receptors	120
A_3AR	$A_{2A}AR$ [3QAK]	Structure-guided design of A_3AR selective nucleosides	262
$5\text{-}HT_7$	$5\text{-}HT_{1B}$ [4IAR], $5\text{-}HT_{2B}$ [4IB4], bRho [1F88]	Analyze the interactions involved in binding of long-chain arylpiperazine derivatives to $5\text{-}HT_7$ and $5\text{-}HT_{1A}$	263
Apelin (APJ)	CXCR4	Design of cyclic peptide analogues (biased agonists) for APJ	264
$5\text{-}HT_6$	β_2AR	Correlate binding pose with $5\text{-}HT_6$ the affinity of designed ligands	265
$5\text{-}HT_{2C}$	Inactive β_2AR [2RH1], active β_2AR [3SN6], $5\text{-}HT_{2B}$ [4IB4]	Probe the binding modes selectives phenylcyclopropylmethylamines $5\text{-}HT_{2C}$ agonists	266
A_1AR, A_3AR	Active $A_{2A}AR$ [3QAK], inactive $A_{2A}AR$ [3UZC]	Understand molecular bases of the A_1AR and A_3AR recognition and activation of 5'-C-ethyl-tetrazolyl derivatives	267
TGR5	S1PR1 [3V2Y], β_2AR [3SN6]	Investigate potential binding sites for naturally occurring bile acid derivatives TGR5 agonists	268
A_3AR	$A_{2A}AR$ [3QAK], β_2AR [3SN6]	Investigate molecular interaction between A_3AR and C2-arylethynyl nucleosides agonists	269
OX_1R, OX_2R	D_3R [3PBL]	Develop binding poses of orexin peptides in the hOX_1R and hOX_2R, with the aim of explaining SDM data and the molecular basis of agonist binding	270
CB1, CB2	$S1P_1$ [3V2Y]	SAR rationalization of biphenylic carboxamides within CB receptors binding sites	271
$A_{2B}AR$	$A_{2A}AR$ [3EML]	Assess structural similarities and differences in the molecular interactions and dynamics of $A_{2A}AR$ and $A_{2B}AR$ using MD	272
$\alpha_{2B}AR$, $\alpha_{2C}AR$, $h5HT_{2C}$, $h5HT_7$, β_3AR	Several templates	Structural probing of off-target GPCR activities within a series of adenosine/adenine congeners	21

[a] PDB code in brackets

[b] RR-Nal-CT-Cit-K-Dcit-PTR-Cit-CR-NH2

and Cit, citrulline] bound to the hCXCR4, which constitutes the first crystallized GPCR target complexed with a peptide-analogue. For each of the three targets, participant groups were allowed to submit up to 5 models. Thirty-five groups took part of the assessment. It was found that achieving accurate homology models requires at least a 35-40% target/template similarity coupled with the use of biochemical and QSAR data. This fact is useful to help prioritize the GPCRs to be crystallized in the future. As with the previous GPCR Dock assessment, modelling the EL_2 represented the biggest challenge, though in both the D_3R-eticlopride and IT1t-CXCR4 complexes, where the binding site is mainly defined by TM residues, ligand pose and contacts may be correctly predicted using a loop-less model. On the contrary, modelling the CXCR4-CVX15 system, where the peptide makes extensive contacts with highly flexible loops and the N-term, represented the most challenging case. The 2010

assessment confirmed that the use of biochemical, biophysical, QSAR and other experimental data is of the utmost importance in high quality homology modelling, even considering the limitation in the interpretation of SDM, where allosteric effects could be mistakenly taken as direct receptor-ligand interaction[35, 273].

In 2013, the last round of GPCR Dock[70] was conducted in coordination with the elucidation of crystal structures of $5HT_{1B}$[40] and $5HT_{2B}$[41], both in complex with the agonist ergotamine, and the TM domain of the human SMO receptor (class F GPCR) complexed with two different small-molecule antagonists, LY-2940680[60] and SANT-1[61]. Forty-four groups were involved in the evaluation. Modellers faced several challenges such as the prediction of activation states (agonism and biased agonism), the allosteric ligand interaction in $5\text{-}HT_{1B}$ and $5\text{-}HT_{2B}$, and homology modelling using remote templates for SMO (less than 15% sequence identity with any of the available template structures). In spite of the high sequence similarity to templates, the prediction of the serotonin-ergotamine complexes achieved a modest accuracy since ergotamine makes extensive and distinct interactions with the ELs. This relative success was in line with the moderate precision in EL predictions. Instead, more accurate predictions resulted for the ergoline core, which interacts mainly with TM regions. The best predictions for the serotonin receptors often used the MODELLER software[180] and multiple templates of aminergic structures, while many of the top-ranking complexes were refined by MD. Model selection by using subfamily-specific receptor-ligand interaction patterns, ligand SAR, and SMD, coupled with visual inspection proved to be a valid strategy. Furthermore, while several submitted models successfully detected the activation state of $5HT_{1B}$, this was not the case for the biased state of $5HT_{2B}$. This situation showed that there is still a need to further expand the crystallization of multiple functional states of GPCRs, and the improvement of computational methods for their prediction. The case of SMO, a class F GPCR with very low sequence similarity to existing structures, illustrated that target-template sequence alignment represents the main obstacle in distant homology modelling. In this sense, composite strategies including threading, fragment assembly, and energy-based refinement (e.g. I-TASSER[274]) showed its benefits for finding the correct residue matching. In addition, whereas alignment uncertainties may be addressed with modern methods, the structural precision of the remote homology models still require further developments.

5. The latest milestone: modelling GPCR classes B and C

Even though attempts were made to model classes B and C GPCRs, including GPRC6A[275], calcitonin gene-related peptide (CGRP) receptor[276], and metabotropic glutamate receptor 8 (mGluR8)[277], based on the crystal structures of class A GPCR, it has only recently been possible to model them using templates of the same family, that is to say, with classes B and C GPCR crystal structures. The construction of the homology models for non-class A GPCR have faced various challenges such as lack of structural data for the helical bundle and low TM sequence identity and 3D similarity for ELs and termini regions (class F GPCR dealt with the same issues[70]) with respect to class A templates[275-277].

However, in 2015, several structural models built on the basis of classes B and C X-ray structures were developed. Homology models of the corticotropin releasing factors receptor-2 (CRF_2R) were constructed using the crystal structure of CRF_1R as template, and both unbiased MD and well-tempered metadynamics simulations were conducted in order to probe the selectivity of an antagonist (CP-376395) towards CRF_2R and CRF_1R[278]. The Authors observed that a hydrogen bond between $His^{3.40}$ and $Tyr^{6.63}$ (using the Wootten et al. universal numbering scheme for class B GPCRs[279]) in CRF_1R, which is not present in CRF_2R, has a key role in explaining the difference of the antagonist selectivity towards both receptors.

With the aim of modelling the glucagon-like peptide-1 (GLP1) bound to the GLP1 receptor (GLP1R), homology models were built by utilizing the crystal structures of the CRF_1R, the glucagon receptor (GCGR), and the ligand-bound ECD of GLP1R and the gastric inhibitory polypeptide receptor (GIPR) as templates[280]. The Authors found that the residues Asp^9 and Gly^4 in GLP-1 interacted with the conserved residues in EL_3, while the binding site of GLP1R is constituted by conserved amino acids in the core domain.

Homology models of TM region of the metabotropic glutamate receptor 5 (mGluR5) were created based on the crystal structure of mGluR1, and refined using an MD-based methodology[281]. Guided by modelling insights, a novel benzoyl-2-benzimidazole scaffold was design and SAR studies were performed. A new positive allosteric modulator (PAM) for mGluR5 was discovered, which exhibited an IC_{50} value of 6.4 μM, *i.e.* about 20 fold more potent than DFB (a known mGluR5 PAM).

Homology modelling and MD simulations were undertaken in six mGluRs (mGluR2, mGluR3, mGluR4, mGluR6, mGluR7, and mGluR8) by using the crystal structure of mGluR5 as template, where the Authors reported predicted allosteric binding sites, and key residues for receptor selectivity[282]. Interesting, most of the findings in mGluR5, for example the "ionic lock" and some amino acid linked with receptor activation, were in accordance with the findings in class A GPCR.

6. Conclusions and Perspectives

With over 800 members in humans, receptors from the GPCR super-family are the target for ~30% of the marketed drugs. The first GPCR structure, bovine rhodopsin covalently bound to retinal was crystallized in 2000. However recent novel crystallization techniques allowed the solution of ~30 different druggable GPCR structures since 2007. This boosted the discovery of novel ligands and new chemical entities through structure-based virtual screening and lead optimization endeavours, using both crystal and modelled structures. This breakthrough also brought new templates for homology modelling, and a wealth of information regarding GPCR-ligand interaction patterns, clues about activation mechanisms, evidence for sequence-induced structural changes at the backbone level, and illustrated conformational loop diversity.

Still, the number of solved GPCR structures represents a very small part of the human GPCRs, and in spite of the tremendous effort and progress in crystallization, a complete coverage of the druggable GPCR structural space in the near- and mid-term does not seem likely. Thus, homology modelling appears as a reliable

and efficient tool to expand the GPCR structural map, and thus the horizons of hit identification and lead optimization in the coming years. Throughout this work, we have shown beyond any doubt from retrospective and prospective studies –including the three GPCR Dock community-wide assessments, that in spite of current limitations of, and standing challenges in homology modelling, *in silico* GPCR models have been invaluable for discovering and optimizing drug leads, characterizing GPCR-ligand interaction, rationalizing existing SAR data, aiding in the design of SDM experiments and SAR studies, and assessing off-target effects.

In the years ahead, the development of more accurate modelling techniques accounting for the wealth of biochemical, biophysical and QSAR data available, coupled with the validation of these methods in retrospective and prospective structure-based drug lead identification and optimization projects, should translate into a growing number of potent, selective and efficient GPCR ligands with high therapeutic value.

Acknowledgments

This work has been supported by the Agencia Nacional de Promoción Científica y Tecnológica, Argentina (PICT-2011-2778) and FOCEM-Mercosur (COF 03/11). The Authors thank Mario Rossi for insighful discussions, and Molsoft LLC for providing an academic license for the ICM program.

Notes and references

[a] *Instituto de Investigación en Biomedicina de Buenos Aires (IBioBA) - CONICET - Partner Institute of the Max Planck Society, Godoy Cruz 2390, C1425FQD, Buenos Aires, Argentina. Tel: +54 11 4899-5500; E-mail*: cnc@cavasotto-lab.net; ccavasotto@ibioba-mpsp-conicet.gov.ar

[*] *Correspondent author.*

1. K. L. Pierce, R. T. Premont and R. J. Lefkowitz, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 639-650.
2. K. Kristiansen, *Pharmacol. Ther.*, 2004, **103**, 21-80.
3. T. Schoneberg, A. Schulz, H. Biebermann, T. Hermsdorf, H. Rompler and K. Sangkuhl, *Pharmacol. Ther.*, 2004, **104**, 173-206.
4. R. Lappano and M. Maggiolini, *Nat. Rev. Drug Discov.*, 2011, **10**, 47-60.
5. J. A. Allen and B. L. Roth, *Annu. Rev. Pharmacol. Toxicol.*, 2011, **51**, 117-144.
6. J. A. Salon, D. T. Lodowski and K. Palczewski, *Pharmacol. Rev.*, 2011, **63**, 901-937.
7. R. C. Stevens, V. Cherezov, V. Katritch, R. Abagyan, P. Kuhn, H. Rosen and K. Wuthrich, *Nat. Rev. Drug Discov.*, 2013, **12**, 25-34.
8. R. Fredriksson, M. C. Lagerstrom, L. G. Lundin and H. B. Schioth, *Mol. Pharmacol.*, 2003, **63**, 1256-1272.
9. Y. Ono, W. Fujibuchi and M. Suwa, *Gene*, 2005, **364**, 63-73.
10. M. C. Lagerström and H. B. Schiöth, *Nat. Rev. Drug Discov.*, 2008, **7**, 339-357.
11. B. G. Tehan, A. Bortolato, F. E. Blaney, M. P. Weir and J. S. Mason, *Pharmacol. Ther.*, 2014, **143**, 51-60.
12. W. M. Oldham and H. E. Hamm, *Nat. Rev. Mol. Cell Biol.*, 2008, **9**, 60-71.
13. D. M. Rosenbaum, S. G. Rasmussen and B. K. Kobilka, *Nature*, 2009, **459**, 356-363.
14. M. Congreve, C. J. Langmead, J. S. Mason and F. H. Marshall, *J. Med. Chem.*, 2011, **54**, 4283-4311.
15. B. K. Shoichet and B. K. Kobilka, *Trends Pharmacol. Sci.*, 2012, **33**, 268-272.
16. M. Congreve, J. M. Dias and F. H. Marshall, *Prog. Med. Chem.*, 2014, **53**, 1-63.
17. S. Costanzi and K. Wang, *Adv. Exp. Med. Biol.*, 2014, **796**, 3-13.
18. F. Fanelli and P. G. De Benedetti, *Chem. Rev.*, 2011, **111**, PR438-535.
19. M. Congreve, C. Langmead and F. H. Marshall, *Adv. Pharmacol.*, 2011, **62**, 1-36.
20. K. A. Jacobson, S. Costanzi and S. Paoletta, *Trends Pharmacol. Sci.*, 2014, **35**, 658-663.
21. S. Paoletta, D. K. Tosh, D. Salvemini and K. A. Jacobson, *PLoS One*, 2014, **9**, e97858.
22. P. A. Hargrave, J. H. McDowell, D. R. Curtis, J. K. Wang, E. Juszczak, S. L. Fong, J. K. Rao and P. Argos, *Biophys. Struct. Mech.*, 1983, **9**, 235-244.
23. G. F. Schertler, C. Villa and R. Henderson, *Nature*, 1993, **362**, 770-772.
24. J. M. Baldwin, *EMBO J.*, 1993, **12**, 1693-1703.
25. K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto and M. Miyano, *Science*, 2000, **289**, 739-745.
26. V. Cherezov, D. M. Rosenbaum, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka and R. C. Stevens, *Science*, 2007, **318**, 1258-1265.
27. P. W. Day, S. G. Rasmussen, C. Parnot, J. J. Fung, A. Masood, T. S. Kobilka, X. J. Yao, H. J. Choi, W. I. Weis, D. K. Rohrer and B. K. Kobilka, *Nat. Methods*, 2007, **4**, 927-929.
28. S. G. Rasmussen, H. J. Choi, D. M. Rosenbaum, T. S. Kobilka, F. S. Thian, P. C. Edwards, M. Burghammer, V. R. Ratnala, R. Sanishvili, R. F. Fischetti, G. F. Schertler, W. I. Weis and B. K. Kobilka, *Nature*, 2007, **450**, 383-387.
29. D. M. Rosenbaum, V. Cherezov, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, X. J. Yao, W. I. Weis, R. C. Stevens and B. K. Kobilka, *Science*, 2007, **318**, 1266-1273.
30. E. Chun, A. A. Thompson, W. Liu, C. B. Roth, M. T. Griffith, V. Katritch, J. Kunken, F. Xu, V. Cherezov, M. A. Hanson and R. C. Stevens, *Structure*, 2012, **20**, 967-976.
31. S. G. Rasmussen, H. J. Choi, J. J. Fung, E. Pardon, P. Casarosa, P. S. Chae, B. T. Devree, D. M. Rosenbaum, F. S. Thian, T. S. Kobilka, A. Schnapp, I. Konetzki, R. K. Sunahara, S. H. Gellman, A. Pautsch, J. Steyaert, W. I. Weis and B. K. Kobilka, *Nature*, 2011, **469**, 175-180.
32. N. Robertson, A. Jazayeri, J. Errey, A. Baig, E. Hurrell, A. Zhukov, C. J. Langmead, M. Weir and F. H. Marshall, *Neuropharmacology*, 2011, **60**, 36-44.
33. S. G. Rasmussen, B. T. DeVree, Y. Zou, A. C. Kruse, K. Y. Chung, T. S. Kobilka, F. S. Thian, P. S. Chae, E. Pardon, D. Calinski, J. M. Mathiesen, S. T. Shah, J. A. Lyons, M. Caffrey, S. H. Gellman, J. Steyaert, G. Skiniotis, W. I. Weis, R. K. Sunahara and B. K. Kobilka, *Nature*, 2011, **477**, 549-555.
34. T. Warne, M. J. Serrano-Vega, J. G. Baker, R. Moukhametzianov, P. C. Edwards, R. Henderson, A. G. Leslie, C. G. Tate and G. F. Schertler, *Nature*, 2008, **454**, 486-491.
35. V. P. Jaakola, M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. Chien, J. R. Lane, A. P. Ijzerman and R. C. Stevens, *Science*, 2008, **322**, 1211-1217.
36. T. Shimamura, M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G. W. Han, T. Kobayashi, R. C. Stevens and S. Iwata, *Nature*, 2011, **475**, 65-70.
37. E. Y. Chien, W. Liu, Q. Zhao, V. Katritch, G. W. Han, M. A. Hanson, L. Shi, A. H. Newman, J. A. Javitch, V. Cherezov and R. C. Stevens, *Science*, 2011, **330**, 1091-1095.
38. K. Haga, A. C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W. I. Weis, T. Okada, B. K. Kobilka, T. Haga and T. Kobayashi, *Nature*, 2012, **482**, 547-551.
39. A. C. Kruse, J. Hu, A. C. Pan, D. H. Arlow, D. M. Rosenbaum, E. Rosemond, H. F. Green, T. Liu, P. S. Chae, R. O. Dror, D. E. Shaw, W. I. Weis, J. Wess and B. K. Kobilka, *Nature*, 2012, **482**, 552-556.
40. C. Wang, Y. Jiang, J. Ma, H. Wu, D. Wacker, V. Katritch, G. W. Han, W. Liu, X. P. Huang, E. Vardy, J. D. McCorvy, X. Gao, X. E. Zhou, K. Melcher, C. Zhang, F. Bai, H. Yang, L. Yang, H. Jiang, B. L. Roth, V. Cherezov, R. C. Stevens and H. E. Xu, *Science*, 2013, **340**, 610-614.
41. D. Wacker, C. Wang, V. Katritch, G. W. Han, X. P. Huang, E. Vardy, J. D. McCorvy, Y. Jiang, M. Chu, F. Y. Siu, W. Liu, H. E.

Xu, V. Cherezov, B. L. Roth and R. C. Stevens, *Science*, 2013, **340**, 615-619.
42. M. A. Hanson, C. B. Roth, E. Jo, M. T. Griffith, F. L. Scott, G. Reinhart, H. Desale, B. Clemons, S. M. Cahalan, S. C. Schuerer, M. G. Sanna, G. W. Han, P. Kuhn, H. Rosen and R. C. Stevens, *Science*, 2012, **335**, 851-855.
43. B. Wu, E. Y. Chien, C. D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F. C. Bi, D. J. Hamel, P. Kuhn, T. M. Handel, V. Cherezov and R. C. Stevens, *Science*, 2010, **330**, 1066-1071.
44. Q. Tan, Y. Zhu, J. Li, Z. Chen, G. W. Han, I. Kufareva, T. Li, L. Ma, G. Fenalti, W. Zhang, X. Xie, H. Yang, H. Jiang, V. Cherezov, H. Liu, R. C. Stevens, Q. Zhao and B. Wu, *Science*, 2013, **341**, 1387-1390.
45. S. Granier, A. Manglik, A. C. Kruse, T. S. Kobilka, F. S. Thian, W. I. Weis and B. K. Kobilka, *Nature*, 2012, **485**, 400-404.
46. A. Manglik, A. C. Kruse, T. S. Kobilka, F. S. Thian, J. M. Mathiesen, R. K. Sunahara, L. Pardo, W. I. Weis, B. K. Kobilka and S. Granier, *Nature*, 2012, **485**, 321-326.
47. H. Wu, D. Wacker, M. Mileni, V. Katritch, G. W. Han, E. Vardy, W. Liu, A. A. Thompson, X. P. Huang, F. I. Carroll, S. W. Mascarella, R. B. Westkaemper, P. D. Mosier, B. L. Roth, V. Cherezov and R. C. Stevens, *Nature*, 2012, **485**, 327-332.
48. A. A. Thompson, W. Liu, E. Chun, V. Katritch, H. Wu, E. Vardy, X. P. Huang, C. Trapella, R. Guerrini, G. Calo, B. L. Roth, V. Cherezov and R. C. Stevens, *Nature*, 2012, **485**, 395-399.
49. J. F. White, N. Noinaj, Y. Shibata, J. Love, B. Kloss, F. Xu, J. Gvozdenovic-Jeremic, P. Shah, J. Shiloach, C. G. Tate and R. Grisshammer, *Nature*, 2012, **490**, 508-513.
50. C. Zhang, Y. Srinivasan, D. H. Arlow, J. J. Fung, D. Palmer, Y. Zheng, H. F. Green, A. Pandey, R. O. Dror, D. E. Shaw, W. I. Weis, S. R. Coughlin and B. K. Kobilka, *Nature*, 2012, **492**, 387-392.
51. K. Zhang, J. Zhang, Z. G. Gao, D. Zhang, L. Zhu, G. W. Han, S. M. Moss, S. Paoletta, E. Kiselev, W. Lu, G. Fenalti, W. Zhang, C. E. Muller, H. Yang, H. Jiang, V. Cherezov, V. Katritch, K. A. Jacobson, R. C. Stevens, B. Wu and Q. Zhao, *Nature*, 2014, **509**, 115-118.
52. D. Zhang, Z. G. Gao, K. Zhang, E. Kiselev, S. Crane, J. Wang, S. Paoletta, C. Yi, L. Ma, W. Zhang, G. W. Han, H. Liu, V. Cherezov, V. Katritch, H. Jiang, R. C. Stevens, K. A. Jacobson, Q. Zhao and B. Wu, *Nature*, 2015, **520**, 317-321.
53. A. Srivastava, J. Yano, Y. Hirozane, G. Kefala, F. Gruswitz, G. Snell, W. Lane, A. Ivetac, K. Aertgeerts, J. Nguyen, A. Jennings and K. Okada, *Nature*, 2014, **513**, 124-127.
54. J. Yin, J. C. Mobarec, P. Kolb and D. M. Rosenbaum, *Nature*, 2015, **519**, 247-250.
55. H. Zhang, H. Unal, C. Gati, G. W. Han, W. Liu, N. A. Zatsepin, D. James, D. Wang, G. Nelson, U. Weierstall, M. R. Sawaya, Q. Xu, M. Messerschmidt, G. J. Williams, S. Boutet, O. M. Yefanov, T. A. White, C. Wang, A. Ishchenko, K. C. Tirupula, R. Desnoyer, J. Coe, C. E. Conrad, P. Fromme, R. C. Stevens, V. Katritch, S. S. Karnik and V. Cherezov, *Cell*, 2015.
56. K. Hollenstein, J. Kean, A. Bortolato, R. K. Cheng, A. S. Dore, A. Jazayeri, R. M. Cooke, M. Weir and F. H. Marshall, *Nature*, 2013, **499**, 438-443.
57. F. Y. Siu, M. He, C. de Graaf, G. W. Han, D. Yang, Z. Zhang, C. Zhou, Q. Xu, D. Wacker, J. S. Joseph, W. Liu, J. Lau, V. Cherezov, V. Katritch, M. W. Wang and R. C. Stevens, *Nature*, 2013, **499**, 444-449.
58. H. Wu, C. Wang, K. J. Gregory, G. W. Han, H. P. Cho, Y. Xia, C. M. Niswender, V. Katritch, J. Meiler, V. Cherezov, P. J. Conn and R. C. Stevens, *Science*, 2014, **344**, 58-64.
59. A. S. Dore, K. Okrasa, J. C. Patel, M. Serrano-Vega, K. Bennett, R. M. Cooke, J. C. Errey, A. Jazayeri, S. Khan, B. Tehan, M. Weir, G. R. Wiggin and F. H. Marshall, *Nature*, 2014, **511**, 557-562.
60. C. Wang, H. Wu, V. Katritch, G. W. Han, X. P. Huang, W. Liu, F. Y. Siu, B. L. Roth, V. Cherezov and R. C. Stevens, *Nature*, 2013, **497**, 338-343.
61. C. Wang, H. Wu, T. Evron, E. Vardy, G. W. Han, X. P. Huang, S. J. Hufeisen, T. J. Mangano, D. J. Urban, V. Katritch, V. Cherezov, M. G. Caron, B. L. Roth and R. C. Stevens, *Nat. Commun.*, 2014, **5**, 4355.
62. T. Warne and C. G. Tate, *Biochem. Soc. Trans.*, 2013, **41**, 159-165.
63. X. Zhang, R. C. Stevens and F. Xu, *Trends Biochem. Sci.*, 2015, **40**, 79-87.
64. A. C. Kruse, A. M. Ring, A. Manglik, J. Hu, K. Hu, K. Eitel, H. Hubner, E. Pardon, C. Valant, P. M. Sexton, A. Christopoulos, C. C. Felder, P. Gmeiner, J. Steyaert, W. I. Weis, K. C. Garcia, J. Wess and B. K. Kobilka, *Nature*, 2013, **504**, 101-106.
65. A. J. Kooistra, L. Roumen, R. Leurs, I. J. de Esch and C. de Graaf, *Methods Enzymol.*, 2013, **522**, 279-336.
66. J. S. Mason, A. Bortolato, M. Congreve and F. H. Marshall, *Trends Pharmacol. Sci.*, 2012, **33**, 249-260.
67. V. Katritch and R. Abagyan, *Trends Pharmacol. Sci.*, 2011, **32**, 637-643.
68. M. Audet and M. Bouvier, *Cell*, 2012, **151**, 14-23.
69. H. Tang, X. S. Wang, J. H. Hsieh and A. Tropsha, *Proteins*, 2012, **80**, 1503-1521.
70. I. Kufareva, V. Katritch, R. C. Stevens and R. Abagyan, *Structure*, 2014, **22**, 1120-1139.
71. I. Kufareva, M. Rueda, V. Katritch, R. C. Stevens and R. Abagyan, *Structure*, 2011, **19**, 1108-1126.
72. M. Michino, E. Abola, C. L. Brooks, 3rd, J. S. Dixon, J. Moult and R. C. Stevens, *Nat. Rev. Drug Discov.*, 2009, **8**, 455-463.
73. S. Shacham, M. Topf, N. Avisar, F. Glaser, Y. Marantz, S. Bar-Haim, S. Noiman, Z. Naor and O. M. Becker, *Med. Res. Rev.*, 2001, **21**, 472-483.
74. N. Vaidehi, W. B. Floriano, R. Trabanino, S. E. Hall, P. Freddolino, E. J. Choi, G. Zamanakos and W. A. Goddard, 3rd, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12622-12627.
75. O. M. Becker, Y. Marantz, S. Shacham, B. Inbal, A. Heifetz, O. Kalid, S. Bar-Haim, D. Warshaviak, M. Fichman and S. Noiman, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 11304-11309.
76. S. Bhattacharya, S. E. Hall, H. Li and N. Vaidehi, *Biophys. J.*, 2008, **94**, 2027-2042.
77. P. L. Freddolino, M. Y. Kalani, N. Vaidehi, W. B. Floriano, S. E. Hall, R. J. Trabanino, V. W. Kam and W. A. Goddard, 3rd, *Proc. Natl. Acad. Sci. U S A*, 2004, **101**, 2736-2741.
78. A. Kirkpatrick, J. Heo, R. Abrol and W. A. Goddard, 3rd, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 19988-19993.
79. S. Shacham, Y. Marantz, S. Bar-Haim, O. Kalid, D. Warshaviak, N. Avisar, B. Inbal, A. Heifetz, M. Fichman, M. Topf, Z. Naor, S. Noiman and O. M. Becker, *Proteins*, 2004, **57**, 51-86.
80. J. Zhang, J. Yang, R. Jang and Y. Zhang, *Structure (Camb.)*, 2015, **In press**.
81. F. Spyrakis and C. N. Cavasotto, *Arch. Biochem. Biophys.*, 2015, submitted.
82. A. Fiser, *Expert Rev. Proteomics*, 2004, **1**, 97-110.
83. C. Chothia and A. M. Lesk, *EMBO J.*, 1986, **5**, 823-826.
84. T. Mirzadegan, G. Benko, S. Filipek and K. Palczewski, *Biochemistry*, 2003, **42**, 2759-2767.
85. T. Kenakin, *Annu. Rev. Pharmacol. Toxicol.*, 2002, **42**, 349-379.
86. B. K. Kobilka and X. Deupi, *Trends Pharmacol. Sci.*, 2007, **28**, 397-406.
87. X. J. Yao, G. Velez Ruiz, M. R. Whorton, S. G. Rasmussen, B. T. DeVree, X. Deupi, R. K. Sunahara and B. Kobilka, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 9501-9506.
88. C. N. Cavasotto, *Curr. Top. Med. Chem.*, 2011, **11**, 1528-1534.
89. D. Palomba and C. N. Cavasotto, in *In Silico Drug Discovery and Design: Theory, Methods, Challenges, and Applications*, ed. C. N. Cavasotto, CRC Press, Taylor & Francis Group, 2015, p. in press.
90. G. H. Lushington, *Methods Mol. Biol.*, 2015, **1215**, 309-330.
91. A. Fiser, *Methods Mol. Biol.*, 2010, **673**, 73-94.
92. S. Costanzi, *Methods Mol. Biol.*, 2012, **857**, 259-279.
93. J. Ballesteros and H. Weinstein, *Methods Neurosci.*, 1995, **25**, 366-428.
94. M. Zhu and M. Li, *Mol. BioSyst.*, 2012, **8**, 1686-1693.
95. K. Bondensgaard, M. Ankersen, H. Thogersen, B. S. Hansen, B. S. Wulff and R. P. Bywater, *J. Med. Chem.*, 2004, **47**, 888-899.
96. C. L. Worth, G. Kleinau and G. Krause, *PLoS One*, 2009, **4**, e7011.
97. J. C. Mobarec, R. Sanchez and M. Filizola, *J. Med. Chem.*, 2009, **52**, 5207-5216.

98. F. Deflorian and K. A. Jacobson, *J. Comput-Aided. Mol. Des.*, 2011, **25**, 329-338.
99. J. Carlsson, R. G. Coleman, V. Setola, J. J. Irwin, H. Fan, A. Schlessinger, A. Sali, B. L. Roth and B. K. Shoichet, *Nat. Chem. Biol.*, 2011, **7**, 769-778.
100. M. M. Mysinger, D. R. Weiss, J. J. Ziarek, S. Gravel, A. K. Doak, J. Karpiak, N. Heveker, B. K. Shoichet and B. F. Volkman, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 5517-5522.
101. S. S. Phatak, E. A. Gatica and C. N. Cavasotto, *J. Chem. Inf. Model.*, 2010, **50**, 2119-2128.
102. S. Vilar, J. Karpiak and S. Costanzi, *J. Comput. Chem.*, 2010, **31**, 707-720.
103. S. Vilar, G. Ferino, S. S. Phatak, B. Berk, C. N. Cavasotto and S. Costanzi, *J. Mol. Graph. Model.*, 2011, **29**, 614-623.
104. S. Vilar, J. Karpiak, B. Berk and S. Costanzi, *J. Mol. Graph. Model.*, 2011, **29**, 809-817.
105. S. Vilar and S. Costanzi, *Methods Mol. Biol.*, 2012, **914**, 271-284.
106. S. Costanzi and S. Vilar, *J. Comput. Chem.*, 2012, **33**, 561-572.
107. K. Rataj, J. Witek, S. Mordalski, T. Kosciolek and A. J. Bojarski, *J. Chem. Inf. Model.*, 2014, **54**, 1661-1668.
108. M. Kolaczkowski, A. Bucki, M. Feder and M. Pawlowski, *J. Chem. Inf. Model.*, 2013, **53**, 638-648.
109. S. Bhattacharya, A. R. Lam, H. Li, G. Balaraman, M. J. Niesen and N. Vaidehi, *Proteins*, 2013, **81**, 729-739.
110. A. Gonzalez, A. Cordomi, G. Caltabiano and L. Pardo, *ChemBioChem*, 2012, **13**, 1393-1399.
111. C. N. Cavasotto and S. S. Phatak, *Drug Discov. Today*, 2009, **14**, 676-683.
112. B. Vroling, M. Sanders, C. Baakman, A. Borrmann, S. Verhoeven, J. Klomp, L. Oliveira, J. de Vlieg and G. Vriend, *Nucleic Acids Res.*, 2011, **39**, D309-319.
113. C. de Graaf and D. Rognan, *Curr. Pharm. Des.*, 2009, **15**, 4026-4048.
114. A. Evers and G. Klebe, *Angew. Chem. Int. Ed. Engl.*, 2004, **43**, 248-251.
115. A. Evers and G. Klebe, *J. Med. Chem.*, 2004, **47**, 5381-5392.
116. A. J. W. Orry and C. N. Cavasotto, in *231st Meeting of the American Chemical Society*, Atlanta, GA, 2006.
117. C. N. Cavasotto, A. J. Orry, N. J. Murgolo, M. F. Czarniecki, S. A. Kocsi, B. E. Hawes, K. A. O'Neill, H. Hine, M. S. Burton, J. H. Voigt, R. A. Abagyan, M. L. Bayne and F. J. Monsma, Jr., *J. Med. Chem.*, 2008, **51**, 581-588.
118. P. Diaz, S. S. Phatak, J. Xu, F. Astruc-Diaz, C. N. Cavasotto and M. Naguib, *J. Med. Chem.*, 2009, **52**, 433-444.
119. P. Diaz, S. S. Phatak, J. Xu, F. R. Fronczek, F. Astruc-Diaz, C. M. Thompson, C. N. Cavasotto and M. Naguib, *ChemMedChem*, 2009, **4**, 1615-1629.
120. R. R. Petrov, L. Knight, S. R. Chen, J. Wager-Miller, S. W. McDaniel, F. Diaz, F. Barth, H. L. Pan, K. Mackie, C. N. Cavasotto and P. Diaz, *Eur. J. Med. Chem.*, 2013, **69**, 881-907.
121. C. N. Cavasotto and R. A. Abagyan, *J. Mol. Biol.*, 2004, **337**, 209-225.
122. C. N. Cavasotto, J. A. Kovacs and R. A. Abagyan, *J. Am. Chem. Soc.*, 2005, **127**, 9632-9640.
123. J. A. Kovacs, C. N. Cavasotto and R. A. Abagyan, *J. Comp. Theor. Nanosci.*, 2005, **2**, 354-361.
124. M. C. Monti, A. Casapullo, C. N. Cavasotto, A. Napolitano and R. Riccio, *ChemBioChem*, 2007, **8**, 1585-1591.
125. M. Rossi, B. Rotblat, K. Ansell, I. Amelio, M. Caraglia, G. Misso, F. Bernassola, C. N. Cavasotto, R. A. Knight, A. Ciechanover and G. Melino, *Cell Death Dis.*, 2014, **5**, e1203.
126. S. Moro, F. Deflorian, M. Bacilieri and G. Spalluto, *Curr. Pharm. Des.*, 2006, **12**, 2175-2185.
127. W. Sherman, T. Day, M. P. Jacobson, R. A. Friesner and R. Farid, *J. Med. Chem.*, 2006, **49**, 534-553.
128. S. Costanzi, *J. Med. Chem.*, 2008, **51**, 2907-2914.
129. F. M. McRobb, B. Capuano, I. T. Crosby, D. K. Chalmers and E. Yuriev, *J. Chem. Inf. Model.*, 2010, **50**, 626-637.
130. T. Thomas, K. C. McLean, F. M. McRobb, D. T. Manallack, D. K. Chalmers and E. Yuriev, *J. Chem. Inf. Model.*, 2014, **54**, 243-253.
131. D. Pala, T. Beuming, W. Sherman, A. Lodola, S. Rivara and M. Mor, *J. Chem. Inf. Model.*, 2013, **53**, 821-835.
132. S. P. Chin, M. J. Buckle, D. K. Chalmers, E. Yuriev and S. W. Doughty, *J. Mol. Graph. Model.*, 2014, **49**, 91-98.
133. V. Katritch, M. Rueda, P. C. Lam, M. Yeager and R. Abagyan, *Proteins*, 2010, **78**, 197-211.
134. ICM, MolSoft, LLC, La Jolla, CA, 2012.
135. S. Wolf, M. Bockmann, U. Howeler, J. Schlitter and K. Gerwert, *FEBS Lett.*, 2008, **582**, 3335-3342.
136. S. R. Kimura, A. J. Tebben and D. R. Langley, *Proteins*, 2008, **71**, 1919-1929.
137. J. Varady, X. Wu, X. Fang, J. Min, Z. Hu, B. Levant and S. Wang, *J. Med. Chem.*, 2003, **46**, 4377-4392.
138. J. Mortier, C. Rakers, M. Bermudez, M. S. Murgueitio, S. Riniker and G. Wolber, *Drug Discov. Today*, 2015.
139. Z. Deng, C. Chuaqui and J. Singh, *J. Med. Chem.*, 2004, **47**, 337-344.
140. M. Schneider, S. Wolf, J. Schlitter and K. Gerwert, *FEBS Lett.*, 2011, **585**, 3587-3592.
141. A. Tarcsay, G. Paragi, M. Vass, B. Jojart, F. Bogar and G. M. Keserü, *J. Chem. Inf. Model.*, 2013, **53**, 2990-2999.
142. V. Katritch, V. Cherezov and R. C. Stevens, *Trends Pharmacol. Sci.*, 2012, **33**, 17-27.
143. M. C. Peeters, G. J. van Westen, Q. Li and I. J. AP, *Trends Pharmacol. Sci.*, 2011, **32**, 35-42.
144. M. A. Hanson and R. C. Stevens, *Structure*, 2009, **17**, 8-14.
145. C. N. Cavasotto, A. J. Orry and R. A. Abagyan, *Proteins*, 2003, **51**, 423-433.
146. G. V. Nikiforovich, C. M. Taylor, G. R. Marshall and T. J. Baranski, *Proteins*, 2010, **78**, 271-285.
147. C. de Graaf, N. Foata, O. Engkvist and D. Rognan, *Proteins*, 2008, **71**, 599-620.
148. D. A. Goldfeld, K. Zhu, T. Beuming and R. A. Friesner, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 8275-8280.
149. D. A. Goldfeld, K. Zhu, T. Beuming and R. A. Friesner, *Proteins*, 2013, **81**, 214-228.
150. S. Kmiecik, M. Jamroz and M. Kolinski, *Biophys. J.*, 2014, **106**, 2408-2416.
151. M. Jamroz, M. Orozco, A. Kolinski and S. Kmiecik, *J. Chem. Theory Comput.*, 2013, **9**, 119-125.
152. A. Kolinski, *Acta Biochim. Pol.*, 2004, **51**, 349-371.
153. M. R. Ali, R. Latif, T. F. Davies and M. Mezei, *J. Biomol. Struct. Dyn.*, 2015, **33**, 1140-1152.
154. A. Fiser and A. Sali, *Bioinformatics*, 2003, **19**, 2500-2501.
155. C. A. Rohl, C. E. Strauss, D. Chivian and D. Baker, *Proteins*, 2004, **55**, 656-677.
156. P. W. Hildebrand, A. Goede, R. A. Bauer, B. Gruening, J. Ismer, E. Michalsky and R. Preissner, *Nucleic Acids Res.*, 2009, **37**, W571-574.
157. T. Schmidt, A. Bergner and T. Schwede, *Drug Discov. Today*, 2014, **19**, 890-897.
158. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Cryst.*, 1993, 283-291.
159. R. W. Hooft, G. Vriend, C. Sander and E. E. Abola, *Nature*, 1996, **381**, 272.
160. I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, 3rd, J. Snoeyink, J. S. Richardson and D. C. Richardson, *Nucleic Acids Res.*, 2007, **35**, W375-383.
161. E. A. Gatica and C. N. Cavasotto, *J. Chem. Inf. Model.*, 2012, **52**, 1-6.
162. V. Katritch, V. P. Jaakola, J. R. Lane, J. Lin, A. P. Ijzerman, M. Yeager, I. Kufareva, R. C. Stevens and R. Abagyan, *J. Med. Chem.*, 2010, **53**, 1799-1809.
163. S. Engel, A. P. Skoumbourdis, J. Childress, S. Neumann, J. R. Deschamps, C. J. Thomas, A. O. Colson, S. Costanzi and M. C. Gershengorn, *J. Am. Chem. Soc.*, 2008, **130**, 5115-5123.
164. I. G. Tikhonova, C. S. Sum, S. Neumann, S. Engel, B. M. Raaka, S. Costanzi and M. C. Gershengorn, *J. Med. Chem.*, 2008, **51**, 625-633.
165. D. Wacker, G. Fenalti, M. A. Brown, V. Katritch, R. Abagyan, V. Cherezov and R. C. Stevens, *J. Am. Chem. Soc.*, 2010, **132**, 11443-11445.
166. C. N. Cavasotto and S. S. Phatak, *Methods Mol. Biol.*, 2011, **685**, 155-174.

167. G. L. Warren, C. W. Andrews, A. M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. F. Semus, S. Senger, G. Tedesco, I. D. Wall, J. M. Woolven, C. E. Peishoff and M. S. Head, *J. Med. Chem.*, 2006, **49**, 5912-5931.
168. C. N. Cavasotto and N. Singh, *Curr. Comput.-Aided Drug Design*, 2008, **4**, 221-234.
169. C. N. Cavasotto and A. J. Orry, *Curr. Top. Med. Chem.*, 2007, **7**, 1006-1014.
170. J. Moult, J. T. Pedersen, R. Judson and K. Fidelis, *Proteins*, 1995, **23**, ii-v.
171. D. Latek, P. Pasznik, T. Carlomagno and S. Filipek, *PLoS One*, 2013, **8**, e56742.
172. C. L. Worth, A. Kreuchwig, G. Kleinau and G. Krause, *BMC Bioinformatics*, 2011, **12**, 185.
173. M. Sandal, T. P. Duy, M. Cona, H. Zung, P. Carloni, F. Musiani and A. Giorgetti, *PLoS One*, 2013, **8**, e74092.
174. D. Rodríguez, X. Bello and H. Gutiérrez-de-Terán, *Molecular Informatics*, 2012, **31**, 334-341.
175. H. Gutierrez-de-Teran, X. Bello and D. Rodriguez, *Biochem. Soc. Trans.*, 2013, **41**, 205-212.
176. G. Launay, S. Teletchea, F. Wade, E. Pajot-Augy, J. F. Gibrat and G. Sanz, *Protein Eng. Des. Sel.*, 2012, **25**, 377-386.
177. M. Y. Shen and A. Sali, *Protein Sci.*, 2006, **15**, 2507-2524.
178. C. A. Rohl, C. E. Strauss, K. M. Misura and D. Baker, *Methods Enzymol.*, 2004, **383**, 66-93.
179. J. Soding, *Bioinformatics*, 2005, **21**, 951-960.
180. N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper and A. Sali, *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]*, 2006, **Chapter 5**, Unit 5 6.
181. L. Willard, A. Ranjan, H. Zhang, H. Monzavi, R. F. Boyko, B. D. Sykes and D. S. Wishart, *Nucleic Acids Res.*, 2003, **31**, 3316-3319.
182. O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455-461.
183. C. Dominguez, R. Boelens and A. M. Bonvin, *J. Am. Chem. Soc.*, 2003, **125**, 1731-1737.
184. V. Le Guilloux, P. Schmidtke and P. Tuffery, *BMC Bioinformatics*, 2009, **10**, 168.
185. L. Martinez, R. Andreani and J. M. Martinez, *BMC Bioinformatics*, 2007, **8**, 306.
186. M. Pawlowski, S. Saraswathi, H. K. Motawea, M. A. Chotani and A. Kloczkowski, *PLoS One*, 2014, **9**, e103099.
187. A. Rodriguez, A. Guerrero, H. Gutierrez-de-Teran, D. Rodriguez, J. Brea, M. I. Loza, G. Rosell and M. Pilar Bosch, *MedChemComm*, 2015, **6**, 1178-1185.
188. D. Rodriguez, A. Ranganathan and J. Carlsson, *J. Chem. Inf. Model.*, 2014, **54**, 2004-2021.
189. P. Kolb, D. M. Rosenbaum, J. J. Irwin, J. J. Fung, B. K. Kobilka and B. K. Shoichet, *Proc. Natl. Acad. Sci. U S A*, 2009, **106**, 6843-6848.
190. S. Topiol and M. Sabio, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 1598-1602.
191. M. Sabio, K. Jones and S. Topiol, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 5391-5395.
192. J. Carlsson, L. Yoo, Z. G. Gao, J. J. Irwin, B. K. Shoichet and K. A. Jacobson, *J. Med. Chem.*, 2010, **53**, 3748-3755.
193. C. de Graaf, A. J. Kooistra, H. F. Vischer, V. Katritch, M. Kuijer, M. Shiroishi, S. Iwata, T. Shimamura, R. C. Stevens, I. J. de Esch and R. Leurs, *J. Med. Chem.*, 2011, **54**, 8195-8206.
194. K. A. Jacobson and S. Costanzi, *Mol. Pharmacol.*, 2012, **82**, 361-371.
195. S. P. Andrews, G. A. Brown and J. A. Christopher, *ChemMedChem*, 2014, **9**, 256-275.
196. A. Evers, G. Hessler, H. Matter and T. Klabunde, *J. Med. Chem.*, 2005, **48**, 5448-5465.
197. R. Kiss, B. Kiss, A. Konczol, F. Szalai, I. Jelinek, V. Laszlo, B. Noszal, A. Falus and G. M. Keserü, *J. Med. Chem.*, 2008, **51**, 3145-3153.
198. J. Bayry, E. Z. Tchilian, M. N. Davies, E. K. Forbes, S. J. Draper, S. V. Kaveri, A. V. Hill, M. D. Kazatchkine, P. C. Beverley, D. R. Flower and D. F. Tough, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 10221-10226.
199. E. Kellenberger, J. Y. Springael, M. Parmentier, M. Hachet-Haas, J. L. Galzi and D. Rognan, *J. Med. Chem.*, 2007, **50**, 1294-1303.
200. B. S. Edwards, C. Bologa, S. M. Young, K. V. Balakin, E. R. Prossnitz, N. P. Savchuck, L. A. Sklar and T. I. Oprea, *Mol. Pharmacol.*, 2005, **68**, 1301-1310.
201. O. M. Salo, K. H. Raitio, J. R. Savinainen, T. Nevalainen, M. Lahtela-Kakkonen, J. T. Laitinen, T. Jarvinen and A. Poso, *J. Med. Chem.*, 2005, **48**, 7166-7171.
202. C. S. Sum, I. G. Tikhonova, S. Neumann, S. Engel, B. M. Raaka, S. Costanzi and M. C. Gershengorn, *J. Biol. Chem.*, 2007, **282**, 29248-29255.
203. M. Bhasin and G. P. Raghava, *Nucleic Acids Res.*, 2004, **32**, W383-389.
204. M. Wistrand, L. Kall and E. L. Sonnhammer, *Protein Sci.*, 2006, **15**, 509-521.
205. W. K. Chan, H. Zhang, J. Yang, J. R. Brender, J. Hur, A. Ozgur and Y. Zhang, *Bioinformatics*, 2015.
206. M. Floris, D. Sabbadin, A. Ciancetta, R. Medda, A. Cuzzolin and S. Moro, *In silico pharmacology*, 2013, **1**, 25.
207. M. Floris, D. Sabbadin, R. Medda, A. Bulfone and S. Moro, *Eur. J. Med. Chem.*, 2012, **58**, 248-257.
208. R. Yan, X. Wang, L. Huang, J. Lin, W. Cai and Z. Zhang, *Mol. BioSyst.*, 2014, **10**, 2495-2504.
209. H. Zhou and J. Skolnick, *Mol. Pharm.*, 2012, **9**, 1775-1784.
210. V. Isberg, B. Vroling, R. van der Kant, K. Li, G. Vriend and D. Gloriam, *Nucleic Acids Res.*, 2014, **42**, D422-425.
211. D. Kozma, I. Simon and G. E. Tusnady, *Nucleic Acids Res.*, 2013, **41**, D524-529.
212. M. W. Beukers, I. Kristiansen, I. J. AP and I. Edvardsen, *Trends Pharmacol. Sci.*, 1999, **20**, 475-477.
213. S. Jayasinghe, K. Hristova and S. H. White, *Protein Sci.*, 2001, **10**, 455-458.
214. L. Skrabanek, M. Murcia, M. Bouvier, L. Devi, S. R. George, M. J. Lohse, G. Milligan, R. Neubig, K. Palczewski, M. Parmentier, J. P. Pin, G. Vriend, J. A. Javitch, F. Campagne and M. Filizola, *BMC Bioinformatics*, 2007, **8**, 177.
215. J. Kazius, K. Wurdinger, M. van Iterson, J. Kok, T. Back and A. P. Ijzerman, *Hum. Mutat.*, 2008, **29**, 39-44.
216. S. P. Alexander, H. E. Benson, E. Faccenda, A. J. Pawson, J. L. Sharman, M. Spedding, J. A. Peters and A. J. Harmar, *Br. J. Pharmacol.*, 2013, **170**, 1459-1581.
217. Y. Okuno, A. Tamon, H. Yabuuchi, S. Niijima, Y. Minowa, K. Tonomura, R. Kunimoto and C. Feng, *Nucleic Acids Res.*, 2008, **36**, D907-D912.
218. J. Zhang and Y. Zhang, *Bioinformatics*, 2010, **26**, 3004-3005.
219. J. J. Irwin and B. K. Shoichet, *J. Chem. Inf. Model.*, 2005, **45**, 177-182.
220. C. J. Langmead, S. P. Andrews, M. Congreve, J. C. Errey, E. Hurrell, F. H. Marshall, J. S. Mason, C. M. Richardson, N. Robertson, A. Zhukov and M. Weir, *J. Med. Chem.*, 2012, **55**, 1904-1909.
221. F. Sirci, E. P. Istyastono, H. F. Vischer, A. J. Kooistra, S. Nijmeijer, M. Kuijer, M. Wijtmans, R. Mannhold, R. Leurs, I. J. de Esch and C. de Graaf, *J. Chem. Inf. Model.*, 2012, **52**, 3308-3324.
222. M. Baroni, G. Cruciani, S. Sciabola, F. Perruccio and J. S. Mason, *J. Chem. Inf. Model.*, 2007, **47**, 279-294.
223. S. Cross, M. Baroni, L. Goracci and G. Cruciani, *J. Chem. Inf. Model.*, 2012, **52**, 2587-2598.
224. S. Cross, F. Ortuso, M. Baroni, G. Costa, S. Distinto, F. Moraca, S. Alcaro and G. Cruciani, *J. Chem. Inf. Model.*, 2012, **52**, 2599-2608.
225. P. Kolb, K. Phan, Z. G. Gao, A. C. Marko, A. Sali and K. A. Jacobson, *PLoS One*, 2012, **7**, e49910.
226. D. R. Weiss, S. Ahn, M. F. Sassano, A. Kleist, X. Zhu, R. Strachan, B. L. Roth, R. J. Lefkowitz and B. K. Shoichet, *ACS Chem. Biol.*, 2013, **8**, 1018-1026.
227. Y. Yoshikawa, S. Oishi, T. Kubo, N. Tanahara, N. Fujii and T. Furuya, *J. Med. Chem.*, 2013, **56**, 4236-4251.
228. D. Schmidt, V. Bernat, R. Brox, N. Tschammer and P. Kolb, *ACS Chem. Biol.*, 2015, **10**, 715-724.
229. A. Gaulton, L. J. Bellis, A. P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani and J. P. Overington, *Nucleic Acids Res.*, 2012, **40**, D1100-1107.

230. M. Pappalardo, N. Shachaf, L. Basile, D. Milardi, M. Zeidan, J. Raiyn, S. Guccione and A. Rayan, *PLoS One*, 2014, **9**, e109340.
231. P. R. Daga, W. E. Polgar and N. T. Zaveri, *J. Chem. Inf. Model.*, 2014, **54**, 2732-2743.
232. M. Vass, E. Schmidt, F. Horti and G. M. Keserü, *Eur. J. Med. Chem.*, 2014, **77**, 38-46.
233. E. Szollosi, A. Bobok, L. Kiss, M. Vass, D. Kurko, S. Kolok, A. Visegrady and G. M. Keserü, *Bioorg. Med. Chem.*, 2015.
234. E. P. Istyastono, A. J. Kooistra, H. F. Vischer, M. Kuijer, L. Roumen, S. Nijmeijer, R. A. Smits, I. J. P. de Esch, R. Leurs and C. de Graaf, *MedChemComm*, 2015.
235. Y. Liu, E. Zhou, K. Yu, J. Zhu, Y. Zhang, X. Xie, J. Li and H. Jiang, *Molecules*, 2008, **13**, 2426-2441.
236. C. S. Tautermann, *Future medicinal chemistry*, 2011, **3**, 709-721.
237. M. Congreve, S. P. Andrews, A. S. Dore, K. Hollenstein, E. Hurrell, C. J. Langmead, J. S. Mason, I. W. Ng, B. Tehan, A. Zhukov, M. Weir and F. H. Marshall, *J. Med. Chem.*, 2012, **55**, 1898-1903.
238. P. A. Procopiou, C. Browning, J. M. Buckley, K. L. Clark, L. Fechner, P. M. Gore, A. P. Hancock, S. T. Hodgson, D. S. Holmes, M. Kranz, B. E. Looker, K. M. Morriss, D. L. Parton, L. J. Russell, R. J. Slack, S. L. Sollis, S. Vile and C. J. Watts, *J. Med. Chem.*, 2011, **54**, 2183-2195.
239. X. Dong, Y. Zhao, X. Huang, K. Lin, J. Chen, E. Wei, T. Liu and Y. Hu, *Eur. J. Med. Chem.*, 2013, **62**, 754-763.
240. R. D. Cramer, D. E. Patterson and J. D. Bunce, *J. Am. Chem. Soc.*, 1988, **110**, 5959-5967.
241. S. Han, F. F. Zhang, H. Y. Qian, L. L. Chen, J. B. Pu, X. Xie and J. Z. Chen, *Eur. J. Med. Chem.*, 2015, **93**, 16-32.
242. S. Paoletta, D. K. Tosh, A. Finley, E. T. Gizewski, S. M. Moss, Z. G. Gao, J. A. Auchampach, D. Salvemini and K. A. Jacobson, *J. Med. Chem.*, 2013, **56**, 5949-5963.
243. D. K. Tosh, S. Paoletta, Z. Chen, S. Crane, J. Lloyd, Z.-G. Gao, E. T. Gizewski, J. A. Auchampach, D. Salvemini and K. A. Jacobson, *MedChemComm*, 2015, **6**, 555-563.
244. V. Yaziji, D. Rodriguez, H. Gutierrez-de-Teran, A. Coelho, O. Caamano, X. Garcia-Mera, J. Brea, M. I. Loza, M. I. Cadavid and E. Sotelo, *J. Med. Chem.*, 2011, **54**, 457-471.
245. V. Yaziji, D. Rodriguez, A. Coelho, X. Garcia-Mera, A. El Maatougui, J. Brea, M. I. Loza, M. I. Cadavid, H. Gutierrez-de-Teran and E. Sotelo, *Eur. J. Med. Chem.*, 2013, **59**, 235-242.
246. H. Gutierrez-de-Teran, H. Keranen, J. Azuaje, D. Rodriguez, J. Aqvist and E. Sotelo, *Methods Mol. Biol.*, 2015, **1272**, 271-291.
247. J. Moult, K. Fidelis, A. Kryshtafovych, T. Schwede and A. Tramontano, *Proteins*, 2014, **82 Suppl 2**, 1-6.
248. M. F. Lensink, R. Mendez and S. J. Wodak, *Proteins*, 2007, **69**, 704-718.
249. M. F. Lensink and S. J. Wodak, *Proteins*, 2010, **78**, 3073-3084.
250. R. A. Bathgate, M. H. Oh, W. J. Ling, Q. Kaas, M. A. Hossain, P. R. Gooley and K. J. Rosengren, *Frontiers in endocrinology*, 2013, **4**, 13.
251. C. G. Gadhe and M. H. Kim, *Mol. BioSyst.*, 2015, **11**, 618-634.
252. Y. Nikaido, Y. Koyama, Y. Yoshikawa, T. Furuya and S. Takeda, *J Biochem*, 2015, **157**, 311-320.
253. N. Montpas, J. Cabana, G. St-Onge, S. Gravel, G. Morin, T. Kuroyanagi, P. Lavigne, N. Fujii, S. Oishi and N. Heveker, *Biochemistry*, 2015, **54**, 1505-1515.
254. H. J. Wittmann and A. Strasser, *Bioorg. Med. Chem. Lett.*, 2015, **25**, 1259-1268.
255. G. Deganutti, A. Cuzzolin, A. Ciancetta and S. Moro, *Bioorg. Med. Chem.*, 2015, **23**, 4065-4071.
256. T. M. Stepniewski and S. Filipek, *Bioorg. Med. Chem.*, 2015, **23**, 4072-4081.
257. S. Mente, E. Guilmette, M. Salafia and D. Gray, *Bioorg. Med. Chem. Lett.*, 2015, **25**, 2106-2111.
258. E. P. Istyastono, S. Nijmeijer, H. D. Lim, A. van de Stolpe, L. Roumen, A. J. Kooistra, H. F. Vischer, I. J. de Esch, R. Leurs and C. de Graaf, *J. Med. Chem.*, 2011, **54**, 8136-8147.
259. G. Shahane, C. Parsania, D. Sengupta and M. Joshi, *PLoS Comput. Biol.*, 2014, **10**, e1004006.
260. N. Jatana, L. Thukral and N. Latha, *Proteins*, 2014, **83**, 867-880.
261. B. D. Hudson, M. E. Due-Hansen, E. Christiansen, A. M. Hansen, A. E. Mackenzie, H. Murdoch, S. K. Pandey, R. J. Ward, R. Marquez, I. G. Tikhonova, T. Ulven and G. Milligan, *J. Biol. Chem.*, 2013, **288**, 17296-17312.
262. D. K. Tosh, F. Deflorian, K. Phan, Z. G. Gao, T. C. Wan, E. Gizewski, J. A. Auchampach and K. A. Jacobson, *J. Med. Chem.*, 2012, **55**, 4847-4860.
263. V. Canale, P. Guzik, R. Kurczab, P. Verdie, G. Satala, B. Kubica, M. Pawlowski, J. Martinez, G. Subra, A. J. Bojarski and P. Zajdel, *Eur. J. Med. Chem.*, 2014, **78**, 10-22.
264. A. L. Brame, J. J. Maguire, P. Yang, A. Dyson, R. Torella, J. Cheriyan, M. Singer, R. C. Glen, I. B. Wilkinson and A. P. Davenport, *Hypertension*, 2015, **65**, 834-840.
265. J. Staron, D. Warszycki, J. Kalinowska-Tluscik, G. Satala and A. J. Bojarski, *RSC Advances*, 2015, **5**, 25806-25815.
266. J. Cheng, P. M. Giguere, O. K. Onajole, W. Lv, A. Gaisin, H. Gunosewoyo, C. M. Schmerberg, V. M. Pogorelov, R. M. Rodriguiz, G. Vistoli, W. C. Wetsel, B. L. Roth and A. P. Kozikowski, *J. Med. Chem.*, 2015, **58**, 1992-2002.
267. R. Petrelli, I. Torquati, S. Kachler, L. Luongo, S. Maione, P. Franchetti, M. Grifantini, E. Novellino, A. Lavecchia, K. N. Klotz and L. Cappellacci, *J. Med. Chem.*, 2015, **58**, 2560-2566.
268. D. D. Yu, K. M. Sousa, D. L. Mattern, J. Wagner, X. Fu, N. Vaidehi, B. M. Forman and W. Huang, *Bioorg. Med. Chem.*, 2015, **23**, 1613-1628.
269. D. K. Tosh, S. Paoletta, Z. Chen, S. M. Moss, Z. G. Gao, D. Salvemini and K. A. Jacobson, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 3302-3306.
270. A. Heifetz, O. Barker, G. B. Morris, R. J. Law, M. Slack and P. C. Biggin, *Biochemistry*, 2013, **52**, 8246-8260.
271. S. Bertini, T. Parkkari, J. R. Savinainen, C. Arena, G. Saccomanni, S. Saguto, A. Ligresti, M. Allara, A. Bruno, L. Marinelli, V. Di Marzo, E. Novellino, C. Manera and M. Macchia, *Eur. J. Med. Chem.*, 2015, **90**, 526-536.
272. D. Rodriguez, A. Pineiro and H. Gutierrez-de-Teran, *Biochemistry*, 2011, **50**, 4194-4208.
273. S. K. Kim, Z. G. Gao, P. Van Rompaey, A. S. Gross, A. Chen, S. Van Calenbergh and K. A. Jacobson, *J. Med. Chem.*, 2003, **46**, 4847-4859.
274. A. Roy, A. Kucukural and Y. Zhang, *Nat. Protoc.*, 2010, **5**, 725-738.
275. D. E. Gloriam, P. Wellendorph, L. D. Johansen, A. R. Thomsen, K. Phonekeo, D. S. Pedersen and H. Brauner-Osborne, *Chem. Biol.*, 2011, **18**, 1489-1498.
276. S. Vohra, B. Taddese, A. C. Conner, D. R. Poyner, D. L. Hay, J. Barwell, P. J. Reeves, G. J. G. Upton and C. A. Reynolds, *J. Royal Soc. Interface*, 2012, **10**.
277. M. Yanagawa, T. Yamashita and Y. Shichida, *J. Biol. Chem.*, 2013, **288**, 9593-9601.
278. X. Sun, J. Cheng, X. Wang, Y. Tang, H. Agren and Y. Tu, *Scientific reports*, 2015, **5**, 8066.
279. D. Wootten, J. Simms, L. J. Miller, A. Christopoulos and P. M. Sexton, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 5211-5216.
280. M. J. Moon, Y. N. Lee, S. Park, A. Reyes-Alcaraz, J. I. Hwang, R. P. Millar, H. Choe and J. Y. Seong, *J. Biol. Chem.*, 2015, **290**, 5696-5706.
281. X. He, S. K. Lakkaraju, M. Hanscom, Z. Zhao, J. Wu, B. Stoica, A. D. MacKerell, Jr., A. I. Faden and F. Xue, *Bioorg. Med. Chem.*, 2015, **23**, 2211-2220.
282. Z. Feng, S. Ma, G. Hu and X. Q. Xie, *AAPS J.*, 2015, **17**, 737-753.