G protein-coupled receptors: structure- and function-based drug discovery

Dehua Yang1,2, Qingtong Zhou3, Viktorija Labroska1,4, Shanshan Qin5, Sanaz Darbalaei1,4, Yiran Wu5, Elita Yuliantie1,4, Linshan Xie5,6, Houchao Tao5, Jianjun Cheng5, Qing Liu1,2, Suwen Zhao5,6, Wening Shui5,6, Yi Jiang2 and Ming-Wei Wang6,7

As one of the most successful therapeutic target families, G protein-coupled receptors (GPCRs) have experienced a transformation from random ligand screening to knowledge-driven drug design. We are eye-witnessing tremendous progresses made recently in the understanding of their structure–function relationships that facilitated drug development at an unprecedented pace. This article intends to provide a comprehensive overview of this important field to a broader readership that shares some common interests in drug discovery.

Signal Transduction and Targeted Therapy (2021) 6:7; https://doi.org/10.1038/s41392-020-00435-w

INTRODUCTION

G protein-coupled receptors (GPCRs) represent the largest protein family encoded by the human genome. Located on the cell membrane, they transduce extracellular signals into key physiological effects.1 Their endogenous ligands include odors, hormones, neurotransmitters, chemokines, etc., varying from photons, amines, carbohydrates, lipids, peptides to proteins. GPCRs have been implicated in a large number of diseases, such as Type 2 diabetes mellitus (T2DM), obesity, depression, cancer, Alzheimer’s disease, and many others.2 Activated by external signals through coupling to different G proteins or arrestins, GPCRs elicit cyclic adenosine 3,5-monophosphate (cAMP) response, calcium mobilization, or phosphorylation of extracellular regulated protein kinases 1/2 (pERK1/2).3 The seven-transmembrane protein property endows them easy to access, while the diversified downstream signaling pathways make them attractive for drug development.4 The human GPCR family is divided into classes A (rhodopsin), B (secretin and adhesion), C (glutamate), and F (Frizzled) subfamilies according to their amino acid sequences (Fig. 1). Of the 826 human GPCRs, approximately 350 non-olfactory members are regarded as druggable and 165 of them are validated drug targets (Fig. 1 and Table S1).5–6 Latest statistical data indicate that 527 Food and Drug Administration (FDA)-approved drugs1,8 and ~60 drug candidates currently in clinical trials target GPCRs (Table S1).5

Started with crystal structure determination and accelerated by cryo-electron microscopy (cryo-EM) technology, three-dimensional (3D) structural studies on a variety of GPCRs in complex with ligands, G proteins/arrestins, or both7–10 (involving 455 structures from 82 different receptors) significantly deepened our knowledge of molecular mechanisms of signal transduction. Novel insights into ligand recognition and receptor activation are gained from inactive, transitional, active, and apo states, thereby offering new opportunities for structure-based drug design (SBDD).11 Pharmacological parameters such as cAMP accumulation, calcium flux, ERK phosphorylation, arrestin recruitment, and G protein interaction12,13 are commonly used to evaluate ligand action and biased signaling. Ligand-binding kinetics and signaling timing render another dimension for interpreting signal bias profiles and link in vitro bioactivities with in vivo effects.14 In this process, a series of biased and allosteric modulators were discovered by rational design, ligand screening, and pharmacological assessment leading to the identification of novel binding sites or action modes.15,16

Apart from crystallography and cryo-EM, the striking advancement in GPCR biology is also attributable to the deployment of powerful technologies such as nuclear magnetic resonance (NMR), hydrogen–deuterium exchange (HDX), fluorescence resonance energy transfer, bioluminescence resonance energy transfer, surface plasmon resonance, single molecule fluorescence, CRISPR/Cas9, artificial intelligence, etc. This review systematically summarizes the latest information on this important drug target family to cover both basic and translational sciences in the context of drug discovery and development.

GPCR AS DRUG TARGET

Class A

Class A GPCRs, the so called “rhodopsin-like family” consisting of 719 members, are divided into several subgroups: aminergic, peptide, protein, lipid, melatonin, nucleotide, steroid, allicarboxylic acid, sensory, and orphan.17 They have a conventional transmembrane domain (TMD) that forms ligand-binding pocket and additional eight helices with a palmitoylated cysteine at the C terminal.18,19 Given the wide range of their physiological functions, this class of receptors is the most targeted...
therapeutically among all other classes. By manually curating Drugs@FDA original New Drug Application (NDA) and Biologic License Application (BLA) database (data extracted from August 2017 to June 2020) and cross-referencing with Drugbank, IUPHAR and ChemBL databases, we were able to find the approved drugs associated with this class.

Over 500 GPCR drugs target class A and many of them act at >1 receptor: 75% are made against aminergic receptors and 10% for peptidic ligand receptors with indications ranging from analgesics, allergies, cardiovascular diseases, hypertension, pulmonary diseases, depression, migraine, glaucoma, Parkinson’s disease to schizophrenia, cancer-related fatigue, etc. Approximately 500 novel drug candidates are in clinical trials. Of them, 134 are for peptide-activated GPCRs, while small molecules still occupy the majority. It is noted that 6% of class A members are sensory and alicarboxylic acid receptors that have broad untapped therapeutic potentials (Table S1). Chemokine, prostanoid and melanocortin receptors constitute >8% clinical trial targets in this class.

In the past 3 years, about 20 NDAs were approved targeting mostly peptide and aminergic receptors (Table 1). Siponimod and ozanimod provide alternatives to fingolimod (approved in 2010) for treating relapsing forms of multiple sclerosis by modulating sphingosine-1-phosphate receptor. Two radiolabeled ligands, gallium 68 dotatoc and lutetium 177 dotate, have been approved for neuroendocrine tumor and pancreatic gastrointestinal cancer diagnosis, respectively. Pitolisant, a selective inverse

---

**Fig. 1** Phylogenetic tree of GPCRs as drug targets. Node represents GPCR named according to its gene name. Receptors with approved drugs on the market are highlighted by color. GPCRs are organized according to GPCR database. Approved drug list was derived from previous publications, complemented by additional search of newly approved entities at Drugs@FDA (accessdata.fda.gov) until June 2020. See Table S2 for details.
| Drug | Brand name (manufacturer) | Indication | Target | GPCR class | FDA approval date |
|------|--------------------------|------------|--------|------------|------------------|
| Gilteritinib | Xospata (Astellas) | Relapsed or refractory acute myeloid leukemia | Serotonin receptors | A, aminergic, 5-hydroxytryptamine | 11/28/2018 |
| Lasmiditan | Reyvow (Eli Lilly) | Migraine | HTR1F | A, aminergic, 5-hydroxytryptamine | 10/11/2019 |
| Revefenacin | Yupeki (Mylan Ireland) | Chronic obstructive pulmonary disease | CHRM1-CHRM5 | A, aminergic, acetylcholine | 11/09/2018 |
| Lumateperone | Caplyta (Intra-Cellular) | Schizophrenia | HTR2A, DRD1, DRD2 | A, aminergic, dopamine, 5-hydroxytryptamine | 12/20/2019 |
| Amisulpride | Barhemys (Acacia Pharma) | Surgery-induced nausea and vomiting prevention | DRD2, DRD3, HTR7, HTR2A | A, aminergic, dopamine, 5-hydroxytryptamine | 02/26/2020 |
| Pitolisant | Wakix (Harmony) | Narcolepsy excessive daytime sleepiness | HRH3 | A, aminergic, histamine | 08/14/2019 |
| Angiotensin II | Giapreza (La Jolla Pharma) | Septic vasoconstrictor for adults | AGTR1 | A, peptide, angiotensin | 12/21/2017 |
| Macimorelin | Macrilen (Novo Nordisk) | Diagnosis of adult growth hormone deficiency | GHSR | A, peptide, ghrelin | 12/20/2017 |
| Elagolix | Orilissa (Abbvie) | Endometriosis-associated moderate-to-severe pain | GNHR | A, peptide, gonadotropin | 07/23/2018 |
| Cysteamine | Procysbi (Horizon Pharma) | Radiation sickness | NPY2R | A, peptide, neuropeptide Y | 02/14/2020 |
| Lemborexant | Davyigo (Eisai) | Insomnia | HCRTR1 | A, peptide, orexin | 12/20/2019 |
| Gallium 68 dotatoc | NA (UIHC-PET Imaging Center) | Diagnostic agent for neuroendocrine tumors | SSTR2 | A, peptide, somatostatin | 08/21/2019 |
| Lutetium 177 dotatate | Lutathera (AAA USA) | Gastroenteropancreatic neuroendocrine tumors | SSTR2 | A, peptide, somatostatin | 01/26/2018 |
| Fosnetupitant/palonosetron | Akynzeo (Helsinn Hlthcare) | Chemotherapy-associated nausea and vomiting prevention | TACR1 | A, peptide, tachykinin | 04/19/2018 |
| Mogamulizumab-kpc | Poteligeo (Kyowa Kirin) | Non-Hodgkin lymphoma | CXCR4 | A, protein, chemokine | 08/08/2018 |
| Siponimod | Mayzent (Novartis) | Relapsing forms of multiple sclerosis | S1PR1, S1PR5 | A, lipid, lymphosphospholipid | 03/26/2019 |
| Ozanimod | Zeposia (Celgene) | Relapsing forms of multiple sclerosis | S1PR1, S1PR5 | A, lipid, lymphosphospholipid | 03/25/2020 |
| Cannabidiol | Epidiolex (GW Research) | Epilepsy | CNR1 | A, lipid, cannabionid | 06/25/2018 |
| Latanoprostene bunod | Vyzulta (Bausch and Lomb) | Glaucoma or ocular hypertension | PTGFR | A, lipid, prostanoid | 11/02/2017 |
| Istradefylline | Nourianz (Kyowa Kirin) | Parkinson's disease | ADORA2A | A, nucleotide, adenosine | 08/27/2019 |
| Fostamatinib | Tavalisse (Rigel Pharma) | Chronic immune thromboctopenia | Multiple targets, including ADORA3 | A, nucleotide, adenosine | 04/17/2018 |

The drugs listed above were identified manually from Drugs@FDA original NDA and BLA database (data extracted from August 2017 to June 2020) and cross-referenced with Drugbank, IUPHAR, and ChemBL databases. ADORA2A (A2AR) adenosine A2a receptor, ADORA3 (A3AR) adenosine A3 receptor, AGTR1 (AT1R) angiotensin II receptor type 1, CHRM1 (M1R) muscarinic acetylcholine receptor M1, CHRM5 (M5R) muscarinic acetylcholine receptor M5, CNR1 (CB1) cannabinoid receptor 1, CXCR4 C-X-C chemokine receptor type 4, DRD1–DRD3 D1–D3 dopamine receptor, GHSR growth hormone secretagogue receptor, GNRHR gonadotropin-releasing hormone receptor, HCRTR1 (OX1R) orexin receptor type 1, HRH3 histamine H3 receptor, HTR1F 5-hydroxytryptamine receptor 1F, HTR2A (5-HT2A) 5-hydroxytryptamine receptor 2A, HTR7 5-hydroxytryptamine receptor 7, NPY2R neuropeptide Y receptor Y2, PTGFR prostaglandin F receptor, S1PR1 sphingosine 1-phosphate receptor 1, S1PR5 sphingosine 1-phosphate receptor 5, SSTR2 somatostatin receptor 2, TACR1 (NK1R) substance-P receptor.
agonist of histamine receptor, is used to treat narcolepsy-related daytime sleepiness, while lemborexant, an orexin receptor antagonist, is used for insomnia management. Gilteritinib (ASP2215) is a small molecule inhibitor of tyrosine kinase. However, it also antagonizes serotonin receptors without any reported pharmacological consequences. Revefenacin is a long-acting antagonist of muscarinic acetylcholine receptors (mACHRs) indicated for chronic obstructive pulmonary disease. Amisulpride, trialed for antiemetic and schizophrenia, was finally approved for antiemetic in 2020. This molecule is acting as an antagonist against dopamine and serotonin receptors. Fosnetupitant, a prodrug of netupitant, was approved for chemotherapy-induced nausea and vomiting. Cysteamine treats radiation sickness via modifying action of neuropeptide Y receptor. Cannabidiol is one of the active constituents of the Cannabis plant and was trialed for schizophrenia, graft versus host disease, and anticonvulsant. It was eventually approved in 2018 for the treatment of severe forms of epilepsy—Lennox–Gastaut syndrome and Dravet syndrome. Meanwhile, fostamatinib, indicated for chronic immune thrombocytopenia, targets >300 receptors and enzymes, including adenosine receptor A3.

Class B

This class of GPCRs is divided into two subfamilies: secretin (B1) and adhesion (B2), containing 15 and 33 members, respectively. Secretin subfamily members are characteristic of large extracellular domains (ECDs) and bind to vasoactive intestinal peptide (VIP), pituitary adenylate cyclase-activating peptide (PACAP), corticotropin-releasing factor (CRF), parathyroid hormone (PTH), growth hormone-releasing hormone (GHRH), calcitonin gene-related peptide (CGRP), glucagon, and glucagon-like peptides (GLPs), respectively. Adhesion subfamily has nine subgroups, possessing unique N-terminal motifs, such as epidermal growth factor, cadherin, and immunoglobulin domains. They are distinguished from other GPCRs due to their roles in cell adhesion and migration. Apart from the long N-terminal domain, other unique features of the B2 subfamily are the GPCR autoproteolysis-inducing domain and the proteolysis site that are responsible for signaling activation through a Stachel sequence (a tethered agonist) and producing N-terminal fragment (NTF) and C-terminal fragment. The hallmarks of the B2 GPCR subfamily are a two-step activation model, the ligand–NTF interaction and the Stachel signaling/basal activity. Adhesion receptors can also signal independently of fragment dissociation and this has complicated pharmacological consequences.

In this class, receptors of glucagon family peptides, followed by CGRP, PTH, GHRH, CRF, VIP, and PACAP, constitute major targets for therapeutic intervention (Table S1) of various diseases, including obesity, T2DM, osteoporosis, migraine, depression, and anxiety.

To date, multiple GLP-1 receptor (GLP-1R) agonists have been developed by a combination of selective amino acid substitutions, enzymatic cleavage blockade, and conjugation to entities that increase binding to plasma proteins. These methods not only slow down fast renal clearance of the peptides but also extend their half-lives. Dose-dependent side effects such as nausea and gastrointestinal adverse events are the main drawbacks that are becoming more of a compliant with dose scaling. For instance, one newly approved GLP-1R agonist, semaglutide, has a noticeable half-life of 168 h thereby allowing weekly subcutaneous administration, while oral semaglutide (approved in 2019) formulated using absorption enhancer shows a similar half-life but is dosed daily with reported side effects (Table 2).

One of the latest approaches to develop more efficacious therapeutics against T2DM and obesity relates to dual- and tri-agonists targeting two or more of GLP-1R, glucagon receptor (GGCR), and glucose-dependent insulinoletic peptide receptor (GIPR). Many of them are currently in different phases of clinical trials (Table 3). Of note, in this receptor family, GLP-2 stimulates intestinal growth and an approved GLP-2R agonist, teduglutide, is used to treat short bowel syndrome.

CGRP family has a considerable clinical relevance. For instance, pramlintide that targets amylin receptor is utilized to treat both type 1 and type 2 diabetes. Salmon calcitonin has been explored as a treatment for Paget’s disease and metabolic disorders. Furthermore, the association of migraine and CGRP elevation led FDA-approved monoclonal antibodies (mAbs) against its receptor, e.g., erenumab and eptinezumab, as well as several small molecule antagonists such as rimegepant and ubrogepant (Table 2). Two approved diagnostic agents are analogs of CRF (corticotelin ovine trifluoride peptide) and GHRH (sermorelin) for diagnosis of Cushing’s disease or ectopic adrenocorticotropic hormone syndrome and hormone deficiency, respectively. Tesamorelin, another synthetic form of GHRH, was approved in 2010 to treat human immunodeficiency virus (HIV)-associated lipodystrophy.

Table 2. Newly approved drugs targeting class B GPCRs in the past 3 years

| Drug                  | Brand name (manufacturer) | Indication                     | Target               | GPCR class | FDA approval date |
|-----------------------|---------------------------|--------------------------------|----------------------|------------|-------------------|
| Erenumab-aooe         | Aimovig (Amgen)           | Migraine (prevention)          | CALCRL B1, peptide, calcitonin | 05/17/2018 |
| Ubrogepant            | Ubrely (Allergan)         | Migraine                       | CALCRL B1, peptide, calcitonin | 12/23/2019 |
| Rimegepant            | Nurtec ODT (Biohaven Pharm)| Migraine                       | CALCRL B1, peptide, calcitonin | 02/27/2020 |
| Eptinezumab-jjmr      | Vyepti (Lundbeck)         | Migraine (prevention)          | CALCRL B1, peptide, calcitonin | 02/21/2020 |
| Semaglutide (injection)| Ozempic (Novo Nordisk)    | Type 2 diabetes mellitus GLP-1R B1, peptide, glucagon-like peptide-1 | 12/05/2017 |
| Semaglutide (oral)    | Rybelsus (Novo Nordisk)   | Type 2 diabetes mellitus GLP-1R B1, peptide, glucagon-like peptide-1 | 09/20/2019 |

The drugs listed above were identified manually from Drugs@FDA original NDA and BLA database (data extracted from August 2017 to June 2020) and cross-referenced with Drugbank, EUPHAR, and ChemBL databases.

CALCRL (CGRP) calcitonin gene-related peptide type 1 receptor.
and ADGRF1 are considered as potential drug targets due to their extensive pathogenetic involvement. Two ADGRG1/ADGRG5 modulators, dihydromunduletone and α-acetoxydihydrodeoxygedunin developed via drug screening efforts, showed disease-related efficacy changes thereby calling for exploration of their activities in a pathological environment. However, associated drug resistance may not only hamper disease but also offer insights into potential mechanisms of such resistance and strategies to tackle it.

Classes C and F
Class C (glutamate) contains 22 receptors, which are further divided into 5 subfamilies including 1 calcium-sensing receptor (CaSR), 2 gamma-aminobutyric acid (GABA) type B receptors (GABAB1 and GABAB2), 3 taste 1 receptors (TS1R1/3), 8 metabotropic glutamate receptors (mGluR1–8), and 8 orphan GPCRs. The distinctive features of glutamate subfamily are their large ECD enabling the homodimers or heterodimers of class F binding site. Two conserved disulfide structural information of ECD indicates the roles of conserved residues for receptor activation. Linking of Wnt with extracellular CRD would activate downstream effectors of Wnt signaling, while the dimerization process and the interaction between CRD and TMD remain elusive. It is known that the downstream effectors of Wnt signaling consist of β-catenin, planar cell polarity, and Ca2+ pathways, whereas receptor activation involves in Wnt, Norrin, Fzd, LDL receptor-related protein S6, and movement disorders. Acamprosate, an antagonist of mGluR5, was launched in 2004 as an anti-neoplastic agent. In fact, mGluRs have been vigorously pursued as therapeutic targets and there are 15 drug candidates undergoing clinical trials at present for pain, migraine, Parkinson’s disease, Fragile X syndrome, etc. Although allosteric modulators of class C have attracted significant development efforts involving 8 clinical trial stage compounds (2 positive (PAM) and 6 negative (NAM) allosteric modulators), the only success is cinacalcet, a small molecule PAM of CaSR approved in 2004 for hyperparathyroidism and calcimimetics.

Only one class F GPCR (smoothed receptor SMO) has been validated as a drug target whose small molecule antagonists were approved as anti-neoplastic agents. Other 10 members of this class are all Frizzled receptors (FZD1–10), which mediate Wnt signaling and are essential for embryonic development and adult organisms. FZDs together with cognate Hedgehog and Wnt signal pathways, whereas receptor activation +

Table 3. Mono-, dual- and tri-agonists targeting GLP-1R, GCGR, and GIPR

| Receptor                  | Drug          | Dose form | Manufacturer      | Status      |
|---------------------------|---------------|-----------|-------------------|-------------|
| GLP-1R mono-agonist       | Exenatide     | SC, twice daily | AstraZeneca      | Approved    |
|                           | Liraglutide   | SC, once daily | Novo Nordisk     | Approved    |
|                           | Exenatide     | SC, once weekly | AstraZeneca      | Approved    |
|                           | Lixisenatide  | SC, once daily | Sanofi-Aventis   | Approved    |
|                           | Albiglutide   | SC, once weekly | GlaxoSmithKline  | Approved    |
|                           | Dulaglutide   | SC, once weekly | Eli Lilly        | Approved    |
|                           | Semaglutide   | SC, once weekly | Novo Nordisk     | Approved    |
|                           | Semaglutide   | Oral, once daily | Novo Nordisk     | Approved    |
| GLP-1R/GCGR dual-agonist  | HM12525A      | SC, once weekly | Hamni Pharmaceuticals | Phase 2 |
|                           | JNJ54728518   | SC          | Janssen Pharmaceuticals | Phase 1 |
|                           | MEDI0382      | SC, once daily | MedImmune        | Phase 2     |
|                           | MK8521        | SC, once daily | Merck            | Phase 2     |
|                           | NN9277        | SC          | Novo Nordisk     | Phase 1     |
|                           | MOD6030       | SC, once monthly | Prolor Biotech, Opko Health | Phase 1 |
|                           | SAR425899     | SC, once daily | Sanofi-Aventis   | Phase 2     |
|                           | VPD107        | SC, once weekly | Spitfire Pharma   | Preclinical |
|                           | TT401         | SC, once weekly | Transition Therapeutics | Phase 2 |
|                           | ZP2929        | SC, once daily | Zealnd           | Phase 1     |
| GLP-1R/GIPR dual-agonist  | CPD86         | SC, once daily | Eli Lilly        | Preclinical |
|                           | LY3298176     | SC, once weekly | Eli Lilly        | Phase 3     |
|                           | NN9709/MAR709/RG7697 | SC, once daily | Novo Nordisk/Marcadia | Phase 2 |
|                           | SAR438335     | —          | Sanofi-Aventis   | Trial discontinued |
|                           | ZP-I-98       | SC, once weekly | Zealnd           | Preclinical |
|                           | ZP-Dl-70      | SC, once weekly | Zealnd           | Preclinical |
| GLP-1R/GIPR/GIPR tri-agonist | HM15211     | SC          | Hamni Pharmaceuticals | Phase 2 |
|                           | MAR423        | SC, once daily | Novo Nordisk/Marcadia | Phase 1 |

Data were retrieved from the literature and updated to Drugs@FDA, ChemBL, and ClinicalTrials.gov databases.
many other co-factors. Key breakthrough is thus required to advance our knowledge of these receptors.

MEDICINAL CHEMISTRY OF GPCR

Agent type

Agents targeting GPCRs continue to expand in the past decades. Among them, exogenous small molecules, including traditionally developed synthetic organics, natural products, and inorganics, still dominate with a total percentage of 64% (Fig. 2). Nevertheless, the proportion of small molecules declines since 2010. In addition to traditional ligand discovery, several new modalities appear, though currently at the stage of academic research. Covalent ligands, with the embedding of reactive moieties that can be covalently linked to receptors, significantly enhance the weak binding of unoptimized leads. Photoactive ligands, developed by the introduction of photo-responsive groups to drug candidates, bring a new interdisciplinary field, photopharmacology. Albeit in its infancy, it has already found in vivo applications.

In comparison, biologicals, such as peptides, antibodies, and metabolites, become more and more visible in the list. Particularly, the number of approved peptide drugs occupies approximately one third of the whole repertoire, with many more in different clinical stages as the pipeline—most of them target classes A and B GPCRs. Naturally occurring peptides have been continually discovered from plants, animals, fungi, and bacteria. Although they act as efficient chemical messengers to modulate cellular functions, these peptides suffer from unfavorable pharmacokinetic and pharmacodynamics properties, such as very short plasma half-lives and low plasma protein binding. Therefore, chemical modifications are required to promote the membrane permeability, brain penetration, and oral bioavailability. Available strategies include peptide cyclization, N-methylation, palmitoylation, unnatural amino acid insertion, peptide–small molecule conjugation, and peptide self-assembly. By the way, developing peptidic agents may offer a new approach to de-orphanize certain orphan GPCRs.

mAbs represent a promising alternative in GPCR drug discovery. Over small molecules, mAbs possess obvious advantages of improved specificity, affinity, and other pharmacological properties. Thus they are being developed against cancers, inflammation, and metabolic disorders. To date, three GPCR-targeting mAbs were approved (mogamulizumab, erenumab, and epinezumab) while bi-specific antibodies, nanobodies, antibody–drug conjugates, and antibody–peptide conjugation are also in the development stage.

The emergence of many conceptually new molecular entities, such as RNA aptamer, provides not only powerful tool for biophysical study but also potential therapeutic candidates. Usually, aptamer has great molecular diversity and little immunogenicity. In addition, GPCRs are known to function by forming dimers (homodimers or heterodimers) and oligomers on the cell membrane. Therefore, strategies to induce receptor dimerization and/or oligomerization have received attention using scaffolds based on DNA (aptamer), small molecule, and physical stimuli.

Structure–activity relationship (SAR)

Studies of SARs are critical to the identification of drug-like molecules, especially when the crystal or cryo-EM structure of a drug target is not available. Given that many 3D GPCR structures have been solved in the past decade, most approved drugs were discovered without relevant structural information. Two examples are reviewed below to show the importance of SAR analysis.

Orrexin-1 and orexin-2 receptors (also known as hypocretin receptors,OX-R and OX-R2) are class A GPCRs for which two endogenous peptide ligands were identified, orexin A and orexin B (also known as hypocretin 1 and hypocretin 2). The orexin signaling system plays a crucial role in regulating the sleep/wake cycle—both OX-R and OX-R2 are involved while the precise contribution of each has yet to be defined. Therefore, dual antagonists were developed as potential treatment for insomnia.

Suvorexant (belsomra), the first-in-class dual orexin receptor antagonist, was launched in 2014. The second, lemborexant/E2006 (dayvigo) developed by Eisai, was approved by the FDA in 2020. It started from hit compound 1 (Fig. 2) with modest binding affinity to OX-R2 (Kd = 8.7 µM) and no affinity for OX-R1. The first round of SAR studies revealed that changing the ketone group to an amide led to a remarkable enhancement (~1000-fold) of binding affinity at both OX-R1 and OX-R2 (compound 2). Substitution of the aniline group with a 2-amino-5-cyano pyridine (compound 3) maintained OX-R2 affinity and reduced OX-R1 activity, but physicochemical properties were improved compared to compound 2. Further SAR studies focused on the modification of all three aromatic substitutions in compound 3. Changing the di-OMe-phenyl substituent to a pyrimidine group resulted in a significant loss of binding affinity, as shown with compound 4, but an improved overall profile due to reduced lipophilicity and enhanced solubility. Then replacing the cyano group to a fluoro regained the binding affinity for both receptors (compound 5), and finally adding a second fluoro to the benzene group significantly improved OX-R2 affinity and led to lemborexant.

Clearly, slight structural modifications may cause significant change of compound activity, and SAR studies coupled with optimization of physicochemical properties are useful steps to obtain drugable candidates.

CGRP is a 37-amino acid neuropeptide and its receptor is implicated in migraine. The benzodiazapinone compound 7 was identified as a hit compound with modest CGRP receptor binding affinity (Ki = 4.8 µM, Fig. 4). Replacing the right-hand

![Fig. 2](https://example.com)  
**Fig. 2** Analysis on agents targeting GPCRs. Distribution of molecule type (left) and action mode (right). Positive, PAM; Negative, NAM.
spirohydantoin structure with piperidyldihydroquinazolinone, a privileged structure for CGRP receptor antagonists, an affinity boost of 100-fold was gained. Further optimization of the benzodiazepinone core resulted in the caprolactam compound 9, which showed a $K_i$ of 25 nM. Changing the piperidyldihydroquinazolinone moiety to a piperidylazabenzimidazolone led to compound 10, with a binding affinity of 11 nM. Then by changing the N-substituent on the caprolactam and adding difluoro substitutions on the lower benzene ring delivered compound MK-0974 (11, $K_i = 0.77$ nM, Fig. 4), which entered clinical trials. Compound 12 (BMS-846372) shares the same piperidylazabenzimidazolone and the lower difluorobenzene substructures with 11 but differs from the latter with a carbamate core structure and a pyridine-fused-cyclopentane in replacement of the caprolactam. Compound 11 displayed high binding affinity while suffered from poor physicochemical properties, such as low solubility. To improve this, a hydroxyl group was attached to the cycloheptane ring and it was discovered that the (S)-isomer 13 was more potent

Fig. 3 SAR studies that led to the discovery of the dual orexin receptor antagonist lemborexant

Fig. 4 SAR studies that resulted in the discovery of CGRP antagonists
than the (R)-OH compound 14. The –OH was finally replaced with an -NH2 group, which led to the clinical compound rimegepant. The latter was further developed for better safety and efficacy profiles and obtained regulatory approval by the FDA in 2020.

The above examples demonstrate that, starting from a modest affinity hit compound, systematic SAR studies could successfully lead to very potent GPCR ligands that qualify as clinical candidates. Slight modifications of chemical structures sometimes cause remarkable changes of binding affinity or potency, which could not always be accurately predicted by conventional methods, such as docking. Therefore, SAR studies will continue to play a critical role in drug discovery.

**GPCR STRUCTURE**

The structure of GPCRs is a crucial determinant for understanding the molecular mechanisms underlying ligand recognition and receptor activation. It provides a foundation for drug discovery. The first crystal structure of inactive state rhodopsin purified from bovine eyes was solved in 2000. Although tremendous efforts have been made, elucidation of GPCR structures remains challenging due to several bottlenecks, including low receptor expression level, difficulties in extraction, highly flexible conformation, lack of crystal contacts, etc. The first crystal structure of GPCR extracted from exogenously expressed host cells, the human β2-adrenergic receptor (β2-AR, gene name: ADRB2) bound to an antagonist, was disclosed in 2007, representing a milestone in GPCR structural biology. Several innovative methods, especially the incorporation of a soluble fusion partner and lipidic cubic phase (LCP) crystallization, facilitated subsequent studies. Further technological breakthroughs in protein expression and purification, receptor engineering,9,9 application of Fab fragment and nanobody,9,55 and GPCR crystallization96 led to an exponential growth of this field.

The crystal structure of β2-AR in complex with stimulatory G protein (Gs) solved in 2011 revealed rhodopsin bound to visual arrestin reported in 2015 revealed the molecular mechanism of GPCR interaction with G protein and arrestin, respectively. Notably, the wave of resolution revolution in the single-particle cryo-EM has brought a significant impact on the determination of GPCR complexes. Over 90% of GPCR-transducer complex structures were solved using cryo-EM (Table 4). To date, a total of 455 structures from 82 GPCRs belonging to all classes except B2 leading to diversification of CRF1R105 and a common binding site for allosteric modulators of GCGR106 and GLP-1R107 located outside the TMD bundle of GCGR and GLP-1R, respectively, outside the TMD bundle of GCGR and GLP-1R, respectively.

A comparison of the full-length active receptor structures with that in the inactive state reveals a general activation mechanism for class B GPCRs (Fig. 6d, e). The binding of a peptidic ligand to mediate the activation. Collectively, these rearrangements and reorganizations of conserved motifs are critical to the activation of class A GPCRs. Class B GPCRs contain a large ECD and a TMD bundle with the peptide ligand recognition by both domains. According to the two-domain-binding model, the C-terminus of the peptide interacts with the ECD and orients the N-terminus of the peptide toward the TMD bundle. It then engages with the TMD core to facilitate receptor activation. The most remarkable structural feature of this class is the swing of ECD, accompanied by the corresponding shift of the peptide C-terminus (Fig. 6a, b). Conversely, the N-terminus inserts into a V-shape cavity within the helix bundle with a similar binding pose. Compared to small molecule-binding pocket of class A, that of class B is more solvent-accessible with higher flexibility and larger volume to accommodate sizeable peptidic ligands. In addition, structural studies also revealed an antagonist-binding pocket deep in the TMD bundle of CBR195 and a common binding site for allosteric modulators of GCGR106 and GLP-1R107 located outside the TMD bundle of GCGR and GLP-1R, respectively.

Finally, “DRY”, one of the most conserved motifs in class A receptors, locates at the bottom of the TTM and forms an interhelical salt bridge between D/E8.49 and R13.50. R13.50 forms an additional inter-helical salt bridge with D/E3.60, known as the ionic lock, which connects the intracellular ends of TM3 and TM6 to stabilize receptors in an inactive state (Fig. 5f). These contacts are eliminated after agonist binding, and R13.50 is released to interact with other residues to facilitate the G protein coupling. It is notable that an acidic residue at position 6.30 is less conserved in 30% of class A receptors. Alternatively, R13.50 may form polar interactions with other polar residues in TM6 (i.e., T6.27 in K-OR and μ-OR) to mediate the activation. Collectively, these rearrangements and reorganizations of conserved motifs are critical to the activation of class A GPCRs.

**Ligand binding pocket to the G protein-coupling region (Fig. 5b–f).** The binding of diverse agonists triggers the rotamer switch of W6.48, a highly conserved residue in the “CWxP” motif, and the concomitant side chain rotation of F6.44 (Fig. 5b). Upon stimulation by an agonist, conformational rearrangement occurs in the PIF (P.950, I.146, and P.444 Fig. 5c) and the Na+ pocket residues (D.2.50, S.3.49, N.4.65, and N.7.49, Fig. 5d). These reorganizations trigger the notable outward displacement of TM6, the hallmark of class A GPCR activation (Fig. 5b). The repacking of Na+ pocket residues initiates the TM7 movement toward TM3. Upon receptor activation, the “NPxY” residue Y5.53 changes its rotamer conformation and points toward TM3, rendering new contact formation between Y5.53 and residues in TM3 (L.3.42, I.4.66, and R.13.50, Fig. 5e) and subsequently the enhanced packing of TM3–TM7. Finally, “DRY”, one of the most conserved motifs in class A receptors, locates at the bottom of the TTM and forms an inter-helical salt bridge between D/E8.49 and R13.50. R13.50 forms an additional inter-helical salt bridge with D/E3.60, known as the ionic lock, which connects the intracellular ends of TM3 and TM6 to stabilize receptors in an inactive state (Fig. 5f). These contacts are eliminated after agonist binding, and R13.50 is released to interact with other residues to facilitate the G protein coupling. It is notable that an acidic residue at position 6.30 is less conserved in 30% of class A receptors. Alternatively, R13.50 may form polar interactions with other polar residues in TM6 (i.e., T6.27 in K-OR and μ-OR) to mediate the activation. Collectively, these rearrangements and reorganizations of conserved motifs are critical to the activation of class A GPCRs.
Table 4. List of GPCR structures

| Receptor   | Number of structures | PDB code (GPCR structure without downstream effector) | PDB code (GPCR structure with downstream effector) |
|------------|----------------------|-----------------------------------------------------|--------------------------------------------------|
| **Class A**|                      |                                                     |                                                  |
| ADORA1     | 3                    | 5N2S, 5UEN                                         | 6D9H                                             |
| ADORA2A    | 49                   | 2YDO, 2YDV, 3EML, 3PWH, 3QAK, 3REY, 3RFM, 3UZA, 3UZC, 3VG9, etc. | 6GDG                                             |
| ADRA2B     | 2                    |                                                     | 6K41, 6K42                                       |
| ADRB1      | 27                   | 2VT4, 2Y00, 2YCW, 3ZPQ, 4AMI, 4BVN, 4GPO, 5A8E, 5F8U, 6H7J, etc. | 6TKO, 7JJO                                       |
| ADRB2      | 33                   | 2R4R, 2RH1, 3D4S, 3K6, 3NY8, 3P0G, 3PDS, 4GBR, 4LDE, 5D5A, etc. | 3SN6, 6NI3                                       |
| AGTR1      | 6                    | 4YAY, 4ZUD, 6DO1, 6OS1, 6OS2, 6OS0                 |                                                  |
| AGTR2      | 5                    | 5UNF, 5UNG, 5UNH, 5XJM, 6JOD,                      |                                                  |
| APLNR      | 1                    |                                                     |                                                  |
| C5AR1      | 3                    | 5O9H, 6C1Q, 6C1R                                   |                                                  |
| CCR2       | 3                    | 5T1A, 6GPS, 6GPX                                   |                                                  |
| CCR5       | 6                    | 4MB5, 5UIW, 6AKX, 6AKY, 6MEO, 6MET                 |                                                  |
| CCR6       | 1                    |                                                     | 6WWZ                                             |
| CCR7       | 1                    |                                                     |                                                  |
| CCR9       | 1                    |                                                     |                                                  |
| CHRM1      | 3                    | 5CXV, 6WJC                                        | 6OJ                                              |
| CHRM2      | 11                   | 3U0N, 4MQS, 4MQT, 5YC8, 5ZK3, 5ZK8, 5ZKB, 5ZKC     | 6OIK, 6U1N, 6UP7                                  |
| CHRM3      | 5                    | 4DAJ, 4U14, 4U15, 4U16, 5ZHP                       |                                                  |
| CHRM4      | 1                    | 5DS5                                              |                                                  |
| CHRM5      | 1                    | 6OL9                                              |                                                  |
| CNR1       | 7                    | 5TGZ, 5U09, 5XR8, 5XRA, 6KQ1                       | 6N4B, 6KPG                                       |
| CNR2       | 4                    | 5ZTY, 6KPC                                        | 6PT0, 6KPF                                       |
| CXC1       | 1                    | 2LN1                                              |                                                  |
| CXC2       | 3                    | 6FL                                               | 6LFM, 6LFO                                       |
| CXC4       | 6                    | 3ODU, 3OE0, 3OE6, 3OE8, 3OE9, 4RWS                 |                                                  |
| CYSLTR1    | 2                    | 6RZ4, 6RZ5                                        |                                                  |
| CYSLTR2    | 4                    | 6RZ6, 6RZ7, 6RZ8, 6RZ9                             |                                                  |
| DRD2       | 2                    | 6CM4                                              | 6VMS                                             |
| DRD3       | 1                    | 3PBL                                              |                                                  |
| DRD4       | 3                    | 5WIU, 5WIV, 6IQL                                   |                                                  |
| EDNRB      | 8                    | 5GLH, 5GLI, 5X93, 5XPR, 6IGK, 6IGL, 6K1O, 6LRY    |                                                  |
| F2R        | 1                    | 3VW7                                              |                                                  |
| F2RL1      | 3                    | 5NDD, 5NDZ, 5N6                                   |                                                  |
| FFAR1      | 4                    | 4PHU, 5KW2, 5TZR, 5TZY                             |                                                  |
| FFAR2      | 2                    | 6LW5                                              | 6OMM                                             |
| GPBAR1     | 2                    |                                                     | 7CFM, 7CFN                                       |
| GPR52      | 4                    | 6LI1, 6LJ2, 6LJ0                                  | 6L13                                             |
| HCRTR1     | 11                   | 4ZJ8, 4ZJC, 6TO7, 6TOD, 6TP3, 6TP4, 6TP6, 6TQ4, 6TQ6, 6TQ7, 6TQ9 |                                                  |
| HCRTR2     | 6                    | 4S0V, 5WQC, 5WS3, 6TPG, 6TPJ, 6TPN                 |                                                  |
| HRH1       | 1                    | 3RZE                                              |                                                  |
| HTR1B      | 4                    | 4IAQ, 4IAR, 5VS4                                  | 6G79                                             |
| HTR2A      | 5                    | 6A93, 6A94, 6WH4, 6WGT                             | 6WHA                                             |
| HTR2B      | 8                    | 4IB4, 4NC3, 5UDN, 5YDN, 6DRX, 6DRY, 6DRZ, 6DS0     |                                                  |
| HTR2C      | 2                    | 6BOG, 6BQH                                        |                                                  |
| LPAR1      | 3                    | 4Z34, 4Z35, 4Z36                                  |                                                  |
| MC4R       | 1                    | 6W25                                              |                                                  |
| MTNR1A     | 5                    | 6ME2, 6ME3, 6ME4, 6ME5, 6PS8                       |                                                  |
| MTNR1B     | 4                    | 6ME6, 6ME7, 6ME8, 6ME9                            |                                                  |
| NPY1R      | 2                    | 5ZBH, 5ZBQ                                        |                                                  |
| Receptor | Number of structures | PDB code (GPCR structure without downstream effector) | PDB code (GPCR structure with downstream effector) |
|----------|----------------------|-----------------------------------------------------|--------------------------------------------------|
| NTSR1    | 11                   | 3ZEV, 4BU0, 4BV0, 4BVW, 4GRV, 4XE0, 4XES, 5T04         | 6OS9, 6OSA, 6PWC                                  |
| OPN1D1   | 4                    | 4E4J, 4N6H, 4RWA, 4RWD                                |                                                  |
| OPRK1    | 2                    | 4DJH, 6B73                                           |                                                  |
| OPR1L1   | 3                    | 4EAS, 5DHH, 5DHH                                     |                                                  |
| OPRM1    | 4                    | 4DKL, SC1M                                           | 6DDE, 6DDF                                      |
| OXTR     | 1                    | 6TPK                                                |                                                  |
| P2RY1    | 2                    | 4XNV, 4XNW                                           |                                                  |
| P2RY12   | 3                    | 4NTJ, 4PXZ, 4PY0                                    |                                                  |
| PTAFR    | 2                    | 5ZKP, S2KQ                                           |                                                  |
| PTGRD2   | 2                    | 6D26, 6D27                                          |                                                  |
| PTGER3   | 2                    | 6AK3, 6M9T                                           |                                                  |
| PTGER4   | 2                    | 5YHL, SYWY                                          |                                                  |
| RHO      | 55                   | 1F88, 1GZM, 1HZX, 1L9H, 1LN6, 1U19, 2G87, 2HPY, 2J35, | 4ZWJ, 5DGY, 5W0P, 6FUF, 6CM0, 6OY9, 6OYA, 6QNO   |
| S1PR1    | 2                    | 3V2W, 3V2Y                                          |                                                  |
| TACR1    | 9                    | 2KS9, 2KSA, 2KSB, 6HLL, 6HLO, 6HLP, 6J20, 6J21, 6E59 |                                                  |
| TBX2A2R  | 2                    | 6IU2, 6IV                                            |                                                  |
| US28     | 4                    | 4XT1, 4XT3, 5WB1, 5WB2                               |                                                  |
| **Class B** |                       |                                                     |                                                  |
| ADCYAP1R1| 4                    | 6PS9Y, 6M1H, 6M1L, 6LBP                              |                                                  |
| CALCRL   | 4                    | 5UZ7, 6NYY                                           |                                                  |
| CALCRL   | 4                    | 6E3Y, 6UVA, 6UUN, 6UUS                               |                                                  |
| CRHR1    | 4                    | 4KSY, 4Z9G                                           |                                                  |
| CRHR2    | 1                    | 6PB0, 6P9X                                           |                                                  |
| GGR01    | 9                    | 4L16R, SEE7, 5XEX, 5XF1L, 5YQZ                        |                                                  |
| GHRHR    | 1                    | 6LMK, 6MLM, 6WHC, 6WPM                              |                                                  |
| GLP-1R   | 12                   | 5N2X, 5V2E, 5VEX, 6KJ7, 6KK2, 6K7K, 6L02             |                                                  |
| GLP-2R   | 12                   | 5V5A, 6B5J, 6ORV, 7C2E, 6VCB                         |                                                  |
| PTH1R    | 4                    | 6FJ3                                                |                                                  |
| SCTR     | 1                    | 6W9G, 6W9I                                          |                                                  |
| VIPR1    | 1                    | 6VNN                                               |                                                  |
| **Class C** |                       |                                                     |                                                  |
| GABBR2   | 8                    | 7C75, 7C7Q, 6U08, 6VJM, 6U0A, 6U09, 6W2X, 6WIV       |                                                  |
| GRM1     | 1                    | 4OR2                                               |                                                  |
| GRM5     | 5                    | 4O09, 5GC5, 5CGD, 6FFH, 6FFI, 6N4X, 6N51, 6N52       |                                                  |
| **Class F** |                       |                                                     |                                                  |
| FZD4     | 1                    | 6BD4                                               |                                                  |
| FZD5     | 1                    | 6WW2                                               |                                                  |
| SMO      | 11                   | 4J7K, 4N4W, 4OR9, 4QIM, 4QIN, 5L7D, 5L7I, 5S56, 5S7, 5O3C, 6OT0 |                                                  |

The structures were updated in September 2020. The PDB codes of GPCR structures determined by cryo-EM are in bold. The structural data were collected from the Protein Data Bank (rcsb.org)²⁹⁴. ADCCYP1R1 is a putative adenylyl cyclase-activating polypeptide type 1 receptor, ADORA1 (A1R) is adenosine A1 receptor, ADRA2B (a2B adrenergic receptor, ADRB1 (B1R) is an adrenergic receptor, AGTR2 (AT2R) is an angiotensin II receptor type 2, APLNR Apelin receptor, CSAR1C5a anaphylatoxin chemotactic receptor 1, CALCRL calcitonin receptor, CCR1-9 C-C chemokine receptor (CCR) type 1–9, CHRM2 (M2R) muscarinic acetylcholine receptor M2, CHRM3 (M3R) muscarinic acetylcholine receptor M3, CHRM4 (M4R) muscarinic acetylcholine receptor M4, CR2 (C2B) cannabinoid receptor 2, CRHR1 (CRF1R) corticotropin-releasing factor receptor 1, CRHR2 (CRF2R) corticotropin-releasing factor receptor 2, CXCR1–2 C-X-C chemokine receptor type 1–2, CYSLTR1–2 cysteinyl leukotriene receptor 1–2, DRD1–4 dopamine receptor, EDRB1 endothelin receptor type B, F2R proteinase-activated receptor 1, F2RL1 proteinase-activated receptor 2, FFAR2 free-fatty-acid receptor 1–2, GABBR2 (GABA_B2) GABA type B receptor subunit 2, GHRHR growth hormone-releasing hormone receptor, GLP-2R glucagon-like peptide-2 receptor, GPBAR1 protein-coupled bile acid receptor, GPR32 G protein-coupled receptor 32, HCNR2 (OxR) orexin receptor type 2, HHR1 histamine H1 receptor, HTR1B 5-hydroxytryptamine receptor 1B, HTR2B (5-HT_2B) 5-hydroxytryptamine receptor 2B, HTR2C (5-HT_2C) 5-hydroxytryptamine receptor 2C, LPAR1 lysophosphatidic acid receptor 1, MC4R melanocortin-4 receptor, MTNR1A melatonin receptor type 1A, MTNR1B melatonin receptor type 1B, NPY1R neuropeptide Y receptor Y1, NTSR1 neurotensin receptor type 1, OPR1 delta-type opioid receptor, OPRM1 nociceptin receptor, OPRK1 (k-OR) kappa-type opioid receptor, OPRM1 (µ-OR) mu-type opioid receptor, OXTR oxytocin receptor, P2RY1 P2Y1 purinoceptor 1, P2RY12 P2Y2 purinoceptor 12, PFAFR platelet-activating factor receptor, PTGDR2 prostaglandin D2 receptor 2, PTGER3 prostaglandin E2 receptor EP3 subtype, PTGER4 prostaglandin E2 receptor EP4 subtype, RHO rhodopsin, SCTR secretin receptor SMO homolog, TBA2R thromboxane A2 receptor, US28 G-protein coupled receptor homolog US28, VIPR1 vasoactive intestinal peptide receptor.
ally, due to lacking CRD in the GABABRs, the relatively shorter stalk
latter only exhibits a negligible conformational change. Addition-
the TMD core of class C GPCRs (Fig. 7e),109
Noteworthy, in contrast to other allosteric modulators that bind to
GB2 subunits52 (Fig. 7d). This novel allosteric binding site may
novel allosteric site at the interface of TMDs in GB1 and
unique allosteric modulation mechanism.113 In fact, two ligand-
d), which predominantly couples to Gi1 heterotrimer. Interestingly,
and TM3/4/5-ICL3 regions at the cytoplasmic part of GB2 (Fig. 7c,
contrast, substantial conformational alterations occur at the stalk
and TM3/4/5-ICL3 regions at the cytoplasmic part of GB2 (Fig. 7c,
d), which predominantly couples to G1 heterotrimer. Interestingly,
cholesterol are observed at the TMD interface of inactive
GABAβRs52 (Fig. 7c), while two chained phospholipids occupy a
cholesterols are observed at the TMD interface of inactive
Class F GPCRs include SMO and 10 FZDs in humans. Besides a
canonical TMD across all classes of GPCRs, class F is characterized
by a large ECD composed of a CRD and an ECD linker domain to
connect with TMD (Fig. 8a, b).112 It was reported that SMO has a
unique allosteric modulation mechanism.114,115 In fact, two ligand-
binding sites have been identified: one in CRD and the other in
TMD (Fig. 8b). SMO is activated by cholesterol via binding to CRD.
The binding of an antagonist to TMD was proposed to trigger its
conformation change thereby propagating to CRD and allosteri-
ically impeding the binding of cholesterol.113 Recent structural
studies reveal that cholesterol and oxysterol that are critical for
SMO activation are located deep within the 7TM domain of SMO
(Fig. 8d, e).114,115 CRD of FZD can interact with lipoglycoprotein
Wnt and Norrin (specific ligand for FZD4) to mediate the Wnt
signaling.116 Structures of CRD in complex with Wnt or Norrin
provided molecular details of how they formed a symmetrical
homodimer (2:2 complex) during ligand recognition (Fig. 8d, e).
In contrast to SMO, the ligand recognition and receptor
activation mechanisms of FZD remains elusive due to the absence
of the full-length FZD structures. So far, only two apo TMD
structures of FZD4 and FZD5 have been reported (Fig. 8f).61,63
Structures of the full-length FZD in a ligand-bound state are
required awaiting to provide mechanistic explanations.

**GPCR PHARMACOLOGY**

The explosion of 3D GPCR structures and computational simula-
tions has revealed the dynamic conformations between inactive,
intermediate, and active states of GPCRs. The detailed structural
information illustrated that cholesterol, ion, lipids, and water also
participate in receptor activation.99,119 The flexibility of receptor-
binding pocket endows the complex pharmacological mechan-
isms of ligand recognition and signal transduction. Biased
signaling, allosteric modulation, and polypharmacology are help-
ing us better understand how GPCRs bind to numerous ligands
and how they transmit diverse signals to elicit physiological
functions.

**Polypharmacology**

Ligand binding to multiple targets leads to antagonism, additive,
so synergism pharmacological responses that could be positive or
negative based on the mechanism of action. The paradigm of one
drug vs. multiple targets has outpaced the time and cost
associated with the conventional therapy.120 Polypharmacology
thus emerges to study acceptable degree of specificity toward
multiple targets, interconnected signaling pathways that result in
clinical benefit or cross-reactivity that may cause adverse
events.121,122 T2DM, obesity, cancer, and Alzheimer’s disease are
major indications for GPCR modulators.4 These polygenic diseases
are not completely treatable by a single agent, while desirable
efficacies may be achieved for certain respiratory conditions,
central nervous system (CNS) disorders, and cardiovascular
diseases through modulators directed against β2AR, DRD2, and AGTR1,
respectively.

It was shown that 5-hydroxytryptamine receptor 2 (5-HT2) binds
to selective inverse (ritanserin) and highly promiscuous
(ergotamine) agonists but the interaction with ergotamine is broad. This feature allows the development of pan serotonin receptor modulators to treat different diseases. For instance, zolmitriptan as an anti-migraine drug is also used for hyperesthesia via binding to off-target site, and lorcaserin (Belviq) is used to treat obesity while its therapeutic potential for depression, schizophrenia, and drug addiction is being investigated.

Fig. 6 Structural features and common activation mechanism of class B GPCRs. a, b Structural features of the peptide-binding pocket. The shift of peptide C-terminus (a) and ECD (b) is indicated as red arrows. The peptides urocortin 1 (UCN1) bound to CRF1R (light blue, PDB code: 6PB0), UCN12 bound to CRF2R (salmon, PDB code: 6PB1), PACAP38 (red, PDB code: 6P9Y), long-acting PTH (LA-PTH, green, PDB code: 6NBF), GLP-1 (cyan, PDB code: 5VAI), sCT (yellow, PDB code: 6NIY), and CGRP (magenta, PDB code: 6PB1) are shown as cartoons. Binding poses of the antagonist (green) and allosteric ligand (salmon) are shown as sticks (c, PDB codes: 4K5Y, 5EE7, 4Z9G, 5VEX, and 5VEX). d, e The common activation mechanism of class B GPCRs as exampled by the structures of inactive GCGR (gray, PDB code 3NYA) and active VIP1R (green, PDB code 6VN7). Side chains of residues in three conserved polar network are shown in stick presentation. The conserved P6.47bXXG6.50b motifs in TM6 are shown as single red spheres.
However, off-target activity, hallucinations, and cardiac valvulopathy related to 5-HT2A and 5-HT2B modulation should be carefully monitored. Atypical antipsychotics are mainly targeting both dopamine and serotonin receptors, usually as antagonist for DRD2 and antagonist or inverse agonist for 5-HT2A. Exemplified by clozapine and aripiprazole, haloperidol, amoxapine, and asenapine display a diverse spectrum of receptor interaction. Additionally, carazolol, a member of aminergic division exerts its effects by interacting with multiple adrenergic receptors as inverse agonist or allosteric antagonist. Istradefylline combined with L-DOPA/dopamine simultaneously target A2AR, DRD1 and DRD2 in animal model of Parkinson’s disease. Amitrytrpyline, a tricyclic compound targeting muscarinic and histamine H1 receptors, is used to treat depression and non-selective muscarinic receptor antagonists are trialed for bladder dysfunction. Lorazepam, indicated for anxiety due to interaction with GABAaR, is also an allosteric modulator of the proton-sensing GPCR (GPR68) and has been repurposed to treat pancreatic cancer. 6'-Guandinonalarindole (6'-GNTI) is an agonist with higher selectivity for δ/κ-opioid receptor heterodimer but not homodimer. Importantly, 6'-GNTI is an analgesic that offers additional benefit. In cardiovascular diseases, β blockers decrease catecholamine-induced heart rate elevation via interaction with valsartan (AT1R-mediated signaling). It is of note that mono-, dual-, and tri-agonists for the glucagon family of receptors (GLP-1R, GCGR, and GIPR) have been developed and trialed for weight loss and glucose control (Table 3). Successful outcome will determine whether unimolecular polypharmacology is a practical approach to translate safety and efficacy of multiple agents into a single molecule.

Biased agonism
Activated GPCRs can recruit multiple transducers (such as heterotrimeric G proteins, GPCR kinases, and β-arrestins) and consequently produce distinct biological responses. Ligands that preferentially engage one signaling pathway over others are regarded as bias and may show improved therapeutic outcomes. Biased signaling that has been applied to drug discovery involve AT2R, µ-OR, κ-OR, β-adrenergic receptors, DRD2, CTR, CCR, and adenosine receptors. µ-OR is the best studied...
receptor for biased agonism. Compounds that stimulate Gq, coupling and cAMP production but not β-arrestin recruitment are preferable to retain analgesia and reduce opioid-related side effects. This G protein bias was also demonstrated with widely used drug tramadol, whose active metabolite, desmetramadol, elicited maximum cAMP production without affecting β-arrestin 2 recruitment compared to fentanyl and morphine. Safety profile is improved with less adverse effect such as respiratory depression.

Another μ-OR-biased ligand, oliceridine (TRV130, Olinvo™), passed phase III clinical trial but did not get the FDA approval for safety concerns. The NDA for oliceridine was resubmitted and a new counterpart, TRV734, is not only suitable for oral administration but also safer due to reduced dependency. A fourth μ-OR-biased ligand, PZM21, cross-reacts with κ-OR and failed to reduce respiratory depression in CS78L and CD-1 mice. Whether this relates to its residual but marked effect on β-arrestin 2 recruitment, as opposed to oliceridine whose action is negligible, remains to be further studied.

Similar situation occurred with κ-OR as well whose agonists possess analgesic property and have a low risk of dependence and abuse but with adverse effects such as sedation, motor dysfunction, hallucination, and dysphoria. G protein-biased agonists of κ-OR, including RB-64, mesyl salvinorin B, triazole 1,1, diphenethylamines and LOR17, were reported to minimize the adverse effects in preclinical settings. One of such, nalfurafine, was approved in Japan (2015) as an anti-pruritic agent for patients with chronic liver diseases.

Carvedilol, known as a β1 and β2 adrenoceptor blocker, was found to be biased toward β-arrestin recruitment, G protein-coupled receptor kinase activation, and ERK1/2 phosphorylation. Joining its rank included alprenolol, bucindolol, and nebivolol, all found to be biased toward the lipidic interface close to the orthosteric site, leading to a hypothesis that G protein-biased agonists may show species preference. For instance, CL316243 was β3PAM-biased, whereas L748337 and SR59230 are ERK/p38 phosphorylation-biased. Another μ-OR-biased ligand, PZM21, cross-reacts with κ-OR and failed to reduce respiratory depression in CS78L and CD-1 mice. Whether this relates to its residual but marked effect on β-arrestin 2 recruitment, as opposed to oliceridine whose action is negligible remains to be further studied.

In contrast to μ-OR, arrestin bias is desirable for AT1R to improve cardiac performance. Nonetheless, clinical development of AT1R modulators either resulted in a phase llb trial failure (TRV027) in 2017 or never reached to clinical stage (SBpa, SVDf, SI, sarmesin, saralasin, and SII). Of note is that biased molecules may show species preference. For instance, CL316243 is more active in mice than in humans, whereas nalfurafine works better in humans vs. rodents. A list of therapeutic agents with biased signaling approved or advanced to clinical trials is shown in Table 5.

### Allosteric modulation

In recent years, studies on allosteric GPCR modulators have gained unprecedented momentum. An allosteric modulator is a ligand binding to a position other than the orthosteric site but can modify responses of a receptor to stimulus. Allosteric modulators that enhance agonist-mediated response are called PAMs, while those attenuate the response are called NAMs. This phenomenon is very common such that the Allosteric Database 2019 (ASD, http://mdl.shsnu.edu.cn/ASD) records 37520 allosteric modulations on 118 GPCR members, covering all four classes.

Allosteric modulation is advantageous in terms of (i) using highly druggable pockets. In some cases, it is easier to design ligands at an allosteric site than the orthosteric site, such as class B GPCRs with orthosteric pockets wide open. For example, both PAMs and NAMs binding to the same position at the TMD of GLP-1R were reported; (ii) improving selectivity. The orthosteric site and cognate ligand are often highly conserved, making it hard to discover very selective orthosteric binders. Meanwhile, non-conserved allosteric sites would be a better choice evidenced by discovery of many subtype selective allosteric modulators of acetylcholine and cannabinoid receptors; (iii) introducing signal bias. Allosteric modulators with biased signaling were developed for prostaglandin F2α receptor and chemokine receptor CXCR4. Although still as an emerging concept, allosteric modulators have exhibited a great potential with some compounds being marketed or in clinical trials.

However, developing allosteric modulators of GPCRs remains challenging—molecules recorded in the ASD largely concentrate on two superfamilies, the mGlurS (8 members, 17,115 modulations), and mACHRs (5 members, 7666 modulations), accounting for nearly 2/3 of the total number. Some individual receptors also contribute a significant proportion, such as CB1 (1948 modulations), GABAβ (1286 modulations), and folliculo-stimulating hormone receptor (1233 modulations). Excluding these “easy cases,” allosteric modulators are few in number. Furthermore, the structural diversity of the allosteric modulators is quite low, for many derivatives would be included soon after a parent compound is identified. The difficulty in developing allosteric modulators is partly due to the limitation of detecting allosteric behavior: Not every newly discovered active compound could be tested for its effect on binding affinity or EC50 of an orthosteric agonist, therefore some allosteric modulators were not correctly identified. For instance, BPTU in P2R1, the first GPCR NAM solved in complex structure (PDB code: 4XNV), was not considered allosteric until the structure was obtained. To make things worse, NAMs may weaken the binding of an endogenous ligand thus behaving like a competitor, such as NDT951 (PDB code: 5O9H).

The most effective way to identify the binding site of an allosteric modulator on a GPCR is solving the complex structure. Crystallography is an effective technique, while rapidly deployment of cryo-EM has started to deliver its promise (PDB codes: 6OIK172 and 6U1N173). To date, 17 GPCRs have reported structures in complex with allosteric modulators. Detailed analysis of complex structures before October 2018 was reported previously, and here we focus on insights provided by newly published results. The most unusual allosteric-binding sites on GPCRs are at the lipidic interface embedded in cell membrane. Five different positions were identified by crystal structures (Fig. 9): UP12, UP34, LOW34, LOW345, and LOW67. Four of them were recently reviewed. The LOW34 site was reported in 2019 for ORG27569 in CB1 (PDB code: 6KQI166; Fig. 10a).

ORG27569 attracted much attention for its distinctive function: increasing the binding of orthosteric agonist CP55940 but making it act as inverse agonist. Many attempts were made to locate the binding site of ORG27569 by mutagenesis but the results are conflicting: one study showed that the effect of ORG27569 on CP55940-induced [35S]GTPγS binding was disturbed by mutations to multiple residues at the orthosteric site, leading to a hypothesis that ORG27569 stays in the same pocket close to CP55940. Another study found that ORG27569 reduced the binding of a fluorescence-labeled orthosteric antagonist, and the effect was only disturbed by mutations at the lipidic interface close to the cytoplasmic end of CB1. Besides, it was reported that the functions of ORG27569 were also affected by breaking a disulfide bond at the N-terminus or by constitutive active/inactive mutations at the cytoplasmic interface.

The crystal structure exhibited that the position of ORG27569 is considerably overlapped with a cholesterol captured in another intermediate state (PDB code: 5XRA179; Fig. 10a), consistent with the site located by the fluorescence-labeled orthosteric antagonist. At this site, the higher selectivity to CB1 over CB2 could be explained. Interestingly, ORG27569 is the only allosteric modulator at lipid interface forming no hydrogen bond to the receptor.

There have been three more complex structures of allosteric modulators at lipidic interface since October 2018, all obtained by crystallography. Two are β2AR, with a NAM AS408 (PDB code: 6OBA186) or a PAM Cmpd-6FA (PDB code: 6N48181). Both allosteric modulators bind to the LOW345 site (Fig. 10b). The NAM stays at a
position very similar to NAMs in C5AR1 (PDB codes: 5O9H, 1716C1R, and 6C1Q182) but the PAM is close to ICL2 and only partially overlaps with PAMs of FFAR1 (PDB codes: 5TZY183 and 5KW2184), showing a complex regulation nature at this site. The other complex structure is full-length GLP-1R with PF-06372222 (PDB code: 6LN2185), a NAM previously used to co-crystallized with GLP-1R TMD (PDB code: 5VEW107).

Even around the position of orthosteric ligands (among the helices and facing extracellular side), another ligand may occupy the space not taken by the endogenous ligand and act as an allosteric modulator. The very abundant PAMs/NAMs of mAChRs function in this mechanism. PAM LY2119620 in M2R (with the orthosteric agonist iperoxo and stabilized by a nanobody) was the first allosteric modulator to obtain complex structure with a class A GPCR (PDB code: 6MQT102). Recently, LY2119620 was also observed in protein complexes of M2R with G protein (PDB code: 6OIK172) or arrestin (PDB code: 6U1N173) by cryo-EM.

CCR5 is a chemokine receptor and an important anti-HIV drug target. A marketed inhibitor, maraviroc, has long been recognized as a NAM of CCR5. There were hypotheses that small molecule NAMs, chemokine, and the HIV-binding protein have separate binding sites.186 However, structures of CCR5 in complex with maraviroc (PDB code: 4MBS187), chemokine analog antagonist (PDB code: 5UIW188), or HIV envelope glycoprotein (PDB code: 6MEO189) show that these ligands highly overlap in CCR5 pocket (Fig. 10c). Therefore, the noncompetitive behavior of maraviroc may be due to a very extensive interface of peptidic CCR5 agonist, thus a small molecule cannot diminish the binding even with this much collision. The results illustrate that allosteric behavior is not equal to totally separated binding positions,

### Table 5. Therapeutic agents with biased signaling approved or in clinical trials

| Ligand       | Receptor (GPCR class)                  | Signaling bias | Indication                                                                 | Development status | Reference |
|--------------|----------------------------------------|----------------|----------------------------------------------------------------------------|-------------------|-----------|
| Bromocriptine | Serotonin receptors, adenosine receptors, dopamine receptors (class A) | At 5HT2, Gq/11, at DRD2 β-arrestin | Acromegaly, Parkinson’s disease, T2DM, idiopathic hyperprolactinemic disorder, neuroleptic malignant syndrome | Approved          | 297–299   |
| Pergolide    | SHT2 (class A)                         | Gq/11          | Parkinson’s disease                                                        | Approved          | 297,296   |
| Ergotamine   | HTR2B (class A)                        | β-arrestin     | Migraine                                                                  | Approved          | 297,300   |
| Atropine     | CHRM3 (class A)                        | Low efficacy agonist for G1b, inverse agonist for Gq, antagonist for G10 | Organophosphorous poison antiodote                                         | Approved          | 297,301   |
| Pilocarpine  | CHRM3 (class A)                        | β-arrestin and pERK1/2 | Xerostomia                                                               | Approved          | 297,302   |
| Capadenoson  | ADRA1A (class A)                       | cAMP           | Atrial fibrillation                                                        | Phase 2           | 297,303   |
| Alpenolol, bucindolol, carvedilol, nebivolol | ADRA1B and ADRA2 (class A) | β-arrestin | Congestive heart failure                                                   | Approved          | 148,304   |
| Isoetharine  | ADRB8 (class A)                        | β-arrestin     | Asthma                                                                    | Approved          | 304,305   |
| Dihydrdixine (DAR-0100A) | DRD2 (class A) | Full agonists for G1b, but partial agonists for Gq signaling | Schizotypal personality disorder                                           | Phase 2           | 297,306   |
| Bifeprunox   | DRD2 (class A)                         | Kinetic bias   | Bipolar disorder, depression, schizophrenia, psychosis                    | Phase 3           | 297,307   |
| Aripiprazole | DRD2 (class A)                         | Kinetic bias   | Psychosis                                                                 | Approved          | 297,307   |
| TRV027 (TRV120027) | AGTR1 (class A) | β-arrestin | Anti-hypertensive with cardio-protection                                  | Phase 2           | 5,304     |
| TRV250       | OPRD1 (class A)                        | G protein      | Migraine                                                                  | Phase 1           | 5,137     |
| Nalfurafine  | OPRK1 (class A)                        | G protein      | Pruritus                                                                  | Approved          | 157       |
| Tramadol     | OPRM1 (class A)                        | G protein      | Pain                                                                      | Approved          | 140       |
| Oliceridine  | OPRM1 (class A)                        | G protein      | Pain                                                                      | Phase 3           | 5,137     |
| TRV734       | OPRM1 (class A)                        | G protein      | Pain                                                                      | Phase 1           | 5,137     |
| Cyt-1010     | OPRM1 (class A)                        | G protein      | Pain                                                                      | Phase 1           | 308       |
| Satavaptan   | AVPR2 (class A)                        | β-arrestin (partial agonist while inverse agonist at G1) | Hyponatremia and ascites                                                  | Phase 3           | 297,309   |
| Atosiban     | OXTR (class A)                         | G1 and G3      | Delaying imminent preterm birth                                           | Approved          | 297,310   |
| BMS-986104   | S1PR1 (class A)                        | cAMP           | Rheumatoid arthritis                                                      | Phase 1           | 297,311   |
| Ly2828360    | CNR2 (class A)                         | G/ERK          | Knee osteoarthritis                                                       | Phase 2           | 297,312   |
| MK-0354      | HCAR2 (class A)                        | G protein      | Dyslipidemia                                                              | Phase 2           | 297,313   |
| TRV027 (TRV120027) | AGTR1 (class A)     | β-arrestin     | Anti-hypertensive with cardio-protection                                  | Phase 2           | 5,304     |
| Exenatide    | GLP-1R (class B)                       | β-arrestin     | T2DM                                                                      | Approved          | 12        |
| TTP273       | GLP-1R (class B)                       | G protein      | T2DM                                                                      | Phase 2           | 190       |

Receptor abbreviations are according to IUPHAR. The list is derived from earlier reports5,137,304,308 with addition based on literature research. The indication and development status are updated from DrugBank.ca and ClinicalTrials.gov database.

ADRA1A alpha-1A adrenergic receptor, AVPR2 (V2R) arginine vasopressin receptor 2, HCAR2 hydroxycarboxylic acid receptor 2.
because partially overlapped sites with different key interactions are also allowed.

The last case of allosteric modulator in extracellular pocket is PAM TT-OAD2 of GLP-1R (PDB code: 6ORV). This small molecule agonist only slightly collides with the endogenous peptide (PDB code: 5VAI, Fig. 10d), consistent with its behavior that only partially displaces an orthosteric probe.190

Fig. 9  Schematic diagram of allosteric sites at the lipidic surface identified by complex structures. The binding sites are manually labeled on the crystal structure of β2AR (PDB code: 6OBA). Solid line, allosteric site at front side; dashed line, allosteric site at back side. UP, upper part aka close to the extracellular end; LOW, lower part aka close to the cytoplasmic end; numbers, main interacting transmembrane helices

Fig. 10  Binding sites of allosteric modulators in GPCRs reported after October 2018, in comparison with related ligands. a NAM ORG27569 in CB1 (PDB code: 6KQI) in comparison with cholesterol (PDB code: 5XRA); b NAM AS408 (PDB code: 6OBA) and PAM Cmpd-6FA (PDB code: 6N48) in β2AR, in comparison with NDT9513727 in C5AR1 (PDB code: 6CTQ) and PAM AP8 (PDB code: 5TZY); c NAM maraviroc in CCR5 (PDB code: 4MBS) in comparison with chemokine analog antagonist [5P7]CCL5 (PDB code: SUIW) and HIV envelope glycoprotein gp120 (PDB code: 6MEO); d PAM TT-OAD2 in GLP-1R (PDB code: 6ORV) in comparison with GLP-1 (PDB code: 5VAI)
proximately share the same binding site (TM1, TM2, TM6, TM7, ICL1, and H8). Their binding position does not overlap with Gq, therefore they may stabilize the inactive state by blocking conformational changes required for receptor activation. This site is generally non-conserved in the GPCR superfamily, thus targeting here may provide some selectivity. Additionally, many nanobodies at the cytoplasmic interface were also developed for several receptors, including AGTR1 (PDB codes: 6D01196 and 6OSO197), β2AR (PDB code: 6B1B198), β2AR (PDB code: 6N48199), and SMO (PDB code: 6O3C198, for information before October 2018, see review161).

Multi-domain regulation is an interesting topic in allosteric modulator discovery. Class C GPCRs use EC2Ds to recognize their cognate ligands, leaving the classic pocket of TMD for allosteric modulating. However, this is the major reason why this class has a large number of allosteric modulators. In the case of mGLURs, both PAMs and NAMs have been widely reported, but only NAMs obtained complex structures—there is no solved active state structure. The full-length structures of mGLU5 (PDB codes: 6NS1 and 6NS2200) displayed how the binding of orthosteric agonist to ECD triggers the change of interaction between two monomers, but the conformational change of TMD remains elusive.

SMO in class F is also a multi-domain receptor. The first reported ligand of SMO cycloamine (an antagonist causing birth defects) binds to the classic TMD pocket (PDB code: 4O9R201) shared by several other antagonists with different chemical scaffolds and an agonist (SAG).202–204 ALLO-1, an antagonist identified as allosteric modulator not competitive to cycloamine, was recently found to bind at a deeper position in the pocket by photo-affinity labeling combined with mass spectrometry (MS).205 SMO has another pocket in ECD that interacts with steroids, including cholesterol (PDB codes: 5L7D203 and 6D35204). Since cholesterol has been the most favored candidate of SMO endogenous ligand, the ECD pocket is treated as orthosteric making the TMD pocket allosteric. However, newly obtained structures demonstrated that cholesterol or its analog can also bind to TMD pocket (PDB codes: 6O3C206 and 6OT0207), leaving the question open for which is the true orthosteric site.

Disease indication

GPCRs are involved in many human diseases and specific drug intervention is one of the most celebrating achievements in the pharmaceutical industry (Table 5 and Fig. 51). Among all available drugs targeting GPCRs, HRT1, DRD2, M1R, and ADRA1A are the most frequently addressed for indications such as hypertension, allergy, pain, and schizophrenia, and 33% of them have >1 indication with an overall average of 1.5. Although CNS diseases are still popular accounting for 26% of all approved indications, development focuses have now been shifted to T2DM, obesity, multiple sclerosis, smoking cessation, short bowel syndrome, and hypocalcemia. Repurposing of existing drugs for new indications also emerged to supplement discovery efforts.

**STRUCTURE-BASED DRUG DESIGN**

As two general types of computer-aided drug design techniques (Table 6),208 SBDD and ligand-based drug design, exploit the structural information of protein targets and the knowledge of known ligands, respectively (Table 6). SBDD, on the basis of crystal/cryo-EM/NMR structures or homology models, first identifies key sites and important interactions responsible for target functions, then screens large virtual library/designated agents that disrupt or enhance such interactions to modulate relevant biological processes and/or signaling pathways by molecular docking, and finally discovers active leads with desired pharmacological properties.

Clearly, the past decade is a golden age for SBDD on GPCR. With the year of 2011 (when LPC crystallization209 fusion proteins, and other key techniques collaborated to launch the outbreak of GPCR structure determination including the landmark β2AR-Gαs for watershed207), SBDD of GPCR evolves two distinct stages: rhodopsin-based homology model and truly authentic structure of individual receptors. Boosted by the fast-increasing number of high-quality GPCR structures, improved accuracy of combinational computational approaches, and better understanding of activation mechanism and pharmacology, SBDD is developing rapidly with fruitful scientific reports and increasing GPCR-targeted drugs contributed by this approach. Considering the length of time required for a drug to be available on the market (10–15 years) and the chance of applying structural biology information to hit discovery and lead optimization in the first 2–3 years of a drug discovery program, it is probably too early to see the approval of GPCR-targeted drugs being developed with the aid of a structure, and such situation is likely to change as the tremendous efforts from both academia and industry start to bear the fruits of successes. The following is a brief account of recent advances in three main aspects of SBDD (chemical space, receptor dynamics, and pose evaluation) in the context of their application in GPCR pharmacology.

Optimized virtual library

Despite the vast chemical space (>10^12 drug-like molecules), only a nominal fraction has been explored by SBDD, where both the compound availability and insufficient diversity limited the number of screened ligands. To overcome these problems, ultra-large208–210 and focused libraries211–213 were employed. Lyu and colleagues206 presented an excellent model of “bigger is better” in virtual drug screening. Based on the 130 well-characterized reactions, they generated 170 million make-on-demand compounds (http://zinc15.docking.org/), the resulting library is remarkably diverse with >10.7 million scaffolds unavailable before. By docking 138 million molecules against DRD4, they discovered 81 new chemotypes (24% hit rate), 30 of them showed submicromolar activity, including a 180-pM subtype-selective and Gβi-biased DRD4 agonist. This ultra-large library docking study provides important information: (i) hit rate fell almost monotonically with docking score; (ii) hit rate vs. score curve of DRD4 predicted that 1 from every 873 compounds may have a minimum affinity of 1 μM; and (iii) human visual evaluation improved the selected compound with higher affinities, efficacies, and potencies but not the hit rate. A follow-up study on MT1 by docking >150 million “lead-like” molecules209 identified 15 active leads (39% hit rate) with potencies ranging from 470 pM to 6 μM. Alternatively, focused libraries identified scaffolding library, natural products, and screening tests. Chemical (i.e., chemicals that have never shown bioactivity in virtual screening (VS) for dozens of receptors. Focused on compound library of traditional Chinese medicine (TCM), Liu et al. found that salvianolic acids A and C antagonized the activity of both P2RY1 and P2RY12 purinoceptors in the low μM range, while salvianolic acid B antagonized the P2RY12 purinoceptor. Remarkably, these three salvianolic acids are major active components of the broadly used hemorheologic TCM Danshen (Salvia miltiorrhiza). Taking NT5R1 as an example, Ranganathan et al. found that the fragment library tended to have higher hit rate than that of the lead-like library (19%) but the affinities were ~100-fold weaker. Collectively, these results demonstrate the importance and advantages of ultra-large and tailored libraries in discovering potent GPCR modulators.

Receptor dynamics

Emerging evidence from crystallography, spectroscopy, and molecular dynamics (MD) simulations have demonstrated the crucial roles of GPCR dynamics involved in ligand recognition,
receptor activation, and allosteric modulation.\textsuperscript{216–218} To consider the protein flexibility during GPCR-related SBDD, many computational approaches\textsuperscript{17} including rotamer sampling, induced-fit docking, and ensemble docking have been employed showing a great promise, especially in the search of biased, bi-topic, or allosteric modulators. During ensemble docking,\textsuperscript{212,217,219–221} ligands are docked into multiple structures representing different possible conformational states rather than a single structure, where the targets could be multiple crystal structures or extracted from MD/Monte Carlo (MC) simulations or normal mode analysis (NMA). By evaluating the known ligand enrichment, as well as selectivity for agonists or antagonists on seven GPCR/ligand co-structures, Coudrat et al. found that small variations in structural features are responsible for their success in VS, while a combination of ligand/receptor interaction patterns and predicted interaction strength is associated with the predictive power of binding pockets in VS.\textsuperscript{220} Compared to the Glide VS workflow, the combination of accelerated MD simulations and Glide induced fit docking of M2R by Miao et al. provided much-improved enrichment factors and identified four PAMs and one NAM with unprecedented chemical diversity.\textsuperscript{221} For S-HT\textsubscript{4A}, whose crystal structure is not available, Warszycki et al. applied MC and NMA to generate an ensemble of binding pockets with the input of a homology template and known active compounds and finally discovered two new active ligands through VS.\textsuperscript{222}

Pose evaluation

Correctly selecting and ranking poses of docked compounds in the ligand-binding pockets have been a challenge for SBDD, especially for GPCR that is embedded in the cell membrane with significant conformational adaptability. To address this problem, many physics-based scoring functions\textsuperscript{223–226} integrated with some user-friendly computer programs (e.g., Dock, GOLD, AutoDock, Glide, and rDock) were routinely adopted in SBDD. Recently, precise computational approaches including free energy calculation methods like molecular mechanics/Poisson–Boltzmann surface area (MM/PB(G)SA),\textsuperscript{227,228} free energy perturbation (FEP),\textsuperscript{229} quantum mechanical/MM calculations,\textsuperscript{230} and fragment molecular orbital\textsuperscript{231,232} have been employed with improved performance. Compared to the empirical scoring functions, MM/PB(G)SA and FEP are physically more rigorous free-energy calculation methods with an increased computational cost and have been adopted in the studies of DNA–ligand, protein–ligand, and protein–protein interactions.\textsuperscript{233,234} By introducing of the minimization-based MM/GBSA refining and rescoring of docked poses, Zhou et al. identified seven 5-HT\textsubscript{2B} antagonists with novel chemical scaffolds and the most potent one has an IC\textsubscript{50} of 27.3 nM in a cellular assay.\textsuperscript{227} Lenselink et al. used FEP to design A\textsubscript{2A}R antagonists and identified a highly potent molecule with K\textsubscript{i} of 1.2 nM.\textsuperscript{232} However, computational investigation across 20 class A crystal structures and 934 known ligands demonstrated that the correlations between predicted binding free energy by MM/PBSA and experimental data varied significantly. The observed variations exist between individual receptors and are highly system specific,\textsuperscript{235} indicating that successful application of MM/PBSA may require additional efforts in validation of experimental data and optimization of simulation/calculation parameters. Alternatively, protein–ligand interaction fingerprints exacted from available crystal structures\textsuperscript{225,236} fueled docking score with protein–ligand-binding mode information and resulted in improved VS virtual hit rates for β\textsubscript{2}AR (53%) and HRH1 (73%) with up to nM affinities and potencies.\textsuperscript{225,236}

Collectively, innovation in VS and pose evaluation, along with evolution of computational hardware, has significantly advanced SBDD and is expected to lift the discovery efficiency to a new height, since GPCRs have multiple downstream signaling pathways responsible for distinct functions or consequences. The high degree of sequence and pocket similarities between different subtypes demands for novel ligands with superior specificity and selectivity. In this regard, allosteric and biased modulators may offer additional pharmacological benefits.

Subtype selectivity

It is known that GPCR subtypes share high sequence identities in orthosteric sites with distinct distribution and downstream signaling profiles. Cross-reactivity among subtypes could cause undesired side effects. For example, five MR subtypes display different G protein coupling features (G\textsubscript{q/11}, M1R, M3R and M5R; G\textsubscript{i/o}, M2R and M4R) and organ distribution (CNS, M1R; peripheral tissues such as heart and colon, M2R), while their sequence identity (64–82%) and similarity (82–92%) in TMD are quite conserved.\textsuperscript{172} Similar observations were seen among dopamine receptors (DRD1 to DRD5), histamine receptors (HRH1 to HRH4), and adenosine receptors (A1AR, A2A/R, A2B/R, and A3AR). Guided by structural information, rational design of subtype selective compounds progresses steadily. Using an extended DRD2-specific binding pocket from the haloperidol-bound DRD2 crystal structure, Fan et al. discovered two highly selective DRD2 antagonists (O4SE6 and O8LE6) that specifically activate DRD2 (EC\textsubscript{50} = 1 μM) after screening of 320 non-olfactory GPCRs.\textsuperscript{237} Through VS of 3.1 million molecules against M2 and M3, Kruse et al. identified a partial M3 agonist without measurable M2 agonism, capable of stimulating insulin release from a mouse β-cell line.\textsuperscript{238} Wei et al. reported a multistage VS of the ChemDiv library (1,492,362 compounds) toward A1AR and discovered four novel antagonists with good affinity and selectively (>100-fold) over A3AR.\textsuperscript{239}

| Table 6. Comparison of the advantages and disadvantages of various computer-aided drug design approaches |
|-----------------|-----------------|-----------------|
| **Approach**    | **Advantage**    | **Disadvantage** |
| Ligand-based drug design (LBDD) | Understanding interactions between functional groups, convenient, does not require the structural information of a target | Descriptor selection, false correlations, enough known ligands |
| Structure-based drug design (SBDD) | Effective model, convenient, does not require the structural information of a target | Known ligands, less novelty; missing conformations |
| Pharmacophore modeling | Novel scaffolds and chemotypes, higher chances of finding potent ligands | Substantial computational resources, compound synthesis |
| Virtual screening of ultra-large libraries | Less demand for computational resources, compound easy to purchase, specific scaffolds or origins | Reduced diversity, less novelty |
| Virtual screening of focused libraries | Considering protein flexibility, improved enrichment factors, rescue of false-negative ligands | Increased computational burden, pose evaluation, false positive |
| Ensemble docking | Improved scoring and ranking ability | Substantial computational burden, method validation target dependency |
| Energy-based pose evaluation | | |

Signal Transduction and Targeted Therapy (2021) 6:7
Biased signaling
Recently, Suomivuori et al. performed extensive MD simulations to identify two major signaling conformations that couple effectively to arrestin or G protein, respectively. They then designed ligands via minor chemical modification resulting in strong arrestin-biased or G<sub>i</sub>-biased signal transduction. 240 Meanwhile, McCorry and colleagues discovered that specific ECL2–ligand contacts are associated with β-arrestin recruitment, whereas blockage of TM5 interaction reduces the G<sub>i/o</sub> signaling. An arrestin-biased DRD2 modulator was thus made exhibiting a calculated bias factor of 20 relative to quinpirole. 241 In addition, Mannel et al. conducted a VS of a tailored virtual library bearing 2,3-dichlorophenylpiperazine for DRD2 and found that 18 compounds occupy both orthosteric and allosteric sites, and 4 of them stimulated β-arrestin recruitment (EC<sub>50</sub> = 320 nM, E<sub>max</sub> = 16%) without detectable G protein signaling. 242

Allosterism
In the past 5 years, an increasing number of receptor–allosteric modulator complex structures revealed diversified positions of allosteric sites and a variety of binding modes, thereby deepening our understanding of allosteric modulation in terms of underlying mechanisms and structural basis. Further to conventional orthosteric ligands, allosteric modulators affect receptor function in different ways. While PAM may enhance maximal efficacy, NAM could reduce agonist signaling strength. 243, 244, 245 It was reported that a PAM of M2R is located above the orthosteric site and interacts with ECL5. Korczyńska et al. screened 4.6 million molecules against the allosteric sites of M2R and identified a PAM that potentiated the action of antagonist N-methyl scopolamine (NMS). Subsequent optimization led to a subtype-selective compound 628 that increased NMS binding with a co-operativity factor of 5.5 and a K<sub>a</sub> of 1.1 μM. 246 Alternatively, Lückmann et al. carried out MD simulations of agonist-removed FFAR1 and found that closure of a potential allosteric site is associated with agonist binding—compounds that bind to this site to prevent the closure functions as allosteric agonists. 247 Obviously, aided by >400 structures from 82 receptors, SBDD is now entering into a new era with substantial knowledge of GPCR signaling 248–249 and drug candidate attributes. 250

NOVEL SCREENING TECHNOLOGY
AsGPCRs represent the most prominent family of therapeutic targets, 251 innumerable efforts have been made in both industry and academia to screen for novel ligands that can modulate the activity of a specified GPCR and serve as lead compounds for drug development.

A diverse array of experimental technologies suitable for assaying protein–ligand interactions have been directly applied or tailored to GPCR-targeted ligand screening, and they can be classified into three main categories: binding-based, stability-based, and cell signaling-based assays (Table 7). Binding-based assays monitor the physical interactions between a GPCR protein typically in a purified recombinant form with individual test compounds. Cell signaling-based assays measure downstream effectors (e.g., cAMP, Ca<sup>2+</sup>, IP1/IP3) of specific intracellular signaling pathways known to be mediated by GPCR, which reflect the functional outcome of ligand binding to the receptor. Stability-based assays assess the variation of thermal stability for a purified protein when treated by test compounds. These different techniques vary in the ligand screening throughput and binding characteristics (Table 7). In the lead discovery stage, both binding- and signaling/activity-based assays are implemented in a parallel or sequential manner, as the multipronged use of complementary techniques would reduce the overall false-positive and false-negative rates. 252

Here we summarize about 20 experimental screening technologies adapted to GPCR ligand discovery (Table 7) and highlight the most recent development of binding-based approaches. Notably, an update of assays assessing GPCR activation and signaling has been provided in a previous review 247 and will not be elaborated here. The structure-based VS is covered in the above section. Structural elucidation technologies (e.g., X-ray crystallography, single-particle cryo-EM, NMR, and HDX-MS) not suitable for high-throughput screening (HTS) are also excluded.

DNA-encoded library (DEL)
Impressive technological advances have been made for binding-based ligand screening over the past decade. Specifically, DEL has emerged as a powerful approach to drug discovery. 248–251 Created by split and pool synthesis, DEL usually contains hundreds of thousands to billions of distinct small molecule–DNA conjugates. A majority of DEL-based HTS reported to date involve incubation of an immobilized target protein with the library before the protein–ligand complexes are isolated. Encoding DNA tags associated with the immobilized target are then amplified and sequenced to assign relevant chemical structures. 250, 251 Although DEL was predominantly applied to ligand screening against soluble proteins such as enzymes, successful adaptation of this technique to GPCRs was reported in a few cases.252–255 Lefkowitz group reported the discovery of a NAM for β<sub>2</sub>AR by screening a DEL of 190 million. 255 This NAM not only has a unique chemotype but also exhibits low μM affinity and inhibits cAMP production as well as β-arrestin recruitment. Later on, the same team discovered the first small molecule PAM for β<sub>2</sub>AR through HTS of >500 million DEL compounds. 253 Both NAM and PAM demonstrated high selectivity. Of note is that the NAM was found using unliganded β<sub>2</sub>AR, whereas the PAM was unmasked via intentional application of β<sub>2</sub>AR with its orthosteric site occupied by an agonist thereby shifting the receptor to the active state. 252, 253 These two studies elegantly demonstrated a proof-of-concept strategy for binding-based screening of allosteric modulators targeting different conformational states. 253

The power of DEL in the discovery of allosteric GPCR modulators was further demonstrated for PAR2. 254 Screening a billion-size library with a thermostabilized PAR2 mutant resulted in the identification of several agonists and antagonists, and some of them bind to an allosteric pocket in the TMD of PAR2. A similar approach was used to discover tachykinin receptor neurokinin-3 (NK3) antagonists of low nM potency involving NK3 overexpressing cells and a library containing tens of millions DNA-encoded compounds. 255 Clearly, DEL-based ligand screening against GPCRs and other integral membrane proteins offers great promises as it circumvents difficulties in receptor purification.

Affinity selection MS
Due to the high sensitivity and high selectivity of modern MS for both protein and small molecule analysis, versatile MS-based technologies have been developed in the past two decades for screening ligands of a given protein target or characterization of ligand-binding properties (Table 7). Almost all of them were originally developed for measuring ligand interactions with soluble proteins, 256–258 and recently they have been adapted to more challenging GPCR drug discovery. The majority of MS-based technologies (e.g., automated ligand identification system (ALIS), ultrafiltration–liquid chromatography/MS, frontal affinity chromatography–MS, membrane-based affinity MS, and competitive MS binding) employ a methodology very similar to DEL as they all capture and detect ligands physically associated with a given GPCR except that native MS analyzes the entire ligand-bound receptor complexes. 269, 269 In general, these methods have several advantages over ligand binding or cell signaling assays: (i) unbiased and direct detection of ligand–receptor binding facilitates the identification of both orthosteric and allosteric modulators; (ii) confirmation of ligand identity with accurate mass
measurement; (iii) no chemical labeling or DNA encoding of test compounds; and (iv) quantitative MS analysis enables ranking of ligand affinity or evaluation of binding characteristics.

ALIS is currently the most prevailing MS-based technique employed in pharmaceutical companies for HTS of large-scale synthetic compound libraries. This system integrates size exclusion chromatography for isolating protein–ligand complexes and reverse-phase chromatography for dissociating bound ligands, which are then identified by high-resolution MS. Not surprisingly, the application of ALIS to ligand screening for GPCRs substantially lagged behind soluble proteins due to difficulties in obtaining membrane receptors of sufficient purity and stability. The earliest application was ligand screening for M2R, in which purified M2R was incubated with a 1500-compound pool in each round of affinity selection. After screening a total of 350,000 compounds, one orthosteric antagonist and one allosteric modulator were identified for AChR. Later on, a similar strategy was implemented to screen ligands for CXCR4 using two libraries comprised of 48,000 and 2.75 million compounds, respectively. Each reaction consumed 250 ng purified receptor incubated with a pool of 100 or 2000 compounds. Out of the 362 primary hits, 34 were subsequently confirmed to be new antagonists.

Membrane-based affinity MS developed by Shui’s group enables ligand screening toward wild-type active GPCRs embedded in the cell membrane. It features isolation of membrane fractions from cells expressing a GPCR at high yield and incubation of the cell membrane with a compound cocktail, thus keeping the receptor in its native conformation and eliminating the need of protein purification. Compounds associated with the receptor were then released and subjected to high-resolution MS for structural

| Category                     | Method                        | Assay principle                                                                 | Readout                  | Activity characterization | Throughput     |
|------------------------------|-------------------------------|--------------------------------------------------------------------------------|--------------------------|--------------------------|-----------------|
| Binding-based assay          | Radiolabeled ligand binding   | Detect binding of a radioisotope-labeled ligand to a target in competition with a test compound | Radioactivity            | Kᵦ, Kᵦ⁻¹, Kᵦ⁻²          | Medium          |
|                              | DEL DNA-encoded compounds bound to a target are affinity selected and their structures revealed by DNA sequencing | DNA sequence              | Affinity ranking         | Ultra-high              |                |
|                              | SPR Detect changes in the refractive index of the gold film surface when ligands interact with a target immobilized on the chip surface | Refractive index (mass on surface) | Kᵦ, stoichiometry (n), Kᵦ⁻¹, Kᵦ⁻² | Medium          |                |
|                              | MST Detect directed movement of molecules through a temperature gradient using covalently attached or intrinsic fluorophores | Fluorescence intensity    | Kᵦ, stoichiometry (n)    | Medium          |                |
|                              | TR-FRET Detect fluorescence resonance energy transfer caused by interaction between target and ligand labeled with specific fluorophores | Fluorescence intensity    | Kᵦ, stoichiometry (n)    | Medium          |                |
| Membrane-based affinity MS   | Target–ligand complexes are isolated by fast SEC, and dissociated ligands are identified by high-res MS | m/z and MS intensity      | Affinity ranking, Kᵦ, ACE₅₀ | High           |                |
|                              | UF-LC/MS Target–ligand complexes are separated from solution by filtration prior to ligand identification by high-res MS | m/z and MS intensity      | Affinity ranking         | High           |                |
|                              | FAC-MS Detect ligands flowing through a protein-immobilized column based on breakthrough curves determined by MS | m/z and MS intensity      | Affinity ranking, Kᵦ, ACE₅₀ | Medium          |                |
|                              | Competitive MS binding        | Detect binding of a non-radioactive ligand to a target in competition with a test compound by MRM-based MS analysis | m/z and MS intensity      | Kᵦ, Kᵦ⁻¹, Kᵦ⁻²          | Low            |
|                              | Native MS Detect intact protein–ligand complexes in the gas phase by MS | m/z and MS intensity      | Kᵦ, stoichiometry (n)    | Low           |                |
| Stability-based assay        | DSF Detect changes in protein fluorescence over a temperature gradient | Fluorescence intensity    | Tₘ                        | Medium          |                |
|                              | DLS Measure changes in the protein aggregate size based on static light scattering properties over a temperature gradient | Light scattering intensity | T₉₉₉              | Medium          |                |
| Cell signaling assay         | GTP/₅S Detect ³²S-GTP/₅S binding to GPCR-expressing cell membranes as a result of receptor activation | Radioactivity             | EC₅₀, IC₅₀              | High           |                |
|                              | cAMP Detect cellular levels of cAMP coupled to Gₓ₅₁ or Gₓ₅₁ activation | Luminescence/fluorescence | EC₅₀, IC₅₀              | High           |                |
|                              | Ca²⁺ Detect cellular levels of free Ca²⁺ coupled to Gₓ₅₁/₁₁ or Gₓ₅₁/₁₆ activation | Fluorescence              | EC₅₀, IC₅₀              | High           |                |
|                              | IP3/IP1 Detect cellular levels of IP3 coupled to Gₓ₅₁ or Gₓ₅₁ activation | Fluorescence              | EC₅₀, IC₅₀              | High           |                |
|                              | Luciferase reporter gene β-arrestin recruitment Detect cellular levels of β-arrestin along with receptor endocytosis | Luminescence              | EC₅₀, IC₅₀              | Medium–High    |                |

DEL DNA-encoded library, SPR surface plasmon resonance, MST microscale thermophoresis, ALIS automated ligand identification system, UF ultrafiltration, FAC frontal affinity chromatography, DSF differential scanning fluorimetry, DLS dynamim light scattering.
Each incubation consumed about 2 µg membrane-embedded GPCR protein with a pool of 480 compounds. Primary hits were selected based on the binding index (BI) derived from quantitative MS signals used to distinguish putative ligands from non-specific binders (Fig. 11b). Screening a small compound library with this approach led to the discovery of an antagonist for the 5-HT2C receptor and four PAMs for GLP-1R that are not reported previously266 (Fig. 11c).

More recently, the same team devised another affinity MS strategy that enabled screening of 20,000 compounds in one pool.275 Specifically, they modified the workflow by performing iterative rounds of affinity selection for compounds associated...
with $A_2A$R. Similar to the previously described single-round affinity MS screening assay, quantitative measurement of BI renders detection of high-affinity ligands in this experiment. By comparing the selection of 16 benchmark $A_2A$R ligands from screening compound pools of 480-mix, 2400-mix, 4800-mix, and 20k-mix, they demonstrated that this accelerated affinity MS screening approach, using either the purified receptor or receptor-expressing cell membranes, allowed detection of most high-affinity $A_2A$R ligands ($K_i < 5 \mu M$) and significant reduction of protein consumption and MS instrument time. 275 Three new antagonists for $A_2A$R were identified as a result. It is likely that the throughput of this method could be further increased to assay close to or above 1 million compounds in one pool. 275

The affinity MS technique has been widely employed to fish out and identify putative ligands toward various enzyme targets from complex extracts of natural products, which could promote lead discovery from TCM. 276–281 Indeed, this technique was successfully extended to GPCR ligand screening from herbal extracts. It involved the optimization of receptor construct and integration of affinity MS with metabolomics data mining workflow for sensitive and accurate ligand identification 282 (Fig. 11d). After screening a panel of herbal extracts, a naturally occurring aroporphine compound (1857) displaying strong subtype selectivity for 5-HT$_{2C}$ without affecting 5-HT$_{2A}$ or 5-HT$_{2B}$ was discovered (Fig. 11e–g). Moreover, this new lead exhibited exclusive bias toward G protein signaling and showed in vivo efficacy for food intake suppression and weight loss. 283

Although not directly applied to GPCRs, a previously reported cell-based assay vascular endothelial growth factor receptor 2 (VEGFR2) is interesting. 283 It used a special one-bead-one-compound library of peptoids and cells expressing VEGFR2. Beads bound to the color-coded VEGFR2-expressing cells were selected under fluorescence microscopy and the attached ligands decoded by tandem MS analysis. Hits with low $K_i$ affinity to the soluble VEGFR2 ectodomain were identified subsequently. We envision that these membrane-based or cell-based screening platforms will make a major impact on GPCR drug discovery, especially when they are fully integrated.

**Competitive MS binding assay** employs a non-radioactive ligand to compete the binding of a test compound to a protein target. It resembles radioligand-binding assays but avoids the use of radioisotope. 284–286 When assaying, the marker ligand liberated from the target is measured by a multiple reaction monitoring-based MS method of high sensitivity and selectivity for compound detection. Not only is the method useful for a number of transporters and ion channels, 287–289 this approach is equally effective in addressing GPCRs as recently exemplified on $A_2A$/AR/A$\beta$R and DPD1/2/5,287,288,289,290,291 It was shown that unlabelled marker compounds could substitute their radiolabelled counterparts in all types of ligand-binding characterization studies, including saturation, displacement, dissociation, and competitive association, yielding results in excellent accordance with classic radioligand-binding assays. 287,290

**EMERGING OPPORTUNITIES AND PROSPECTS**

Recent scientific and technological advancements in GPCR biology have provided an enormous amount of information that will benefit our current and future efforts in rational drug design. Integration and refinement of massive data by artificial intelligence is a clear direction to guide both virtual and experimental screening of efficacious therapeutic agents with new scaffolds and of novel chemotypes for all classes of GPCRs.

However, as described in this review, factors that influence GPCR drug discovery include, but not limited to, therapeutic target, chemical diversity, mechanism of signaling, ligand-binding site, mode of action, clinical indication, polypharmacology, etc. Future opportunities may arise from: (i) de-orphanization of orphan GPCRs to provide novel targets; (ii) new indication for drug intervention via discovery and/or repurposing efforts; (iii) development of lead compounds targeting classes B and F GPCRs to address unmet medical needs; and (iv) validation of polypharmacology may lead to improved drug therapies.

**ACKNOWLEDGEMENTS**

The authors acknowledge funding support from the National Natural Science Foundation of China 81872915 (to M.-W.W.), 81773792 (to D.Y.), 81973373 (to D.Y.), 21704064 (to Q.Z.), 31971362 (to W.S.), 31971178 (to S.Z.), and 31770796 (to Y.J.); National Science & Technology Major Project of China—Key New Drug Creation and Manufacturing Program 2018ZX09735–001 (to M.-W.W.), 2018ZX09711002–002–005 (to D.Y.), and 2018ZX09711002–002–002 (to Y.J.); the National Key Basic Research Program of China 2018YFA0507000 (to M.-W.W., S.Z., W.S., and H.T.); Novo Nordisk-CAS Research Fund grant NNCS-2017-1-CC (to D.Y.); The Belt and Road Master Fellowship program (to V.L.); UCAS Scholarship for International Students (to S.D.); and The CAS-TWAS President’s Fellowship for International Doctoral Students (to E.V.).

**ADDITIONAL INFORMATION**

The online version of this article (https://doi.org/10.1371/journal.pone.004354-w) contains supplementary material, which is available to authorized users.

**Competing interests:** The authors declare no competing interests.

**REFERENCES**

1. Insel, P. A. et al. GPCRomics: an approach to discover GPCR drug targets. Trends Pharmacol. Sci. 40, 378–387 (2019).
2. Sriram, K. & Insel, P. A. G protein-coupled receptors as targets for approved drugs: how many targets and how many drugs? Mol. Pharmacol. 93, 251–258 (2018).
3. Wootten, D. et al. Allostery and biased agonism at class B G protein-coupled receptors. Nat. Rev. Mol. Cell Biol. 19, 638–653 (2018).
4. Hauser, A. S. et al. Trends in GPCR drug discovery: new agents, targets and indications. Nat. Rev. Drug Discov. 16, 829–842 (2017).
5. Shimada, I. et al. GPCR drug discovery: integrating solution NMR data with crystal and cryo-EM structures. Nat. Rev. Drug Discov. 18, 59–82 (2019).
6. Dalesio, N. M., Barrete Ortit, S. F., Pluznick, J. L. & Berkowitz, D. E. Olfactory, taste, and vision. Science 363, 1330–1331 (2019).
7. Cheerezov, V. et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318, 1258–1265 (2007).
8. Rosenbaum, D. M. et al. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. Science 318, 1266–1273 (2007).
9. Liang, Y. L. et al. Phase-plate cryo-EM structure of a class B GPCR-G-protein complex. Nature 546, 118–123 (2017).
10. Saldari, H. A., Pandey, S., Shukla, A. K. & Dutta, S. Illuminating GPCR signaling by cryo-EM. Trends Cell Biol. 28, 591–594 (2018).
11. Congreve, M. & Dravet, C. Dravet syndrome and epilepsy. Nat. Rev. Neurol. 10, 331–342 (2014).
12. Wootten, D. et al. Allosteric selectivity in protein-protein interactions and beyond. Nat. Rev. Drug Discov. 14, 522–534 (2015).
13. Inoue, A. et al. G protein-coupled receptor signaling. Annu. Rev. Pharmacol. Toxicol. 57, 399–417 (2017).
14. Lane, J. R. et al. A kinetic view of GPCR allosteric and biased agonism. Nat. Rev. Drug Discov. 13, 929–937 (2014).
15. Manglik, A. et al. Structure-based discovery of opioid analgesics with reduced side effects. Nature 537, 185–190 (2016).
16. Korczynska, M. et al. Structure-based discovery of selective positive allosteric modulators of antagonists for the M2 muscarinic acetylcholine receptor. Proc. Natl Acad. Sci. USA 115, E2419–E2428 (2018).
17. Foster, S. R. et al. Discovery of human signaling networks: pairing peptides to G protein-coupled receptors. Cell 179, 895–908, e821 (2019).
18. Hu, G. M., Mai, T. L. & Chen, C. M. Visualizing the GPCR network: classification and evolution. Sci. Rep. 7, 15495 (2017).
19. Basith, S. et al. Exploring G protein-coupled receptors (GPCRs) ligand space via cheminformatics approaches: impact on rational drug design. Front. Pharmacol. 9, 128 (2018).
20. Wishart, D. S. et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucl. Acids Res. 46, D1074–D1082 (2018).
21. Alexander, S. P. H. et al. The concise guide to pharmacology 2019/20: G protein-coupled receptors. Br. J. Pharmacol. 176, 521–514 (2019).
22. Bhudia, N. et al. G protein-coupling of adhesion GPCRs ADGRE2/EMR2 and ADGRE5/CD97, and activation of G protein signalling by an anti-EMR2 antibody. Sci. Rep. 10, 10004 (2020).
G protein-coupled receptors: structure- and function-based drug discovery

Yang et al.

85. Rudolf, K. et al. Development of human calcitonin gene-related peptide (CGRP) receptor antagonists. 1. Potent and selective small molecule CGRP antagonists.1-[N-2-[3,5-dibromo-N-[4-[3,4-dihydro-2H]-1H-oxazoin-3-yl]-1-piperidinyl]carboxyl]-D-tyrosyl]-L-lysyl-[4-(4-pyridyl)piperazin: the first CGRP antagonist for clinical trials in acute migraine. J. Med. Chem. 48, 5921–5931 (2005).

86. Shaw, A. W. et al. Caprolactams as potent CGRP receptor antagonists for the treatment of migraine. Bioorg. Med. Chem. Lett. 17, 4795–4798 (2007).

87. Paoane, D. V. et al. Potent, orally bioavailable calcitonin gene-related peptide receptor antagonists for the treatment of migraine: discovery of N-(3R,6S)-6-(2,3-difluorophenyl)-2-oxo-1-(2,2,2-trifluoroethyl)azepan-3-yl-4-(2-oxo-2,3-dihydro-1H-imidazo[4,5-b]pyridin-1-yl)piperidine-1-carboxamide (MK-9974). J. Med. Chem. 50, 5564–5567 (2007).

88. Liu, G. et al. Discovery of BMS-864672, a potent and orally active human CGRP receptor antagonist for the treatment of migraine. ACS Med. Chem. Lett. 3, 337–341 (2012).

89. Luo, G. et al. Discovery of (5S,6S,9R)-5-amino-6-(2,3-diarylpyridin-1-yl)piperidine-1-carboxylate (BMS-927711): an oral calcitonin gene-related peptide receptor antagonist for the treatment of migraine. J. Med. Chem. 55, 10644–10651 (2012).

90. Palczewski, K. et al. Crystal structure of rhodopsin: a G-protein-coupled receptor. Science 289, 739–745 (2000).

91. Hanson, M. A. et al. Profiling of membrane protein variants in a baculovirus system by coupling cell-surface detection with small-scale parallel expression. Protein Expr. Purif. 56, 85–92 (2007).

92. Chae, P. S. et al. Maltose-nonylglucoside (MNG) amphiphiles for solubilization, stabilization, and crystallization of membrane proteins. Nat. Methods 7, 1003–1008 (2010).

93. Chun, E. et al. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. Structure 20, 967–976 (2012).

94. Rasmussen, S. G. et al. Structure of a nanobody-stabilized active state of the beta 2 adrenergic G-protein-coupled receptor. Cell 155, 175–180 (2013).

95. Rasmussen, S. G. et al. Crystal structure of the human beta 2 adrenergic G-protein-coupled receptor. Nature 450, 383–387 (2007).

96. Caffrey, M. Crystallizing membrane proteins for structure-function studies using lipidic mesophases. Biochim. Biophys. Acta 175, 47–59 (2006).

97. Zhang, H. et al. Structural basis for ligand recognition and functional selectivity of class Frizzled receptors. Science 337, 383–386 (2012).

98. Kang, Y. et al. Crystal structure of rhodopsin bound to arrestin by femtosecond light switching revealed by a NanoBiT tethering strategy. J. Mol. Biol. 421, 1680–1690 (2019).

99. Hirai, H. et al. Crystal structure of a mammalian Wnt-frizzled complex. Nature 496, 561–567 (2013).

100. Kruse, A. C. et al. Activation and allosteric modulation of a muscarinic acetylecholine receptor. Annu. Rev. Pharmacol. Toxicol. 53, 531–556 (2013).

101. Zhang, H. et al. Structure and dynamics of the M3 muscarinic acetylcholine receptor. Nature 482, 552–556 (2012).

102. Schulte, G. International Union of Basic and Clinical Pharmacology. LXXX. The class Frizzled receptors. Pharmacol. Rev. 62, 632–667 (2010).

103. Byrne, E. F. X. et al. Structural basis of the human cannabinoid CB1 receptor. Nature 477, 549–555 (2011).

104. Golshani, P. et al. Structural basis of human cannabinoid CB1 receptor activation. Nature 477, 549–555 (2011).

105. Hollenstein, K. et al. Structure of class B GPCR corticotropin-releasing factor receptor. Nature 499, 438–443 (2013).

106. Nasuto, L. et al. Human GLP-1 receptor transmembrane domain structure in complex with allosteric modulators. Nature 546, 312–315 (2017).

107. Duan, J. et al. Cryo-EM structure of an activated V1P1 receptor-G protein complex revealed by a NanoBIT tethering strategy. Nat. Commun. 11, 4121 (2020).

108. Dore, A. S. et al. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. Nature 511, 557–562 (2014).

109. Swaroop, V. et al. Structural insights into G-protein-coupled receptor allosterism. Nature 559, 45–53 (2018).

110. Wu, H. et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. Science 344, 58–64 (2014).

111. Schulte, G. International Union of Basic and Clinical Pharmacology. LXX. The class Frizzled receptors. Pharmacol. Rev. 62, 632–667 (2010).

112. Byrne, E. F. X. et al. Structural basis of smooth muscle regulation by its extracellular domains. Nature 515, 517–522 (2016).

113. Deshpande, I. et al. Smoothened stimulation by membrane sterols drives Hedgehog pathway activity. Nature 571, 284–288 (2019).

114. Qi, X. et al. Cryo-EM structure of oxyester-bound human smoothened coupled to a heterotrimeric Gi. Nature 571, 279–283 (2019).
168. Quoyer, J. et al. Pepducin targeting the C-X-C chemokine receptor type 4 acts as
a dispute for shutting CB1 among different membrane microenvironments. Sci. Rep. 5, 15453 (2015).

169. Zhang, D. et al. Two disparate ligand-binding sites in the human P2Y1 receptor. 

172. Maeda, S. et al. Structures of the M1 and M2 muscarinic acetylcholine receptor./ 

175. Stornaiuolo, M. et al. Endogenous vs exogenous allosteric modulators in GPCRs: 

176. Fay, J. F. & Farrens, D. L. The membrane proximal region of the cannabinoid 

178. Baillie, G. L. et al. CB(1) receptor allosteric modulators display both agonist 

179. Hua, T. et al. Crystal structures of agonist-bound human cannabinoid receptor 

186. Muniz-Medina, V. M. et al. The relative activity of “function sparing” HIV-1 entry 

190. Liu, X. et al. An allosteric modulator binds to a conformational hub in the beta2 

191. Zhang, Y. et al. Cryo-EM structure of the activated GLP-1 receptor in complex 

192. Zheng, Y. et al. Structure of CC chemokine receptor 2 with orthosteric and 

194. Oswald, C. et al. Intracellular allosteric antagonism of the CCR9 receptor. 

196. Wingler, L. M. et al. Distinctive activation mechanism for angiotensin receptor 

197. Lyu, J. et al. Ultra-large library docking for discovering new chemotypes. 

198. Erlandson, S. C., McMahon, C. & Kruse, A. C. Structural basis for G protein-

200. Weierstall, U. et al. Lipidic cubic phase injector facilitates membrane protein 

203. Zhou, F. et al. Colocalization strategy unveils an underside binding site in the trans-

210. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic meso-

211. Erlandson, S. C., McMahon, C. & Kruse, A. C. Structural basis for G protein-

215. Stauss, D. P. et al. Structure of the M2 muscarinic receptor-beta-arrin complex 

217. Shore, D. M. et al. Allosteric modulation of a cannabinoid G protein-coupled receptor: binding site elucidation and relationship to G protein signaling. J. Biol. Chem. 289, 5828–3845 (2014).
G protein-coupled receptors: structure- and function-based drug discovery

Yang et al.

209. Stein, R. M. et al. Virtual discovery of melatonin receptor ligands to modulate circadian rhythms. Nature 579, 609–614 (2020).

210. Ballantine, F. et al. Docking finds GPCR ligands in dark chemical matter. J. Med. Chem. 63, 613–620 (2020).

211. Ranganathan, A. et al. Ligand discovery for a peptide-binding GPCR by structure-based screening of fragment- and lead-like chemical libraries. ACS Chem. Biol. 12, 735–745 (2017).

212. Liu, X. et al. Salvianolic acids from antiarthritic Traditional Chinese Medicine Danshen are antagonists of human P2Y1 and P2Y12 receptors. Sci. Rep. 8, 8084 (2018).

213. Stein, R. M. et al. Virtual discovery of melatonin receptor ligands to modulate circadian rhythms. Nature 579, 609–614 (2020).

214. Mannel, B. et al. Structure-guided screening for functionally selective D2 dopamine receptor ligands from a virtual chemical library. ACS Chem. Biol. 12, 2652–2661 (2017).

215. Guo, T. & Hobbs, D. W. Privileged structure-based combinatorial libraries targeting G protein-coupled receptors. Assy Drug Dev. Technol. 1, 579–592 (2003).

216. Latonaca, N. R., Venkatakrishnan, A. J. & Dör, R. G. GPCR dynamics: structures in motion. Chem. Rev. 117, 139–155 (2017).

217. Lee, Y., Lazim, R., Macalino, S. J. Y. & Choi, S. Importance of protein dynamics in the structure-based drug discovery of class A G protein-coupled receptors (GPCRs). Curr. Opin. Struct. Biol. 53, 147–153 (2019).

218. Hilger, D., Masureel, M. & Kobikila, B. K. Structure and dynamics of GPCRs signaling complexes. Nat. Struct. Mol. Biol. 25, 4–12 (2018).

219. Vilar, S. & Costanzi, S. In G Protein Coupled Receptors: Modeling, Activation, Interactions and Virtual Screening. Methods in Enzymology, Vol. 522 (ed. Conn P. M.) 263–278 (Elsevier Academic Press, 2013).

220. Coudrat, T., Christopoulos, A., Sexton, P. M. & Wootten, D. Structural features of the structure-based drug discovery of class A G protein-coupled receptors (GPCRs). Curr. Opin. Struct. Biol. 53, 147–153 (2019).

221. Zhou, Y. et al. Accelerated structure-based design of chemically diverse allosteric modulators of a muscarinic G protein-coupled receptor. Proc. Natl Acad. Sci. USA 113, E5675–E5684 (2016).

222. Warszyczki, D. et al. From homology models to a set of predictive binding pockets-a 5-HT1A receptor case study. J. Chem. Inf. Model. 57, 311–321 (2017).

223. de Graaf, C. et al. Crystal structure-based virtual screening for fragment-like ligands of the human histamine H1 receptor. J. Med. Chem. 54, 8195–8206 (2011).

224. David, L., Nielsen, P. A., Hedstrom, M. & Norden, B. Scope and limitation of ligand docking: methods, scoring functions and protein targets. Curr. Comput. Aided Drug Des. 1, 275–306 (2005).

225. Kooistra, A. J. et al. Function-specific virtual screening for GPCR ligands using a combined scoring method. Sci. Rep. 6, 28288 (2016).

226. Bartuzi, D., Kaczor, A. A., Targowska-Duda, K. M. & Matoiuik, D. Recent advances and applications of molecular docking to G protein-coupled receptors. Molecules 22, 23 (2017).

227. Zhou, Y. et al. Structure-based discovery of novel and selective 5-hydroxytryptamine 2B receptor antagonists for the treatment of irritable bowel syndrome. J. Med. Chem. 59, 707–720 (2016).

228. Rastelli, G. & Pinzi, L. Recent applications of molecular docking to G protein-coupled receptors. Molecules 22, 23 (2017).

229. Ahn, S. et al. Allosteric “beta-blocker” isolated from a DNA-encoded small molecule library. Proc. Natl Acad. Sci. USA 114, 1708–1713 (2017).

230. Ahn, S. et al. Small-molecule positive allosteric modulators of the beta2-adrenergic receptor identified from DNA-encoded libraries. Mol. Pharmacol. 94, 850–861 (2018).

231. Brown, D. G. et al. Agonists and antagonists of protease-activated receptor 2 discovered within a DNA-encoded chemical library using mutational stabilization of the target. SLAS Discov. 23, 429–436 (2018).

232. Wu, Z. et al. Cell-based selection expands the utility of DNA-encoded small-molecule library technology to cell surface drug targets: identification of novel antagonists of the Nk3 tachykinin receptor. ACS Comb. Sci. 17, 722–731 (2015).

233. Annis, A., Chuang, C. C. & Nazef, N. In Mass Spectrometry in Medicinal Chemistry: Applications in Drug Discovery, Methods and Principles in Medicinal Chemistry (eds Wanner, K. T. & Höfner, G.) 121–156 (Wiley, 2007).

234. O’Connell, T. N. et al. Solution-based indirect affinity selection mass spectrometry—a general tool for high-throughput screening of pharmaceutical compound libraries. Anal. Chem. 86, 7413–7420 (2014).

235. Chen, X. et al. Identification of inhibitors of the antibiotic-resistance target New Delhi metallo-beta-lactamase 1 by both nanolaser electrospray ionization mass spectrometry and ultrafiltration liquid chromatography-mass spectrometry approaches. Anal. Chem. 85, 7957–7965 (2013).

236. Chen, X. et al. A ligand-observed mass spectrometry approach integrated into the fragment based lead discovery pipeline. Sci. Rep. 5, 8361 (2015).

237. Qin, S. et al. Multiple ligand detection and affinity measurement by ultrafiltration and mass spectrometry analysis applied to fragment mixture screening. Anal. Chem. Acta 806, 98–106 (2015).

238. Gesmund, N. J. et al. Nanoscale synthesis and affinity ranking. Nature 557, 228–232 (2018).

239. Whitehurst, C. E. et al. Application of affinity selection-mass spectrometry assays to purification and affinity-based screening of the chemokine receptor CXCR4. Comb. Chem. High Throughput Screen. 15, 473–485 (2012).

240. Ma, J. et al. Ligand identification of the adenosine A2A receptor in self-assembled nanodiscs by affinity mass spectrometry. Anal. Methods 8, 5851–5858 (2017).

241. Calleri, E. et al. Frontal affinity chromatography-mass spectrometry useful for characterization of new ligands for GPR17 receptor. J. Med. Chem. 53, 3489–3501 (2010).

242. Temporini, C. et al. Development of new chromatographic tools based on A2A adenosine receptor subtype for ligand characterization and screening by FAC-MSC. Anal. Bioanal. Chem. 405, 837–845 (2013).

243. Qin, S. et al. High-throughput identification of G protein-coupled receptor ligands through affinity mass spectrometry screening. Chem. Sci. 9, 3192–3199 (2018).
