GPCRdb in 2023: state-specific structure models using AlphaFold2 and new ligand resources

Gáspár Pándy-Szekeres[1,2,†], Jimmy Caroli[1,†], Alibek Mamyrbekov[1], Ali A. Kermani[3], György M. Keserű[2], Albert J. Kooistra [ID][1,*,‡] **and David E. Gloriam** [ID][1,*,‡]

[1]Department of Drug Design and Pharmacology, University of Copenhagen, Universitetsparken 2, 2100 Copenhagen, Denmark, [2]Medicinal Chemistry Research Group, Research Center for Natural Sciences, Budapest H-1117, Hungary and [3]Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA

ABSTRACT

G protein-coupled receptors (GPCRs) are physiologically abundant signaling hubs routing hundreds of extracellular signal substances and drugs into intracellular pathways. The GPCR database, GPCRdb supports >5000 interdisciplinary researchers every month with reference data, analysis, visualization, experiment design and dissemination. Here, we present our fifth major GPCRdb release setting out with an overview of the many resources for receptor sequences, structures, and ligands. This includes recently published additions of class D generic residue numbers, a comparative structure analysis tool to identify functional determinants, trees clustering GPCR structures by 3D conformation, and mutations stabilizing inactive/active states. We provide new state-specific structure models of all human non-olfactory GPCRs built using AlphaFold2-MultiState. We also provide a new resource of endogenous ligands along with a larger number of surrogate ligands with bioactivity, vendor, and physiochemical descriptor data. The one-stop-shop ligand resources integrate ligands/data from the ChEMBL, Guide to Pharmacology, PDSP Ki and PubChem database. The GPCRdb is available at https://gpcrdb.org.

INTRODUCTION

G protein-coupled receptors (GPCRs) account for 799 (1) out of 20 383 (2) (4%) human genes and transduce the responses of two-thirds (342/515) of signaling substances (3). Their predominant role in physiology is mirrored in medicine, as 34% of marketed drugs target GPCRs (4). The GPCR database, GPCRdb serves ~5000 researchers every month with reference data, analysis, visualization, experiment design and data deposition. GPCRdb is open access and open source, and compliant with the FAIR principles (5). In 2022, what was previously a G protein section of GPCRdb grew into a dedicated database, GproteinDb adding, e.g. a coupling atlas integrating major datasets (6), structure complexes with GPCRs, and interface interactions (7). Also ArrestinDb (8) and Biased Signaling Atlas (9) have been much expanded and are brought forward as separate resources covering complementary aspects of signal transduction while serving dedicated research communities and use cases.

GPCRdb has long provided state-of-the-art homology models from a unique multi-template pipeline. However, current advances call for a transition towards machine-learning based models taking advantage of recent breakthroughs. AlphaFold2 was, together with RoseTTA-fold, awarded the method and breakthrough of the year by *Nature* (10) and *Science* (11), respectively. While pre-generated AlphaFold2 models can be downloaded from the European Bioinformatics Institute (12), they have limitations of both general and GPCR-specific scope. Furthermore, AlphaFold models are provided for only one structural/functional state of each protein. To provide both inactive or active state GPCR models, a prediction protocol AlphaFold-MultiState (13) has been developed using two distinct sets of templates for the inactive and active states, respectively (13) based on state classifications from GPCRdb (14). However, the AlphaFold-MultiState models do not cover all GPCRs and were published without mention of long-term regular updates incorporating new templates (latest templates from June 2021).

Ligands of GPCRs span endogenous ligands, tool compounds, drug, and agents in clinical trials. The endogenous ligands are curated by experts coordinated by the authoritative Nomenclature Committee of the International Union

---

of Basic and Clinical Pharmacology (NC-IUPHAR, https://www.guidetopharmacology.org/nomenclature.jsp). They are provided in the Guide to Pharmacology (GtP) database (15) along with a classification of a receptor's principal and secondary ligand (when multiple), binding affinity (pK_i) and potency (pEC_{50}) values. The largest source of tool compounds is ChEMBL (16), which contains over 206 000 GPCR ligands with binding (pK_i, pK_d) or functional (pEC_{50}, (pIC_{50})) data. Information about drugs approved by the US Food and Drug Administration is available from DrugCentral and DrugBank. Another source of consistently determined binding affinities – for psychoactive endogenous, tool and drug ligands—is the K_i database from the Psychoactive Drug Screening Program (https://pdsp.unc.edu/databases/kidb.php). Furthermore, commercial availability at vendors and physicochemical properties are provided for nearly all ligands in the PubChem database (17). GPCRdb has previously integrated ChEMBL ligands and PubChem data (18,19).

In this article, we present an overview of all data and tool resource in our fifth major GPCRdb release, state-specific GPCR structure models (using AlphaFold2-MultiState and greatly expanded ligand resources (integrating all above ligand databases). These updates will increase GPCRdb's utility as a one-stop-shop for the GPCR community across basic research and drug discovery areas.

METHODS

Building state-specific GPCR structure models

Due to AlphaFold's limitation regarding lack of activation state specificity, we built 844 active or inactive state AlphaFold-MultiState models for 422 GPCRs in GPCRdb. We used structural templates published up until 4 July 2022 (published AlphaFold-MultiState models were based on templates until June 2021). The use of AlphaFold-MultiState removed the need for alternative backbone templates in our previously published GPCRdb's structure modelling pipeline (18,19). All generated models are based on the full-length wildtype sequence, except for AGRV1 for which partial models were generated since this receptor comprises 6306 amino acids which exceeded our computing capabilities.

Building refined GPCR structures

Refined structures revert mutations to wildtype. Missing regions are filled in from an AlphaFold model based solely only the given experimental structure and fitted by superposition at junction sites.

Updating of ligands and bioactivities

GPCR ligands and their biological activities were imported from the ChEMBL (v. 30) (16), Guide to Pharmacology (v. 2022.2) (15) and PDSP K_i databases (accessed 7 September 2022, https://pdsp.unc.edu/databases/kidb.php), as in (19). Ligand physicochemical properties and commercial availability data were retrieved from the PubChem database (17), as in (18). Information about FDA approval of drugs was imported from DrugBank. To enable data integration and links to external drug (DrugBank, and DrugCentral)

and ligand databases (ChEMBL, Guide to Pharmacology, and PubChem), we developed a ligand search functionality. This approach allows for searching ligands using information from different sources of information, such as ligand names, database identifiers matched via UniChem (20), ligand InChIKeys (calculated with RDKit, http://www.rdkit.org), CAS numbers via the Entrez E-utilities (21), UniProt accession numbers (2) and sequences.

Building an endogenous ligand browser

Endogenous ligand bioactivities and associated references were downloaded from the Guide to Pharmacology (https://www.guidetopharmacology.org/download.jsp). These data contain a 'principal' and 'secondary' classification of many, but not all, endogenous ligands that is important in many scientific studies, for example when selecting a reference ligand for physiology-biased signaling (22). For receptors lacking this classification and having only a single endogenous ligand we assigned the 'principal' category. To estimate the physiologically most important ligand for receptors with multiple unclassified ligands, we calculated a potency ranking for all endogenous ligands of each GPCR. This calculation was performed based on the mean pEC_{50}, and if not available pK_i, across studies.

CURRENT DATA, ANALYSIS TOOLS AND EXPERIMENT DESIGN TOOLS

Given that this is our fifth publication of a major GPCRdb release (18,19,23,24), and additional resources have published (14,25) after the latest publication (19), we provide an overview of the current data, analysis, data-driven experiment design, and data deposition in Figures 1 and 2. Data: The updated data types span receptor sequences, isoforms (26), genetic variants (27), structures, drugs and mutations used to pinpoint ligand binding residues or (thermo-)stabilize structures (28). The sequences and generic residue numbers were recently expanded to cover the Class D of Ste2 fungal pheromone receptors (25). For refined structures, structure models, and ligands (endogenous and surrogate) we describe major updates of data or functionality in the following sections. Analysis: The analysis tools cover all of the sequence, structural and ligand/drug data spaces. Recently, we added an online GPCR structure analysis platform featuring a tool to identify functional determinants from comparative structure analysis, and trees clustering receptors by their 3D conformation (14). Data-driven experiment design tools: The current set of data-driven tools serve to (i) design constructs for crystallography/cryo-EM studies (28), (ii) identify functional determinants from comparative sequence analysis, (iii) provide mutations stabilizing inactive/active states, and mutate ligand binding sites (29).
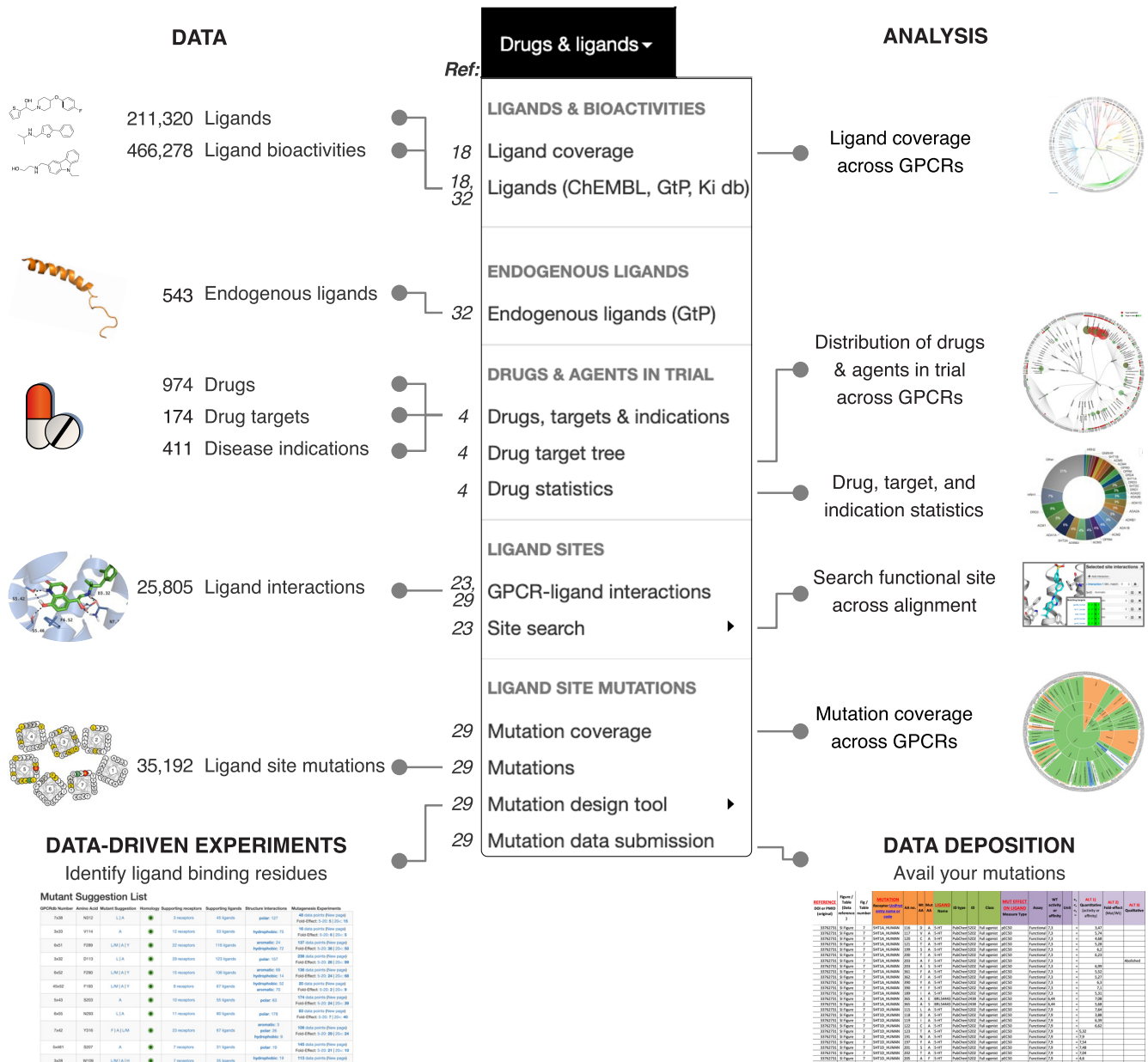
STRUCTURES AND MODELS

Structures and refined structures

GPCRdb makes monthly updates of GPCR structures from the Protein Data Bank (PDB) (33) that are complemented with manual annotation of additional information about the receptors, endogenous ligands, bound ligands, bound transducer proteins, and auxiliary proteins. This allows the

Figure 1. GPCR data, analysis resources and data-driven experiment tools in the GPCRdb 2023 release. The numbers to the left of each menu item are publication references with more information (including 30–32).

Figure 2. Drug/ligand data, analysis, data-driven experiment tool, and data deposition in the GPCRdb 2023 release. The numbers to the left of each menu item are publication references with more information.

Structures page to aid structure browsing and selection based on e.g. receptor classes, ligand, or structure properties, such as resolution. Since the last GPCRdb update publication (19), the number of annotated structures has increased from 488 to 788. Structures are also offered in a refined version wherein mutated residues are reverted to wildtype. Furthermore, missing regions are modelled in using AlphaFold models as swap-in templates for the missing coordinates.

State-specific structure models using AlphaFold-Multistate

This release contains 844 inactive or active state GPCR structure models built using AlphaFold-Multistate (exemplified in Figure 3). The models can be downloaded from the *Structure models* page with or without loop and termini segments. In a model details page, users can view the 3D structure models color-coded by the AlphaFold confidence scores. All are full length models (include loops and termini) of wildtype sequence, except for AGRV1 which had to be reduced due to its extraordinary length of 6,306 amino acids. This includes 41 new and 71 longer receptors (82 and 142 models, respectively) compared to the AlphaFold-Multistate archive, and 5 receptors (10 models) that are not available from AlphaFold Protein Structure Database (12) website. To provide an initial validation, we calculated root-mean-square (RMSD) values for five target structures that were published after our models were built. The validation is available in the *Structure model validation* page and yielded an average RMSD of 1.6 Å for the back-
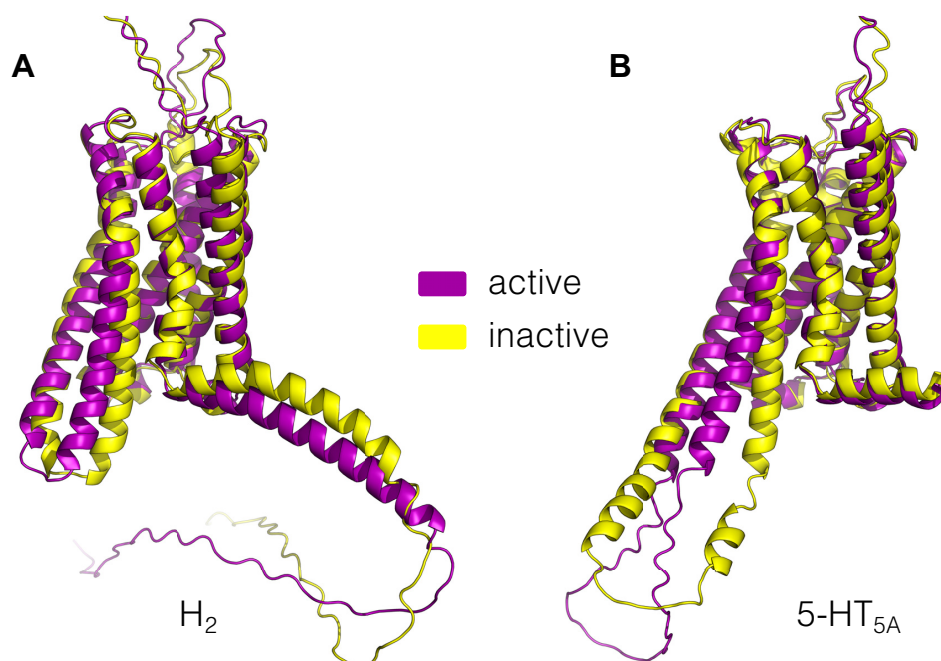
Figure 3. Structure models of the histamine H_2 and serotonin 5-HT$_{5A}$ receptors in both the active and inactive states built with AlphaFold-Multistate.

bone of transmembrane domain. This supports the use of AlphaFold-Multistate for the specific modelling of activation state of GPCRs.

The models will be updated twice a year including newly published structure templates and their determined activation state, thereby constantly improving the quality of the provided receptor models. In some cases, AlphaFold models do not represent a fully inactive or active state because there are no closely related structural templates. We will filter these models and sample more receptor conformations and explore rotamer optimization to improve docking of receptor ligands.

LIGANDS

Ligands and their coverage across GPCR targets

The GPCRdb ligand resource has been updated to incorporate the most recent version of ChEMBL (16) and add ligands from the Guide to Pharmacology (15) and PDSP K_i (https://pdsp.unc.edu/databases/kidb.php) databases. It currently contains 211 320 ligands and 466 278 binding affinity or potency values (6% (>20% if considering duplicates) and a 19% increase, respectively compared to GPCRdb's 2021 publication). The coverage of ligands across the receptors in the different human GPCR classes is A: 236 (75%), B1: 15 (100%), B2: 5 (15%), C: 12 (55%), F: 8 (73%) and T2: 25 (100%). Furthermore, the average number of ligands per receptor in each GPCR class is A: 612, B1: 400, C: 372, F: 90 and T2: 19.

Ligand browser

The receptor target selection page has added selection of receptors that are drug targets through two new columns

'Target of an approved drug', and 'Target in clinical trials', as defined in DrugBank, and DrugCentral. After selecting the receptor target, users can choose between two ligand bioactivity browsers. The 'Compact (1 row/ligand)' browser collates all binding or functional bioactivities of a given ligand and source database on one row by calculating minimum, average and maximum activity across studies. The 'Extended (1 row/activity)' browser instead lists the specific binding affinity or potency value of each study. This GPCRdb release has restructured potency and affinity data into separate tabs. It has also added fold selectivity values, along with the underlying number of experiments, allowing ligand selection based on their selectivity for the target of interest relative to all other stored GPCR targets. Both browsers also present information about vendors from which one can purchase the given ligand along with key physicochemical descriptors (18).

Ligand info page

This GPCRdb release contains a new page for individual ligands and its targets (Figure 4). This can be accessed directly by a ligand query or from the results page of the above ligand browser. The top of the ligand info page displays information about ligand structure (2D image, SMILES, InChI key), names (common, and chemical names, and aliases), physicochemical properties (molecular weight, logP and counts for hydrogen bond acceptors/donors, and rotatable bonds), molecule type (small molecule/peptide/protein, drug status and endogenous/surrogate), and database links (internal and external). For ligands that have this information, two additional boxes provide GPCR-ligand crystal/cryo-EM structure complexes and mutations affecting ligand activity. The bottom of the page shows bioactivities for the given ligand

Dopamine

SMILES	NCCc1ccc(c(c1)O)O
InChIKey	VYFYYTLLBUKUHU-UHFFFAOYSA-N
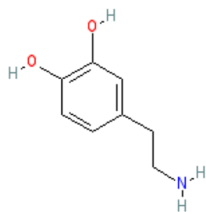
Chemical properties

Hydrogen bond acceptors	3
Hydrogen bond donors	3
Rotatable bonds	2
Molecular weight (Da)	153.1

Drug properties

Molecular type	Small molecule
Endogenous/Surrogate	Endogenous
Approved drug	Yes

Database connections

Guide To Pharmacology DrugBank Drug Central ChEMBL_compound_ids PubChem GPCRdb	
Structure pdb	7LJD 7CKZ
Ligand site mutations	α_{1B} α_{2A} β_2 D_1 D_2 D_3 D_4 D_5

Bioactivity

Export to CSV | Reset All Filters | Get Vendors

Receptor					Type	Activity			Source
GTP	Uniprot	Species	Family	Class	Type	Min	Avg	Max	Database
5-HT$_7$	5HT7R	Human	5-Hydroxytryptamine	A	pKi	4.0	4.0	4.0	PDSP Ki database
5-HT$_{1A}$	5HT1A	Human	5-Hydroxytryptamine	A	pKi	4.08	4.08	4.08	PDSP Ki database
β_2	ADRB2	Human	Adrenoceptors	A	pKd	4.19	4.19	4.19	ChEMBL
β_2	ADRB2	Human	Adrenoceptors	A	pKi	4.21	4.21	4.21	ChEMBL
D$_1$	DRD1	Human	Dopamine	A	pKi	4.3	4.95	5.6	Guide to Pharmacology

Figure 4. Ligand info page. The Ligands subsection of GPCRdb has been extended to 211 320 ligands and 466 278 binding affinity or potency values from the ChEMBL (16), Guide to Pharmacology (15) or PDSP K_i databases (https://pdsp.unc.edu/databases/kidb.php). The new ligand info page shows ligand structure, names, chemical properties, molecule type, drug status, and endogenous/surrogate status, and database links. When available, GPCR-ligand crystal/cryo-EM structure complexes and mutations affecting ligand activity are also shown. Bioactivities across receptor targets are shown in the browser at the bottom.

Data showed in this table is continuously collected from the Guide to Pharmacology database. Ref

Export to Excel | Reset filters

Receptor					Endogenous ligands				pEC50			pKi			References
Class	Receptor family	UniProt	IUPHAR	Species	Ligand	GtP Classification	Potency Ranking	Type	Min	Mid	Max	Min	Mid	Max	
B1 (Secretin)	Calcitonin	CALCR	CT	Human	Calcitonin	Principal		Peptide	9.0	10.1	11.2				Refs
B1 (Secretin)	VIP and PACAP	PACR	PAC$_1$	Human	Pacap-38	Principal		Peptide	9.5	9.8	10.0	6.9	8.2	9.0	Refs
B1 (Secretin)	Glucagon recep...	SCTR	secretin	Human	Secretin		1	Peptide		9.7	9.7				Refs
B1 (Secretin)	Glucagon recep...	GLR	glucagon	Human	Glucagon			Peptide		9.0	9.0				Refs
B1 (Secretin)	VIP and PACAP	VIPR1	VPAC$_1$	Human	Vip	Principal		Peptide	7.9	9.0	10.0	8.5	9.2	9.8	Refs
B1 (Secretin)	VIP and PACAP	PACR	PAC$_1$	Human	Pacap-27	Principal		Peptide	7.4	8.9	10.6	6.9	8.0	8.5	Refs
B1 (Secretin)	Calcitonin	CALCR	CT	Human	Amylin	Principal		Peptide	8.0	8.6	9.2				Refs
B1 (Secretin)	VIP and PACAP	VIPR1	VPAC$_1$	Human	Pacap-38	Principal		Peptide	7.4	8.6	9.7		8.2	8.2	Refs
B1 (Secretin)	VIP and PACAP	VIPR1	VPAC$_1$	Human	Pacap-27	Principal		Peptide	7.6	8.6	9.9		8.9	8.9	Refs
B1 (Secretin)	VIP and PACAP	VIPR2	VPAC$_2$	Human	Pacap-27	Principal		Peptide	7.6	8.5	9.4				Refs
B1 (Secretin)	VIP and PACAP	VIPR2	VPAC$_2$	Human	Pacap-38	Principal		Peptide	7.7	8.5	9.3				Refs
B1 (Secretin)	VIP and PACAP	VIPR2	VPAC$_2$	Human	Vip	Principal		Peptide	7.3	8.3	9.3	7.8	8.4	8.8	Refs
B1 (Secretin)	Calcitonin	CALCR	CT	Human	A-Cgrp	Secondary		Peptide	6.2	7.3	8.4				Refs
B1 (Secretin)	Calcitonin	CALCR	CT	Human	Adrenomedullin	Secondary		Peptide	6.7	7.2	7.7				Refs
B1 (Secretin)	Calcitonin	CALCR	CT	Human	B-Cgrp	Secondary		Peptide		7.2	7.2				Refs
B1 (Secretin)	VIP and PACAP	PACR	PAC$_1$	Human	Vip	Secondary		Peptide	6.0	7.0	8.7	6.0	6.9	8.4	Refs
B1 (Secretin)	Calcitonin	CALCR	CT	Human	Adrenomedullin ...	Secondary		Peptide		6.5	6.5				Refs

Figure 5. The endogenous ligand resource offers browsing across both multiple ligand and receptors. Four sections present (i) receptors (names and classification), (ii) endogenous ligands (name with link to detailed ligand info, ligand type (small molecule, peptide, protein), classification into principal, secondary or none categories, and potency rank), (iii) bioactivities (pEC$_{50}$ and pK$_i$ minimum, average, and maximum values across studies) and (iv) references (publication link). All endogenous ligands and their primary/secondary classification were derived from the Guide to Pharmacology database (15).

across receptor targets. The bioactivity browser allows filtering by the receptor classification, bioactivity, and source database. The information on endogenous ligands and target FDA approval status have been derived from Guide to Pharmacology (15) and DrugBank (34) databases, respectively.

Endogenous ligand browser

The endogenous ligand-GPCR system spans different relationships ranging (ligand:receptor) 1:1, 1:many, many:1 and many:many relationships (3). To facilitate browsing across either ligands or receptors, we developed an endogenous ligand browser (Figure 5). This browser contains data for 543 distinct endogenous ligands for 253 human GPCRs, and 157 mouse, rat, or guinea pig receptors. For each receptor, alternative endogenous ligands are classified as principal or secondary, as defined by the nomenclature committee of the International Union of Basic and Clinical Pharmacology (https://www.guidetopharmacology.org/nomenclature.jsp) and have an additional ranking by potency. For ligand-receptor pairs with multiple potency (pEC$_{50}$) or affinity (pKi) values, the browser provides minimum, mean and maximum values, with grayscale background aiding comparison. Finally, the browser contains information about the ligand type, ligand name with a direct link to the ligand info page in GPCRdb, receptor information as family, species, IUPHAR and UniProt name and a popup showing the original references for bioactivities.

DATA AVAILABILITY

GPCRdb is available at https://gpcrdb.org and can also be accessed via a RESTful API, which complies with the OpenAPI specification using Swagger (code examples are available at https://docs.gpcrdb.org/web_services.html). The source code, the underlying data and a virtual machine configuration are all available in the repositories at https://github.com/protwis/.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Gloriam,D.E., Fredriksson,R. and Schioth,H.B. (2007) The g protein-coupled receptor subset of the rat genome. *BMC Genomics*, **8**, 338.
2. UniProt Consortium (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, **49**, D480–D489.
3. Foster,S.R., Hauser,A.S., Vedel,L., Strachan,R.T., Huang,X.P., Gavin,A.C., Shah,S.D., Nayak,A.P., Haugaard-Kedstrom,L.M., Penn,R.B. *et al.* (2019) Discovery of human signaling systems: pairing peptides to g protein-coupled receptors. *Cell*, **179**, 895–908.
4. Hauser,A.S., Attwood,M.M., Rask-Andersen,M., Schioth,H.B. and Gloriam,D.E. (2017) Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discov.*, **16**, 829–842.
5. Wilkinson,M.D., Dumontier,M., Aalbersberg,I.J., Appleton,G., Axton,M., Baak,A., Blomberg,N., Boiten,J.-W., da Silva Santos,L.B., Bourne,P.E. *et al.* (2016) The FAIR guiding principles for scientific data management and stewardship. *Sci Data*, **3**, 160018.
6. Hauser,A.S., Avet,C., Normand,C., Mancini,A., Inoue,A., Bouvier,M. and Gloriam,D.E. (2022) Common coupling map advances GPCR-G protein selectivity. *Elife*, **11**, e74107
7. Pandy-Szekeres,G., Esguerra,M., Hauser,A.S., Caroli,J., Munk,C., Pilger,S., Keseru,G.M., Kooistra,A.J. and Gloriam,D.E. (2022) The g protein database, GproteinDb. *Nucleic Acids Res.*, **50**, D518–D525.
8. Caroli,J., Pandy-Szekeres,G., Hauser,A.S., Keseru,G.M., Kooistra,A.J. and Gloriam,D.E. (2022) The arrestin database. *ArrestinDb*. https://arrestindb.org/.
9. Caroli,J., Mamyrbekov,A., Harpsøe,K., Gardizi,S., Dörries,L., Ghosh,E., Hauser,A.S., Kooistra,A.J. and Gloriam,D.E. (2022) An online biased signaling atlas. In Review.
10. Editorial (2022) Method of the year 2021: protein structure prediction. *Nat. Methods*, **19**, 1.
11. Humphreys,I.R., Pei,J., Baek,M., Krishnakumar,A., Anishchenko,I., Ovchinnikov,S., Zhang,J., Ness,T.J., Banjade,S., Bagde,S.R. *et al.* (2021) Computed structures of core eukaryotic protein complexes. *Science*, **374**, eabm4805.
12. Varadi,M., Anyango,S., Deshpande,M., Nair,S., Natassia,C., Yordanova,G., Yuan,D., Stroe,O., Wood,G., Laydon,A. *et al.* (2021) AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic. Acids. Res.*, **50**, D439–D444.
13. Heo,L. and Feig,M. (2022) Multi-state modeling of G-protein coupled receptors at experimental accuracy. *Proteins*, **90**, 1873–1885.
14. Kooistra,A.J., Munk,C., Hauser,A.S. and Gloriam,D.E. (2021) An online GPCR structure analysis platform. *Nat. Struct. Mol. Biol.*, **28**, 875–878.
15. Harding,S.D., Armstrong,J.F., Faccenda,E., Southan,C., Alexander,S.P.H., Davenport,A.P., Pawson,A.J., Spedding,M., Davies,J.A. and Nc,I. (2022) The IUPHAR/BPS guide to PHARMACOLOGY in 2022: curating pharmacology for COVID-19, malaria and antibacterials. *Nucleic Acids Res.*, **50**, D1282–D1294.
16. Mendez,D., Gaulton,A., Bento,A.P., Chambers,J., De Veij,M., Felix,E., Magarinos,M.P., Mosquera,J.F., Mutowo,P., Nowotka,M. *et al.* (2019) ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.*, **47**, D930–D940.
17. Kim,S., Chen,J., Cheng,T., Gindulyte,A., He,J., He,S., Li,Q., Shoemaker,B.A., Thiessen,P.A., Yu,B. *et al.* (2021) PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res.*, **49**, D1388–D1395.
18. Pandy-Szekeres,G., Munk,C., Tsonkov,T.M., Mordalski,S., Harpsoe,K., Hauser,A.S., Bojarski,A.J. and Gloriam,D.E. (2018) GPCRdb in 2018: adding GPCR structure models and ligands. *Nucleic Acids Res.*, **46**, D440–D446.
19. Kooistra,A.J., Mordalski,S., Pandy-Szekeres,G., Esguerra,M., Mamyrbekov,A., Munk,C., Keseru,G.M. and Gloriam,D.E. (2021) GPCRdb in 2021: integrating GPCR sequence, structure and function. *Nucleic Acids Res.*, **49**, D335–D343.
20. Chambers,J., Davies,M., Gaulton,A., Hersey,A., Velankar,S., Petryszak,R., Hastings,J., Bellis,L., McGlinchey,S. and Overington,J.P. (2013) UniChem: a unified chemical structure cross-referencing and identifier tracking system. *J. Cheminform.*, **5**, 3.
21. Sayers,E.W., Beck,J., Bolton,E.E., Bourexis,D., Brister,J.R., Canese,K., Comeau,D.C., Funk,K., Kim,S., Klimke,W. *et al.* (2021)

Database resources of the national center for biotechnology information. *Nucleic Acids Res.*, **49**, D10–D17.

22. Kolb,P., Kenakin,T., Alexander,S.P.H., Bermudez,M., Bohn,L.M., Breinholt,C.S., Bouvier,M., Hill,S.J., Kostenis,E., Martemyanov,K.A. *et al.* (2022) Community guidelines for GPCR ligand bias: IUPHAR review 32. *Br. J. Pharmacol.*, **179**, 3651–3674.

23. Isberg,V., Mordalski,S., Munk,C., Rataj,K., Harpsoe,K., Hauser,A.S., Vroling,B., Bojarski,A.J., Vriend,G. and Gloriam,D.E. (2016) GPCRdb: an information system for g protein-coupled receptors. *Nucleic Acids Res.*, **44**, D356–D364.

24. Isberg,V., Vroling,B., van der Kant,R., Li,K., Vriend,G. and Gloriam,D. (2014) GPCRDB: an information system for g protein-coupled receptors. *Nucleic Acids Res.*, **42**, D422–D425.

25. Velazhahan,V., Ma,N., Pandy-Szekeres,G., Kooistra,A.J., Lee,Y., Gloriam,D.E., Vaidehi,N. and Tate,C.G. (2021) Structure of the class d GPCR ste2 dimer coupled to two g proteins. *Nature*, **589**, 148–153.

26. Marti-Solano,M., Crilly,S.E., Malinverni,D., Munk,C., Harris,M., Pearce,A., Quon,T., Mackenzie,A.E., Wang,X., Peng,J. *et al.* (2020) Combinatorial expression of GPCR isoforms affects signalling and drug responses. *Nature*, **587**, 650–656.

27. Hauser,A.S., Chavali,S., Masuho,I., Jahn,L.J., Martemyanov,K.A., Gloriam,D.E. and Babu,M.M. (2018) Pharmacogenomics of GPCR drug targets. *Cell*, **172**, 41–54.

28. Munk,C., Mutt,E., Isberg,V., Nikolajsen,L.F., Bibbe,J.M., Flock,T., Hanson,M.A., Stevens,R.C., Deupi,X. and Gloriam,D.E. (2019) An online resource for GPCR structure determination and analysis. *Nat. Methods*, **16**, 151–162.

29. Munk,C., Harpsoe,K., Hauser,A.S., Isberg,V. and Gloriam,D.E. (2016) Integrating structural and mutagenesis data to elucidate GPCR ligand binding. *Curr. Opin. Pharmacol.*, **30**, 51–58.

30. Isberg,V., de Graaf,C., Bortolato,A., Cherezov,V., Katritch,V., Marshall,F.H., Mordalski,S., Pin,J.P., Stevens,R.C., Vriend,G. *et al.* (2015) Generic GPCR residue numbers - aligning topology maps while minding the gaps. *Trends Pharmacol. Sci.*, **36**, 22–31.

31. Munk,C., Isberg,V., Mordalski,S., Harpsoe,K., Rataj,K., Hauser,A.S., Kolb,P., Bojarski,A.J., Vriend,G. and Gloriam,D.E. (2016) GPCRdb: the g protein-coupled receptor database - an introduction. *Br. J. Pharmacol.*, **173**, 2195–2207.

32. Pándy-Szekeres,G., Caroli,J., Mamyrbekov,A., Kermani,A.A., Keserű,G.M., Kooistra,A.J. and Gloriam,D.E. (2022) GPCRdb in 2023: state-specific structure models using alphafold2 and expansion of ligand resources. *Nucleic Acids Res.*, https://doi.org/10.1093/nar/gkac1013.

33. Burley,S.K., Bhikadiya,C., Bi,C., Bittrich,S., Chen,L., Crichlow,G.V., Christie,C.H., Dalenberg,K., Di Costanzo,L., Duarte,J.M. *et al.* (2021) RCSB protein data bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res.*, **49**, D437–D451.

34. Wishart,D.S., Feunang,Y.D., Guo,A.C., Lo,E.J., Marcu,A., Grant,J.R., Sajed,T., Johnson,D., Li,C., Sayeeda,Z. *et al.* (2018) DrugBank 5.0: a major update to the drugbank database for 2018. *Nucleic Acids Res.*, **46**, D1074–D1082.