



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Accessing expert-curated pharmacological data in the IUPHAR/BPS Guide to PHARMACOLOGY

Citation for published version:

NCIUPHAR 2018, 'Accessing expert-curated pharmacological data in the IUPHAR/BPS Guide to PHARMACOLOGY: Guide to PHARMACOLOGY and Guide to IMMUNOPHARMACOLOGY', *Current Protocols in Bioinformatics*. <https://doi.org/10.1002/cpbi.46>

Digital Object Identifier (DOI):

[10.1002/cpbi.46](https://doi.org/10.1002/cpbi.46)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Current Protocols in Bioinformatics

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





Accessing expert-curated pharmacological data in the IUPHAR/BPS Guide to PHARMACOLOGY

Journal:	<i>Current Protocols</i>
Manuscript ID	CP-17-0310.R1
Wiley - Manuscript type:	Protocol
Date Submitted by the Author:	n/a
Complete List of Authors:	Sharman, Joanna; University of Edinburgh, Deanery of Biomedical Sciences Harding, Simon; University of Edinburgh, Deanery of Biomedical Sciences Southan, Christopher; University of Edinburgh, Deanery of Biomedical Sciences Faccenda, Elena; University of Edinburgh, Deanery of Biomedical Sciences Pawson, Adam; University of Edinburgh, Deanery of Biomedical Sciences Davies, Jamie; University of Edinburgh, Deanery of Biomedical Sciences
Keywords:	Pharmacology, immunology, medicinal chemistry, curated database, drug target, ligand, receptor
Abstract:	The IUPHAR/BPS Guide to PHARMACOLOGY is an expert-curated, open-access database of information on drug targets and the substances that act on them. This paper describes the procedures for searching and downloading ligand-target binding data and for finding detailed annotations and the most relevant literature. The database includes concise overviews of the properties of 1,700 data-supported human drug targets and related proteins, divided into families, and 9,000 small molecule and peptide experimental ligands and approved drugs that bind to those targets. More detailed descriptions of pharmacology, function, and pathophysiology are provided for a subset of important targets. The information is reviewed regularly by expert subcommittees of the IUPHAR Committee on Receptor Nomenclature and Drug Classification. A new immunopharmacology portal has recently been added, drawing together data on immunological targets, ligands, cell types, processes and diseases. The data are available for download and can be accessed computationally via web services.

SCHOLARONE™
Manuscripts

Unit Title

Accessing expert-curated pharmacological data in the IUPHAR/BPS Guide to PHARMACOLOGY

Running Title

Guide to PHARMACOLOGY and Guide to IMMUNOPHARMACOLOGY

Author(s)

Joanna L Sharman, Simon D Harding, Christopher Southan, Elena Faccenda, Adam J Pawson, Jamie A Davies and NC-IUPHAR.

Contact information

Deanery of Biomedical Sciences, University of Edinburgh, Hugh Robson Building, George Square, Edinburgh, EH8 9XD, UK

Significance Statement

Online databases contain millions of chemical structures, a fraction of which have been tested for biological activity in the literature. Activity data are often hidden in dense text or supporting material making them difficult to find. The IUPHAR/BPS Guide to PHARMACOLOGY uses an expert-driven process to select the most useful data on drug targets and to recommend ligands for experimental use, presenting the information in a standard database format for easy searching.

ABSTRACT

The IUPHAR/BPS Guide to PHARMACOLOGY is an expert-curated, open-access database of information on drug targets and the substances that act on them. This paper describes the procedures for searching and downloading ligand-target binding data and for finding detailed annotations and the most relevant literature. The database includes concise overviews of the properties of 1,700 data-supported human drug targets and related proteins, divided into families, and 9,000 small molecule and peptide experimental ligands and approved drugs that bind to those targets. More detailed descriptions of pharmacology, function, and pathophysiology are provided for a subset of important targets. The information is reviewed regularly by expert subcommittees of the IUPHAR Committee on Receptor Nomenclature and Drug Classification. A new immunopharmacology portal has recently been added, drawing together data on immunological targets, ligands, cell types, processes and diseases. The data are available for download and can be accessed computationally *via* web services.

Keywords:

Pharmacology, immunology, medicinal chemistry, curated database, drug target, ligand, receptor

INTRODUCTION

The International Union of Basic and Clinical Pharmacology (IUPHAR)/British Pharmacological Society (BPS) Guide to PHARMACOLOGY database (GtoPdb, <http://www.guidetopharmacology.org/>) is unique in receiving expert contributions and review by >500 scientists in 96 subcommittees of the IUPHAR Committee on Receptor Nomenclature and Drug Classification (NC-IUPHAR) (Harding et al., 2018). The resource aims to summarise key information on drug targets and the substances that act on them from literature and present this in a searchable, user-friendly format. The scope of coverage includes any human protein whose activity can be modulated by a small molecule or peptide ligand. Ligands are either approved drugs, natural ligands, antibodies, or other recommended compounds available for experimental use and, where possible, quantitative binding data are provided along with references. Presently, GtoPdb includes ~1,700 data-supported human drug targets and some additional related proteins of emerging interest to drug discovery. These are organised hierarchically by family: Basic Protocol 1 shows how to navigate the family tree to locate a target of interest and find ligands that modulate it. Alternate Protocol 1 shows how to perform a text-based search to find the same target; this is a quick way to find a specific target, or can be useful when the target classification is unknown. Support Protocol 1 introduces the powerful functionality that can be gained by using the GtoPdb-UniProtKB reciprocal links to combine and filter across hundreds of data types and cross-references in UniProtKB (The UniProt Consortium, 2017).

The ~9,000 ligands are organised by chemical class, e.g. synthetic organic compounds and natural products. Basic Protocol 2 shows how to navigate the ligand list to find a drug and locate information about its mechanism of action and known targets. For small molecule ligands and drugs, searching by chemical structure is also possible. Alternate Protocol 2 shows how to use the GtoPdb chemical drawing tool and search function to find ligands by structural similarity. This has a number of applications, for instance, to identify possible targets for a novel structure, or to explore sets of similar compounds. Support Protocol 2 demonstrates how to use the GtoPdb cross-references to 92 million compounds in PubChem (Kim *et al.*, 2016), to find further sets of related structures, to which filters can then be applied; for instance, to find bioactive compounds from the medicinal chemistry literature. The protocol also shows how the results can be pivoted back to the GtoPdb “small-data” resource.

Much of the data are available to download in a choice of formats. Basic Protocol 3 describes how to download CSV files with lists of ligands, targets and interactions. Alternate Protocol 3 gives an introduction to accessing data *via* the web services, which can be used in computer programs to give access to the data in a machine-readable format.

A recent expansion of the database has focussed on new types of data pertaining to immunopharmacology, the immune system and inflammation. GtoPdb targets have been linked to immunological cell types, diseases and processes, accessible *via* a new portal, the IUPHAR Guide to IMMUNOPHARMACOLOGY (GtoImmuPdb, www.guidetoimmunopharmacology.org). Immunity, inflammation and infection are priority areas for drug discovery (Carvalho *et al.*, 2016; Ishii, 2017; Landolina *et al.*, 2016; Tiligada *et al.*, 2015). Many chronic diseases, including aspects of ageing, have immune and/or inflammatory components (Licastro *et al.*, 2005). Auto-immunity is a serious clinical problem (McInnes *et al.*, 2011; O'Shea *et al.*, 2013), with its progression dependent on immune and inflammatory responses (Perry *et al.*, 2003). A main driver of the GtoImmuPdb project is to support the discovery of more efficacious and selective drugs to modulate these types of responses. GtoImmuPdb aims to do this by facilitating a

1 stronger connection between the immunological and pharmacological research communities, with easy access to data in
2 both directions. In particular, pharmacological information has hitherto been difficult for immunologists without
3 cheminformatics expertise to access. The development of GtoImmuPdb, with an immunologist-friendly portal, aims to
4 address this gap in information exchange.
5
6

7
8 The first protocol focussed on GtoImmuPdb, Basic Protocol 4, shows how to navigate the immunological disease data
9 to identify the approved drugs and primary targets for a given disease. It introduces the GtoImmuPdb portal and
10 explains how to view and understand the disease to target association data. Alternate Protocol 4 acts as a companion to
11 Basic Protocol 4 by illustrating how to browse the disease to ligand association data. With both protocols, the aim is to
12 identify approved drugs for a given immunological disease, and illustrate how navigating between disease association
13 data and ligand summary pages provides access to more details about the approved drugs. Support Protocol 3 provides a
14 more general illustration of searching for diseases in GtoPdb/GtoImmuPdb using the search function to find associated
15 targets/ligands.
16
17
18

19
20 Basic Protocols 5 and 6 deal with exploring the GtoImmuPdb cell type (Basic Protocol 5) and process (Basic Protocol
21 6) data. They explain the way GtoImmuPdb maps external ontology terms to the data and how this underpins the data
22 curation and website navigation. The two protocols demonstrate how to browse and view targets associated with cell
23 types and processes, how to navigate from these pages to find further information, and how to interpret the data.
24
25
26

27 Finally, Basic Protocol 7 illustrates the power of GtoPdb curation in connecting literature on biology and bioactive
28 chemistry *via* the triad of GtoPdb-PubMed-PubChem reciprocal linking.
29
30

31 **BASIC PROTOCOL 1**

32 **FIND A DRUG TO MODULATE A KNOWN TARGET**

33
34 The GtoPdb home page is located at <http://www.guidetopharmacology.org/>. There are several ways to locate targets.
35 Clicking on the Targets link in the main navigation menu on any page goes to the targets home page at
36 <http://www.guidetopharmacology.org/targets.jsp>. Mousing-over the Targets menu reveals direct links to the top-level
37 target classes: G protein-coupled receptors (GPCRs), ion channels, nuclear hormone receptors (NHRs), kinases,
38 catalytic receptors, transporters, enzymes and other protein targets. These are also available as quick links from the left-
39 hand side of the home page.
40
41
42
43
44
45

46 Necessary resources

47
48 *Software*

49
50 An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.
51
52

53 ***Browsing the target hierarchy to locate a target of interest***

- 54 1. Begin at the Targets home page (<http://www.guidetopharmacology.org/targets.jsp>).

1
2
3 *Targets are divided into phylogenetically-related classes, with an additional Other Protein Targets class to*
4 *capture targets not classified as one of the other types. Some targets may be classified under more than one*
5 *class, for example some Kinases are also found under Catalytic Receptors.*
6
7

- 8
9 2. Under the section on Ion Channels click on the link “View the list of ion channel families”.

10
11 *Targets are organised into families, whose members tend to be similar in structure and/or share an*
12 *endogenous ligand or family of ligands. By default, the tree loads with two levels expanded. Intermediate level*
13 *branches (indicated by a plus or minus symbol to their left and in black text), can be expanded/condensed by*
14 *clicking on their names. Clicking on the family names of leaf nodes (blue links) takes users to a summary page*
15 *about that family. The whole tree can be expanded by clicking on the “Expand all nodes” button at the top.*
16 *Clicking on the purple “OVERVIEW” link next to top and intermediate level families (such as Ligand-gated*
17 *ion channels) will take the user to a summary page with an overview of the family’s properties and a list of*
18 *subfamilies.*
19
20
21
22

- 23
24 3. In the ion channel family tree under Ligand-gated ion channels, click on “5-HT₃ receptors” to go to their
25 family page (Figure 1).
26

27 *The first section on the family page contains an overview with background information on the family which is*
28 *a useful summary for newcomers to the field. At the top right is a link to a more detailed introduction in the*
29 *style of a review article, introducing important structural, functional and pharmacological aspects.*
30
31
32

33 **Finding ligands known to modulate a target**

- 34
35
36 4. Under the Channels and Subunits section, click on “Show summary” next to 5-HT₃AB to show basic
37 information about this target and key recommended ligands. Ligands are listed according to their action at the
38 target, in this case “Channel blockers”. The affinity and reference are given next to the ligand name, where
39 available (Figure 1).
40
41

42 *In some cases, such as 5-HT₃ receptors, the functional targets can consist of multiple subunits. Multimeric*
43 *targets are listed under the sub-heading “Complexes”, with a separate sub-heading below for “Subunits”. All*
44 *data unless otherwise stated refer to the human proteins. Other species are sometimes included where human*
45 *data are unavailable or significant species differences exist. Genetic information and database links are*
46 *provided for human, mouse and rat.*
47
48
49

- 50
51 5. Click on a ligand name to access further information about the ligand (described in Basic Protocol 2).
52
53
54 6. Return to the 5-HT₃AB summary and click on “More detailed page” to go to the detailed target page with more
55 information on the structure, function, clinical relevance, and pharmacology.
56
57
58
59
60

Detailed pages are provided for a subset of important targets. Information is provided by NC-IUPHAR subcommittees and reviewed and updated as necessary every two years. The data cover a range of different types, including genetics, protein structure, functional assays, tissue distribution, clinically-relevant mutations, associated diseases, and mouse gene knockouts. Bespoke information is also provided for particular target types, such as ion channel conductance, GPCR transduction mechanisms, and NHR target genes. Detailed pages typically contain longer lists of ligands, which include experimentally-used compounds, approved drugs, and endogenous ligands.

7. Scroll down (or click on the link to Agonists in the Contents list) to the tables of ligands, which are divided by ligand action, such as Agonists, Antagonists, Channel Blockers *etc.* (Figure 2). Here, there are several antagonists listed, with a comment below providing further context. These compounds are not listed on the summary view because they are not selective for the 5-HT₃AB receptor, but they are useful for characterising the 5-HT₃ receptor family. Ligands marked with a medicine bottle icon are clinically approved drugs. Another icon (a number one in a bullseye) indicates where this target is thought to be the primary target of the drug (*i.e.* through which the drug exerts its therapeutic action).

Ligands are ordered by affinity type (pK_d , pK_i etc), and then by decreasing potency. Tables can be reordered by clicking on the column headings.

8. Under “Allosteric Modulators” click on the arrow at the right of the row for trichloroethanol. The row expands to reveal further data about the interaction.

ALTERNATE PROTOCOL 1

USE SEARCH TOOLS TO FIND A DRUG TARGET

Basic Protocol 1 is useful if users know how the target they are looking for is categorised in GtoPdb. However, if a user does not know where to find it in the family hierarchy, or simply wishes to go directly to the detailed target page, they may wish to try a keyword search or use one of a number of recognised identifiers to search. This protocol uses the Advanced Target Search function, but most searches can also be done using the quick search box at the top of every page. The Target Search page includes various types of search: text-based, identifier, wildcard matches, literature-based (*e.g.* by PubMed ID), and BLAST sequence searches.

Necessary resources

Software

An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.

1. Go to the home page at <http://www.guidetopharmacology.org/>.
2. Mouse over the “Advanced search” menu of the main navigation bar and click on “Target search”.

- 1
2
3 3. The top section is a keyword-based text search. In the search box type in “5-HT3”. Whilst typing, the search
4 function will automatically try to match the keyword to a list of target and family names and synonyms (Figure
5 3). In the drop-down list click on one of the target names to go directly to its detailed target page, or on the
6 family name to go to the family summary page.
7
8

9
10 *The text search can be further restricted to specific database fields, species or target classes, and the search*
11 *can be completed by clicking on the “Search the database” button.*
12

- 13
14 4. Return to the search page by hitting the back button on the browser.
15
16
17 5. In the third search box paste in the UniProtKB accession P46098. The source database is “UniProt” by default,
18 and this should be changed if a different type of identifier is used. Click “Search to database”.
19
20
21 6. On the results page click on the name “5-HT3A” to go to its detailed target page.
22
23

24 SUPPORT PROTOCOL 1

25 NAVIGATING RECIPROCAL LINKS TO UNIPROT

26
27
28
29 The provision of UniProtKB (Pundir *et al.*, 2015; The UniProt Consortium, 2017) links in GtoPdb for protein targets
30 (and some ligands) offers the utility to users of not only browsing the full entries but also using the protein accession for
31 many types of query. All our human entries are Swiss-Prot (*i.e.* fully annotated and manually expert reviewed) but some
32 rodent and other species are automated TrEMBL records (pending completion of their complete proteomes). Note also,
33 all our ligand-to-target links are manually curated. For each release of GtoPdb we refresh the set of UniProtKB cross-
34 references that point back to us (representing the set of proteins that have a recorded ligand interaction). This reciprocity
35 has powerful advantages, just a few of which can be exemplified here. In essence, users can navigate from ligand space
36 to target space and vice versa. This protocol will also demonstrate how to query across proteins in UniProt and then
37 come “in” to GtoPdb.
38
39
40
41

42 Necessary resources

43 *Software*

44
45 An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.
46
47

48
49 A text editor, such as Microsoft WordPad, for temporary paste-across storage.
50
51

52
53 A program capable of archiving lists, such as Microsoft Excel or LibreOffice Calc.
54
55

56 For set intersecting, the Venny tool is useful (<http://bioinfogp.cnb.csic.es/tools/venny/>)
57
58
59

Getting a UniProtKB list of all the human GtoPdb targets and sorting them

1. Go to the UniProtKB website at (<http://www.uniprot.org/>) and open the “Advanced Query” menu by clicking on the "Advanced" button beside the search bar at the top of the page. You will be presented with the options shown in Figure 4. Set the first selection step as “Cross-references”. Then look for the subset under “Chemistry” and select “GuidetoPHARMACOLOGY”. Executing the query brings up 1999 results; the results can be filtered using the quick links on the left-hand side of the page. Restricting to reviewed (*i.e.* Swiss-Prot only) gives 1995, then selecting Human gives 1483.

The combination of any selects can be copied, stored and re-used as a direct query to paste into the search bar (database:(type:guidetopharmacology) AND reviewed:yes AND organism:"Homo sapiens (Human) [9606]") (when doing this copy all the text in bold), as well as storing/sharing the execution as a URL.

2. For the 1483 entries, select “Columns” and a “Customize results table” page should open. For the example shown in Figure 5 the selected columns are Entry, Entry name, Protein names, Gene names, Length, EC number, Gene Ontology (molecular function) and Subcellular location [CC].

The query and display options in UniProtKB are extensive and cover combinations of many annotation features and keywords, as well as over 100 cross-references that are mainly pointers to other databases. Note, the queries can be combined in Boolean sets and toggling which features and cross-references are shown in the table adds a second detailed level of selection. Yet a third layer of analysis is enabled by downloading the results from the display selections for sorting, sub-setting and ranking in Excel.

3. As an example, find the GtoPdb entries for proteases (with ligands) by combining GtoPdb with MEROPS (proteases and inhibitors database) (Rawlings *et al.*, 2014; Rawlings *et al.*, 2016) in the Advanced query section. Adding in “hydrolase” as a keyword gives a clean listing of 152, as this excludes endogenous protease inhibitor proteins from MEROPS. The storable query has this form (**database:(type:guidetopharmacology) database:(type:merops) keyword:"Hydrolase [KW-0378]" AND reviewed:yes AND organism:"Homo sapiens (Human) [9606]"**). The column display filters should be set as above, except that Gene Ontology can be dropped because “hydrolase” has been added as general function, and Tissue specificity (found under the Expression category) and 3D (found under the Structure category) should be added.

For ease of use, we suggest pasting the above query directly into the search box. If users wish to see how this query can be constructed within the Advanced query menu, clicking on Advanced after running the query will show the selections. The column display filters still need to be set manually, however.

4. Use the Download button to download the result sheet for the 152 proteases to Excel. As the format select Excel and set to Uncompressed. It may be possible to rank the results in Excel by the number of Protein Data Bank (PDB) structures (Rose *et al.*, 2017), but only after manipulating the text in the 3D column to remove the text strings and brackets. Nonetheless, for each protein, the number of 3D structures is given in parentheses after the name of each experimental method. A snapshot ranked by most populated in PDB is shown in Figure 6.

1
2
3 The examples in 2 and 3 indicate the wide flexibility to “slice and dice” the GtoPdb subset via the extensive
4 query options in UniProtKB. This particular sequence has ordered each of the proteases by the number of
5 PDB entries. Unsurprisingly, thrombin ranks top with 383 (but note the entry is technically prothrombin) but it
6 is of interest that the second on the list is BACE1 with 357 because this is not yet an approved drug target. We
7 can look at this in more detail in the next section but some caveats can be pointed out for Figure 6. Firstly
8 there are some false negatives in the form of six missing tissue distribution annotations. These are probably
9 just curation catch-up issues (i.e. these will eventually be added), but in the interim other cross-referenced
10 sources, such as the Human Protein Atlas (Uhlen et al., 2015), will have expression data. The other caveat is
11 that from the PDB counts alone, those with bound ligand structures cannot be filtered from apoprotein
12 structures without ligands (although this splitting can be done via the PDB web portal). Nonetheless, for all
13 the proteases in Figure 6 the vast majority of the PDB entries cross-referenced via UniProtKB will represent
14 (mostly distinct) small-molecule ligands. GtoPdb will only curate those that are both clinically advanced and
15 have documented activity data. The inconvenience of the Excel sheet is the absence of live links in the
16 download. There is a workaround in the form of pasting directly from the UniProtKB result displays into an
17 Excel sheet where the URLs are maintained in their original format.

23 24 ***Navigating from UniProtKB back to GtoPdb***

- 25
26
27 5. Inspect the BACE1 entry at <http://www.uniprot.org/uniprot/P56817> that has the most PDB structures for a
28 protease target without any approved drugs. Using the Display menu on the left-hand side of the page, click on
29 "Cross-references" to jump to that section of content. Scroll down and the cross-reference to GtoPdb (that was
30 used for the query at the beginning of this section) can be found in the “Chemistry databases” section. Click on
31 the GuidetoPHARMACOLOGY ID (2330)
32 (<http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=2330>). This gives the detailed
33 view page for BACE1; from the Contents menu click on Inhibitors (or scroll-down) to navigate to the inhibitor
34 list.

35
36
37
38
39 Note that one of the ligands for BACE1, verubecestat, is used as the PubChem navigation example in Support
40 Protocol 2.

41 42 43 ***Querying across multiple databases to find specific subsets***

- 44
45
46 6. For this example, search for human proteins (with GtoPdb links) containing a domain signature for aspartyl
47 proteases (Asp proteases). Return to the BACE1 UniProtKB entry (<http://www.uniprot.org/uniprot/P56817>)
48 and select "Family & Domains" from the left-hand menu. Open up the InterPro (Finn et al., 2017) graphical
49 display by clicking on “View protein in InterPro” under "Family and domain databases". One of the domain
50 signatures specific for Asp proteases in general is “IPR001461 Aspartic_peptidase_A1”. Back on the
51 UniProtKB site, run a query with the format (**database:(type:interpro id:IPR001461)**
52 **database:(type:guidetopharmacology) AND reviewed:yes AND organism:"Homo sapiens (Human)**
53 **[9606]"**) by pasting this query string into the search box. Note, a protease select is no longer needed since this
54 is implicit within the aspartyl proteases (MEROPS class A1).
55
56
57
58
59
60

1
2
3 *As above, if users wish to see how the query can be constructed, click on the Advanced menu button after*
4 *running it, and check that the cross-references have been set to InterPro and GtoPdb, including the identifier*
5 *for InterPro, with reviewed set to yes and the organism to human.*
6
7

- 8
9 7. Customise the results table as before, with Entry, Entry name, Protein names, Gene names, Length, EC
10 number, and this time include a cross-reference to the rare disease database, Orphanet (www.orpha.net)
11 [Orphanet: an online database of rare diseases and orphan drugs. Copyright, INSERM 1997. Available at
12 <http://www.orpha.net> Accessed (12/10/2017)]. Orphanet is found under the "Organism-specific" category. The
13 results should look similar to Figure 7, which shows the table for seven Asp proteases.
14
15

16 *This shows the use of detailed selects for specific subsets of GtoPdb targets in UniProtKB. Note, if the GtoPdb*
17 *cross-reference was left out of the select, it would bring back 10 Asp proteases since NAPSA, PGA3 and PGA4*
18 *have no curated ligand interactions in GtoPdb. Another example could have used the Beta-secretase BACE*
19 *PRINTS signature as IPR009119, to retrieve just BACE1 and BACE2.*
20
21
22

23 **BASIC PROTOCOL 2**

24 **FIND THE MECHANISM OF ACTION AND KNOWN TARGETS FOR A DRUG**

25
26
27
28
29 GtoPdb ligands are classified into groups based on their chemical properties, including synthetic organic compounds,
30 natural products and their derivatives, peptides that are endogenous in mammals, peptides from other organisms or
31 synthetically derived, metabolites, which includes small biogenic compounds such as hormones and neurotransmitters,
32 antibodies, and inorganic compounds. In addition, ligands have also been grouped by virtue of a common property such
33 as all approved drugs, and compounds labelled with, for example, a radioactive or fluorescent group. All these
34 categories are available to browse *via* the GtoPdb ligand list available by mousing over the Ligands submenu of the
35 main navigation bar. Quick links to these categories are also found on the left panel of the home page. This protocol
36 describes how to find a drug *via* the list and locate information on its mechanism of action and known targets.
37
38
39

40
41 Necessary resources

42
43 *Software*

44
45
46 An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.
47
48

- 49 1. From the home page <http://www.guidetopharmacology.org/>, mouse over the Ligands submenu of the main
50 navigation bar and click on "Ligand list".
51
52
53 2. The ligand list opens at the first tab which is an alphabetical list of approved drugs. Click on the letter "E" to
54 go to drug names starting with that letter.
55
56
57 3. Scroll down and click on the name "erlotinib" to go to its ligand page.
58
59
60

1
2
3 *The ligand page opens with the Summary tab open (Figure 8), which shows the drug approval status,*
4 *nomenclature, and database links.*

- 5
6
7 4. Click on the “Clinical data” tab.

8
9 *Information on this tab summarises the clinical use and therapeutic mechanism of action for the drug erlotinib,*
10 *which targets the epidermal growth factor receptor (EGFR).*

- 11
12
13
14 5. To access specific binding data at targets, click on the “Biological activity” tab (Figure 9). The data presented
15 here confirm EGFR as the primary target of this compound, and this tab also provides information on other
16 proteins erlotinib is known to interact with.

17
18 *Figure 9 also shows a link near the top: “View interactive charts of activity data from ChEMBL and GtoPdb*
19 *across species”. The charts available by clicking on this link provide a method of visualizing ligand activity at*
20 *different targets and easily comparing across species. The tool combines data from GtoPdb and the medicinal*
21 *chemistry database, ChEMBL (Bento et al., 2014), which covers a wider range of species. A table provided*
22 *below the charts lists the source databases, data values and links to the references cited.*

23 24 25 26 **ALTERNATE PROTOCOL 2**

27 28 29 **SEARCH FOR LIGANDS SIMILAR IN STRUCTURE TO A KNOWN DRUG AND FIND THEIR TARGETS**

30
31
32 It is also possible to search for ligands using various methods, which is useful where different isomers exist, the name-
33 structure relationships are complex, or users wish to find ligands similar to a particular structure or substructure. This
34 example demonstrates how to use the ligand structure search tool to find ligands sharing a substructure. It then takes
35 one of the matching ligands and shows how to submit this structure to a further similarity search. The structure drawing
36 tool can be accessed from the “Ligand search” link under both the “Advanced search” and the “Ligands” menus of the
37 main navigation bar. It can also be opened with a ligand structure pre-loaded into the editor by clicking on the structure
38 image on ligand pages (Figure 8). Ligands can also be searched by name (including wildcard searches), or chemical
39 identifier such as PubChem CID, International Union of Pure and Applied Chemistry (IUPAC) International Chemical
40 Identifier (InChI), or the shorter InChIKey.
41
42
43

44
45
46 Necessary resources

47
48 *Software*

49
50
51 An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.

52
53 *Identifiers*

54
55
56 A chemical structure in simplified molecular-input line-entry system (SMILES) format.
57
58
59

Perform a substructure search

1. Go to the home page at <http://www.guidetopharmacology.org/>
2. Mouse over the “Advanced search” menu of the main navigation bar and click on “Ligand search”.
3. Under “Chemical structure search” paste the SMILES string “N(c1ccccc1)c1ncnc2ccccc12” into the box and click “Import SMILES/SMARTS” (Figure 10).

“SMILES” refers to chemical structures described in Simplified Molecular-Input Line-Entry System.

“SMARTS” refers to SMILES Arbitrary Target Specification, a way of defining chemical structures with variable atoms of side groups.

4. The chemical structure loads into the Marvin JS chemical editor, version 14.10.27, 2014, ChemAxon (<http://www.chemaxon.com>).

Chemical structures can also be drawn into the editor by hand, optionally using one of the templates shown at the bottom, and adding bonds with the button on the left and atoms from the list on the right. A user guide to Marvin JS is available at <https://docs.chemaxon.com/display/docs/Marvin+JS+User's+Guide>.

5. In the first drop-down menu choose “Substructure” search. Leave the chemical class set to “Small molecules” and click on “Search the database”. Searches are performed using the Pinpoint chemical cartridge, version 2.4.2, Dotmatics (<https://www.dotmatics.com/>).
6. The search results are shown in a grid and the default ordering is by increasing number of rotatable bonds and molecular weight. Examine the matching structures and notice how they all contain the input structure displayed at the top of the page.
7. Change the order of results by selecting “Alphabetical” in the drop-down menu next to “Order results by”. Click on “Update”.
8. The order of results is now alphabetical by ligand name. Scroll down to find erlotinib.

Perform a similarity search

9. On the erlotinib box click on “Use in search”.
10. The erlotinib structure is loaded into the chemical drawing tool. Change the type of search to “Similarity: moderate (>70%)”. Click on “Search the database”.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

When GtoPdb ligand structures are loaded into the drawing tool, any chiral specification is removed (i.e. chiral bonds are shown in the flat form) so that substructure or similarity searches are not confounded by the presence of different isomers.

11. The results now contain only structures that are at least 70% similar to erlotinib. The targets of each compound are summarised underneath the structures. Click on a ligand name or image to access the ligand page, or on a target name to access the target page.

Similarity searches use the Tanimoto coefficient to compare the similarity between fingerprints generated from chemical structures and return all results above a certain threshold.

SUPPORT PROTOCOL 2

NAVIGATING RECIPROCAL LINKS TO PUBCHEM: SIMILARITY SEARCHING AND SOURCE INTERSECTING

The search for ligands similar in structure to a known drug, as described in Alternate Protocol 2, is internal and consequently encompasses only those entries that GtoPdb has SMILES for. Most of these can be classified as small-molecules but include small to moderate-sized peptides up to our largest example, ligand ID 7387, lixisenatide, with 44 amino acids. Using the database links provided for ligands, users can go “out” to many other databases that are chosen for various types of complementarity (*i.e.* they can add value over and above what our ligand entries offer). In many cases these links are reciprocal, meaning users of the external resource can come “in” to GtoPdb. These external sources are each different and users cannot thus easily become familiar with all of their individual features and content. For bioactive chemistry, PubChem is the most comprehensive, so this will be used as an example. In addition, as explained, it is our primary source in most cases so we make considerable effort to map our ligand IDs to the correct PubChem structures. It is also important to note that we submit not just structures but also PubMed literature links. In addition, PubChem is an entry point (along with those for proteins, genes, PubMed articles, MeSH terms and others) into the NCBI Entrez system (Gibney *et al.*, 2011). We would recommend anyone who regularly needs to engage with small molecule ligands beyond an introductory level in GtoPdb (whatever their background discipline) should become acquainted with PubChem at least to some level. The associated challenge is that just a click instantly takes a user from our modest chemical space neighbourhood of ~7,000 structures, out into one of over 90 million. There is a very wide range of operations that can be executed in PubChem but there is only space here to introduce a few basic protocols. This means there are many features and operation choices in the figures that we cannot expand on here, but these are detailed in the PubChem Help documentation (<https://pubchem.ncbi.nlm.nih.gov/help.html>).

Necessary resources

Software

An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.

A text editor, such as Microsoft WordPad, for temporary paste-across storage.

1
2
3 A program capable of archiving lists, such as Microsoft Excel or LibreOffice Calc.

4
5 For set intersecting, the Venny tool is useful (<http://bioinfogp.cnb.csic.es/tools/venny/>).

6 7 8 *Identifiers*

9
10 PubChem substance identifier (SID) and PubChem compound identifier (CID)

11 12 13 ***Navigating across to PubChem records***

- 14
15
16 1. Find the ligand entry for “verubecestat” by searching for its name in the search box at the top right of the
17 GtoPdb site, or go directly to
18 <http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?ligandId=8699>.

19
20
21
22 *This is a Merck development compound against BACE1 for Alzheimer’s disease (AD), originally declared as*
23 *MK-8931 but now with the International Nonproprietary Name (INN) verubecestat. The first Phase 3 clinical*
24 *study has been stopped because of lack of efficacy. Note, the searches below can be carried out with any of the*
25 *other 16 ligand entries for BACE1.*

- 26
27
28 2. Check the Similar ligands tab. This produces just one match (within GtoPdb) with a compound (example 20)
29 from a patent (ligand ID 7853).

30
31
32 *This comes from a Roche patent (WO2010128058) for BACE2 inhibitors claimed for diabetes. As the*
33 *paralogue of BACE1, GtoPdb includes nine BACE2 inhibitor ligands*
34 *(<http://www.guidetopharmacology.org/GRAC/ObjectDisplayForward?objectId=2330>). Significantly, potent*
35 *verubecestat inhibition is common to both targets.*

- 36
37
38 3. Click back to the Summary tab. As shown in Figure 11, click on the link to PubChem CID within the Database
39 Links section. This will now open the link <https://pubchem.ncbi.nlm.nih.gov/compound/51352361> in a new
40 tab.

41
42
43
44 *Many links, features and statistics are shown on the CID home page. Their density is roughly proportional to*
45 *compound “popularity”. Generally, older approved drugs will have many pages whereas patent-only*
46 *compounds have less information available. Verubecestat has a middle level of popularity with facet 5.3*
47 *(“Substances”) indicating 21 submissions of the same structure as SIDs from different sources (including*
48 *GtoPdb).*

49 50 51 52 ***Using PubChem pre-set similarity search and clustering***

4. Use the Contents link “5 Related Records” to navigate the CID entry to reach the 5.2 facet for “Related Compounds” and the subheading “Similar Compounds” (Figure 12). Then click on the highlighted “526 records”. The results, features and options for this page are shown in Figure 13.

Clicking “Similar Compounds” executes, on the fly, a 90% Tanimoto similarity search. This actually shows 684 results (which is presumably because the search executes as a live-link, but the number shown, 526, was pre-computed and more compounds have since been added). It is important to note that % similarity settings (Tanimoto-based or by other methods) are not identically parameterized between different databases. This means the GtoPdb and PubChem structural “neighbours” may not give the same results but there are clearly likely to be more in the latter. Many filters and sub-sets can be accessed via the right-hand side facets shown in Figure 13.

5. Click on the “Structure Clustering” logo near the top right (Figure 13). This will generate the tree result seen in Figure 14. Note, it is sometimes necessary to expand the tree to get the full display. To do this, click on the image on the blue dots between the two blue lines representing the top and bottom branches of the tree. On the pop-up links, click on “Expand Subtree”.

There are many options that can be explored on this PubChem clustering display. One of those shown in Figure 14 is “Common substructure” that can be toggled in for any branch of the tree by clicking on the branch node blue spots. Note also the entire scoring matrix can be downloaded using the Export links.

Narrowing the search, 3D display and substructure picking

6. Return to the list of 684 similar compounds to verubecestat (CID 51352361).

Triaging 684 similar structures to CID 51352361 as a well-populated chemotype is a challenge. It can be established, by checking intersects with PubChem sources such as SureChEMBL (Papadatos et al., 2016) and/or IBM (<https://www.ibm.com/>), that most of these are patent-extracted structures from the many pharmaceutical companies (and some academic groups) working on BACE1 inhibitors. Filtration can of course be informative. As an example, we can select (from the right facet) “BioAssays, Active”, to view 17 entries where the activity has been extracted from published papers by ChEMBL (Bento et al., 2014). Two follow up approaches can be taken here, the first is simply to narrow down the search stringency to get a smaller set of matches. This can then be 3D clustered as an adjunct approach to 2D. The second option is to use substructure searches. While this may expand the hit list in PubChem, the protocol will show how to transfer this across to the smaller search space of GtoPdb.

7. As an alternative to the pre-set 90% similarity search used above, for any individual CID it is possible to set up a customised search as shown in Figure 15. From the PubChem Compound home page (<https://pubchem.ncbi.nlm.nih.gov/>), click on Structure Search on the right, and then on the tab Identity/Similarity. Underneath, click on the tab for CID, SMILES/SMARTS, InChI. Paste “51352361” into the box, and edit the Options to match those in Figure 15. The result display is the same as in Figure 13 but the result numbers in stringency order are: 93% = 288, 95% = 151, 96% = 110, 97% = 74. In the same way as

1 performed above, run the clustering for the 74, but this time set the display for 3D rather than 2D. The result is
2 shown in Figure 16 and gives good shape-based indications on similarities within the series. Note that in most
3 cases it is possible to find SAR data specified in the patent documents that the structures were extracted from.
4
5

- 6
7 8. Mouse over the red branch lines to display the “common substructure” for a set of ligands (an example of
8 which is shown in Figure 14). A SMILES representation of this substructure can be obtained by using either
9 the PubChem sketcher or MarvinJS in GtoPdb (described in Alternate Protocol 2) to edit the structure for
10 verubecestat down. The example is represented by the SMILES string
11 FC2=C(C)C=C(N(C(=O)C1=NC=CC=C1)[H])C=C2. On the PubChem Structure Search, click on
12 Substructure/Superstructure, then on CID, SMILES/SMARTS, InChI, and paste in the SMILES string. At the
13 time of writing, the example returns 6160 CIDs, and these can be filtered for structures of particular interest.
14
15

16
17 *A link to the PubChem sketcher can be found under the Draw a Structure tab, which is next to the CID,*
18 *SMILES/SMARTS, InChI tab.*
19
20

- 21
22 9. Go to the GtoPdb Chemical structure search at <http://www.guidetopharmacology.org/GRAC/chemSearch.jsp>.
23 Paste FC2=C(C)C=C(N(C(=O)C1=NC=CC=C1)[H])C=C2 into the input box, click “Import
24 SMILES/SMARTS” and hit “Search the database”. The result is shown Figure 17.
25
26

27 *Users can transition any standardised query between PubChem and GtoPdb. In this case we are running the*
28 *equivalent substructure search from PubChem described above to search against the smaller, highly curated*
29 *ligand set of GtoPdb. Here we see the complementarity of working between the small-data and big-data*
30 *resources by exploiting different search stringencies, filtration and display options. In the PubChem 6160*
31 *results, clicking on “BioAssays, Active” reduces the number to 339, while clicking on “Protein 3D Structures”*
32 *reduces this to 17, and 16, respectively. In the GtoPdb results we can see the substructure matches both*
33 *BACE1 and BACE2 inhibitors.*
34
35
36
37

38 **BASIC PROTOCOL 3**

39 **LEARN HOW TO DOWNLOAD DATA AND VIEW IT IN MICROSOFT EXCEL**

40
41 GtoPdb includes various file download and data access options. This includes data files for targets, ligands and
42 interactions in comma-separated value (CSV) format. These represent the most requested datasets; other datasets and
43 file formats are available by contacting the team (enquiries@guidetopharmacology.org). There is also the option of
44 downloading an SQL file containing the whole PostgreSQL database (<http://www.postgresql.org/>), which would require
45 a recent version of PostgreSQL to be installed. Basic Protocol 3 explains how to download a CSV file and open it with a
46 program capable of viewing this file format. GtoPdb recently added web services to facilitate computational access to
47 the data, and Alternate Protocol 3 describes how to build the URLs needed to access data on targets, ligands and
48 interactions in a computer program.
49
50
51
52
53

54
55 Necessary resources
56
57
58
59
60

Software

An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.

A program capable of viewing CSV format files, such as Microsoft Excel or LibreOffice Calc.

1. Go to the download page at <http://www.guidetopharmacology.org/download.jsp>.
2. Scroll down to the section called “Download data files”.

A description of each data file can be accessed by clicking on the link above the file list. This is available at http://www.guidetopharmacology.org/DATA/file_descriptions.txt. Users are encouraged to read this document whilst looking at the downloaded files for the first time as it describes the content of each column. This document will be updated if the file format changes.

3. Click on the “Download CSV file” link next to “target and family list”.
4. Either save the file or select the program to open it with. If it is being opened with LibreOffice a Text Import wizard will open. Select “Separated by” and “comma” under Separator Options. With Microsoft Excel, if the file doesn’t open automatically, it can be saved and then loaded using the Open command under the File menu, choosing “Text Files (*.csv)” as the file type. Alternatively, if Excel does not open the CSV file correctly, try importing the file using the “From text” option in “Get External Data” under the Data menu. There, select “Delimited”, and set the file origin as “Unicode (UTF-8)”. In the second step, select “Comma” as the delimiter. In the third step, the column type can be set, which can be useful to prevent Excel from changing certain numbers (such as many identifiers) to scientific format or even to date format (to do this click on the column and choose format “Text”).

Some users have reported problems opening CSV files with Microsoft Excel 2016 where their default list separator is a semicolon. In this case, all text after a semicolon will be lost. If this has happened, there will be no text in any of the columns after column 1 (“synonyms”) for several rows of the target and family CSV file (see, for example, rows 3 and 5). This is because synonyms sometimes contain semicolons as part of the HTML codes for special characters such as the Greek alpha, beta etc.. The list separator can be changed under Windows in “Control Panel / Region / Additional Settings / List separator:”.

ALTERNATE PROTOCOL 3

USE WEB SERVICES TO GAIN COMPUTATIONAL ACCESS TO TARGET AND LIGAND DATA

GtoPdb offers a range of REST (Representational State Transfer) web services described in full at <http://www.guidetopharmacology.org/webServices.jsp>. Requests to web services will return data in JavaScript Object Notation (JSON) format, which is both easily understood by humans and machine-readable. REST APIs (Application

1
2 Programming Interface) use HTTP to transfer data to clients, providing a simple way to incorporate external data in
3 applications.
4

5
6 The GtoPdb API includes options to retrieve lists of targets, ligands, interactions, and references, as well as data for
7 individual entries. Various parameters can be included in the URL which have the effect of filtering data in various
8 ways. For instance, interaction lists can be filtered by ligand action (*e.g.* agonist, inhibitor), affinity type (*e.g.* pK_i, pK_d),
9 target type (*e.g.* GPCR, NHR), approved drugs, or minimum affinity value. Ligand lists can be filtered by chemical
10 property (*e.g.* number of rotatable bonds or hydrogen bond acceptors), type (*e.g.* synthetic organic, antibody), or
11 searches performed by name, external database identifier, or SMILES string (in the same way as described in Alternate
12 Protocol 2). Targets can be filtered by type or searched by name, gene symbol, Enzyme Commission (EC) number, or
13 another database identifier. Alternate Protocol 3 explains how to build URLs to fetch data for sets of targets, ligands and
14 interactions, and includes examples for accessing specific data on individual targets and ligands. Further examples are
15 provided at <http://www.guidetopharmacology.org/webServices.jsp#examples>.
16
17
18
19

20 Necessary resources

21 *Software*

22
23 An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.
24
25

26
27 (Optional) a tool for optimally viewing JSON format data (modern Web browsers can display JSON data, but it can be
28 helpful to view it in a “pretty” format). For example, the Firefox add-on JSONView (<https://jsonview.com/>) pretty-
29 prints JSON data and provides collapsible array and object sections.
30
31
32

- 33
34 1. All URLs are built by extending the base URL: <http://www.guidetopharmacology.org/services/>
35

36
37 *Note, accessing the base URL itself doesn't display anything.*
38

39 *Accessing target information*

- 40
41
42 2. View the target web services documentation at <http://www.guidetopharmacology.org/webServices.jsp#targets>,
43 which lists all the services that are available for targets and families.
44
45
- 46
47 3. To access target data append “targets” to the base URL. Thus, the full target list is available at
48 <http://www.guidetopharmacology.org/services/targets>. To restrict the list to GPCRs, use
49 <http://www.guidetopharmacology.org/services/targets?type=GPCR>. Type can be one of: 'GPCR', 'NHR',
50 'LGIC', 'VGIC', 'OtherIC', 'Enzyme', 'CatalyticReceptor', 'Transporter', 'OtherProtein', or 'AccessoryProtein'.
51
52

53 *In this case, the URL contains a parameter (“type”) and a value (“GPCR”), which are added after a question*
54 *mark at the end of the URL. Multiple parameters can be included by separating them with an ampersand*
55 *character. The returned JSON results include basic information on each target, covering GtoPdb ID,*
56
57
58
59

1
2 *nomenclature, family IDs, and the ID of any subunits this protein has, or complexes it forms part of (Figure*
3 *18).*

- 4
5
6 4. Search the target list by appending one of a few different parameters to the end of the URL. All possible
7 parameters are listed in the documentation, but one useful example is to do a UniProtKB accession look-up.
8 This is achieved using a URL of the form
9 <http://www.guidetopharmacology.org/services/targets?database=UniProt&accession=P08908>. For UniProt
10 searches, the inclusion of the “database=UniProt” part (and thus the ampersand) is redundant, because UniProt
11 is the default database used for accession searches. However, the database name for any other type of
12 accession look-up would need to be specified and a full list of accepted names is given in the documentation:
13 <http://www.guidetopharmacology.org/webServices.jsp#databases>.

14
15
16
17 *UniProtKB accession searches can normally be done using any one of the human, mouse or rat SwissProt*
18 *accessions which form the triad normally mapped to each GtoPdb target (if available). Therefore, any of the 3*
19 *accessions (P08908, Q64264, P19327) would return the same result: the 5-HT_{1A} receptor (target ID 1).*

- 20
21
22
23 5. Access specific targets by appending the target ID to the URL. For example, for target ID 1 use
24 <http://www.guidetopharmacology.org/services/targets/1>. The JSON result shows a summary of the target data,
25 which is similar to the list view.
26
27
28 6. Access further information on an individual target by appending the relevant service name after the target id.
29 For example, to access information on the interactions that target ID 1 is involved in use
30 <http://www.guidetopharmacology.org/services/targets/1/interactions>.

31
32
33
34 *The interaction data and references are listed in full, and empty fields will usually have either a null value, or*
35 *an empty string. Affinity data are reported as $-\log_{10}$ values, and raw data (where known) are reported as the*
36 *molar values using HTML codes for superscripts to specify the power value. If the value for “conciseView” is*
37 *true, this interaction has been selected to show on the website summary view of the target (described in Basic*
38 *Protocol 1 and Figure 1).*

- 39
40
41
42 7. Experiment with adding different service names to the end of the URL. For example, target ID 1’s functional
43 data can be retrieved using <http://www.guidetopharmacology.org/services/targets/1/function>.
44
45
46 8. Access family-specific information in a similar manner by adding “targets/families” to the base URL. Thus,
47 <http://www.guidetopharmacology.org/services/targets/families> gives the complete family list in GtoPdb. This
48 can also be filtered by type as described in step 3. Try obtaining a list of Enzyme families with the URL
49 <http://www.guidetopharmacology.org/services/targets/families?type=Enzyme>.

50
51
52
53 *The returned results include each family’s ID, name, the target member IDs, and the family IDs of any parent*
54 *families or subfamilies in the family hierarchy, which models the display of families on the website (see Basic*
55 *Protocol 1).*

Accessing ligand information

9. Access a description of the ligand web services at <http://www.guidetopharmacology.org/webServices.jsp#ligands>.
10. Ligand URLs are constructed in a similar way to target URLs; to access ligand data append “ligands” to the base URL. The full list of ligands is available at <http://www.guidetopharmacology.org/services/ligands> and this can be filtered and searched in several ways. For example, retrieve a list of natural products by adding the type parameter: [http://www.guidetopharmacology.org/services/ligands?type=Natural product](http://www.guidetopharmacology.org/services/ligands?type=Natural%20product).

The type parameter can be used with any of the following options: 'Synthetic organic', 'Metabolite', 'Natural product', 'Endogenous peptide', 'Peptide', 'Antibody', 'Inorganic', 'Approved', 'Withdrawn', 'Labelled', or 'INN'. These correspond to the subsets available via the website ligand list (see Basic Protocol 2), except for “INN” which is an additional list of all GtoPdb ligands with INNs, and “Withdrawn”, which is a set of previously approved, but now withdrawn drugs (at least by some jurisdictions; they may still be available in other countries).

11. The ligand list is searchable in a similar way to that for targets described in step 4. However, there are a few differences, as described in the documentation. This example finds ligands by InChIKey. Compare the results for a full InChIKey look-up at <http://www.guidetopharmacology.org/services/ligands?inchikey=USQOVYLRWBOSQC-OAHLLOKOSA-N> with a “backbone” look-up at <http://www.guidetopharmacology.org/services/ligands?inchikey=USQOVYLRWBOSQC>.

The first URL uses the full InChIKey for exact matching, while the second uses the first 14 characters of the InChIKey to perform a backbone match (i.e. matches to both the isomers).

12. Ligands are also searchable by SMILES strings based on exact match, substructure or similarity in a similar way to that described in Alternate Protocol 2. For example, type the following URL to look for ligands similar to the structure of 5-fluorouridine (ID 4614): [http://www.guidetopharmacology.org/services/ligands/similarity?smiles=OC\[C@H\]1O\[C@H\]\(\[C@@H\]\(\[C@@H\]1O\)O\)n1cc\(F\)c\(=O\)\[nH\]c1=O](http://www.guidetopharmacology.org/services/ligands/similarity?smiles=OC[C@H]1O[C@H]([C@@H]([C@@H]1O)O)n1cc(F)c(=O)[nH]c1=O). This uses the default minimum of at least 85% similarity, which can be changed by adding the “similarityGt” parameter and specifying a value. Try changing this to 75% to see how it affects the results by typing this URL: [http://www.guidetopharmacology.org/services/ligands/similarity?smiles=OC\[C@H\]1O\[C@H\]\(\[C@@H\]\(\[C@@H\]1O\)O\)n1cc\(F\)c\(=O\)\[nH\]c1=O&similarityGt=75](http://www.guidetopharmacology.org/services/ligands/similarity?smiles=OC[C@H]1O[C@H]([C@@H]([C@@H]1O)O)n1cc(F)c(=O)[nH]c1=O&similarityGt=75).

13. An individual ligand entry is obtainable by adding the ligand ID at the end of the URL. As an example, retrieve the ligand with ID 4789 (5-fluorouracil) with the URL <http://www.guidetopharmacology.org/services/ligands/4789>. Looking at the results for this query, notice that there are two values under the field “prodrugIds”. These are the ligand IDs for the prodrugs capecitabine (ID 6799) and floxuridine (ID 4801), which are enzymatically converted into the active metabolite 5-fluorouracil.

1 Retrieve further data on these prodrugs by using their IDs in the ligand URL, i.e.

2 <http://www.guidetopharmacology.org/services/ligands/6799> and

3 <http://www.guidetopharmacology.org/services/ligands/4801>.

- 4
5
6
7 14. Additional data types for ligands are accessible by appending the relevant web service name to the URL. For
8 example, access the interaction list for ligand ID 5000 (human IL-9) using
9 <http://www.guidetopharmacology.org/services/ligands/5000/interactions>. Note that one of the interaction
10 results (interaction ID 83153) has a zero value in the “targetId” field, but 5000 in the “ligandAsTargetId” field
11 which indicates that this ligand is actually the target of the interaction (by ligand 8934).
12
13
14

15 *Accessing interactions*

- 16
17
18 15. View the interaction web service documentation at
19 <http://www.guidetopharmacology.org/webServices.jsp#interactions>.
20
21
22 16. Access a list of interactions for approved drugs by using the “approved” parameter
23 (<http://www.guidetopharmacology.org/services/interactions?approved=true>). The full list of interactions is
24 available at <http://www.guidetopharmacology.org/services/interactions>, however, note this list is very long so
25 can take a long time to load. There are several ways to filter the list, such as by target type (e.g. for GPCRs use
26 <http://www.guidetopharmacology.org/services/interactions?targetType=GPCR>) or by ligand action (e.g. for
27 channel blockers use [http://www.guidetopharmacology.org/services/interactions?type=Channel blocker](http://www.guidetopharmacology.org/services/interactions?type=Channel%20blocker)).
28 Interactions can also be filtered directly by ligand structure using SMILES or InChIKeys.
29
30
31

32
33 *The interaction list results include a summary view of the interaction data for a target-ligand pair. The*
34 *individual data point interaction IDs are given in the “dataPointInteractionIds” field. These can then be used*
35 *to obtain the full data about an interaction by putting the interaction ID at the end of the URL. For example,*
36 *interaction ID 1697 is retrievable at <http://www.guidetopharmacology.org/services/interactions/1697>.*
37
38

39 **BASIC PROTOCOL 4**

40 41 42 **BROWSE TO FIND APPROVED DRUGS AND THEIR PRIMARY TARGETS FOR AN IMMUNOLOGICAL** 43 **DISEASE** 44 45

46 A major data extension of the Guide to IMMUNOPHARMACOLOGY (GtoImmuPdb) has been to associate targets and
47 ligands to diseases of immunological relevance. By doing this, users can search or browse by disease to identify the key
48 targets and ligands involved. Information about disease associations can be readily accessed from the GtoImmuPdb
49 portal home page, *via* the Disease panel. Disease data can also be accessed from anywhere on GtoImmuPdb from the
50 Disease menu of the main navigation bar. The GtoImmuPdb disease association information contains curator
51 comments, giving details about why an association exists between a given disease and targets or ligands, and any
52 relevant literature references.
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Wherever possible, disease terms are mapped to external disease resources. Disease mapping is a manual process identifying equivalent terms in external resources, adding the unique identifiers and using these to hyperlink to the resources. For example, the disease chronic lymphocytic leukemia (CLL) maps to the Disease Ontology (DOID:1040) term, "chronic lymphocytic leukemia", the OMIM (151400) term "Leukemia, chronic lymphocytic; CLL"; and the Orphanet (ORPHA67038) term "B-cell chronic lymphocytic leukemia". Mapping disease terms to other resources in this way is important for several reasons. Firstly, it provides clarity in understanding which disease is being referred to. In many cases, the same disease may be named differently in different papers, studies or data resources. Secondly, the mappings aid interoperability, meaning that data can be more easily linked between different resources. Finally, it enables us to provide external links allowing users to quickly navigate to other resources of disease information, bringing added value to the data. For example, GtoImmuPdb does not currently include detailed disease definitions. For CLL, clicking the external link to OMIM provides a disease description, details of clinical features, pathogenesis, and more. Linking to Orphanet also gives a disease description, and other external disease identifiers to ICD and MeSH (for example), while the Disease Ontology provides a hierarchical, ontological view of the disease, placing it into context with similar disease classifications.

The GtoImmuPdb disease section presents a list of diseases associated to targets and ligands. For targets, it also displays approved drugs that interact with the target (where the target is considered a primary target), For ligands, it includes summaries of their clinical use and distinguishes the approved drugs from those under development. The aim here is to provide a useful summary of the molecular targets and key ligands that modulate a disease, without requiring a user to navigate away from the disease page. Further details about the individual targets and ligands can be found on their database pages by clicking on their respective names.

Necessary resources

Software

An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.

Finding known targets and drugs for an immunological disease (example, Chronic lymphocytic leukemia (CLL)).

1. Begin at the GtoImmuPdb portal home page (<http://www.guidetoimmunopharmacology.org>) (Figure 19).

The GtoImmuPdb portal shares a similar style with GtoPdb, but has its own header, menu bar, colour scheme and logo to distinguish it. There are main panels that can be used as routes to browse for process, cell type, target, ligand, and disease data. GtoImmuPdb uses the same underlying database as GtoPdb, so users can find the same data in both resources. When viewing pages via GtoImmuPdb, however, different browsing options are available, data of immunological relevance are highlighted, and are also prioritised in searches.

2. Under the section Disease, select "Immuno Disease to Target Associations". Alternatively, select "Immuno Disease to Targets" from the "Disease" drop-down menu of the navigation bar.

1
2
3
4
5
6
7
8
9
10
All immunological diseases which are associated to targets in the database are listed, alphabetically, in the table (Figure 20). Each section contains the details for one disease. By default, the details of the associated targets are hidden (in order to provide an at-a-glance view of the diseases when the page first loads). Details for a given disease can be shown by clicking the "more details" link. The two grey buttons in the top right can be used to expand (show) or hide the details for every disease. For each disease, the total number of curated target associations is displayed, showing the names of the first few targets.

3. Scroll down the table to find the disease of interest, "Chronic lymphocytic leukemia".

11
12
13
14
15
16
17
18
19
20
21
The row for Chronic lymphocytic leukemia (CLL) shows "Disease X-Refs"; these are cross-mappings to other disease resources, in this case OMIM, Orphanet and the Disease Ontology for CLL. The importance of these mappings is explained in the introduction to this protocol. Clicking on the ID next to OMIM (151400) opens a new tab displaying the OMIM entry for #151400, "Leukemia, Chronic Lymphocytic; CLL". OMIM pages typically provide details about the disease, its clinical features, genetics, and more.

4. The summary shows that there is only 1 target associated with CLL. Click on the "more details" link to see full information about this target association (Figure 20).

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
In the expanded section it shows the target name(s) associated with the disease and curator comments about the association. The comments indicate that CD20 is a molecular target of ofatumumab and obinutuzumab, both of which are noted as CLL therapeutics. It also includes a list of ligands that interact with the target, where it is considered to be a primary target. There are four ligands listed against the primary target of CD20, which are associated with CLL. (Note, there may be other ligands curated as being associated with CLL not listed here, but their target would be different to CD20 and its association with CLL may be unclear, which is why it doesn't appear in the list. Other CLL ligands not associated to CLL targets would be included under the Ligands tab, as described in Alternate Protocol 4.) The icons next to the ligand name indicate whether the ligand is an approved drug (medicine bottle) or if the target is a primary target of the ligand (number 1 in a bullseye).

5. Click on the ofatumumab link in the ligand column to view the GtoPdb ligand summary page for ofatumumab. The link opens on the Clinical data tab which reveals further information about the ligand's therapeutic associations.

39
40
41
42
43
44
45
46
47
48
49
Clicking through to the ligand summary allows users to look more closely at what is known about the ligand. The GtoPdb ligand summary page presents data about the immunopharmacology ligands under five main tabs: Summary, Biological activity, Clinical data, References, and Immunopharmacology. The clinical use summary under the Clinical data tab provides details about ofatumumab's use to treat CLL.

6. Click on the Biological activity tab for ofatumumab to view a list of targets for this drug. The data show that ofatumumab is an antibody only known to interact with one target: CD20.

50
51
52
53
54
55
56
57
58
59
60
The data table includes the recorded affinity of the interaction shown on the $-\log_{10}$ scale, and the reference(s). Clicking the arrow on the right gives a bit more information about the assay including the raw affinity data.

ALTERNATE PROTOCOL 4

BROWSE TO FIND APPROVED DRUGS FOR AN IMMUNOLOGICAL DISEASE VIA LIGAND-DISEASE ASSOCIATIONS

Basic Protocol 4 takes a route whereby the user browses a list of target-disease associations, looking to identify approved drugs that modulate these targets for a known immunological disease. Similar data can be accessed by browsing lists of ligand-disease associations as an alternative, which is described in this alternate protocol.

Necessary resources

Software

An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.

Finding known drugs for an immunological disease (example, Chronic lymphocytic leukemia (CLL)).

1. Begin at the GtoImmuPdb portal home page (<http://www.guidetoimmunopharmacology.org>) (Figure 19).
2. Under the section Disease, select “Immuno Disease to Ligand Associations”. Alternatively, select "Immuno Disease to Ligand" from the "Disease" drop-down menu of the navigation bar.

All immunological diseases which are associated to ligands in the database are listed, alphabetically, in the table. Each section contains the details for one disease. By default, the details of the associated ligands are hidden (in order to provide an at-a-glance view of the diseases). Details for a given disease can be shown by clicking the "more details" link. The two grey buttons in the top right can be used to expand (show) or hide the detailed sections for every disease. For each disease, the total number of curated ligand associations is displayed, showing the names of the first few ligands. Users may notice that there are more disease-ligand associations than there are disease-target associations (described in Basic Protocol 4). This is because, in some cases, the actual therapeutic target and mechanism of action for the (potential) drug against the disease has yet to be confirmed, or information is not readily available in the literature or clinical trial databases.

3. Scroll down the table to find the disease of interest, “Chronic lymphocytic leukemia”.

The row for Chronic lymphocytic leukemia shows “Disease X-Refs”; these are cross-mappings to other disease resources, OMIM, Orphanet and the Disease Ontology, for the same disease. The number of ligands associated to the disease is shown, with the names of the first few.

4. Click on the “more details” link to expand the information on ligand associations to Chronic lymphocytic leukemia (Figure 21).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The expanded section for CLL shows the ligand name(s) associated with the disease in the first column. If the ligand is an approved drug this is indicated by the small medicine bottle icon and text stating which regulatory body approved the drug. Curator comments about the disease-ligand association are shown in the second column, and the ligand's clinical use comment is summarised in column 3. References are listed in column 4.

5. Observe information on ligand associations with CLL.

When compared to the list of ligands against target CD20 which is associated with CLL as described in Basic Protocol 4, there are additional ligands listed here. Idelalisib, ibrutinib and milatuzumab are listed in addition to the four ligands for CD20. This means that they are associated with CLL, but don't interact with CD20 (as a primary target). The information shows that idelalisib and ibrutinib are additional approved drugs for CLL, and milatuzumab is an orphan drug for CLL undergoing a phase 2 trial.

6. Click on the idelalisib link in the Ligand column to view the ligand summary page for idelalisib. The link opens at the Immunopharmacology tab which shows all immunological disease associations for this ligand.
7. Click on the clinical data tab to find more information about idelalisib's status as an approved drug for CLL in combination with rituximab (Figure 22). Click on the biological activity tab to find data for PI3K δ as the primary target of idelalisib, as well as other secondary targets.

SUPPORT PROTOCOL 3

USE SEARCH TOOLS TO FIND TARGETS AND LIGANDS ASSOCIATED TO A DISEASE

Basic and Alternate Protocols 4 describe how to find targets and ligands associated to an immunological disease, *via* the GtoImmuPdb portal. Many diseases are also described in the main GtoPdb database, but there is no specific disease listing page through which to access this information directly. Many targets contain related information in their detailed pages, covering disease-causing mutations, drugs that have been developed to target them, rodent models of diseases, and changes in gene expression in disease states. Drugs that have been approved to treat a disease and those showing potential in trials will have clinical use comments mentioning the disease on their summary pages. The easiest way to locate information on targets and ligands related to a specific disease in GtoPdb is to search for it by name. This protocol shows an example of searching for Alzheimer's disease, and explains how to interpret the results.

Necessary resources

Software

An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.

1. Go to the home page at www.guidetopharmacology.org/.
2. In the search box in the top right, enter the term "Alzheimer's" and click "Search Database".

1
2
3 *The search function can accept various different spellings of this term: Alzheimer's, Alzheimers and Alzheimer*
4 *all work. "Alzheimer's disease", however, brings back 5 fewer results, as sometimes the term "disease" does*
5 *not appear next to "Alzheimer's".*
6
7

- 8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
3. Examine the results of the search. The results page shows 109 results, mainly matching to targets and ligands in the database (Figure 23). The term is also found in a few family overviews and introductions (for example, muscarinic acetylcholine receptors on page 5). Each page shows 10 results. Move between pages by clicking on the numbers at the bottom of the page.

The results are ranked by how many times they match the search term and how common the term is within a particular database field. Additional weightings are also applied if a term matches to a target or ligand name or synonym, so these results appear near the top of the list. The order of results can be changed by clicking on the drop-down menu in the top right. The same search can also be run from the GtoImmuPdb home page at <http://www.guidetoimmunopharmacology.org>. The number of returned results is the same, but the order is slightly different, because targets and ligands that are relevant to immunopharmacology are up-weighted so they appear higher up the list.

4. Look at the matches under the target GPR3. The term has matched to text in the database tables "Consequences of altering gene expression" and "Mutations/pathophysiology". More information can be found by clicking on the highlighted target name to go to its database page.
5. Narrow the search to just targets by clicking on the "Target search page" link at the bottom.
6. Enter "Alzheimer's" into the Target text search box and click "Search the database".

The results can be narrowed to certain database fields by selecting one or more field names in the list.

7. Examine the results. There are 30 matches to target pages, which includes the family overviews and introductions.

A similar search can be conducted just for ligands by clicking on "Ligand search page" at the bottom of the result list (or going to Advanced search > Ligand search in the navigation bar) and entering "Alzheimer's" in the Ligand name/text search box.

BASIC PROTOCOL 5

BROWSE TO EXPLORE DATA FOR IMMUNOLOGICAL CELL TYPES

The GtoImmuPdb project has enriched GtoPdb with data on immunological cell types. Targets have been mapped to several top-level immunological cell-type categories, providing a basis for organising, navigating and searching GtoPdb data in ways that are intuitive to immunologists. Cell type data are accessible from the GtoImmuPdb portal home page,

1 via the Cell Types panel. They can also be reached from anywhere on the GtoImmuPdb site using the Cell Types menu
2 in the main navigation bar. This protocol introduces how to browse and view cell type data and link this to
3 pharmacological targets.
4

5
6
7 Necessary resources

8
9
10 *Software*

11
12 An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.
13
14

15 ***Exploring cell type data in the Guide to IMMUNOPHARMACOLOGY***

- 16
17 1. Begin at the GtoImmuPdb portal home page (<http://www.guidetoimmunopharmacology.org>) (Figure 19).
18
- 19
20 2. Under the section Cell Types, select “View immuno cell types home page”. Alternatively, click on Cell Types
21 in the menu-bar, or directly navigate to
22 <http://www.guidetoimmunopharmacology.org/GRAC/CelltypesForward>.
23
24

25
26 *This page describes the top-level cell type categories used in GtoImmuPdb (Figure 24). Quick links at the top*
27 *jump to the start of each cell type category section on the page. For each category, there is a brief description,*
28 *alongside the Cell Ontology (CL) terms included under that category. Cell type data in GtoImmuPdb are*
29 *underpinned by the use of CL terms (<http://obofoundry.org/ontology/cl.html>) (Bard, 2005, genome biology 6)),*
30 *comprising a set of cell type terms and identifiers that are mapped to top-level categories. Using the CL*
31 *provides a controlled vocabulary and external identifiers, which supports interoperability of the data. Another*
32 *benefit is that it supports high-resolution annotation of cell types to targets while also enhancing searches*
33 *across cell type data. This means users can search using parent or child terms (which can be specific cell type*
34 *terminology), and the search results will bring back all relevant data mapped to that term in GtoImmuPdb.*
35
36
37

- 38
39 3. Scroll down to the section on Dendritic cells.
40
41

42 *The description explains that this category includes cell types that are children of the parent term "dendritic*
43 *cell (CL:0000451)" from the Cell Ontology. Clicking on the CL ID links to EMBL-EBI's Ontology Lookup*
44 *Service (Jupp et al., 2015), which provides more information on the term and displays it in the context of the*
45 *CL hierarchy.*
46
47

- 48
49 4. Click on the heading for the "Dendritic cells" section. Alternatively, select "Dendritic cells" from the "Cell
50 Types" menu item of the navigation bar.
51
52

53 *This page shows all targets associated to the immunological cell type category "Dendritic cells" (DCs) (Figure*
54 *25 panel A). The table is split into sections for each target class (e.g. GPCRs, Enzymes), and quick links are*
55 *available in the top right corner. The pull-down menu in the top left can be used to switch cell type category.*
56 *The first column shows the target name (linking to its GtoPdb detailed view page) and family name (in*
57
58
59

1
2 *parentheses), the second shows curator comments about the association, the third gives Cell Ontology terms*
3 *associated to the target, and the final column shows general immunopharmacology comments about the target.*

- 4
5
6 5. Scroll down the page to the bottom of the GPCRs section. The final row shows the target XCR1 (Figure 25
7 panel B).

8
9
10 *The information for XCR1 shows that it is part of the chemokine receptor family. The curator comments give*
11 *specific details about XCR1's expression in immunological cell types; in this case highlighting the specificity of*
12 *XCR1 to a specific subset of dendritic cells (CD141-positive DCs). The third column shows Cell Ontology*
13 *terms associated with XCR1, namely "CD141-positive myeloid dendritic cell (CL:0002394)". This is an*
14 *example of a specific, high resolution CL term annotated to a target.*

- 15
16
17
18 6. Click on the 'XCR1' name in the first column to view its detailed page.

19
20
21 *This displays the GtoPdb detailed page for the target with the GtoImmuPdb view switched on. Under this view,*
22 *in the Contents section and lower down the page, tables with data of immunological relevance are highlighted*
23 *in blue. A toggle button above the Contents section can be used to switch off or on the GtoImmuPdb view.*

24 *There is no difference in data content between the original GtoPdb view and the new GtoImmuPdb view; the*
25 *latter merely highlights content of specific immunopharmacological relevance.*

- 26
27
28
29 7. Click on Immunopharmacology Comments in the Contents section. The page focus jumps down to the relevant
30 table.

31
32
33 *The general curator comments here indicate that XCR1 is one of >20 chemokine receptors. It states that it is*
34 *expressed in a subset of dendritic cells and a reference is provided - click on the reference number to view*
35 *details. The comments may also contain hyperlinks to relevant ligands. The Immuno Cell Type Associations*
36 *section below this contains more specific comments about XCR1 expression in dendritic cells, with additional*
37 *references.*

40 41 **BASIC PROTOCOL 6**

42 43 **BROWSE TO EXPLORE DATA FOR IMMUNOLOGICAL PROCESSES**

44
45
46 The GtoImmuPdb project has enriched GtoPdb with data on biological processes of relevance to immunology. Targets
47 have been mapped to several top-level immunological process categories, which provides another way for organising
48 and searching immunopharmacological data. This is complementary to the cell type data described in Basic Protocol 5.
49 Process data are accessible from the GtoImmuPdb portal home page, *via* the Processes/Pathways panel. They are also
50 available from the Processes menu of the main navigation bar on every page. This protocol introduces how to browse
51 and view immunological process data and find pharmacological targets.
52
53

54
55
56 Necessary resources

Software

An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.

Exploring process data in the Guide to IMMUNOPHARMACOLOGY

1. Begin at the GtoImmuPdb portal home page (<http://www.guidetoimmunopharmacology.org>) (Figure 19).
2. Under the section Processes/Pathways, select “View immuno processes home page”. Alternatively, click on the Processes menu at the top, or directly navigate to <http://www.guidetoimmunopharmacology.org/GRAC/ProcessesForward>.

This page shows more details about the main process categories used in GtoImmuPdb (Figure 26). Quick links at the top move to the start of each process category section on the page. For each category, there is a brief description, alongside lists of the Gene Ontology (GO) terms included under that category. The process data in GtoImmuPdb are underpinned by the use of GO terms (www.geneontology.org) (Ashburner et al., 2000; The Gene Ontology Consortium, 2017), comprising a set of biological process terms and identifiers that are mapped to top-level immunological categories. Using GO provides a controlled vocabulary and external identifiers, which supports interoperability of the data. Similar to cell types, ontology mapping enables high resolution annotation of targets to process terms, and facilitates searches by parent or child terms to bring back all the data.

3. Click on the quick link for “Chemotaxis & migration” (Figure 26).

The summary for Chemotaxis & migration explains that this category includes (but is not limited to) processes that are children of four different GO parent terms: leukocyte cell-cell adhesion (GO:0007159), chemokine production (GO:0032602), leukocyte migration (GO:0050900), and response to chemokine (GO:1990868). Clicking on a GO term links to AmiGO 2 (Carbon et al., 2009), a GO browser, which provides more details for the term including its definition, context and annotations.

4. Click on the heading for the “Chemotaxis & migration” section. Alternatively, select this name from the “Processes” menu of the navigation bar.

The page now shows all targets associated to the immunological process category Chemotaxis & migration (Figure 27 panel A). The table is split into sections for each target class; quick links are available in the top right corner. The pull-down menu in the top left can be used to switch process category. The first column shows the target name (linking to its detailed view page) and family name (in parentheses), the second shows curator comments about the association, the third gives GO terms associated to the target, and the final column shows general immunopharmacology comments about the target.

5. Click on “Catalytic Receptors” to move quickly to that section, and scroll down to find the Integrins (this is the family name shown in parenthesis after the target name).

1
2
3 *There are three Integrins associated with the Chemotaxis & migration process (Figure 27 panel B). The third*
4 *column lists GO terms that are annotated against these targets. Looking at the data for "integrin, alpha 4*
5 *subunit", it shows six GO terms with their IDs and 3-letter evidence codes. The evidence code is provided by*
6 *the GO annotation source (the GO annotations used in GtoImmuPdb have been downloaded from UniProt*
7 *(The UniProt Consortium, 2017), and it indicates the level of support for the annotation. More information on*
8 *evidence codes is available from <http://www.geneontology.org/page/guide-go-evidence-codes>. The evidence*
9 *code helps to determine the veracity of the annotation.*

- 10
11
12
13
14 6. Click on the 'integrin, alpha 4 subunit' name in the first column to view its detailed view page. The detailed
15 page provides further information on the target, with data of immunological relevance highlighted in blue.
16
17
18 7. Click on Immunopharmacology Comments under the Contents section to jump to this section on the page.
19

20 *The comments indicate that this target promotes cell migration by interacting with a signalling adaptor*
21 *protein. They also indicate that alpha 4-containing integrins are the molecular targets of the antibody*
22 *natalizumab, a migration-inhibiting drug. Below this section is a detailed table of "Immuno Process*
23 *Associations" listing all the processes that this target is associated with.*

24 25 26 27 **BASIC PROTOCOL 7**

28 29 **NAVIGATING RECIPROCAL LINKS TO PUBMED AND PUBCHEM**

30
31
32 The provenance of entity records, out-links and contextual comments from curators and NC-IUPHAR members is
33 supported by document citations. While the database does have small numbers of book references, patent numbers and
34 URL-only links, by far the majority have PubMed Identifiers (PMIDs). GtoPdb release 2017.5 contains ~31,000 unique
35 PMIDs, of which ~22,000 are associated with targets and families. The rest are associated with ligand data, of which
36 ~6,000 include the quantitative binding interactions. These PMIDs form part of the substance records we submit to
37 PubChem (SIDs). This means that users can not only query into GtoPdb with a ligand or a PMID but can also navigate
38 PubChem-to-PubMed reciprocal linking as is shown below (Kim *et al.*, 2015).
39
40
41
42

43 Necessary resources

44
45 *Software*

46
47
48 An up-to-date Web browser, such as Firefox or Google Chrome, with JavaScript enabled.

49
50
51 Wordpad, or equivalent for temporary paste-across storage, particularly for query strings

52
53
54 Access to PubMed <https://www.ncbi.nlm.nih.gov/pubmed/>

55
56
57 *Identifiers*

PubMed ID (PMID), PubChem substance identifier (SID) and PubChem compound identifier (CID)

Protocol steps—Step annotations

Locating a reference and navigating to PubMed

1. For this example, use the reference: “Shibata K, Foglar R, Horie K, Obika K, Sakamoto A, Ogawa S, Tsujimoto G. (1995) KMD-3213, a novel, potent, alpha 1a-adrenoceptor-selective antagonist: characterization using recombinant human alpha 1-adrenoceptors and native tissues. *Mol. Pharmacol.*, 48: 250-258. [PMID:7651358]”.
2. From the GtoPdb home page click on Advanced Search in the main navigation bar or go directly to the target search tools at <http://www.guidetopharmacology.org/GRAC/searchPage.jsp>. In the fourth box, choose PubMed Id from the field options, enter 7651358 into the search box and hit “Search the database”.
3. The result of the search is shown in Figure 28. The matches include 13 ligands and three targets. Each of these entries has a reference section with multiple PMIDs (*e.g.* <http://www.guidetopharmacology.org/GRAC/LigandDisplayForward?tab=refs&ligandId=499>), but each reference includes a link-out (with a blue URL) to PubMed. Locate the example reference in one of the linked entries or move directly to the PubMed/NCBI Entrez system at <https://www.ncbi.nlm.nih.gov/pubmed/7651358>. Many linking options are displayed but Figure 29 shows the three that link to GtoPdb targets directly (plus a family link), and our associated PubChem SID links (shown on the right).

The alternative literature-related query options in GtoPdb can appear complex, particularly since targets and ligands can be citation-linked not only to the same but multiple papers. However, the circularity (document <>target <> ligand) is actually useful in that users can navigate to what they need by different routes.

Navigating from PubMed to PubChem and back to GtoPdb

4. The right-hand (RH) facets of Figure 29 show the linking options within the NCBI Entrez system. Note, these are distinct from the external LinkOut resources that are “pushed” to PubMed (as we have done for the target links). To square the circle by navigating back to GtoPdb try the SID links first (by definition, some of these will be from GtoPdb because we have curated the PMID link). The top section of the SID link results is shown in Figure 30. These are clearly labelled as IUPHAR/BPS Guide to PHARMACOLOGY, and each entry will link back to GtoPdb. Alternatively, the corresponding CIDs can be accessed by selected this option from the PubMed RH facet (Figure 29).

Link navigation within the NCBI systems is powerful but dauntingly complex in the many presented options, most of which have elements of circularity. Notwithstanding, the examples above show users can make useful queries from both the GtoPdb and NCBI sides and join “across” both resources. The entry points can be via

1 *PMIDs or CIDs and SIDs. Once users have crossed over to PubChem, all the options described above in*
2 *Support Protocol 2 (navigating reciprocal links to PubChem) become available to extend the user's*
3 *exploration of ligand relationships (Kim et al., 2015).*
4
5
6

7 **COMMENTARY**

8 **Background Information**

9
10 The IUPHAR/BPS Guide to PHARMACOLOGY is intended to be reference resource for researchers, teachers and
11 those studying pharmacology. It aims to provide information at two levels of detail: the first gives an overview of the
12 nomenclature, key properties and recommended experimental ligands for a wide-range of targets. These can be found
13 on the family summary pages (Figure 1 and described in Basic Protocol 1 steps 3 and 4). The summary information is
14 published every two years as the Concise Guide to PHARMACOLOGY (CGTP,
15 <http://www.guidetopharmacology.org/concise/>) (Alexander *et al.*, 2015; Alexander *et al.*, 2017; Alexander *et al.*, 2013).
16 The information is checked and updated by the CGTP collaborators (members of the NC-IUPHAR subcommittees or
17 independent experts where no subcommittee exists) prior to publication. The CGTP provides a citable, point-in-time
18 snapshot of the data, extracted into publication format intended as a quick desktop reference and introduction to
19 pharmacology. It is available online as a HTML and PDF document, but the paper copies remain very popular, as it has
20 the advantage that it can be used at the laboratory bench and kept on a shelf for rapid access. This is particularly useful
21 in locations where there is unreliable internet access, such as in some developing countries.
22
23
24
25
26
27

28 The second, more comprehensive, level of detail provides additional database pages for a selection of established
29 targets, which covers a wide range of information on function, tissue distribution, and assays, through to clinically-used
30 drugs and variants affecting pharmacology or causing disease. Data on these pages are also reviewed biennially by the
31 wider NC-IUPHAR subcommittees. These pages include longer lists of ligands, which may be endogenous ligands,
32 approved drugs, or ligands currently recommended for experimental use, while some others may have been used in the
33 past and are maintained for a historical perspective. All data tables have the option for longer free text comments, where
34 more details about the experiments and any caveats are described. Detailed data for the GPCRs, ion channels and NHRs
35 have been collected and revised over a 14-year period, originating in the IUPHAR-DB (which was a precursor for
36 GtoPdb) (Harmar *et al.*, 2009; Sharman *et al.*, 2013; Sharman *et al.*, 2011). Hence, these pages contain in-depth
37 information on a range of topics, while the newer targets tend to focus on key features related to pharmacology,
38 function and clinical relevance. Some of the data types originally included, such as variants and tissue expression, are
39 now readily available from other resources, so we do not tend to curate these ourselves. Much of this information is
40 available *via* the database link-outs provided on target pages. An example is the Human Protein Atlas (Uhlen *et al.*,
41 2015) which maps protein expression in cells, tissues, and organs, and now includes the Pathology Atlas (Uhlen *et al.*,
42 2017).
43
44
45
46
47
48
49

50 The IUPHAR Guide to IMMUNOPHARMACOLOGY has been created as an extension to the existing GtoPdb site,
51 with a new portal and data types, which draws together information on immunological cell types, processes and
52 diseases, with data on pharmacological targets and their ligands. In doing so, we have deliberately not created an
53 entirely new resource, but have extended the existing database schema and website for GtoPdb. The
54 immunopharmacology data (and all other data types) are available both from the original GtoPdb site and the
55 GtoImmuPdb site. The latter alerts users to specific immunopharmacology data and relevant targets and ligands through
56
57
58
59
60

1 content highlighting. The GtoImmuPdb portal provides new paths to access the data *via* the cell type, process and
2 disease pages described in Protocols 4, 5 and 6. These have been developed in collaboration with immunologists to
3 capture the relevant categories and data fields. The site is currently under development, and is undergoing rounds of
4 user testing, which will identify any areas for improvement. Therefore, although the main data types and navigation
5 systems are in place, the interface and data content may change slightly from the description presented here.
6
7

8
9
10 The latest content statistics and a detailed breakdown of each target/ligand category are described in our latest *Nucleic*
11 *Acids Research* Database Issue article (Harding et al., 2018). This two-yearly update includes a detailed explanation of
12 new features and content added to the database since the previous descriptions were published in 2014 and 2016
13 (Pawson *et al.*, 2014; Southan *et al.*, 2016). We also regularly publish technical reports, announcements of database
14 updates and invited commentaries on hot topics in pharmacology on our blog (<https://blog.guidetopharmacology.org/>).
15 These and other news items, such as IUPHAR review articles (see publication list at
16 <http://www.guidetopharmacology.org/nciupharPublications.jsp>), are promoted *via* our social media accounts on Twitter,
17 Facebook and LinkedIn. We also maintain sets of public slides on SlideShare
18 (<https://www.slideshare.net/GuidetoPHARM>), which include generic slides for use in teaching with GtoPdb.
19
20
21
22

23 **Critical Parameters and Troubleshooting**

24

25
26 The GtoPdb and GtoImmuPdb websites generally do not require direct user input of parameters, apart from web service
27 access described in Alternate Protocol 3. The search terms will accept any alphanumeric character, and recognise many
28 special characters such as Greek letters commonly used in target names (*e.g.* β_1 -adrenoceptor). Keyword searches use
29 the full text search algorithm provided as part of PostgreSQL. This involves converting user-provided search terms to
30 their lexemes and performing searches against pre-calculated lexemes for the text fields in the database, with a ranking
31 applied based on the strength of the match to a particular field. In practice, this means that searches for “melatonin
32 receptors” will return results matching to “melatonin receptors”, “melatonin receptor”, or any other possible variant.
33 However, this type of search does not always work well for some types of names used in biology and chemistry,
34 particularly those including punctuation. This commonly affects compounds, since research code names are often
35 recorded in literature or online databases in different formats, some containing punctuation with a hyphen or comma.
36 For example, ligand ID 2592 in GtoPdb is L735821, however this compound has also been referred to in the literature
37 as L-735,821. Searches of GtoPdb using the latter term unfortunately fail because of the punctuation, which causes the
38 name to be split into separate parts. Therefore, if users encounter a ligand name with punctuation, and initially get no
39 results from a name search, they are encouraged to try a search without punctuation, or to try a wildcard search for part
40 of the name. Wildcard searches are available *via* the Advanced Target Search or Ligand Search pages. Queries should
41 be entered in the relevant box with a percentage (%) sign used to represent the missing characters. For example, a
42 ligand name wildcard search for “%821” returns 4 results, one of which is L735821.
43
44
45
46
47
48
49

50 The GtoPdb web services do require URLs and parameters to be constructed correctly. Misspelling a parameter name in
51 an otherwise correct web service call would lead to the default behaviour of all results being returned. If an incorrect
52 target or ligand ID is provided, this would usually lead to a 404 Not Found error response code. The same is true for
53 trying to call a web service that does not exist, for example by misspelling the name. If a correct call is made to web
54 service that does not return any results because that particular database table is empty (for example, asking for a list of
55 variants for target ID 400 using <http://www.guidetopharmacology.org/services/targets/400/variants>) a 204 No Content
56
57
58
59
60

1 response code would be returned. The list of response codes and their meanings are described in the documentation at
2 <http://www.guidetopharmacology.org/webServices.jsp#responses>.
3
4

5
6 Whilst older entries are generally not removed from the database as they serve as a useful historical record of
7 pharmacological development, occasionally data are removed if they are found to be incorrect or new information has
8 come to light which reduces the value of their inclusion. Sometimes links to the deleted data will still exist on external
9 sites, so users may encounter a “not found” error for deprecated entries. Occasionally, new information results in
10 reclassification of targets and/or reorganisation of the family hierarchy. The database tracks changes in target and
11 family nomenclature and deprecated URLs will redirect to the new pages where possible.
12
13

14
15 GtoPdb is manually compiled by a team of curators with advice from a network of over 500 expert scientists organised
16 into 96 subcommittees overseen by NC-IUPHAR. The majority of references have thus been recommended by experts
17 with the data extracted and normalised into a database format by the curators. In some newer areas, curators select the
18 initial ligand reports while the subcommittees are being set up. Despite this level of double checking, as with any
19 database errors can occur, and we encourage users to contact us by email if they are unsure about any data. The best
20 way to do this is *via* the email address enquiries@guidetopharmacology.org. If the query is complex, we often refer this
21 on to the relevant subcommittee chair for comment.
22
23
24

25
26 We provide Support Protocols 1 and 2 and Basic Protocol 7 as examples of the powerful data mining possibilities that
27 arise from combining GtoPdb data with other sources, *via* the UniProtKB, PubChem and PubMed interfaces. However,
28 these sites have their own complexities and we cannot go into detail about their usage here. Each resource provides their
29 own detailed help pages and support desk, and users with specific queries about these examples are welcome to contact
30 us for help.
31
32

33 ACKNOWLEDGEMENT

34
35 The authors wish to thank all members of NC-IUPHAR for their continued support
36 (<http://www.guidetopharmacology.org/nciuphar.jsp#membership>). We also thank the global network of NC-IUPHAR
37 subcommittees and all the CGTP contributors (a full list of subcommittee members and contributors can be viewed at
38 <http://www.guidetopharmacology.org/GRAC/ContributorListForward>). We thank T. Wigglesworth for administrative
39 support. This work was supported by IUPHAR and the BPS. In addition to our primary funding, we are grateful for
40 sponsorship from the American Society for Pharmacology and Experimental Therapeutics (ASPET), Laboratoires
41 Servier and The University of Edinburgh.
42
43
44
45

46 LITERATURE CITED

47
48 Alexander SP, Kelly E, Marrion N, Peters JA, Benson HE, Faccenda E, *et al.* (2015). The Concise Guide to
49 PHARMACOLOGY 2015/16: Overview. *Br J Pharmacol* 172(24): 5729-5743.
50
51

52
53 Alexander SP, Kelly E, Marrion NV, Peters JA, Faccenda E, Harding SD, *et al.* (2017). The Concise Guide to
54 PHARMACOLOGY 2017/18: Overview. *Br J Pharmacol* 174(Suppl 1): S1-S16 .
55

56
57 Alexander SPH, Benson HE, Faccenda E, Pawson AJ, Sharman JL, McGrath JC, *et al.* (2013). THE CONCISE GUIDE
58 TO PHARMACOLOGY 2013/14: OVERVIEW. *Br J Pharmacol* 170(8): 1449-1458.
59

1
2
3 Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, *et al.* (2000). Gene ontology: tool for the
4 unification of biology. The Gene Ontology Consortium. *Nat Genet* 25(1): 25-9.

5
6
7 Bento AP, Gaulton A, Hersey A, Bellis LJ, Chambers J, Davies M, *et al.* (2014). The ChEMBL bioactivity database: an
8 update. *Nucleic Acids Res* 42(Database issue): D1083-1090.

9
10
11 Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, *et al.* (2009). AmiGO: online access to ontology and
12 annotation data. *Bioinformatics* 25(2): 288-289.

13
14
15 Carvalho S, Levi-Schaffer F, Sela M, Yarden Y (2016). Immunotherapy of cancer: from monoclonal to oligoclonal
16 cocktails of anti-cancer antibodies: IUPHAR Review 18. *Br J Pharmacol* 173(9): 1407-1424.

17
18
19 Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, *et al.* (2017). InterPro in 2017-beyond protein
20 family and domain annotations. *Nucleic Acids Res* 45(D1): D190-D199.

21
22
23 Gibney G, Baxeivanis AD (2011). Searching NCBI databases using Entrez. *Current protocols in bioinformatics* Chapter
24 1: Unit 1 3.

25
26
27 Harding SD, Sharman JL, Faccenda E, Southan C, Pawson AJ, Ireland S, *et al.* (2018). The IUPHAR/BPS Guide to
28 PHARMACOLOGY in 2018: Updates and expansion to encompass the new Guide to IMMUNOPHARMACOLOGY.
29 *Nucleic Acids Res* 46(Database issue) doi: 10.1093/nar/gkx1121. [Epub ahead of print]

30
31
32 Harmor AJ, Hills RA, Rosser EM, Jones M, Buneman OP, Dunbar DR, *et al.* (2009). IUPHAR-DB: the IUPHAR
33 database of G protein-coupled receptors and ion channels. *Nucleic Acids Res* 37(Database issue): D680-685.

34
35
36 Ishii M (2017). Immunology proves a great success for treating systemic autoimmune diseases - a perspective on
37 immunopharmacology: IUPHAR Review 23. *Br J Pharmacol* 174(13): 1875-1880.

38
39
40 Malone J, Stevens R, Forsberg K, Splendiani A (eds) (2015). A new Ontology Lookup Service at EMBL-EBI.
41 *Proceedings of the 8th International Conference on Semantic Web Applications and Tools for Life Sciences*; December
42 7-10; Cambridge, UK.

43
44
45 Kim S, Han L, Yu B, Hahnke VD, Bolton EE, Bryant SH (2015). PubChem structure-activity relationship (SAR)
46 clusters. *J Cheminform* 7: 33.

47
48
49 Kim S, Thiessen PA, Cheng T, Yu B, Shoemaker BA, Wang J, *et al.* (2016). Literature information in PubChem:
50 associations between PubChem records and scientific articles. *J Cheminform* 8: 32.

51
52
53 Landolina N, Levi-Schaffer F (2016). Monoclonal antibodies: the new magic bullets for allergy: IUPHAR Review 17.
54 *Br J Pharmacol* 173(5): 793-803.

55
56
57 Licastro F, Candore G, Lio D, Porcellini E, Colonna-Romano G, Franceschi C, *et al.* (2005). Innate immunity and
58 inflammation in ageing: a key for understanding age-related diseases. *Immunity & ageing : I & A* 2: 8.

59
60
61 McInnes IB, Schett G (2011). The pathogenesis of rheumatoid arthritis. *The New England journal of medicine* 365(23):
2205-2219.

1
2
3 O'Shea JJ, Holland SM, Staudt LM (2013). JAKs and STATs in immunity, immunodeficiency, and cancer. *The New*
4 *England journal of medicine* 368(2): 161-170.

5
6
7 Papadatos G, Davies M, Dedman N, Chambers J, Gaulton A, Siddle J, *et al.* (2016). SureChEMBL: a large-scale,
8 chemically annotated patent document database. *Nucleic Acids Res* 44(D1): D1220-1228.

9
10
11 Pawson AJ, Sharman JL, Benson HE, Faccenda E, Alexander SP, Buneman OP, *et al.* (2014). The IUPHAR/BPS Guide
12 to PHARMACOLOGY: an expert-driven knowledgebase of drug targets and their ligands. *Nucleic Acids Res*
13 42(Database issue): D1098-1106.

14
15 Perry VH, Newman TA, Cunningham C (2003). The impact of systemic infection on the progression of
16 neurodegenerative disease. *Nature reviews. Neuroscience* 4(2): 103-112.

17
18
19 Pundir, S., Magrane, M., Martin, M.J., O'Donovan, C., and The UniProt Consortium. (2015). Searching and navigating
20 UniProt databases. *Curr. Protoc. Bioinform.* 50:1.27.1-1.27.10. doi: 10.1002/0471250953.bi0127s50

21
22 Rawlings, N.D., Barrett, A.J., and Bateman, A. (2014). Using the MEROPS Database for Proteolytic Enzymes and
23 Their Inhibitors and Substrates. *Curr. Protoc. Bioinform.* 48:1.25.1-1.25.33. doi: 10.1002/0471250953.bi0125s48

24
25
26 Rawlings ND, Barrett AJ, Finn R (2016). Twenty years of the MEROPS database of proteolytic enzymes, their
27 substrates and inhibitors. *Nucleic Acids Res* 44(D1): D343-350.

28
29
30 Rose PW, Prlic A, Altunkaya A, Bi C, Bradley AR, Christie CH, *et al.* (2017). The RCSB protein data bank: integrative
31 view of protein, gene and 3D structural information. *Nucleic Acids Res* 45(D1): D271-D281.

32
33
34 Sharman JL, Benson HE, Pawson AJ, Lukito V, Mpamhanga CP, Bombail V, *et al.* (2013). IUPHAR-DB: updated
35 database content and new features. *Nucleic Acids Res* 41(Database issue): D1083-1088.

36
37
38 Sharman JL, Mpamhanga CP, Spedding M, Germain P, Staels B, Dacquet C, *et al.* (2011). IUPHAR-DB: new receptors
39 and tools for easy searching and visualization of pharmacological data. *Nucleic Acids Res* 39(Database issue): D534-
40 538.

41
42
43 Southan C, Sharman JL, Benson HE, Faccenda E, Pawson AJ, Alexander SP, *et al.* (2016). The IUPHAR/BPS Guide to
44 PHARMACOLOGY in 2016: towards curated quantitative interactions between 1300 protein targets and 6000 ligands.
45 *Nucleic Acids Res* 44(D1): D1054-1068.

46
47
48 The Gene Ontology Consortium. (2017) Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids*
49 *Res* 45(D1): D331-D338.

50
51
52 The UniProt Consortium (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45(D1): D158-D169.

53
54
55 Tiligada E, Ishii M, Riccardi C, Spedding M, Simon HU, Teixeira MM, *et al.* (2015). The expanding role of
56 immunopharmacology: IUPHAR Review 16. *Br J Pharmacol* 172(17): 4217-4227.

57
58
59 Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, *et al.* (2015). Proteomics. Tissue-based
60 map of the human proteome. *Science* 347(6220): 1260419.

1
2
3 Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhorji G, *et al.* (2017). A pathology atlas of the human cancer
4 transcriptome. *Science* 357(6352).
5

6 7 FIGURE LEGENDS

8
9
10 Figure 1. Screenshot of the 5-HT₃ receptor family summary page with the summary section for the target 5-HT₃AB
11 open. The summary includes NC-IUPHAR's recommended nomenclature, alternative names by which the target has
12 been known, component subunits (clicking on these takes users to the subunit section, which includes gene and protein
13 identifiers), key selective ligands listed by action with affinities and references, and the functional characteristics of this
14 ion channel.
15

16
17
18 Figure 2. Screenshot of the ligand tables on the 5-HT₃AB detailed target page. The Channel Blockers table contains
19 human and mouse data and species-specific tables are available by clicking on the link below it. Tables showing the
20 chemical structures of the compounds are available by clicking on the "View all chemical structures" links.
21

22
23 Figure 3. The target search tools, showing a text search for "5-HT3" matching to the names of 6 targets and 1 family.
24

25
26 Figure 4. Making an initial select for GtoPdb cross-references using the UniProtKB Advanced search menu bar.
27

28
29 Figure 5. The results of changing the column display settings for 1483 human protein GtoPdb cross-references in
30 UniProtKB.
31

32
33 Figure 6. A sort on a post-download Excel sheet for GtoPdb cross-references and proteases in UniProt, ranked by
34 number of PDB entries.
35

36
37 Figure 7. Seven UniProt entries matching the InterPro domain "IPR001461 Aspartic_peptidase_A1" with GtoPdb cross-
38 references, and showing links to the rare disease resource, Orphanet.
39

40
41 Figure 8. Screenshot of the ligand summary page for erlotinib. Nomenclature information includes commonly-used
42 synonyms and identifiers from the literature, the IUPAC name, English trade names, and the International
43 Nonproprietary Name (INN). The ligand page also includes a depiction of the chemical structure, and under the
44 Structure tab various structural identifiers such as the SMILES string, International Chemical Identifier (InChI), and
45 InChIKey. Clicking on the structure image loads the structure into the chemical drawing tool which can be used for
46 structure-based searches of the database.
47
48

49
50 Figure 9. Data on quantitative ligand binding at targets are displayed under the Biological activity tab of the erlotinib
51 ligand page. The primary target is indicated as epidermal growth factor receptor, with an additional target of the
52 OATP2B1 transporter. This drug is also known to be a substrate of the CYP3A4 enzyme.
53

54
55 Figure 10. A SMILES structure has been loaded into the ligand structure drawing tool. The tool can perform searches
56 by substructure, high or moderate similarity, exact match, or SMARTS pattern matching.
57
58

1
2
3 Figure 11. Connection between GtoPdb and PubChem. The GtoPdb ligand entry is 8699. The arrow indicates the click
4 crossing to PubChem CID 51352361.
5

6
7 Figure 12. Navigation from CID 51352361 “Contents” (left hand facets) across to “Similar Compounds”, highlighted in
8 yellow.
9

10
11 Figure 13. Showing the result of the “Similar Compounds” link from CID 51352361. Additional options and filters are
12 shown in the right-hand facets.
13

14
15 Figure 14. Clustering of all 684 “Similar Compounds”. One CID is highlighted to show the substructure common to all
16 members of the set. Note, the button “Similarity Data” produces a CSV matrix file of all the Tanimoto scores.
17

18
19 Figure 15. The PubChem structure search page set up with CID 51352361. The Tanimoto options for similarity
20 searching are shown with 96% set as the threshold.
21

22
23 Figure 16. 3D clustering of the 74 compounds 97% similar to CID 51352361. The options were set to display the
24 conformer renderings and large-size 2D/3D juxtaposed displays for the mouseovers.
25

26
27 Figure 17. Showing the results of editing CID 51352361 and searching GtoPdb by substructure using the SMILES
28 string FC2=C(C)C=C(N(C(=O)C1=NC=CC=C1)[H])C=C2. Six of the seven matches are shown.
29

30
31 Figure 18. A section of the JSON output from a web service call to generate a list of GPCRs. Each target is shown with
32 summary information on its nomenclature, family IDs, component subunit IDs if it is a complex (*e.g.* the GABA_B
33 receptor), and complex IDs if it is a component subunit (*e.g.* GABA_{B1} and GABA_{B2}). Names include HTML codes to
34 represent formatting such as subscript (the sub tag), and special characters such as the Greek alpha.
35
36

37
38 Figure 19. The Guide to IMMUNOPHARMACOLOGY portal, showing the entry points to processes, cell types, targets
39 and ligands.
40

41
42 Figure 20. The GtoImmuPdb disease-target associations page. Users can switch to view disease-ligand associations
43 using the tab at the top.
44

45
46 Figure 21. GtoImmuPdb expanded view of the disease-ligand associations for Chronic lymphocytic leukemia.
47 Approved drugs are indicated with the medicine bottle icon, and the main drug approval bodies (usually FDA and/or
48 EMA, although others are given where the drug is only approved outside the US and Europe) and first approval dates
49 are given.
50

51
52
53 Figure 22. A summary of clinical data for the drug idelalisib, from its GtoPdb ligand page.
54

55
56 Figure 23. A section of the search results for the disease term “Alzheimer’s”. The results are ordered by best match,
57 with the top result (target GPR3) having 6 matches across 5 database fields.
58

1
2
3 Figure 24. The GtoImmuPdb immunological cell types home page, showing the list of top-level cell type categories.
4

5
6 Figure 25. (A) The list of GtoImmuPdb target associations to dendritic cells. Users can switch cell type using the drop
7 down menu at the top left. (B) A close-up of the association between the target XCR1 and dendritic cells. The second
8 column shows curator comments, the third column gives the Cell Ontology terms, and the final column shows an
9 extract of the general immunopharmacology comments on XCR1.
10

11
12 Figure 26. The immunological processes home page, showing the definitions and mapped GO terms for the first few
13 top-level process categories. Quick links to the categories are provided at the top.
14

15
16 Figure 27. (A) The list of target associations for chemotaxis and migration. The third column shows the GO
17 immunological process terms mapped to the targets named in the left column. The final column summarises the general
18 immunopharmacology comments for the target. (B) The process GO terms and comments for the integrins.
19
20

21
22 Figure 28. The literature search box from the GtoPdb Advanced search page is shown at the top. The initial section of
23 the result list is in the centre, and the reference as shown in each of the linked entries is at the bottom. The PMID in blue
24 is live-linked.
25

26
27 Figure 29. Link displays from <https://www.ncbi.nlm.nih.gov/pubmed/7651358>. The GtoPdb targets are linked *via*
28 Molecular Biology Databases under LinkOut resources. PubChem links for GtoPdb ligands can be found on the right
29 under Related information.
30

31
32 Figure 30. Some of the PubChem SID links from PMID 7651358 (as shown in Figure 29). All four of these structures
33 link directly back to GtoPdb.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

5-HT₃ receptors

Unless otherwise stated all data on this page refer to the human proteins. Gene information is provided for human (Hs), mouse (Mm) and rat (Rn).

GtoImmuPdb view: **OFF** [Toggle GtoImmuPdb View](#) [Expand all sections](#) [Collapse all sections](#)

Overview

[Show »](#) [More detailed introduction !\[\]\(bfe64b3b99d726c20cb41da66e0bcb5a_img.jpg\)](#)

Channels and Subunits

Complexes

5-HT₃AB [Hide summary](#) [More detailed page !\[\]\(15a01675795da3c0924461c44b8bf202_img.jpg\)](#)

Target Id	378
Nomenclature	5-HT₃AB
Previous and unofficial names	5-HT _{3A/B}
Subunits	5-HT _{3A} 5-HT _{3B}
Channel blockers	picrotoxinin pIC ₅₀ 4.2 [46] bilobalide pIC ₅₀ 2.5 [46] ginkgolide B pIC ₅₀ 2.4 [46]
Functional characteristics	$\gamma = 0.4\text{--}0.8$ pS [+ 5-HT _{3B} , $\gamma = 16$ pS]; inwardly rectifying current [+ 5-HT _{3B} , rectification reduced]; n _H 2-3 [+ 5-HT _{3B} 1-2]; relative permeability to divalent cations reduced by co-expression of the 5-HT _{3B} subunit

[5-HT_{3A} Show summary »](#) [More detailed page !\[\]\(5c21dc776ae4aea2d2bd6ea020387e51_img.jpg\)](#)

Subunits






5-HT_{3A} Show summary »	More detailed page 
5-HT_{3B} Show summary »	More detailed page 
5-HT_{3C} Show summary »	More detailed page 
5-HT_{3D} Show summary »	More detailed page 
5-HT_{3E} Show summary »	More detailed page 

Figure 1. Screenshot of the 5-HT₃ receptor family summary page with the summary section for the target 5-HT₃AB open. The summary includes NC-IUPHAR's recommended nomenclature, alternative names by which the target has been known, component subunits (clicking on these takes users to the subunit section, which includes gene and protein identifiers), key selective ligands listed by action with affinities and references, and the functional characteristics of this ion channel.

250x268mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Agonists						
Key to terms and symbols		View all chemical structures			Click column headers to sort	
Ligand	Sp.	Action	Affinity	Units	Reference	
quipazine	Hs	Agonist	9.0	pK _i	1	
meta-chlorophenylbiguanide	Hs	Agonist	7.0	pK _i	1	
5-hydroxytryptamine	Hs	Agonist	6.0	pK _i	1	
1-phenylbiguanide	Hs	Agonist	4.9	pK _i	1	
meta-chlorophenylbiguanide	Hs	Agonist	5.7	pEC ₅₀	5	
5-hydroxytryptamine	Hs	Agonist	4.8 – 5.8	pEC ₅₀	2,5-6,14	
2-methyl-5-HT	Hs	Agonist	4.9	pEC ₅₀	5	

Agonist Comments
 Apparent affinities of agonists are for ligand binding to the recombinant 5-HT₃AB receptor expressed in mammalian cells, or pEC₅₀ values determined under voltage-clamp for the receptor expressed in *Xenopus laevis* oocytes. Selectivity refers to the 5-HT₃ receptor family; the agents listed do not discriminate between 5-HT₃A and 5-HT₃AB receptors, although in some cases they demonstrate lower potency at the latter. Comments concerning efficacy relate to data obtained from voltage-clamp studies of the human 5-HT₃AB receptor expressed in *Xenopus laevis* oocytes and from Ca²⁺ imaging studies of the receptor expressed in HEK 293 cells [6].

Antagonists						
Key to terms and symbols		View all chemical structures			Click column headers to sort	
Ligand	Sp.	Action	Affinity	Units	Reference	
[³ H]granisetron	Hs	Antagonist	8.8	pK _d	1	
(S)-zacopride	Hs	Antagonist	8.8	pK _i	1	
azasetron	Hs	Antagonist	8.4	pK _i	1	
ondansetron	Hs	Antagonist	7.8	pK _i	1	
(R)-zacopride	Hs	Antagonist	7.7	pK _i	1	
metoclopramide	Hs	Antagonist	5.7	pK _i	1	
cocaine	Hs	Antagonist	4.8	pK _i	1	
tubocurarine	Hs	Antagonist	4.5	pK _i	1	

Antagonist Comments
 Data tabulated are for ligand binding to the human recombinant 5-HT₃AB receptor expressed in mammalian cells. Selectivity refers to the 5-HT₃ receptor family; the agents listed do not discriminate between 5-HT₃A and 5-HT₃AB receptor subtypes in radioligand binding studies. However, in electrophysiological studies, (+)-tubocurarine demonstrates modest selectivity for human 5-HT₃A (IC₅₀ = 3 μM) versus human 5-HT₃AB (IC₅₀ = 1.4–2.1 μM) receptors [5]. A more potent blockade by (+)-tubocurarine, although with reduced selectivity, is apparent for the rat 5-HT₃A and 5-HT₃AB receptors [7].

Channel Blockers						
Key to terms and symbols		View all chemical structures			Click column headers to sort	
Ligand	Sp.	Use-dependent	Affinity	Units	Voltage-dependent (mV)	Reference
picrotoxinin	Hs	no	4.2	pIC ₅₀	no	15
picrotoxin	Mm	yes	2.9	pIC ₅₀	no	3-4
bilobalide	Hs	no	2.5	pIC ₅₀	no	15
ginkgolide B	Hs	no	2.4	pIC ₅₀	no	15

[View species-specific channel blocker tables](#)

Channel Blocker Comments
 Although picrotoxin is approximately 27-fold more potent in blocking mouse 5-HT₃AB versus mouse 5-HT₃A receptors, the degree of discrimination between equivalent human receptor orthologues is substantially smaller, most probably due to differences in the structure of the TM2 domain [4].

Allosteric Modulators						
Key to terms and symbols		View all chemical structures			Click column headers to sort	
Ligand	Sp.	Action	Concentration range (M)	Voltage-dependent (mV)	Reference	
trichloroethanol	Mm	Positive	2.5 × 10 ⁻⁴ - 1 × 10 ⁻²	no	9	

Allosteric Modulator Comments
 Ethanol is a positive allosteric modulator of the 5-HT₃A receptor but, at concentrations up to 200 mM, has no effect on currents mediated by the 5-HT₃AB receptor [9]. Chloroform, halothane and small volume n-alcohols enhance the gating of 5-HT₃A receptors and incorporation of the 5-HT₃B subunit to form 5-HT₃AB receptors suppresses this action [13-14].

Figure 2. Screenshot of the ligand tables on the 5-HT₃AB detailed target page. The Channel Blockers table contains human and mouse data and species-specific tables are available by clicking on the link below it. Tables showing the chemical structures of the compounds are available by clicking on the "View all chemical structures" links.

247x406mm (96 x 96 DPI)

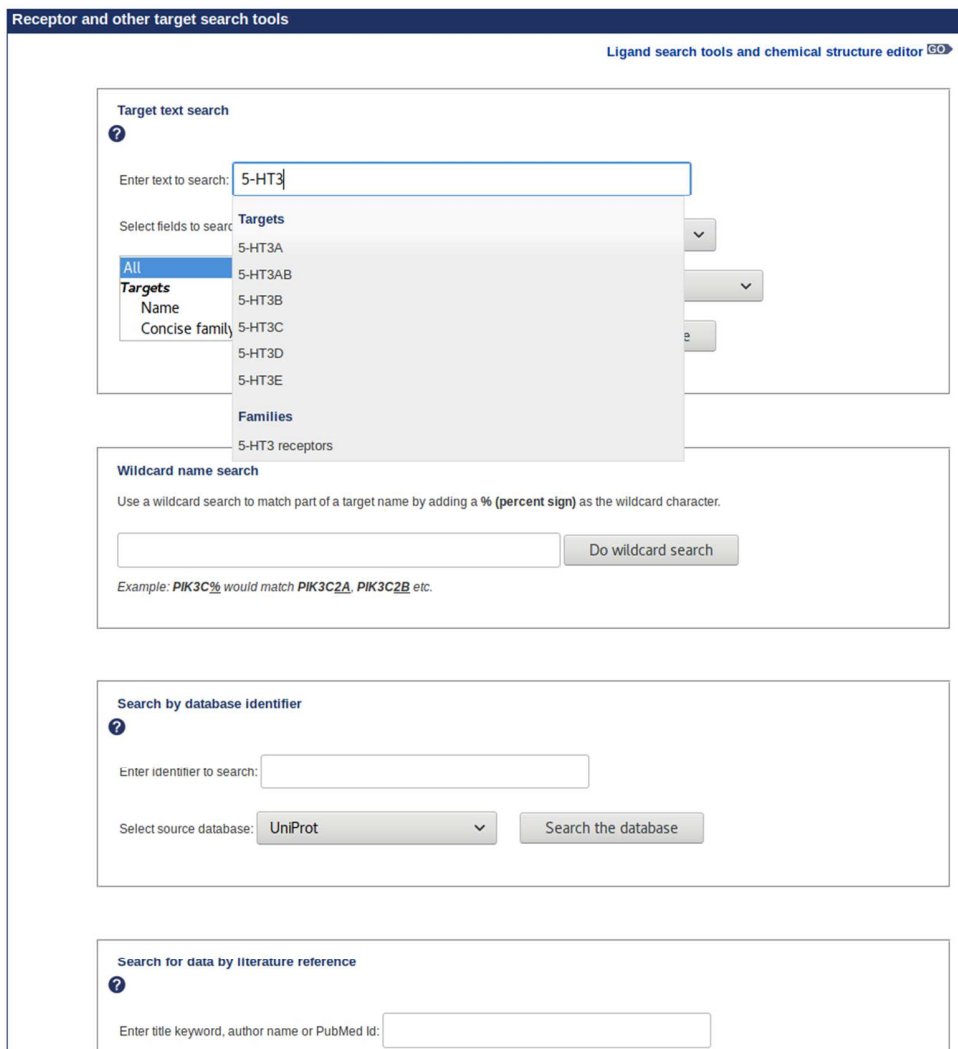


Figure 3. The target search tools, showing a text search for "5-HT3" matching to the names of 6 targets and 1 family.

251x271mm (96 x 96 DPI)

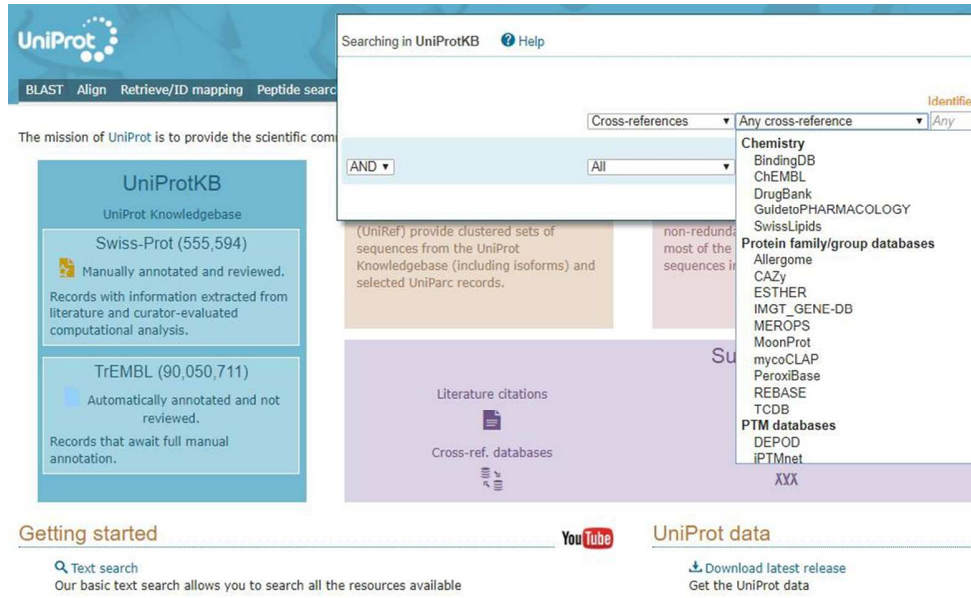


Figure 4. Making an initial select for GtoPdb cross-references using the UniProtKB Advanced search menu bar.

257x160mm (96 x 96 DPI)

Entry	Entry name	Protein names	Gene names	Length	EC number	Gene ontology (molecular function)	Subcellular location [CC]
Q14524	SCNSA_HUMAN	Sodium channel protein type 5 subun...	SCNSA	2,016		ankyrin binding; calmodulin binding; enzyme binding; fibroblast growth factor binding; ion channel binding; nitric-oxide synthase binding; protein domain specific binding; protein kinase binding; scaffold protein binding; ubiquitin protein ligase binding; voltage-gated sodium channel activity; voltage-gated sodium channel activity involved in AV node cell action potential; voltage-gated sodium channel activity involved in bundle of His cell action potential; voltage-gated sodium channel activity involved in cardiac muscle cell action potential; voltage-gated sodium channel activity involved in Purkinje myocyte action potential; voltage-gated sodium channel activity involved in SA node cell action potential	<ul style="list-style-type: none"> Cell membrane 1 4 Publications 1; Multi-pass membrane protein 1 4 Publications 1; By similarity 1
P00533	EGFR_HUMAN	Epidermal growth factor receptor	EGFR ERBB, ERBB1, HER1	1,210	2.7.10.1	actin filament binding; ATP binding; cadherin binding; calmodulin binding; chromatin binding; double-stranded DNA binding; enzyme binding; epidermal growth factor-activated receptor activity; epidermal growth factor binding; identical protein binding; integrin binding; MAP kinase kinase activity; phosphatidylinositol-4,5-bisphosphate 3-kinase activity; protein heterodimerization activity; protein kinase binding; protein phosphatase binding; protein tyrosine kinase activity; Ras guanyl-nucleotide exchange factor activity; signal transducer, downstream of receptor, with protein tyrosine kinase activity; transmembrane receptor protein tyrosine kinase activity; transmembrane signalling receptor activity; ubiquitin protein ligase binding; virus receptor activity	<ul style="list-style-type: none"> Cell membrane 1; Single-pass type I membrane protein 1 Endoplasmic reticulum membrane 1; Single-pass type I membrane protein 1 Golgi apparatus membrane 1; Single-pass type I membrane protein 1 Nucleus membrane 1; Single-pass type I membrane protein 1 Endosome 1 Endosome membrane 1 Nucleus 1 <p>Note: In response to EGF, translocated from the cell membrane to the nucleus via Golgi and ER. Endocytosed upon activation by ligand. Colocalized with GPER1 in the nucleus of estrogen agonist-induced cancer-associated fibroblasts (CAF).</p> <p>Isoform 2:</p> <ul style="list-style-type: none"> Secreted 1
P10275	ANDR_HUMAN	Androgen receptor	AR DHTF, NR3C4	920		androgen binding; androgen receptor activity; ATPase binding; beta-casein binding; chromatin binding; DNA binding; enzyme binding; POU domain binding; protein dimerization activity; receptor binding; RNA polymerase II core promoter proximal region sequence-specific DNA binding; RNA polymerase II transcription factor activity; ligand-activated sequence-specific DNA binding; RNA polymerase II transcription factor binding; steroid binding; transcriptional activator activity; RNA polymerase II core promoter proximal region sequence-specific binding; transcription factor activity; sequence-specific DNA binding; transcription factor binding; transcription regulatory region DNA binding; zinc ion binding	<ul style="list-style-type: none"> Nucleus 1 4 Publications 1 Cytoplasm 1 3 Publications 1 <p>Note: Predominantly cytoplasmic in unligated form but translocates to the nucleus upon ligand-binding. Can also translocate to the nucleus in unligated form in the presence of RACK1. 2 Publications 1</p>

Figure 5. The results of changing the column display settings for 1483 human protein GtoPdb cross-references in UniProtKB.

402x197mm (96 x 96 DPI)

Prothrombin EC 3.4.11.2	622 3.4.21.5	383	TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
Beta-secretase 1 EC 3.4.23.46	501 3.4.23.46	357	TISSUE SPECIFICITY: Expressed at high levels in the brain and pancreas.
Caspase-3 CASP-3 EC CASP3 CPP32	277 3.4.22.56	100	TISSUE SPECIFICITY: Highly expressed in lung, spleen, heart, liver and
Dipeptidyl peptidase DPP4 ADCP2 CD26	766 3.4.14.5	98	TISSUE SPECIFICITY: Expressed specifically in lymphatic vessels but not
Renin EC 3.4.23.15 A1 REN	406 3.4.23.15	81	
Glutamate carboxypeptidase FOLH1 FOLH NAALAC	750 3.4.17.21	63	TISSUE SPECIFICITY: Highly expressed in prostate epithelium. Detected
Leukotriene A-4 hydrolase LTA4H LTA4	611 3.3.2.6	52	TISSUE SPECIFICITY: Isoform 1 and isoform 2 are expressed in monocytes
Cathepsin K EC 3.4.21.29 CTSK CTSO CTSO2	329 3.4.22.38	52	TISSUE SPECIFICITY: Predominantly expressed in osteoclasts bones.
Insulin-degrading enzyme IDE	1019 3.4.24.56	40	
Angiotensin-converting enzyme ACE DCP DCP1	1306 3.2.1.17; 3.4.15.1	35	TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in lung
Cathepsin L1 EC 3.4.21.10 CTSL CTSL1	333 3.4.22.15	33	
Methionine aminopeptidase METAP1 KIAA0094	386 3.4.11.18	29	
Neutrophil collagenase MMP8 CLG1	467 3.4.24.34	25	TISSUE SPECIFICITY: Neutrophils.
Methionine aminopeptidase METAP2 MNPEP P67	478 3.4.11.18	24	
Proprotein convertase PCSK9 NARC1 PSECOX	692 3.4.21.10	21	TISSUE SPECIFICITY: Expressed in neuro-epithelioma, colon carcinoma
Complement factor 1 CF1 CF1A CF1B	253 3.4.21.46	20	
Neutrophil elastase ELANE ELA2	267 3.4.21.37	19	TISSUE SPECIFICITY: Bone marrow cells.
Chymase EC 3.4.21.33 CMA1 CYH CYM	247 3.4.21.39	18	TISSUE SPECIFICITY: Mast cells in lung, heart, skin and placenta. Expressed

Figure 6. A sort on a post-download Excel sheet for GtoPdb cross-references and proteases in UniProt, ranked by number of PDB entries.

284x94mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Entry	Entry name	Protein names	Gene names	Length	EC number	3D	Cross-reference (Orphanet)
P56817	BACE1_HUMAN	Beta-secretase 1	BACE1 BACE, KIAA1149	501	3.4.23.46	X-ray crystallography (357)	
P00797	RENI_HUMAN	Renin	REN	406	3.4.23.15	X-ray crystallography (81)	217330. Hyperuricemia - anemia - renal failure. 97369. Renal tubular dysgenesis of genetic origin.
P07339	CATD_HUMAN	Cathepsin D	CTSD CPSD	412	3.4.23.5	X-ray crystallography (6)	228337. CLN10 disease.
Q9Y5Z0	BACE2_HUMAN	Beta-secretase 2	BACE2 AEPLC, ALP56, ASP21, CDA13, UNQ418/PRO852	518	3.4.23.45	X-ray crystallography (12)	
P14091	CATE_HUMAN	Cathepsin E	CTSE	401	3.4.23.34	Model (1) X-ray crystallography (1)	
P20142	PEPC_HUMAN	Gastricsin	PGC	388	3.4.23.3	X-ray crystallography (2)	
P0DJ09	PEPA5_HUMAN	Pepsin A-5	PGA5	388	3.4.23.1		


Figure 7. Seven UniProt entries matching the InterPro domain "IPR001461 Aspartic_peptidase_A1" with GtoPdb cross-references, and showing links to the rare disease resource, Orphanet.

291x111mm (96 x 96 DPI)

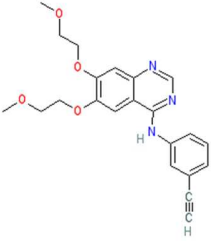
erlotinib


Ligand id: 4920

Name: erlotinib

 View more information in the IUPHAR Pharmacology Education Project: [erlotinib](#)

Structure and Physico-chemical Properties

2D Structure 

Calculated Physico-chemical Properties 

Hydrogen bond acceptors	5
Hydrogen bond donors	1
Rotatable bonds	10
Topological polar surface area	74.73
Molecular weight	393.17
XLogP	3.14
No. Lipinski's rules broken	0

Molecular properties generated using the CDK

Summary [Biological activity](#) [Screens](#) [Clinical data](#) [References](#) [Structure](#) [Similar ligands](#)

Classification

Compound class	Synthetic organic
Approved drug?	Yes (FDA (2004), EMA (2005))

IUPAC Name

N-(3-ethynylphenyl)-6,7-bis(2-methoxyethoxy)quinazolin-4-amine

International Nonproprietary Names

INN number	INN
8133	erlotinib




Synonyms

NSC 718781 | OSI 744 | OSI-774 | R 1415 | Tarceva®

Figure 8. Screenshot of the ligand summary page for erlotinib. Nomenclature information includes commonly-used synonyms and identifiers from the literature, the IUPAC name, English trade names, and the International Nonproprietary Name (INN). The ligand page also includes a depiction of the chemical structure, and under the Structure tab various structural identifiers such as the SMILES string, International Chemical Identifier (InChI), and InChIKey. Clicking on the structure image loads the structure into the chemical drawing tool which can be used for structure-based searches of the database.

249x247mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

Summary	Biological activity	Screens	Clinical data	References	Structure	Similar ligands
 View interactive charts of activity data from ChEMBL and GtoPdb across species (New!)						
Enzymes Catalysing Reactions with this Compound as a Substrate or Product						
Enzyme					Reference	
CYP3A4					3	
Selectivity at human catalytic receptors						
Key to terms and symbols Click column headers to sort						
Target	Type	Action	Affinity	Units	Reference	
epidermal growth factor receptor	 Inhibitor	Inhibition	9.2	pK _d	2	▼
epidermal growth factor receptor	 Inhibitor	Inhibition	9.0	pIC ₅₀	7	▼
Selectivity at human transporters						
Key to terms and symbols Click column headers to sort						
Target	Type	Action	Affinity	Units	Reference	
OATP2B1	Inhibitor	Inhibition	6.3	pK _i	5	▼
Ligand mentioned in the following text fields						
Type I RTKs: ErbB (epidermal growth factor) receptor family comments						

28 Figure 9. Data on quantitative ligand binding at targets are displayed under the Biological activity tab of the erlotinib ligand page. The primary target is indicated as epidermal growth factor receptor, with an additional
29 target of the OATP2B1 transporter. This drug is also known to be a substrate of the CYP3A4 enzyme.
30

31 250x161mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

Chemical structure search

1. Load or draw a structure into the editor below

To return all relevant hits please ensure that your input structure **does not** include chiral specification.

Marvin JS courtesy of Marvin JS user's guide [SMARTS help](#)

2. Choose type of search to perform

3. Limit results by chemical class

Search powered by from knowledge solutions

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 10. A SMILES structure has been loaded into the ligand structure drawing tool. The tool can perform searches by substructure, high or moderate similarity, exact match, or SMARTS pattern matching.

222x286mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Database Links	
CAS Registry No.	1286770-55-5
ChEMBL Ligand	CHEMBL3301601
PubChem CID	51352361
RCSB PDB Ligand	66F
Search Google for chemical match using the InChIKey	YHYKUSGACIYRMLKRWDZBOOSA-N
Search Google for chemicals with the same backbone	YHYKUSGACIYR
Search PubMed clinical trials	verubecestat
Search PubMed titles	verubecestat
Search PubMed titles/abstracts	verubecestat
Search UniChem for chemical match using the InChIKey	YHYKUSGACIYR
Search UniChem for chemicals with the same backbone	YHYKUSGACIYR

PubChem Compound Summary for CID 51352361

Verubecestat

STRUCTURE VENDORS DRUG-INFO LITERATURE PATENTS BIOACTIVITIES

PubChem CID: 51352361

Chemical Names: Verubecestat; UNII-J10PW7T7; J10PW7T7; 1286770-55-5; Verubecestat [USAN]; Verubecestat [USAN/INN] More...

Molecular Formula: $C_{14}H_{16}N_2O_5$

Molecular Weight: 408.412 g/mol

InChI Key: YHYKUSGACIYRMLKRWDZBOOSA-N

Drug Information: Clinical Trials FDA UNII

PUBCHEM > COMPOUND > VERUBECESTAT

Contents

- 1 2D Structure
- 2 3D Conformer
- 3 Names and Identifiers
- 4 Chemical and Physical Properties
- 5 Related Records
- 6 Chemical Vendors
- 7 Drug and Medication Information
- 8 Literature
- 9 Patents
- 10 Biomolecular Interactions and Pathways
- 11 Biological Test Results
- 12 Classification
- 13 Information Sources

1 2D Structure

Search Download Get Image

Stagnify

Figure 11. Connection between GtoPdb and PubChem. The GtoPdb ligand entry is 8699. The arrow indicates the click crossing to PubChem CID 51352361.

259x205mm (96 x 96 DPI)

+ Contents		5.2 Related Compounds	
1 2D Structure		Same Connectivity	3 records
2 3D Conformer		Same Stereo	2 records
3 Names and Identifiers		Same Isotope	2 records
4 Chemical and Physical Properties		Same Parent, Connectivity	4 records
● 5 Related Records		Same Parent, Stereo	3 records
6 Chemical Vendors		Same Parent, Isotope	3 records
7 Drug and Medication Information		Same Parent, Exact	2 records
8 Literature		Mixtures, Components, and Neutralized Forms	1 record
9 Patents		Similar Compounds	526 records
10 Biomolecular Interactions and Pathways		Similar Conformers	575 records
11 Biological Test Results			
12 Classification			
13 Information Sources			

5.3 Substances	
5.3.1 Related Substances	
All	21 records
Same	20 records
Mixture	1 record


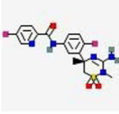
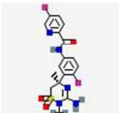
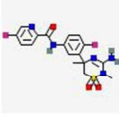
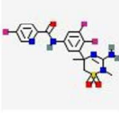
Figure 12. Navigation from CID 51352361 "Contents" (left hand facets) across to "Similar Compounds", highlighted in yellow.

203x198mm (96 x 96 DPI)

Links from PubChem Compound

Items: 1 to 20 of 684

<< First < Prev Page 1 of 35 Next > Last >>

1.  [Verubecestat: UNII-J1I0P6WT7T; J1I0P6WT7T ...](#)
 MW: 409.412 g/mol MF: C₁₇H₁₇F₂N₅O₃S
 IUPAC name: N-[3-[(5R)-3-amino-2,5-dimethyl-1,1-dioxo-6H-1,2,4-thiadiazin-5-yl]-phenyl]propanamide
 Create Date: 2011-05-16
 CID: 51352361
[Summary](#) [Similar Compounds](#) [Same Parent, Connectivity](#) [Mixture/Component Compounds](#) [PubMed \(MeSH Keyword\)](#)
2.  new record, indexing in progress
 CID: 124662750
[Summary](#)
3.  [SCHEMBL10328791](#)
 MW: 412.430 g/mol MF: C₁₇H₁₇F₂N₅O₃S
 IUPAC name: N-[3-[(5R)-3-amino-5-methyl-1,1-dioxo-2-(trideuteriomethyl)-5,6-dihydro-2H-1,2,4-thiadiazin-5-yl]-phenyl]propanamide
 Create Date: 2015-02-13
 CID: 88594172
[Summary](#) [Similar Compounds](#) [Same Parent, Connectivity](#)
4.  MW: 409.412 g/mol MF: C₁₇H₁₇F₂N₅O₃S
 IUPAC name: N-[3-(3-amino-2,5-dimethyl-1,1-dioxo-6H-1,2,4-thiadiazin-5-yl)-phenyl]propanamide
 Create Date: 2014-07-16
 CID: 75629988
[Summary](#) [Similar Compounds](#) [Same Parent, Connectivity](#)
5.  new record, indexing in progress
 CID: 77412075
[Summary](#)

Actions on your results

 **BioActivity Analysis**
 Analyze the BioActivities of the compounds

 **Structure Clustering**
 Cluster structures based on structural similarity

 **Structure Download**
 Download the structures in various formats

Refine your results

• What's this?

Chemical Properties

Rule of 5 (455)

BioActivity Experiments

BioAssays, Active (17)

BioAssays, Tested (17)

Protein 3D Structures (1)

Bace1 In Complex With (r)-n-(3-(3-amino-2,5-dimethyl-1,1-dioxido-5,6-dihydro-2h-1,2,4-thiadiazin-5-yl)-4-fluorophenyl)-5-fluoropicolinamide (1)

Depositor Category

Chemical Vendors (3)

Find related data

Database: PubMed

Option:

PubMed Citations

Related PubMed Citation

Find items

Figure 13. Showing the result of the "Similar Compounds" link from CID 51352361. Additional options and filters are shown in the right-hand facets.

216x189mm (96 x 96 DPI)

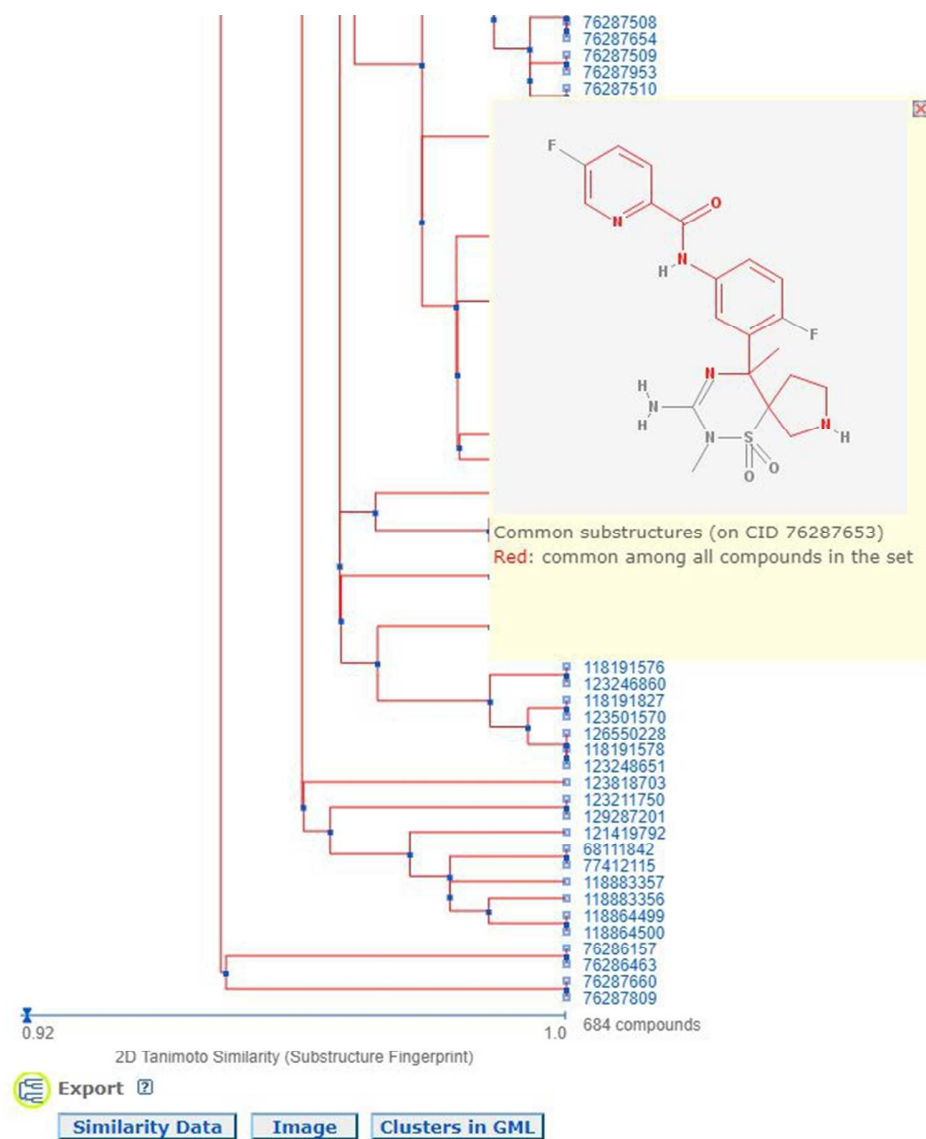


Figure 14. Clustering of all 684 "Similar Compounds". One CID is highlighted to show the substructure common to all members of the set. Note, the button "Similarity Data" produces a CSV matrix file of all the Tanimoto scores.

186x218mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

PubChem Structure Search PubChem Compound Limits Advanced search

Search By: Name/Text Identity/Similarity Substructure/Superstructure Molecular Formula

Draw a Structure CID, SMILES, InChI Structure File

51352361 Edit Search Preview

Enter single structure identifier

Options

Similar Structures threshold >= 95% 99% 98% 97% 96% 95% 93% 90% 85% 80% 70% 60%

Time Limit(seconds): unlimited hit 2,000,000

Filters

Clear Save Queue

35 Figure 15. The PubChem structure search page set up with CID 51352361. The Tanimoto options for
36 similarity searching are shown with 96% set as the threshold.

37 213x186mm (96 x 96 DPI)

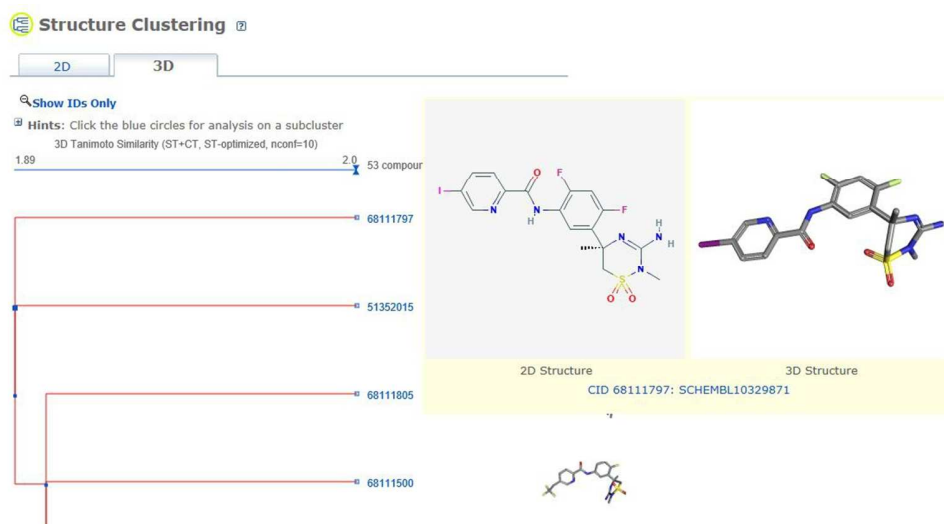
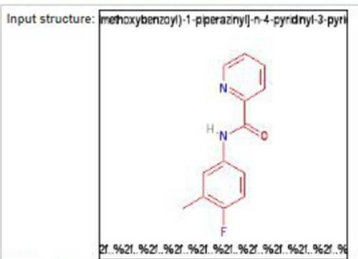


Figure 16. 3D clustering of the 74 compounds 97% similar to CID 51352361. The options were set to display the conformer renderings and large-size 2D/3D juxtaposed displays for the mouseovers.

297x160mm (96 x 96 DPI)

Chemical structure search results

Input structure: **methoxybenzoyl-1-piperazinyl-n-4-pyridyl-3-pyr**



Input SMILES:
[H]N(C(=O)C1=NC=CC=C1)C1=CC(C)=C(F)C=C1

Your query returned 7 matches:

Order results by: **Default** Limit results by: **None** <=> 0.0 **Update**

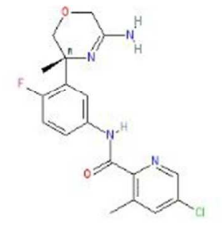
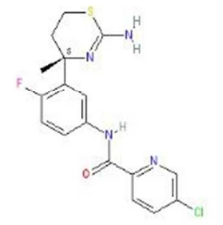
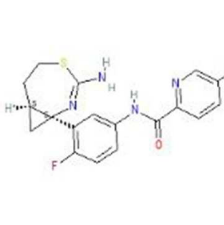

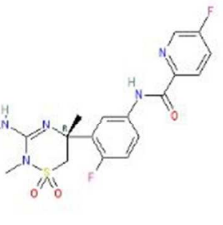
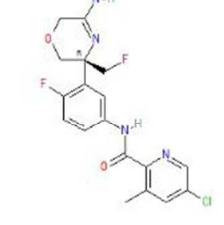
<p>example 98 [WQ2011020808]</p>  <p>Targets beta-site APP-cleaving enzyme 2</p> <p>Use in search</p>	<p>compound J [PMID: 21907142]</p>  <p>Targets beta-secretase 1; beta-site APP-cleaving enzyme 2</p> <p>Use in search</p>	<p>example 2 [WQ2013004878]</p>  <p>Targets beta-secretase 1; beta-site APP-cleaving enzyme 2</p> <p>Use in search</p>
<p>LY2889721</p>  <p>Targets beta-secretase 1</p> <p>Use in search</p>	<p>verubecestat</p>  <p>Targets beta-secretase 1; beta-site APP-cleaving enzyme 2</p> <p>Use in search</p>	<p>example 92 [WQ2012065521]</p>  <p>Targets beta-site APP-cleaving enzyme 2</p> <p>Use in search</p>

Figure 17. Showing the results of editing CID 51352361 and searching GtoPdb by substructure using the SMILES string FC2=C(C)C=C(N(C(=O)C1=NC=CC=C1)[H])C=C2. Six of the seven matches are shown.

184x231mm (96 x 96 DPI)

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

```

```

  {
    targetId: 240,
    name: "GABA<sub>B1</sub>",
    abbreviation: null,
    systematicName: null,
    type: "GPCR",
    familyIds: [
      26
    ],
    subunitIds: [ ],
    complexIds: [
      242
    ]
  },
  {
    targetId: 241,
    name: "GABA<sub>B2</sub>",
    abbreviation: null,
    systematicName: null,
    type: "GPCR",
    familyIds: [
      26
    ],
    subunitIds: [ ],
    complexIds: [
      242
    ]
  },
  {
    targetId: 242,
    name: "GABA<sub>B</sub> receptor",
    abbreviation: null,
    systematicName: null,
    type: "GPCR",
    familyIds: [
      26
    ],
    subunitIds: [
      1917,
      1918,
      1919,
      1920,
      240,
      241
    ],
    complexIds: [ ]
  },

```

Figure 18. A section of the JSON output from a web service call to generate a list of GPCRs. Each target is shown with summary information on its nomenclature, family IDs, component subunit IDs if it is a complex (e.g. the GABAB receptor), and complex IDs if it is a component subunit (e.g. GABAB1 and GABAB2). Names include HTML codes to represent formatting such as subscript (the sub tag), and special characters such as the Greek alpha.

93x199mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

IUPHAR
Guide to IMMUNOPHARMACOLOGY *beta v2.0*

Search Database

Home About Targets Ligands Processes Cell Types Disease Resources Guide to PHARMACOLOGY

This is beta-release v2.0 of the GtoImmuPdb. It contains the majority of features and functionality expected in the full public release. However, it remains under development and while it should not contain any critical bugs, some portions are not yet optimised and may lack full functionality or content. [Report a bug](#)

Processes/Pathways

- Barrier integrity
- Inflammation
- Antigen presentation
- T cell (activation)
- B cell (activation)
- Immune regulation
- Tissue repair
- Immune system development
- Cytokine production & signalling
- Chemotaxis & migration
- Cellular signalling
- [View immuno processes home page](#)

Cell Types

- B-cells
- T-cells (alpha/beta)
- Dendritic cells
- Other T-cells (NKT, MAIT, TRM etc.)
- Macrophages & monocytes
- Granulocytes
- Natural killer (NK) cells
- Mast cells
- Innate lymphoid cells
- Stromal cells
- [View immuno cell types home page](#)

Latest Updates & Help

Latest updates:

Beta-release v2.0 released 22nd Aug 2017. See the blog post for full details.
[View beta v2.0 release notes.](#)

Guide

[View our tutorial on navigating the Guide to IMMUNOPHARMACOLOGY.](#)

Help

[View our help page for information about the data in and using the Guide to IMMUNOPHARMACOLOGY.](#)

Targets

- G protein-coupled receptors
- Ion channels
- Nuclear hormone receptors
- Kinases
- Catalytic receptors
- Transporters
- Enzymes
- Other protein targets
- [Search for targets](#)

Ligands

- Immuno ligands
- Antibodies
- Approved drugs
- Synthetic organics
- Metabolites
- Natural products
- Endogenous peptides
- Other peptides
- Inorganics
- Labelled ligands
- [Search for ligands](#)

Disease

- Immuno Disease to Target Associations
- Immuno Disease to Ligand Associations

News

- [View GtoImmuPdb blog posts](#)
- [Introducing GtoImmuPdb at Pharmacology 2017 \(slideshare\)](#)
- [Latest Guide to PHARMACOLOGY news](#)

Figure 19. The Guide to IMMUNOPHARMACOLOGY portal, showing the entry points to processes, cell types, targets and ligands.

258x269mm (96 x 96 DPI)

The IUPHAR Guide to IMMUNOPHARMACOLOGY Disease list

Targets Ligands

List of immunological diseases and the biological targets contained in GtoPdb they are associated with. [Show all targets](#) [Hide all targets](#)

Disease: Acute lymphocytic leukemia (ALL) **2 target associations:** CD3e; CD19 [more details »](#)
Disease X-Refs: OMIM: 613065

Disease: Asthma **5 target associations:** MMP2; MMP9; MMP8 ... [more details »](#)
Disease X-Refs: Disease Ontology: [DOID:2841](#)
 OMIM: 600807

Disease: Atopic dermatitis **1 target association:** INPP5D [more details »](#)
Disease X-Refs: Disease Ontology: [DOID:3310](#)
 OMIM: 603165

Disease: Autoimmune lymphoproliferative syndrome; ALPS **1 target association:** Fas [more details »](#)
Disease X-Refs: Disease Ontology: [DOID:6688](#)
 OMIM: 601859
 Orphanet: [ORPHA3261](#)

Disease: Chronic lymphocytic leukemia **1 target association:** CD20 (membrane-spanning 4-domains, subfamily A, member 1) [« hide details](#)
Disease X-Refs: Disease Ontology: [DOID:1040](#)
 OMIM: 151400
 Orphanet: [ORPHA67038](#)






Target	Comments	Ligands <i>key to symbols</i>	References												
CD20 (membrane-spanning 4-domains, subfamily A, member 1)	CD20 is the molecular target of the CLL therapeutics ofatumumab and obinutuzumab.	<table border="1"> <tr> <td>ofatumumab</td> <td></td> <td></td> </tr> <tr> <td>rituximab</td> <td></td> <td></td> </tr> <tr> <td>obinutuzumab</td> <td></td> <td></td> </tr> <tr> <td>veltuzumab</td> <td></td> <td></td> </tr> </table>	ofatumumab			rituximab			obinutuzumab			veltuzumab			
ofatumumab															
rituximab															
obinutuzumab															
veltuzumab															

Figure 20. The GtoImmuPdb disease-target associations page. Users can switch to view disease-ligand associations using the tab at the top.

251x212mm (96 x 96 DPI)

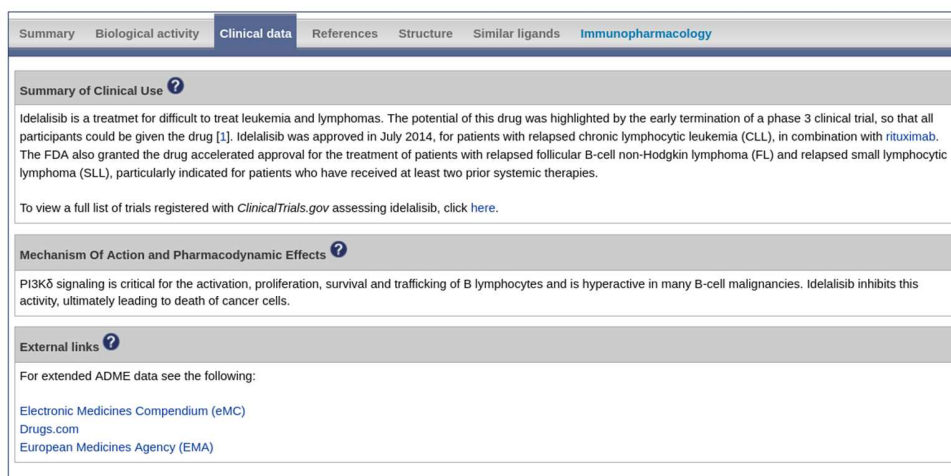
Disease: **Chronic lymphocytic leukemia** **7 ligand associations:** ofatumumab; veltuzumab; idelalisib; rituximab; ibrutinib; obinutuzumab; milatuzumab [« hide details](#)

Disease X-Refs: Disease Ontology: DOID:1040
OMIM: 151400
Orphanet: ORPHA67038

Ligand <i>To view more details about a ligand, and its activity, click on its name to view the ligand summary page</i>	Disease Association Comments	Clinical Use	References
ofatumumab  [FDA (2009), EMA (2010) (Drugs.com)]	Approved drug for some CLL patients.	Approved to treat chronic lymphocytic leukemia that is refractory to fludarabine ...	
veltuzumab	FDA and EMA orphan drug for CLL.	The EMA and FDA have granted veltuzumab orphan designation for the treatment of the rare disease chronic lymphocytic leukemia (CLL).	
idelalisib  [FDA & EMA (2014) (Drugs.com)]	Approved in combination with rituximab for CLL. Also used to treat relapsed small lymphocytic leukemia (SLL).	Idelalisib is a treatment for difficult to treat leukemia and lymphomas. The potential of this drug was highlighted by the early termination of	
rituximab  [FDA (1997), EMA (1998) (Drugs.com)]	An anti-CD20 therapy approved for CLL and non-Hodgkins lymphoma.	Used to treat CD20-positive non-Hodgkins lymphoma, chronic lymphocytic leukemia, and rheumatoid arthritis. Rituximab is also used to	
ibrutinib  [FDA and EMA (2013) (Drugs.com)]	Approved drug for CLL.	Ibrutinib is approved to treat patients with mantle cell lymphoma (MCL), a rare and aggressive type of leukemia, especially	
obinutuzumab  [FDA (2013), EMA (2014) (Drugs.com)]	Approved for use in combination with chemotherapy in patients who have received no prior therapy.	Initially approved to treat chronic lymphocytic leukemia (CLL) in combination with chemotherapy in patients who have received	
milatuzumab	FDA and EMA orphan drug for CLL and multiple myeloma; in Phase 2 trial.	Both the FDA and EMA have granted milatuzumab orphan designation for the treatment of the rare diseases multiple	

299x218mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22



The screenshot displays the 'Clinical data' tab of the GtoPdb ligand page for idelalisib. The page is organized into three main sections: 'Summary of Clinical Use', 'Mechanism Of Action and Pharmacodynamic Effects', and 'External links'. The 'Summary of Clinical Use' section provides a brief overview of the drug's application in treating leukemia and lymphomas, mentioning its approval in July 2014 and its use in combination with rituximab. The 'Mechanism Of Action and Pharmacodynamic Effects' section explains that PI3Kδ signaling is critical for B lymphocyte activity and that idelalisib inhibits this pathway. The 'External links' section offers resources for extended ADME data, including the Electronic Medicines Compendium (eMC), Drugs.com, and the European Medicines Agency (EMA).

Summary	Biological activity	Clinical data	References	Structure	Similar ligands	Immunopharmacology
Summary of Clinical Use ?						
Idelalisib is a treatment for difficult to treat leukemia and lymphomas. The potential of this drug was highlighted by the early termination of a phase 3 clinical trial, so that all participants could be given the drug [1]. Idelalisib was approved in July 2014, for patients with relapsed chronic lymphocytic leukemia (CLL), in combination with rituximab. The FDA also granted the drug accelerated approval for the treatment of patients with relapsed follicular B-cell non-Hodgkin lymphoma (FL) and relapsed small lymphocytic lymphoma (SLL), particularly indicated for patients who have received at least two prior systemic therapies.						
To view a full list of trials registered with <i>ClinicalTrials.gov</i> assessing idelalisib, click here .						
Mechanism Of Action and Pharmacodynamic Effects ?						
PI3Kδ signaling is critical for the activation, proliferation, survival and trafficking of B lymphocytes and is hyperactive in many B-cell malignancies. Idelalisib inhibits this activity, ultimately leading to death of cancer cells.						
External links ?						
For extended ADME data see the following:						
Electronic Medicines Compendium (eMC)						
Drugs.com						
European Medicines Agency (EMA)						

23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 22. A summary of clinical data for the drug idelalisib, from its GtoPdb ligand page.

315x156mm (96 x 96 DPI)

Search results Page 1 of 11

Your search for **Alzheimer's** returned 109 results

Order results by:

Target: GPR3 (Class A Orphans)

Consequences of altering gene expression: implicated in **Alzheimer's** disease and is elevated in the sporadic **Alzheimer's** disease brain.
 Consequences of altering gene expression: accumulation of the amyloid-beta peptide in vitro and in an **Alzheimer's** disease mouse model
 Mutations/pathophysiology - disease: **Alzheimer** disease
 Mutations/pathophysiology - disease: **Alzheimer's** disease aka Alzheimer disease
 Mutations/pathophysiology - role: neurons and may be a potential therapeutic target for the treatment of **Alzheimer's** disease

Target: programmed cell death 1 (CD279) (Other immune checkpoint proteins, CD molecules)

Mutations/pathophysiology - disease: **Alzheimer** disease
 Mutations/pathophysiology - disease: **Alzheimer's** disease aka Alzheimer disease
 Mutations/pathophysiology - role: **Alzheimer's** disease anti-PD1 antibody treatment was used to induce immune checkpoint blockade. The biological
 Receptor Immuno Comment: **Alzheimer's** disease (AD): In mouse models of AD anti-PD-1 antibody treatment was used

Target: presenilin 2 (A22: Presenilin)

Synonym: presenilin 2 (**Alzheimer** disease 4)
 Mutations/pathophysiology - disease: **Alzheimer** disease 4
 Mutations/pathophysiology - disease: Early-onset autosomal dominant **Alzheimer** disease aka Alzheimer disease 4
 Mutations/pathophysiology - disease: **Alzheimer's** disease aka Alzheimer disease 4

Ligand: leucomethylthionium

Comments: **Alzheimer's** disease (AD) and other tau-opathies. We believe this is the compound being
 Clinical use: **Alzheimer's** Association International Conference in Toronto (July 2016) and reported on the NewScientist website
 Mechanism of action: **Alzheimer's** plaques and filaments respectively. Inhibition of the formation of tau filaments/tangles is one strategy

Ligand: flortaucipir (18F)

Comments: P10636) in **Alzheimer's** disease brains. Compare with the approved agent, , used to visualise beta-amyloid.
 Clinical use: NCT02016560, in cognitively healthy volunteers and patients with mild cognitive impairment and **Alzheimer's** disease.
 Bioactivity comments: **Alzheimer's** disease patients . Flortaucipir (18F) signal co-localises with beta amyloid, but does not bind

Ligand: florbetaben

Comments: **Alzheimer's** disease (AD) . and are the two other PET reagents approved for Abeta detection
 Clinical use: Abeta) as a diagnostic biomarker in the brains of patients with **Alzheimer's** disease (AD) .
 Mechanism of action: **Alzheimer's** disease. The 18F radioactive tracer can then be detected by PET scanning and its accumulation

Target: presenilin 1 (A22: Presenilin)

Mutations/pathophysiology - disease: **Alzheimer** disease 3
 Mutations/pathophysiology - disease: Early-onset autosomal dominant **Alzheimer** disease aka Alzheimer disease 3
 Mutations/pathophysiology - disease: **Alzheimer's** disease aka Alzheimer disease 3

Figure 23. A section of the search results for the disease term "Alzheimer's". The results are ordered by best match, with the top result (target GPR3) having 6 matches across 5 database fields.

250x238mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Immuno Cell Types

The Guide to IMMUNOPHARMACOLOGY provides lists of targets associated to top-level cell type categories. These categories are described below, showing Cell Ontology terms (and definitions) that have been assigned to them. Click on the category title to view the detailed list of target associations. The GtoImmuPdb help pages provide more information.

Jump to: [B cells](#) | [T cells](#) | [Dendritic cells](#) | [Other T cells](#) | [Macrophages & monocytes](#) | [Granulocytes](#) | [Natural killer cells](#) | [Mast cells](#) | [Innate lymphoid cells](#) | [Stromal cells](#) |

B cells

The *B cells* category includes the following Cell Ontology parent terms:

lymphocyte of B lineage (CL:0000945) - A lymphocyte of B lineage is a lymphocyte that expresses CD19 on the cell surface. An additional defining characteristic is the commitment to express an immunoglobulin complex. Types of B lineage lymphocytes include B cells and antibody secreting cells (plasmablasts and plasma cells). Lymphocytes of B cell lineage can be distinguished from those of T cell lineage by their lack of CD3e (as part of the T cell receptor complex).

T cells

The *T cells* category includes the following Cell Ontology parent terms:

alpha-beta T cell (CL:0000789) - A T cell that expresses an alpha-beta T cell receptor complex.
effector T cell (CL:0000911) - A differentiated T cell with ability to traffic to peripheral tissues and is capable of mounting a specific immune response.
regulatory T cell (CL:0000815) - A T cell which regulates overall immune responses as well as the responses of other T cell subsets through direct cell-cell contact and cytokine release.

This should encompass most cytotoxic and helper T cells, as well as regulatory T cells.

Dendritic cells

The *Dendritic cells* category includes the following Cell Ontology parent terms:

dendritic cell (CL:0000451) - A cell of hematopoietic origin, typically resident in particular tissues, specialized in the uptake, processing, and transport of antigens to lymph nodes for the purpose of stimulating an immune response via T cell activation. These cells are lineage negative (CD3-negative, CD19-negative, CD34-negative, and CD56-negative).

29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 24. The GtoImmuPdb immunological cell types home page, showing the list of top-level cell type categories.

314x212mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

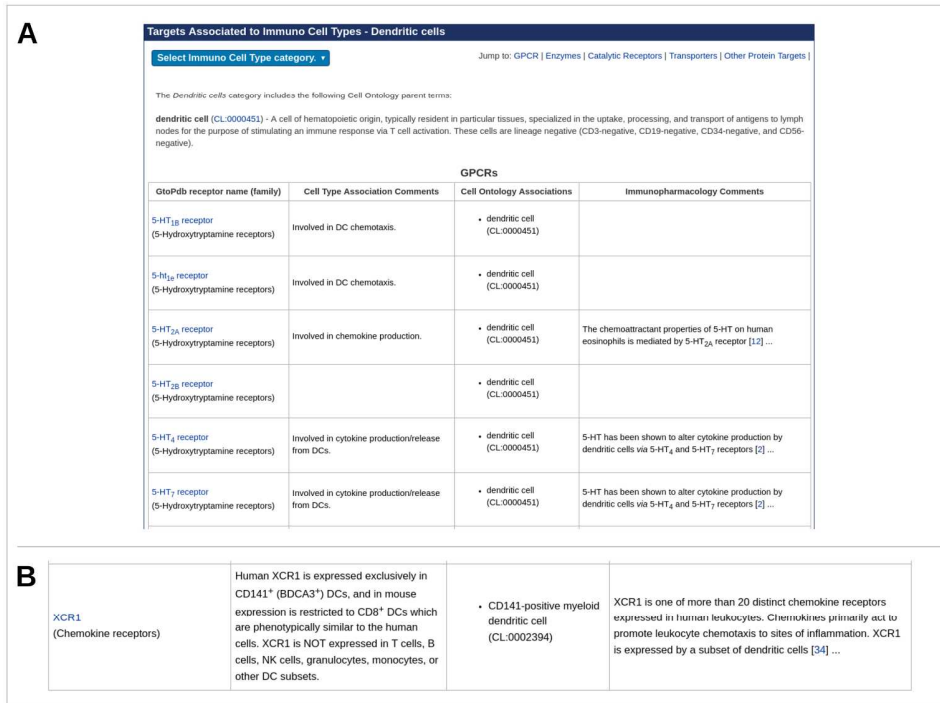


Figure 25. (A) The list of GtoImmuPdb target associations to dendritic cells. Users can switch cell type using the drop down menu at the top left. (B) A close-up of the association between the target XCR1 and dendritic cells. The second column shows curator comments, the third column gives the Cell Ontology terms, and the final column shows an extract of the general immunopharmacology comments on XCR1.

624x468mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Immuno Processes

The Guide to IMMUNOPHARMACOLOGY provides lists of targets associated to top-level process categories. These categories are described below, showing Gene Ontology biological process terms that have been assigned to them. Click on the category title to view the detailed list of target associations. The [GtoImmuPdb help](#) pages provide more information.

Jump to: [Barrier integrity](#) | [Inflammation](#) | [Antigen presentation](#) | [T cell \(activation\)](#) | [B cell \(activation\)](#) | [Immune regulation](#) | [Tissue repair](#) | [Immune system development](#) | [Cytokine production & signalling](#) | [Chemotaxis & migration](#) | [Cellular signalling](#) |

Barrier integrity

Barrier integrity encompasses biological processes included under, but not limited to, the Gene Ontology terms:

- [entry into host \(GO:0044409\)](#)
- [epithelial cell-cell adhesion \(GO:0090136\)](#)

Inflammation

Inflammation encompasses biological processes included under, but not limited to, the Gene Ontology terms:

- [plasmacytoid dendritic cell activation \(GO:0002270\)](#)
- [myeloid leukocyte activation \(GO:0002274\)](#)
- [leukocyte activation involved in immune response \(GO:0002366\)](#)
- [T cell cytokine production \(GO:0002369\)](#)
- [production of molecular mediator involved in inflammatory response \(GO:0002532\)](#)
- [phagocytosis \(GO:0006909\)](#)
- [inflammatory response \(GO:0006954\)](#)
- [pattern recognition receptor activity \(GO:0038187\)](#)
- [innate immune response \(GO:0045087\)](#)
- [regulation of inflammatory response \(GO:0050727\)](#)
- [leukocyte migration \(GO:0050900\)](#)
- [leukocyte adhesive activation \(GO:0050902\)](#)

Inflammation also includes biological process relating to inflammation, including pattern recognition, macrophage activation, phagocytosis and granulopoiesis.

Antigen presentation

Antigen presentation encompasses biological processes included under, but not limited to, the Gene Ontology terms:

- [antigen sampling in mucosal-associated lymphoid tissue \(GO:0002404\)](#)
- [endocytosis \(GO:0006897\)](#)
- [antigen processing and presentation \(GO:0019882\)](#)

Figure 26. The immunological processes home page, showing the definitions and mapped GO terms for the first few top-level process categories. Quick links to the categories are provided at the top.

251x207mm (96 x 96 DPI)

A

Targets Associated to Immuno Processes - Chemotaxis & migration

Select Immuno Process category... Jump to: GPCR | Ion Channels | NHRs | Enzymes | Catalytic Receptors | Other Protein Targets |

Chemotaxis & migration

GPCRs

GoPdb receptor name (family)	Process Association Comments	GO Associations	Immunopharmacology Comments
ACKR1 (Chemokine receptors)		<ul style="list-style-type: none"> regulation of chemokine production (GO:0032642) IBA 	ACKR1 is one of more than 20 distinct chemokine receptors expressed in human leukocytes. Chemokines primarily act to promote leukocyte chemotaxis to sites of inflammation.
ADGRE2 (Adhesion Class GPCRs)		<ul style="list-style-type: none"> granulocyte chemotaxis (GO:0071621) IMP 	ADGRE2 is included in GtoImmuPdb because its tissue expression profile and its ability to activate secretion of inflammatory cytokines (IL-8, TNF) by monocytes and macrophages [81] ...
A _{2A} receptor (Adenosine receptors)		<ul style="list-style-type: none"> negative regulation of alpha-beta T cell activation (GO:0046636) IEA 	A _{2A} receptor is discussed in this immuno-oncology review [2] ...
A _{2B} receptor (Adenosine receptors)		<ul style="list-style-type: none"> positive regulation of chemokine production (GO:0032722) IEA 	A _{2B} receptor is discussed in this immuno-oncology review [2] ...

B

integrin, alpha 4 subunit (antigen CD49D, alpha 4 subunit of VLA-4 receptor) (Integrins)		<ul style="list-style-type: none"> leukocyte cell-cell adhesion (GO:0007159) IDA leukocyte migration (GO:0050900) TAS leukocyte tethering or rolling (GO:0050901) IMP diapedesis (GO:0050904) IEA positive regulation of leukocyte tethering or rolling (GO:1903238) IEA positive regulation of T cell migration (GO:2000406) IEA 	Integrin subunit alpha 4 is the alpha subunit of the α4β1 lymphocyte homing receptor. The cytoplasmic domain of α4 binds tightly to paxillin, a signaling adaptor protein, and this interaction promotes increased cell migration and inhibits cell spreading [73] ...
integrin, alpha L subunit (antigen CD11A (p180), lymphocyte function-associated antigen 1; alpha polypeptide) (Integrins)		<ul style="list-style-type: none"> leukocyte cell-cell adhesion (GO:0007159) IMP leukocyte migration (GO:0050900) TAS 	ITGAL associates with the beta 2 chain (ITGB2) integrin and CD18 to form the lymphocyte function-associated antigen-1 (LFA-1) complex. LFA-1 is crucial for leukocyte intercellular adhesion and costimulatory signaling in the immune system. Ligands for LFA-1 are the intercellular adhesion molecules (ICAMs) 1-3. ITGAL is the target of the withdrawn psoriasis monoclonal antibody drug etalizumab ...
integrin, beta 2 subunit (complement component 3 receptor 3 and 4 subunit) (Integrins)		<ul style="list-style-type: none"> leukocyte migration involved in inflammatory response (GO:0002523) IEA leukocyte cell-cell adhesion (GO:0007159) IDA neutrophil chemotaxis (GO:0030593) IDA cellular extravasation (GO:0045123) IEA leukocyte migration (GO:0050900) TAS 	ITGB2 is the beta component of the β2 integrins. It forms αβ heterodimers with the CD11 alpha subunits ITGAL (αL aka CD11a forming integrin LFA-1), ITGAM (αM aka CD11b, forming integrin Mac-1) and ITGAX (αX aka CD11c, forming integrin p150/95). These integrins are essential for leukocyte extravasation to sites of infection. Their absence causes a leukocyte adhesion deficiency, which manifests clinically as recurrent severe infections, defective wound healing and neutrophilia [102] ...

Figure 27. (A) The list of target associations for chemotaxis and migration. The third column shows the GO immunological process terms mapped to the targets named in the left column. The final column summarises the general immunopharmacology comments for the target. (B) The process GO terms and comments for the integrins.

673x642mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

Search for data by literature reference

Enter title keyword, author name or PubMed Id:

Select field to search:

Home Target search Search results

Search results Page 1 of 2

Your search for **7651358** returned 17 results

Order results by:

Ligand: WB 4101
Reference: Shibata K, Foglar R, Horie K, Obika K, Sakamoto A, Ogawa S, Tsujimoto G (1995) KMD-3213, a novel, potent, alpha 1a-adrenoceptor-selective antagonist: characterization using recombinant human alpha 1-adrenoceptors and native tissues. *Mol. Pharmacol.* 48: 250-258 [PMID:7651358]

Ligand: tamsulosin
Reference: Shibata K, Foglar R, Horie K, Obika K, Sakamoto A, Ogawa S, Tsujimoto G (1995) KMD-3213, a novel, potent, alpha 1a-adrenoceptor-selective antagonist: characterization using recombinant human alpha 1-adrenoceptors and native tissues. *Mol. Pharmacol.* 48: 250-258 [PMID:7651358]

Ligand: silodosin
Reference: Shibata K, Foglar R, Horie K, Obika K, Sakamoto A, Ogawa S, Tsujimoto G (1995) KMD-3213, a novel, potent, alpha 1a-adrenoceptor-selective antagonist: characterization using recombinant human alpha 1-adrenoceptors and native tissues. *Mol. Pharmacol.* 48: 250-258 [PMID:7651358]

Ligand: S(+)-niguldipine
Reference: Shibata K, Foglar R, Horie K, Obika K, Sakamoto A, Ogawa S, Tsujimoto G (1995) KMD-3213, a novel, potent, alpha 1a-adrenoceptor-selective antagonist: characterization using recombinant human alpha 1-adrenoceptors and native tissues. *Mol. Pharmacol.* 48: 250-258 [PMID:7651358]

Ligand: prazosin
Reference: Shibata K, Foglar R, Horie K, Obika K, Sakamoto A, Ogawa S, Tsujimoto G (1995) KMD-3213, a novel, potent, alpha 1a-adrenoceptor-selective antagonist: characterization using recombinant human alpha 1-adrenoceptors and native tissues. *Mol. Pharmacol.* 48: 250-258 [PMID:7651358]

6. Shibata K, Foglar R, Horie K, Obika K, Sakamoto A, Ogawa S, Tsujimoto G. (1995)
KMD-3213, a novel, potent, alpha 1a-adrenoceptor-selective antagonist: characterization using recombinant human alpha 1-adrenoceptors and native tissues. *Mol. Pharmacol.*, 48: 250-258. [PMID:7651358]

36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 28. The literature search box from the GtoPdb Advanced search page is shown at the top. The initial section of the result list is in the centre, and the reference as shown in each of the linked entries is at the bottom. The PMID in blue is live-linked.

227x203mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The screenshot displays a PubMed article interface. On the left, under 'LinkOut - more resources', there are sections for 'Full Text Sources' (HighWire - PDF), 'Other Literature Sources' (Cited by Patents in - The Lens), and 'Molecular Biology Databases'. The latter section lists several links to 'Hazardous Substances Data Bank' for various substances like alpha1A-adrenoceptor, alpha1B-adrenoceptor, alpha1D-adrenoceptor, and Adrenoceptors, along with a 'Guide to Pharmacology' link. A yellow box highlights the 'Related PubChem Substance' link. On the right, the article title is 'The Persistence of Silodosin Monotherapy and the Reasons for Withdrawal from [Adv Urol. 2017]'. Below the title, there is a 'Review' link and another article reference. A 'See all...' link is also present. The 'Related information' section includes links for 'Articles frequently viewed together', 'Cited in systematic reviews', 'MedGen', 'PubChem Compound', 'PubChem Compound (MeSH Keyword)', 'PubChem Substance', and 'Cited in PMC'. At the bottom right, there is a 'Recent Activity' section with 'Turn Off' and 'Clear' options.

Figure 29. Link displays from <https://www.ncbi.nlm.nih.gov/pubmed/7651358>. The GtoPdb targets are linked via Molecular Biology Databases under LinkOut resources. PubChem links for GtoPdb ligands can be found on the right under Related information.

310x118mm (96 x 96 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



1. [norepinephrine, noradrenaline, Levorfenol...](#)
Source: [IUPHAR/BPS Guide to PHARMACOLOGY](#)
Deposit Date: 2012-04-25 Available Date: 2012-04-25 Modify Date: 2017-05-30
SID: 135651527 [CID: 439260]
[Summary](#) [PubChem Same Compound](#)

2. [\(S\)-adrenaline, \(+\)-adrenaline, Racepinefrine...](#)
Source: [IUPHAR/BPS Guide to PHARMACOLOGY](#)
Deposit Date: 2012-04-25 Available Date: 2012-04-25 Modify Date: 2017-05-30
SID: 135651431 [CID: 247704]
[Summary](#) [PubChem Same Compound](#)

3. [epinephrine, l-Adrenaline, l-epinephrine...](#)
Source: [IUPHAR/BPS Guide to PHARMACOLOGY](#)
Deposit Date: 2012-04-25 Available Date: 2012-04-25 Modify Date: 2017-05-30
SID: 135651430 [CID: 5816]
[Summary](#) [PubChem Same Compound](#)

4. [TAMSULOSIN, Flomax, Omnic...](#)
Source: [IUPHAR/BPS Guide to PHARMACOLOGY](#)
Deposit Date: 2012-04-25 Available Date: 2012-04-25 Modify Date: 2017-05-30
SID: 135651350 [CID: 129211]
[Summary](#) [PubChem Same Compound](#)

Actions on your results

- BioActivity Analysis
Analyze the BioActivities of the substances
- Structure Clustering
Cluster structures based on structural similarity
- Structure Download
Download the structures in various formats

Refine your results - What's new?

BioActivity Experiments
BioAssays, Active (1)
BioAssays, Tested (1)

Find related data

Database | Select

Find Items

Recent activity

Figure 30. Some of the PubChem SID links from PMID 7651358 (as shown in Figure 29). All four of these structures link directly back to GtoPdb.

310x125mm (96 x 96 DPI)