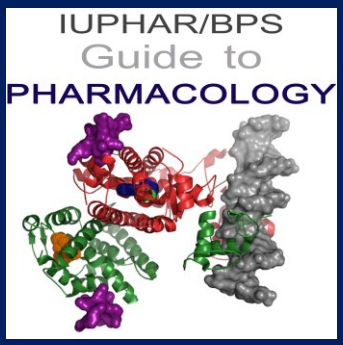
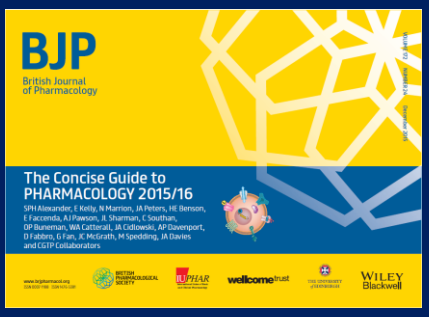
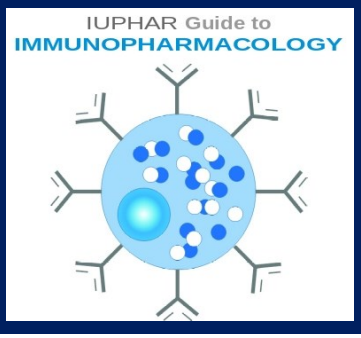
IUPHAR
International Union of Basic and Clinical Pharmacology

BRITISH PHARMACOLOGICAL SOCIETY







DATABASE Report

April 2017

enquiries@guidetopharmacology.org

TABLE OF CONTENTS

INTRODUCTION

This April 2017 database report provides an overview of recent progress and the current status of the IUPHAR/BPS Guide to PHARMACOLOGY (GtoPdb) since the last NC-IUPHAR meeting held in Paris in October 2016.

We are now 17 months into our three year Wellcome Trust funded project to develop the "The Guide to IMMUNOPHARMACOLOGY (GtoImmuPdb): Integration of targets, diseases and therapies into an expert-driven database". This grant began on the 1st of November 2015. More details can be found in our blog, which includes technical blog posts highlighting aspects of the development of the resource.

This report (along with the accompanying slide set) will detail our progress on the GtoPdb and GtoImmuPdb projects. Please talk to us to follow-up any points raised in this document, issues and suggestions.

GENERAL OVERVIEW OF DATABASE TEAM ACTIVITIES

PUBLIC ENGAGEMENT – PROMOTING OUR RESOURCES

CONFERENCES/MEETINGS (SINCE OCTOBER 2016 AND UPCOMING)

- All Africa Congress on Pharmacology and Pharmacy, South Africa, October 2016, Adam Pawson
- 12th German Conference on Chemoinformatics, November 2016, Christopher Southan (slides)
- Global Health Compound Design Webinar, Nov 2016, Christopher Southan (slides)
- Webinar for ELIXIR UK weekly group conference, December 2016 (slides)
- BPS Pharmacology 2016, December 2016, two abstracts from the team, Simon Harding (poster & flash poster presentation), Christopher Southan (talk)
- Annual lecture (with practical) to the Edinburgh 4th year Pharmacology students, Jan 2017, Chris Southan (invited by Prof Mark Evans slides)
- Members of the Edinburgh and IUPHAR teams presented a webinar (IUPHAR Web Resources – Simplifying Complexity for Medicine and Education) in the ICSU World Data System series on 28 Feb 2017 (slides and recording)
- ITMAT in Edinburgh, March 2017, Adam Pawson, Christopher Southan (poster)
- BiVi, the Biological Visualisation Community, 3rd Annual Meeting in Edinburgh, April 2017, Joanna Sharman
- BPS, In silico and in vitro methods in modern drug discovery, Nottingham, April 2017, Steve Alexander (poster)
- BioIT World, Boston, May 2017, Christopher Southan (patent mining workshop and poster)

Our slideshare account includes/will include slide sets and posters presented by team members. Some are also posted on Christopher Southan's own slideshare

PUBLICATIONS

PUBLISHED (SINCE OCTOBER 2016)

- Book chapter: **Small-molecule Bioactivity Databases**. Sean Ekins, Alex M. Clark, Christopher Southan, Barry A. Bunin and Antony J. Williams. Chapter 16 in: High Throughput Screening Methods: Evolution and Refinement, 2017, P344 – 365, Nathan Ross and Joshua Bittker, Eds, Royal Society of Chemistry, DOI:10.1039/9781782626770-00344.
- Review: **Last rolls of the yoyo: Assessing the human canonical protein count**. Christopher Southan, F1000Research 2017, 6:448 (doi: 10.12688/f1000research.11119.1) version 1, awaiting open peer review.

IN PRESS/SUBMITTED/IN PREPARATION

- Book chapter**: Examples of SAR-centric patent mining using open resources**, Christopher Southan. In: Comprehensive Medicinal Chemistry III, Andy Davies and Colin Edge, Eds, Elsevier, in press, due July 2017

- **When will systems pharmacology impact upon drug development? A study on the cholesterol biosynthesis pathway**. Helen Benson, Steven Watterson, Joanna Sharman, Chido Mpamhanga, Andrew Parton, Christopher Southan, Peter Ghazal, Anthony Harmar, British Journal of Pharmacology, submitted.
- **SynPharm, a database of drug-responsive protein sequences**. Sam Ireland, Simon Harding, Joanna Sharman, Christopher Southan and Jamie Davies, in preparation
- **Virtual versus reality: An analysis of deuterated drugs** (Southan), in preparation
- **Will the real drugs please stand up**? Comparing approved structures in PubChem (Southan, et al.), in preparation.
- **Advances in proteases and α/β hydrolases inhibition for human disease** (Turner, et al.) in preparation

SOCIAL MEDIA

We use mainstream social media outlets for five primary purposes 1) outreach to potential new users and/or followers 2) informing on new features or releases 3) enhancing awareness of our publications and presentations 4) keeping collaborators and other followers (including many major databases) aware of our activities. 5) establishing reciprocity with our key followers and collaborators (n.b. our strategic exploitation of Social Media and Scholarly Portals overlap in practice but the latter has its own section below).

FACEBOOK

The number of 'likes' increased to 3204 from 3157 in October 2016.

TWITTER

@GuidetoPHARM has just pipped 1,500 tweets, our followers have increased to 1127 from 945 in October 2016 and our re-tweet rate is gradually increasing. This medium helps with rapid technical interchanges with teams from other resources. It is also an increasingly useful alerting system for our blog posts, key papers, including from BJP, other pharmacology journals, immunology, biochemistry and medicinal chemistry, weekly PDB structures etc. Consequently, most of our Hot Topics are now first picked up from Twitter. We also engage in a discrete level of diplomatic re-tweeting for reciprocal reach extensions (*e.g.* with @BritPharmSoc , @BrJPharmacol, @PharmRevJournal , @cdsouthan and @mqzspa (NC-IUPHAR chair).

LINKEDIN

The Curation Team have been encouraging Subcommittee Chairs and collaborators to increase their reciprocal connectivity as individual LinkedIN users, thereby gradual expanding our collective inter-network outreach for posting updates, new papers ect. (n.b. interested readers of this report are encouraged make connection requests from GtoPdb and IUPHAR scientists they know). Our own LinkedIN group page now has 143 followers, up 20 from Oct 2016. We are also now reciprocally cross-pointing to the IUPHAR LinkedIN page and the IUPHAR/ASPET Pharmacology Education Project page.

BLOGGING

Our Edinburgh blog (http://blog.guidetopharmacology.org/) is receiving over 300 views on average per month, which has increased because we are now posting more content. This has now become our primary news feed and includes database release updates, new features, technical items or articles (we also now have our first guest post). We also post all Hot Topics that have comments and announcements of IUPHAR reviews. This replaces our old RSS feed which we no longer maintain. Team member Chris Southan maintains his own (http://cdsouthan.blogspot.com/) where relevant posts include cross-pointers to GtoPdb.

HOT TOPICS

As an established and popular feature our Hot Topics are seeded in the form of new significant pharmacology, drug discovery and key human genomics papers. These are communicated to us from Subcommittee members but increasingly picked up from Twitter. We have experienced such an increase in 2017 that we have moved them to their own website page. For a selection, as before, we commission concise commentaries from our expert contacts. When these come in they are moved across to the blog. Where pertinent, we add new links to GtoPdb entities and in some cases updates to these (*e.g.* when ligands we

already have appear in new GPCR or ion channel PDB structures). Guest posts have also been introduced (*e.g.* from Jörg Striessnig and David Gloriam).

SLIDESHARE

Slideshare (http://www.slideshare.net/GuidetoPHARM) allows the database team to share slide sets and posters with the community thereby extending the reach way beyond conference session direct attendees. Our slides are proving to be popular with over 4,703 views in a year. We have also added a set of generic slides which can be used by anyone presenting or teaching on GtoPdb and a generic poster which can be printed out in various sizes and taken to meetings or handed out as flyers.

PUBMED COMMONS

We continue our exploitation by adding comments on relevant PubMed IDs. We restrict these to judicious cross-pointers either between our publications and/or to draw attention to particular database sets. Examples include pointing forwards from the 2013 NAR and 2014 to the 2016 NAR, adding an open access link for our GPCR Database article and coupling between a BACE1 inhibitor ligand featured in an F1000 recommendation and a blog post.

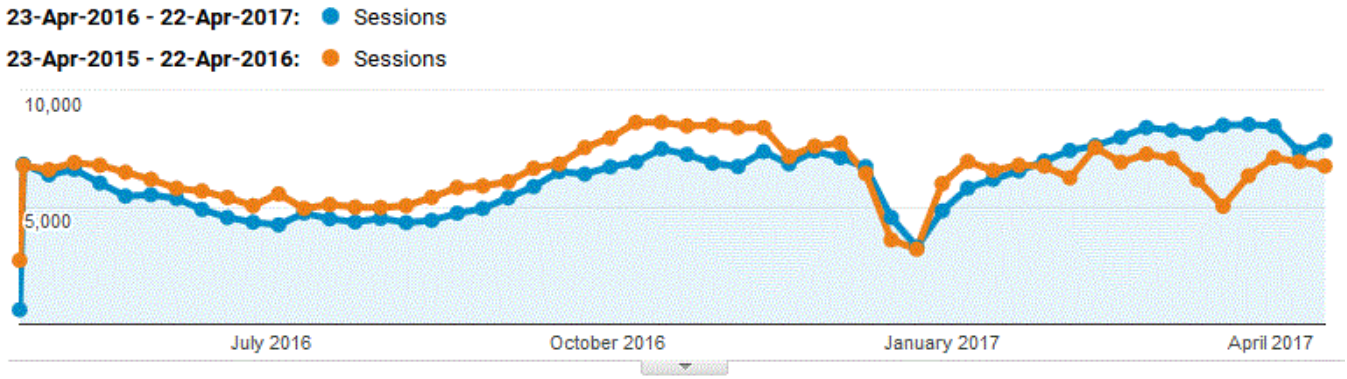
ENQUIRIES RECEIVED FROM USERS

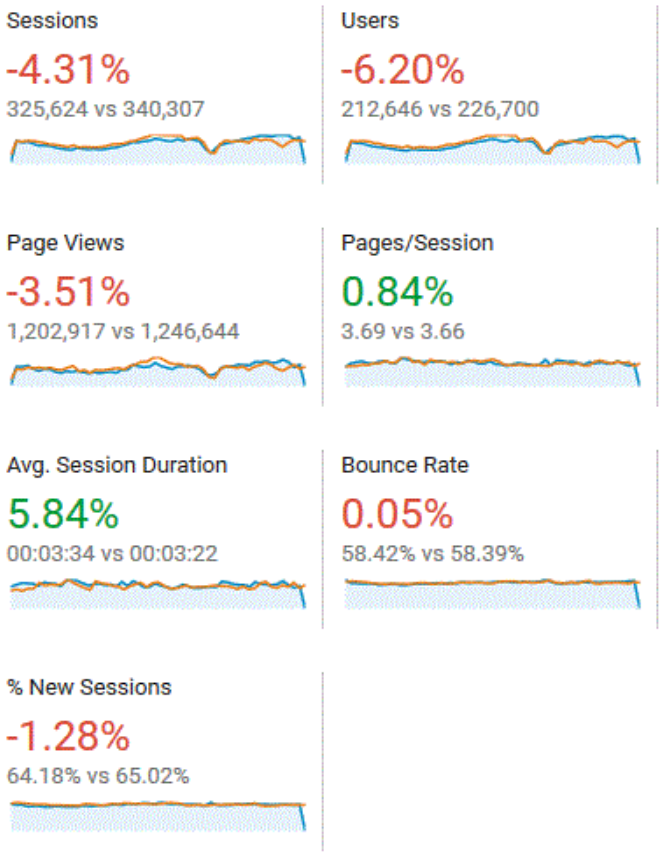
During 2016 we had noticed more user communications coming in to enquiries@guidetopharmacology.org, This upswing has continued into 2017, currently up to about one a week, include our first one direct from a Dundee University Twitter follower. These pose a variety of questions, some of which are quite challenging. We have also had database errors pointed out (at a low frequency we should add), covering a spectrum from wrong names to disputed mechanism of action to complex stereochemistry issues. Fielding this feedback has had a number of positive consequences, including the included compliments and alerts to additional papers. We have also highlighted selected corrections in release notes. Note also from correcting these individual instances we can sometimes execute an internal consistency check to pick up other errors of the same type.

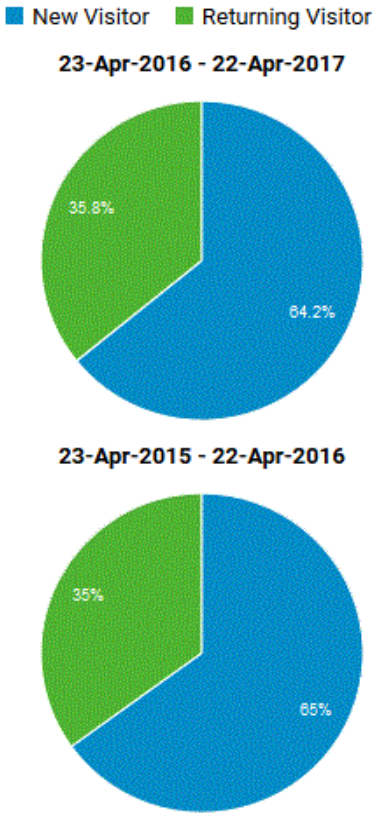
ENGAGING WITH US

As is implicit from the Social Media section above, it is crucial to extend our external "presence". Thus, the more readers of this document who "connect" with us, (*via* whichever of the channels above they use for their own professional profile) the more our outreach extends. This also has reciprocal mutual advantages. In particular re-tweets and LinkedIN likes are useful for extending the alerting network for new releases, new publications, meeting slide sets and blog posts. Note also that each time you either save one of our publications to your own Mendeley or CiteUlike accounts or mention it in a tweet, blog or PubMed commons comment (but make sure you specify a DOI or PubMed link for the auto-indexing) the Altmetrics score (see below in Portals) notches up for that paper (n.b. we are only advocating professionally considered and low-key social media engagement).

GtoPdb WEB SITE ACCESS STATISTICS







Graphs comparing visitors to guidetopharmacology.org for the 12 months from April 2016 to April 2017, with the previous 12 months.

Monthly statistics	*April 2016-April 2017 (previous 12 months)*
Sessions	27,135 (28,358)
Users	17,205 (18,891)
Page views	100,243 (103,887)
Pages / Session	3.69 (3.66)
Avg. Session Duration	00:03:34 (00:03:22)

GtoPdb CONTENT

These stats were compiled for the database on 14/10/16, on the day of the October 2016 public release. All database stats can be found at http://www.guidetopharmacology.org/about.jsp#content.

Targets	*Number of UniProt IDs*
7TM receptors	395
Nuclear hormone receptors	48
Catalytic receptors	243
Ligand-gated ion channels	81
Voltage-gated ion channels	144
Other ion channels	47
Enzymes	1181
Transporters	509
Other protein targets	160
Targets with ligand interactions	1660
Targets with quantitative ligand interactions	1408
Targets with approved drug interactions	596
Primary targets with approved drug interactions	312
Total number of targets	**2808**
Ligands	Number of ligands
Synthetic organics	5729
Metabolites	584
Endogenous peptides	772
Other peptides including synthetic peptides	1291
Natural products	246
Antibodies	212
Inorganics	38
Approved drugs	1322
Withdrawn drugs	67
Ligands with INNs	2099
Labelled ligands	607
PubChem CIDs (SIDs)	6813 (8831)
Ligands with target interactions	7576
Ligands with quantitative interactions (approved drugs)	6630 (813)
Ligands with clinical use summaries (approved drugs)	2046 (1320)
Total number of ligands	**8872**
Number of binding constants	45906
Number of binding constants curated from the literature	15080

DOWNLOAD STATISTICS

Yearly period 21st April Year 1 to 20th April Year 2.

GOOGLE ANALYTICS COMPARISON OF DOWNLOADS

Event Category: Downloads
Event Label: Downloaded

	Count
2015-2016	2,477
2016-2017	2,481
Change	+0.16%

This corresponds to files downloaded from our main downloads page:
http://www.guidetopharmacology.org/download.jsp
and the slides page: http://www.guidetopharmacology.org/slides.jsp

A more specific breakdown is shown here:

	2015-2016	2016-2017	Change
Targets CSV file	877	976	10%
Interactions CSV file	255	273	6.6%
Ligands CSV file	232	230	-0.87%
UniProt Mapping file	121	166	27%
HGNC mapping file	72	79	8.9%
Peptides CSV file	86	86	0%
PostgreSQL*	146	169	16%
Generic slides (PPT & PDF)	280	282	25.33%
Generic poster	132	106	-25%

* Total downloads of PostgreSQL database dump files (versions 2015.1-2017.2)

OTHER FILES

Our website tutorial has been downloaded 401 times in the last year. The 'Terms and Symbols' PDF has been downloaded 293 times in the last year (~20% decrease on previous year).

GtoPdb INTERACTIONS WITH OTHER RESOURCES

For more details of the teams we collaborate with please see April 2016 and October 2016 reports. Only significant changes since October 2016 are reported below.

GPCRDB

With David Gloriam and his team at the University of Copenhagen we intend to resolve our human GPCR numbers in 3Q2017 as we have small discrepancies.

ELIXIR

Engagements continue with this important Europe-wide initiative. As reported have an entry in the ELIXIR tools and resources directory as one of the official UK ELIXIR Node Resources and part of the Excelerate initiative. This gives us local Edinburgh connections to Prof. Chris Ponting as Head of the UK ELIXIR Node and Prof. Richard Baldock the PI for the EMAP mouse expression database. We attend weekly UK node teleconferences and presented our resource at one of these in Dec 2016. In Nov 2016 we completed an application to be promoted to a core European Node and hope to hear the outcome in 2Q2017.

OPENPHACTS

We have set up collaboration with Alasdair Gray, a Linked data and RDF expert from Heriot-Watt University. His undergraduate student, Liam Bruce, has been working with us on a project to produce an RDF version of the GtoPdb data to facilitate eventual loading into the OpenPHACTS API. Other important

direct consumers of our content, including pharmaceutical companies, increasingly prefer RDF format for data uptake. The final results of this project will be available soon.

PUBCHEM

We continue to interact with PubChem including Skype calls with Evan Bolton, mainly around enhancing visibility and utility of our ligand entries (see ligand drug link section below). The PubChem team not only gave us a citation but also a detailed and important mention in their paper "Literature information in PubChem: associations between PubChem records and scientific articles" (PMID 27293485).

BINDINGDB

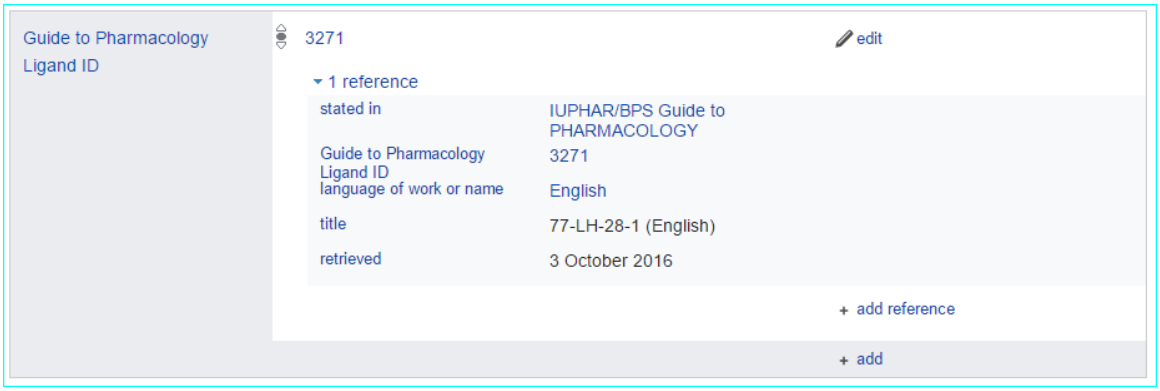
We continue to interact with Michael Gilson the PI for BindingDB where Christopher Southan is on their SAB. We are trying out a form of patent exploitation where we will sent USPTO document nos for recent ligands curated for GtoImmuPdb where extensive analogue SAR in the patent that exceeds that in the paper.

STRUCTURAL GENOMICS CONSORTIUM PROBE PORTAL

Christopher Southan is on the SAB for the SGC/Wellcome Chemical Probes Portal for which he has to comment on assigned probe reports. At their launch in July 2016 we established we had ligand entries for ~ 90% of their first set of 84 structures. As proof of concept we have added number of outlinks (*e.g.* SGX523 is our Ligand 5709 and their SGX-523). Their news list includes interesting initiatives including arranging compound availability from Tocris and requests for new GPCR probes. We have been in contact with Amy Donner the Portal Director, to explore interactions going forward, including reciprocal connectivity options. They have now expanded to 147 probes with PubChem CIDs and we have 110 of these.

WIKIDATA

As another broadening of our reach, we have consolidated contacts with Wikidata team members to subsume our content. Work is continuing with Sebastian Burgstaller from Scripps along with exploitation support from Egon Willighagen of Maastricht University. We now have 5914 links, for example Q4643539 shows our cross-reference below.



This new integration step will bring us into the wonderful world of SPARQL connectivity for our data.

JOURNAL-TO-DATABASE CONNECTIVITY

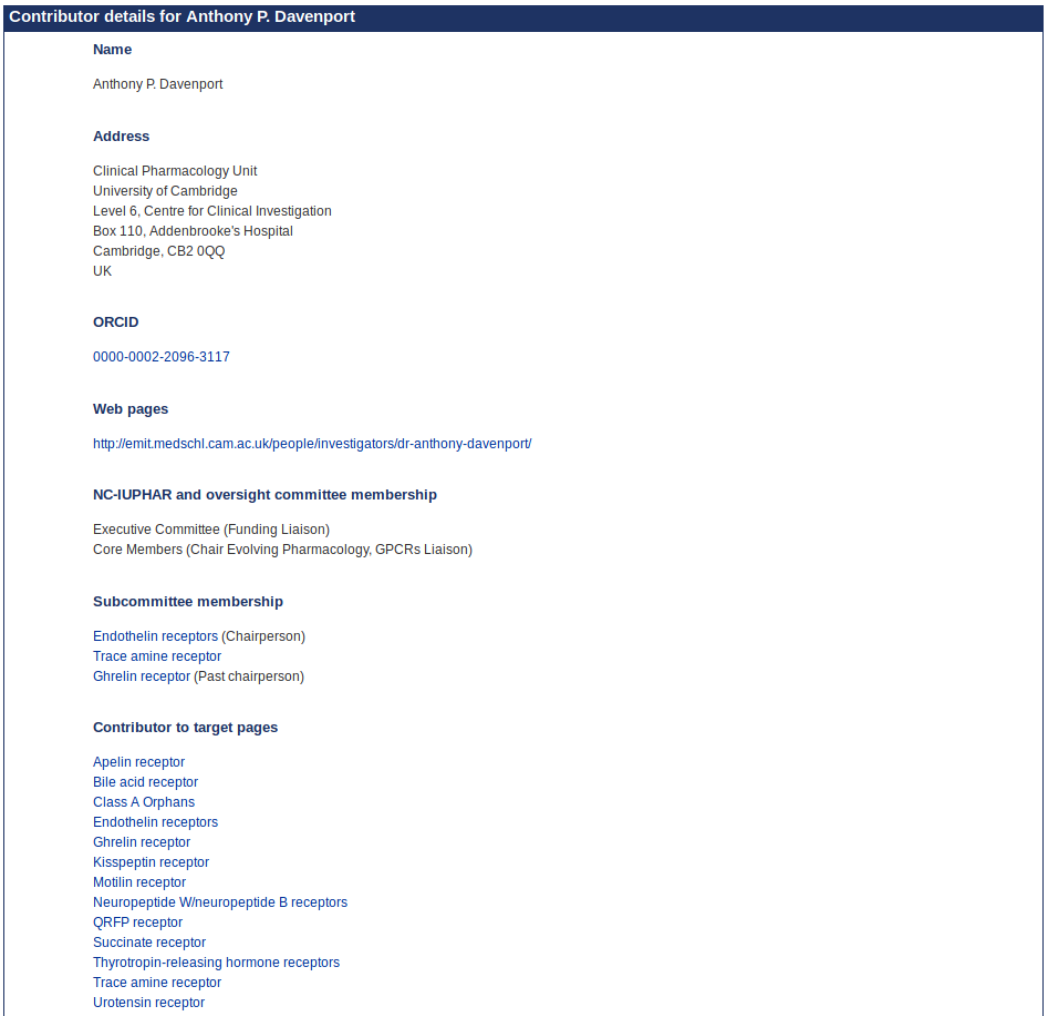
Engagements in this important area continue. These include that targets and ligands specified in The Concise Guide to PHARMACOLOGY are hyperlinked directly to the database records. The marking-up tables of links (ToLs) directly to GtoPdb records continues for both regular papers as well as reviews in BJP and has recently been expanded for BJCP . IUPHAR Review 20 is one of the largest examples where the ToLs link to 10 target and 46 ligands. We have consolidated the provision of through-links from ligands with referenced BJP articles through to PubChem and PubMed. An example is C108297, ligand 9180, a selective glucocorticoid receptor modulator that attenuates inflammation. Since it featured in BJP as PMID 26990179 this was a secondary reference (*i.e.* in addition to the first binding data report). The consequence of this is that, *via* submission of our ligand entry into PubChem CID 25110774, this entry now points to the BJP paper as well as the primary citation (and backwards as PubMed > PubChem). An analogous virtuous connectivity example for BJCP is exemplified by reciprocal links between PMID 27730665 our ligand ID

for ACT-389949 and PubChem CID 49834265 (n.b. both C108297 and ACT-389949 are now in GtoImmuPdb). These surfacing loops benefit us, PubChem, the journal and the author (a 4x win). BJP is now transitioning to direct inline GtoPdb links embedded in text rather than ToLs to provide smoother reader navigation. Two reviews have appeared where this has been piloted, PMID 28299772 and PMID 28369768. Going forward, we will explore options whereby we can arrange to get relevant new BJP ligands into GtoPdb/GtoImmuPdb (and new BJP references into PubChem) in time for the inlink publishing to ensure complete capture.

NEW GTOPDB WEBSITE FEATURES (SINCE OCTOBER 2016)

NEW CONTRIBUTOR FACULTY PAGES

We have introduced new contributor faculty pages which are accessible by clicking on names in the full contributor list and from individual target page contributor lists. Every database contributor now has an individual page (see screenshot) which includes their address, the target pages they contribute to and subcommittee membership, and optional additional details including ORCIDs and external home page and profile links (we can accommodate any external profile link such as LinkedIn). For NC-IUPHAR members, profile pages and the NC-IUPHAR membership list provide summaries of their research interests. Our administrator, Toni Wigglesworth, sent out >600 requests to all the contributors and NC-IUPHAR members we have contact details for asking them to provide the additional information. As of April 2017 she has had over 100 responses back but many still need to be chased, and there were also a significant number of 'bounce backs' indicating our records are not up-to-date (it may be difficult to find the new email addresses). There are some members of the NC-IUPHAR committee and oversight committees who have also not responded, we'd be grateful if you could please contact Toni so we can complete the NC-IUPHAR faculty pages.



Contributor details for Anthony P. Davenport

Name

Anthony P. Davenport

Address

Clinical Pharmacology Unit
University of Cambridge
Level 6, Centre for Clinical Investigation
Box 110, Addenbrooke's Hospital
Cambridge, CB2 0QQ
UK

ORCID

0000-0002-2096-3117

Web pages

http://emit.medschl.cam.ac.uk/people/investigators/dr-anthony-davenport/

NC-IUPHAR and oversight committee membership

Executive Committee (Funding Liaison)
Core Members (Chair Evolving Pharmacology, GPCRs Liaison)

Subcommittee membership

Endothelin receptors (Chairperson)
Trace amine receptor
Ghrelin receptor (Past chairperson)

Contributor to target pages

Apelin receptor
Bile acid receptor
Class A Orphans
Endothelin receptors
Ghrelin receptor
Kisspeptin receptor
Motilin receptor
Neuropeptide W/neuropeptide B receptors
QRFP receptor
Succinate receptor
Thyrotropin-releasing hormone receptors
Trace amine receptor
Urotensin receptor

Example of an NC-IUPHAR faculty profile page

WEB SERVICES

The REST web services have been updated and now include interactions web services providing lists of target-ligand pairs which can be filtered by target/ligand type and properties, binding affinity etc., and references web services which can retrieve references by id or the full interaction reference set. Improvements were also made to the speed of loading, especially for long lists of data. We have also introduced Google Analytics tracking, but since this only went live a few weeks ago we will report on the access statistics at the next meeting. In the meantime we have had some good feedback from web service users and many of the changes were done in response to requests from users.

DATABASE LINKS REVIEW AND ECOSYSTEM AWARENESS

When IUPHAR-DB first launched the only comparable databases in the ligand small-molecule arena were ChEBI, DrugBank. PubChem, ChemSpider and BindingDB, although many protein and genomics databases were already well established. Even in the time elapsed since our last report the "ecosystem" of relevant resources in which we operate become significantly more crowded. Consequently, we are increasingly impinged w.r.t. content, coverage, externally perceived utility and actual utility (as defined by our real-world usage reported here). This has important implications (including for funding) that cannot be expanded on here. What we can outline briefly is our engagement in link review and awareness of newer intersecting resources. Addressing the first of these, two crucial database features are what we link out to (OLs) and those who link in to us (ILs) and where these reciprocally intersect (*i.e.* ecosystem connectivity).

Our assessment of ILs was reported in Mar 2016. The extent was positively surprising, with over 15 major resources pointing explicitly to us. However, since we are open, new resources generally do not inform us directly about new ILs (although they may cite us) so it's difficult to keep up. In 1Q17 we initiated reviews of our OLs on both the target and ligand side, with a primary objective to simplify (since databases constitutively suffer from over-linking, much of which is circular) but also considering possible value-added additions. The process was informative (*e.g.* we found the HPRD protein database had not been updated since 2010) and a number of other OLs were removed, including those easily accessible as cross references in other databases we maintained links to. In terms of additions, we have added new links from transporter pages to the Bioparadigms SLC Tables database. This site aggregates lots of information relevant to the Solute Carrier superfamily. We look forward to collaborating with the developers of the SLC tables in future, as their site grows. A representative example is family 165. Other sources are being technically assessed for inclusion. The first of these is the Human Protein Atlas (as a comprehensive cell/tissue expression resource *via* Abs and transcript profiling) where we have established collaborative contact with Prof Uhlen at the recent ITMAT meeting. The second is Exome Aggregation Consortium (ExAC and gnomAD) as a deep genomic and protein variant resources. The third, particularly relevant to GtoImmuPdb, is the Cellosaurus knowledge resource on cell lines. Our initial statistics collected for ligand OLs presented a complex picture we are still resolving. We will consult the DRUTACS committee on this but we are assessing UniChem as a possible substitute for multiple EBI chemistry databases One outcome was that we have now automated the addition of PDB ligand links, since, *via* manual curation, we were unable to keep up either with a backlog of 100s of missing connections or the increase of co-crystallized ligand structures, including recent ones from GPCRs and ion channels.

For many reasons we need to keep abreast of our ecosystem and, crucially, engage to influence our positioning in it. We maintain personal contacts and publication awareness where we can (and travel budget permitting) but it's difficult to keep up. One way is to check who is citing our 2014 and 2016 NAR papers. Because of the reference citation issue (see bibliometrics section) we need to filter out the BJP and BJCP cites. This still leaves us with an impressive list of 53 papers citing us in the last three years, many of which are new resources and have added ILs to us. The list, in reverse date order, includes ELIXIR , mutLBSgeneDB, Pharos, ChEMBL, PubChem BioAssay, DrugCentral, OpenPhacts, Wikidata and MEROPS. We have contacts in most of these teams and for the most recent (April 2017) The Drug Repurposing Hub we are specified as a source and have consequently contacted with the PI in regard to updates.

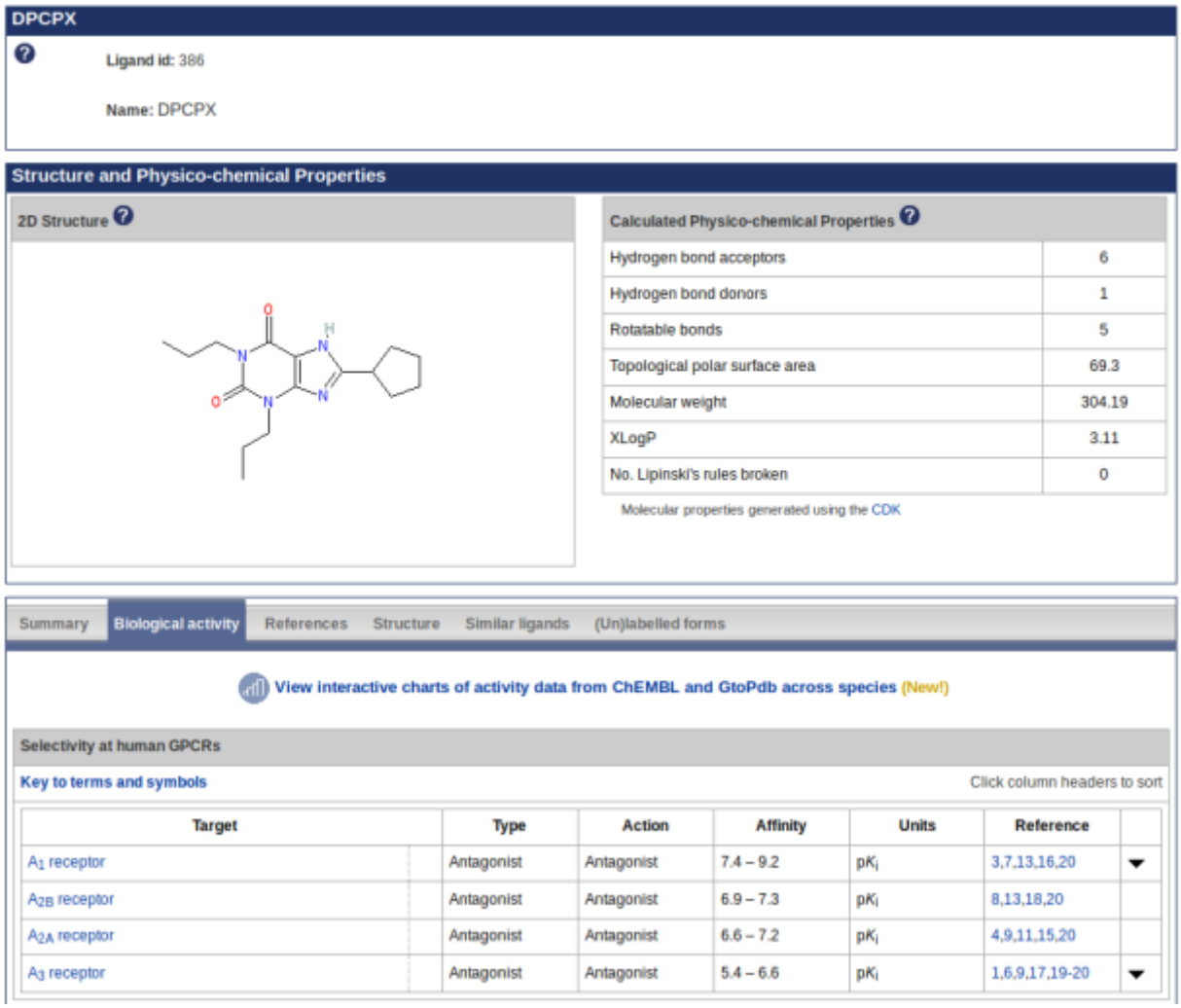
GRAPHS COMPARING LIGAND ACTIVITY DATA ACROSS SPECIES

We have developed new ligand activity graphs comparing activity ranges across species using data extracted from GtoPdb and ChEMBL. The graphs are available *via* the ligand page biological activity tab (see first screenshot), but note this feature is currently ***only for ligands that we have links to in ChEMBL***. Unfortunately this excludes most large peptides, due to the difficulty in identifying equivalent database entries for these complex structures.

The charts are presented as box plots showing the median, interquartile range, low and high data points. Mouse-over a plot to see the data range, with one chart per target and a different colour used for each species (see second screenshot). There is also a separate plot for each activity type (pKi, pIC50 etc), although we may consider combining these in future. The data extracted from ChEMBL (version 22) have been standardised where possible to one of the 5 main IUPHAR types (pKd, pKi, pIC50, pEC50, and a few cases of pA2). Only binding ('B') and functional ('F') assays are included. ChEMBL and GtoPdb targets were compared by name and UniProt id. For further details about the selection criteria see the GtoPdb help page. The third screenshot below shows the ligand palosuran and visualises how the data differ significantly between human and rat UT receptors.

In addition a data table is provided below the charts showing the source databases, the assay details, recorded values and references, with PubMed ids (the fourth screenshot).

Full details are available in our blog post: https://blog.guidetopharmacology.org/2017/03/23/database-release-2017-2/



Example of a ligand page showing link to the interactive activity charts from GtoPdb and ChEMBL.

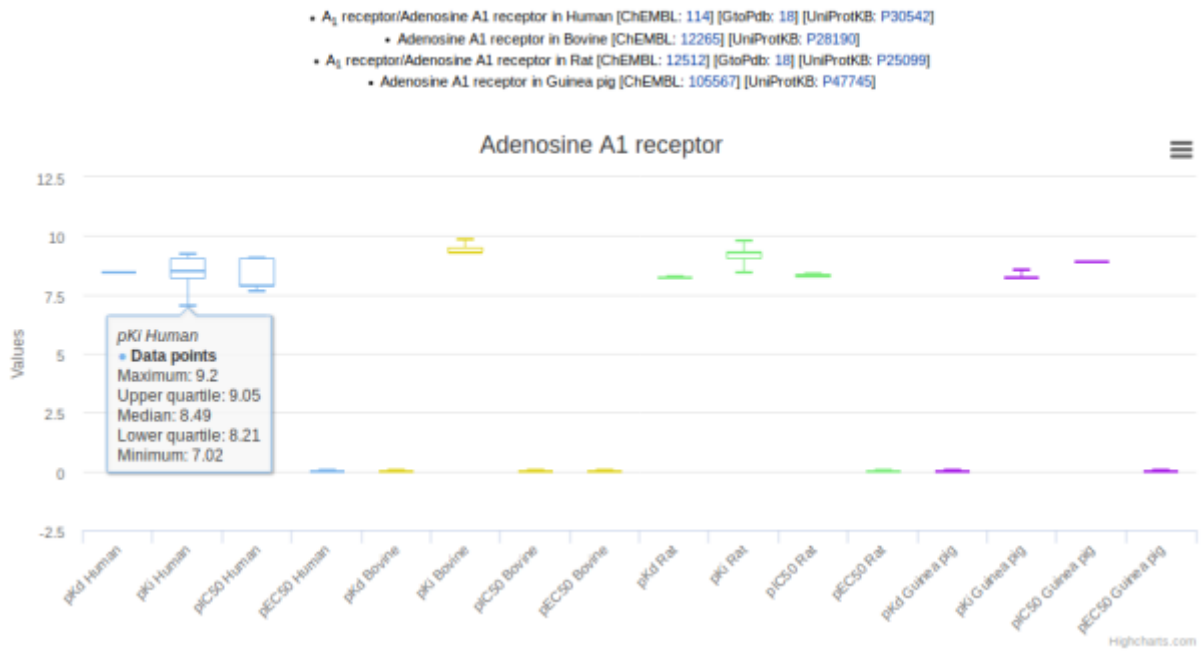


Chart showing DPCPX ligand activity data from ChEMBL and GtoPdb across 4 species. Mouse-over a plot to see the median, lower and upper quartiles, and minimum and maximum data points for each activity type.

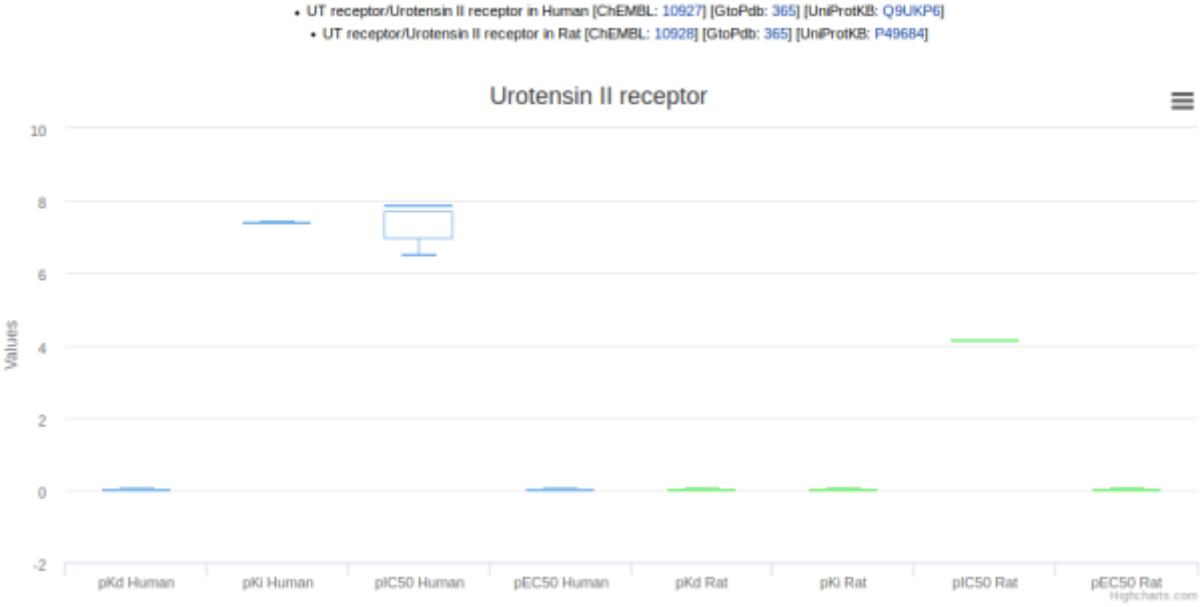


Chart showing palosuran activity at human and rat UT receptors.

DB	Assay description	Assay Type	Standard value	Standard activity	Original value	Original units	Original activity	Reference
UT receptor/Urotensin II receptor in Human (target type: SINGLE PROTEIN) [ChEMBL: 10927] [GtoPdb: 365] [UniProtKB: Q9UKP6]								
ChEMBL	Displacement of [125I]urotensin 2 from urotensin 2 receptor in human RMS13 cells by scintillation proximity assay	B	7.39	pKi	41	nM	Ki	J. Med. Chem. (2009) 52: 7432-7445 [PMID:19731961]
ChEMBL	Antagonist activity at urotensin 2 receptor in human RMS13 cells assessed as inhibition of urotensin 2-induced intracellular calcium mobilization after 1 hr by FLIPR assay	F	6.49	pIC50	320	nM	IC50	J. Med. Chem. (2009) 52: 7432-7445 [PMID:19731961]
ChEMBL	Antagonist activity at recombinant human urotensin2 receptor expressed in CHO cells assessed as inhibition of urotensin2-stimulated Ca2+ mobilization incubated 10 mins prior to urotensin2 stimulation measured after 10 mins by fluorometric analysis	B	7.41	pIC50	39	nM	IC50	Bioorg. Med. Chem. Lett. (2013) 23: 2177-2180 [PMID:23453841]
ChEMBL	Displacement of [125I]-U2 from recombinant human urotensin2 receptor expressed in human U2OS cells after 30 mins by gamma counting analysis	B	8.3	pIC50	<5	nM	IC50	Bioorg. Med. Chem. Lett. (2013) 23: 2177-2180 [PMID:23453841]
GtoPdb	-	-	7.1	pIC50	-	-	IC50	J. Pharmacol. Exp. Ther. (2004) 311: 204-12 [PMID:15146030]
UT receptor/Urotensin II receptor in Rat (target type: SINGLE PROTEIN) [ChEMBL: 10928] [GtoPdb: 365] [UniProtKB: P49684]								
ChEMBL	Antagonist activity at rat urotensin 2 receptor expressed in CHOK1 cells assessed as inhibition of urotensin 2-induced intracellular calcium mobilization after 1 hr by FLIPR assay	F	4.12	pIC50	>75000	nM	IC50	J. Med. Chem. (2009) 52: 7432-7445 [PMID:19731961]

Table of activity data from ChEMBL and GtoPdb.

OTHER SUGGESTED NEW FEATURES IN OCTOBER 2016

Other suggestions for new website features discussed at the October Paris meeting were:

- **Enhancing mobile accessibility of site:** This is potentially a large amount of work which would need careful planning and design work in order to turn the existing site into an optimised mobile and tablet site. Therefore, significant time would need to be scheduled for this. An option is to pay a professional mobile site design company or contractor.
- **Converting to HTTPS:** Using HTTPS (secure connection) on websites is becomingly increasing important (browsers and search engines are starting to warn users when they access an insecure site). JS has been discussing with UofE Information Services about installing HTTPS on the server, and we are hoping the work can be progressed this summer, once budgets have been agreed.

GtoPdb ENTITY GROWTH

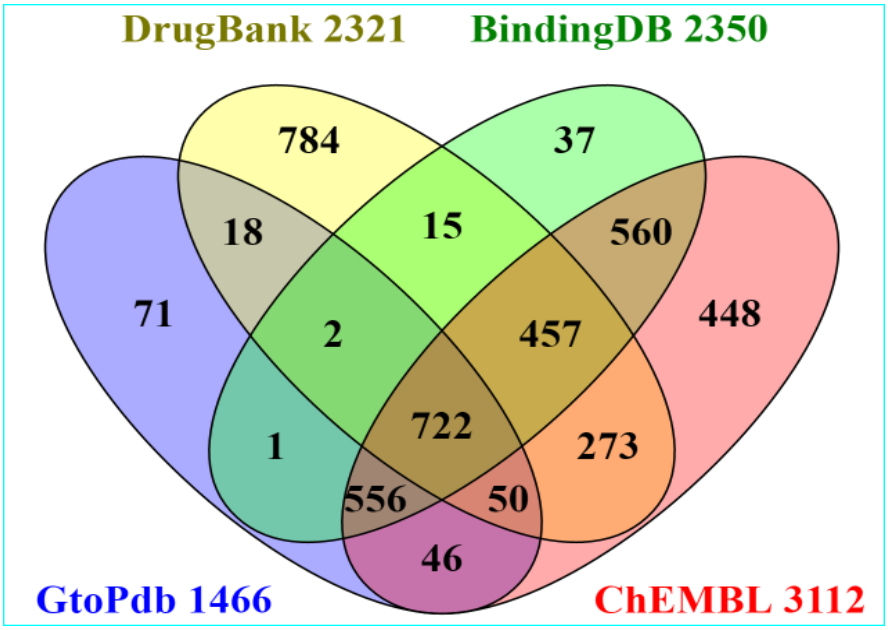
Growth rates over the span of the previous Wellcome Trust grant are documented in earlier reports and our 2016 NAR paper. Notwithstanding, it is important to note that the staff changes associated with the new Wellcome grant resulted in the loss of one curatorial FTEs from the team as a whole, plus the effective transfer of two FTEs to GtoImmuPdb. We consequently cannot sustain the previous overall growth rate (*i.e.* 2013-15 below). While the committees have submitted Concise Guide updates, most new entities are being added *via* the population of GtoImmuPdb. However, significant curation effort goes towards tagging pre-existing targets and ligands with GtoImmuPdb relevant comments and new references.

	Oct 2013	*Oct 2015*	*April 2016*	*Oct 2016*	*Apr 2017*
Target protein IDs	2485	2761	2775	2794	2808
Ligands total	6064	8024	8400	8674	8872
Approved drugs	559	1233	1273	1291	1322
Antibodies	10	138	172	205	212
Peptides	1776	1981	2007	2039	2063
Synthetic small molecules	3504	5055	5363	5563	5729
PubChem SIDs	3107	8024	8328	8674	8831
PubChem CIDs	2694	6057	6163	6337	6813
Binding constants	41076	44691	45534	45908	46287
References	21774	27880	29247	30251	31239

GtoPdb TARGET UPDATES (SINCE OCTOBER 2016)

- The major part of the work to update the target family summary pages has been completed in advance of producing the Concise Guide to PHARMACOLOGY 2017/18 from the database (see below), which is due out in September this year. Since most of our curation effort has gone into these updates, relatively fewer detailed page updates have been made in the period since the October 2016 meeting.
- GPCR updates: Gonadotropin-releasing hormone receptors, Glucagon receptors, Melanin-concentrating hormone receptors, Neuromedin U receptors, Orexin receptors, Prostanoid receptors, Relaxin family peptide receptors.
- Ion channel updates: Major revisions for the Detailed introductions have been provided for Calcium-activated potassium channels Cyclic nucleotide-regulated channels, Inwardly rectifying potassium channels, Two P domain potassium channels, Voltage-gated potassium channels, Transient Receptor Potential channels, Voltage-gated calcium channels.
- Catalytic receptor: Reorganisation and updating of the Pattern recognition receptors.
- Enzyme updates: Enzymes involved in Adenosine turnover, reorganisation and updating of the enzymes involved in Endocannabinoid turnover, the Guanylyl cyclases and Gasotransmitters (ongoing nomenclature issues).
- Other protein targets updates: RGS proteins
- **New targets:**
    - **CD molecules:** CD6, CD300a, LAG3 (CD223)
    - **2-Acylglycerol ester turnover:** αβ-Hydrolase 6
    - **RIG-I-like receptor family:** DExD/H-box helicase 58, interferon induced with helicase C domain 1, DExH-box helicase 58
    - **Absent in melanoma (AIM)-like receptors (ALRs):** absent in melanoma 2, interferon gamma inducible protein 16, pyrin and HIN domain family member 1, myeloid cell nuclear differentiation antigen
    - **C-type lectin-like receptors (CLRs):** C-type lectin domain family 7 member A, C-type lectin domain family 6 member A, C-type lectin domain family 4 member E, CD209 molecule, C-type lectin domain family 4 member A
- Relative target growth and coverage. This can be assessed by comparing our own UniProt cross-references (for targets with quantitative interactions) against the other major chemogenomic resources

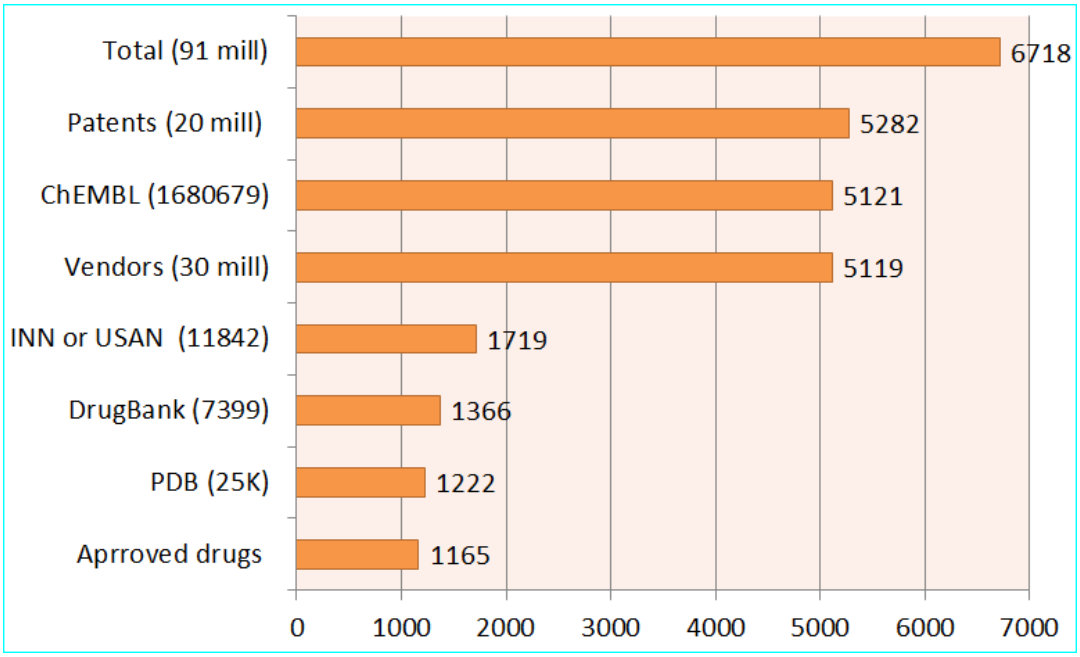
with UniProt cross-references, DrugBank, BindingDB and ChEMBL.



These intercepts and differences in the above figure are informative, but note that DrugBank includes targets based on literature co-occurrence rather than data-supported mechanism of action. Our own cross-reference are based on quantitative interactions and indicate we have 71 targets not in the other three databases. ChEMBL has expanded human target coverage in their new release 22.

GtoPdb LIGAND AND PUBCHEM STATS

New curated ligand interactions have been added as a result of both target updates from subcommittees and curators' efforts on new targets and ligands. As a consequence our total number of curated interactions now stands at 14327 (*i.e.* number of curated binding constants). Note that most new ligands have been added *via* GtoImmuPdb (see below). These add to our current substances (SID) that we submit to PubChem (refreshing previous submissions) giving us 8845 ligand entries from release 2017.2 . Those that have defined chemical structures are merged into 6718 Compound Identifiers, CIDs. The new feature allowing the selection of just approved drugs was introduced in the last report. The select is "IUPHAR/BPS Guide to PHARMACOLOGY"[SourceName] AND approved [Comment] which can be pasted into the Substance query box to retrieve 1312 SIDs. Of these 1165 have CIDs (use the "Find Related Data" operator and select "same CIDs". The chart below indicates selected PubChem overlap statistics.



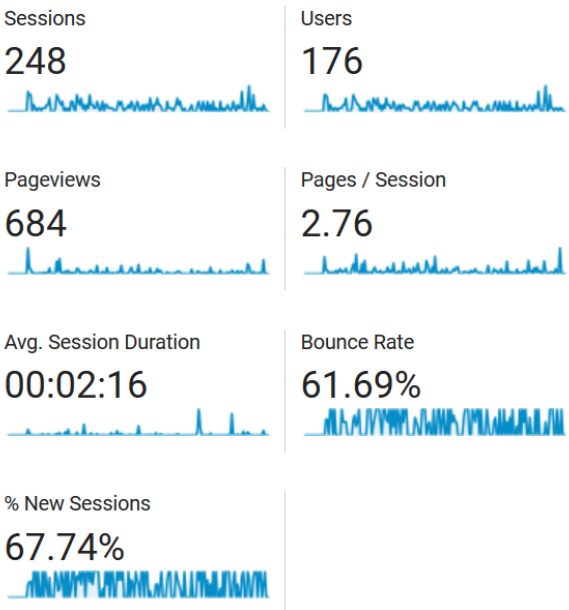
The high patent overlap allows users to drill-back *via* SureChEMBL (see PMID 26194581) in many cases to first-filings (from pharma companies or academics) that are likely to have extended SAR data (n.b. we include references to such useful patents for more recent ligands but coverage is low). We monitor ChEMBL intersects to compare with our curated numbers. The vendor availability of compounds is obviously key information for experimentalists and indeed remains an explicit GtoPdb objective. The next three are figures we need to keep an eye on (*i.e.* we have structures for 15% of nonproprietary drug names, a low overlap with DrugBank and we monitor the steady expansion of PDB coverage). In terms of utility and competitive positioning, what we do not overlap with is more important than structures-in-common. For example we have 1597 structures that ChEMBL does not have, 5352 not in DrugBank and 5458 not in DrugCentral. In addition we have 325 structures unique to us. Figures such as these four, emphasising the complementarity of our content to related databases, ensures that our integration by other resources (academic or commercial) becomes imperative rather than optional.

PRODUCING THE CONCISE GUIDE TO PHARMACOLOGY 2017/18

The major part of the work to update the target family summary pages has been completed in advance of producing the Concise Guide to PHARMACOLOGY 2017/18 from the database, which is due out in September this year. For the next version, we have been working towards trying to make the information more concise, and limiting both ligands and further reading to the 5 most useful in many cases. Obviously there are some targets where it makes sense to have more or less than 5 displayed on the summary page, but in any case, all the ligands can still be viewed on the detailed target page, and the website contains more further reading references than are included in the published Concise Guide. We are very grateful to all the contributors and the editors who have provided information.

SynPHARM: A NEW DATABASE OF SMALL MOLECULES AND THEIR DRUG-RESPONSIVE PROTEIN SEQUENCES LINKED TO GtoPdb

For a detailed description of SynPHARM please see the October 2016 report or the website: http://synpharm.guidetopharmacology.org/. It is a database of drug-responsive protein sequences derived from GtoPdb interaction data. A paper describing SynPHARM is being drafted.



SynPharm access statistics for 6 months to April 2017

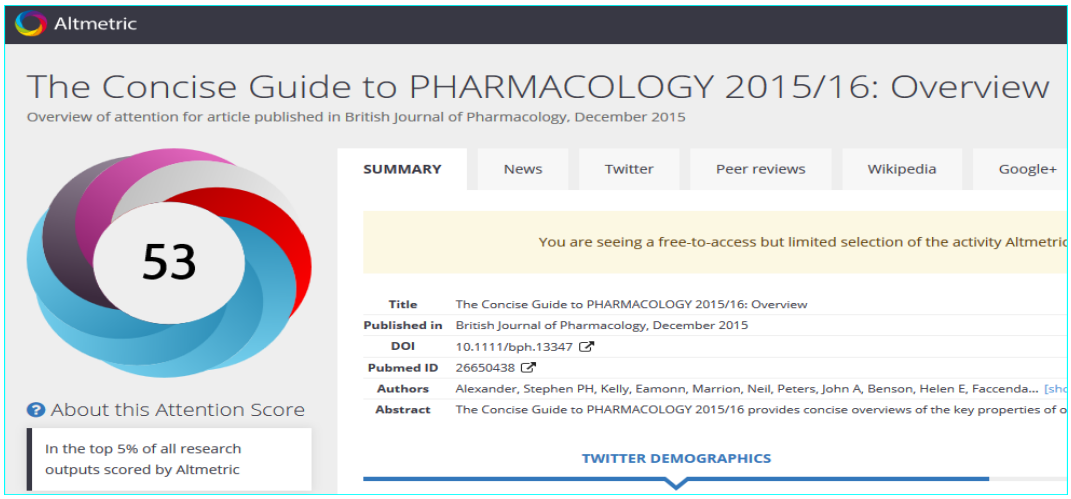
BIBLIOMETRICS AND SCHOLARLY PORTALS

As outlined in previously we now track various impact metrics for the GtoPdb team and NC-IUPHAR affiliated papers in PubMed, PubMed Central , European Pub Med Central (EPMC). In addition to being our mandatory source for funder linking and (their own) assessments EPMC has introduced a number of new features relevant to impact assessment. One of these was Altmetric badges which is now joined by in-line

citation graphs, the inclusion of the WOS citation figures (as well as the EPMC ones) and pointers to Kudos entries. These feature are available for the "external links" for each publication. (n.b. we have historically included PMID links in this document for citation counts so we now have a de facto two-stop shop, note also that each of the four public citation sources, GS, PubMed, EPMC and WOS, record different citation counts in that order).

As of April 2017 these sources record the following:

- Database team member cumulative co-authored publications have increased to 151.
- All team members now have ORCID IDs (from which an individual author citation graph can be generated in EPMC) as well as Edinburgh Research Explorer profiles.
- Most also have updated Google Scholar pages (*e.g.* Adam Pawson, Chris Southan and Joanna Sharman) as well as ResearchGate entries.
- IUPHAR reviews in BJP stand at 23 with an (EPMC) total citation count of 307
- The IUPHAR Pharmacological Reviews have reached 91 (some older ones do not retrieve with this simple query) with the (EPMC) citations adding up to 18,504.
- The BJP "Concise Guide" sets from 2013 and 2015 add up to 17 papers with an (EPMC) citation total of 2374.
- We now have five publications in the NAR Database issue which add up to 822 EPMC citations and 1036 in PubMed
- Our Wellcome Trust mandated grant linking score for EPMC full-text Gold Open Access (GOA) for our finished 099156/Z/12/Z award, now stands at 24 (but includes four false-positives).
- We are seeing high citation rates in some of our grant-linked articles because the BJP selected these as reference citations from the tables of links. These are topped by our NAR 2014 Database Issue (PMID 24234439) that now has 665 PubMed citations (we used to also track download statistics but unfortunately, Oxford University Press reset all the NAR access counts on their new website so our historical metrics are lost). The equivalent 2016 Database Issue (PMID 26464438) is now catching up at 225 PubMed citations. The Concise Guide to PHARMACOLOGY 2013/14: G-Protein Coupled Receptors (PMID 24517644) now has 403 PubMed citations.
- Altmetrics tracks mentions of our papers in a variety of online sources and institutions. Funders (including the Wellcome Trust) are taking increasing notice of the resulting interest "scores". Since Oct 2016 we note increases for our 2014 NAR article to 17 (< 6) while the 2016 article has 13 and the latest Concise Guide increased from 51 to 53 (the Wiley news outlets boosted the count).



- These Altmetric scores generally rank our papers around or above the 90th percentile compared to articles of the same age and source but 95th for the above.

THE GUIDE TO IMMUNOPHARMACOLOGY DATABASE (GtoImmuPdb)

GtoImmuPdb WEB INTERFACE DEVELOPMENT STATUS

The Guide to IMMUNOPHARMACOLOGY (GtoImmuPdb) has been developed as an extension to the existing Guide to PHARMACOLOGY (GtoPdb). The development of GtoImmuPdb aims to provide improved data exchange between immunology and pharmacology expert communities, so to better support research and development of drugs targeted at modulating immune, inflammatory or infectious components of disease. The underlying GtoPdb schema has been extended to incorporate new immune system specific data types (such as processes and cell types). It also means the existing GtoPdb website has been further developed to surface this new data and incorporate it into the existing search and browse mechanisms. GtoImmuPdb does not therefore have its own website, although it is being developed so that it can run from a unique domain (guidetoimmunopharmacology.org), with its own portal, which serves as a unique immunological access-point to the Guide to PHARMACOLOGY.

The first alpha-release (v1.0) of the Guide to IMMUNOPHARMACOLOGY (GtoImmuPdb) was made on 6th October 2016, since then 3 further releases have been made, the most recent being v4.0 on 23rd March 2017. Alpha-releases are only internal and run on our development server. Full technical details on the development progress of GtoImmuPdb can be found on our blog. The first public beta-release is currently being prepared for release in May 2017.

GtoImmuPdb PORTAL AND WEB-INTERFACE

The portal has its own unique branding (header bar, logo and colour scheme) to distinguish it, but retains many of the layout features from the main GtoPdb site. This consistency should help users already familiar with GtoPdb to orientate themselves with the new GtoImmuPdb. A tutorial is provided on navigating from the new portal (http://dev.guidetoimmunopharmacology.org/immuno/docs/GtoImmuPdb_Tutorial_v4.0.pdf)

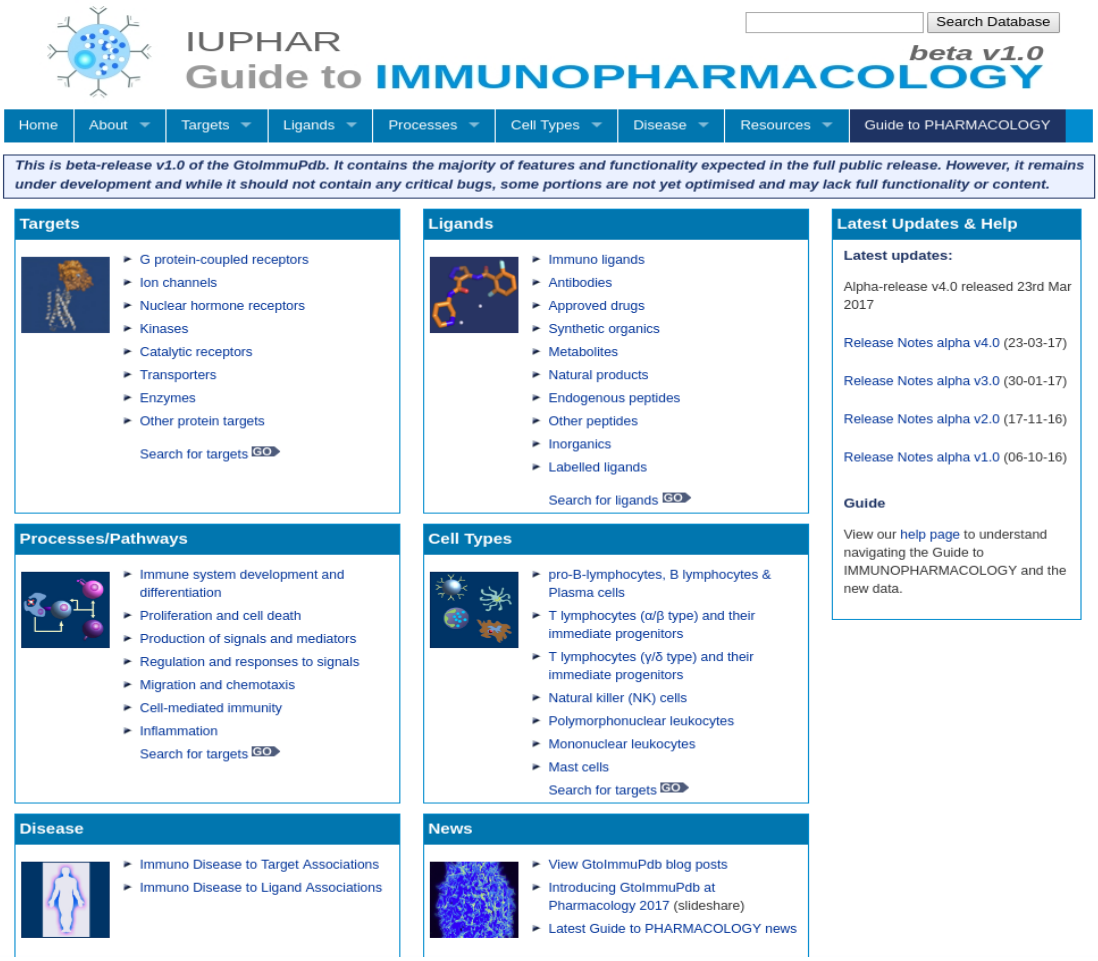
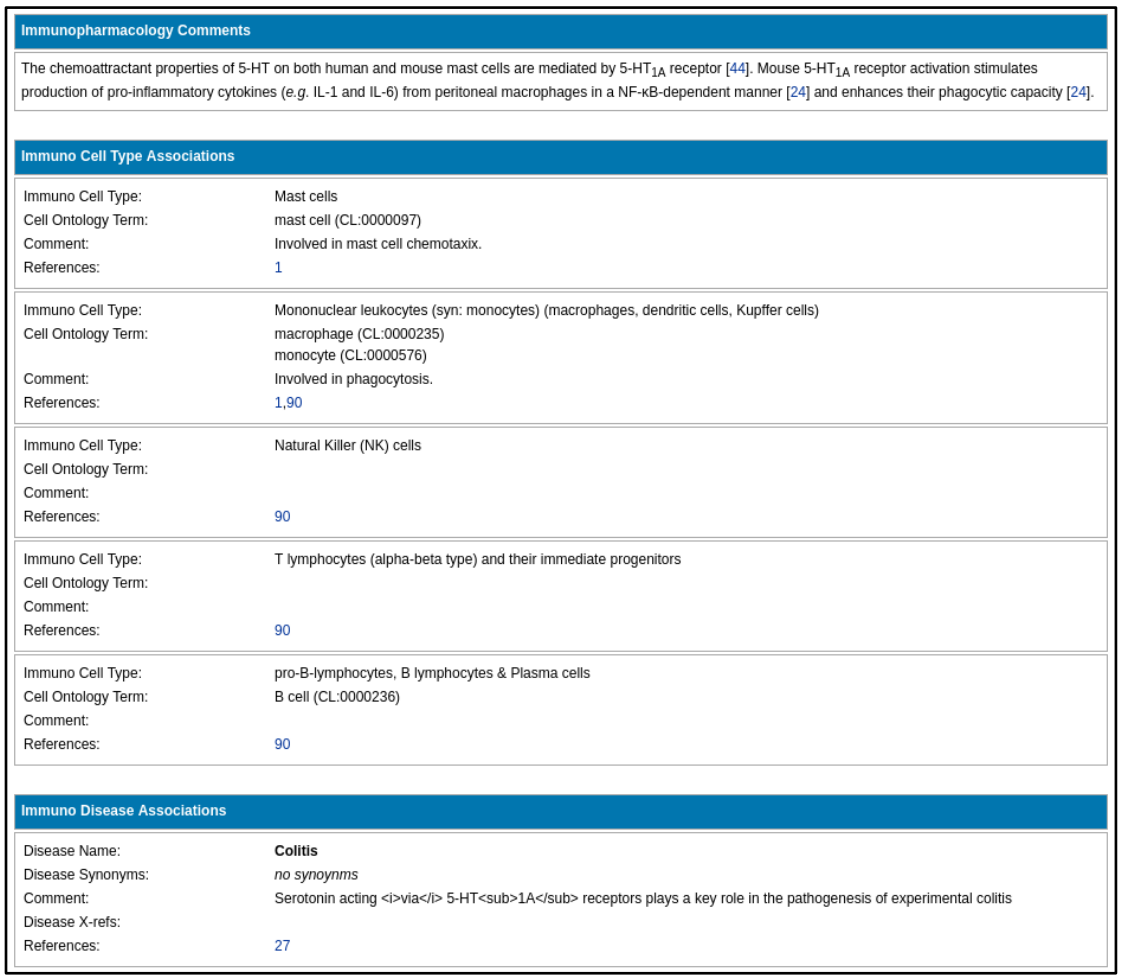


Figure A: GtoImmuPdb Portal.

The portal provides a starting-point for accessing data in GtoImmuPdb, tailored to the requirements of users with a specific interest in immunopharmacology. Browsing by target, ligand, process, cell-type & disease have been implemented in the current alpha-release (v4.0).

In preparation for the public beta-release (v1.0), the GtoImmuPdb site has been developed to run from its own URL domain, guidetoimmunopharmacology.org. Although this URL points to the same website at GtoPdb, using this specific domain provides the user with a distinct GtoImmuPdb view. So, if users are on the guidetoimmunology.org URL they'll see the same data as in GtoPdb, but with the GtoImmuPdb header and menus, plus the site will highlight items of immunological relevance on the appropriate pages. In most cases the GtoImmuPdb view can be toggled on and off switching between the GtoImmuPdb view and the GtoPdb view.

Targets - previously, on the target class and target family pages, targets of immunological relevance (those tagged as being in GtoImmuPdb in the database) were highlighted. This view could be toggled on and off. The detailed target pages had section surfacing data on general immunopharmacology comments, immunological process associations and cell type associations. A section on disease association data has now been added, giving disease names, synonyms, comments and external references (see Figure B). In the beta release the detailed view will have a toggle to switch between GtoImmuPdb view and GtoPdb, which effectively removes the highlighting on the immunological sections, but doesn't remove them altogether.



Immunopharmacology Comments

The chemoattractant properties of 5-HT on both human and mouse mast cells are mediated by 5-HT_{1A} receptor [44]. Mouse 5-HT_{1A} receptor activation stimulates production of pro-inflammatory cytokines (*e.g.* IL-1 and IL-6) from peritoneal macrophages in a NF-κB-dependent manner [24] and enhances their phagocytic capacity [24].

Immuno Cell Type Associations

Immuno Cell Type:	Mast cells
Cell Ontology Term:	mast cell (CL:0000097)
Comment:	Involved in mast cell chemotaxix.
References:	1
Immuno Cell Type:	Mononuclear leukocytes (syn: monocytes) (macrophages, dendritic cells, Kupffer cells)
Cell Ontology Term:	macrophage (CL:0000235) monocyte (CL:0000576)
Comment:	Involved in phagocytosis.
References:	1,90
Immuno Cell Type:	Natural Killer (NK) cells
Cell Ontology Term:	
Comment:	
References:	90
Immuno Cell Type:	T lymphocytes (alpha-beta type) and their immediate progenitors
Cell Ontology Term:	
Comment:	
References:	90
Immuno Cell Type:	pro-B-lymphocytes, B lymphocytes & Plasma cells
Cell Ontology Term:	B cell (CL:0000236)
Comment:	
References:	90

Immuno Disease Associations

Disease Name:	**Colitis**
Disease Synonyms:	*no synoynms*
Comment:	Serotonin acting via 5-HT_{1A} receptors plays a key role in the pathogenesis of experimental colitis
Disease X-refs:	
References:	27

Figure B - display of GtoImmuPdb data on detailed target view pages.

Ligands - Ligand list pages have now been developed, these are are linked to from the portal 'ligand' panel (Figure C1). They include an immuno tab that when selected lists all ligands tagged in the database as being included in GtoImmuPdb. The page has a toggle button to switch between the GtoImmuPdb and GtoPdb views. Under the GtoImmuPdb view, only ligands tagged in GtoImmuPdb are displayed under each ligand category. A new 'immuno ligand' icon has been created to be displayed in the table with the other icons when the ligand has been tagged in GtoImmuPdb. This icon is also used in the target detailed view pages, in

the interactions sections. The ligand summary page have also been developed to contain a section called ‘immunopharmacology’ which display and specific immunopharmacology comments and disease associations (Figure C2)

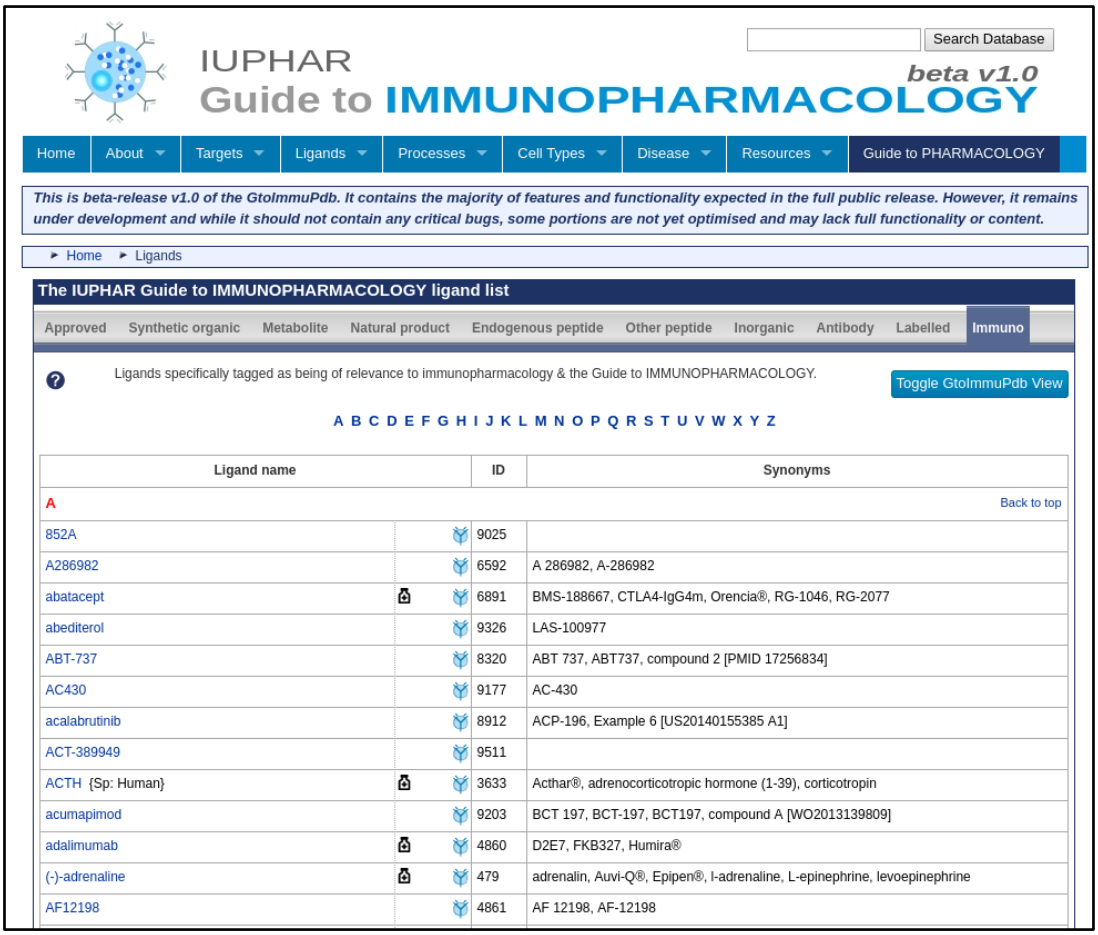


Figure C1 - GtoImmuPdb ligand list view, showing immuno tab selected.

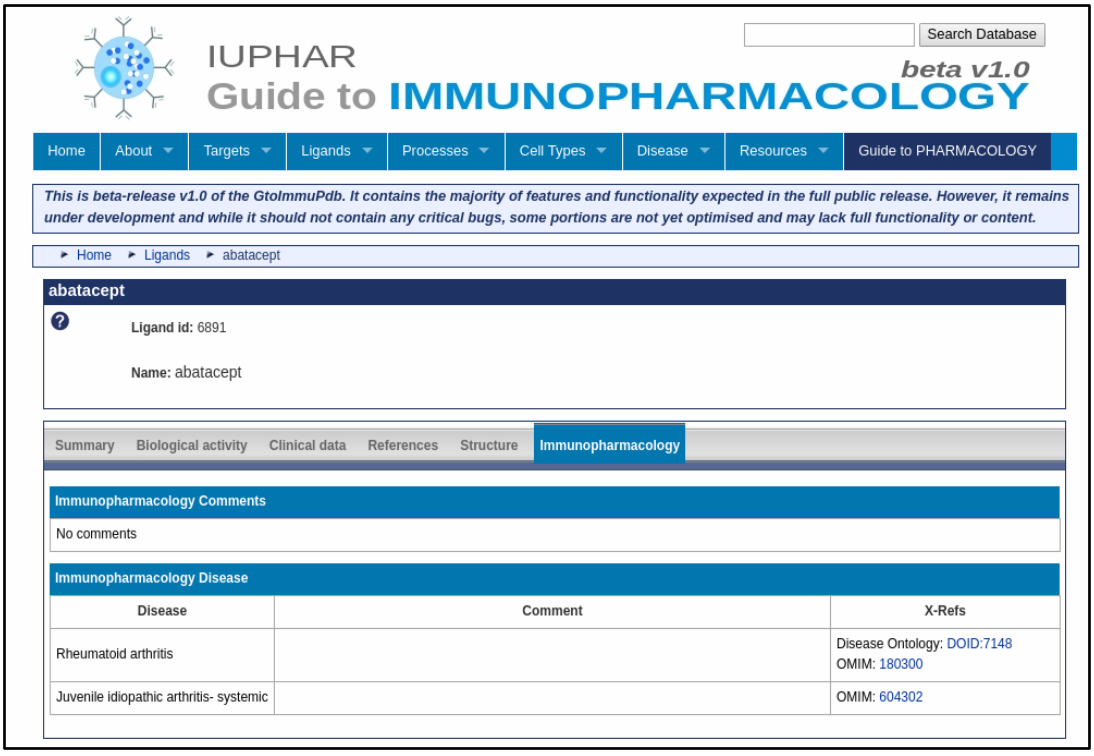


Figure C2. Immunopharmacology section (tab) of ligand summary page for abatacept.

Processes & Cell Types - previously pages that list all targets associated to either a process or a cell type category were developed. These are accessed by browsing *via* one of these categories from the portal. The pages are arranged as a series of tables, displaying associations to one of the main process or cell type categories at a time (selected *via* tabs). The tables are then split into section by target class. The data in the tables lists Gene Ontology biological process terms or Cell Ontology terms annotated to the target along with any specific curator comments about the association (Figure D). Some refinement of the display has been made to condense comments fields and improved ontology term display.

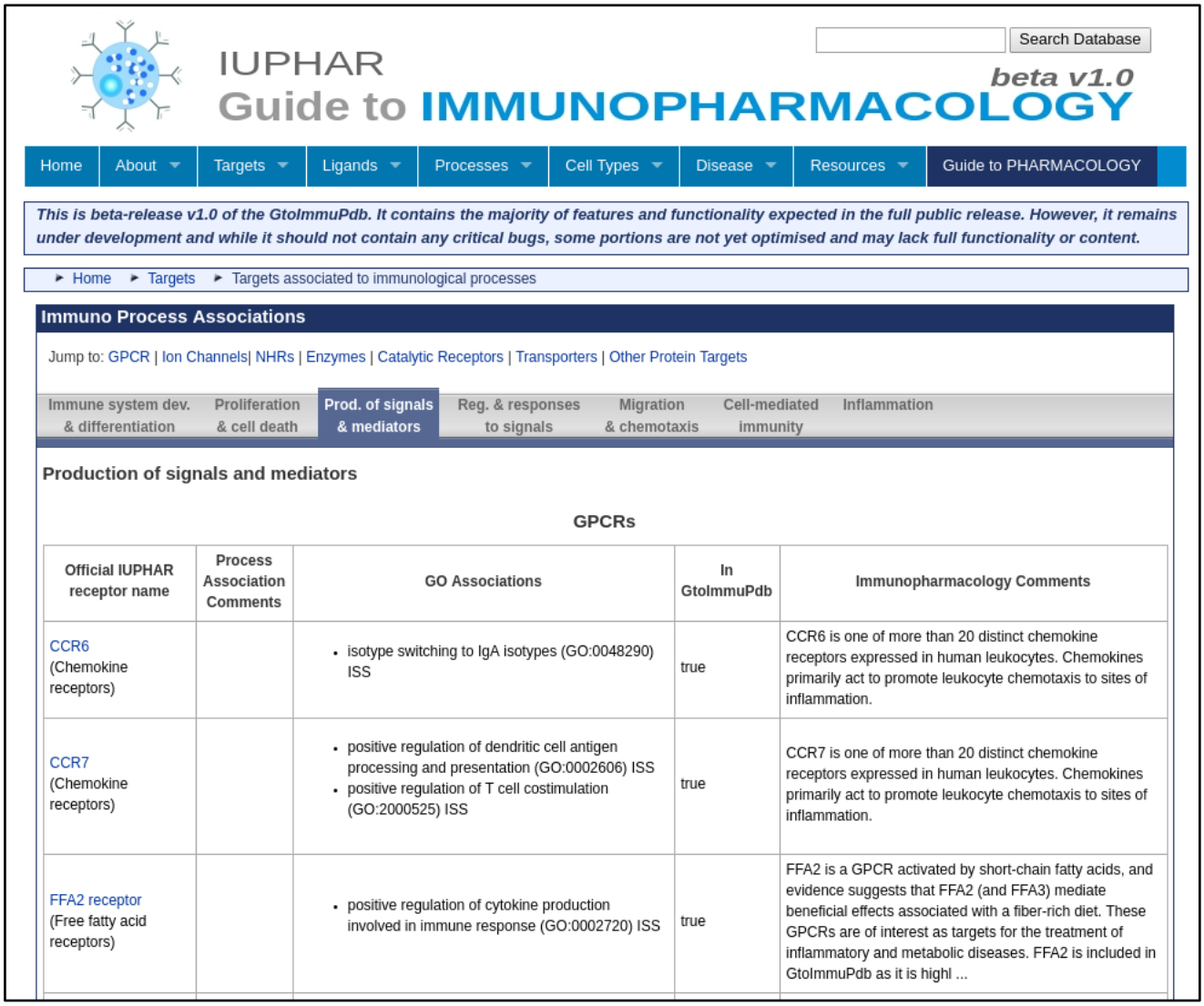


Figure D - GtoImmuPdb Process Associations page.

Diseases - a new page has been developed to display lists of diseases associated to either targets or ligands where the association or disease is of immunological relevance (Figure E). The page is accessed from the 'Disease' panel on the portal. The disease list page is designed to display all disease associations curated as part of the GtoImmuPdb and is divided (by a tab) into targets and ligand associations. The format of the list of disease associations is similar for both targets and ligands, showing one section or row per disease. Each section gives the disease name, external references and a count of the number of targets or ligands associated to it. The full list of targets or ligand is hidden but can be selected for display. When associations are displayed, curator comments are visible, and for targets the table indicates whether there are ligands for which the target is a primary target and if the ligands are approved drugs.

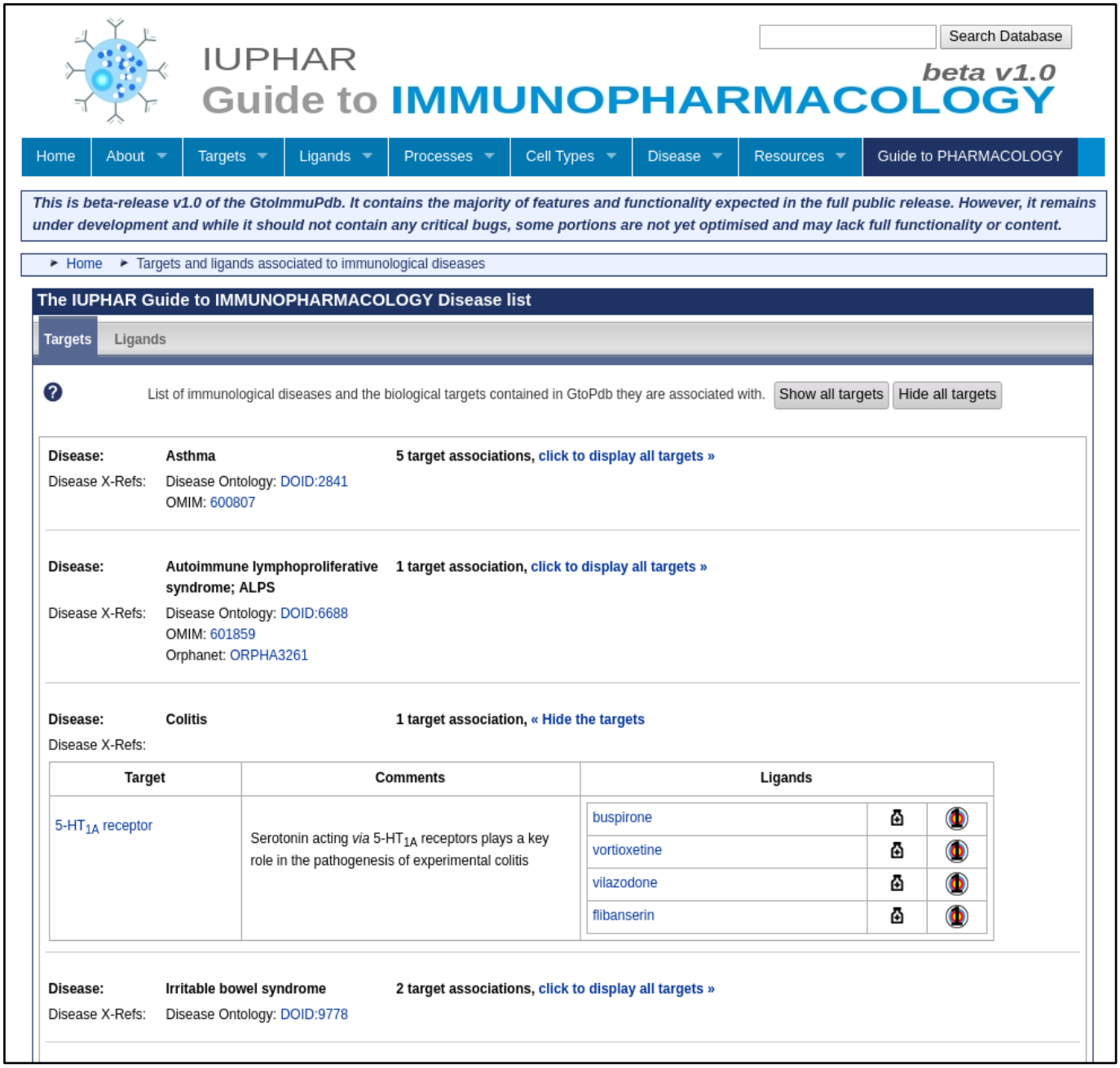


Figure E - GtoImmuPdb Disease Association page.

Searching - The ranking of search results has been developed to apply a weighting to targets and ligands that are returned from a search that are consider of greater immunological relevance. The weighting is only applied when searching from a GtoImmuPdb page (guidetoimmunopharmacology URL), not from the standard GtoPdb pages. The criteria used to determine the immune-relevancy (and weighting factor) of a given target or ligand is based on the amount of immunological data curated against it. For example, targets that have process, cell type and disease data annotated against them will rank higher than targets with only process data. The weighting factor is applied in addition to existing search weightings – so exact matches (to target or ligand name for example) will still score highest. We will be refining this relevancy scoring during testing of the beta release.

GtoImmuPdb DATA

GtoImmuPdb uses the same underlying database as GtoPdb and this has been extended to include and integrate GtoImmuPdb data. In addition to the process and cell type data already included are disease associations. The database schema has been extended to accommodate these data-types and to associate them with targets and ligands in the database.

IMMUNO PROCESS DATA

GtoImmuPdb has defined its own set of top-level immunological process categories against which targets in the database can be annotated and which form the basis of organising, navigating and searching for immunological processes and associations.
These categories are:

- Immune system development and differentiation
- Proliferation and cell death
- Production of signals and mediators
- Regulation and responses to signals
- Migration and chemotaxis
- Cell-mediated immunity
- Inflammation

We have associated sets of Gene Ontology (GO) terms with each of these categories. This enables us to auto-curate targets annotated to any of those terms (or their children) by GO into our top-level immunological categories. GO data is obtained *via* an OBO file (http://purl.obolibrary.org/obo/go.obo) for the ontology, which is edited to restrict it to immuno-specific terms. This file is then parsed to populate our database tables and further processing scripts are used to auto-curate targets to the top-level process terms. This step uses GO annotation information from UniProt, for human targets. This gave a total of 2,221 annotation to 724 targets.

The table below summaries the unique targets (UniProt) annotated under each category.

GtoImmuPdb 'High-Level' Process	*Distinct UniProt*
Immune System Development and Differentiation	200
Proliferation and Cell Death	54
Production of Signals and Mediators	110
Regulation and Responses to Signals	647
Migration and Chemotaxis	145
Cell-Mediated Immunity	254
Inflammation	549

Provision has been made in the database schema to capture curator comments against process information and annotations and the design is fully-adaptable to future changes.

CELL TYPE DATA

The Cell Ontology provides the formalised vocabulary against which we annotated target to cell type associations. GtoImmuPdb has defined its own set of top-level immunological cell type categories against which targets in the database can be annotated and which form the basis of organising, navigating and searching for immunological cell types and associations.
These categories are:

- pro-B-lymphocytes, B lymphocytes & Plasma cells
- T lymphocytes (alpha-beta type) and their immediate progenitors
- T lymphocytes (gamma-delta type) and their immediate progenitors
- Natural Killer (NK) cells
- Polymorphonuclear leukocytes (neutrophils, eosinophils, basophils)
- Mononuclear leukocytes (syn: monocytes) (macrophages, dendritic cells, Kupffer cells)
- Mast cells
- Innate lymphoid cells (added April 2017)

We have assigned one or more Cell Ontology terms (and IDs) to each of these categories. The assigned CO terms represents the highest level parent term(s) within the ontology for that category. For the purposes of annotation, it is these CO terms and their children that can be used when annotating a target to a given category.

Curators can add/remove/edit cell type association to targets using the submission tool. A target can be annotated to one or more of the top-level immuno cell type categories. A comment can be applied, as can any literature references. Additionally, the association can have any CO terms added. This is a way to annotating the association with ontology terms, and making the annotation higher resolution. For each association a list of CO terms is therefore stored and these are surfaced to the user on the detailed target page. Searches against CO terms will detect any cell type associations annotated with those terms (or their children). Curators can also edit a description on each top-level cell type category.

GtoImmuPdb DISEASE ASSOCIATION DATA

We have built on the existing disease data and tables in Guide to PHARMACOLOGY. These include a set of around 2,000 diseases and disease synonyms, stored in the disease and disease2synonym tables. The database currently maps these disease terms to 3 external resources: OMIM, Orphanet and the Disease Ontology. These resource help provide a formalised vocabulary against which we can annotated targets and ligands to diseases Annotating disease associations, curators can add/remove/edit disease association to targets and ligands using the submission tool. These can be annotated to one or more of the diseases in the database. Diseases themselves can be update and added. Both X-Refs and synonyms for disease can be added, edited and deleted. Comments can be applied, as can any literature references.

GtoImmuPdb SUBMISSION INTERFACE

The submission tool (used by the curators to insert and modify data in GtoPdb) had previously been modified to include tools to flag targets and ligands as being relevant to GtoImmuPdb. Extensions have been made so that the process, cell type and disease associations can be added, modified and deleted. In addition references (literature) can be added to each process or cell type association. Improvements have been made to how these new elements of the submission interface function so that data cannot be duplicated and can more accurately modified. This includes extension to provide rich text editing for comments so ligand and references links can be embedded within the general immunopharmacology comments.

GtoImmuPdb TARGET AND LIGAND CURATION STATUS

GtoImmuPdb CURATION STATS

- 428 targets tagged as in GtoImmuPdb:
  - 137 catalytic receptors
  - 128 enzymes
  - 83 gpcrs
  - 17 voltage-gated ion channels
  - 51 other proteins
  - 6 nuclear hormone receptors
  - 3 ligand-gated ion channels
- 719 ligands tagged as in GtoImmuPdb:
  - 382 synthetic organic
  - 101 antibodies
  - 204 peptides
  - 8 natural products
  - 1 inorganic
- Detailed lists on:
  - dev.guidetoimmunopharmacology.org/immuno/immunoHelpPage.jsp

DATA MODEL FOR POPULATING

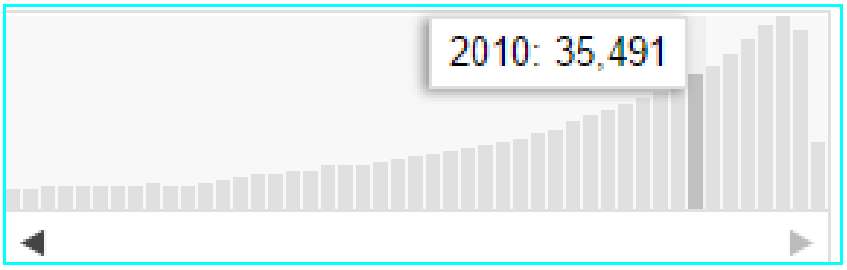
This is being strategically developed by the team with committee inputs. Technicalities can be provided but the design principle is to share the same basic database architecture between GtoPdb and GtoImmuPdb (*i.e.* with the latter being “forked” from the former) but add layers of immunology-relevant indexing. This includes protein-specific Gene Ontology (GO) terms related to immunological systems and inflammation, and mapping of these to a high level systems classification recommended by the committee.

GO INTERSECTS

Assessments are being updated with human protein lists associated with the terms GO:0002376 immune system process and GO:0006954 for the inflammatory response. These are being intersected with the new release of GtoPdb to establish which proteins we already have. These can be tagged if the GO link is directly publication-supported and the targets have immunopharmacology relevant ligands. Those we don't have can be triaged for eventual capture.

GETTING TO CONTENT: SEARCHING, COLLATION, EXTRACTION, ALERTING

With the objective of expanding GtoImmuPdb, different strategies have been explored to retrieve papers (and some other sources) of appropriate quality for extraction. They all worked to some extent but obviously varied in their balance of specificity vs recall as well as practical efficiency. One approach was to find "pre cooked" compendia, as exemplified by the Immunology Product Listing (Tocris), Immunology_Inflammation (MCE), Inflammation Inhibitors (conference) Medicines in Development for Autoimmune Diseases 2016 (54 page report). These have been useful but entity lists are difficult to extract from the PDFs. As mentioned previously we have also been focusing on review articles as the next high-curation-density documents after compendia and these provided a good database seeding yeild. We have now moved on to the primary the literature across the inflamm/immuno domains. Predictably this is a large corpus, as evidenced by an extended multi-term PubMed query that brings back 0.95 million papers (see terms in ribbon or facet).



However, we can slice 'n dice, for example 1494 of these were BJP papers, 336 in the last 5 years, were linked to 63 PubChem compounds (CIDs), of which 43 were in GtoPdb and about ½ of which had GtoImmuPdb links.

While we had already exploited it for GtoPdb CiteULike, as described in the previous report, has become a key resource for GtoImmuPdb in the curation team. We now have 435 publications tagged as immphar" (n.b. this is a public collection so anyone can engage with us). A selection is shown in the snapshot below.

✓ **Kinase inhibition, competitive binding and proteasomal degradation: resolving the molecular function of the suppressor of cytokine signaling (SOCS) proteins.**
Immunological reviews, Vol. 266, No. 1. (July 2015), pp. 123-133
by Edmond M. Linossi, Sandra E. Nicholson
posted to immpharm immunology kinase_inhibitors by efaccenda on 2017-04-12 13:06:40 ★★★
■ Abstract ■ Notes ■ Copy

✓ **Binding site elucidation and structure guided design of macrocyclic IL-17A antagonists.**
Scientific reports, Vol. 6 (16 August 2016)
by Shenping Liu, Leslie A. Dakin, Li Xing, et al.
posted to immpharm by cdsouthan keyed Liu2016Binding on 2017-04-12 11:43:23 ★★/
■ Abstract ■ Copy ■ My Copy

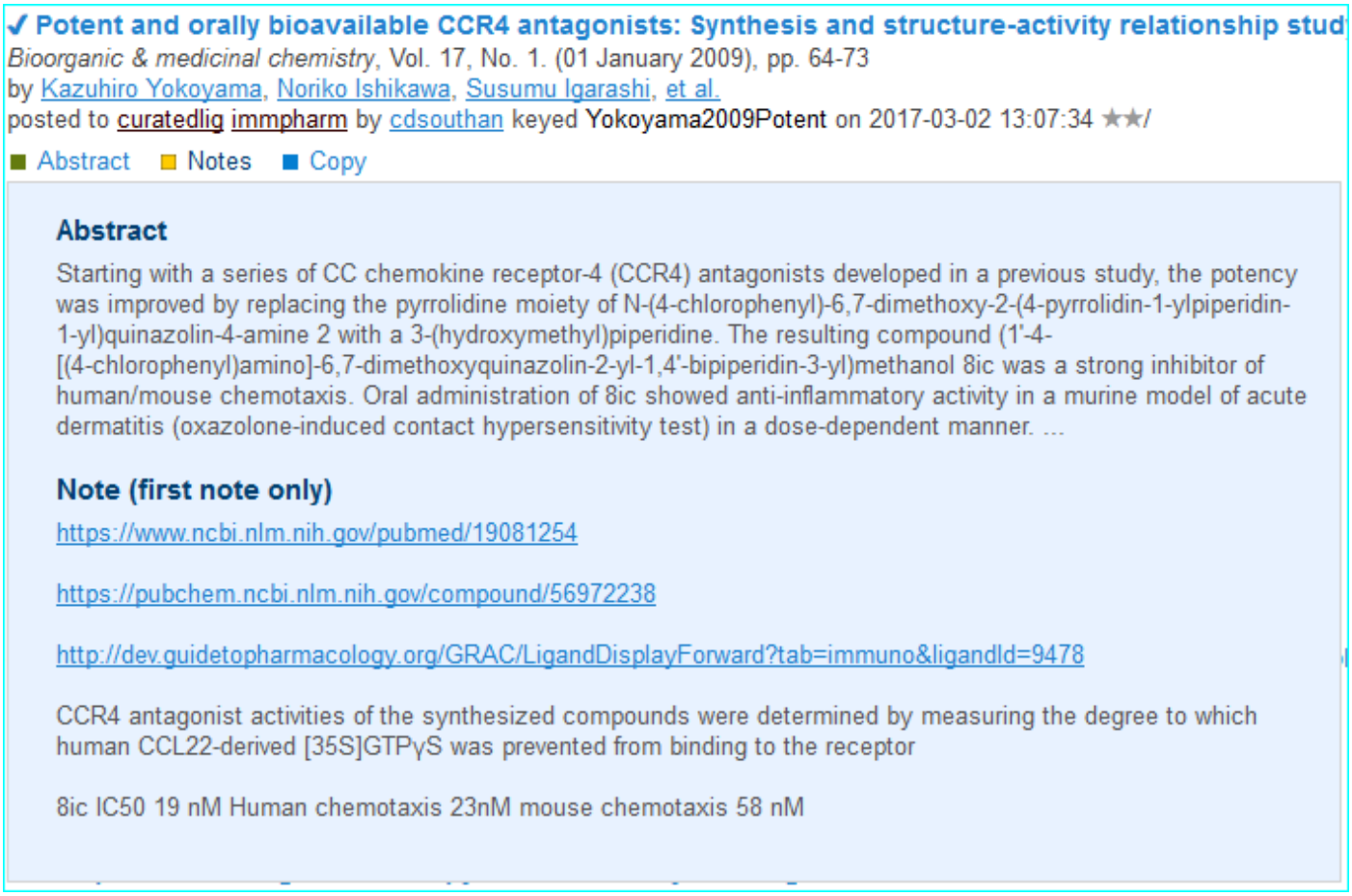
✓ **Drug Discovery Targeting Bromodomain-Containing Protein 4.**
Journal of medicinal chemistry (02 March 2017)
by Zhiqing Liu, Pingyuan Wang, Haiying Chen, et al.
posted to immpharm by cdsouthan keyed Liu2017Drug on 2017-04-10 21:35:11 ★★/
■ Abstract ■ Copy ■ My Copy

✓ **Discovery of a First-in-Class Receptor Interacting Protein 1 (RIP1) Kinase Specific Clinical Candidate (GSK2982772) for the Treatment of Inflammatory Diseases.**
Journal of medicinal chemistry, Vol. 60, No. 4. (23 February 2017), pp. 1247-1261
by Philip A. Harris, Scott B. Berger, Jae U. Jeong, et al.
posted to immpharm by cdsouthan keyed Harris2017Discovery on 2017-04-10 21:21:49 ★★/
■ Abstract ■ Copy ■ My Copy

✓ **Development of highly selective Kv1.3-blocking peptides based on the sea anemone peptide ShK.**
Marine drugs, Vol. 13, No. 1. (16 January 2015), pp. 529-542
by Michael W. Pennington, Shih Chieh C. Chang, Satendra Chauhan, et al.
posted to immpharm by cdsouthan keyed Pennington2015Development on 2017-04-09 22:38:44 ★★/
■ Abstract ■ Copy ■ My Copy

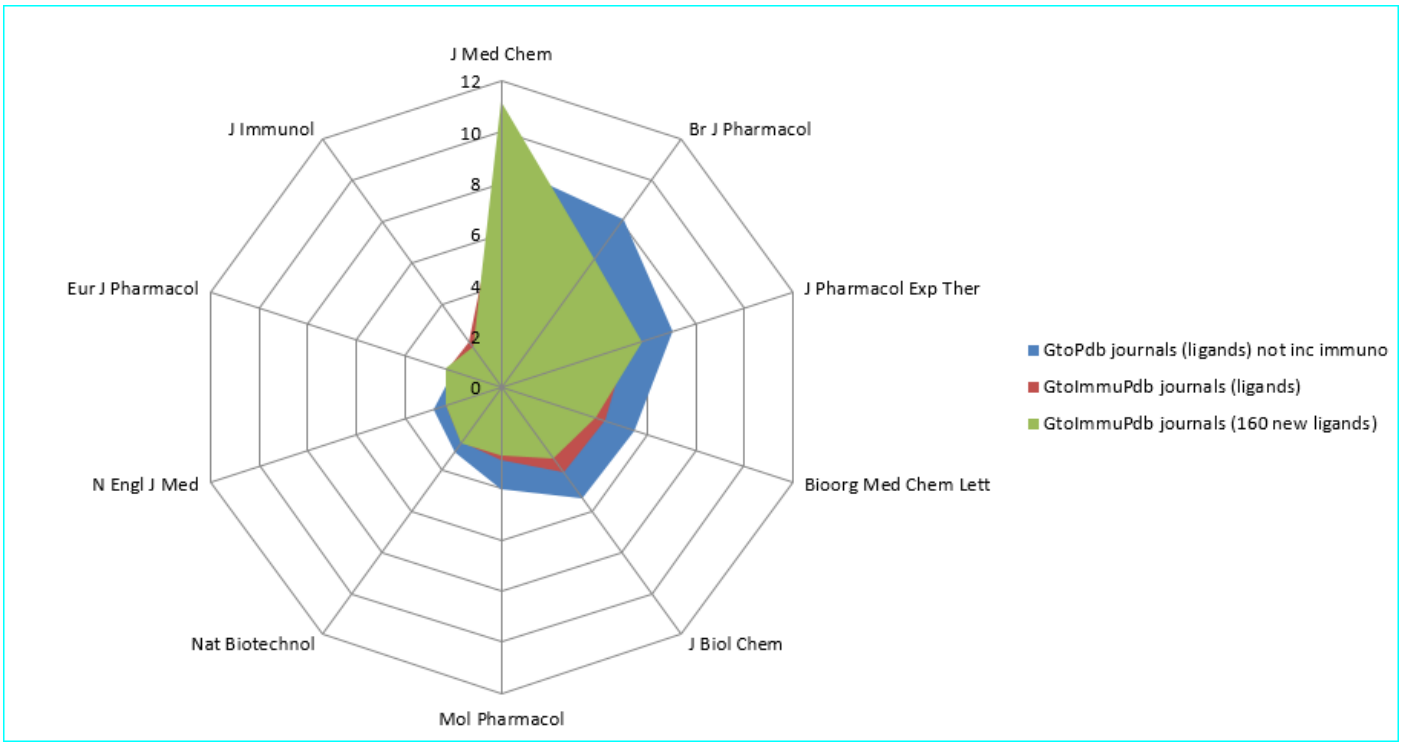
The list is approximately split between curatable papers and non-curatable background articles. We use the notes feature as a useful pre-curation tool by adding cross pointers, including for the Swiss-Prot targets,

PubChem IDs for ligands and binding constants for the interaction. We also add notes for the GtoImmuPdb fields, disease mappings and, post-curation, add in the ligand and target links. An example of a set of entry notes is shown below.



As we curate papers and/or explore the information space around entities (*e.g.* from the corpora mentioned), we have developed an expansion strategy that exploits familiar PubMed functionality. Simply put we "walk" backwards and forwards in time using "Similar articles". We also check "Cited by" as an alternative forward point (*i.e.* that may connect to different articles than the similarity heuristic). This not only gives us a quality inference for the paper in question but can pick up a more recently published ligand with improved properties. Note also we often pick up a target review (especially for a new target) and add it to the CiteUlike list.

In terms of alerting sources, Twitter has become increasingly effective. We get regular news distillations from Immune Regulation News , Human Immune News British Society for Immunology, Edinburgh Centre for Inflammation Research as well as journals such as Nature Immunology Journal of Immunology and others.. While expansion and extraction continues we would really welcome committee evaluation of our immphar list and the database content to see if we had significant gaps and or biases in curation triage. As an assessment of journal selectivity we have been looking at differential capture in the two databases. This can be seen in the radar plot below.



Details of the data can be provided but we can see, as expected, the beginnings of divergence for GtoImmuPdb with a relative increase in primary references (*i.e. in vitro* activity data) from J. Med. Chem. but less secondary references, such as BJP (*i.e. in vivo* or clinical reports).