

The ChEMBL Database

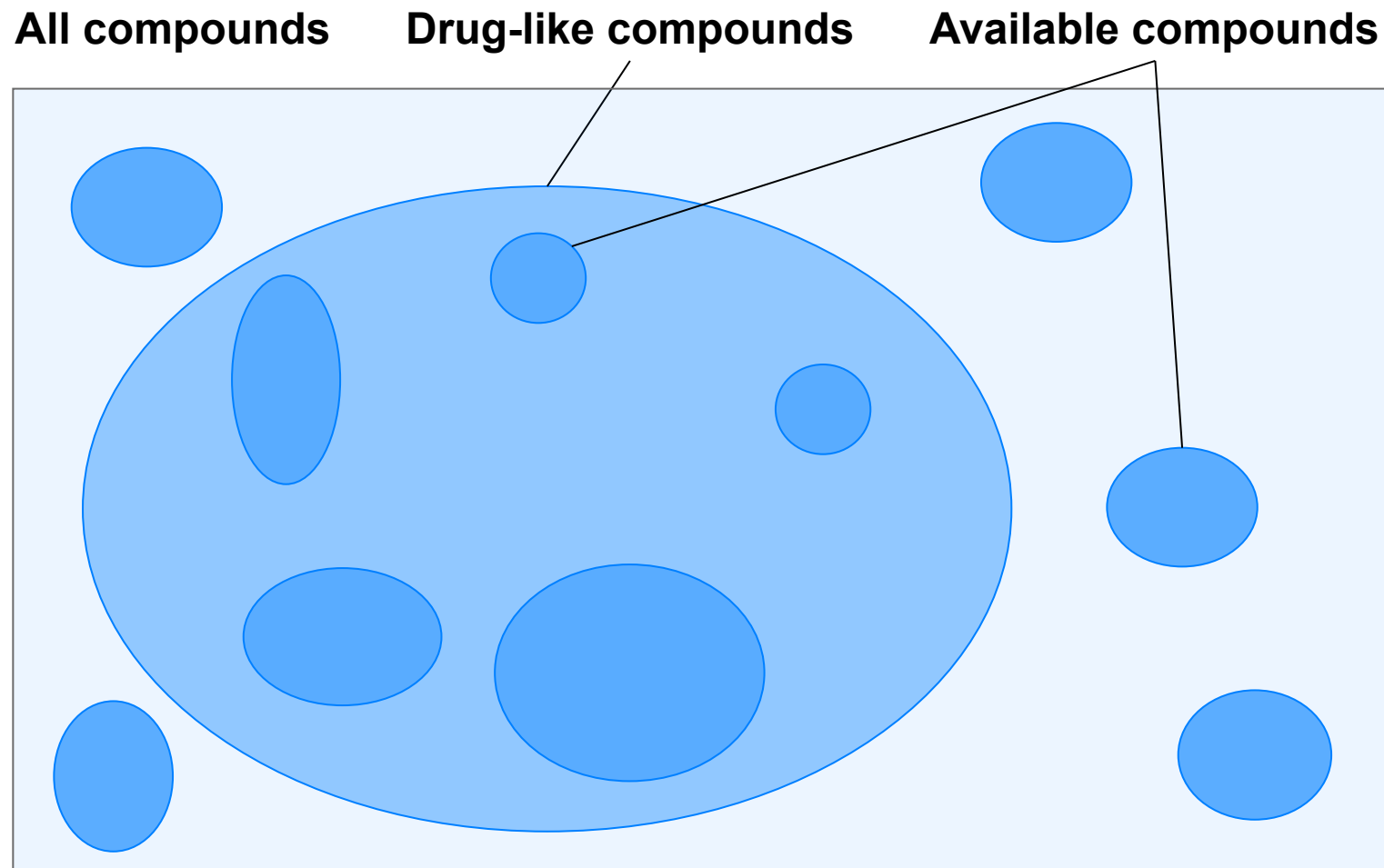
ICIC 2012
Berlin, Germany
October 2012

John P. Overington
EMBL-EBI

jpo@ebi.ac.uk

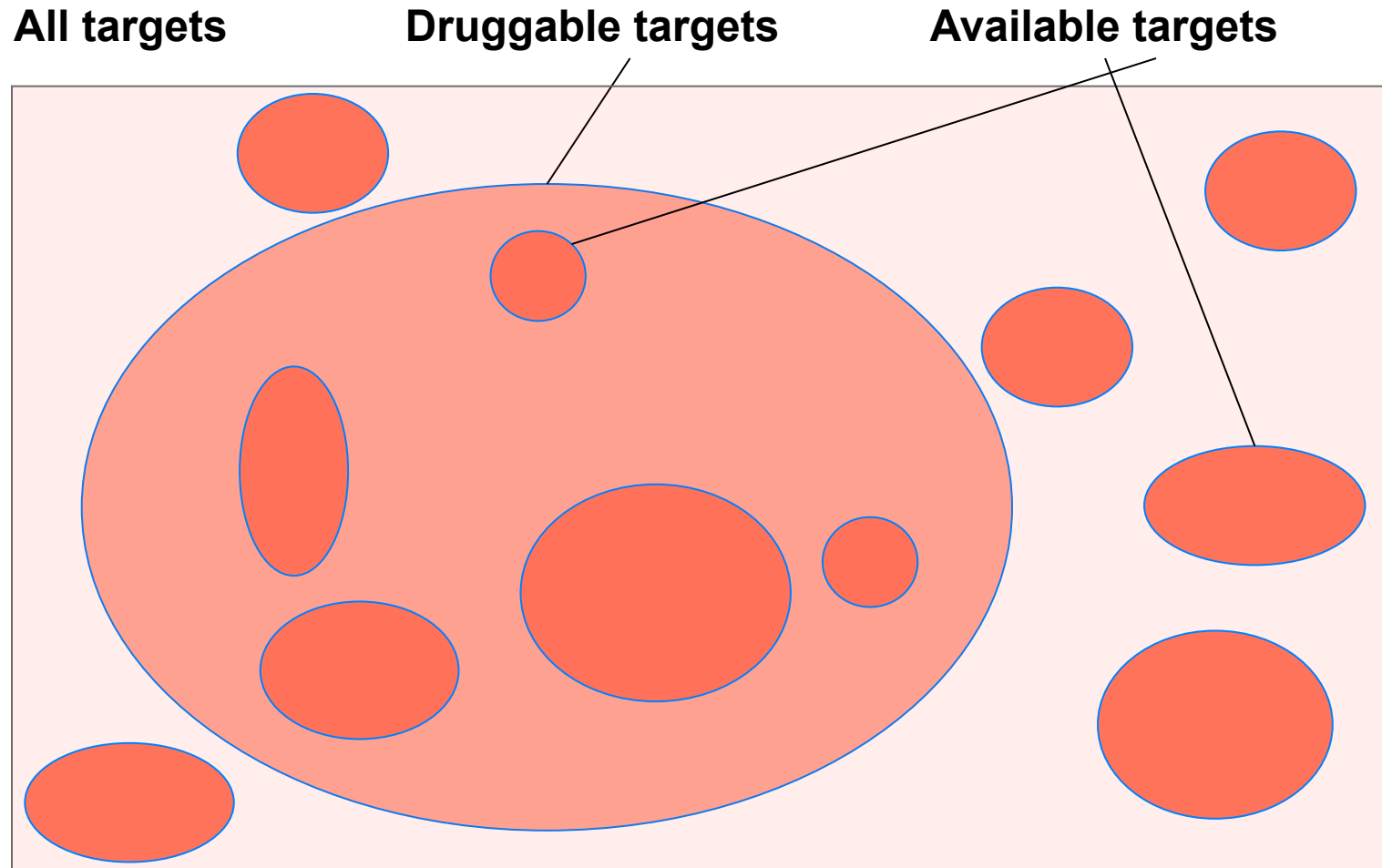


Chemical Space



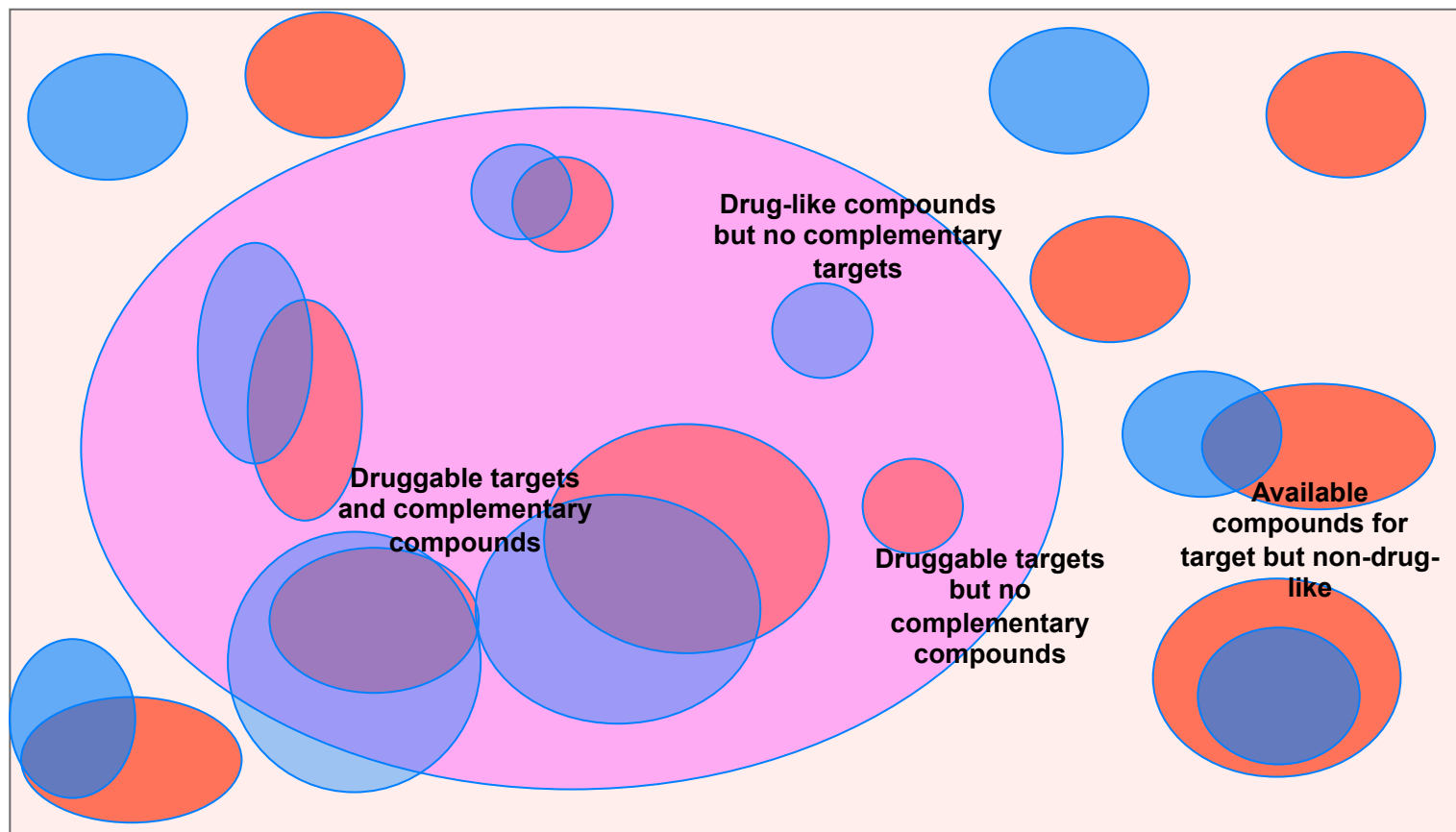
Only certain molecules have features consistent with good pharmacological properties

Target Space



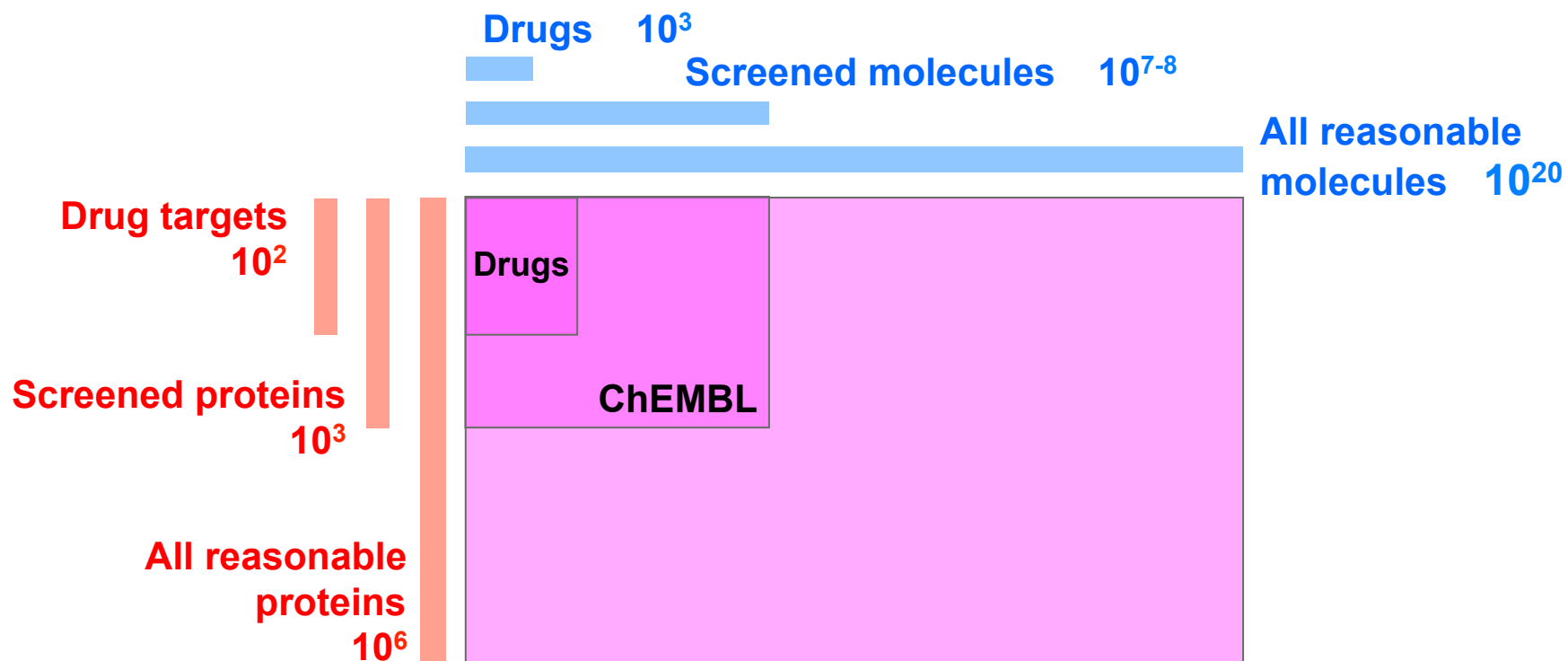
Only certain targets have binding sites capable of ligand efficient binding of drug-like ligands

Accessible Pharmacological Space



Chemogenomics

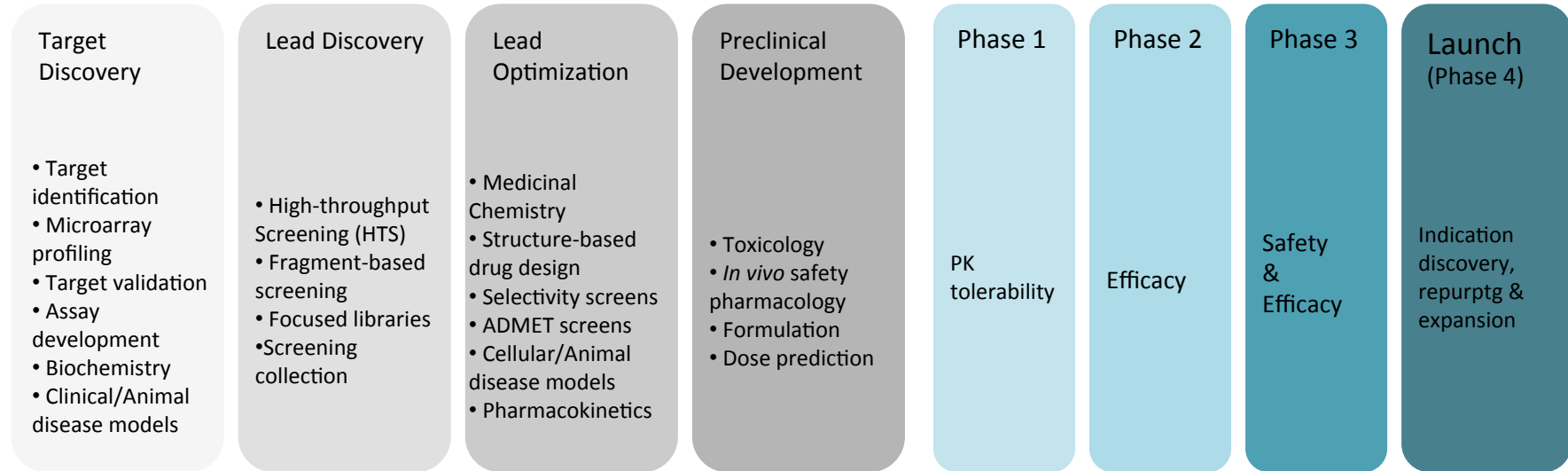
Exploration of bioactivity space at genomic scale
Structure Activity Relationship (SAR)



ChEMBL Database

- <http://www.ebi.ac.uk/chembl>
- Funded by a Strategic Award from the Wellcome Trust
- World's largest primary source of Open pharmacology/drug discovery data
 - Contains synthetic small molecules, natural products and biologicals
 - Strong integration and annotation of chemical *and* biological data
 - OSINT approach to data gathering
 - Tight integration with other EBI resources
 - Ensembl, 1000 Genomes, UniProt, PDBe, ArrayExpress, Atlas....
 - Data sharing agreements in place with key public resources, *e.g.* PubChem
- Open Data – CC-BY-SA licence
- Free downloads, secure private searching,...
- REST web service API

Drug Discovery



Discovery

Development

Use

Med. Chem. SAR

Clinical Candidates

Drugs

ChEMBL
content

~1,400,000 compound records
>10,000,000 bioactivities
~46,000 abstracted papers
~9,000 targets

~12,000 clinical
candidates

~1,600
drugs

Only ~1% of Genome is a Drug Target

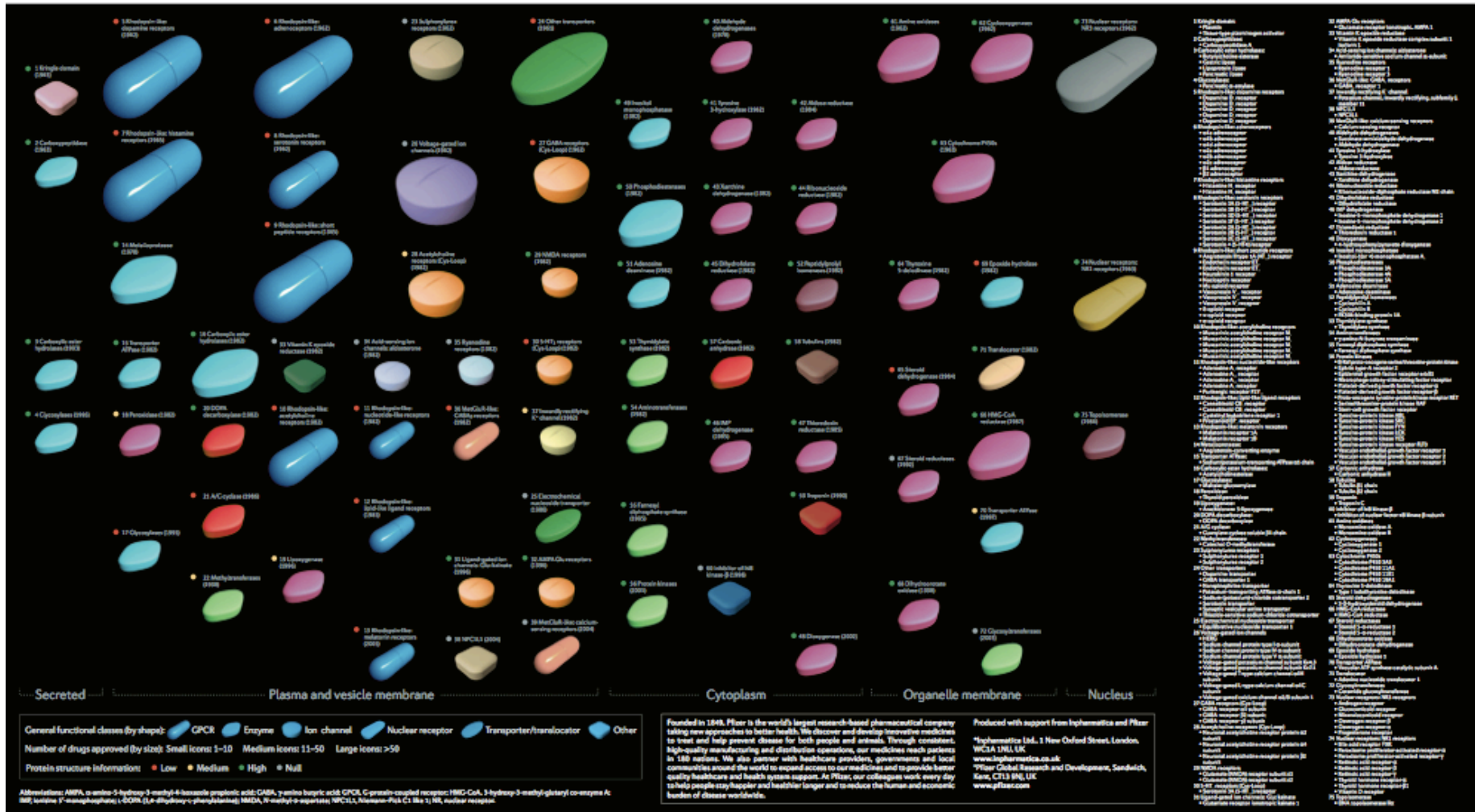


The molecular pharmacopoeia

The human targets of FDA-approved oral drugs

John P. Overington*, Bissan Al Lazikani* and Andrew L. Hopkins†

The molecular pharmacopoeia of 186 human protein drug targets for FDA-approved oral drugs. The drug targets are grouped into target subfamilies where there are multiple related drug targets. The size of the icons represents the number of drugs approved for that target or target subfamily. The horizontal axis illustrates the cellular location of a drug target. The vertical axis illustrates an approximate timeline depicting when the first drug for a target or target subfamily was approved (older drug targets are at the top of the chart). The dates next to the targets illustrate the year in which the first USAN (United States Adopted Name) was assigned for the first drug against that target or any target in the subfamily (which usually occurs in the late stages of clinical development). The availability of protein structural information for the target or target subfamily is illustrated by the coloured dot next to the target name. The shapes of the icons represent the general functional classes of drug targets and related groups within a functional class are coloured the same.



Drug Approvals

ChEMBLdb

https://www.ebi.ac.uk/chembl/db/drug/approvals

ChEMBL ChEMBL-og Twitter Google+ Citations Analytics Facebook LinkedIn F1000 Wikipedia PubMed BioMedBridges Science (6) EBI Computing Finance Travel

EMBL-EBI Enter Text Here Find Terms of Use Privacy Cookies

Databases Tools Research Training Industry About Us Help Site Index

ChEMBL

ChEMBLdb ChEMBL-NTD Kinase SARfari GPCR SARfari DrugEBility ChEMBL Group Downloads Web Services FAQ

ChEMBLdb Statistics

- DB: ChEMBL_13
- Targets: 8,845
- Compound records: 1,296,266
- Distinct compounds: 1,143,682
- Activities: 6,933,068
- Publications: 44,682

ChEMBL Blog

- New Drug Approvals 2012 - Pt. XII - pertuzumab (Perjeta™)
- Drug Side Effect Prediction and Validation

EBI > Databases > Small Molecules > ChEMBL Database > Home

Search ChEMBLdb... Compounds Targets Assays Activity Source Filter

ChEMBLdb Compound Search Protein Target Search Browse Targets Browse Drugs **Drug Approvals**

The following table lists US Food and Drug Administration (FDA) drug approvals for New Molecular Entities (NMEs) in 2009, 2010, and 2011. Links are to the corresponding Drug Approval Monographs on the ChEMBL-og.

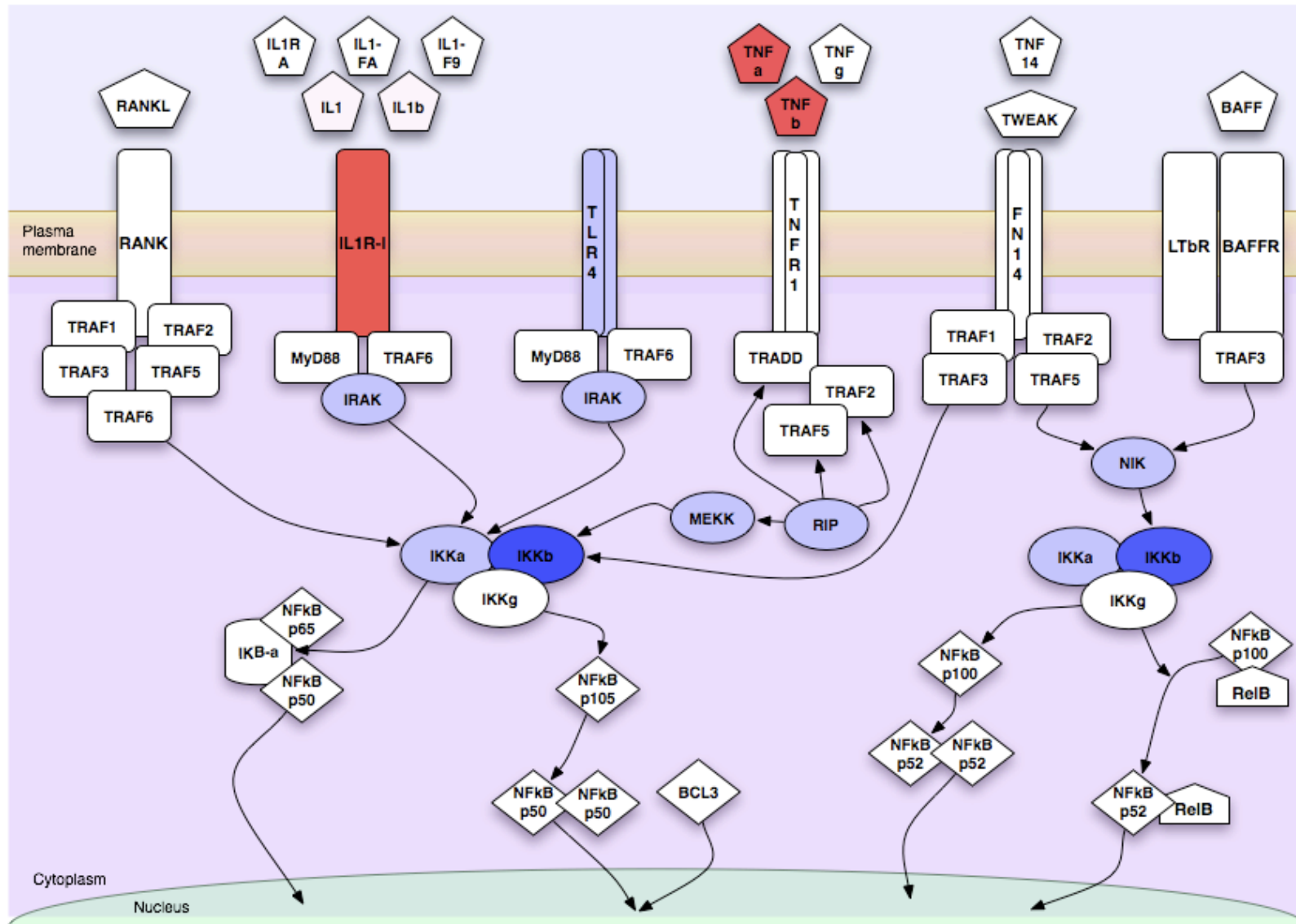
Subscribe to the RSS feed to receive new drug approvals [RSS](#)

Browse Drug Approvals

Generic Name	Trade Names	ATC Code	Date of Approval	Drug Monograph	Icon
Taiglucerase alfa	Elelyso	A16AB11	01-May-2012	http://chembl.blogspot.co.uk/2012/05/new-drug-approvals-2012-pt-xi.html	
Avanafil	Stendra	Not Assigned	27-Apr-2012	http://chembl.blogspot.co.uk/2012/05/new-drug-approvals-2012-pt-x-avanafil.html	
Florbetapir F 18	Amyvid	Not Assigned	06-Apr-2012	http://chembl.blogspot.co.uk/2012/04/new-drug-approvals-2012-pt-ix.html	
Peginesatide	Omontys	Not Assigned	27-Mar-2012	http://chembl.blogspot.co.uk/2012/04/new-drug-approvals-2012-pt-viii.html	
Lucinactant	Surfaxin	R07AA30	06-Mar-2012	http://chembl.blogspot.co.uk/2012/03/new-drug-approvals-2012-pt-vii.html	
Tafuprost	Zioptan	S01EE05	13-Feb-2012	http://chembl.blogspot.com/2012/02/new-drug-approvals-2012-pt-vi.html	
Ivacaftor	Kalydeco	R07AX02	31-Jan-2012	http://chembl.blogspot.com/2012/02/new-drug-approvals-2012-pt-v-ivacaftor.html	
Vismodegib	Erivedge	Not Assigned	31-Jan-2012	http://chembl.blogspot.com/2012/02/new-drug-approvals-2012-pt-iii.html	
Axitinib	Inlyta	L01XE17	27-Jan-2012	http://chembl.blogspot.com/2012/01/new-drug-approvals-2012-pt-iii-axitinib.html	
Ingenol Mebutate	Picato	Not Assigned	23-Jan-2012	http://chembl.blogspot.com/2012/01/new-drug-approvals-2012-pt-ii-ingenol.html	
Glucarpidase	Voraxaze	V03AF09	17-Jan-2012	http://chembl.blogspot.com/2012/01/new-drug-approvals-2012-pt-i.html	
Asparaginase Erwinia chrysanthemi	Erwinaze	L01XX02	18-Nov-2011	http://chembl.blogspot.com/2011/11/new-drug-approvals-2011-pt-xxx.html	
Aflibercept	Eylea	S01LA05	18-Nov-2011	http://chembl.blogspot.com/2011/11/new-drug-approvals-2011-pt-xxx.html	
Ruxolitinib	Jakafi	L01XE18	16-Nov-2011	http://chembl.blogspot.com/2011/11/new-drug-approvals-2011-pt-xxix.html	

View 1 - 87 of

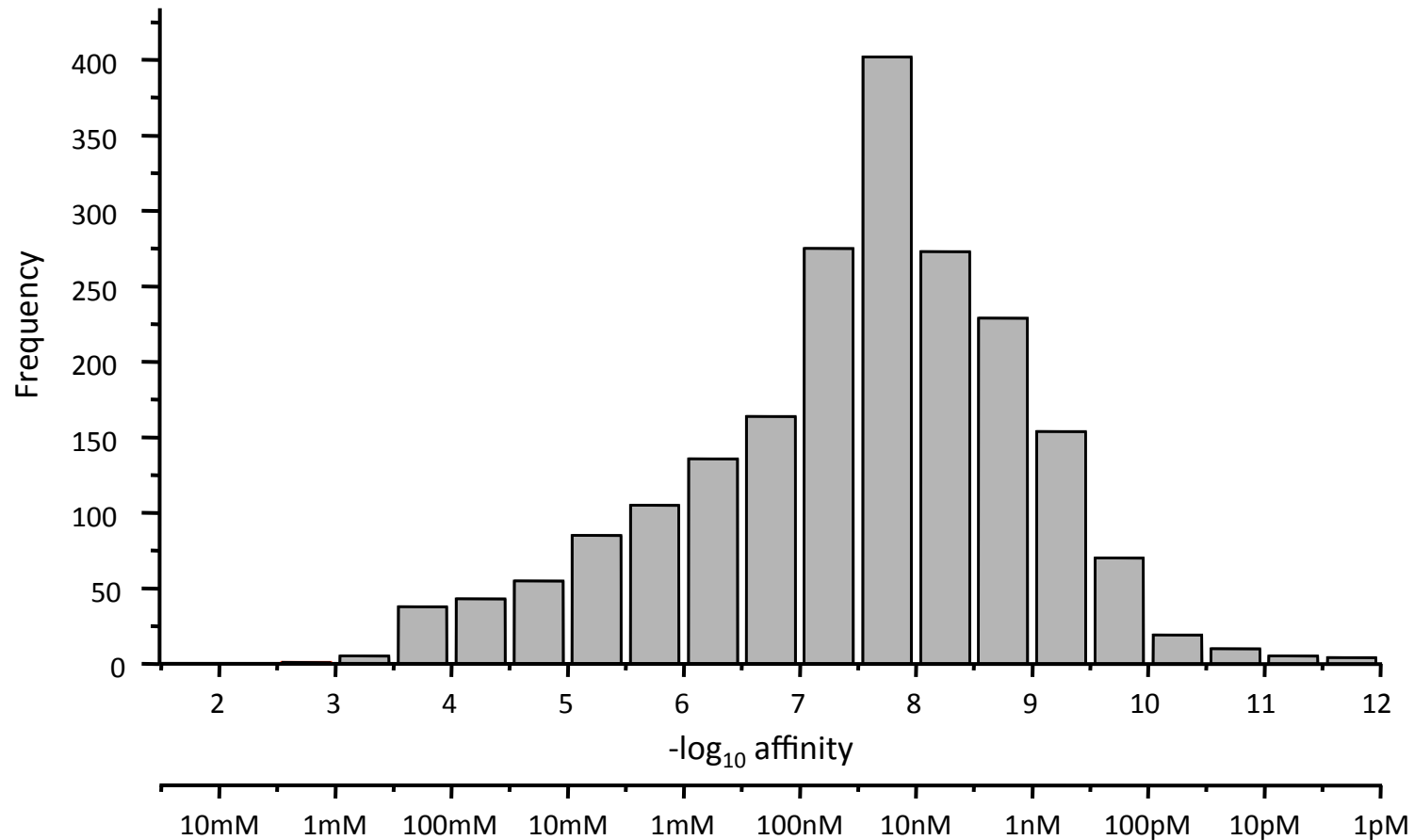
FDA Approved Drugs



NFκB Pathway – key control mechanism for inflammation

Affinity of Drugs for their 'Targets'

K_i , K_d , IC_{50} , EC_{50} , & pA_2 endpoints for drugs against their 'efficacy targets'

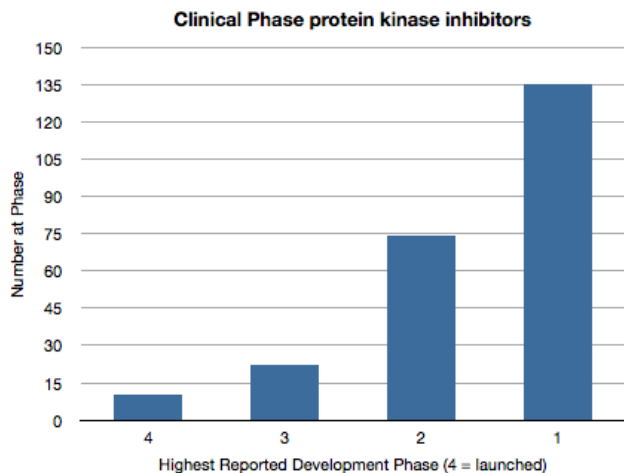


Overington, et al, *Nature Rev. Drug Discov.* **5** pp. 993-996 (2006)

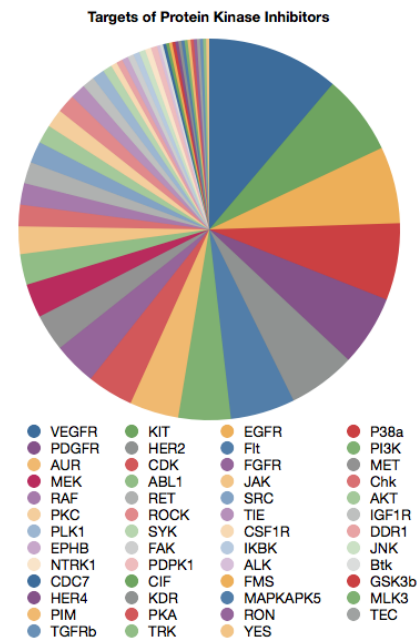
Gleeson et al, *Nature Rev. Drug Discov.* **10** pp. 197-208 (2011)

Clinical Candidates

- Collection of clinical development candidates
 - Contains ~12,000 2-D structures/sequences
 - Estimated size ~35-45,000 compounds
 - Work in progress
 - *e.g.* Protein kinases, 393 distinct clinical candidates

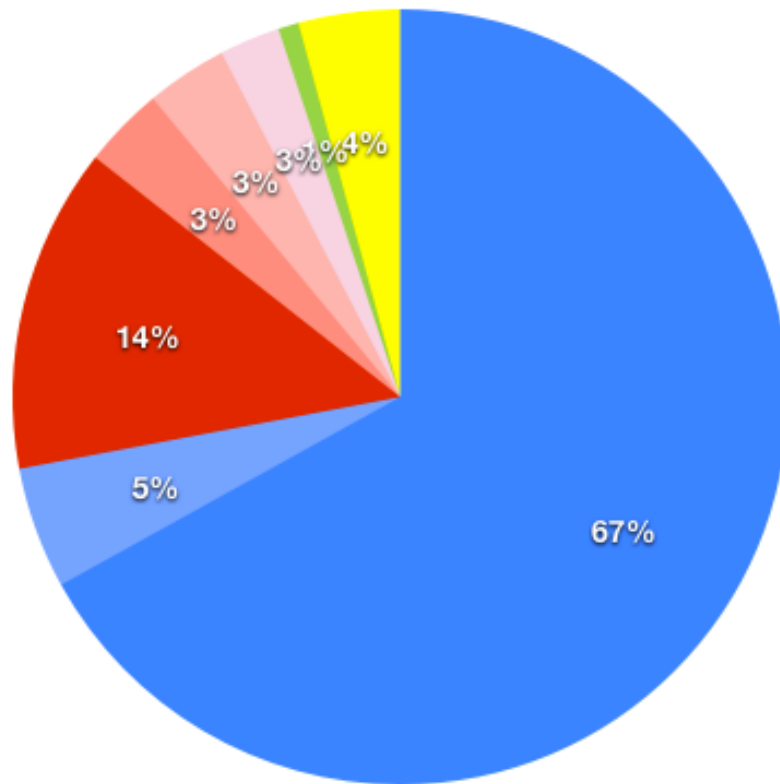


Compound Code	INN	Trade Name	Highest Phase	Company	Primary Targets	Primary Approved Indications	Approved Year (US)	CAS Numbers	Mol. wt.
AB-1000	None	None	2	AB Science	c-KIT, PDGFR			41 79229-74-5	328
ABT-899	None	None	2	Abbott	www.abbott.com PDGFR, VEGFR2			37 796801-16-3	328
AC-220	None	None	4	AstraZeneca	www.astrazeneca.com c-KIT, CSF1R, PDGFR, PDGFR, Ret, VEGFR1, VEGFR2			85	328
AMG-708	None	None	3	Amgen	www.amgen.com c-KIT, PDGFR, PDGFR, Ret, VEGFR1, VEGFR2			2 807476-20-3	328
AMG-548	None	None	2	Amgen	www.amgen.com			47	328
AMG-707	None	None	2	Amgen	www.amgen.com			88	328
APRY-162	None	None	1	Array BioPharma	www.arraybio.com MEK			85	328
APRY-100	None	None	1	Array BioPharma	www.arraybio.com ERK2			92	328
APRY-104	None	None	1	Array BioPharma	www.arraybio.com ERK1, ERK2			92	328
APRY-101	None	None	1	Array BioPharma	www.arraybio.com MEK			92	328
APRY-102	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-106	None	None	2	Array BioPharma	www.arraybio.com MEK			92	328
APRY-107	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-108	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-109	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-110	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-111	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-112	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-113	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-114	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-115	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-116	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-117	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-118	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-119	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-120	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-121	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-122	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-123	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-124	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-125	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-126	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-127	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-128	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-129	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-130	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-131	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-132	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-133	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-134	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-135	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-136	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-137	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-138	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-139	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-140	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-141	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-142	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-143	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-144	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-145	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-146	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-147	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-148	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-149	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-150	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-151	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-152	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-153	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-154	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-155	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-156	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-157	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-158	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-159	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-160	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-161	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-162	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-163	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-164	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-165	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-166	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-167	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-168	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-169	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-170	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-171	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-172	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-173	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-174	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-175	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-176	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-177	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-178	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-179	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-180	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-181	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-182	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-183	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-184	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-185	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-186	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-187	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-188	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-189	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-190	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-191	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-192	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-193	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-194	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-195	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-196	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-197	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-198	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-199	None	None	2	Array BioPharma	www.arraybio.com			92	328
APRY-200	None	None	2	Array BioPharma	www.arraybio.com			92	328



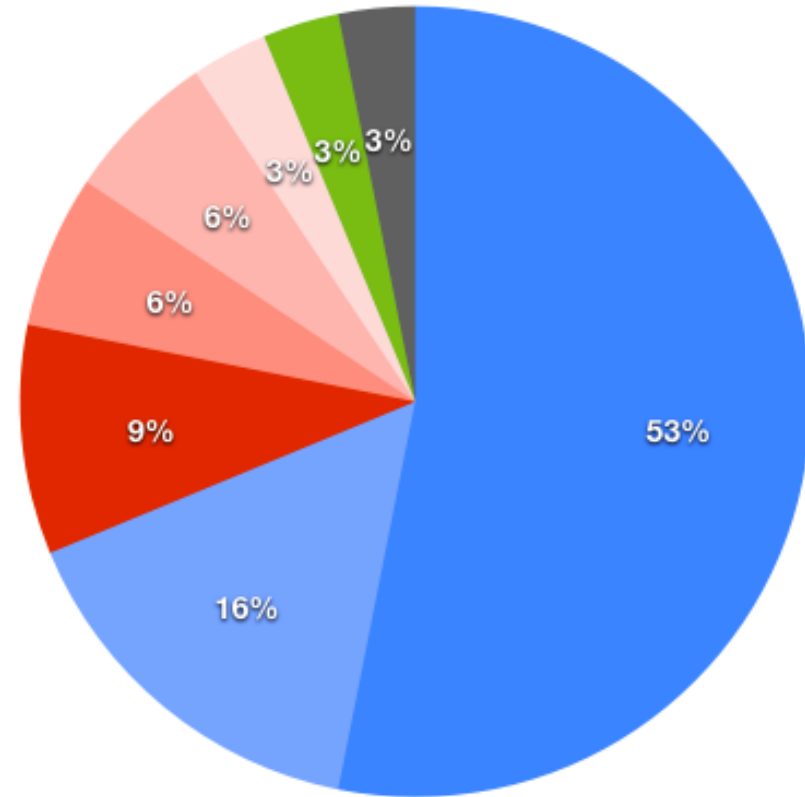
Different Types of Drugs

Drugs entering late stage development in 2011



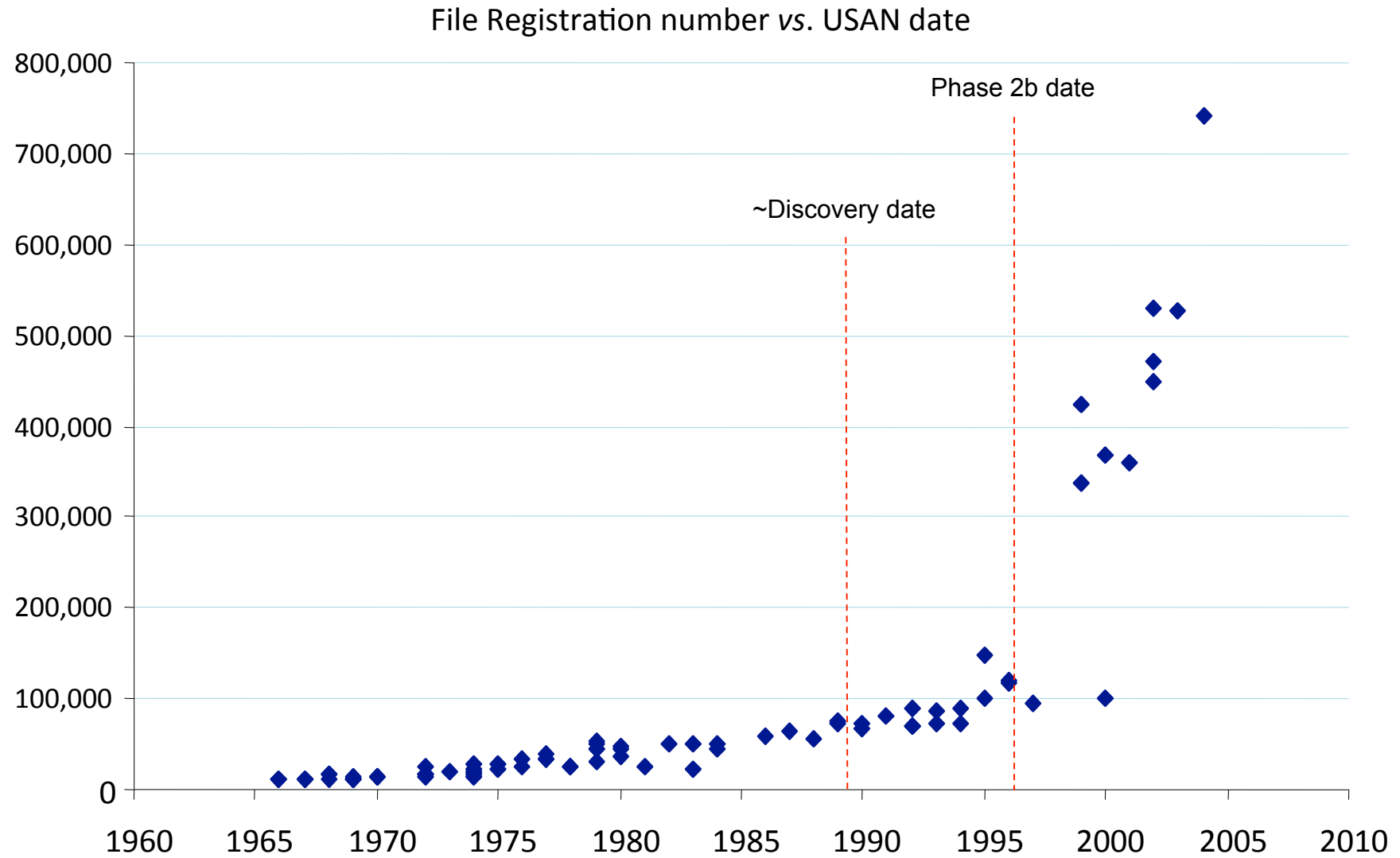
- Synthetic Small molecule
- mab
- Protein
- Cell
- Natural Product derived
- Enzyme
- Peptide
- Oligonucleotide

Drugs launched in US in 2011



- Synthetic Small molecule
- mab
- Protein
- Cell
- Inorganic
- Natural Product derived
- Enzyme
- Peptide
- Oligonucleotide

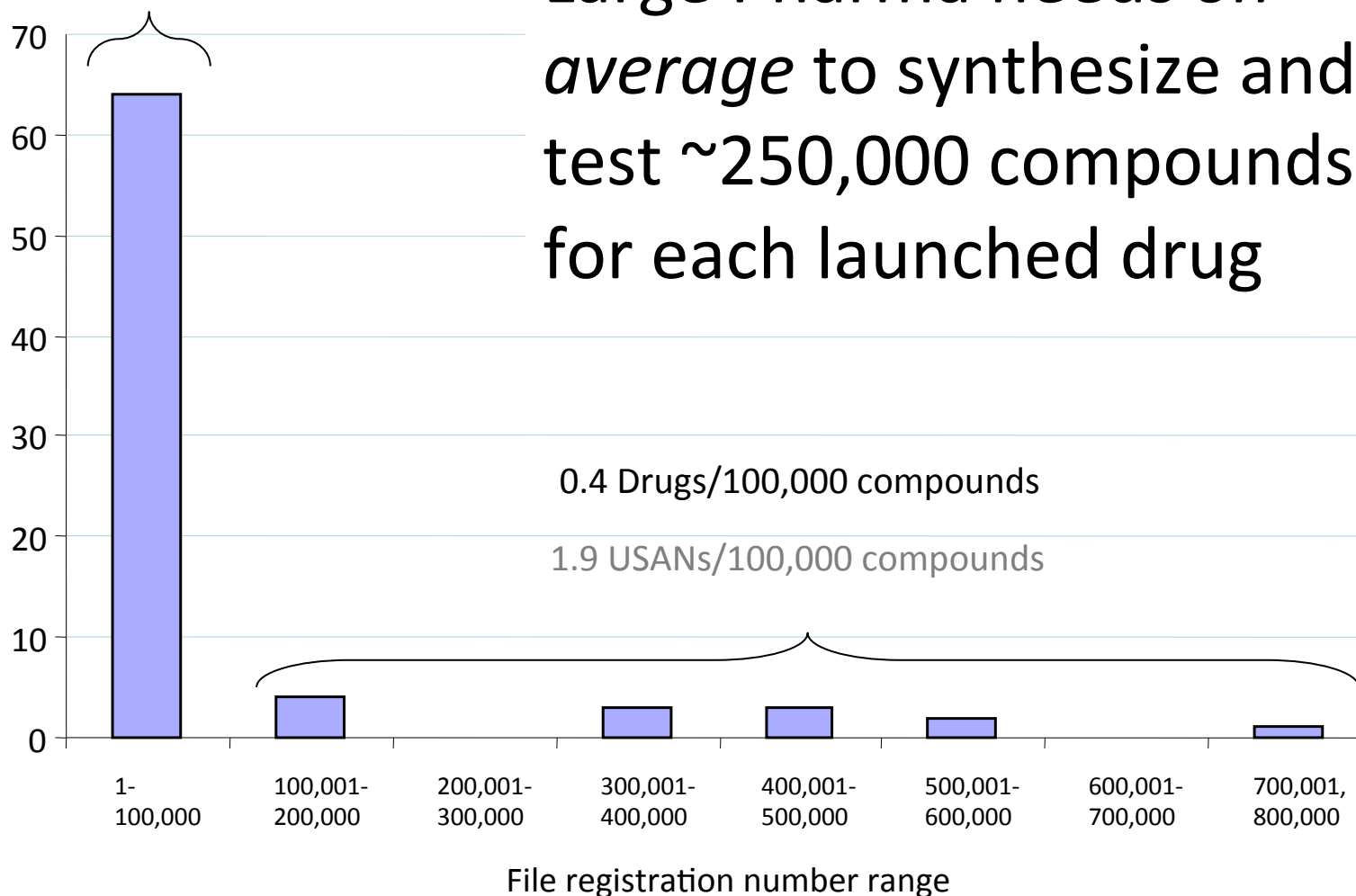
Pharma Industry Productivity



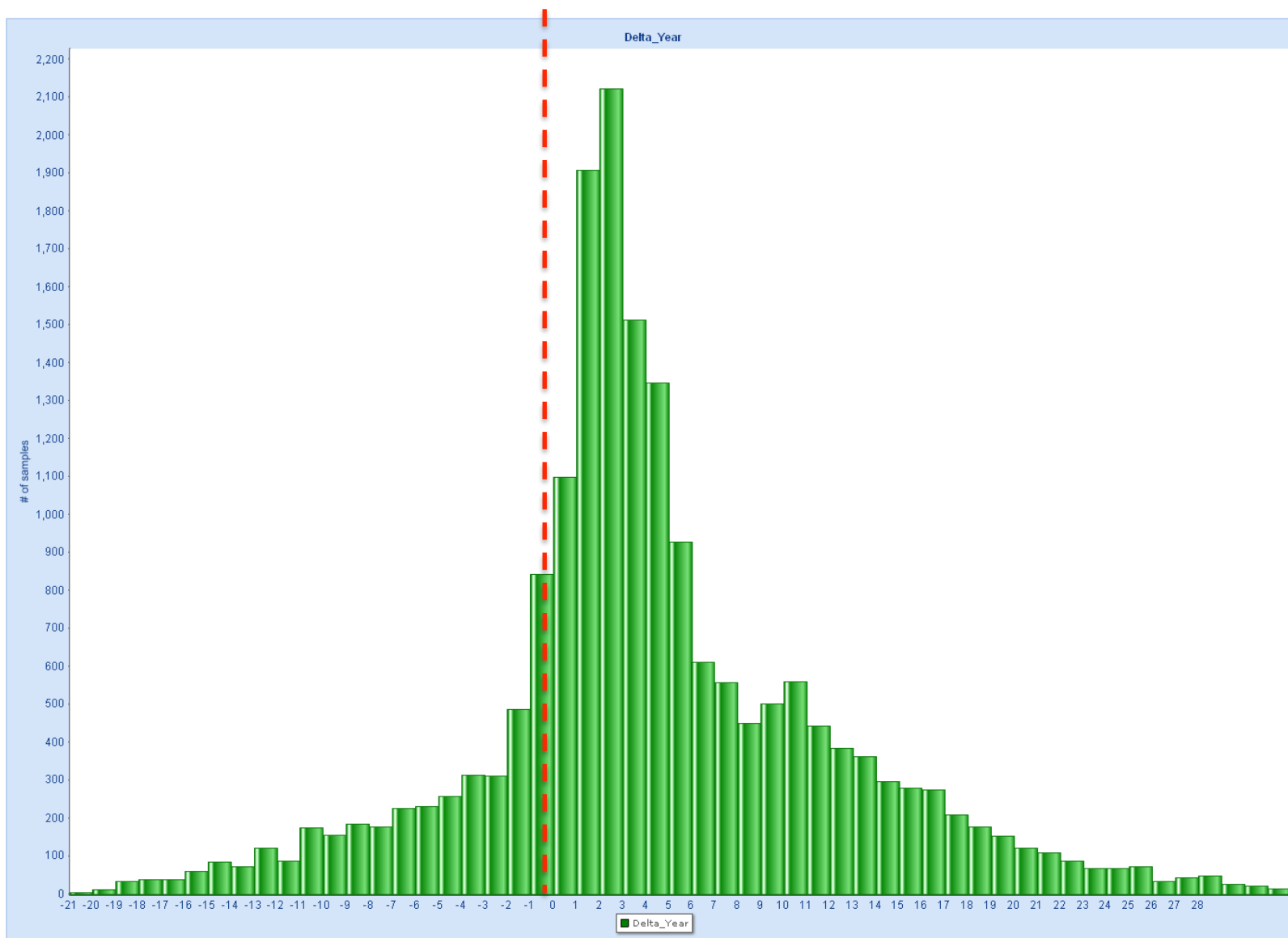
Pharma Industry Productivity

16 Drugs/100,000 compounds

64 USANs/100,000 compounds

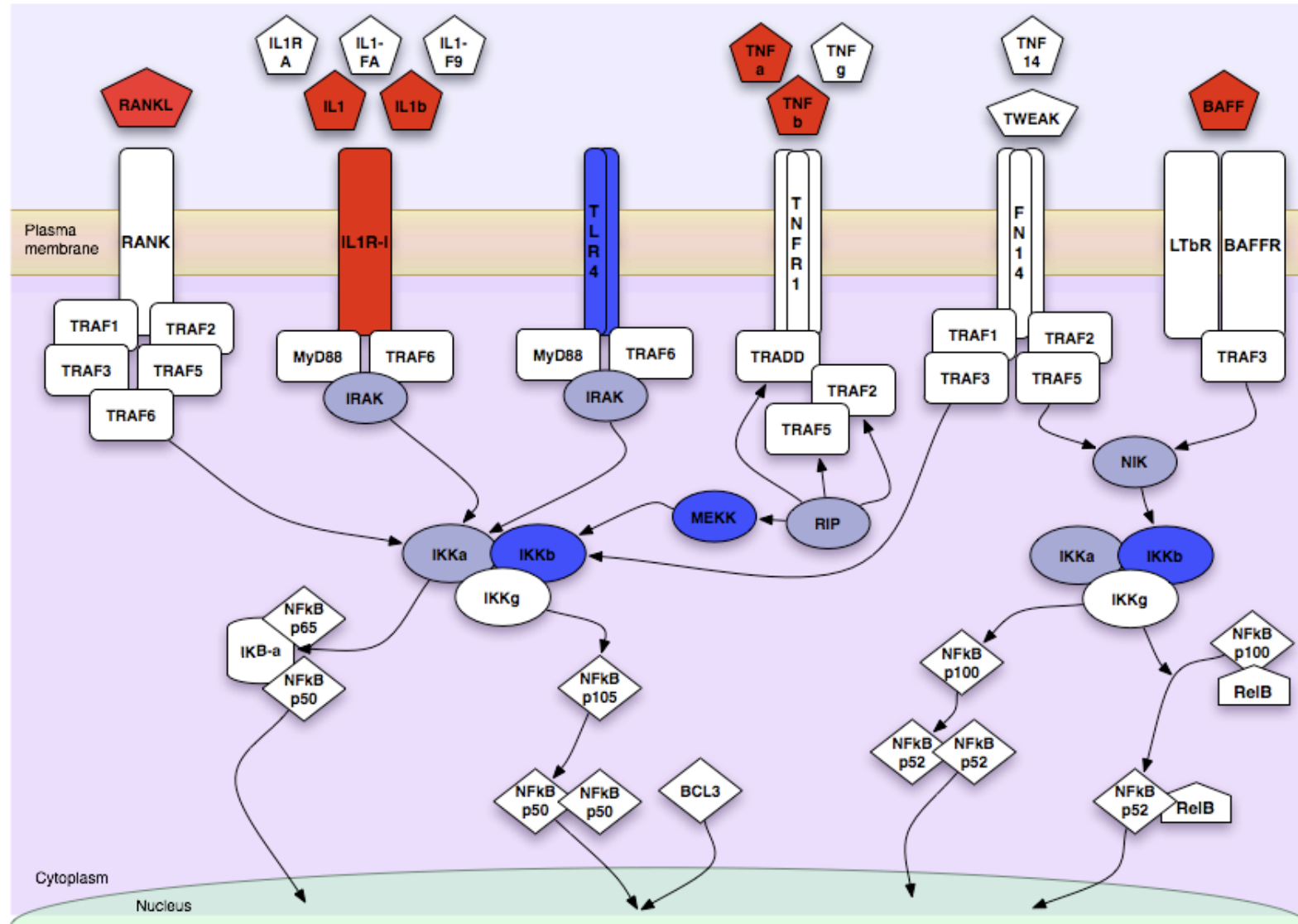


Patent and Publication Lag



IBM Patent data and ChEMBL

Clinical Candidates



What Is the ChEMBL Data?

2432

J. Med. Chem. 2002, 45, 2432–2453

Design of Selective Thrombin Inhibitors Based on the (R)-Phe-Pro-Arg Sequence

John C. Danilewicz,^{1,2} Stuart M. Abel,³ Alan D. Brown,² Paul V. Fish,^{4,5} Edward Hawkeswood,³ Stephen J. Holland,³ Keith James,³ Andrew B. McElroy,³ John Overington,³ Michael J. Powling,⁶ and David J. Rance³

Departments of Discovery Chemistry, Drug Metabolism, Discovery Biology, and Molecular Informatics Structure and Design, Pfizer Global Research and Development, Sandwich, Kent CT13 9NJ, United Kingdom

Received December 21, 2001

Potent and selective inhibitors of thrombin were sought based on the (R)-Phe-Pro-Arg sequence. The objective was to generate similar binding interactions to those achieved by potent competitive inhibitors of the argatroban type, so eliminating the need for covalent interaction with the catalytic serine function, as utilized by aldehyde and boronic acid type inhibitors. Improving the S₁ subsite interaction by substitution of arginine with a 4-alkoxybenzamidine residue provided potent lead 2 (*K*_i = 0.37 nM). Though an amide bond, which H-bonds to the active site, is lost, modeling indicated that a new H-bond is generated between the alkoxy oxygen atom and the catalytic Ser-195 hydroxyl group. Substitution of the benzamidine system by 1-amidinopiperidine then gave compound 4, which provided a further gain in selectivity over trypsin. However, previous work had shown that these compounds were likely to be too lipophilic (Log *D* +0.4 and +0.2, respectively) and to suffer rapid hepatic extraction, presumably via biliary elimination. Accordingly, both proved short-acting when administered intravenously to rats and showed poor activity when given intraduodenally. The aim was then to reduce lipophilicity below a log *D* of -1.2, which in a previously reported series had been effective in preventing rapid clearance. It was anticipated that compounds of this type would rely on the cation selective paracellular route of absorption from the gastrointestinal tract. Potent polar analogues with selectivity >1000 over trypsin were obtained. The best *in vivo* activity was shown by compound 12. However, in the final analysis, its oral bioavailability proved poor, relative to analogues with similar physicochemical properties derived from argatroban, consistent with the hypothesis that molecular shape is an additional important determinant of paracellular absorption.

Introduction

The search for potent selective and orally active thrombin inhibitors has gathered momentum in recent years.¹ Thrombin is the last in a cascade of trypsin-like plasma serine proteases, which by catalyzing the conversion of fibrinogen to fibrin, activation of FXIII and inducing platelet aggregation is a key enzyme in haemostasis and thrombus formation. The inhibition of a single enzyme in the cascade, and in particular thrombin, has been an attractive goal in that it could also provide superior antithrombotic therapy by increasing efficacy and safety as compared to heparin and the coumarins. Additionally, by keeping molecular size small, the opportunity exists for obtaining oral activity.^{1,2}

Two small molecular weight inhibitor types are emerging as structure-activity relationships are explored. The first is of the argatroban³ and NAPAP⁴ type (Chart 1), where lipophilic groups on either side of the

basic P₁ side chain pack together to interact with the hydrophobic S₂ site.^{5–7} Napsagatran (Ro 46-6240), developed by Hilpert et al.,⁸ though having a more complex P₁ residue, can nevertheless be viewed as belonging to this group. The only interaction with the catalytic serine residue is via a hydrogen bond to the carboxylate function in both argatroban and napsagatran. Unfortunately, none of these compounds is orally active due to either poor absorption from the gastrointestinal tract and/or rapid clearance via the bile.^{4,9}

A second inhibitor type is based on the substrate-derived irreversible chloromethyl ketone inhibitor PPACK and includes compounds such as DuP-714^{10,11} and efegatran (GYKI-14 766).¹² These compounds interact covalently with the hydroxyl group of the catalytic serine residue. The neighboring proline ring and (R)-Phe side chain cooperate to fill the S₂ site in a similar fashion as the two distal lipophilic groups of the first series.⁵ Though oral activity has been claimed for these compounds, we were concerned that high enzyme selectivity might not be obtainable when substantial affinity is derived by interacting covalently with the ubiquitous active site serine function. In the case of aldehyde type inhibitors, there is also the potential problem of achieving adequate optical and chemical stability.

* To whom correspondence should be addressed. Tel.: 44 1304 644589. Fax: 44 1304 656433. E-mail: paul_fish@sandwich.pfizer.com.

¹ Senior author.

² Department of Discovery Chemistry.

³ Department of Drug Metabolism.

⁴ Department of Discovery Biology.

⁵ Molecular Informatics Structure and Design.

What is the ChEMBL Data?

Design of Novel Thrombin Inhibitors

Fragment residues 55-65 of the compounds were generated by soaking crystals with mother liquor containing 1 mM inhibitor. Data were collected on a MAR Research imaging plate (X-ray Research, Hamburg, Germany) mounted on a Rigaku R2000 rotating anode generator and processed and scaled with HKL2000. Molecular models were generated and refined with SHELX and CNS. The X-ray structures of compounds 2 and 24 were deposited in the Brookhaven data bank. The accession numbers are HTT and HTS, respectively.

Measurement of Thrombin Inhibition: The thrombin inhibitory effects (IC₅₀) of the compounds were determined with a commercially available chromogenic assay (Boehr. Mannheim, Germany). Human thrombin (Boehr. 0.042 IU/ml) was preincubated for 10 min at 37 °C with 10 different dilutions (concentration range of 0.002–100 μM) of the test compounds dissolved in DMSO or with DMSO as control. Upon addition of the prothrombin mixture to the chromogenic substrate, total glycolipid arginyl-4-nitroanilic acetate, its reaction is cleaved by thrombin and the increase in absorbance at 405 nm, related to the free nitroaniline, is measured in a spectrophotometer (Spectromax, Molecular Devices, Sunnyvale, CA). By plotting the absorbance at 405 nm vs the concentration of the test compound, the concentration that induced a 50% thrombin inhibition (IC₅₀) was calculated. All measurements were performed in duplicate, and the mean values of both determinations are represented.

Measurement of the aPTT: aPTT was measured in a coagulometer (Biotest, B10, Seefeld, Germany) using the PTT reagent of Biotest, Mannheim, Germany, according to the manufacturer's instructions as a measure for the anticoagulant effect of the respective compound. Blood samples were collected in sodium citrate solution (final concentration: 0.11%) in vacuo. Each test sample (0.1 ml) was pipetted into test tubes prewarmed to 37 °C. The PTT reagent (0.1 ml) was added, mixed, and incubated for exactly 3 min. Calcium chloride solution (0.1 ml), prewarmed to 37 °C, was added to initiate the coagulation cascade, and the time (aPTT) in seconds was determined that elapsed from the addition of calcium chloride to the onset of clotting.

Acknowledgment: We thank Ingrid Christ, Josef Elband, Herbert Fluchbach, Monika Kock-Elband, Vera Koch, Michael Koshler, Angela Schmidt, Johannes Schare, and Elisabeth Waldmann for their skilful technical assistance.

Supporting Information Available: X-ray crystallographic data of compounds 2 and 24 with human thrombin. Data collection and refinement statistics. Combination analysis of all new compounds. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Journal of Medicinal Chemistry, 2002, Vol. 45, No. 9, 1765

1 Wang, J. L.; Hahn, J. *New Anticoagulant Drugs*, 2001, 107-108.

2 Hahn, J.; Fuster, V. *Guide to Anticoagulant Therapy*; Part 1: Synthetic. *Circulation* 1994, 90, 1449-1469. **3** Hahn, J.; Fuster, V. *Guide to Anticoagulant Therapy*, Part 2: Oral Anticoagulants. *Circulation* 1994, 90, 1469-1486.

4 Eriksson, B. J.; Ehren, S.; Lefström, S.; Nair, M.; Rank, S.; Schmidt, C.; Schuler, P.; Cimo, P. *Preparation of a thrombin inhibitor 2*. *J. Org. Chem.* 1997, 62, 1079-1082.

5 Hahn, J.; Fuster, V.; Rank, S.; Nair, M.; Rank, S.; Lefström, S.; Cimo, P. *A comparison of recombinant fibrin with molecular weight heparin to prevent thrombotic complications after intracoronary stent implantation*. *Am. J. Cardiol.* 1997, 79, 1029-1033.

6 Wang, J. L.; Fuster, V. *Stent thrombosis: direct thrombin inhibition*. *Expert Opin. Ther. Targets* 2002, 2, 225-236.

7 Hahn, J.; Fuster, V.; Rank, S.; Nair, M.; Rank, S.; Lefström, S.; Cimo, P. *Comparison of recombinant fibrin with molecular weight heparin to prevent thrombotic complications after intracoronary stent implantation*. *Am. J. Cardiol.* 1997, 79, 1029-1033.

Design of Novel Thrombin Inhibitors

Fragment residues 55-65 of the compounds were generated by soaking crystals with mother liquor containing 1 mM inhibitor. Data were collected on a MAR Research imaging plate (X-ray Research, Hamburg, Germany) mounted on a Rigaku R2000 rotating anode generator and processed and scaled with HKL2000. Molecular models were generated and refined with SHELX and CNS. The X-ray structures of compounds 2 and 24 were deposited in the Brookhaven data bank. The accession numbers are HTT and HTS, respectively.

Measurement of Thrombin Inhibition: The thrombin inhibitory effects (IC₅₀) of the compounds were determined with a commercially available chromogenic assay (Boehr. Mannheim, Germany). Human thrombin (Boehr. 0.042 IU/ml) was preincubated for 10 min at 37 °C with 10 different dilutions (concentration range of 0.002–100 μM) of the test compounds dissolved in DMSO or with DMSO as control. Upon addition of the prothrombin mixture to the chromogenic substrate, total glycolipid arginyl-4-nitroanilic acetate, its reaction is cleaved by thrombin and the increase in absorbance at 405 nm, related to the free nitroaniline, is measured in a spectrophotometer (Spectromax, Molecular Devices, Sunnyvale, CA). By plotting the absorbance at 405 nm vs the concentration of the test compound, the concentration that induced a 50% thrombin inhibition (IC₅₀) was calculated. All measurements were performed in duplicate, and the mean values of both determinations are represented.

Measurement of the aPTT: aPTT was measured in a coagulometer (Biotest, B10, Seefeld, Germany) using the PTT reagent of Biotest, Mannheim, Germany, according to the manufacturer's instructions as a measure for the anticoagulant effect of the respective compound. Blood samples were collected in sodium citrate solution (final concentration: 0.11%) in vacuo. Each test sample (0.1 ml) was pipetted into test tubes prewarmed to 37 °C. The PTT reagent (0.1 ml) was added, mixed, and incubated for exactly 3 min. Calcium chloride solution (0.1 ml), prewarmed to 37 °C, was added to initiate the coagulation cascade, and the time (aPTT) in seconds was determined that elapsed from the addition of calcium chloride to the onset of clotting.

Acknowledgment: We thank Ingrid Christ, Josef Elband, Herbert Fluchbach, Monika Kock-Elband, Vera Koch, Michael Koshler, Angela Schmidt, Johannes Schare, and Elisabeth Waldmann for their skilful technical assistance.

Supporting Information Available: X-ray crystallographic data of compounds 2 and 24 with human thrombin. Data collection and refinement statistics. Combination analysis of all new compounds. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Journal of Medicinal Chemistry, 2002, Vol. 45, No. 9, 1765

1 Wang, J. L.; Hahn, J. *New Anticoagulant Drugs*, 2001, 107-108.

2 Hahn, J.; Fuster, V. *Guide to Anticoagulant Therapy*; Part 1: Synthetic. *Circulation* 1994, 90, 1449-1469. **3** Hahn, J.; Fuster, V. *Guide to Anticoagulant Therapy*, Part 2: Oral Anticoagulants. *Circulation* 1994, 90, 1469-1486.

4 Eriksson, B. J.; Ehren, S.; Lefström, S.; Nair, M.; Rank, S.; Schmidt, C.; Schuler, P.; Cimo, P. *Preparation of a thrombin inhibitor 2*. *J. Org. Chem.* 1997, 62, 1079-1082.

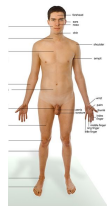
5 Hahn, J.; Fuster, V.; Rank, S.; Nair, M.; Rank, S.; Lefström, S.; Cimo, P. *A comparison of recombinant fibrin with molecular weight heparin to prevent thrombotic complications after intracoronary stent implantation*. *Am. J. Cardiol.* 1997, 79, 1029-1033.

6 Wang, J. L.; Fuster, V. *Stent thrombosis: direct thrombin inhibition*. *Expert Opin. Ther. Targets* 2002, 2, 225-236.

7 Hahn, J.; Fuster, V.; Rank, S.; Nair, M.; Rank, S.; Lefström, S.; Cimo, P. *Comparison of recombinant fibrin with molecular weight heparin to prevent thrombotic complications after intracoronary stent implantation*. *Am. J. Cardiol.* 1997, 79, 1029-1033.

>Thrombin

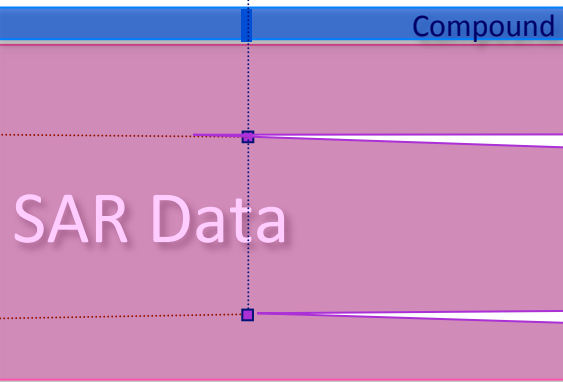
MAHVRGLQLPGLALALCSLVSQHVFLPAPQSSQLQVRNRAFLLEVRKGNLRECEVETCS
YEEAFALSSATDVFVWYKATACETARTPKDKLAACLEGACGGSTNYGHVNIITRSGIEQLGL
RSLRPFKPEINSTFEGALGEMCRNFSSITDFPFCVQSGVQDFVYRPF
SGSSVNSLPLDCCPDRGQVQGLAVTTHGLPCLAWASARALSKHDFNSAVQLVYENCRN
PQDDEGVGWCYKAGPQDFGCDLVNCEAVEEETDGLDESDRAIEGRATSEYQSRVNRVAT
SGEDAGCLRLPEFKKLEDKTERELLESYIDGRIVEGSDAFGMSFWQVMLFRKSPQELCGSL
SDRWVLAHGLIYPPWDFNPTENDLVLGICXHSRATYRLEKISMLKLIYTHPRVNRNDRD
IALMKLKKVFAFSEYHPCVLDKFSALSLTQAGYGRVYKQGNLSEYKLVANVKGQVYKQVW
PIYRPFVCKDSTRITIDMPCAGYRQDEKRGDACEGSGGDFVFKSPFNINRYQMGIVSNGEGC
DRDQKGYPTVHFLRKLQIWKVLDQEG



PTT (partial thromboplastin time)

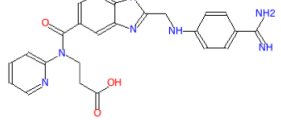
Inhibition of human Thrombin

Assay



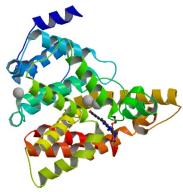
K_i=4.5 nM

ED₂=230 nM



ChEMBL Target Types

Protein



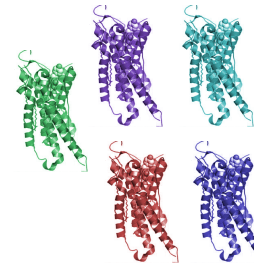
e.g. PDE5

Protein complex



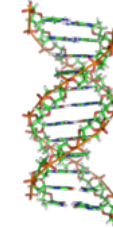
e.g. Nicotinic acetylcholine receptor

Protein family



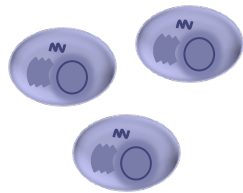
e.g. Muscarinic receptors

Nucleic Acid



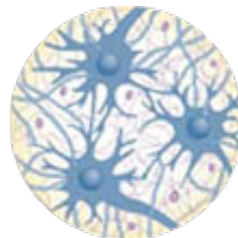
e.g. DNA

Cell line



e.g. HEK293 cells

Tissue



e.g. Trachea

Sub-cellular fraction



e.g. Mitochondria

Organism



e.g. Drosophila

Compound Searching

ChEMBLdb

Search ChEMBLdb... **Compounds** Targets Assays [Activity Source Filter](#)

ChEMBLdb **Compound Search** Protein Target Search Browse Targets Browse Drugs Drug Approvals

File Edit View Insert Atom Bond Structure Calculations Tools Help

100%

List Search

SMILES Search ChEMBL ID Search Keyword Search

Please enter a list of Compound IDs, keywords, or SMILES separated by newlines

Fetch Compounds

ChEMBLdb Statistics

- DB: ChEMBL_13
- Targets: 8,845
- Compound records: 1,296,266
- Distinct compounds: 1,143,682
- Activities: 6,933,068
- Publications: 44,682

ChEMBL Blog

- [New Drug Approvals 2012 - Pt. XII- pertuzumab \(Perjeta™\)](#)
- [Drug Side Effect Prediction and Validation](#)

Compound Sketcher: Please select....

Substructure Search 100% Fetch Compounds

Terms of Use | Privacy | Cookies | EBI Funding | Contact EBI | © European Bioinformatics Institute 2012. EBI is an Outstation of the European Molecular Biology Laboratory.

Spreadsheet Views

ChEMBLdb

https://www.ebi.ac.uk/chembl/index.php/compound/results/1/chemblid/asc/tab/substructure

Can't Update...ommunities ChEMBL ChEMBL-og Twitter Google+ Citations Analytics Facebook LinkedIn F1000 Wikipedia PubMed BioMedBridges Science (5) EBI Computing

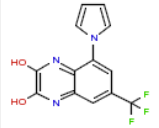
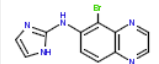
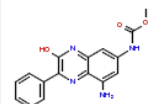
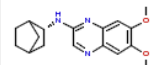
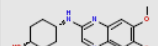
EMBL-EBI Enter Text Here Find Terms of Use Privacy Cookies

Databases Tools Research Training Industry About Us Help Site Index

EBI > Databases > Small Molecules > ChEMBL Database > Compound Search > Results

ChEMBL Compound Search Results: 5055 Hits

1 2 3 4 5 6 [Next] [End] Please select....

<input checked="" type="checkbox"/>	Compound	Synonyms	Parent Mol Weight	ALogP	PSA	HBA	HBD	#Ro5 Vio.	#Rotatable Bonds	Passes Rule of Three	Med Chem Friendly	ACD_APKA	ACD_BPKA	ACD_LOGP	ACD_LOGS
<input checked="" type="checkbox"/>	 ChEMBL100999		295.2	3.71	71.2	4	2	0	2	No	Yes	5.415		3.773	1.632
<input checked="" type="checkbox"/>	 ChEMBL1019		290.1	1.38	66.5	4	2	0	2	No	Yes	13.218	6.348	1.104	1.029
<input checked="" type="checkbox"/>	 ChEMBL103363		310.3	2.56	110.4	6	3	0	3	No	Yes	12.431	8.74	3.431	2.896
<input checked="" type="checkbox"/>	Chiral  ChEMBL104011		299.4	2.74	56.3	5	1	0	4	No	Yes		2.86	3.757	3.757
<input checked="" type="checkbox"/>	Chiral 		303.4	1.76	76.5	6	2	0	4	No	Yes		2.86	1.999	1.999

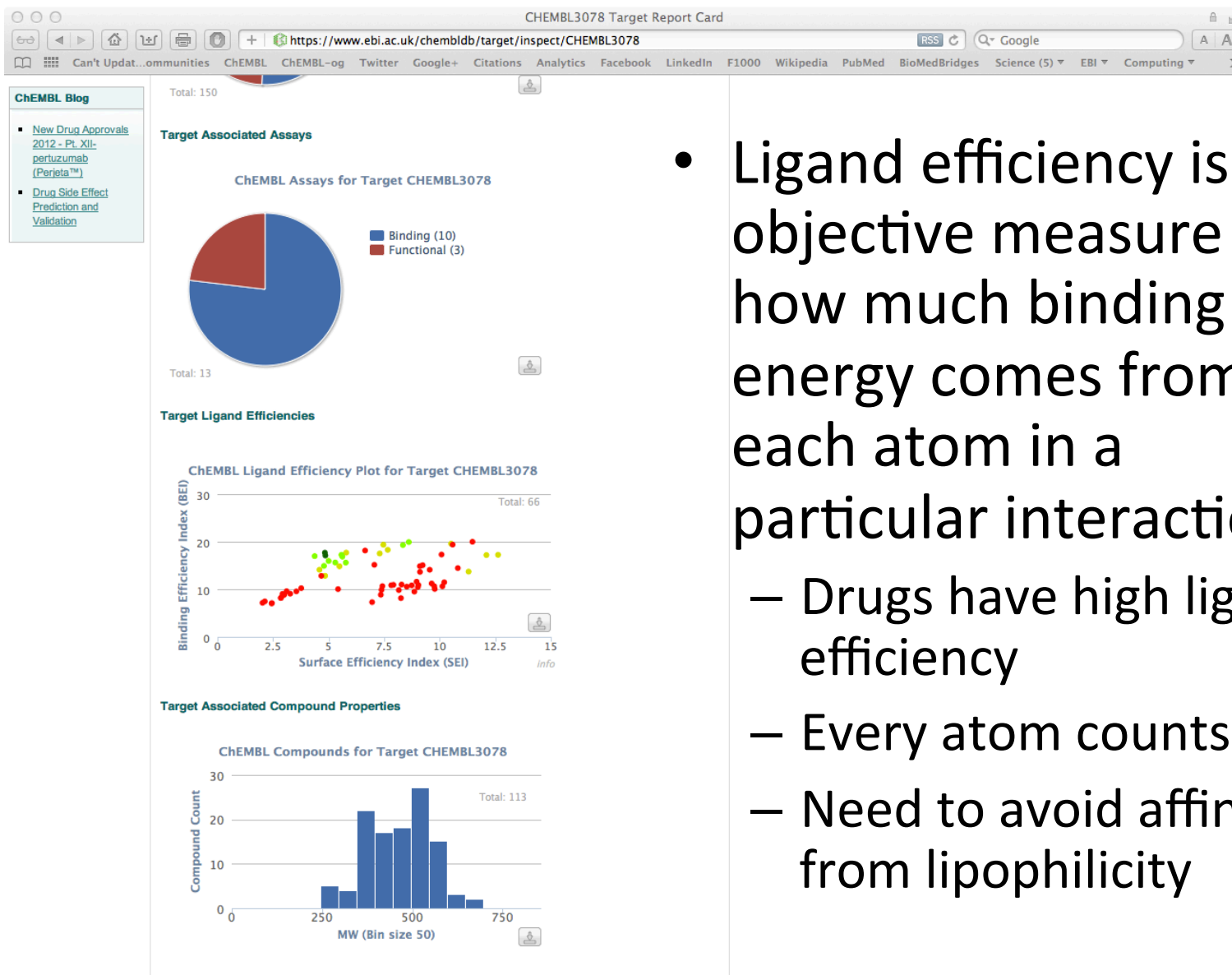
ChEMBLdb Statistics

- DB: ChEMBL_13
- Targets: 8,845
- Compound records: 1,296,266
- Distinct compounds: 1,143,682
- Activities: 6,933,068
- Publications: 44,682

ChEMBL Blog

- New Drug Approvals 2012 - Pt. XII: pertuzumab (Perjeta™)
- Drug Side Effect Prediction and Validation

Ligand Efficiency



- Ligand efficiency is an objective measure of how much binding energy comes from each atom in a particular interaction
 - Drugs have high ligand efficiency
 - Every atom counts
 - Need to avoid affinity from lipophilicity

Target Class Data

ChEMBL

http://www.ebi.ac.uk/chembl/index.php/target/browser/classification

Analytics ChEMBL Twitter LinkedIn Facebook Wikipedia Import to Mendeley Flickr! Doodle jpo iDisk ChEMBL EBI Google Science

EMBL-EBI EB-eye Search All Databases Enter Text Here Go Reset Advanced Search Give us feedback

Databases Tools EBI Groups Training Industry About Us Help Site Index

ChEMBL

ChEMBLdb

- Search Compounds
- Search Targets

ChEMBL-NTD

Kinase SARfari

ChEMBL Group

Downloads

FAQ

ChEMBLdb Statistics

- Targets: 8,054
- Compound records: 726,872
- Distinct compounds: 600,625
- Activities: 2,925,588
- Publications: 38,029
- DB: ChEMBL_06

ChEMBL Blog

- Lecture: Development and Applications of Computational Chemogenomics
- Books: Germ Stories

EBI > Databases > Small Molecules > ChEMBL Database > Target Search > Target Classification Hierarchy

Search ChEMBLdb... Compounds Targets Assays

ChEMBLdb Compound Search Protein Target Search **Browse Targets**

Browse Protein Target Tree Taxonomy Tree

Click arrows to navigate tree

- Enzyme (2443)
- Membrane receptor (554)
- Ion channel (338)
- Transporter (136)
- Transcription Factor (103)
- Cytosolic other (102)
- Secreted (58)
- Structural (29)
- Surface antigen (26)
- Membrane other (16)
- Adhesion (14)
- Nuclear other (13)

Target Class	Count
Enzyme	2443
Membrane receptor	554
Ion channel	338
Transporter	136
Transcription Factor	103
Cytosolic other	102
Secreted	58
Structural	29
Surface antigen	26
Membrane other	16
Adhesion	14
Nuclear other	13

Legend:

- Enzyme
- Membrane receptor
- Ion channel
- Transporter
- Transcription Factor
- Cytosolic other
- Secreted
- Structural
- Surface antigen
- Membrane other
- Adhesion
- Nuclear other

Assay Organism Data

ChEMBL

http://www.ebi.ac.uk/chembl/index.php/target/browser/organism/

Analytics ChEMBL Log Twitter LinkedIn Facebook Wikipedia Import to Mendeley Flickr! Doodle jpo iDisk ChEMBL EBI Google Science

EMBL-EBI EB-eye Search All Databases Enter Text Here Go Reset Advanced Search Give us feedback

Databases Tools EBI Groups Training Industry About Us Help Site Index

EBI > Databases > Small Molecules > ChEMBL Database > Target Search > Target Organism Hierarchy

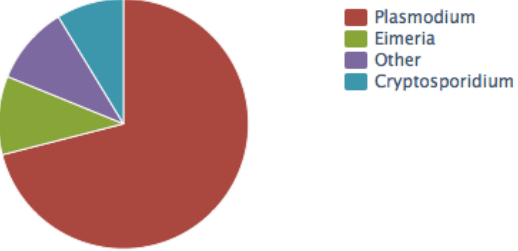
Search ChEMBLdb... Compounds Targets Assays

ChEMBLdb Compound Search Protein Target Search **Browse Targets**

Browse Protein Target Tree Taxonomy Tree

Click arrows to navigate tree

- Eukaryotes (5448)
 - Mammalia (5076)
 - Apicomplexa (69)
 - Plasmodium (49)
 - Eimeria (7)
 - Other (7)
 - Cryptosporidium (6)
 - Viridiplantae (68)
 - Kinetoplastida (57)
 - Arthropoda (52)
 - Aves (35)
 - Nematoda (26)
 - Eukaryotes (other) (19)
 - Teleostei (13)
 - Amphibia (12)
 - Lepidosauria (11)
 - Platyhelminthes (7)
 - Echinodermata (3)
- Bacteria (614)
- Unclassified (561)
- Fungi (229)
- Viruses (211)
- Archaea (6)



Legend:

- Plasmodium
- Eimeria
- Other
- Cryptosporidium

ChEMBLdb Statistics

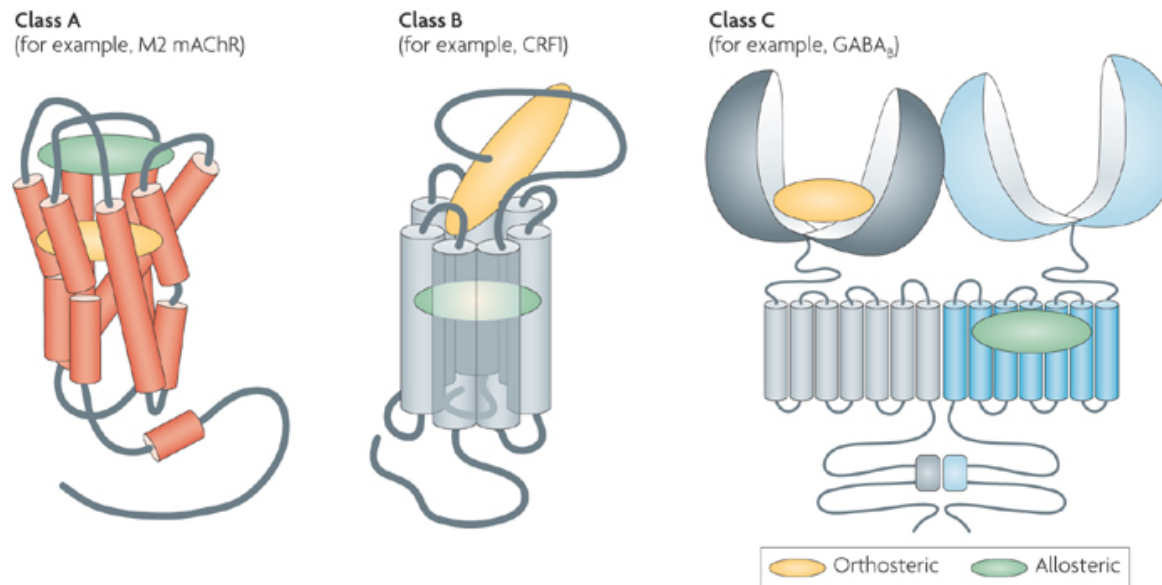
- Targets: 8,054
- Compound records: 726,872
- Distinct compounds: 600,625
- Activities: 2,925,588
- Publications: 38,029
- DB: ChEMBL_06

ChEMBL Blog

- Lecture: Development and Applications of Computational Chemogenomics
- Books: Germ Stories

Allosteric Regulators

- Allosteric drugs can have some advantages over orthosteric drugs
 - Selectivity
 - Orthosteric site may be undruggable



Allosteric/Orthosteric sites for GPCRs

http://www.chemblog.org

The ChEMBL-og - Open Data For Drug Discovery: Query Privacy in ChEMBL

chembl.blogspot.de/2012/09/query-privacy-in-chembl.html

The Great Tre... | Visual.ly BCS - The Ch... itute for IT Home: Path...on Database LINCS - Home John P. Overington Cambridge t...bume.co.uk MTI: Home

john.p.overington@gmail.com New Post Design Sign Out

The ChEMBL-og - Open Data For Drug Discovery

The news, progress, whereabouts, and ephemera from the Computational Chemical Biology group at the EMBL-EBI.

ChEMBL database ChEMBL-NTD GPCR SARfari Kinase SARfari DrugEBility ADMET SARfari

+1 13

ABOUT CHEMBL

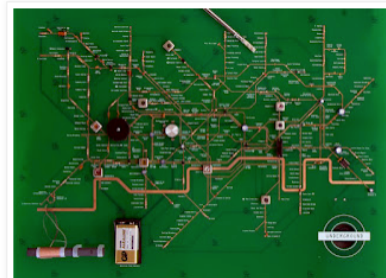


The ChEMBL-og covers the activities of the [Computational Chemical Biology Group](#) at the EMBL-EBI in Hinxton. Our interests cover **Drug Discovery, Computational Chemical Biology, Chemogenomics, Chemoinformatics, Bioinformatics, Structural Biology, Pharmacogenomics, Open Data, Knowledge Management, Semantic Web, and Data Integration** and include...

- [ChEMBLdb](#) - a drug discovery SAR and bioactivity database.
- [SARfari](#) - a sequence, binding site, structure, SAR integration platform.
- [DrugEBility](#) - Drug target annotation & prioritisation.

SATURDAY, 15 SEPTEMBER 2012

Query Privacy in ChEMBL



We have been asked several times for all the user-generated queries of [ChEMBL](#) - *i.e.* the structures sketched in to the interface that are then searched against the database. We will not (and in fact, physically can't) share these. Sorry. It is against both our institutional privacy policy, and standard Terms of Use, and also we've engineered the app to avoid us 'storing' any of this information where at all possible (*e.g.* in avoiding `/tmp` type fluff, minimizing residency time in caches, etc.).

There are clearly some advantages in pooling or analysing website search data - it highlights interesting trends, something becoming more interesting to a user community can spot emerging events, etc. It can alert to flu outbreaks (there was a Science paper from google on this, don't have the reference handy though - you may be able to find it with [google](#) though....). There is a huge interest in many sites that I use in tracking and analysing query terms and usage patterns, and in some contexts this is just the thing to do - like when [ebay](#) teases me (and surely of all the tortured obsessive souls on the planet, it is just me and me alone) with a rare phosphor or perforation machin variant I don't have.

The types of query that people perform can clearly also be used to develop ways of improving a website, or specifically the