

*Ferramentas e bases de dados
online para desenho de fármacos*

- Portais: locais de acesso a recursos de vários tipos
- Ferramentas on-line: conversão de formatos, cálculo de propriedades moleculares, visualização, docking,...
- Bases de dados: contêm estruturas moleculares de macromoléculas e moléculas pequenas que podem ser pesquisadas de múltiplas formas
- Podem ser serviços de acesso livre, ou sites comerciais com custos de utilização

Vantagens da utilização dos serviços online

- Disponíveis em qualquer local
- Custos de manutenção reduzidos
- Custos de licenciamento reduzidos
- Integração de diferentes tipos de software
- Fácil monitorização da utilização
- Computação em *cloud*
- Compatibilidade com múltiplos plataformas informáticas (Win, Mac, Linux, Android, etc)

O portal Click2Drug

- <http://www.click2drug.org>
- O portal faz parte do Swiss Institute of Bioinformatics
- Contem ~800 links divididos em categorias, incluindo diferentes tipos de software e bases de dados
- Cada link contem uma descrição resumida do serviço oferecido
- Está em permanente actualização

0 portal Click2Drug

Directory of in silico x

www.click2drug.org

Apps Enzymology Piano Music Production Bioinformatics Databases Bioinformatics T... Misc Programming D pmartel Other bookmarks

Click2Drug | SwissDock | SwissParam | SwissSidechain | SwissBioisostere | SwissTargetPrediction | About us

SIB Swiss Institute of Bioinformatics

Click2Drug

Directory | Bibliography | Encyclopedia | Citations | Contacts | Disclaimer

Directory of Tools

- Databases
- Chemical structure rep.
- Molecular modeling
- Homology modeling
- Binding site prediction
- Docking
- Screening
- Target prediction
- Ligand design
- Binding free energy estimation
- QSAR
- ADME Toxicity

Mobile applications

Last additions

Tag cloud

FAQ

Directory of computer-aided Drug Design tools

Click2Drug contains a comprehensive list of computer-aided drug design (CADD) software, databases and web services. These tools are classified according to their application field, trying to cover the whole drug design pipeline. If you think that an interesting tool is missing in this list, please contact us.

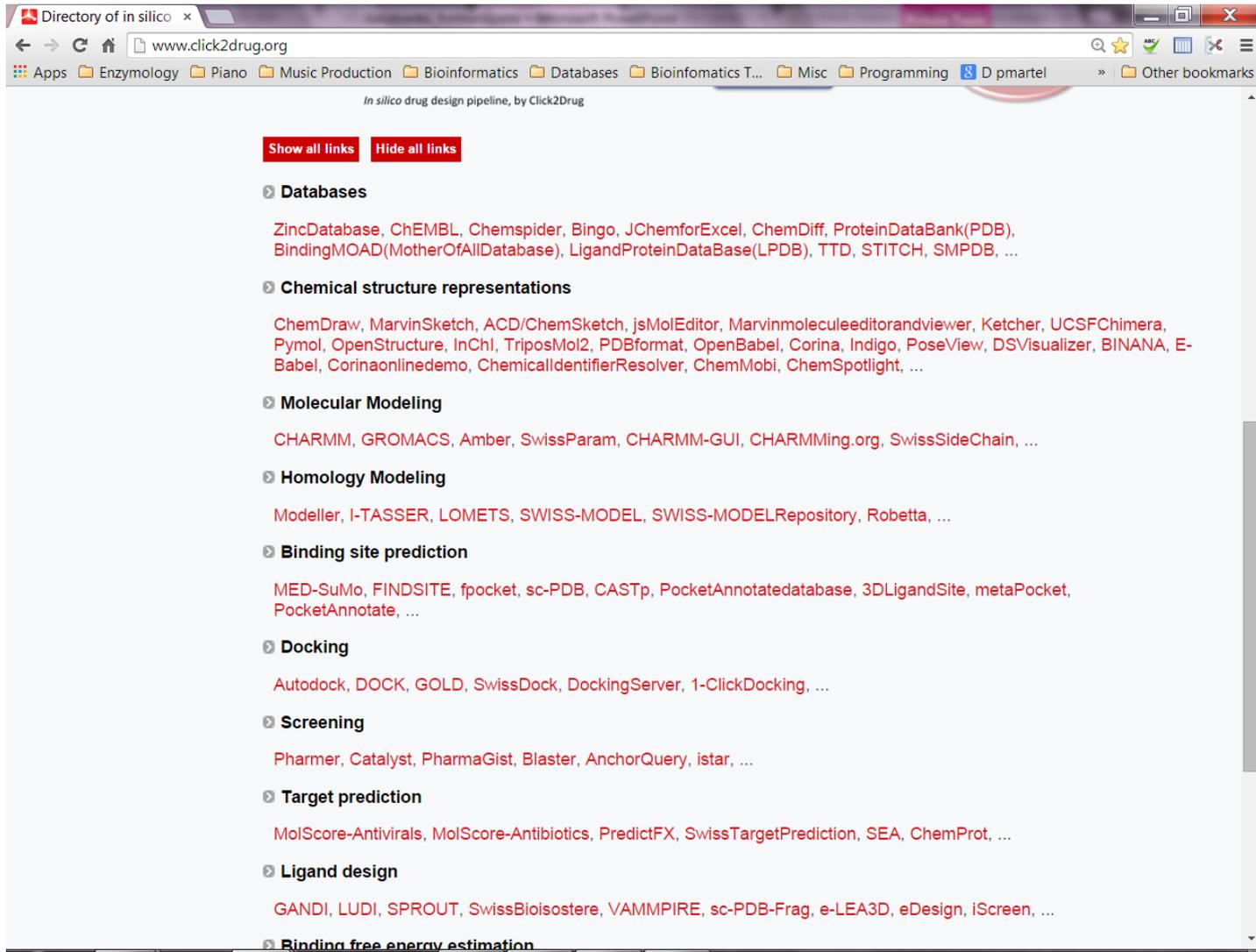
8+1 47 Updated on 7/18/2014. Currently 777 links. [Show all links](#) / [Hide all links](#).

Click on the following picture to select tools related to a given activity:

In silico drug design pipeline, by Click2Drug

[Show all links](#) [Hide all links](#)

Portal Click2Drug



Directory of in silico x
www.click2drug.org
Apps Enzymology Piano Music Production Bioinformatics Databases Bioinformatics T... Misc Programming D pmarcel Other bookmarks

In silico drug design pipeline, by Click2Drug

[Show all links](#) [Hide all links](#)

- Databases**
ZincDatabase, ChEMBL, Chempider, Bingo, JChemforExcel, ChemDiff, ProteinDataBank(PDB), BindingMOAD(MotherOfAllDatabase), LigandProteinDataBase(LPDB), TTD, STITCH, SMPDB, ...
- Chemical structure representations**
ChemDraw, MarvinSketch, ACD/ChemSketch, jsMolEditor, Marvinmoleculeeditorandviewer, Ketcher, UCSFChimera, Pymol, OpenStructure, InChI, TriposMol2, PDBformat, OpenBabel, Corina, Indigo, PoseView, DSVisualizer, BINANA, E-Babel, Corinaonlinedemo, ChemicalIdentifierResolver, ChemMobi, ChemSpotlight, ...
- Molecular Modeling**
CHARMM, GROMACS, Amber, SwissParam, CHARMM-GUI, CHARMMing.org, SwissSideChain, ...
- Homology Modeling**
Modeller, I-TASSER, LOMETS, SWISS-MODEL, SWISS-MODELRepository, Robetta, ...
- Binding site prediction**
MED-SuMo, FINDSITE, fpocket, sc-PDB, CASTp, PocketAnnotatedatabase, 3DLigandSite, metaPocket, PocketAnnotate, ...
- Docking**
Autodock, DOCK, GOLD, SwissDock, DockingServer, 1-ClickDocking, ...
- Screening**
Pharmer, Catalyst, PharmaGist, Blaster, AnchorQuery, istar, ...
- Target prediction**
MolScore-Antivirals, MolScore-Antibiotics, PredictFX, SwissTargetPrediction, SEA, ChemProt, ...
- Ligand design**
GANDI, LUDI, SPROUT, SwissBioisostere, VAMMPIRE, sc-PDB-Frag, e-LEA3D, eDesign, iScreen, ...
- Binding free energy estimation**

Portal Click2Drug

Directory of in silico x

www.click2drug.org

In silico drug design pipeline, by Click2Drug

Show all links Hide all links

▼ Databases

ZincDatabase, ChEMBL, ChempSpider, Bingo, JChemforExcel, ChemDiff, ProteinDataBank(PDB), BindingMOAD(MotherOfAllDatabase), LigandProteinDataBase(LPDB), TTD, STITCH, SMPDB, ...

Chemical databases

- **Zinc Database.** Curated collection of commercially available chemical compounds, with 3D coordinates, provided by the Shoichet Laboratory in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF).
- **ChEMBL.** Curated database of small molecules. Includes interactions and functional effects of small molecules binding to their macromolecular targets, and series of drug discovery databases.
- **ChempSpider.** Collection of chemical compounds maintained by the Royal Society of Chemistry. Includes the conversion of chemical names to chemical structures, the generation of SMILES and InChI strings as well as the prediction of many physicochemical parameters.
- **CoCoCo.** Free suite of multiconformational molecular databases for High-Throughput Virtual Screening. It has single and multi conformer databases prepared for HTVS in different formats like Phase, Catalyst, Unity and SDF. Provided by the Department of Pharmaceutical Sciences of the University of Modena and Reggio Emilia.
- **DrugBank.** Bioinformatics and cheminformatics resource combining detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. Allows searching for similar compounds.
- **PubChem.** Database of chemical compounds maintained by the National Center for Biotechnology Information (NCBI), along with bioassays results. Allows similar compounds search (2D and 3D).
- **PubChem Mobile.** Free application to search PubChem databases using chemical names, synonyms, and keywords. For Android.
- **TCM.** Free small molecular database on traditional Chinese medicine, for virtual screening. It is currently the world's largest TCM database, and contains 170'000 compounds, with 3D mol2 and 2D cdx files, which passed ADMET filters.
- **Mcule database.** Commercial database of commercially available small molecules. Allows filtering by chemical supplier data (stock availability, price, delivery time, chemical suppliers, catalogs, minimum purity, etc.) and export the whole Mcule database including supplier and procurement related properties. Reduced prices for academic. Provided by Mcule.
- **WOMBAT.** (World of Molecular Bioactivity). Database of 331,872 entries (268,246 unique SMILES), representing 1,966 unique targets, with bioactivity annotations. Compiled by Sunset Molecular Discovery LLC.
- **Approved Drugs.** The Approved Drugs app contains over a thousand chemical structures and names of small molecule drugs approved by the US Food & Drug Administration (FDA). Structures and names can be browsed in a list, searched by name, filtered by structural features, and ranked by similarity to a user-drawn structure. The detail view allows viewing of a 3D conformation as well as tautomers. Structures can be exported in a variety of ways, e.g. email, twitter, clipboard. For iPad and iPhone. Developed by Molecular Materials Informatics, Inc.
- **ChemSpider Mobile.** Allows searching the ChemSpider chemical database, provided by the Royal Society of Chemistry. Compounds can be searched by structure or by name, and browsed within the app. Results can be examined by jumping to the web page. Search structures are drawn using the powerful MMDS molecular diagram editor. For iPad. Provided by Molecular Materials Informatics, Inc.
- **e-Drug3D.** Database mirroring the current content of the U.S. pharmacopeia of small drugs. Contains 1533 molecular structures with a molecular weight < 2000 (last update: February 2012). Provides SD files (single conformer, tautomers or multiple conformers). Maintained by the Institut de Pharmacologie Moléculaire et Cellulaire, France.
- **ChemDB/ChemicalSearch.** Find chemicals by various search criteria.
- **Structural Database (CSD).** Repository for small molecule crystal structures in CIF format. The CSD is compiled and maintained by the Cambridge Crystallographic Data Centre
- **SPRESI^{web}.** Integrated database containing over 8.7 million molecules, 4.1 million reactions, 658,000 references and 164,000 patents covering the years 1974 - 2009. Developed by InfoChem.
- **MMSiNC.** Database of non-redundant, annotated and biomedically relevant chemical structures. Includes the analysis of chemical properties, such as ionization and tautomerization processes, and the in silico prediction of 24 important molecular properties in the biochemical profile of each structure. MMSiNC supports various types of queries, including substructure queries and the novel 'molecular scissoring' query. MMSiNC is interfaced with other primary data collectors, such as PubChem, Protein Data Bank (PDB), the Food and Drug Administration (FDA) database of approved drugs and ZINC, provided by the CRS4 - Bioinformatics Laboratory, Parco Sardegna Ricerche, Italy.

www.click2drug.org/directory_MolecularModeling.html

Virtual Computational Chemistry Laboratory - VCCLAB

Virtual Laboratory Sc x

www.vcclab.org/lab/

Apps Enzymology Piano Music Production Bioinformatics Databases Bioinformatics T... Misc Programming D pmartel Other bookmarks

http://www.vcclab.org

Virtual Computational Chemistry Laboratory

Home About Partners Software Articles Servers Download Web Services How to cite? Contact

on-line software

- ALOGPS 2.1* is the most accurate program to predict lipophilicity and aqueous solubility of molecules
- ASNN* calculates highly predictive non-linear neural network models
- E-BABEL is molecular structure information interchange hub
- PNN produces clearly interpretable analytical non-linear models
- PCLIENT generates more than 3000 descriptors
- E-DRAGON calculates DRAGON molecular indices
- PLS implements original two-step descriptors selection procedure
- UFS produces a reduced data set that contains no redundancy and a minimal amount of multicollinearity

If you have any questions, problems to run applets, please, contact

PREV TOP

Home About Partners Software Articles Servers Download Web Services How to cite? Contact

Virtual Computational Chemistry Laboratory - VCCLAB

The screenshot displays the VCCLAB Servers website interface. At the top, the browser address bar shows www.vcclab.org/servers/. The page title is "Virtual Computational Chemistry Laboratory". A navigation menu includes links for Home, About, Partners, Software, Articles, Servers, Download, Web Services, How to cite?, and Contact. A sidebar on the left lists: Home, About, Partners, Software, Articles, Servers, Download, Web Services, How to cite?, and Contact. The main content area features a map of Europe with server locations marked: Portsmouth, Erlangen, Munich, Milano, Moscow, and Kyiv. An "External User" is shown connecting to the network. The central diagram, titled "VCCLAB Servers", illustrates a network architecture. A central server node in Munich (SuperServer, VCCLAB site, ASNN, ALOGPS, PNN, BABEL, ParamClient) is connected to various client nodes and software services: Portsmouth (UFS, data reduction algorithm), Erlangen (CORINA, 2D => 3D), Moscow (MSU, fragmental descriptors, PLS), Kyiv (IBPC, E-state indices), Milano (DRAGON, molecular descriptor calculation), and External programs (KOWWIN, ChemEXper, etc.). A box labeled "Applet Clients" lists: ALOGPS 2.1, ASNN, BABEL, PNN, PARAMETER, CLIENT, E-DRAGON1.0, PLS, and UFS. On the right, a vertical list titled "ON-LINE SOFTWARE" includes: ALOGPS 2.1, ASNN, E-BABEL, PNN, PCLIENT, E-DRAGON 1.0, PLS, UFS, and SPC. Navigation arrows for "PREV" and "TOP" are visible at the bottom of the diagram area.

iDrug: on-line Drug Design Workbench

The screenshot displays the iDrug web interface. At the top, the browser address bar shows lilab.ecust.edu.cn/idrug/. The navigation bar includes links for Documentation, Register, Login, Guest, and Load Session. The main interface is divided into three sections:

- Tasks:** A tree view showing a hierarchy of tasks: Demo (No editable) > Pharmacophore > Target Navigator (4OH-tamoxifen) > Hit Explorer (CDK2: 1AQ1) > 18820, 18822. Other tasks include Similarity and HYZ_2RGP.mol2 (EGFR).
- 3D Model:** A central 3D visualization of a protein structure (blue) with a ligand (red) and several colored spheres (green, purple, pink) representing different pharmacophore features. The JSmol logo is visible in the bottom right of the model area.
- Results:** A table with the following data:

Pocket	Volume	RankScore	Druggability
+ 1	1386.75	5.21	932.00
+ 2	117.87	4.64	593.00
+ 3	142.25	2.64	-880.00
+ 4	170.75	1.98	-954.00

Below the results table, it indicates "Showing 1 to 4 of 4 entries". At the bottom of the interface, there is a footer with system requirements and contact information: "IE 9.0+, Firefox, Chrome, Safari is recommended for viewing this web site. JavaScript should be enabled for viewing chemical structures. For any problems, please contact: lilab_ecust@163.com Prof. Honglin Li's Group, School of Pharmacy, East China University of Science & Technology".

Load Add Feature Clear

Save Results

iDrug: a web-accessible and interactive drug discovery and design platform

Xia Wang¹, Haipeng Chen², Feng Yang², Jiayu Gong², Shiliang Li¹, Jianfeng Pei^{3*}, Xiaofeng Liu^{1*}, Hualiang Jiang¹, Luhua Lai³ and Honglin Li^{1,2*}

Wang *et al.* *Journal of Cheminformatics* 2014, **6**:28

<http://lilab.ecust.edu.cn/idrug/>

Drug Design Workshop

Drug Design Workshop

Swiss Institute of Bioinformatics

FONDS NATIONAL SUISSE DE LA RECHERCHE SCIENTIFIQUE

Drug Design Workshop

Workshop Biological context Help Medias More Disclaimer

How do researchers design tomorrow's drugs?

a workshop on
DRUG DESIGN,
or how to design
tomorrow's medicine

Try and design a drug...

<http://www.drug-design-workshop.ch/>

iDrug-Target

iDrug-Target: A package of web-... x +

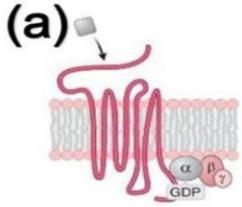
Not secure | jci-bioinfo.cn/iDrug-Target/

Apps | Bookmarks | Settings | Extensions | Ualg | Tools | Code Tools | LibGen | Cell Bits | 10FastFingers | Other bookmarks

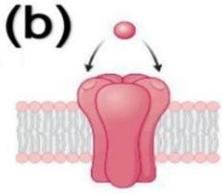
iDrug-Target: A package of web-services for predicting drug-target interaction

[Read Me](#) | [Data](#) | [Supporting information](#) | [Citation](#) |

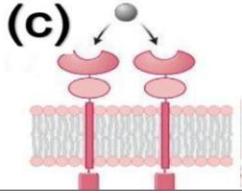
(a) [iDrug-GPCR:](#)
The web-server for predicting the interaction between GPCRs and drugs in cellular networking.



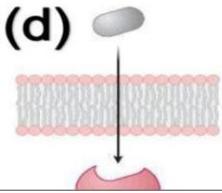
(b) [iDrug-ChI:](#)
The web-server for predicting the interaction between ion channels and drugs in cellular networking.



(c) [iDrug-Ezy:](#)
The web-server for predicting the interaction between enzymes



(d) [iDrug-NR:](#)
The web-server for predicting the interaction between nuclear receptors



<http://www.jci-bioinfo.cn/iDrug-Target/>

*Bases de dados e formatos de
representação de moléculas*

Bases de dados

- Macromoléculas (Target):
 - Estrutura (Protein Data Bank, PLD, TTD, ModBase)
 - Sequência (Uniprot, Genbank, ...)

- Moléculas pequenas:
 - (PubChem, Drugbank, Cambridge Database, ZINC, ChEMBL, TCM, WOMBAT,)

Contêm muita informação além da *estrutura/sequência* propriamente dita.

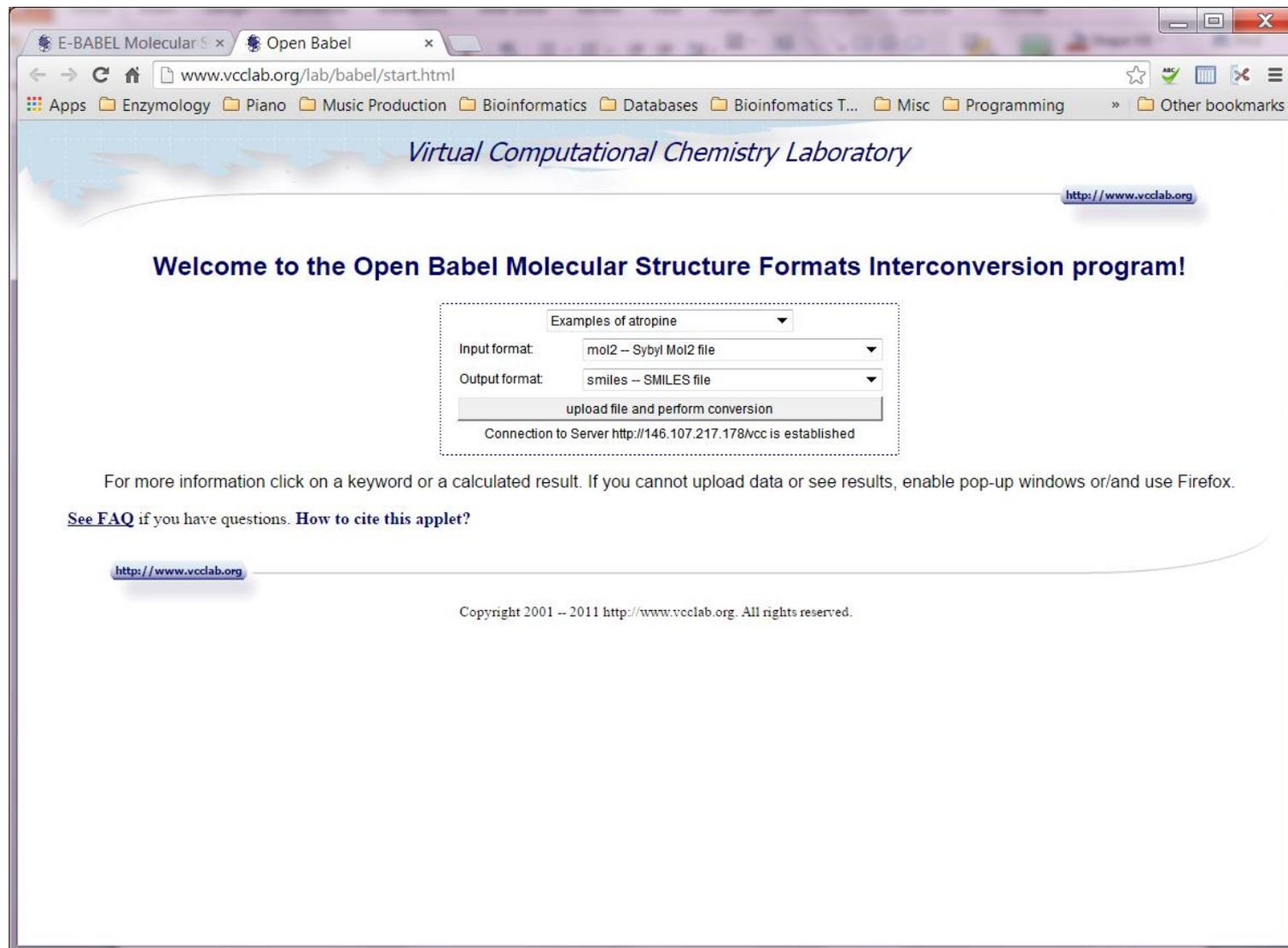
Formatos de representação

- Estrutura:
 - PDB, MDL, SDF, MOL2, CIF, ASN.1, HIN, Trypos, Sybil, Gaussian, XYZ, CML, XML, SMILES
- Sequência:
 - Fasta, SWISSPROT, ASN.1, GCG, GenBank, PIR, Phylip,

Ferramenta de conversão entre formatos:

OpenBabel (<http://openbabel.org>)

E-Babel: conversão de formatos online



The screenshot shows a web browser window with the following elements:

- Browser Tabs:** "E-BABEL Molecular S" and "Open Babel".
- Address Bar:** www.vcclab.org/lab/babel/start.html
- Bookmarks Bar:** Apps, Enzymology, Piano, Music Production, Bioinformatics, Databases, Bioinformatics T..., Misc, Programming, Other bookmarks.
- Page Header:** "Virtual Computational Chemistry Laboratory" and a link to <http://www.vcclab.org>.
- Main Content:**
 - Header:** "Welcome to the Open Babel Molecular Structure Formats Interconversion program!"
 - Form:** A dashed box containing:
 - Dropdown menu: "Examples of atropine"
 - Input format: "mol2 -- Sybyl Mol2 file"
 - Output format: "smiles -- SMILES file"
 - Button: "upload file and perform conversion"
 - Status: "Connection to Server http://146.107.217.178/vcc is established"
- Text:** "For more information click on a keyword or a calculated result. If you cannot upload data or see results, enable pop-up windows or/and use Firefox. See [FAQ](#) if you have questions. [How to cite this applet?](#)"
- Footer:** <http://www.vcclab.org> and "Copyright 2001 -- 2011 <http://www.vcclab.org>. All rights reserved."

OpenBabel

The screenshot displays the OpenBabelGUI interface. The window title is "OpenBabelGUI". The menu bar includes "File", "View", "Plugins", and "Help".

--- INPUT FORMAT ---
sdff -- MDL MOL format

--- OUTPUT FORMAT ---
pdb -- Protein Data Bank format

Input Panel:

- Use this format for all input files (ignore file extensions)
- Input below (ignore input file)

Output Panel:

- Output below only (no output file)
- Display in firefox

Conversion Options:

- Add hydrogens appropriate for this pH
- Convert dative bonds e.g. -[N+](O)=O to -N(=O)=O
- Make dative bonds e.g. -[N+](O)=O from -N(=O)=O
- Remove all but the largest contiguous fragment
- Center Coordinates
- Combine mols in first file with others by name

Filter: convert only when tests are true:

- Add properties from descriptors
- Delete properties in list

Append properties or descriptors in list to title:

- Join all input molecules into a single output molecule
- Output disconnected fragments separately
- add or replace a property (SDF)
- Add or replace molecule title
- Append text to title
- Output multiple conformers separately
- Append output index to title
- Additional file output
- Append input filename to title
- Append input index to title
- Adds hydrogen to nonpolar atoms only
- Adds hydrogen to polar atoms only
- Align coordinates to the first molecule
- Canonicalize the atom order
- Change cell size:
- Confab, the diverse conformer generator

Input Data:

```
2244
-OEChem-10171816273D

21 21 0 0 0 0 0 0999 V2000
 1.2333 0.5540 0.7792 O 0 0 0 0 0 0 0 0 0 0
 0 0
-0.6952 -2.7148 -0.7502 O 0 0 0 0 0 0 0 0 0 0
 0 0 0
 0.7958 -2.1843 0.8685 O 0 0 0 0 0 0 0 0 0 0
 0 0
 1.7813 0.8105 -1.4821 O 0 0 0 0 0 0 0 0 0 0
 0 0
-0.0857 0.6088 0.4403 C 0 0 0 0 0 0 0 0 0 0
 0 0
-0.7927 -0.5515 0.1244 C 0 0 0 0 0 0 0 0 0 0
 0 0
-0.7288 1.8464 0.4133 C 0 0 0 0 0 0 0 0 0 0
 0 0
-2.1426 -0.4741 -0.2184 C 0 0 0 0 0 0 0 0 0 0
 0 0 0
-2.0787 1.9238 0.0706 C 0 0 0 0 0 0 0 0 0 0
 0 0
-2.7855 0.7636 -0.2453 C 0 0 0 0 0 0 0 0 0 0
 0 0
-0.1409 -1.8536 0.1477 C 0 0 0 0 0 0 0 0 0 0
```

Output Data:

```
HETATM 7 C UNL 1 -0.729 1.846 0.413 1.00 0.00 C
HETATM 8 C UNL 1 -2.143 -0.474 -0.218 1.00 0.00 C
HETATM 9 C UNL 1 -2.079 1.924 0.071 1.00 0.00 C
HETATM 10 C UNL 1 -2.785 0.764 -0.245 1.00 0.00 C
HETATM 11 C UNL 1 -0.141 -1.854 0.148 1.00 0.00 C
HETATM 12 C UNL 1 2.109 0.671 -0.311 1.00 0.00 C
HETATM 13 C UNL 1 3.531 0.600 0.164 1.00 0.00 C
HETATM 14 H UNL 1 -0.185 2.755 0.659 1.00 0.00 H
HETATM 15 H UNL 1 -2.725 -1.361 -0.456 1.00 0.00 H
HETATM 16 H UNL 1 -2.580 2.887 0.051 1.00 0.00 H
HETATM 17 H UNL 1 -3.837 0.824 -0.509 1.00 0.00 H
HETATM 18 H UNL 1 3.729 1.418 0.859 1.00 0.00 H
HETATM 19 H UNL 1 4.205 0.697 -0.692 1.00 0.00 H
HETATM 20 H UNL 1 3.711 -0.366 0.643 1.00 0.00 H
HETATM 21 H UNL 1 -0.256 -3.592 -0.734 1.00 0.00 H
CONNECT 1 5 12
CONNECT 2 11 21
CONNECT 3 11
CONNECT 4 12
CONNECT 5 1 6 7
CONNECT 6 5 8 11
CONNECT 7 5 9 14
CONNECT 8 6 10 15
CONNECT 9 7 10 16
```

Formato FASTA

- É um formato de representação de sequências biológicas (DNA ou proteína)
- Consiste numa linha de cabeçalho, seguida de linhas contendo a sequência de aminoácidos ou nucleótidos representada em códigos de 1 letra
- Contem muito pouca informação para além da sequência

Formato FASTA

Cabeçalho

>gi|19151|emb|Z14088.1| L.esculentum mRNA for 108 protein

```
AACAATCATGGCATCTGTGAAGTCGTCGTCGTCATCATCATCATTTTATTTTCCTTGTT  
GTTGTTGATTTTGGCTTGTGATTGTAAGTCAAGTTATCGAGTGTCAACCTCAACAGT  
CATGCACCGCGTCACTTACTGGCCTGAACGTCTGCGCCCCATTCTGGTCCCAGGCTCACCTAC  
TGCAAGTACGGAGTGTTGCAA TGCAGTACAGTCGATTAATCATGACTGTATGTGCAACACT  
ATGCGCATTGCAGCTCAAATTCCAGCTCAG TGCAACCTCCCTCCACTCTCTTGTTCTGCAAAT  
TGAGTTTGAGATCAGTGGCCAGCAAGTTTACATCTGC TACATGAGCAAATTAATAATATC  
GTAACAATAAATTAAGTTGTCTTTTTTTTTTTTTTTGGTTATGCAAC AGACCAAGGGGGTCA  
TGAGAAAAGAGTTTGTACTATCATATGATTATCAATAAAAAAAATTATGAG
```

>Q43495|108_SOLLC Protein 108 precursor - Solanum lycopersicum

```
MASVKSSSSSSSSSFISLLLLLILLVIVLQSQVIECQPQQSCTASLTGLNVCAPFLVPGSP  
TASTECCNAVQSINHDCMCNTMRIA AQIPAQC�LPPLSCSAN
```

Sequência

Formato SWISSPROT

- Representação de sequências de proteína
- Sintaxe complexa com uma variedade de *campos*
- Contem muita informação além da sequência

Formato SWISSPROT

ID TRY1_HUMAN Reviewed; 247 AA.
AC P07477; A1A509; A6NJ71; B2R5I5; Q5NV57; Q7M4N3; Q7M4N4; Q92955;
AC Q9HAN4; Q9HAN5; Q9HAN6; Q9HAN7;
DT 01-APR-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-APR-1988, sequence version 1.
DT 18-SEP-2013, entry version 154.
DE RecName: Full=Trypsin-1;
DE EC=3.4.21.4;
DE AltName: Full=Beta-trypsin;
DE AltName: Full=Cationic trypsinogen;
DE AltName: Full=Serine protease 1;
DE AltName: Full=Trypsin I;
DE Contains:
DE RecName: Full=Alpha-trypsin chain 1;
DE Contains:
DE RecName: Full=Alpha-trypsin chain 2;
DE Flags: Precursor;
GN Name=PRSS1; Synonyms=TRP1, TRY1, TRYP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
OC Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX PubMed=3011602; DOI=10.1016/0378-1119(86)90111-3;
RA Emi M., Nakamura Y., Ogawa M., Yamamoto T., Nishide T., Mori T.,
RA Matsubara K.;
RT "Cloning, characterization and nucleotide sequences of two cDNAs
RT encoding human pancreatic trypsinogens.";
RL Gene 41:305-310(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].

(continua)

Formato SWISSPROT

RX PubMed=8650574; DOI=10.1126/science.272.5269.1755;
RA Rowen L., Koop B.F., Hood L.;
RT "The complete 685-kilobase DNA sequence of the human beta T cell
RT receptor locus.";
RL Science 272:1755-1762(1996).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Prostate;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,

• • • • •

T STRAND 183 187
FT STRAND 192 194
FT STRAND 203 206
FT STRAND 209 216
FT STRAND 218 222
FT STRAND 227 231
FT HELIX 232 235
FT HELIX 236 245
SQ SEQUENCE 247 AA; 26558 MW; DD49A487B8062813 CRC64;
MNPLLILTFV AAALAAPFDD DDKIVGGYNC EENSVPYQVS LNSGYHFCGG
SLINEQWVVS
AGHCYKSRIQ VRLGEHNIEV LEGNEQFINA AKIIRHPQYD RKTLLNDIML IKLSSRAVIN
ARVSTISLPT APPATGKCL ISGWGNTASS GADYPDELQC LDAPVLSQAK CEASYPGKIT
SNMFVCGFLE GGDSCQGDS GGPVVCNGQL QGVVSWGDGC AQKNKPGVYT
KVYNYVKWIK
NTIAANS
//

UniProt, a referência universal para sequências de proteínas

- A fusão das bases de dados PIR, TrEMBL e Swiss-Prot numa única base de dados vem constituir uma referência definitiva para a pesquisa de sequências de proteína.
- Uniprot contem as seguintes subsecções:
 - UniProtKB: contem SwissProt e TrEMBL (translated EMBL)
 - UniParc: contem sequências não-annotadas de várias fontes
 - UniRef: contem sequências agrupadas por similaridade

<http://uniprot.org>



UniProt - Mozilla Firefox

File Edit View History Bookmarks ScrapBook Tools Help del.icio.us

http://beta.uniprot.org/ Uniprot

Google Calendar Gmail Wiley InterScience: J... Tony Schreiner's We... Prediction of Second... Gmail - Inbox Marés - Portos Princi...

UniProt Downloads · Contact · Help

Search in **Query**

Protein Knowledgebase (UniProtKB) Search Clear Fields »

Core Data

- Protein Knowledgebase (UniProtKB)
- Sequence Clusters (UniRef)
- Sequence Archive (UniParc)

Supporting Data

- Literature citations
- Taxonomy
- Keywords

Information

- News
- Documents
- FAQ
- Help

the scientific community with a
ely accessible resource of protein

What we provide

UniProtKB	Protein knowledgebase, consists of two sections: <ul style="list-style-type: none"> ★ Swiss-Prot, which is manually annotated and reviewed. ★ TrEMBL, which is automatically annotated and is not reviewed.
UniRef	Sequence clusters, used to speed up similarity searches.
UniParc	Sequence archive, used to keep track of sequences and their identifiers.
Supporting data	Literature citations , taxonomy , keywords and more .

NEWS

Release 12.6 – Dec 4, 2007
Complete proteome for Arabidopsis thaliana in UniProtKB

- › Statistics for UniProtKB: [Swiss-Prot](#) · [TrEMBL](#)
- › [Forthcoming changes](#)
- › [News archives](#)

SITE TOUR

Learn how to make best use of the tools and data on this site.

PROTEIN SPOTLIGHT

Downloads brochure.pdf MAPMBSaug200... surface.csh T00710308TQ1... pass.pdf

Done FoxyProxy: Ualg

insulin in UniProtKB - Mozilla Firefox

File Edit View History Bookmarks ScrapBook Tools Help del_jcio.us

http://beta.uniprot.org/uniprot/?query=insulin&sort=score uniprot

Google Calendar Gmail Wiley InterScience: J... Tony Schreiner's We... Prediction of Second... Gmail - Inbox Marés - Portos Princ... Index of /cd oranger

UniProt UniProtKB Downloads Contact Help

Search in Protein Knowledgebase (UniProtKB) Query insulin Search Clear Fields »

Search Blast Align Retrieve ID Mapping *

1 - 25 of 2,876 results for **insulin** in UniProtKB sorted by **score** descending

Browse by taxonomy, keyword, gene ontology, enzyme class or pathway | Reduce sequence redundancy to 100%, 90% or 50% | Customize [Download...](#) display

› Show only **reviewed** (UniProtKB/Swiss-Prot) or **unreviewed** (UniProtKB/TrEMBL) entries

› Restrict term "insulin" to **protein family**, **gene name**, **gene ontology**, **protein name**, **strain**, **taxonomy**, **tissue**, **web resource**

Page 1 of 116 | Next »

	Accession	Entry Name	Status	Protein Names	Genes	Organism	Length
<input type="checkbox"/>	P06213	INSR_HUMAN	★	Insulin receptor precursor (EC 2.7.10.1) (IR) (CD220 antigen) [Cleaved into: Insulin receptor subunit alpha; Insulin receptor subunit beta]	INSR	Homo sapiens (Human)	1,382
<input type="checkbox"/>	P01308	INS_HUMAN	★	Insulin precursor [Cleaved into: Insulin B chain; Insulin A chain]	INS	Homo sapiens (Human)	110
<input type="checkbox"/>	P35568	IRS1_HUMAN	★	Insulin receptor substrate 1 (IRS-1)	IRS1	Homo sapiens (Human)	1,242
<input type="checkbox"/>	P09208	INSR_DROME	★	Insulin-like receptor precursor (EC 2.7.10.1) (DIR) (DInr) (dIRH) [Cleaved into: Insulin-like receptor subunit alpha; Insulin-like receptor subunit beta 1; Insulin like receptor subunit beta 2]	InR (dInr) (Dir-a) (Inr-a) (CG18402)	Drosophila melanogaster (Fruit fly)	2,144

Downloads brochure.pdf MAPMBSaug200... surface.csh T00710308TQ1... pass.pdf Clear

Done FoxyProxy: Ualg

Insulin receptor precursor - Homo sapiens (Human) - Mozilla Firefox

File Edit View History Bookmarks ScrapBook Tools Help del_jcio.us

http://beta.uniprot.org/uniprot/P06213 uniprot

Google Calendar Gmail Wiley InterScience: J... Tony Schreiner's We... Prediction of Second... Gmail - Inbox Marés - Portos Princi... Index of /cd oranger

★ Reviewed, UniProtKB/Swiss-Prot **P06213** (INSR_HUMAN)

Last modified November 13, 2007. Version 123. [History...](#)

Contribute
Send feedback
WikiProteins

Clusters with 100%, 90%, 50% identity | Documents (7) | Third-party data | Customize display

TEXT XML RDF/XML GFF FASTA

[Names and origin](#) · [General annotation \(Comments\)](#) · [Ontologies](#) · [Binary interactions](#) · [Alternative products](#) · [Sequence annotation \(Features\)](#) · [Sequences](#) · [References](#) · [Web resources](#) · [Cross-references](#) · [Entry information](#) · [Relevant documents](#)

Names and origin Hide | Top

Protein names	Insulin receptor [Precursor] <i>Also known as:</i> EC 2.7.10.1 IR CD220 antigen <i>Cleaved into:</i> Insulin receptor subunit alpha Insulin receptor subunit beta
Gene names	Name: INSR
Organism	Homo sapiens (Human)
Taxonomic identifier	9606 [NCBI]
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo
Protein existence	Evidence at protein level.

General annotation (Comments) Hide | Top

Function	This receptor binds insulin and has a tyrosine-protein kinase activity. Isoform Short has a higher affinity for insulin. Mediates the metabolic functions of insulin. Binding to insulin stimulates association of the receptor with downstream mediators including IRS1 and phosphatidylinositol 3'-kinase (PI3K). Can activate PI3K either directly by binding to the p85 regulatory subunit, or indirectly via IRS1.
Catalytic activity	ATP + a [protein]-L-tyrosine = ADP + a [protein]-L-tyrosine phosphate.
Enzyme regulation	Autophosphorylation activates the kinase activity.
Subunit structure	Tetramer of 2 alpha and 2 beta chains linked by disulfide bonds. The alpha chains contribute to the formation of the ligand-binding domain, while the beta chains carry the kinase domain. Interacts with SORBS1 but dissociates from it following insulin stimulation. Binds SH2B2. Interacts with the PTB/PID domains of IRS1

Done

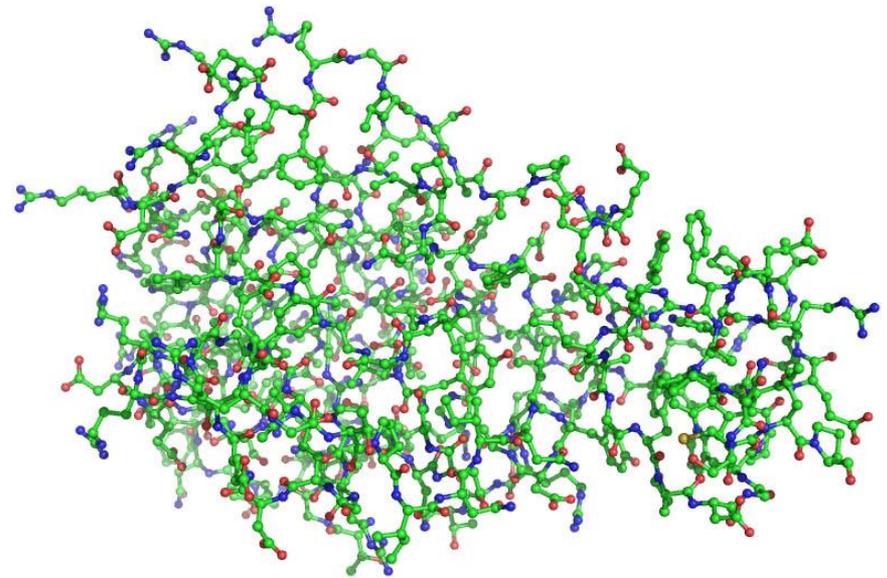
FoxyProxy: Ualg

A representação da estrutura é muito mais complexa que a sequência

Enquanto a sequência de uma proteína ou ácido nucleico é caracterizada simplesmente pela base ou aminoácido que ocorre em cada posição, a descrição duma estrutura molecular implica a indicação da posição de cada átomo no espaço tridimensional, bem como a especificação das ligações químicas entre todos os átomos que constituem a molécula

...AVAGGATILVHNQDAGEPAIVLAFG...

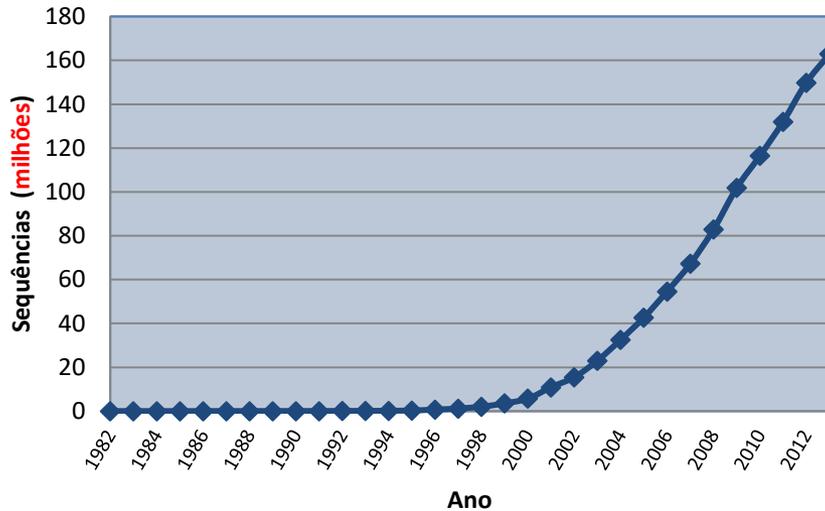
Sequência



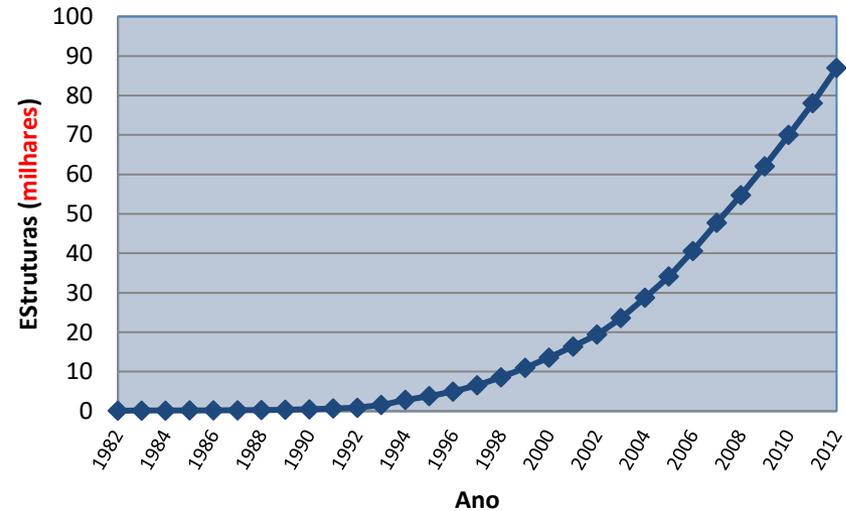
Estrutura

Sequência versus estrutura

Crescimento do GenBank



Crescimento do Protein Databank



milhões de sequências versus milhares de estruturas!

Em 1982: conhecidas 172 estruturas e 315 sequências ...

Hoje (Nov 2014): conhecidas 104,866 estruturas e 274,414,298 sequências!!

Conclusão: A determinação das sequências faz-se a um ritmo muito superior ao das estruturas (cada vez temos mais proteínas de **sequência conhecida** e **estrutura desconhecida**)!

Formatos de representação da estrutura

- A representação da estrutura molecular em bancos de dados passa pela descrição das **coordenadas atómicas**, do **tipo de átomo**, e das **ligações químicas** presentes.
- No caso das proteínas, a topologia de ligação dos 20 aminoácidos standard pode ser assumida *a priori*
- A topologia de outras moléculas, tais como grupos prostéticos, deverá ser especificada
- O formato “tradicional” de representação de estruturas de proteínas é o formato **PDB** (Protein Data Bank file format).
- Para moléculas pequenas usam-se muitos outros formatos, tais como: **cif, asn.1, mol, mdl, mol2, sdf, hin, ...**, ...

```
@<TRIPOS>MOLECULE
C9H8O4
 21 21 1 0 0
SMALL
NO_CHARGES
```

Representação da Aspirina em formato MDL2

```
@<TRIPOS>ATOM
```

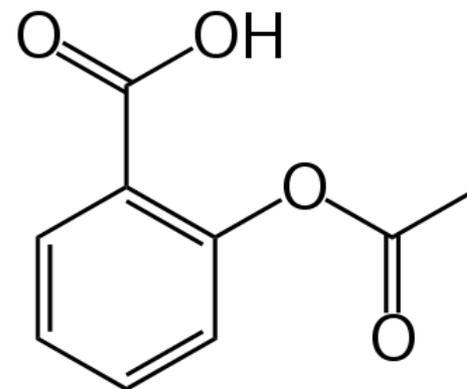
1 C1	2.2393	-0.3791	0.2630	C.ar	1 <1>	0.0000
2 C2	0.8424	1.9231	-0.4249	C.ar	1 <1>	0.0000
3 C3	2.8709	0.8456	0.2722	C.ar	1 <1>	0.0000
4 C4	2.1751	1.9935	-0.0703	C.ar	1 <1>	0.0000
5 C5	-3.4838	0.4953	-0.0896	C.3	1 <1>	0.0000
6 C6	0.8910	-0.4647	-0.0939	C.ar	1 <1>	0.0000
7 C7	0.1908	0.6991	-0.4402	C.ar	1 <1>	0.0000
8 O1	-0.9633	-1.8425	-0.4185	O.2	1 <1>	0.0000
9 O2	-1.6531	0.8889	1.3406	O.2	1 <1>	0.0000
10 O3	0.8857	-2.8883	0.2267	O.3	1 <1>	0.0000
11 C8	0.2090	-1.7720	-0.1069	C.2	1 <1>	0.0000
12 C9	-2.0185	0.6853	0.2071	C.2	1 <1>	0.0000
13 O4	-1.1189	0.6285	-0.7886	O.3	1 <1>	0.0000
14 H1	0.3962	-3.7219	0.2035	H	1 <1>	0.0000
15 H2	2.7867	-1.2719	0.5268	H	1 <1>	0.0000
16 H3	0.3069	2.8224	-0.6911	H	1 <1>	0.0000
17 H4	3.9130	0.9108	0.5482	H	1 <1>	0.0000
18 H5	2.6781	2.9492	-0.0604	H	1 <1>	0.0000
19 H6	-3.7360	-0.5623	-0.0120	H	1 <1>	0.0000
20 H7	-4.0763	1.0637	0.6273	H	1 <1>	0.0000
21 H8	-3.6988	0.8471	-1.0986	H	1 <1>	0.0000

Coordenadas

```
@<TRIPOS>BOND
```

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

Ligações

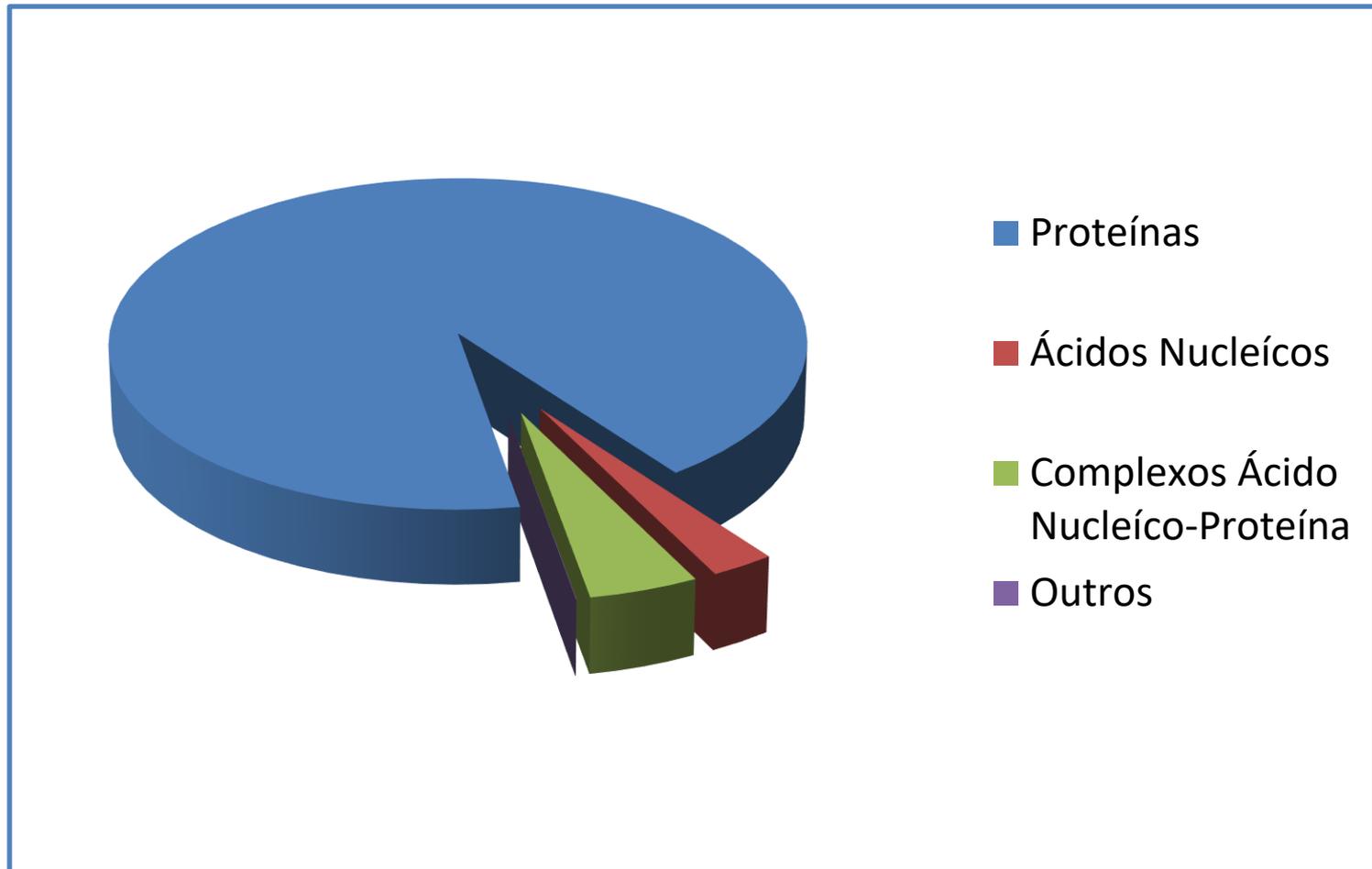


O Protein Data Bank

- O Protein Data Bank (PDB) foi criado em 1971 por E.Meyer e W.Hamilton, do Brookhaven National Laboratory (USA), contendo no início 7 estruturas!
- A gestão do PDB foi transferida em 1998 para os membros do RCSB (Research Collaboratory in Structural Bioinformatics) dos quais a Universidade de Rutgers é o site principal. O PDB (<http://www.rcsb.org>) é um banco de dados de acesso **livre**.
- Contendo inicialmente estruturas de proteínas, o PDB contém hoje em dia outros tipos de moléculas, tais como ácidos nucleicos, lípidos e polissacáridos.
- Número total de estruturas em 9/1/2022: **185610**

Técnica experimental	Proteínas	Ácidos nucleicos	Complexos Ac.Nuc./Proteína	Outros	Total
Cristalografia de raios X	151958	2387	7575	161	162081
NMR	11881	1391	274	37	13583
Microscopia electrónica	7477	61	2101	3	9642
Outras	102	3	3	4	109
Combinação	183	8	8	1	195
Total	171601	3850	9953	206	185610

O Protein Data Bank contem vários tipos de macromoléculas



De onde provêm a informação estrutural ?

Combinação de vários tipos de conhecimento:

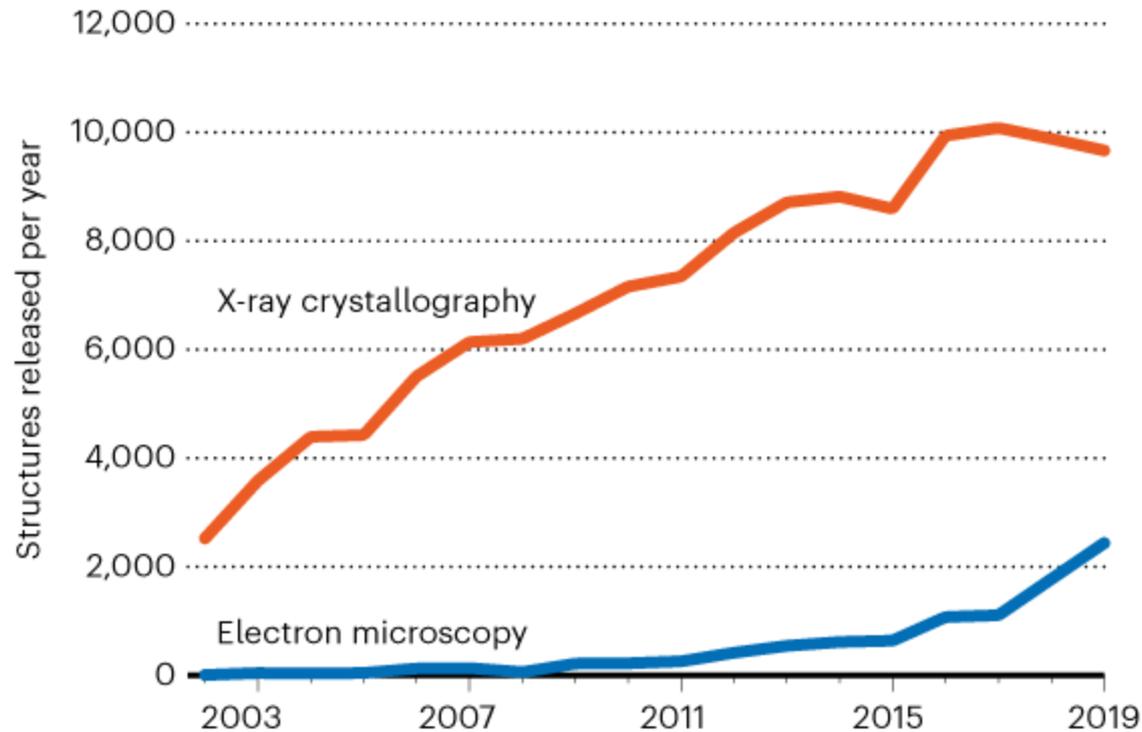
- Teoria da ligação química
- Geometria de moléculas pequenas
- Métodos experimentais para a determinação da estrutura:
 - ❖ Cristalografia de raios X
 - ❖ Ressonância Magnética Nuclear (NMR)
 - ❖ Outros métodos (microscopia, difracção de neutrões, etc...)

Métodos experimentais

- **Cristalografia de raios X:** a molécula a estudar é purificada e cristalizada a partir de uma solução concentrada. Um feixe de raios X é projectado através do cristal da molécula e o padrão de difracção obtido é usado para resolver a estrutura.
- **Ressonância magnética Nuclear:** a molécula purificada é colocada numa solução aquosa bastante concentrada. A acção de um campo magnético muito intenso provoca o desdobramento dos níveis de energia do spin nuclear de alguns elementos (H, ^{13}C , ^{15}N), permitindo o estudo do seu ambiente químico e a determinação da estrutura da macromolécula.
- **Crio-microscopia electrónica:** a amostra da molécula a estudar é congelada rapidamente a cerca de $-180\text{ }^\circ\text{C}$ e um feixe de electrões é usado para criar imagens de um enorme número de moléculas da amostra. A análise combinada destas imagens permite resolver a estrutura 3D da molécula.

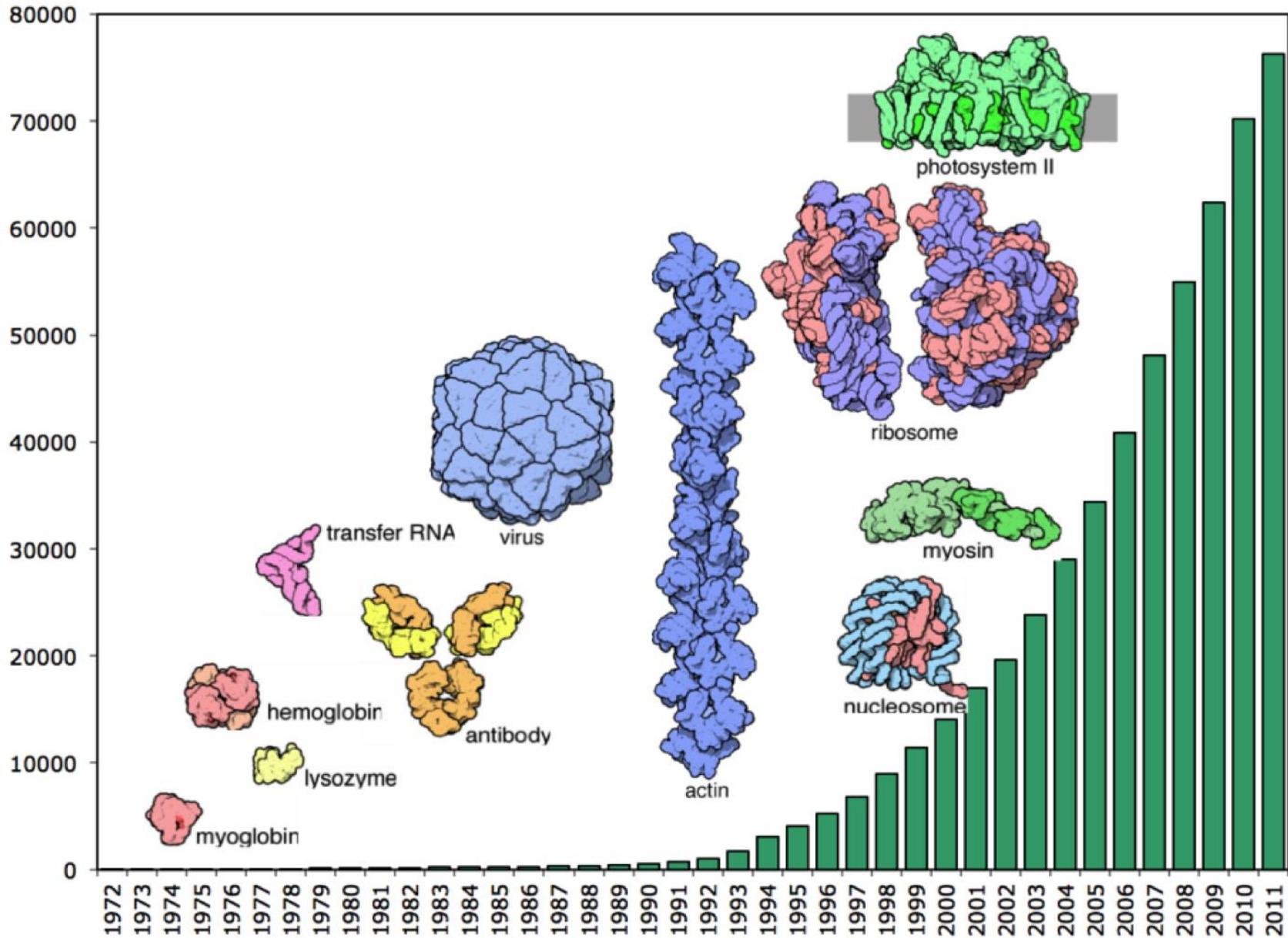
STRUCTURE SLEUTHS

Most structures of proteins and other biological molecules are still solved with X-ray crystallography. But a revolutionary technique called cryo-electron microscopy is catching up, as it becomes more sensitive and widely available.



The electron microscopy line shows structures submitted to the Electron Microscopy Data Bank. Nearly all use cryo-EM.

Estruturas no Protein Data Bank



Formato da informação no Protein Data Bank

- A informação contida no Protein Databank inclui coordenadas atómicas, topologias de ligação (descrição das ligações químicas), nomes dos átomos e grupos químicos, e outros dados associados ao processo de determinação experimental da estruturas
- Presentemente a informação no PDB está disponível nos seguintes formatos:
 - **pdb file:** O formato “flat file”, um tipo de ficheiro chamado “ficheiro PDB”. Estes ficheiros são os mais utilizados pelos softwares de manipulação e visualização de estruturas e têm geralmente a extensão “.pdb”
 - **mmCIF:** - um formato mais poderoso e estruturado que o ficheiro PDB, ainda não tendo sido largamente adoptado
 - **XML:** - extended mark-up language, um formato muito geral de representação de informação, compatível com um vasto número de aplicações de software.

Formato do ficheiro PDB

```

HEADER      METAL BINDING PROTEIN                      21-AUG-03   1Q8H
TITLE       CRYSTAL STRUCTURE OF PORCINE OSTEOCALCIN
COMPND      MOL_ID: 1;
COMPND      2 MOLECULE: OSTEOCALCIN;
COMPND      3 CHAIN: A
SOURCE      MOL_ID: 1;
SOURCE      2 ORGANISM_SCIENTIFIC: SUS SCROFA;
SOURCE      3 ORGANISM_COMMON: PIG
KEYWDS      HELIX-TURN-HELIX-TURN-HELIX, PAPER-CLIP, HYDROXYAPATITE
KEYWDS      2 CRYSTAL SURFACE BINDING PROTEIN, CALCIUM BINDING PROTEIN,
KEYWDS      3 BONE GLA PROTEIN
EXPDTA      X-RAY DIFFRACTION
AUTHOR      Q.Q.HOANG,F.SICHERI,A.J.HOWARD,D.S.YANG
REVDAT      1 11-NOV-03 1Q8H 0
JRNL        AUTH  Q.Q.HOANG,F.SICHERI,A.J.HOWARD,D.S.YANG
JRNL        TITL  BONE RECOGNITION MECHANISM OF PORCINE OSTEOCALCIN
JRNL        TITL 2 FROM CRYSTAL STRUCTURE.
JRNL        REF  NATURE                      V. 425   977 2003
JRNL        REFN ASTM NATUAS  UK ISSN 0028-0836
REMARK      1
REMARK      2
REMARK      2 RESOLUTION. 2.00 ANGSTROMS.
REMARK      3
REMARK      3 REFINEMENT.
REMARK      3 PROGRAM      : CNS 1.1
REMARK      3 AUTHORS      : BRUNGER,ADAMS,CLORE,DELANO,GROS,GROSSE-
.....
ATOM        1  N   PRO A  13      10.210  29.966  44.935  1.00 38.06
ATOM        2  CA  PRO A  13      9.718  29.013  43.919  1.00 37.33
ATOM        3  C   PRO A  13      9.566  29.662  42.541  1.00 37.52
ATOM        4  O   PRO A  13      9.275  30.855  42.444  1.00 38.00
ATOM        5  CB  PRO A  13      8.383  28.488  44.434  1.00 37.68
ATOM        6  CG  PRO A  13      7.919  29.624  45.336  1.00 36.60
ATOM        7  CD  PRO A  13      9.196  30.126  45.995  1.00 36.47
ATOM        8  N   ASP A  14      9.777  28.879  41.483  1.00 36.83
ATOM        9  CA  ASP A  14      9.671  29.384  40.116  1.00 36.13
.....
MASTER     299  0  6  3  0  0  0  6 378  1 38  4
END

```

Header

Coordenadas

Portal de acesso ao PDB

- Acesso ao repositório de estruturas do Protein Databank
- Pesquisa por nomes, sequência, estruturas, ligandos, organismo, método experimental, etc...
- Ferramentas integradas para visualização, comparação de estruturas, análise, etc...

New: More Computed Structure Models (CSM) available [Learn more](#)

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

RCSB Protein Data Bank (RCSB PDB) enables breakthroughs in science and education by providing access and tools for exploration, visualization, and analysis of:

- Experimentally-determined 3D structures from the **Protein Data Bank (PDB)** archive
- Computed Structure Models (CSM)** from AlphaFold DB and ModelArchive

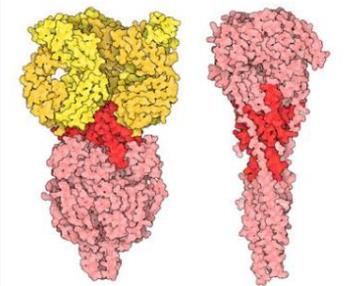
These data can be explored in context of external annotations providing a structural view of biology.

Explore
NEW
Features



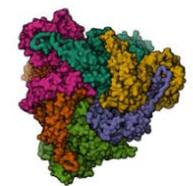
Virtual Crash Course
Leveraging RCSB PDB APIs for Bioinformatics Analyses and Machine Learning
October 12 | Register Now!

October Molecule of the Month



RSV Fusion Glycoprotein

Latest Entries As of Tue Oct 03 2023



8C0V

Structure of the peroxisomal Pex1/Pex6 ATPase complex bound to a substrate in single seam state

Features & Highlights

-  ASBMB Members: Register Now for Virtual Event
Learn about *Teaching enzymology with the Protein Data Bank: from pandemic to Paxlovid* on October 19.
-  Register Now for October Virtual Crash Courses on RCSB PDB APIs
Learn about *Leveraging RCSB PDB APIs for Bioinformatics Analyses and Machine Learning*. Part 1 will introduce the Data and Search APIs. Part 2 will offer hands-on exercises.
-  Coming Soon: PDB Entries with Novel Ligands Distributed Only in PDBx/mmCIF and PDBML File Formats
PDB will run out of 3-character CCD IDs before the end of 2023.

News Publications -

-  **Fall Newsletter Published**
In this issue: Fall Training Opportunities; DNS Name Changes; IUCr Report; and more. In the Education Corner, learn about *MedChemBlog: An Innovative Distance Learning Experience for Teaching Medicinal Chemistry* » 10/08/2023
-  **The Nobel Prize in Physiology or Medicine 2023**
Congratulations to Katalin Karikó and Drew Weissman on the award for their *discoveries concerning nucleoside base modifications that enabled the development of effective mRNA vaccines against COVID-19*. PDB-101 offers resources to learn more about

New: More Computed

- Welcome
- Deposit
- Search
- Visualize
- Analyze
- Download
- Learn

RCSB Protein Data Bank (RCSB PDB) enables breakthrough science and education by providing access and tools for visualization, and analysis of:

Experimentally-determined 3D structures from the Protein Data Bank (PDB) archive

Computed Structure Models (CSM) from AlphaFold ModelArchive

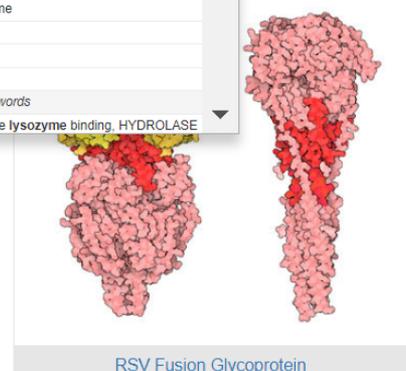
These data can be explored in context of external annotations providing a structural view of biology.

Explore NEW Features

Virtual Crash Course
Leveraging RCSB PDB APIs for Bioinformatics Analyses and Machine Learning
October 12 | Register Now!

SEARCH API DATA API

- in Structure Keywords
- LYSOZYME
 - Immune system, Lysozyme
- in UniProt Molecule Name
- Lysozyme
 - Lysozyme 1
 - Lysozyme C
 - Lysozyme C I
 - Lysozyme C II
 - Lysozyme C, milk isozyme
 - Lysozyme C-1
 - Lysozyme C-2
 - Lysozyme g
- in Additional Structure Keywords
- lysozyme inhibitor, g-type lysozyme binding, HYDROLASE



Latest Entries As of Tue Oct 03 2023

8C0V

Structure of the peroxisomal Pex1/Pex6 ATPase complex bound to a substrate in single seam state

Features & Highlights

ASMB Members: Register Now for Virtual Event
Learn about *Teaching enzymology with the Protein Data Bank: from pandemic to Paxlovid* on October 19.

Register Now for October Virtual Crash Courses on RCSB PDB APIs
Learn about *Leveraging RCSB PDB APIs for Bioinformatics Analyses and Machine Learning*. Part 1 will introduce the Data and Search APIs. Part 2 will offer hands-on exercises.

Coming Soon: PDB Entries with Novel Ligands Distributed Only in PDBx/mmCIF and PDBML File Formats
PDB will run out of 3-character CCD IDs before the end of 2023.

News Publications

Fall Newsletter Published
In this issue: Fall Training Opportunities; DNS Name Changes; IUCCr Report; and more. In the Education Corner, learn about *MedChemBlog: An Innovative Distance Learning Experience for Teaching Medicinal Chemistry* » 10/06/2023

The Nobel Prize in Physiology or Medicine 2023
Congratulations to Katalin Karikó and Drew Weissman on the award for their discoveries concerning nucleoside base modifications that enabled the development of effective mRNA vaccines against COVID-19. PDB-101 offers resources to learn more about

Search Summary This query matches 5,868 Structures.

Refinements

Structure Determination Methodology

- experimental (5,868)

Scientific Name of Source Organism

- Gallus gallus (1,155)
- Homo sapiens (1,114)
- Tequatrovirus T4 (810)
- Mus musculus (170)
- Escherichia coli (157)
- Escherichia coli K-12 (123)
- synthetic construct (85)
- Bos taurus (83)
- Listeria monocytogenes EGD-e (58)
- Camelus dromedarius (57)
- More...

Taxonomy

- Eukaryota (3,192)
- Bacteria (1,723)
- Duplodnaviria (899)
- Archaea (142)
- Riboviria (129)
- other sequences (85)
- Monodnaviria (7)
- Varidnaviria (6)
- unclassified bacterial viruses (3)
- unclassified sequences (2)
- More...

Experimental Method

- X-RAY DIFFRACTION (5,721)
- ELECTRON MICROSCOPY (72)
- SOLUTION NMR (35)
- ELECTRON CRYSTALLOGRAPHY (20)
- NEUTRON DIFFRACTION (15)
- POWDER DIFFRACTION (14)
- EPR (4)
- SOLID-STATE NMR (1)

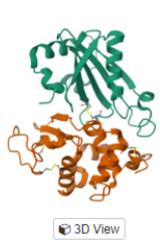
Polymer Entity Type

- Protein (5,845)
- DNA (64)
- RNA (33)
- NA-hybrid (4)

Refinement Resolution (Å)

- 0.5 - 1.0 (76)

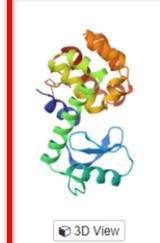
1 to 25 of 5,868 Structures Page 1 of 235 2 Sort by Score



1UUZ Download File View File

IVY:A NEW FAMILY OF PROTEIN
 Abergel, C., Lembo, F., Byrne, D., Maza, C., Claverie, J.M.
 (2007) Proc Natl Acad Sci U S A **104**: 6394

Released 2004-01-14
Method X-RAY DIFFRACTION 1.8 Å
Organisms Gallus gallus
 Pseudomonas aeruginosa
Macromolecule INHIBITOR OF VERTEBRATE LYSOZYME (protein)
 LYSOZYME C (protein)



1L58 Download File View File

ANALYSIS OF THE INTERACTION BETWEEN CHARGED SIDE CHAINS AND THE ALPHA-HELIX DIPOLE USING DESIGNED THERMOSTABLE MUTANTS OF PHAGE T4 LYSOZYME
 Nicholson, H., Matthews, B.W.
 To be published

Released 1991-10-15
Method X-RAY DIFFRACTION 1.65 Å
Organisms Tequatrovirus T4
Macromolecule T4 LYSOZYME (protein)
Unique Ligands BME



1L36 Download File View File

TOWARD A SIMPLIFICATION OF THE PROTEIN FOLDING PROBLEM: A STABILIZING POLYALANINE ALPHA-HELIX ENGINEERED IN T4 LYSOZYME
 Zhang, X.-J., Baase, W.A., Matthews, B.W.
 (1991) Biochemistry **30**: 2012-2017

Released 1991-10-15
Method X-RAY DIFFRACTION 1.7 Å
Organisms Tequatrovirus T4
Macromolecule LYSOZYME (protein)
Unique Ligands BME, CL



1L54 Download File View File

THE STRUCTURAL AND THERMODYNAMIC CONSEQUENCES OF BURYING A CHARGED RESIDUE WITHIN THE HYDROPHOBIC CORE OF T4 LYSOZYME
 Daopin, S., Matthews, B.W.
 (1991) Biochemistry **30**: 11521-11529

Released 1991-10-15



Structure Summary

3D View

Annotations

Experiment

Sequence

Genome

Versions

Display Files

Download Files

Data API

Biological Assembly 1



3D View: [Structure](#) | [1D-3D View](#) |
[Electron Density](#) | [Validation Report](#) |
[Ligand Interaction](#)

Global Symmetry: Asymmetric - C1
Global Stoichiometry: Monomer - A1

[Find Similar Assemblies](#)

Biological assembly 1 assigned by authors.

Macromolecule Content

- Total Structure Weight: 18.71 kDa
- Atom Count: 1,462
- Modelled Residue Count: 164
- Deposited Residue Count: 164
- Unique protein chains: 1

1L58

ANALYSIS OF THE INTERACTION BETWEEN CHARGED SIDE CHAINS AND THE ALPHA-HELIX DIPOLE USING DESIGNED THERMOSTABLE MUTANTS OF PHAGE T4 LYSOZYME

PDB DOI: <https://doi.org/10.2210/pdb1L58/pdb>Classification: **HYDROLASE (O-GLYCOSYL)**

Organism(s): Tequatrovirus T4

Mutation(s): No

Deposited: 1991-05-06 Released: 1991-10-15

Deposition Author(s): Nicholson, H., Matthews, B.W.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

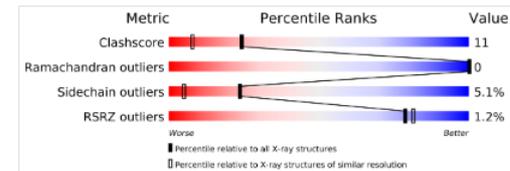
Resolution: 1.65 Å

R-Value Observed: 0.167

wwPDB Validation

3D Report

Full Report

This is version 1.3 of the entry. See complete [history](#).

Literature

Download Primary Citation

[Nicholson, H., Becktel, W., Matthews, B.W.](#)*To be published.*

Macromolecules

Find similar proteins by: [Sequence](#) (by identity cutoff) | [3D Structure](#)

Entity ID: 1

Molecule	Chains	Sequence Length	Organism	Details	Image

Macromolecules

Find similar proteins by: [Sequence](#) (by identity cutoff) | [3D Structure](#)

Entity ID: 1

Molecule	Chains	Sequence Length	Organism	Details	Image
T4 LYSOZYME	A	164	Tequatrovirus T4	Mutation(s): 0 Gene Names: E EC: 3.2.1.17	

UniProt

Find proteins for [P00720](#) (*Enterobacteria phage T4*)Explore [P00720](#)Go to UniProtKB: [P00720](#)

Entity Groups

Sequence Clusters

[30% Identity](#)[50% Identity](#)[70% Identity](#)[80% Identity](#)[95% Identity](#)[100% Identity](#)

UniProt Group

[P00720](#)

Protein Feature View

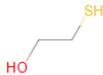
Expand

Reference Sequence

[1L58_1](#)

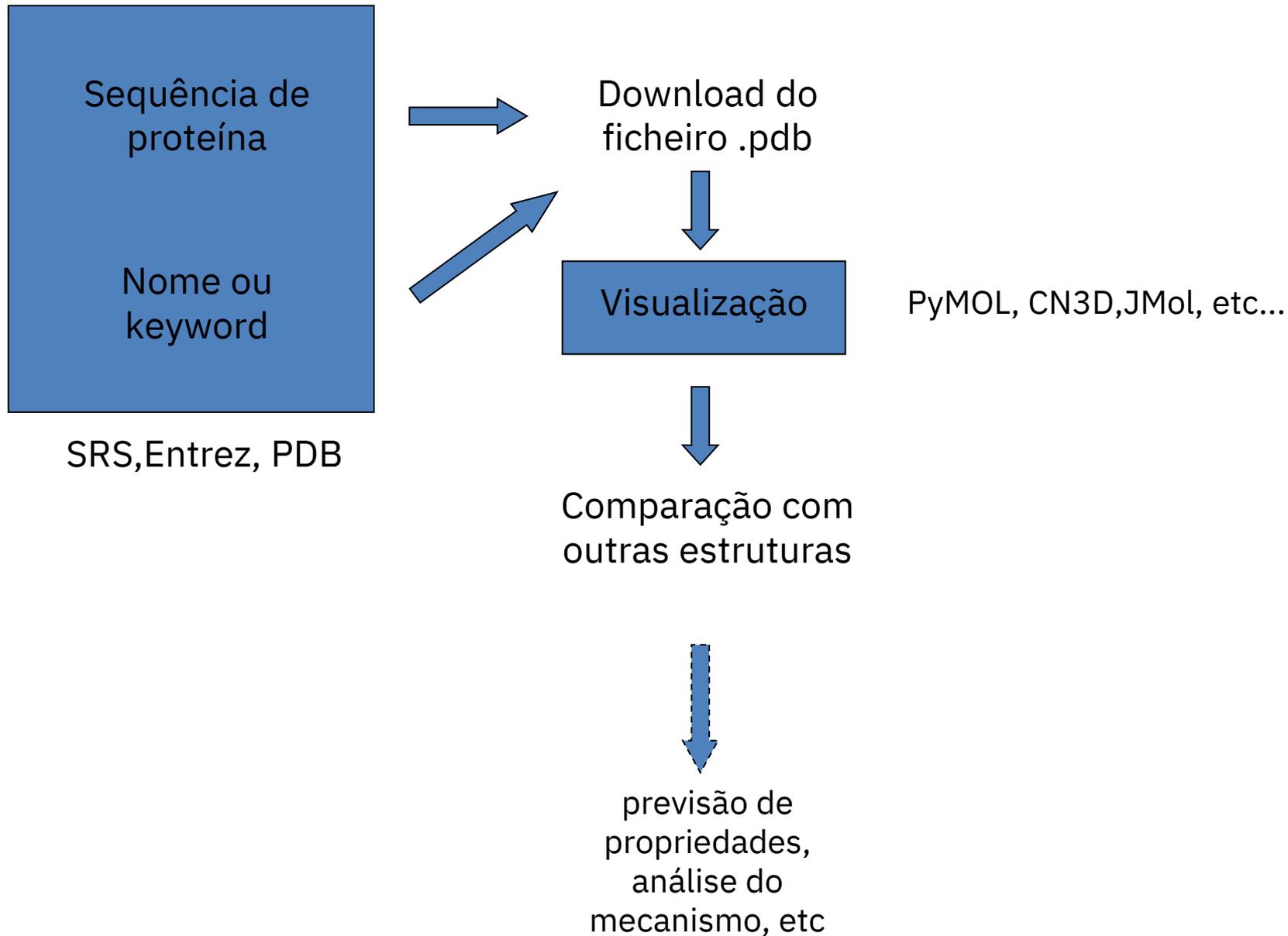
Small Molecules

Ligands **1 Unique**

ID	Chains	Name / Formula / InChI Key	2D Diagram	3D Interactions
BME Query on BME	B [auth A]	BETA-MERCAPTOETHANOL C ₂ H ₆ O S DGWVWUTYPXICAM-UHFFFAOYSA-N		Ligand Interaction

[Download Ideal Coordinates CCD File](#)[Download Instance Coordinates](#)

Visualização de estruturas moleculares



Software para visualização molecular

Aplicações de software que permitem a visualização de ficheiros de estrutura molecular (ficheiros PDB e outros formatos), permitindo a análise e cálculo de propriedades moleculares e a comparação de diferentes estruturas

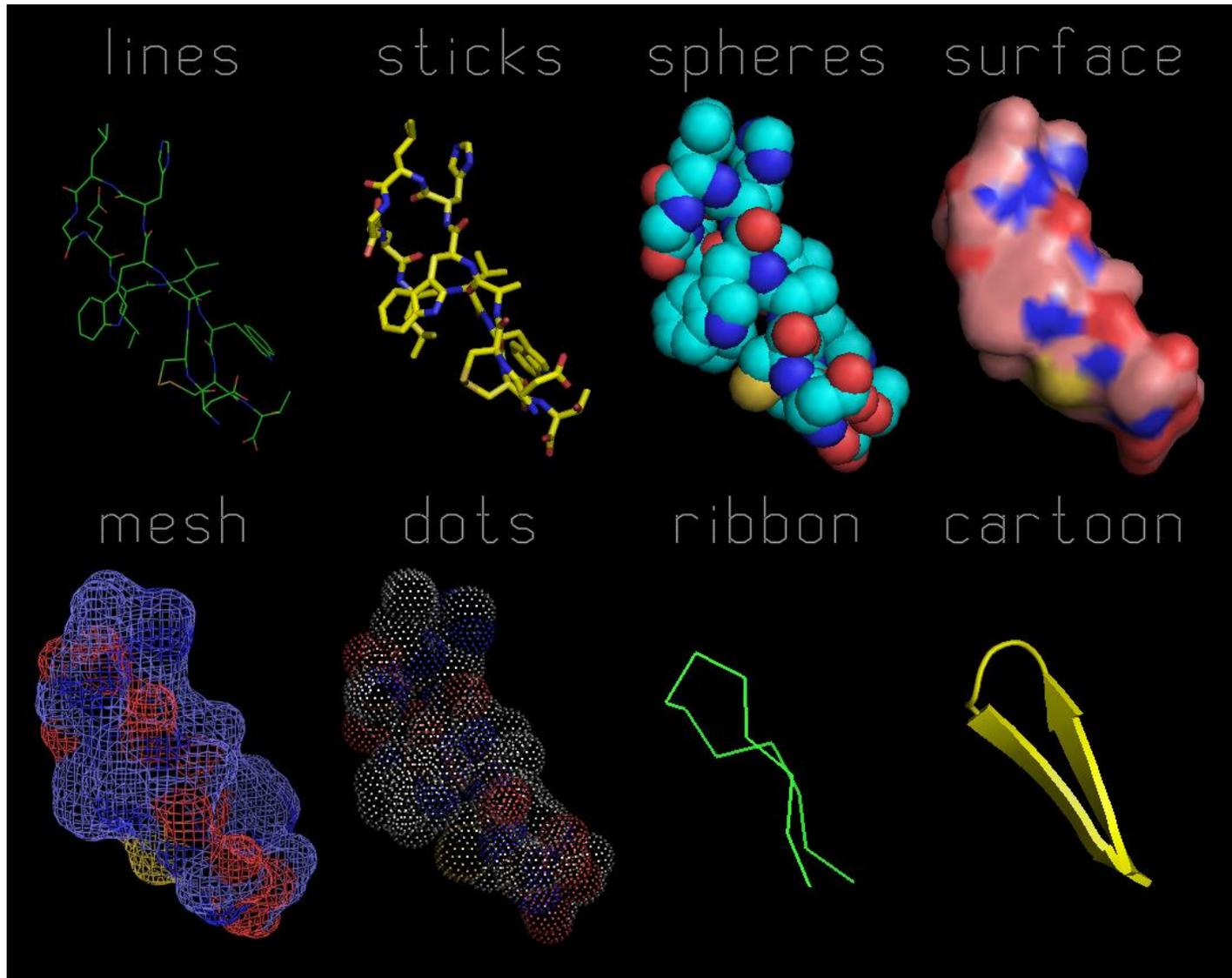
Instaláveis:

- **PyMOL:** <http://www.pymol.org>
- **ICM :** <http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml>
- **QuteMol:** <http://qutemol.sourceforge.net/>
- **SwissPDB viewer:** <http://www.expasy.org/spdbv/>

On-line:

- **nglviewr:** <http://nglviewer.org/>
- **ICMJS:** <http://www.molsoft.com>
- **Jmol/JS Mol:** <http://jmol.sourceforge.net/>

Modos de representação de estruturas



Bases de dados de pequenas moléculas

- Bases de dados que contêm estruturas de milhares ou milhões de pequenas moléculas , na sua maioria orgânicas
- Ferramenta essencial para o *screening* virtual
- Contêm uma variedade de *descritores* e propriedades das moléculas, umas experimentais, outras calculadas.

Bases de datos de pequeñas moléculas

- PubChem - <http://pubchem.ncbi.nlm.nih.gov>
- DrugBank - <http://www.drugbank.ca>
- ChEMBL - <https://www.ebi.ac.uk/chembl>
- ZINC (purchasable compounds) - <http://zinc.docking.org>
- TCM (traditional chinese medicine) - <http://tcm.cmu.edu.tw>
- CSD (Cambridge Structural Database) - <http://webcsd.ccdc.cam.ac.uk>
- ChemDB (database+tools) - <http://www.chemdb.com>
- MOLE DB (molecular descriptors) - http://michem.disat.unimib.it/mole_db

PubChem



- Conjunto de bases de dados mantido pelo National Institute for Biotechnology Information (NCBI), parte da rede dos National Institutes of Health (NIH), nos EUA.
 - Três bases de dados centrais contendo substâncias, compostos químicos e ensaios de actividade para diferentes sistemas biológicos
 - Contem moléculas com menos de 1000 átomos e menos de 1000 ligações químicas
 - 3 bases de dados interligadas:
 - Compound (**111,050,847**)
 - Substance (**277,194,318**)
 - Bioassay (**1,391,562**)
- } 21/11/2021
- Permite pesquisa por estrutura, similaridade, etc...

PubChem Data (21/11/2021)

Data Collection	Live Count	Description
Compounds	111,050,847	Unique chemical structures extracted from contributed PubChem Substance records
Substances	277,194,318	Information about chemical entities provided by PubChem contributors
BioAssays	1,391,562	Biological experiments provided by PubChem contributors
Bioactivities	292,633,795	Biological activity data points reported in PubChem BioAssays
Genes	103,715	Gene targets tested in PubChem BioAssays and those involved in PubChem Pathways
Proteins	96,561	Protein targets tested in PubChem BioAssays and those involved in PubChem Pathways
Taxonomy	531,241	Organisms of targets tested in PubChem BioAssays and those involved in PubChem Pathways
Pathways	238,597	Interactions between chemicals, genes, and proteins
Literature	33,307,005	Scientific publications with links in PubChem
Patents	28,543,965	Patents with links in PubChem
Data Sources	824	Organizations contributing data to PubChem

21/11/2021

<https://pubchemdocs.ncbi.nlm.nih.gov/statistics>

Bases de dados



- **PubChem Substance:** cada entrada nesta base de dados contém informação sobre uma *amostra química* de proveniência bem definida, que pode conter um ou mais compostos. Cada entrada possui referências cruzadas para bibliografia, ensaios biológicos, estruturas de compostos, proteínas, etc... A informação sobre substâncias é fornecida pelas organizações que contribuem para o PubChem
- **PubChem Compound:** base de estruturas químicas validadas e agrupadas por similaridade. Contém vários descritores e propriedades moleculares pré-calculados (eg: XlogP, MW) que podem ser usados para filtrar as pesquisas. Cada **substância** pode conter um ou mais compostos.
- **PubChem Bioassay:** ensaios de actividade biológicas relativos às entradas de **PubChem Substance**, contendo as descrições e resultados dos ensaios.



- Depositor-provided
- Unique Identifier: **SID**



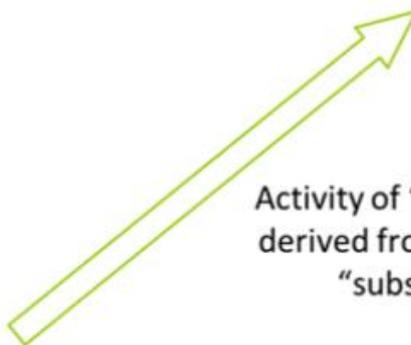
- Unique chemical structures
- Unique Identifier: **CID**

Activity of
tested
"substances"



- Biological activity test results
- Depositor-provided
- Unique Identifier: **AID**

Activity of "compounds"
derived from associated
"substances"





❖ **Validate chemical contents**

- Atoms defined/real
- Implicit hydrogen
- Functional group
- Atom valence



❖ **Normalize representation**

- Tautomer invariance
- Aromaticity detection
- Stereochemistry
- Explicit hydrogen



❖ **Calculate**

- Coordinates
- Properties
- Descriptors



❖ **Detect components**

- Isolate covalent units
- Neutralize (by $\pm H^+$ or e^-)
- **Reprocess**
- Detect unique components



- **Compound:** nomes, sinónimos ou keywords.
- **Substance:** nomes, sinónimos, keywords
- **Bioassay:** pesquisa de termos nas descrição do ensaio
- **Entrez:** pesquisar usando as ferramentas do NCBI
- **Estrutura:** pesquisar por similaridade de estrutura
- **Ferramentas de análise:** SAR maps, tabelas customizáveis, etc...

Explore Chemistry

Quickly find chemical information from authoritative sources

Try covid-19 aspirin EGFR C9H8O4 57-27-2 C1=CC=C(C=C1)C=O InChI=1S/C3H6O/c1-3(2)4/h1-2H3

Use Entrez Compounds Substances BioAssays



Draw Structure



Upload ID List



Browse Data



Periodic Table

Explore Chemistry

Quickly find chemical information from authoritative sources

aspirin

Compound	Gene	Taxonomy
aspirin	asporin	Aspergillus viridinutans
Aspirine	pirin	
Aspirin sodium	akirin	
Aspirin anhydride	akirin 1	
Aspirin DL-lysine	akirin 2	
Aspirin methyl ester	akirin 1 pseudogene	
Aspirin calcium	Aspn	
Aspirin acetaminophen ester	agrin	
Aspirin-alanine	HASPIN	
Aspirin copper	Ankyrin	

111M Compounds 277M Substances

[See More Statistics >](#)

824 Data Sources

[Explore Data Sources >](#)

- Compounds (120)
- Substances (615)
- Pathways (25)
- BioAssays (1,998)
- Literature (69,504)
- Patents (2,134)

Searching chemical names and synonyms including IUPAC names and InChIKeys across the compound collection. Note that annotations text from compound summary pages is not searched. [Read More...](#)

120 results | Filters | SORT BY: Relevance

Download

Search in Entrez

ACTIONS ON RESULTS WITH ID TYPE: Compounds

- Push to Entrez
- Save for Later
- Linked Data Sets

 **Aspirin; ACETYSALICYLIC ACID; 50-78-2; 2-Acetoxybenzoic Acid; 2-(Acetyloxy)Benzoic Acid; ...**

Compound CID: 2244
MF: C₉H₈O₄ MW: 180.16g/mol
IUPAC Name: 2-acetoxybenzoic acid
Isomeric SMILES: CC(=O)OC1=CC=CC(=O)O
InChIKey: BSYNRYMUTXBXSQ-UHFFFAOYSA-N
InChI: InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12)
Create Date: 2004-09-16

Summary | Similar Structures Search | Related Records | PubMed (MeSH Keyword)

 **Aspirin Calcium; Calcium Aspirin; Ascal; 69-46-5; Solprin; ...**

Compound CID: 6247
MF: C₁₈H₁₄CaO₈ MW: 398.4g/mol
IUPAC Name: calcium;2-acetoxybenzoate
Isomeric SMILES: CC(=O)OC1=CC=CC(=O)[O-].CC(=O)OC1=CC=CC(=O)[O-].[Ca+2]
InChIKey: KRALOLGXHLZTCW-UHFFFAOYSA-L
InChI: InChI=1S/2C9H8O4.Ca/c2*1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2*2-5H,1H3,(H,11,12);/q;+2/p-2
Create Date: 2005-08-08

PubChem Compound

aspirin - PubChem C x

https://www.ncbi.nlm.nih.gov/pccompound/?term=aspirin

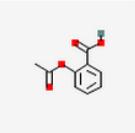
Apps Enzymology Piano Music Production Bioinformatics Databases Bioinformatics T... Misc Programming D pmarcel Other bookmarks

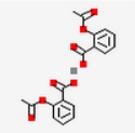
NCBI Resources How To Sign in to NCBI

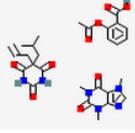
PubChem Compound PubChem Compound aspirin Search PubChem Compound. Use up and down arrows to choose an item from the autocomplete. Save search Limits Advanced Help

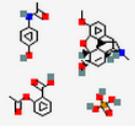
Display Settings: Summary, 20 per page, Sorted by Default order Send to: Filters: Manage Filters

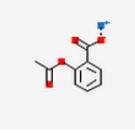
Results: 1 to 20 of 88 << First < Prev Page 1 of 5 Next > Last >>

1.  [aspirin; ACETYL SALICYLIC ACID; 2-Acetoxybenzoic acid ...](#)
MW: 180.157420 g/mol MF: C₉H₈O₄
IUPAC name: 2-acetoxybenzoic acid
CID: 2244
[Summary](#) [Similar Compounds](#) [Same Parent, Connectivity](#) [Mixture/Component Compounds](#) [PubMed \(MeSH Keyword\)](#) [Active in 125 of 3501 BioAssays](#)

2.  [Calcascorbin; Calcium aspirin; Calcascorbate ...](#)
MW: 398.376960 g/mol MF: C₁₈H₁₄CaO₈
IUPAC name: calcium;2-acetoxybenzoate
CID: 6247
[Summary](#) [Similar Compounds](#) [Same Parent, Connectivity](#) [Mixture/Component Compounds](#) [PubMed \(MeSH Keyword\)](#)

3.  [Axotal; BUTALBITAL ASPIRIN AND CAFFEINE; BUTAL COMPOUND ...](#)
MW: 598.604360 g/mol MF: C₂₈H₃₄N₆O₉
IUPAC name: 2-acetoxybenzoic acid;5-(2-methylpropyl)-5-prop-2-enyl-1,3...
CID: 24847961
[Summary](#) [Similar Compounds](#) [Mixture/Component Compounds](#) [PubMed \(MeSH Keyword\)](#)

4.  [CODEINE, ASPIRIN, APAP FORMULA NO. 2; CODEINE, ASPIRIN, APAP FORMULA NO. 3; CODEINE, ASPIRIN, APAP FORMULA NO. 4 ...](#)
MW: 728.679402 g/mol MF: C₃₅H₄₁N₂O₁₃P
IUPAC name: (4R,4aR,7S,7aR,12bS)-9-methoxy-3-methyl-2,4,4a,7,7a,13-hexah...
CID: 24847798
[Summary](#) [Similar Compounds](#) [Mixture/Component Compounds](#)

5.  [Aspirin sodium; Sodium aspirin; Sodium acetylsalicylate ...](#)
MW: 202.139249 g/mol MF: C₉H₇NaO₄
IUPAC name: sodium;2-acetoxybenzoate
CID: 23666729
[Summary](#) [Similar Compounds](#) [Same Parent, Connectivity](#) [Mixture/Component Compounds](#)

Actions on your results

-  **BioActivity Analysis**
Analyze the BioActivities of the compounds
-  **Structure Clustering**
Cluster structures based on structural similarity
-  **Structure Download**
Download the structures in various formats
-  **Pathways**
Analyze pathways containing the compounds

Refine your results • What's this?

Chemical Properties
Rule of 5 (22)

BioActivity Experiments

- BioAssays, Active (13) 
- BioAssays, Tested (19) 
- Protein 3D Structures (3)
 - Human Transthyretin (ttr) Complexed With Diflunisal (1)

BioMedical Annotation

- Pharmacological Actions (25)
 - Anti-Inflammatory Agents, Non-Steroidal (21)
- BioSystems (3)

Depositor Category

- Biological Properties (75)
- Chemical Vendors (62)
- Journal Publishers (32)

PubChem CID 2244

Structure





2D 3D Crystal

[Find Similar Structures](#)

Chemical Safety



Irritant

[Laboratory Chemical Safety Summary \(LCSS\) Datasheet](#)

Molecular Formula $C_9H_8O_4$ or $CH_3COOC_6H_4COOH$ or $HC_9H_7O_4$

Synonyms

aspirin
 ACETYSALICYLIC ACID
 50-78-2
 2-Acetoxybenzoic acid
 2-(Acetyloxy)benzoic acid

[More...](#)

Molecular Weight 180.16

Dates

Modify	Create
2021-11-20	2004-09-16

Aspirin or acetylsalicylic acid is perhaps the most commonly used analgesic and antipyretic medication worldwide, having been in clinical use for over 100

[Cite](#) [Download](#)

CONTENTS

- Title and Summary
- 1 Structures
- 2 Names and Identifiers
- 3 Chemical and Physical Properties
- 4 Spectral Information
- 5 Related Records
- 6 Chemical Vendors
- 7 Drug and Medication Information
- 8 Pharmacology and Biochemistry
- 9 Use and Manufacturing
- 10 Identification
- 11 Safety and Hazards
- 12 Toxicity
- 13 Associated Disorders and Diseases
- 14 Literature
- 15 Patents

Explore Chemistry

Quickly find chemical information from authoritative sources

Try covid-19 aspirin EGFR C9H8O4 57-27-2 C1=CC=C(C=C1)C=O InChI=1S/C3H6O/c1-3(2)4/h1-2H3

Use Entrez Compounds Substances BioAssays



Draw Structure



Upload ID List



Browse Data



Periodic Table

111M Compounds 277M Substances 293M Bioactivities 33M Literature 29M Patents

[See More Statistics >](#)

824 Data Sources

[Explore Data Sources >](#)

Isomeric SMILES: CC(=O)OC1=CC=CC=C1C(=O)O
InChIKey: BSYNRYMUTXBXSQ-UHFFFAOYSA-N
InChI: InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12)
Create Date: 2004-09-16

Summary Similar Structures Search Related Records PubMed (MeSH Keyword)

Compounds (120)

Substances (615)

Pathways (25)

BioAssays (1,998)

Literature (69,504)

Patents (2,134)

Searching chemical names and synonyms in the substance records submitted by PubChem's contributors. [Read More...](#)

615 results

Filters

SORT BY Relevance

Download

Search in Entrez

ACTIONS ON RESULTS WITH ID TYPE:

- Substances
- Compounds

Push to Entrez

Save for Later

Linked Data Sets



Aspirin; ACETYLSALICYLIC ACID; 2-Acetoxybenzoic Acid; Acetylsalicylate; Acylpyrin; ...

Substance SID: 49854366 Compound CID: 2244

Data Source: LeadScope External ID: LS-143

Data Source Category: Legacy Depositors; Subscription Services

Deposit Date: 2008-07-09 Last Modified Date: 2011-04-18



Aspirin; ACETYLSALICYLIC ACID; 2-Acetoxybenzoic Acid; Acetylsalicylate; Acylpyrin; ...

Substance SID: 319061566 Compound CID: 2244

Data Source: ToxPlanet External ID: ToxPlanet-NTAtNzgtMjlyNDQ=

Data Source Category: Subscription Services

Deposit Date: 2016-11-25 Last Modified Date: 2019-02-10

SID 319061566 - PubChem

pubchem.ncbi.nlm.nih.gov/substance/319061566

Chrome Ualg Thank you Paulo! - C... NAS Tools Resources Covid-19 Modelling Dali Server - Job que... Other bookmarks Reading list

SUBSTANCE RECORD

2-(Acetyloxy)benzoic acid

Cite Download

CONTENTS

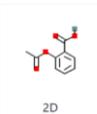
Title and Summary

1 2D Structure

2 Identity

3 Related Records

4 Information Sources

PubChem SID	319061566
Structure	 <p>2D</p>
Source	ToxPlanet
External ID	ToxPlanet-NTAtNzgtMjlyNDQ=
Source Category	Subscription Services
Version	1 Revision History
Status	Live
Related Compounds	PubChem CID CID 2244 (Aspirin)
Dates	Available Deposit 2016-11-25 2016-11-25

Please note that the substance record is presented as provided to PubChem by the source (depositor). For standardized chemical structure and/or annotation information, please visit the summary page for [CID 2244](#).

PubChem Substance

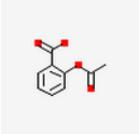
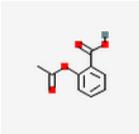
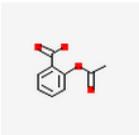
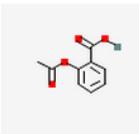
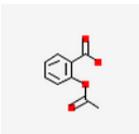
aspirin - PubChem St x
https://www.ncbi.nlm.nih.gov/pcsubstance/?term=aspirin

NCBI Resources How To Sign in to NCBI

PubChem Substance PubChem Substance aspirin Search Save search Limits Advanced Help

Display Settings: Summary, 20 per page, Sorted by Default order Send to: Filters: Manage Filters

Results: 1 to 20 of 547 << First < Prev Page 1 of 28 Next > Last >>

-  **aspirin: ACETYLSALICYLIC ACID; Ecotrin ...**
Source: [LeadScope \(LS-143\)](#)
SID: 49854366 [CID: 2244]
[Summary](#) [PubChem Same Compound](#) [Same Parent, Connectivity](#) [PubMed \(MeSH Keyword\)](#)
-  **aspirin: ACETYLSALICYLIC ACID; Ecotrin ...**
Source: [Comparative Toxicogenomics Database \(D001241\)](#)
SID: 53788943 [CID: 2244]
[Summary](#) [PubChem Same Compound](#) [Same Parent, Connectivity](#) [PubMed \(MeSH Keyword\)](#)
-  **aspirin: ACETYLSALICYLIC ACID; Ecotrin ...**
Source: [Therapeutic Targets Database \(DAP000843\)](#)
SID: 134338122 [CID: 2244]
[Summary](#) [PubChem Same Compound](#) [Same Parent, Connectivity](#) [PubMed \(MeSH Keyword\)](#)
-  **aspirin: ACETYLSALICYLIC ACID; Ecotrin ...**
Source: [Human Metabolome Database \(HMDB01879\)](#)
SID: 126524194 [CID: 2244]
[Summary](#) [PubChem Same Compound](#) [Same Parent, Connectivity](#) [PubMed \(MeSH Keyword\)](#)
-  **aspirin: ACETYLSALICYLIC ACID; Ecotrin ...**
Source: [ChemIDplus \(0000050782\)](#)
SID: 134971785 [CID: 2244]
[Summary](#) [PubChem Same Compound](#) [Same Parent, Connectivity](#) [PubMed \(MeSH Keyword\)](#)
-  **aspirin: ACETYLSALICYLIC ACID; Ecotrin ...**

Actions on your results

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

Refine your results • What's this?

Chemical Properties
Rule of 5 (289)

BioActivity Experiments

- BioAssays, Active (13)
- BioAssays, Tested (42)
- Protein 3D Structures (38)
 - Structural Basis Of The Prevention Of Nsaid-induced Damage Of The Gastrointestinal Tract By C-terminal Half (c-lobe) Of Bovine Colostrum Protein Lactoferrin: Binding And Structural Studies Of The C-lobe Complex With Aspirin (10)

BioMedical Annotation

- Pharmacological Actions (361)
 - Anti-Inflammatory Agents, Non-Steroidal (327)
- BioSystems (1)
- Depositor Category
- Biological Properties (156)

Compounds (120) Substances (615) Pathways (25) **BioAssays (1,998)** Literature (69,504) Patents (2,134)

Searching descriptions and metadata of bioassay records submitted by PubChem's contributors. [Read More...](#)

1,998 results

Filters

SORT BY

Relevance

Download

Search in Entrez

Percentage of aspirin formed during hydrolysis in 10% human plasma at pH 7.4 at 37 degrees celsius

BioAssay AID: 15845 BioAssay Type: Literature-derived

Tested Compounds Count: 27 Tested Substances Count: 27

Data Source: ChEMBL External ID: ChEMBL875111

Data Source Category: Curation Efforts; Research and Development

Modified Date: 2018-10-12

Description: Title: Evaluation of glycolamide esters and various other esters of aspirin as true aspirin prodrugs. Abstract: A series of glycolamide, glycolate, (acyloxy)methyl, alkyl, and aryl esters of acetylsalicylic acid (aspirin) were synthesized and evaluated as potential prodrug forms of aspirin. N,N-Disubstituted glycolamide esters were found to be rapidly hydrolyzed in human plasma, resulting in the formation of aspirin as well as the corresponding salicylate esters. These in turn hydrolyzed rapidly to salicylic acid. The largest amount of aspirin formed from the esters were 50 and 55% in case of the N,N-dimethyl- and N,N-diethylglycolamide esters, respectively. Similar results were obtained in blood with the N,N-dimethyl- and N,N-diethylglycolamide esters. Unsubstituted and monosubstituted glycolamide esters as well as most other esters previously suggested to be aspirin prodrugs were shown to hydrolyze exclusively to the corresponding salicylic acid esters. Lipophilicity parameters and water solubilities of the esters were determined, and structural factors favoring ester prodrug hydrolysis at the expense of deacetylation to yield salicylate ester are discussed. The properties of some N,N-disubstituted glycolamide esters of aspirin are highlighted with respect to their use as potential aspirin prodrugs.

ACTIONS ON RESULTS WITH ID TYPE:

- BioAssays
- Substances
- Compounds

Push to Entrez

Save for Later

Linked Data Sets

The percentage of aspirin formed during hydrolysis in 100% human plasma at pH 7.4 at 37 degrees celsius

BioAssay AID: 15847 BioAssay Type: Literature-derived

Tested Compounds Count: 1 Tested Substances Count: 1

Data Source: ChEMBL External ID: ChEMBL629204

Data Source Category: Curation Efforts; Research and Development

Browser tabs: AID 15845 - Percentage of aspirin fo x

Address bar: pubchem.ncbi.nlm.nih.gov/bioassay/15845

Browser extensions: Chrome, Ualg, Thank you Paulo! - C..., NAS, Tools, Resources, Covid-19, Modelling, Dali Server - Job que..., Other bookmarks, Reading list

BIOASSAY RECORD

Percentage of aspirin formed during hydrolysis in 10% human plasma at pH 7.4 at 37 degrees celsius

PubChem AID	15845
Primary Citation	Evaluation of glycolamide esters and various other esters of aspirin as true aspirin prodrugs [PMID: 2918521]
Source	ChEMBL
External ID	CHEMBL875111
Tested Substances	All (27) Data Table
Tested Compounds	All (27)
Version	4.1 Revision History
Status	Live
Dates	Modify: 2018-10-12 Deposit: 2010-05-21

This bioassay record (AID 15845) reports results from the above primary citation. Additional data from the same publication are reported in a total of [19 BioAssay records](#) in PubChem.

[PubChem](#)

[Cite](#) [Download](#)

CONTENTS

- Title and Summary
- 1 Description
- 2 Comment
- 3 Result Definitions
- 4 Data Table
- 5 Target
- 6 Entrez Crosslinks
- 7 Identity
- 8 Same-Publication BioAssays
- 9 BioAssay Annotations
- 10 Information Sources

1 Description

Title: Evaluation of glycolamide esters and various other esters of aspirin as true aspirin prodrugs.

Abstract: A series of glycolamide, glycolate, (acyloxy)methyl, alkyl, and aryl esters of acetylsalicylic acid (aspirin) were synthesized and evaluated as potential prodrug forms of aspirin. N,N-Disubstituted glycolamide esters were found to be rapidly hydrolyzed in human plasma, resulting in the formation of aspirin as well as the corresponding salicylate esters. These in turn hydrolyzed rapidly to salicylic acid. The largest amount of aspirin formed from the esters were 50 and 55% in case of the N,N-dimethyl- and N,N-diethylglycolamide esters, respectively. Similar results were obtained in blood with the N,N-dimethyl- and N,N-diethylglycolamide esters. Unsubstituted and monosubstituted glycolamide esters as well as most other esters previously suggested to be aspirin prodrugs were shown to hydrolyze exclusively to the corresponding salicylic acid esters. Lipophilicity parameters and water solubilities of the esters were determined, and structural factors favoring ester prodrug hydrolysis at the expense of deacetylation to yield salicylate ester are discussed. The properties of some N,N-disubstituted glycolamide esters of aspirin are highlighted with respect to their use as potential aspirin prodrugs.

[PubChem](#)

PubChem BioAssay

AID 444512 - PubChem x PubChem PC3D View x

pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=444512&loc=ea_ras

NCBI

PubChem BioAssay

Search

Limits Advanced search Help

SHARE

BioAssay: [AID 444512](#)

CSV ASN.1 XML

>> Additional Info & Links

Antiplatelets aggregatory activity in human platelets rich plasma assessed as inhibition of collagen-induced platelets aggregation by aggregometry

Aspirin prodrugs and related nitric oxide releasing compounds hold significant therapeutic promise, but they are hard to design because aspirin esterification renders its acetate group very susceptible to plasma esterase mediated hydrolysis. Isosorbide-2-aspirinate-5-salicylate is a true aspirin prodrug in human blood because it can be effectively hydrolyzed to aspirin upon interaction with [more ..](#)

Table of Contents

- [BioActive Compounds](#)
- [Description](#)
- [Comment](#)
- [Categorized Comment](#)
- [Result Definitions](#)
- [Data Table \(Concise\)](#)

AID: [444512](#)

Data Source: ChEMBL (595690)

Depositor Category: Literature, Extracted

BioAssay Version: 5.1

Deposit Date: 2010-07-08

Modify Date: 2013-07-13

Data Table (Complete): [Active](#) [All](#)

BioActive Compounds: 3

[BioActivity Summary](#)

[Structure-Activity Analysis](#)

[Structure Clustering](#)

Follow us on

[f](#) [t](#) [g+](#) [RSS](#) [m](#)

Tested Compounds

All(5)

Active(3)

Unspecified(2)

Tested Substances

All(5)

Active(3)

Unspecified(2)

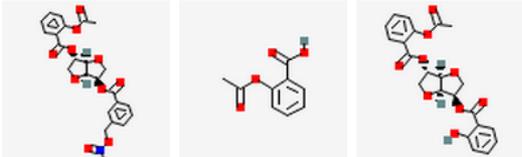
Links

[PubMed \(1\)](#)

[Taxonomy \(1\)](#)

Related BioAssays

[Activity Overlap \(105\)](#)



PubChem – Pesquisa por “Tag”

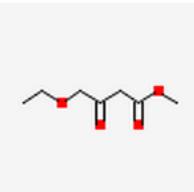
0:500[mw] 0:5[hbdc] 0:10[hbac] -5:5[logp]

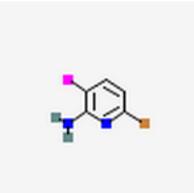
NCBI Resources How To Sign in to NCBI

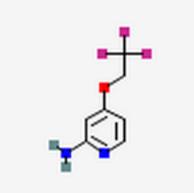
PubChem Compound PubChem Compound 0:500[mw] 0:5[hbdc] 0:10[hbac] -5:5[logp] Search Help

Display Settings: Summary, 20 per page, Sorted by Default order Send to: Filters: Manage Filters

Results: 1 to 20 of 34559871 Lipinski's rule of 5 << First < Prev Page 1 of 1727994 Next > Last >>

1.  [Methyl 4-ethoxy-3-oxobutanoate; AK141825; 415678-65-8](#)
MW: 160.167780 g/mol MF: C₇H₁₂O₄
IUPAC name: methyl 4-ethoxy-3-oxobutanoate
CID: 54303951
[Summary](#)

2.  [6-bromo-3-iodopyridin-2-amine; AK142103; 1245643-34-8](#)
MW: 298.907130 g/mol MF: C₅H₄BrIN₂
IUPAC name: 6-bromo-3-iodopyridin-2-amine
CID: 52987942
[Summary](#)

3.  [AK138368; 4-\(2,2,2-Trifluoroethoxy\)pyridin-2-amine; 1379361-82-6](#)
MW: 192.138490 g/mol MF: C₇H₇F₃N₂O
IUPAC name: 4-(2,2,2-trifluoroethoxy)pyridin-2-amine
CID: 15724964
[Summary](#)

Actions on your results

- BioActivity Analysis
Analyze the BioActivities of the compounds
- Structure Clustering
Cluster structures based on structural similarity
- Structure Download
Download the structures in various formats
- Pathways
Analyze pathways containing the compounds

Refine your results

- What's this?

Chemical Properties

Rule of 5 (34,559,871)

BioActivity Experiments

BioAssays, Probes (142)

Explore Chemistry

Quickly find chemical information from authoritative sources

Try covid-19 aspirin EGFR C9H8O4 57-27-2 C1=CC=C(C=C1)C=O InChI=1S/C3H6O/c1-3(2)4/h1-2H3

Use Entrez Compounds Substances BioAssays



Draw Structure



Upload ID List



Browse Data



Periodic Table

111M Compounds

277M Substances

293M Bioactivities

33M Literature

29M Patents

824 Data Sources

[See More Statistics >](#)

[Explore Data Sources >](#)

PubChem

pubchem.ncbi.nlm.nih.gov/#draw=true

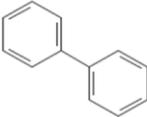
NIH National Library of Medicine
National Center for Biotechnology Information

PubChem

DRAW STRUCTURE

SMILES **C1=CC=CC(=C1)C2=CC=CC=C2**

SMILES



Export MDL Molfile Done

Hydrogen Keep AsIs Help

Import Choose File No file chosen

Search for This Structure

Detailed description: The image shows a screenshot of the PubChem web interface. A 'DRAW STRUCTURE' dialog box is open, displaying the SMILES string 'C1=CC=CC(=C1)C2=CC=CC=C2' in a red box. A red arrow points from the word 'SMILES' to this box. The dialog also shows a chemical structure of biphenyl (two benzene rings connected by a single bond) and a periodic table. The browser's address bar shows 'pubchem.ncbi.nlm.nih.gov/#draw=true' and the NIH logo is visible in the background.

A.

Bioassay Data Source Name	Bioassay count	Substance count
BioAssay Data Deposited by NIH MLPPCN and MLSCN		
NCGC (NIH)	485	398,461
The Scripps Research Institute Molecular Screening Center	483	357,929
Burnham Center for Chemical Genomics	397	400,255
NMMLSC (University of Mexico)	230	348,231
Broad Institute of MIT and Harvard	179	334,761
Vanderbilt Screening Center for GPCRs, Ion Channels & Transporters	101	223,904
SRMLSC (Southern Research Institute)	89	226,666
Johns Hopkins Ion Channel Center	74	305,806
University of Pittsburgh Molecular Library Screening Center	70	222,637
Southern Research Specialized Biocontainment Screening Center	63	339,742
PCMD (Penn Center for Molecular Discovery)	57	226,345
Emory University Molecular Libraries Screening Center	54	370,189
Columbia University Molecular Screening Center	33	197,177
BioAssay Data Deposited by Other Sources		
ChEMBL (European Bioinformatics Institute, EBI)	446,639	551,496
DTP/NCI (NIH)	173	189,809
ChemBank (Broad Institute of Harvard & MIT/Chemical Biology)	106	5,329
SGCOxCompounds (SGC Oxford)	43	319
NINDS Approved Drug Screening Program	34	1,040
BindingDB (CARB)	20	3,285
Diabetic Complications Screening (NIDDK/JDRF)	14	1,040
EPA DSSTox (National Center for Computational Toxicology)	12	4,099
GLIDA, GPCR-Ligand Database	6	19,474
GlaxoSmithKline (GSK)	6	13,533
ProbeDB (NCBI)	5	279
MTDP (CCR, NCI, NIH)	4	99,933
IUPHAR-DB	4	104
Structural Genomics Consortium	2	28
The Genomics Institute of the Novartis Research Foundation (GNF)	1	33,364
Shanghai Institute of Organic Chemistry	1	3,073
Circadian Research, Kay Laboratory (UCSD)	1	1,279
Thermo Scientific Dharmacon RNAi Technologies	1	840
ChemBlock	1	122
CC_PMLSC	1	47
SGCS to Compounds	1	17
Total: 41	449,402	4,985,224

<http://pubchem.ncbi.nlm.nih.gov/sources/>

B.

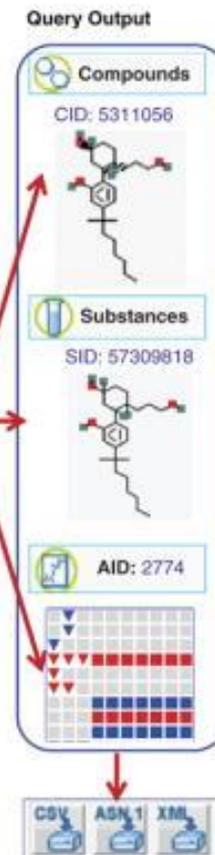


<http://pubchem.ncbi.nlm.nih.gov/>



Literature

C.



Download

ZINC database



- Base de dados de acesso livre
 - Contem cerca de 230 milhões de compostos comercialmente disponíveis (purchasable compounds) com as respectivas estruturas 3D em formatos de fácil uso para docking e screening virtual
 - Contem cerca de 740 milhões de compostos comercialmente disponíveis com estruturas 2D, que podem ser usados para pesquisar análogos.
 - Possui alguns sub-conjuntos especiais:
 - ZDD – compostos puros aprovados como fármacos pela FDA
 - ZMD – metabolitos primários
 - ZND – derivados de compostos naturais
 - ZBC – compostos biogénicos
-
-

ZINC database

ZINC

Substances

Catalogs

Tranches

Biological

More

About

ZINC15

Welcome to ZINC, a free database of commercially-available compounds for virtual screening. ZINC contains over 230 million purchasable compounds in ready-to-dock, 3D formats. ZINC also contains over 750 million purchasable compounds you can search for analogs in under a minute.

ZINC is provided by the [Irwin](#) and [Shoichet](#) Laboratories in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF). We thank [NIGMS](#) for financial support (GM71896).

To cite ZINC, please reference: Sterling and Irwin, *J. Chem. Inf. Model*, 2015 <http://pubs.acs.org/doi/abs/10.1021/acs.jcim.5b00559>. You may also wish to cite our previous papers: Irwin, Sterling, Mysinger, Bolstad and Coleman, *J. Chem. Inf. Model*, 2012 DOI: [10.1021/ci3001277](https://doi.org/10.1021/ci3001277) or Irwin and Shoichet, *J. Chem. Inf. Model*. 2005;45(1):177-82 [PDF](#), [DOI](#).

Getting Started

- [Getting Started](#)
- [What's New](#)
- [About ZINC 15 Resources](#)
- [Current Status / In Progress](#)
- [Why are ZINC results "estimates"?](#)

Explore Resources

Chemistry

[Tranches](#), [Substances](#), [3D](#)

[Representations](#), [Rings](#), [Patterns](#)

And More

[Catalogs](#), [Genes](#), [ATC Codes](#)

Ask Questions

You can use ZINC for **general** questions such as

- [How many substances in current clinical trials have PAINS patterns? \(150\)](#)
- [How many natural products have names in ZINC and are not for sale? \(9296\) get them as SMILES, names and calculated logP](#)
- [How many endogenous human metabolites are there? \(47319\) and how many of these can I buy? \(8271\) How many are FDA approved drugs? \(94\)](#)
- [How many compounds known to aggregate are in current clinical trials? \(60\)](#)
- [How many epigenetic targets have compounds known? \(53\) and Which of these substances can I buy? \(278\)](#)
- [How many ligands are there for the NMDA 1 ion channel GRIN1? \(662\) and How many of these are for sale? \(60\)](#)
- [More...](#)

ZINC15 News

- 2018-02-14 - ZINC reaches 213,235,528 purchasable leadlike 3D!
- 2018-02-13 - ZINC reaches 736,001,654 purchasable molecules 2D!
- 2018-01-14 - Klara Anu is born! Welcome Klara Anu, sister to Lisa!
- 2018-01-01 - Chinzo Dandar joins our team. Welcome Chinzo! Follow us on [twitter @chem4biology](#) [Known limitations](#) [What's new](#)

Caveat Emptor: We do not guarantee the quality of any molecule for any purpose and take no responsibility for errors arising from the use of this database. ZINC is provided in the hope that it will be useful, but you must use it at your own risk.

ZINC tranches

zinc.docking.org/tranches/ho...

Chrome Ualg Thank you Paulo! - C... NAS Tools Resources Other bookmarks Reading list

ZINC Substances Catalogs Tranches Biological More About

Rep. 2D 3D React. Standard Purch. Wait OK pH N/A Charge N/A

Molecular Weight (up to, Daltons)

	200	250	300	325	350	375	400	425	450	500	>500	Totals, by LogP
-1	29,293	204,598	784,279	1,125,069	2,321,356	854,208	300,607	128,558	99,872	86,323	5,615	5,939,778
0	142,690	1,067,035	3,992,760	5,372,590	10,975,901	3,784,188	1,767,726	775,279	606,137	558,305	3,798	29,046,409
1	376,413	3,284,847	13,196,175	17,023,840	34,876,129	12,665,806	7,279,946	3,517,752	2,892,839	2,688,060	7,987	97,809,794
2	497,750	5,391,816	25,622,912	32,914,848	67,733,100	28,989,280	19,267,814	10,563,987	9,000,177	8,721,010	20,894	208,723,588
2.5	189,326	2,643,678	14,831,118	19,486,349	40,600,693	20,281,809	15,126,147	9,325,848	8,120,159	7,879,918	21,325	138,506,270
3	108,266	2,075,334	13,281,388	18,060,096	37,030,641	22,002,857	17,838,728	12,045,788	10,696,251	10,674,949	33,982	143,848,280
3.5	48,705	1,336,320	10,135,959	14,349,999	29,671,752	21,055,698	18,737,428	13,954,286	12,511,433	12,736,846	54,896	134,593,322
4	15,100	613,109	6,131,454	8,128,568	12,531,547	15,472,307	16,892,846	14,129,429	12,864,529	13,378,049	82,058	100,238,996
4.5	1,993	170,043	2,873,064	4,632,339	7,889,352	10,959,424	12,773,295	12,356,636	11,562,431	12,208,182	113,230	75,539,989
5	94	21,765	852,691	1,919,530	3,985,092	6,416,455	8,397,678	9,021,197	8,818,548	9,321,913	139,087	48,894,050
>5	28	884	44,519	175,953	549,357	1,226,923	2,066,211	2,628,127	3,062,143	3,771,646	735,850	14,261,641
Totals, by Weight	1,409,658	16,809,429	91,746,319	123,189,181	248,164,820	143,708,955	120,448,426	88,446,887	80,234,519	82,025,201	1,218,722	997,402,117 Substances 1.7K Tranches

Acknowledgements Usage Why are ZINC results "estimates"? Terms of use Privacy policy Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests Irwin and Shoichet Labs and UC Regents.

Originally generated at 2021-11-21 13:45:55.970735 in 0.03536s on zinc.docking.org using ZINC15 0.20210303.1

ZINC substance search

zinc15.docking.org/substances/home/

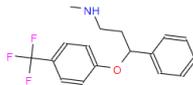
Substances

Help Examples Browse Table Subsets Shopping List

Search for Substances Search

Search Using One

CNCCC(c1ccccc1)Oc2ccc(cc2)C(F)(F)F



JSME Molecular Editor by Peter Ertl and Bruno Bienfait

Search with Nothing selected

Search Using Many

One Identifier per Line

ZINC ID, SMILES, SMARTS, InChI or Supplier Code

OR Upload a File

Choose File No file chosen

Allow Lookups

ZINC ID Structure Names Suppliers Analogs Slow!

Match Tolerance

Retired IDs Charge Scaffold Full Text Accept Multiple Results

Subsets to Check

Nothing selected

Results

Output Format Summary Table

Search Many

[ZINC](#) × +

[zinc15.docking.org/substances/resolved/](#)

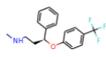
[Chrome](#) [Ualg](#) [Thank you Paulo! - C...](#) [NAS](#) [Tools](#) [Resources](#) [Covid-19](#) [Modelling](#) [Other bookmarks](#) [Reading list](#)

[ZINC](#) [Substances](#) [Catalogs](#) [Tranches](#) [Biological](#) [More](#) About

1 [/substances](#) [Filters](#) [Lookup](#)

ZINC **Summary**

ZINC1530637



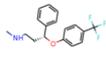
CNCCC(c1ccccc1O)C1=CC=C(C=C1)C(F)(F)F

InChIKey: Same-Scala...

fluoxetine

Tranche **DIAA**
Subsets anodyne bb fda for-sale in-stock natural-products
Purchasability In-Stock (145 vendors, 82 annotated)
Activity 18 activities from 156 observation(s)
in 5 class(es) with 17 gene(s)
Studies 74 citations
186 clinical trial(s)

ZINC1530638



CNCCC(c1ccccc1O)C1=CC=C(C=C1)C(F)(F)F

InChIKey: Same-Scala...

fluoxetine

Tranche **DIAA**
Subsets anodyne bb fda for-sale in-stock
Purchasability In-Stock (146 vendors, 80 annotated)
Activity 18 activities from 157 observation(s)
in 5 class(es) with 17 gene(s)
Studies 75 citations
186 clinical trial(s)

1 [/substances](#) [Filters](#) [Lookup](#)

Acknowledgements Usage Why are ZINC results "estimates"? Terms of use Privacy policy Supported by NIGMS via GM71896 Questions, Discussion, Bug reports, Feature requests Irwin and Shoichet Labs and UC Regents.

Originally generated at 2021-11-22 00:43:23.714411 in 0.38491s on zinc15.docking.org using ZINC15 0.20210303.1

/ substances / ZINC000001530637

ZINC1530637 (Fluoxetine)

In: anodyne bb fda for-sale in-stock natural-products

Google Wikipedia PubMed

Added	Availability	Since	Mwt	logP	Download
2004-10-06	In-Stock	2015-08-07	309.331	4.435	Download

Mol Formula	Rings	Heavy Atoms	Hetero Atoms	Fraction sp ³	Tranche
C17H18F3NO	2	22	5	0.29	DIAA

SMILES	<chem>CNCC[C@H](Oc1ccc(C(F)(F)F)cc1)c1ccccc1</chem>	Download
InChI	InChI=1S/C17H18F3NO/c1-21-12-11-16(13-5-3-2-4-6-13)22-15-9-7-14(8-10-15)17(18,19)20/h2-10,16,21H,11-12H2,1H3/t16/m/0/s1	Download
InChI Key	RTHCYVBBDHJXIQ-INIZCTEOSA-N	Download

Available 3D Representations

pH range	Net charge	H-bond donors	H-bond acceptors	TPSA	Rotatable bonds	Apolar desolvation	Polar desolvation	Download
Reference	1	1	1	25	6	9.47	-44.22	Download

Vendors (79 Total)

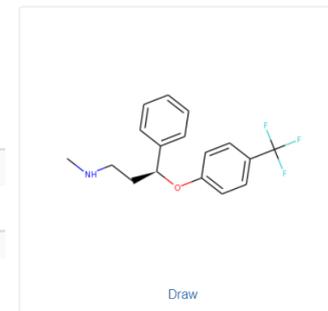
145 Items Total

AK Scientific	J10277
Chemodex Economical	F0141
Chem Scene	CS-1838
KeyOrganics	KS-1061
MedChem Express	HY-B0102A
Molport SC Economical	MolPort-001-683-482, MolPort-003-666-535, MolPort-009-194-198
Oakwood Economical	375072

Annotated Catalogs (44 Total)

82 Items Total

NIH Clinical Collection	MLS002589965
Prestwick Chemical	Prestw-511
SMDC Iconix	131498
SMDC Pharmakon	131498
Tocriscreen	0927
Ambinter	Amb17614490, Amb2608851, Amb534674, Amb6297426
BindingDB.org	30130, 50331514, 81875
CHEBI	CHEBI:86990, CHEBI:86992, CHEBI:86995, CHEBI:86997
CHEMBL20	CHEMBL1169388, CHEMBL1201082, CHEMBL1257031, CHEMBL41



🏠 / substances / ZINC000000000053

ZINC53 (Aspirin)

In: [anodyne](#) [bb](#) [fda](#) [for-sale](#) [in-stock](#) [natural-products](#)

[Google](#) [Wikipedia](#) [PubMed](#)

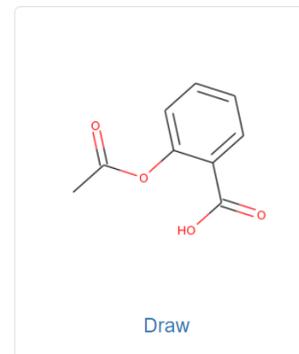
Added	Availability	Since	Mwt	logP	Download
2005-09-27	In-Stock	2015-08-07	180.159	1.31	↓

Mol Formula	Rings	Heavy Atoms	Hetero Atoms	Fraction sp ³	Tranche
C ₉ H ₈ O ₄	1	13	4	0.11	ADAA

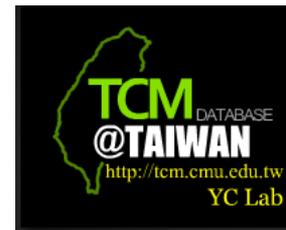
SMILES	<chem>CC(=O)Oc1ccccc1C(=O)O</chem>	📄
InChI	InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12)	📄
InChI Key	BSYNRYMUTXBXSQ-UHFFFAOYSA-N	📄

Available 3D Representations

pH range	Net charge	H-bond donors	H-bond acceptors	tPSA	Rotatable bonds	Apolar desolvation	Polar desolvation	Download
Reference	-1	0	4	66	2	6.58	-56.82	↓



Traditional Chinese Medicine (TCM)



- Contem substâncias derivadas de plantas, extractos animais e minerais
- Estruturas tri-dimensionais de compostos presentes nos extractos
- Estruturas disponíveis em formatos 2D e 3D, pré-minimizadas e prontas para usar em docking e screening virtual
- Acesso Livre

Username

Login

加入會員

- ◊ 簡介
- ◊ 最新消息
- ◊ 普通查詢
- ◊ 進階查詢
- ◊ 計算服務
- ◊ 中草藥對標
- ◊ 檔案下載
- ◊ 檔案上傳
- ◊ FAQ 問與答
- ◊ YC 實驗室
- ◊ 影片



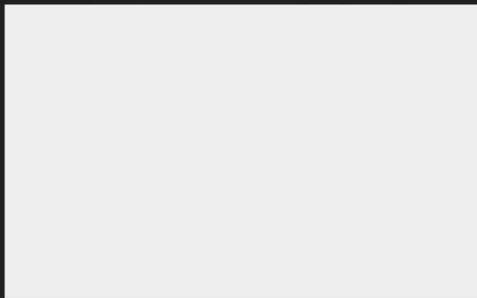
化學成份 Chemical Compound

首頁 > 中醫藥瀏覽 > 化學成份

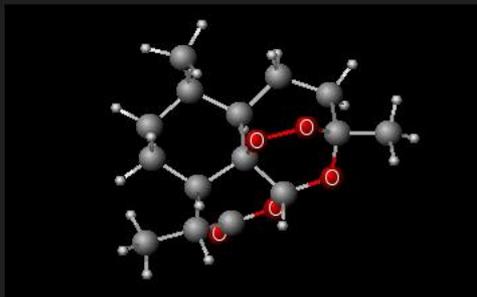
檔案下載 [2D圖](#) [MOL2](#)

Compound **artemisinin**

2D結構圖



3D結構圖

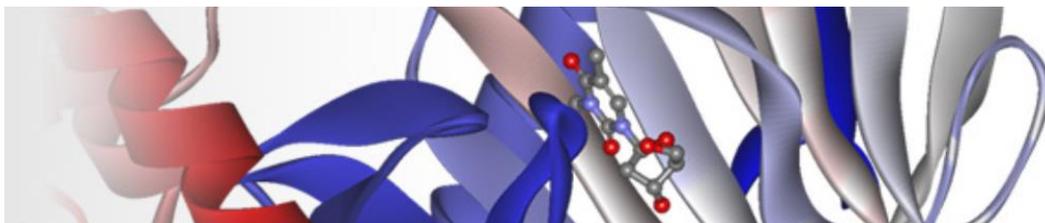


(Generate by MarvinView)

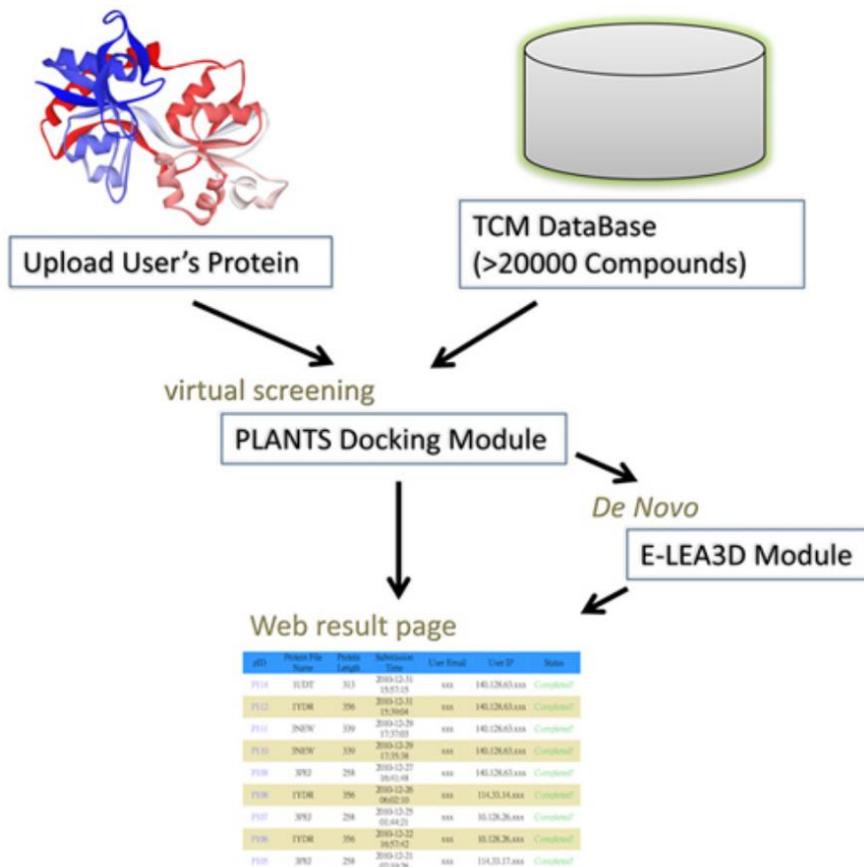
Chemical Formula	C ₁₅ H ₂₂ O ₅
Molecular Weight	282.332
Molecular Volume	197.91
ALogP	1.998
Molecular Polar Surface Area	53.99
Number of Hydrogen Bond Acceptors	0
Number of Hydrogen Bond Donors	0
Number of Rotatable Bonds	0

相關種名 [青蒿素](#) ;

檔案下載 [2D圖](#) [MOL2](#)



iScreen



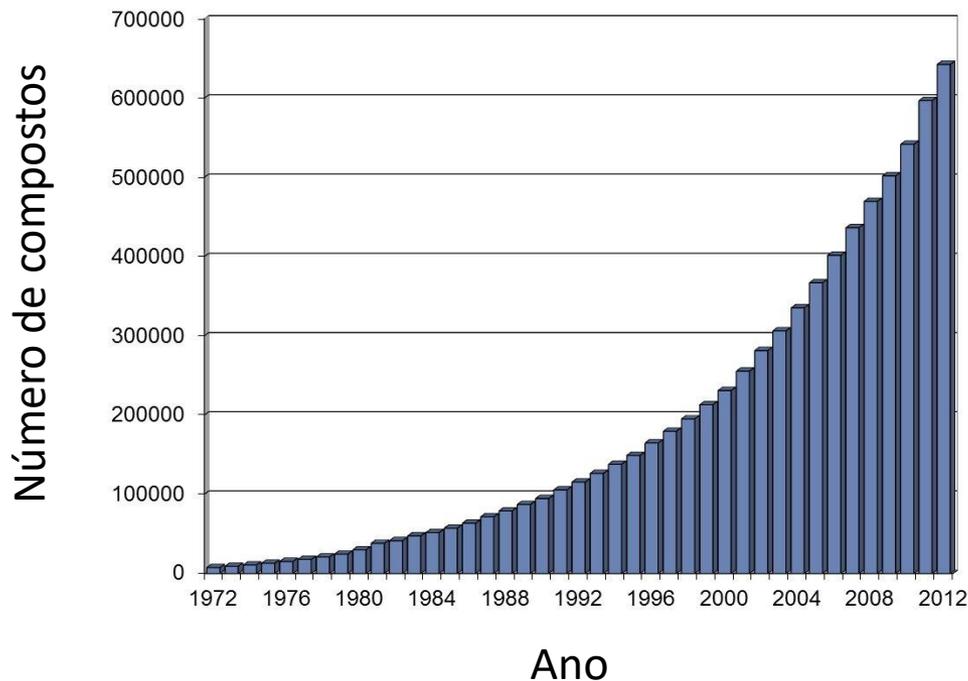
<http://iscreen.cmu.edu.tw/intro.php>

Cambridge Structural Database (CSD)



- Base de dados de estruturas **experimentais** de moléculas pequenas, mantida pelo Cambridge Crystallographic Data Center, UK
- Contem cerca de 700000 compostos **orgânicos** e **organometálicos** determinados por difracção de raios X e de neutrões em cristais individuais e pós.
- É um produto comercial, sem acesso livre. No entanto é possível obter estruturas através de pedidos individuais, desde que para fins não-comerciais.
- É vendida juntamente com o software necessário para a pesquisa, análise e visualização das estruturas (e também o software de docking GOLD)
- Não contém:
 - Polipéptidos e polissacáridos com mais de 24 unidades (ver PDB)
 - Oligonucleótidos
 - Compostos inorgânicos

CSD - Estatísticas



	Structures	%CSD
Total No. of structures	686 944	100.0
No. of different compounds	628 684	-
No. of literature sources	1 578	-
Organic structures	292 661	42.6
Transition metal present	369 682	53.8
Li – Fr or Be – Ra present	34 433	5.0
Main group metal present	41 711	6.1
3D coordinates present	643 032	93.3
Error-free coordinates	630 329	98.0†
Neutron studies	1 616	0.2
Powder diffraction studies	2 930	0.4
Low/high temp. studies	306809	44.7
Absolute configuration determined	14 752	2.1
Disorder present in structure	158 127	23.0
Polymorphic structures	20 753	3.0
R-factor < 0.100	645 809	94.0
R-factor < 0.075	585 333	85.2
R-factor < 0.050	378 391	55.1
R-factor < 0.030	78 594	11.4
No. of atoms with 3D coordinates	53 563 990	-

CSD – Interface WEB

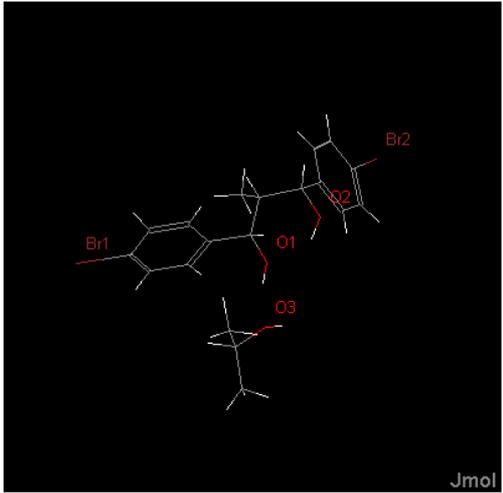
Interactive WebCSD x

webcsd.ccdc.cam.ac.uk/teaching_database_demo.php

This interactive demo allows you to browse through all 733 entries in the [CSD Teaching Database](#) using the standard [WebCSD](#) interface. A number of example [teaching exercises](#) (including VSEPR and stereochemistry) can be completed using this demo version. A free [hyperlink generator tool](#) is now available.

ABALEV : (1S,3S)-1,3-bis(4-Bromophenyl)-2-methylpropane-1,3-diol isopropanol solvate
V.Gnanadesikan, Y.Horiuchi, T.Ohshima, M.Shibasaki; *J.Am.Chem.Soc.* (2004), **126**, 7782, doi:10.1021/ja047906f

Diagram Details Viewer Export Options Help



Hide Viewer

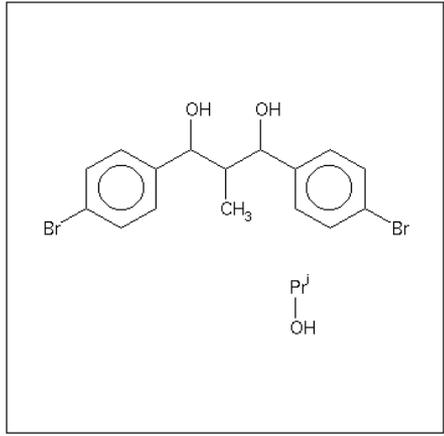
Wireframe All but C/H

Hydrogens Bond types Disorder

Packing Options

None Unit Cell 3x3x3

Launch External Viewer



View Group Symbols Key

C₁₆ H₁₆ Br₂ O₂.C₃ H₈ O

Space Group: P 2₁

a 10.692(2) b 8.858(2) c 11.968(2)

α 90 β 114.40(1) γ 90

R-Factor: 3.8%

Temperature (K): 200

733 Hits

100%

Stop Search

Entry loaded

Added to CSD: 10th November 2004; Last modified: 10th November 2004; Published in WebCSD: 17th September 2009 14:43:54

CSD – Pedido de estrutura

The screenshot shows a web browser window with the URL www.ccdc.cam.ac.uk/Community/Requestastructure/Pages/DataRequestResponse.aspx. The page features a navigation menu with links for COMMUNITY, RESEARCH & CONSULTANCY, SOLUTIONS, NEWS & EVENTS, SUPPORT & RESOURCES, and THE CCDC. A blue banner is present below the navigation. The main content area displays the following information:

Home / Community / Request a Structure / Data Request Results Summary

Your query was: 243822 and returned 1 successful record(s)

Publications
Journal of Organic Chemistry (2004), 69, 4500, doi:10.1021/jo049716t Hongbin Li, Hua Yang, J.L.Petersen, Kung K.Wang

CCDC Structure Summary for All Successful Requests:

Selected	CCDC No	a	b	c	Space Group	Download CIF	View in WebCSD	Created On
<input checked="" type="checkbox"/>	243822	11.1366(12)	6.9872(7)	15.3869(16)	P21/c	Download	ABABEL	06/07/2004

Download Mode

- Deposited CIF
- Deposited CIF without Structure Factor data
- Deposited CIF and Structure Factor Files if available

You can also download all the selected files at once [Download Selected](#)

View Selected in WebCSD [View in WebCSD](#)

[New Request](#)

[Email Failed Requests](#)

Conditions of Use of CIFs provided from the CCDC CIF archive

Individual CIF data sets are provided freely by the CCDC on the understanding that they are used for bona fide research purposes only. They may contain copyright material of the CCDC or of third parties, and may not be copied or further disseminated in any form, whether machine-readable or not, except for the purpose of generating routine backup copies on your local computer system.

I agree to the conditions of use.*

Your Name *

Your Email *

Your Affiliation *

Drug Bank

- Base de dados bioinformática e cheminformática
- Version 5.1.4 (2019-07-02):
 - 13441 compostos
 - 2618 fármacos aprovados pela FDA
 - 1340 fármacos biológicas (proteínas/péptidos)
 - 130 nutraceutuais
 - 6335 fármacos em fase experimental
 - 5157 proteínas (alvos/enzimas/transporters/carriers)
- Cada entrada (DrugCard) contém mais de 200 campos
- As entradas combinam informação sobre o fármaco (química, farmacológica e farmacêutica) com informação sobre o alvo (sequência, estrutura e via metabólica)

DrugBank

https://www.drugbank.ca

DRUGBANK

WHAT ARE YOU LOOKING FOR?

Aspirin

Drugs Targets Pathways Indications

DRUGBANK

The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug data with comprehensive drug target information.

The latest release of DrugBank (version 5.1.1, released 2018-07-03) contains 11,885 drug entries including 2,528 approved small molecule drugs, 1,184 approved biotech (protein/peptide) drugs, 129 nutraceuticals and over 5,755 experimental drugs. Additionally, 5,132 non-redundant protein (i.e. drug

Acetylsalicylic acid - DrugBank x +

← → ↻ 🏠 <https://www.drugbank.ca/drugs/DB00945> 🔍 ☆

📱 Apps 📦 AWS 📖 LibGen 🌐 Tut2017 📄 Acad 🌐 D pmartel 📄 Notepad 📄 Pasteboard 📄 dpaste 🎮 Trinket » 📁 Other bookmarks

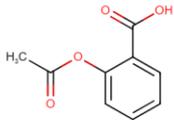
DRUGBANK Browse ▾ Search ▾ Downloads About ▾ Help ▾ Blog Contact Us

Drugs ▾ 🔍

Acetylsalicylic acid

Targets (11) Enzymes (3) Carriers (1) Transporters (3) Biointeractions (16) <<

IDENTIFICATION

Name	Acetylsalicylic acid
Accession Number	DB00945 (APRD00264, EXPT00475)
Type	Small Molecule
Groups	Approved, Vet approved
Description	The prototypical analgesic used in the treatment of mild to moderate pain. It has anti-inflammatory and antipyretic properties and acts as an inhibitor of cyclooxygenase which results in the inhibition of the biosynthesis of prostaglandins. Acetylsalicylic acid also inhibits platelet aggregation and is used in the prevention of arterial and venous thrombosis. (From Martindale, The Extra Pharmacopoeia, 30th ed, p5)
Structure	 <chem>CC(=O)Oc1ccccc1C(=O)O</chem>
	🔍 3D Download ▾ 🔗 Similar Structures
Synonyms	2-Acetoxybenzenecarboxylic acid

ChEMBL

- Base de dados mantida e curada manualmente pelo European Bioinformatics Institute (EBI), parte do European Molecular Biology Laboratory (EMBL).
- Contem informação sobre a acção de compostos bioactivos em alvos farmacológicos (drug targets). A informação inclui K_i , K_d , IC_{50} e EC_{50} .
- Entradas separadas para compostos e alvos.
- A versão mais recente (v. 29, 1/07/2021) contém 2,105,464 compostos, 14,454 alvos e 18,635,916 ensaios de actividade derivados de 81,544 publicações.
- Contém uma série de ferramentas para análise e filtragem da informação contida na base de dados

Exemplo de pesquisa estrutural em ChEMBL

ChEMBL Database

European Bioinformatics Institute [GB] | <https://www.ebi.ac.uk/chembl/b...>

Apps AWS LibGen Tut2017 Acad D pmartel Notepad Pasteboard dpaste Trinket Desmos Other bookmarks

EMBL-EBI Services Research Training About us

EMBL-EBI

ChEMBL

Search in ChEMBL

Example: Dopamine Aspirin NCCc1ccc(O)c(O)c1 Liver

UniChem | ChEMBL-NTD | SureChEMBL | Downloads | Web Services | More

Drugs by Usan Year (4015)

Legend:

- 0
- 1
- 2
- 3

ChEMBL

A manually curated database of bioactive molecules with drug-like properties. It brings together chemical, bioactivity and genomic data to aid the translation of genomic information into effective new drugs. See the [interface documentation](#).

[B](#) [T](#)

This website requires cookies, and the limited processing of your personal data in order to function. By using the site you are agreeing to this as outlined in our [Privacy Notice](#) and [Terms of Use](#). [I agree, dismiss this banner](#)

Exemplo de pesquisa estrutural em ChEMBL

The image shows a screenshot of a web browser displaying the ChEMBL Database interface. The browser's address bar shows the URL [https://www.ebi.ac.uk/...](https://www.ebi.ac.uk/). The ChEMBL logo is visible in the top left corner. The main area of the interface features a large white canvas where a chemical structure of biphenyl (two benzene rings connected by a single bond) is displayed. To the left of the canvas is a vertical toolbar with various drawing tools. To the right is a vertical menu with a list of chemical elements: H, C, N, O, S, F, P, Cl, Br, I, *, and A. At the bottom of the interface, there are three teal buttons: "Connectivity", "Similarity", and "Substructure". The "Similarity" button is highlighted with a red box. Below the "Similarity" button is a slider control, also highlighted with a red box, showing a value of $\geq 100\%$ and a green indicator at the right end of the slider.

Exemplo de pesquisa estrutural em ChEMBL

The screenshot displays the ChEMBL search interface. At the top, a search bar contains the query C1=CC=C(C=C1)C1C=CC=CC=1 with a threshold of 100%. Below the search bar, a chemical structure of biphenyl is shown. The search results section indicates 1 compound found, with options to download in CSV, TSV, or SDF format. The filters panel on the left shows various criteria such as Type (Small molecule), Max Phase, #RO5 Violations, Molecular Weight, and AlogP. The main results area shows a single record for CHEMBL14092, identified as BIPHENYL with a similarity of 100.

ChEMBL

Search in ChEMBL

Query: C1=CC=C(C=C1)C1C=CC=CC=1 Threshold: 100% [Edit Search](#)

1 Compounds
0 Selected - [Select All](#)
[Browse Activities](#) [CSV](#) [TSV](#) [SDF](#)

Table **Cards** Infinite Graph Heatmap

Showing 1-1 out of 1 records

Records per page: 6 Select All

Filters

- Type
 - Small molecule 1
- Max Phase
 - 0 1
- #RO5 Violations
 - 0 1
- Molecular Weight
 - [154.21 to 154.31] 1
- AlogP

Similarity Maps
On
Off

CHEMBL14092
Name: BIPHENYL
Similarity: 100

Exemplo de pesquisa estrutural em ChEMBL

ChEMBL

Compound Report Card

Europea...te [GB] | [https://www.ebi.ac.uk/...](https://www.ebi.ac.uk/)

EMBL-EBI | Services | Research | Training | About us

Search in ChEMBL

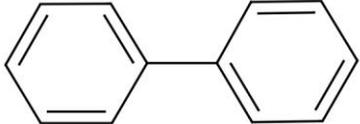
Example: Dopamine Aspirin NCCc1ccc(O)c(O)c1 Liver

UniChem | ChEMBL-NTD | SureChEMBL | Downloads | Web Services | More

EBI > Databases > Chemical Biology > ChEMBL Database > CHEMBL14092

Compound Report Card

Name And Classification



ID: CHEMBL14092

Name: BIPHENYL

Max Phase: 0 Research 

Molecular Formula: C₁₂H₁₀

Molecular Weight: 154.21

ChEMBL Synonyms: E230

Molecule Type: Small molecule

- Name And Classification
- Representations
- Sources
- Clinical Data
- Activity Charts
- Literature
- Calculated Properties
- Cross References
- UniChem Cross References
- UniChem Connectivity
- Layer Cross References
- Alternative Forms



Exemplo de pesquisa estrutural em ChEMBL

The image shows a web browser window displaying the ChEMBL Compound Report Card for biphenyl. The browser's address bar shows the URL <https://www.ebi.ac.uk/...>. The ChEMBL logo is visible in the top left, and a search bar is in the top right. The page title is "Compound Report Card".

A 3D ball-and-stick model of biphenyl is shown in a black viewer window in the center. The model consists of two benzene rings connected by a single carbon-carbon bond. Atoms are represented by spheres: carbon (grey), hydrogen (white), and oxygen (red).

Below the viewer, the 2D skeletal structure of biphenyl is shown. To the right of the structure, the following properties are listed:

- Max Phase:** 0 Research [i](#)
- Molecular Formula:** C₁₂H₁₀
- Molecular Weight:** 154.21
- ChEMBL Synonyms:** [E230](#)
- Molecule Type:** Small molecule

On the right side of the page, a vertical menu lists various sections: Name And Classification, Representations, Sources, Clinical Data, Activity Charts, Literature, Calculated Properties, Cross References, UniChem Cross References, UniChem Connectivity, Layer Cross References, and Alternative Forms.

Exemplo de pesquisa estrutural em ChEMBL

ChEMBL Compound Report Card

Europe... [GB] | https://www.ebi.ac...

Activity Charts

Bioactivity Summary

Exclude Alternate Forms Data

Activity Types for Compound CHEMBL14092 (including alternate forms)



Activity Type	Count
Potency	6
LogP	4
LogP app	2
Papp	2
Ratio	2
Binding aff...	1
Concentration	1
IC50	1
Other	5
(Unlabeled)	1

- Potency
- LogP
- LogP app
- Papp
- Ratio
- logD
- Binding aff...
- Concentration
- IC50
- Other

Assay Summary

Exclude Alternate Forms Data

Assays for Compound CHEMBL14092 (including alternate forms)



Assay Type	Count
P - Physic...	7
B - Binding	6
F - Functional	6
A - ADME	5

- P - Physic...
- B - Binding
- F - Functional
- A - ADME

Target Summary

Exclude Alternate Forms Data

Target Classes for Compound CHEMBL14092 (including alternate forms)



Target Class	Count
Enzyme	3
Secreted p...	2
Ion channel	1
Transcripti...	1
Unclassifie...	1

- Enzyme
- Secreted p...
- Ion channel
- Transcripti...
- Unclassifie...

Literature

Exemplo de pesquisa estrutural em ChEMBL

UniChem Cross References

ACToR	56481-93-7, 92-52-4, 68409-73-4
BindingDB	50168002
Brenda	1885, 107817
ChEBI	17097
ChemicalBook	CB2491271
eMolecules	481835
EPA CompTox Dashboard	DTXSID4020161
FDA SRS	2L9GJK6MGN
Human Metabolome Database	HMDB0034437
IBM Patent System	390A3BB9FB86D9D98D36A1679728E770
KEGG Ligand	C06588
Mcule	MCULE-2274387658
MolPort	MolPort-001-738-537
Nikkaji	J3.929B
NMRShiftDB	10006018
PDBe	BNL

Exemplo de pesquisa estrutural em ChEMBL

ChEMBL | Compound Report Card | PDBeChem: Ligand Dictionary (PI X)

Not secure | www.ebi.ac.uk/pdbe-srv/pdb...

EMBL-EBI | Protein Data Bank in Europe | Chemical Components in the PDB

Bringing Structure to Biology

Summary

BNL : Summary

Code BNL
One-letter code X
Molecule name BIPHENYL

Systematic names

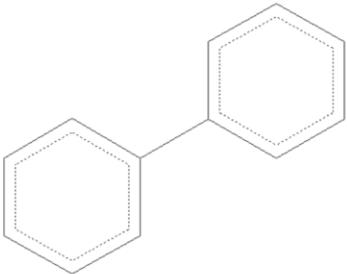
Program	Version	Name
ACDLabs	11.02	biphenyl
OpenEye OEToolkits	1.6.1	1,1'-biphenyl

Formula C12 H10
Formal charge 0
Molecular weight 154.208 Da

SMILES

Type	Program	Version	Descriptor
SMILES	ACDLabs	11.02	c1cc(ccc1)c2ccccc2
SMILES	CACTVS	3.352	c1ccc(cc1)c2ccccc2
SMILES	OpenEye OEToolkits	1.7.0	c1ccc(cc1)c2ccccc2
Canonical SMILES	CACTVS	3.352	c1ccc(cc1)c2ccccc2
Canonical SMILES	OpenEye OEToolkits	1.7.0	c1ccc(cc1)c2ccccc2

IUPAC InChI InChI=1S/C12H10/c1-3-7-11(8-4-1)12-9-5-2-6-10-12/h1-10H
IUPAC InChI key ZUOUZKKEUPVFJK-UHFFFAOYSA-N



wwPDB Information

Atom count 22 (12 without Hydrogen)
Polymer type Bound ligand
Type description NON-POLYMER
Type code HETAIN
Is modified No
Standard parent Not Assigned
Defined at 2003-09-15
Last modified at 2011-06-04
Status Released
Obsoleted Not Assigned

Download Links | Related compounds | 3D-Views | PDB Links

PDBe is a member of PDB | EMDatabank

Exemplo de pesquisa estrutural em ChEMBL

ChEMBL Compound Report Card PDBeChem: Ligand Dicti PDB 1ulj structure summ

Not secure | www.ebi.ac.uk/pdbe-srv/pdb...

This website uses cookies. By continuing to browse this site, you are agreeing to the use of our site cookies. To find out more, see our [Terms of Use](#).

EMBL-EBI Protein Data Bank in Europe Bringing Structure to Biology

Services Research Training About us

Chemical Components in the PDB

Share Feedback

PDBeChem : Used in PDB Entries

Molecule : BNL

The PDB entries where the chemical component is used

Total Number of PDB Entries: 5
(Download list of entries for this compound)

1/1 50 per page

Ligand Code	PDB Entry ID	Type	Total	Distinct
BNL	1ulj	Bound ligand	3	1
BNL	2gbx	Bound ligand	3	1
BNL	2xrx	Bound ligand	12	1
BNL	3gzz	Bound ligand	1	1
BNL	5aew	Bound ligand	9	1

Download Links

Related compounds

3D-Views

PDB Links

PDBe is a member of PDB EMDatabank

Exemplo de pesquisa estrutural em ChEMBL

ChEMBL | Compound Repo | PDBChem: Ligand | PDB 1ulj structure | PDB 1ulj structure

Not secure | www.ebi.ac.uk/pdbe/entry/p... | Search | Examples: hemoglobin, BRCA1_HUMAN | Feedback

EMBL-EBI Protein Data Bank in Europe

Bringing Structure to Biology

PDBe > 1ulj

Biphenyl dioxygenase (BphA1A2) in complex with the substrate
Source organism: *Rhodococcus jostii* RHA1

Primary publication:
Crystal structure of the terminal oxygenase component of biphenyl dioxygenase derived from *Rhodococcus* sp. strain RHA1.
Furusawa Y, Nagarajan V, Tanokura M, Masai E, Fukuda M, Senda T
J. Mol. Biol. **342** 1041-52 (2004)
PMID: 15342255

X-ray diffraction
2.6Å resolution
Released: 28 Sep 2004
Model geometry: [Slider] Fit model/data: [Slider]

Quick links

- 1ulj overview
 - Citations
 - Structure analysis
 - Function and Biology
 - Ligands and Environments
 - Experiments and Validation
- View
- Downloads
- 3D Visualisation

Function and Biology [Details]

Reaction catalysed:
 $\text{Biphenyl} + \text{NADH} + \text{O}_2 = (1\text{S},2\text{R})\text{-3-phenylcyclohexa-3,5-diene-1,2-diol} + \text{NAD}^+$

Biochemical function: biphenyl 2,3-dioxygenase activity

Biological process: oxidation-reduction process

Cellular component: not assigned

Sequence domains:

- Aromatic-ring-hydroxylating dioxygenase, alpha subunit
- Ring-hydroxylating dioxygenase beta subunit
- Rieske [2Fe-2S] iron-sulphur domain
- Rieske [2Fe-2S] iron-sulphur domain superfamily
- NTF2-like domain superfamily
- Aromatic-ring-hydroxylating dioxygenase, alpha subunit, C-terminal domain
- Aromatic-ring-hydroxylating dioxygenase, 2Fe-2S-binding site

Structure domains:

- Ring hydroxylating alpha subunit ISP domain
- Ring hydroxylating alpha subunit catalytic domain

Ligands and Environments

3 bound ligands:

- 3 x FE2
- 3 x FES
- 3 x BNL

No modified residues

Experiments and Validation [Details]

Metric	Percentile Ranks	Value
Rfsec	[Slider]	0.237
Clashscore	[Slider]	3
Ramachandran outliers	[Slider]	0.2%
Sidechain outliers	[Slider]	5.8%
RSRZ outliers	[Slider]	0.9%

Citations

8 review citations

Prospects for using combined engineered bacterial enzymes and plant systems to rhizoremediate polychlorinated biphenyls.
Sylvestre M. (2013) [7 more]

PDB_REDO

The sliders below show the change in model quality between original PDB entry and the PDB_REDO entry

Model Geometry: [Slider] Fit model/data: [Slider]

PDB-REDO

Pesquisa de targets em ChEMBL



[Browse all ChEMBL](#)

[See all visualisations](#)

Current Release: ChEMBL 25

Provided under a [Creative Commons Attribution-ShareAlike 3.0 Unported license](#)

Last Update on 2018-12-10 | [Release notes](#)



12,482
Targets



1,879,206
Distinct compounds



15,504,603
Activities



72,271
Publications



54
Deposited Datasets



Citing ChEMBL

Pesquisa de targets em ChEMBL

EMBL-EBI Services Research Training About us

EMBL-EBI

ChEMBL

Search in ChEMBL

Examples: Imatinib erbB2 brain MDCK c1ccccc1N Draw a Structure | Enter a Sequence

UniChem ChEMBL-NTD SureChEMBL Downloads Web Services Old Interface More Share

EBI > Databases > Chemical Biology > ChEMBL Database > Targets > Query

Browse Targets

[Edit Querystring](#) [Show Full Query](#)

12,482 Targets
0 Selected - Select All
Browse Activities

Table Heatmap CSV TSV

Records per page: 20 Show/Hide Columns

Showing 1-20 out of 12,482 records

* [Search]

1 2 3 4 5 ... >

ChEMBL ID	Name	UniProt Accessions	Type	Organism	Compounds	Activities
<input type="checkbox"/>	CHEMBL3390823	<i>Disialoganglioside GD2</i>	SMALL MOLECULE	Homo sapiens	0	No Data
<input type="checkbox"/>	CHEMBL3833503	<i>tRNA</i>	NUCLEIC-ACID	No Data	0	No Data
<input type="checkbox"/>	CHEMBL3559389	<i>Triglyceride</i>	LIPID	No Data	0	No Data
<input type="checkbox"/>	CHEMBL2366037	<i>Radioactive metals</i>	METAL	No Data	0	No Data
<input type="checkbox"/>	CHEMBL2363056	<i>Zinc</i>	METAL	No Data	0	No Data
<input type="checkbox"/>	CHEMBL2363058	<i>Iron</i>	METAL	No Data	0	No Data

Filters

Organism Taxonomy L1

- Eukaryotes 9592
- Bacteria 1417
- Fungi 686
- Viruses 414
- N/A - 364
- Archaea 7
- Unclassified 2

Organism Taxonomy L2

- Mammalia 8093
- Gram-Negative 740
- Gram-Positive 652

Pesquisa de targets em ChemBL

EMBL-EBI Services Research Training About us

ChEMBL

Search in ChEMBL

Examples: Imatinib erbB2 brain MDCK c1ccccc1N Draw a Structure | Enter a Sequence

UniChem ChEMBL-NTD SureChEMBL Downloads Web Services Old Interface More Share

EBI > Databases > Chemical Biology > ChEMBL Database > Targets > Query

Browse Targets

[Edit Querystring](#) [Show Full Query](#)

12,482 Targets
0 Selected - Select All
Browse Activities

Table Heatmap CSV TSV

Filters

Records per page: 20 Show/Hide Columns

Showing 1-20 out of 12,482 records

trypsin

ChEMBL ID	Name	UniProt Accessions	Type	Organism	Compounds	Activities
<input type="checkbox"/>	CHEMBL3390823	<i>Disialoganglioside GD2</i>	SMALL MOLECULE	Homo sapiens	0	No Data
<input type="checkbox"/>	CHEMBL3833503	<i>tRNA</i>	NUCLEIC-ACID	No Data	0	No Data
<input type="checkbox"/>	CHEMBL3559389	<i>Triglyceride</i>	LIPID	No Data	0	No Data
<input type="checkbox"/>	CHEMBL2366037	<i>Radioactive metals</i>	METAL	No Data	0	No Data

Organism Taxonomy L1

- Eukaryotes 9592
- Bacteria 1417
- Fungi 686
- Viruses 414
- N/A - 364
- Archaea 7
- Unclassified 2

Organism Taxonomy L2

- Mammalia 8093

Pesquisa de targets em ChEMBL

ChEMBL Search in ChEMBL

Records per page: 20 Show/Hide Columns trypsin

Showing 1-20 out of 116 records

<input type="checkbox"/>	ChEMBL ID	Name	UniProt Accessions	Type	Organism	Compounds	Activities
<input type="checkbox"/>	CHEMBL4472	<i>Trypsin II</i>	Q29463	SINGLE PROTEIN	Bos taurus	213 By Mol. Wt.:	226 By Std. Type:
<input type="checkbox"/>	CHEMBL4611	<i>Complement C1r</i>	P00736	SINGLE PROTEIN	Homo sapiens	117 By Mol. Wt.:	224 By Std. Type:
<input type="checkbox"/>	CHEMBL3063	<i>Beta-chymotrypsin</i>	P00767	SINGLE PROTEIN	Bos taurus	23 By Mol. Wt.:	25 By Std. Type:
<input type="checkbox"/>	CHEMBL2111424	<i>Coagulation factor IX and X</i>	P00740, P00742	SELECTIVITY GROUP	Homo sapiens	90 By Mol. Wt.:	91 By Std. Type:
<input type="checkbox"/>	CHEMBL5610	<i>Prostasin</i>	Q16651	SINGLE PROTEIN	Homo sapiens	25 By Mol. Wt.:	25 By Std. Type:
<input type="checkbox"/>	CHEMBL3243910	<i>Acrosin</i>	P08001	SINGLE PROTEIN	Sus scrofa	53 By Mol. Wt.:	53 By Std. Type:

Filters

Organism Taxonomy L1

- Eukaryotes 114
- Bacteria 2

Organism Taxonomy L2

- Mammalia 114
- Gram-Negative 1
- Gram-Positive 1

Organism Taxonomy L3

- Primates 67
- Rodentia 25
- Cetartiodactyla 17
- Lagomorpha 3
- Carnivora 2
- Lysobacter 1
- Staphylococcus 1

Organism

- Bos taurus 9
- Canis lupus familiaris 2
- Homo sapiens 67
- Lysobacter enzymogenes 1
- Mus musculus 12
- Oryctolagus cuniculus 3
- Rattus norvegicus 13
- Staphylococcus aureus 1

Pesquisa de targets em ChEMBL

Target Report Card

Name And Classification



ID:	CHEMBL4472
Type:	SINGLE PROTEIN
Preferred Name:	Trypsin II
Synonyms:	Anionic trypsin
Organism:	Bos taurus
Species Group:	No
Protein Target Classification:	- Enzyme > Protease > Serine protease > Serine protease PA clan > Serine protease S1A subfamily



Name And Classification

Components

Activity Charts

Ligand Efficiencies

Associated Compounds

Gene Cross References

Protein Cross References

Domain Cross References

Structure Cross References

Components

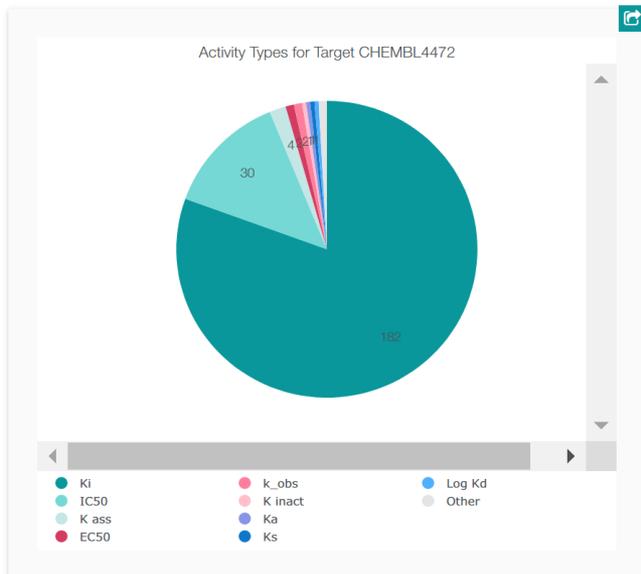


Pesquisa de targets em ChEMBL

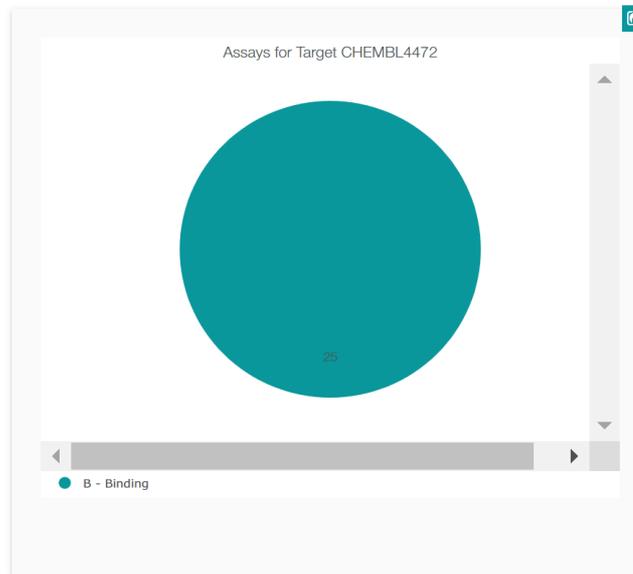
Activity Charts

- Name And Classification
- Components
- Activity Charts
- Ligand Efficiencies
- Associated Compounds
- Gene Cross References
- Protein Cross References
- Domain Cross References
- Structure Cross References

Associated Bioactivities



Associated Assays



Pesquisa de targets em ChEMBL



Pesquisa de targets em ChEMBL

Search in ChEMBL



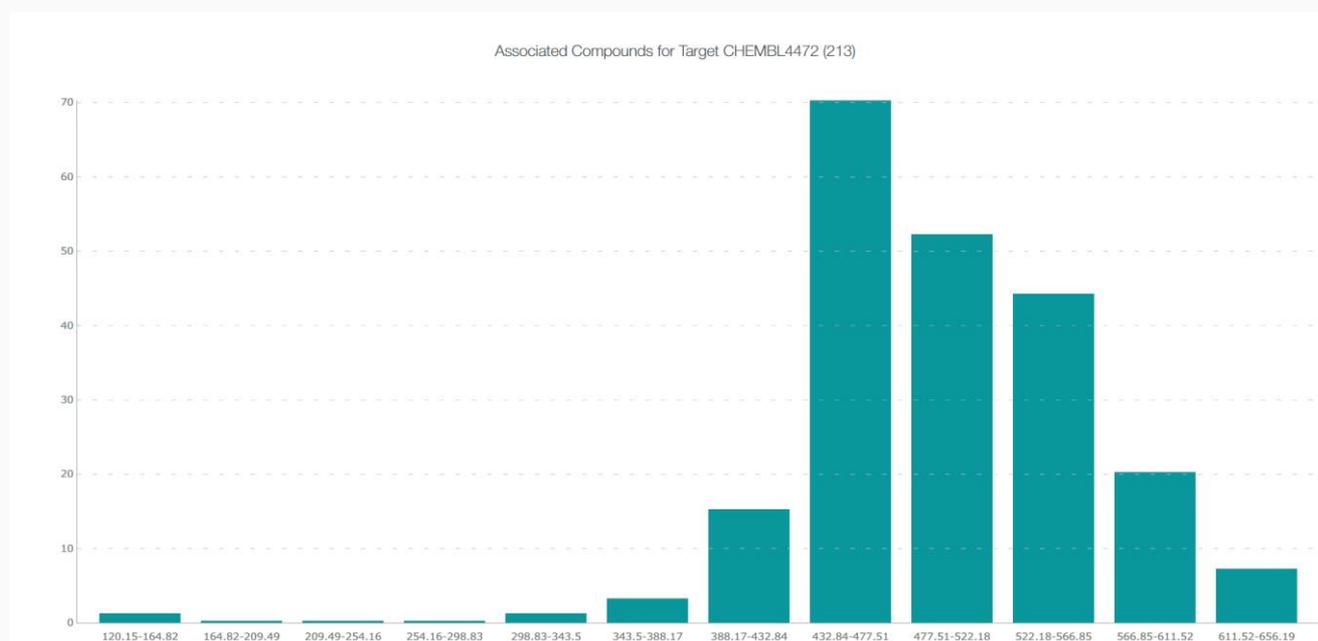
Associated Compounds

X axis:

Parent Molecular Weight

Number of bins: 12

Bin size: 44.67

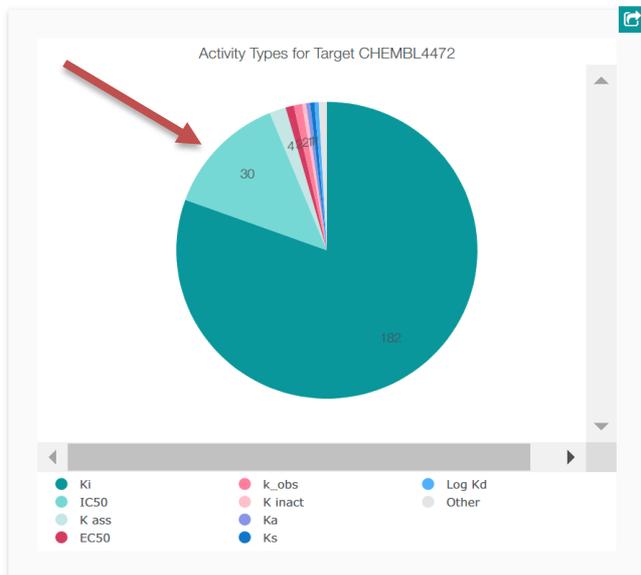


Nan
Con
Acti
Lige
Assi
Gen
Prol
Don
Stru

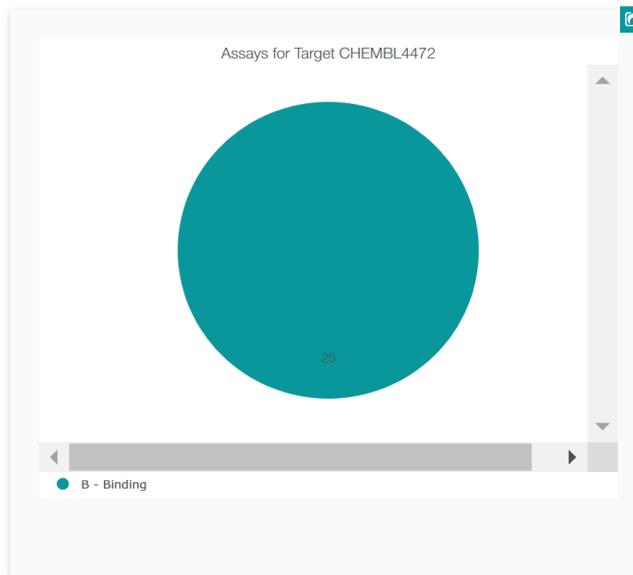
Activity Charts

- Name And Classification
- Components
- Activity Charts
- Ligand Efficiencies
- Associated Compounds
- Gene Cross References
- Protein Cross References
- Domain Cross References
- Structure Cross References

Associated Bioactivities



Associated Assays



Browse Activities

[Edit Querystring](#)
[Show Full Query](#)
 Table

30 Activities

 0 Selected - [Select All](#)
[Browse Compounds](#)
 CSV

 TSV

Filters

Standard Type

IC50 30

Target Type

SINGLE PROTEIN 30

Organism Taxonomy L1

Eukaryotes 30

Organism Taxonomy L2

Mammalia 30

Organism Taxonomy L3

Cetartiodactyla 30

Target Organism

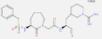
Records per page:

20

[Show/Hide Columns](#)

Showing 1-20 out of 30 records

< 1 2 >

<input type="checkbox"/>	Molecule ChEMBL ID	Compound Key	Standard Type	Standard Relation	Standard Value	Standard Units	pChEMBL Value	Comment	Assay ChEMBL ID	Assay Descri
<input type="checkbox"/>	 CHEMBL342914	6a	IC50	=	15100	nM	4.82	No Data	CHEMBL815149	Compu evaluate inhibito amidoly activity chromo substrat
<input type="checkbox"/>	 CHEMBL141424	5b	IC50	=	64800	nM	4.19	No Data	CHEMBL815149	Compu evaluate inhibito amidoly activity chromo substrat

Filters

- Standard Type
 - IC50 30
 - Target Type
 - SINGLE PROTEIN 30
 - Organism Taxonomy L1
 - Eukaryotes 30
 - Organism Taxonomy L2
 - Mammalia 30
 - Organism Taxonomy L3
 - Cetartiodactyla 30
 - Target Organism
 - Bos taurus 30
 - BAO Format
 - single protein format 30
 - pCHEMBL Value
 - [4.19 to 4.50] 3
 - [4.50 to 5] 1
 - [5 to 5.50] 0
 - [5.50 to 6] 3
 - [6 to 6.50] 2
 - [6.50 to 7] 2
 - [7 to 7.50] 1
 - [7.50 to 8] 0
 - [8 to 8.50] 0
 - [8.50 to 9] 1
 - [9 to 9.02] 0
 - Max Phase
 - 0 30
 - #RoS Violations
 - 0 13
 - 1 9
 - 2 8
 - ALogP
 - [-0.96 to 0] 8
 - [0 to 1] 6
 - [1 to 2] 0
 - [2 to 3] 0
 - [3 to 4] 2
 - [4 to 5] 6
 - [5 to 6] 5
 - [6 to 6.60] 3
 - Molecular Weight
 - [429.52 to 440] 2

Records per page:
20

Show/Hide Columns

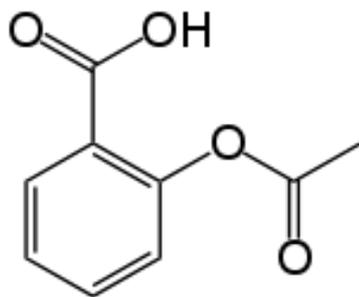
Showing 1-20 out of 30 records

Molecule CHEMBL ID	Compound Key	Standard Type	Standard Relation	Standard Value	Standard Units	pCHEMBL Value	Comment	Assay CHEMBL ID	Assay Description	BAO Label	Assay Organism	Target CHEMBL ID	Target Name	Target Organism	Target Type	Document CHEMBL ID	Source Description	Cell CHEMBL ID
 6a CHEMBL342914		IC50	=	15100	nM	4.82	No Data	CHEMBL815149	Compound was evaluated for inhibition of amidolytic activity for chromogenic substrate trypsin	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1129391	Scientific Literature	No Data
 5b CHEMBL141424		IC50	=	64800	nM	4.19	No Data	CHEMBL815149	Compound was evaluated for inhibition of amidolytic activity for chromogenic substrate trypsin	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1129391	Scientific Literature	No Data
 3i CHEMBL421760		IC50	=	538	nM	6.27	No Data	CHEMBL815150	compound was tested in vitro for inhibition of serine protease Trypsin.	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1129392	Scientific Literature	No Data
 1; CVS 1123 CHEMBL141676		IC50	=	1.2	nM	8.92	No Data	CHEMBL815149	Compound was evaluated for inhibition of amidolytic activity for chromogenic substrate trypsin	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1129391	Scientific Literature	No Data
 14 CHEMBL113418		IC50	No Data	No Data	No Data	No Data	Not Determined	CHEMBL817750	In vitro inhibition of bovine trypsin; Not determined.	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1132421	Scientific Literature	No Data
 3, CVS 1778 CHEMBL344204		IC50	=	329	nM	6.48	No Data	CHEMBL815149	Compound was evaluated for inhibition of amidolytic activity for chromogenic substrate trypsin	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1129391	Scientific Literature	No Data
 4b CHEMBL342838		IC50	=	69.7	nM	7.16	No Data	CHEMBL815149	Compound was evaluated for inhibition of amidolytic activity for chromogenic substrate trypsin	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1129391	Scientific Literature	No Data
 13 CHEMBL111745		IC50	No Data	No Data	No Data	No Data	Not Determined	CHEMBL817750	In vitro inhibition of bovine trypsin; Not determined.	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1132421	Scientific Literature	No Data
 10 CHEMBL110986		IC50	No Data	No Data	No Data	No Data	Not Determined	CHEMBL817750	In vitro inhibition of bovine trypsin; Not determined.	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1132421	Scientific Literature	No Data
 3e CHEMBL263924		IC50	=	1020	nM	5.99	No Data	CHEMBL815150	compound was tested in vitro for inhibition of serine protease Trypsin.	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1129392	Scientific Literature	No Data
 3d CHEMBL140545		IC50	=	2500	nM	5.6	No Data	CHEMBL815150	compound was tested in vitro for inhibition of serine protease Trypsin.	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1129392	Scientific Literature	No Data
 9 CHEMBL109601		IC50	No Data	No Data	No Data	No Data	Not Determined	CHEMBL817750	In vitro inhibition of bovine trypsin; Not determined.	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1132421	Scientific Literature	No Data
 7 CHEMBL110746		IC50	No Data	No Data	No Data	No Data	Not Determined	CHEMBL817750	In vitro inhibition of bovine trypsin; Not determined.	single protein format	No Data	CHEMBL4472	Trypsin II	Bos taurus	SINGLE PROTEIN	CHEMBL1132421	Scientific Literature	No Data

SMILES

SMILES - Simplified Inter Molecular Enter Specification

Linguagem que permite a representação de estruturas moleculares 2D na forma de uma sequência (“string”) de caracteres.



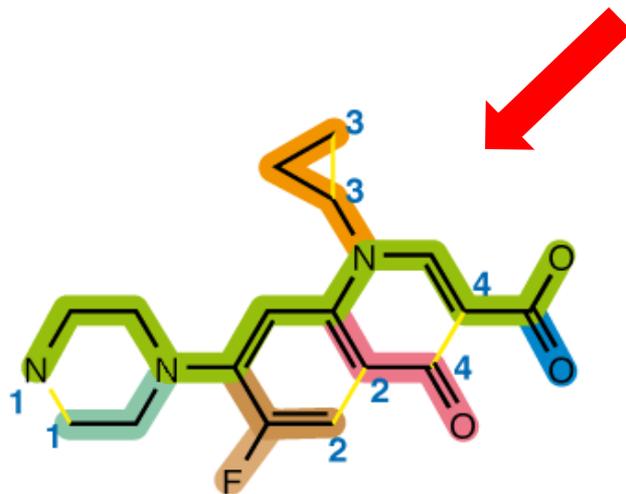
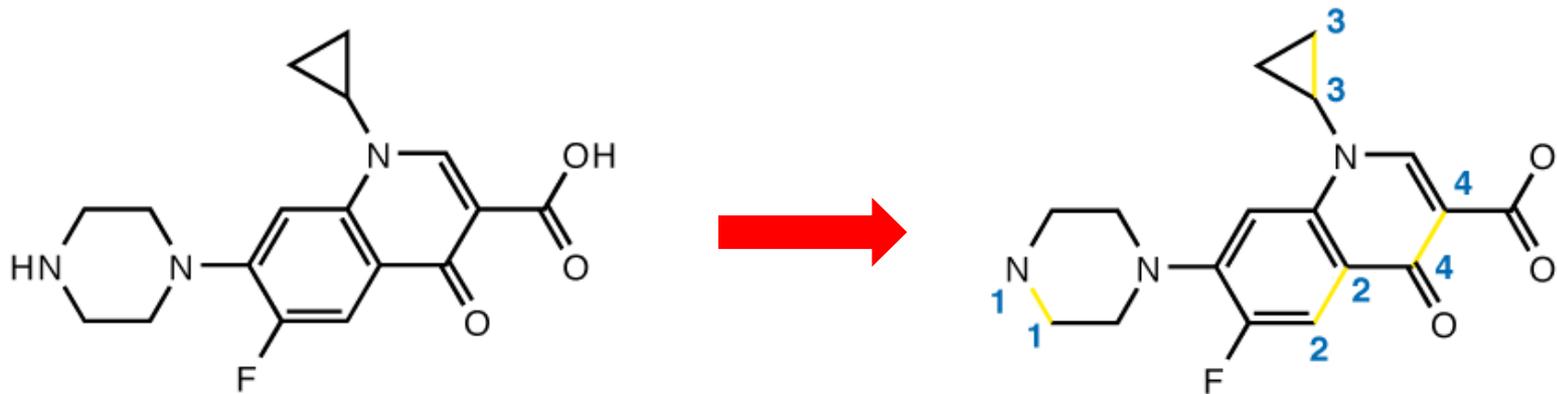
Estrutura 2D



O=C(Oc1ccccc1C(=O)O)C

SMILES

Tutorial SMILES: <http://www.daylight.com/>



N1CCN(CC1)C(C(F)=C2)=CC(=C2C4=O)N(C3CC3)C=C4C(=O)O

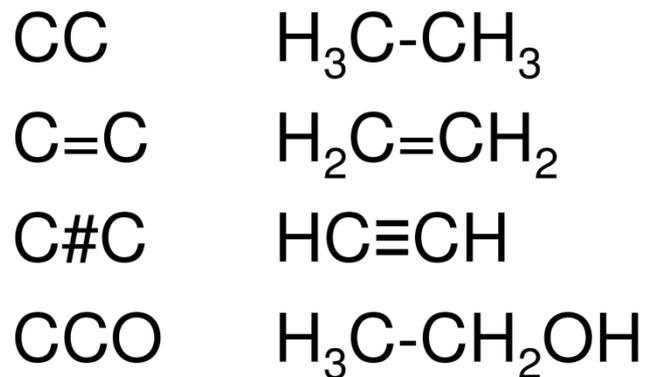


SMILES – Regras(1)

Os átomos são representados pelos seus nomes elementais:

C B N O P S Cl Br I H (compostos orgânicos)

- Outros elementos – **[Si] [Fe] [Co]**
- O hidrogénio é geralmente ignorado: CH₄ → **C**



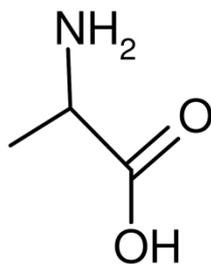
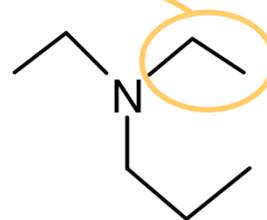
SMILES – Regras(2)

Átomos e ligações:

- CC as ligações simples não são representadas
- C=C ligações duplas
- C#C ligações triplas
- c:c ligações entre carbonos aromáticos
(geralmente não se representam)
- C@C qualquer tipo de ligação num anel
- C~C qualquer tipo de ligação

SMILES – Regras(3)

As ramificações denotam-se com parêntesis:

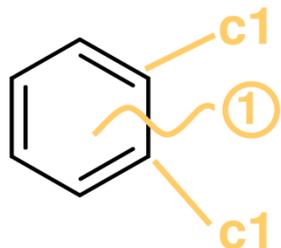


(determinar primeiro a sequência mais longa de ligações)

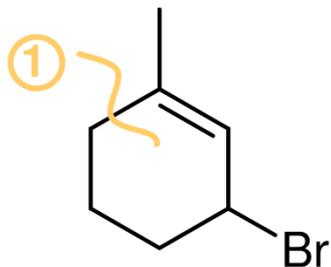
SMILES – Regras(4)

Compostos cíclicos:

- Encontrar cadeia mais longa
- “abrir” o anel para obter uma cadeia
- numerar carbonos no pontos de abertura



c1ccccc1



CC1=CC(Br)CCC1

SMILES – Regras(5)

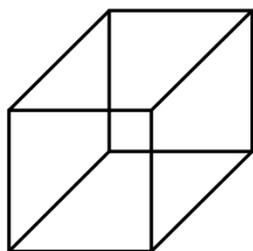
Compostos policíclicos:

- Múltiplos pontos de quebra

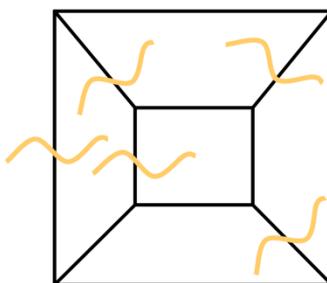


c1cc2ccccc2cc1

Pode ocorrer fecho de mais do que um anel no mesmo átomo:



cubano

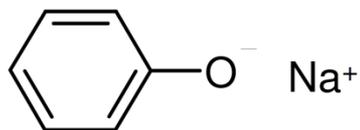


c12c3c4c1c5c4c3c25

Números maiores que 9 são antecidos por um '%' : %11

SMILES – Regras(6)

Compostos ligados não-covalentemente são separados por um “.”



[Na+].[O-]c1ccccc1

Isótopos:

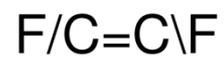
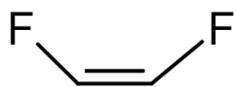
¹³C [13C]

¹³CH₄ [13CH4]

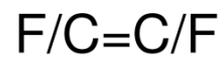
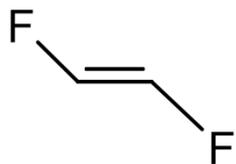
D₂O [2H]O[2H]

SMILES – Regras(7)

Configuração em torno de uma ligação dupla:



cis



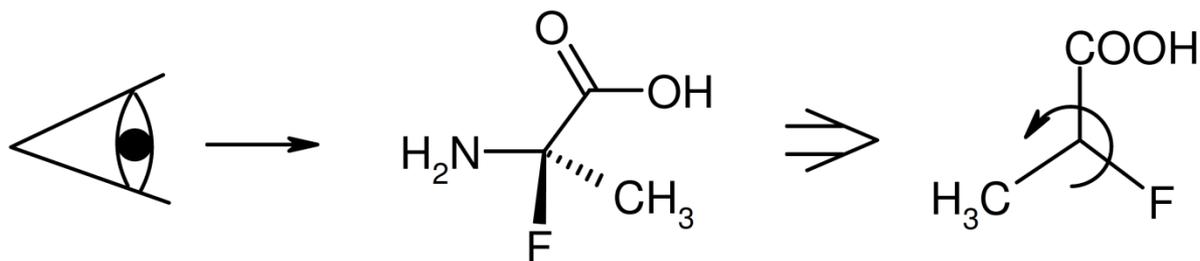
trans



Indeterminada

SMILES – Regras(8)

Quiralidade:



N[C@](C)(F)C(=O)O

@ - sequência anti-horária de substituintes

@@ - sequência horária de substituintes

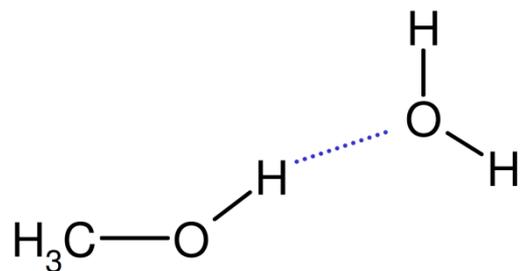
N.B. – Ausência de conformidade com o sistema (r,s) de representação absoluta da configuração

SMILES – Regras(9)

Hidrogénios explícitos:

H+ [H+] protão

H₂ [H][H]



CO[H][OH₂]

ligação de hidrogénio

SMILES – Regras(9)

As reacções químicas são representadas usando símbolo “>” :

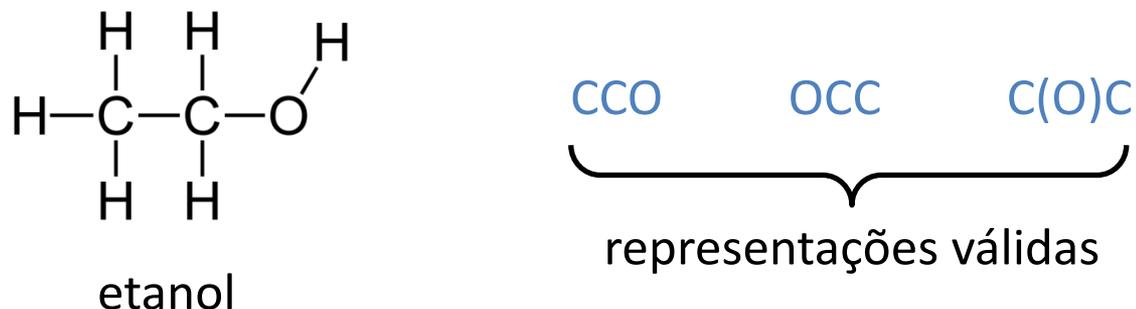
Reagentes > Agentes > Produtos

C.O=O > O=[O+]-[O-]> O=C=O.O combustão de metano na presença de ozono

CC(=[O:1])[OH:2].CC[OH:3] > [H+] > CC(=[O:1])[O:3]CC.[OH2:2] esterificação ácida do ácido acético e etanol

SMILES – Software

O problema da geração de SMILES a partir de estruturas não é trivial, pois geralmente existe mais do que uma representação SMILES válida para uma dada estrutura. Exemplo:



Para resolver este problema foram criados algoritmos de *canonização* que permitem gerar um SMILES único para cada molécula – SMILES canónico. Existem diversos packages de software que permitem gerar estes SMILES canónicos:

- Daylight Chemical Information Systems
- OpenEye Scientific Software
- Chemical Computing Group
- Chemistry Development Kit

SMARTS (1)

SMARTS (SMILES Arbitrary Target Specification):

generalização de SMILES que permite a representação de padrões moleculares. Os padrões são representados dentro de “[]”

Exemplo:

[F,Cl,Br,I] átomo que pode ser um F, Cl, Br ou I

Átomos:

c carbono aromático

a átomo aromático (C, N, O, S, ...)

A átomo alifático (não-aromático)

***** qualquer átomo (ou nenhum)

[#16] elemento nº 16 (qualquer tipo de enxofre)

[rn] átomo num anel de *n* membros

[SX2] enxofre com 2 substituintes —S— mas não $\begin{array}{c} \parallel \\ \text{—S—} \\ \parallel \end{array}$ ou =S

[Fe] átomo de ferro (carga arbitrária)

SMARTS (2)

Operadores lógicos:

A,B A ou B

A&B A e B

A;B A e B

!A não A

exemplos:

[F, Cl, Br, I]

F ou Cl ou Br ou I

[!C;R]

átomo aromático e não-alifático num anel

[CH2]

carbono alifático com 2 hidrogénios (metileno)

[c,n&H1]

carbono aromático ou NH aromático

[c,n;H1]

azoto ou carbono aromático e exactamente um

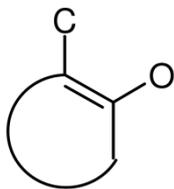
hidrogénio

[#7;r5]

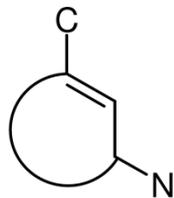
qualquer azoto num anel de 5 membros

SMARTS (3)

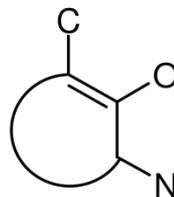
Configuração de substituintes:



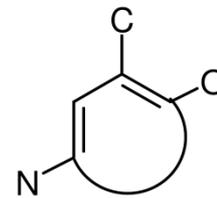
[CaaO]



[CaaaO]



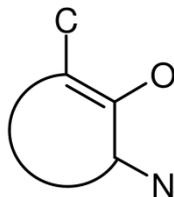
[Caa(O)aN]



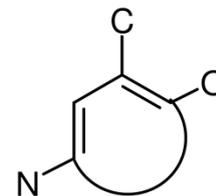
[Ca(aO)aaN]

O ambiente químico de um átomo pode ser especificado da seguinte forma:

C[\$(aaO);\$(aaaN)]



ou

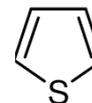


SMARTS (3)

Configuração de substituintes:

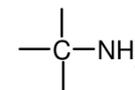
[s,o]1cccc1

tiofenos e furanos



[CX4][NH2]

aminas alifáticas primárias



[C1OC1]

epóxidos



C(=O)[OH,O-,O-.+]

ácido carbónico, carboxilato ou catião

C(=O)[NH1]

ligação peptídica

=[OH]

ácidos e enóis

F.F.F.F.F

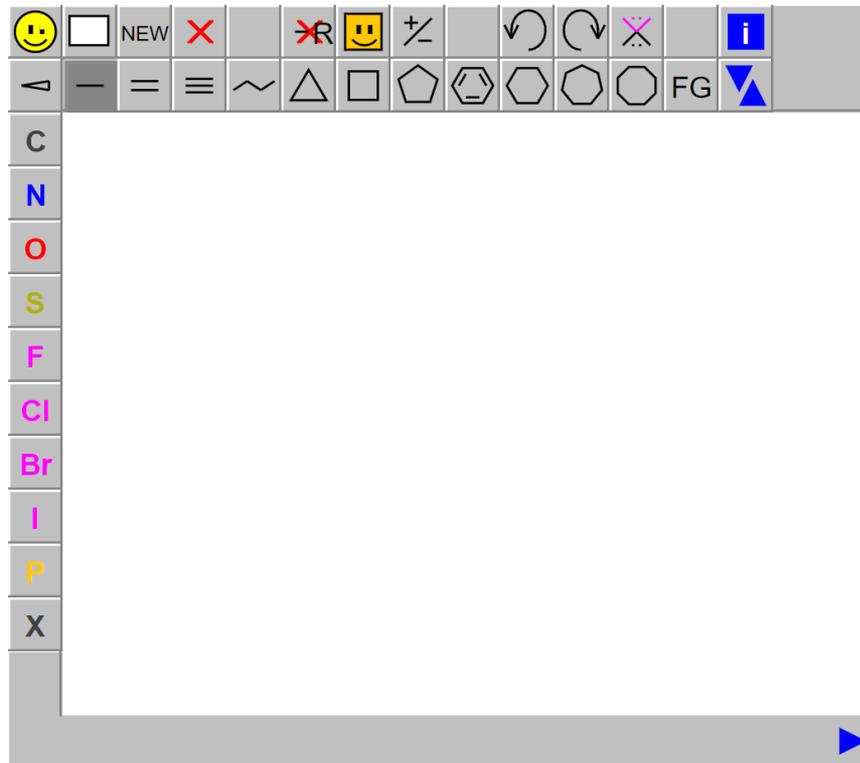
um total de 5 átomos de fluor as

SMIRKS

- Especificação de reacções
- Superset de SMILES
- Subset de SMARTS
- Possui mecanismos que não existem nas outras duas linguagens

SMIRKS Depiction	Reaction	SMIRKS and Note
	Reacting carbon	[C:1]>>[C:1] Agents aren't allowed in SMIRKS. The format is " reactants >> products ".
	Reacting Carbon (2-Connected)	[C;X2:1]>>[C;X2:1] SMIRKS allows atomic SMARTS expressions. The syntax is: [<SMILES_PART>;<SMARTS_PART>:<MAP>]
	No Reaction	[C;X4H3]-[CH2]C>>[C][CH2]C SMARTS atom specifications may be used for mapped atoms only (i.e. unmapped atoms must be valid SMILES expressions).
[NO REACTION]	No Reaction	[C;X2:1]~C>>[C;X2:1]=C SMIRKS doesn't allow SMARTS Bond Queries (e.g. ~). Bonds expressions must be valid SMILES.
	Just add water	>>O Upon transformation, all unmapped product-side SMILES get created.

JSME Molecular Editor & SMILES reader/generator



- Leitura de SMILES, SMARTS, SMIRKS, MOL, SDF
- Geração de SMILES canônicos
- Geração de InChI e InChKey
- Pesquisa de moléculas on-line através da InChKey

https://jsme-editor.github.io/dist/JSME_test.html

InChI Representation

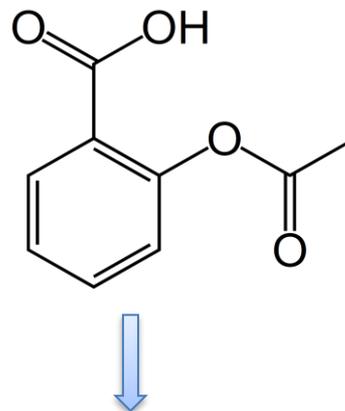
InChI – IUPAC International Chemical Identifier

Developed by IUPAC and NIST 2000-2005 (pronounced “In Key”)

InChI is a text-based identifier for chemical substances, designed to offer a standard way to provide molecular information

The InChI Identifier describes molecules in terms of different layers of information:

- Main Layer
 - Chemical Formula
 - Atom connections
 - Hydrogen Atoms
- Charge Layer
- Stereochemical Layer
- Isotopic Layer
- Fixed-H Layer
- Reconnected layer



InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12)

1-version number
S-standardized InChI

Chemical
Formula

Connectivity

Hydrogen Atoms

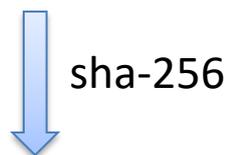
InChI and InChIKey

InChI's are too long and complex to reliably work as search keywords in database/internet searches.

InChIKey – Compressed form of the InChI, using a hashing algorithm (sha-256) to produce an quasi-unique alphabetic string with shorter length.

Different InChI's can produce the same InChIKey, but that's an extremely rare event.

InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12)



InChIKey

BSYNRYMUTXBXSQ-UHFFFAOYSA-N

Connectivity

Other layers

Protonation state

Version

Standard /Non-standard

InChI and InChIKey

InChI's are too
keywords in da

InChIKey – Con
algorithm (sha-
with shorter len

Different InChI'
extremely rare

InChI=1S/C9H8O

string

an

(1,11,12)

CHAR	PROTONS	CHAR	PROTONS
N	0		
M	-1	O	+1
L	-2	P	+2
K	-3	Q	+3
J	-4	R	+4
I	-5	S	+5
H	-6	T	+6
G	-7	U	+7
F	-8	V	+8

InChIKey

Connectivity

Other layers

Protonation state

Version

Standard /Non-standard

Exercícios

