

CoRS: Dynamic Information System for Small Molecules

Telukunta KK, Lucas X, Döring K, Flemming S, Günther, S

Pharmaceutical Bioinformatics, Institute of Pharmaceutical Sciences, University of Freiburg, Germany

kiran.telukunta@pharmazie.uni-freiburg.de

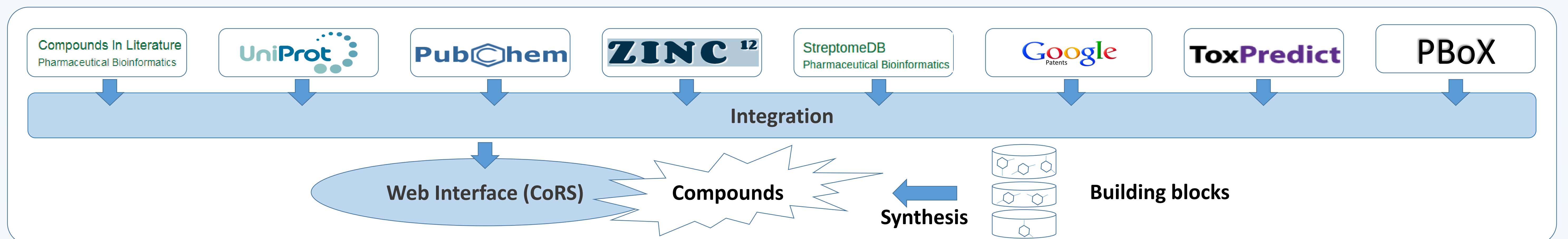


What is CoRS ?

An important task in the field of pharmaceutical sciences is the analysis of the biological effects of small molecules[1]. To identify potential new drugs or to assess health risks from chemicals requires prior knowledge of compounds. The ongoing Comprehensive Research information System (CoRS) for small molecules project is aiming at the **integration** of existing data resources of molecules that are combined with tools for the **prediction** of molecular effects. CoRS integrates databases such as PubChem[1], StreptomeDB[2], CIL[3], PBoX[4], ToxPredict[5] and UniProt[6], and has the capability to search for compounds that can be easily **synthesized** via building blocks.

Aim

To have a system which gathers data of compounds from publicly available resources, and **displays a digest of information** at one place. The information is utilized for pharmaceutical researchers working in the field of drug discovery.



“Type in compound name or draw a chemical structure ...”

Has the compound been tested in bioassays ?

Bioassays data is extracted from the publicly available database PubChem, which consists of three interconnected databases : Substance, BioAssay and Compound[7]. Here the information is displayed with assay identifier and details of bioassays can be retrieved upon browsing.

Is the compound purchasable ?

Millions of molecules are commercially available. The **Purchasability** gives information about vendors. These data in CoRS is integrated from ZINC[8] and PBoX[4] sources and displayed in this tab.

Is the compound patented ?

For a newly developed compound, patent information will be crucial. CoRS takes information from Google Patents and displays the **Patents** information. All the documents available through Google Patents originate from US Patent and Trademark Office.

Similar compounds difference is highlighted by yellow color

Proteins | Bioassays | Toxicity | Purchasability | Patents

compounds

proteins

Granulocyte-macrophage colony-stimulating factor (P04141)

Elafin (P19957)

UDP-galactose 4-epimerase (P78381)

Images from &

Which proteins co-occurred with compounds in literature?

Proteins displays information of CIL[3], which gives the related proteins of the given compound which are mentioned in the literature database PubMed. This information accelerates the time-consuming process of literature research.

Proteins | Bioassays | Toxicity | Purchasability | Patents

Oxazepam 2D - 3D

#	AB	AC1q[M]	AC1q[M]	AC Range	Bioassay (Bioactive Target)	Protein Target
1	720637	1		0.38176 [µM]	qRTS assay for small molecule disruption of the ataxochlorella membrane potential (Summary) (Summary)	Ataxochlorella (Chlorella sp.) (Homo sapiens) (P174994)
2	1487	1		0.4487 [µM]	qRTS Assay for Modulators of Lamin A Splicing (Continuation)	Lamin A/C isoform 3 (Homo sapiens) (P174994)
3	606070			20.8554 - 48.5577 [µM]	qRTS for inhibitors of human tyrosyl-DNA phosphodiesterase 1 (TDP1) qRTS in cells in absence of CPT (Continuation)	TDP1 protein (Homo sapiens) (P191404)
4	606070			20.8554 - 48.5577 [µM]	qRTS for inhibitors of human tyrosyl-DNA phosphodiesterase 1 (TDP1) qRTS in cells in presence of CPT (Continuation)	TDP1 protein (Homo sapiens) (P191404)
5	1760			38.8387 [µM]	qRTS Assay for Inhibitors Targeting the Human MLL Domains in MLL Related Leukemias - Competition With Test Bed Labeled MLL-derived Muramyl Peptide (Continuation)	MLL2 gene product (Homo sapiens) (P191404)

Extracted from PubChem

“Information retrieval of small molecules made easy!”

Proteins | Bioassays | Toxicity | Purchasability | Patents

Nordiazepam 2D - 3D

Predictions Datasets

Run All

[CADASTER FP7] Algae_UIDRAGON_SPLIT Calculate

EC50 aquatic_INAPDOMAIN YES

EC50 aquatic_log[M] 4.60

EC50 aquatic_ACCURACY 0.40

[CADASTER FP7] HMGU: EC50 Algae (TAZ & BTAZ) Calculate

EC50 aquatic_ACCURACY 0.40

EC50 aquatic_INAPDOMAIN YES

EC50 aquatic_log[mol/L] -4.40

[CADASTER FP7] HMGU: LC50 Fish (TAZ & BTAZ) Calculate

Image from ToxPredict

Is the compound toxic?

The **Toxicity** prediction information is provided by the OpenTox community through ToxPredict web service.

Proteins | Bioassays | Toxicity | Purchasability | Patents

ZINC00000431

In ZINC since July 23rd, 2004

Heavy atoms 21

Benign functionality Yes

Popular Name: Lorazepam

PubMed - Wikipedia - Google

CAS Numbers: 110032-65-0, 846-49-1, 907200-04-8

Vendors

Active BioPharma 696

Amadis Chemical A840886

BePharm Building Blocks B20035

ChiralBlock BioScience BB 736

Sequoia Research Products 18878-17-6

Annotations

BindingDB.org

ChEMBL DrugSt

ChEMBL12

ChEMBL13

Image from ZINC

The priority of the display of similar compounds is done based on users preferences, which can be suggested by the user **dynamically** using the tool and prioritization method takes into account the number of bioassays, toxicity prediction, purchasability data and patenting information.

Proteins | Bioassays | Toxicity | Purchasability | Patents

Patents Application Grant

Transnasal microemulsions containing diazepam

CA 2529489 C

Publication number CA2529489 C

Application type Grant

Application number CA 2529489

PCT number PCT/US2004/01424

Publication date Jun 15, 2005

Filing date Jun 15, 2004

Priority date (2) Jun 27, 2003

Also published as CA2529489A1, 6 More +

Inventors Yong Moon Choi, Kwan-Ho Kim

Applicant Yong Moon Choi, Kwan-Ho Kim

Classifications (17), Legal Events (2)

External Links: CIPAC, Espacenet

ABSTRACT

Diazepam is administered intranasally in the form of specific microemulsions having advantageous properties. The microemulsions are comprised of about equal quantities of a fatty acid and ester with the remainder being a hydrophilic surfactant, a polar solvent and an alcohol in a weight ratio such that alcohol is present in a greater quantity by weight than either of the other two. Nasal administration of the subject microemulsions produces a high plasma concentration of diazepam nearly as fast as intravenous administration. The present microemulsions are particularly suitable for a pre-empt and timely treatment of patients in the acute and/or emergency treatment of status epilepticus and other fever-induced seizures.

Image from Google



Pharmazeutische Bioinformatik

http://www.pharmazeutische-bioinformatik.de

- [1] Bolton E. et al.: **Integrated Platform of Small Molecules and Biological Activities**. *Annual Rep Comput Chem* 2008, 4: 217-241.
- [2] Lucas X, Senger C. et al.: **StreptomeDB: a resource for natural compounds isolated from Streptomyces species**. *Nucleic acids res* 2013, 41(D1): D1130-D1136.
- [3] Grüning BA, Senger C. et al.: **Compounds In Literature (CIL): screening for compounds and relatives in PubMed**. *Bioinformatics* 2011, 27(9): 1341-2.
- [4] Lucas X., et al.: (submitted).
- [5] www.toxpredict.org
- [6] The UniProt Consortium: **Update on activities at the Universal Protein Resource**. *Nucleic acids res.* 2013, 41: D43-D47
- [7] Wang Y. et al.: **An overview of the PubChem BioAssay resource**. *Nucleic acids res.* 2012, 38: D255-D266.
- [8] Irwin JJ. et al.: **ZINC: A Free Tool to Discover Chemistry for Biology**. *J. Chem Inf Model* 2012, 52(7): 1757-1768.



UNI-FREIBURG