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

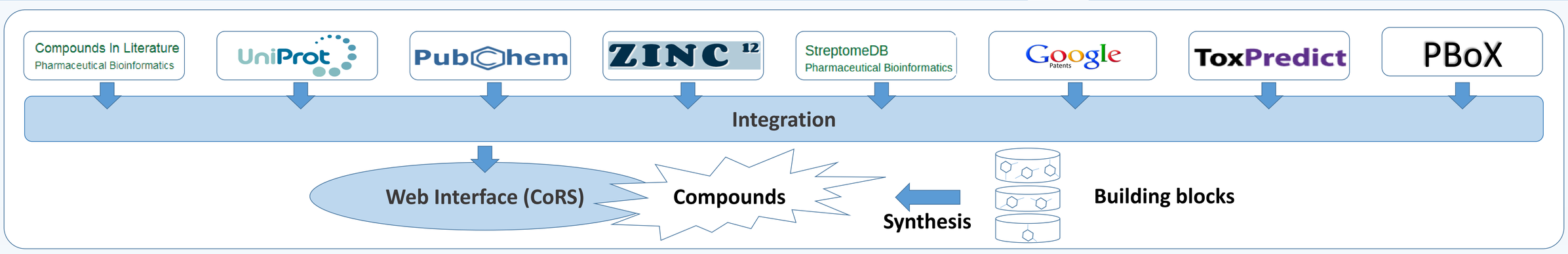
kiran.telukunta@pharmazie.uni-freiburg.de

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.

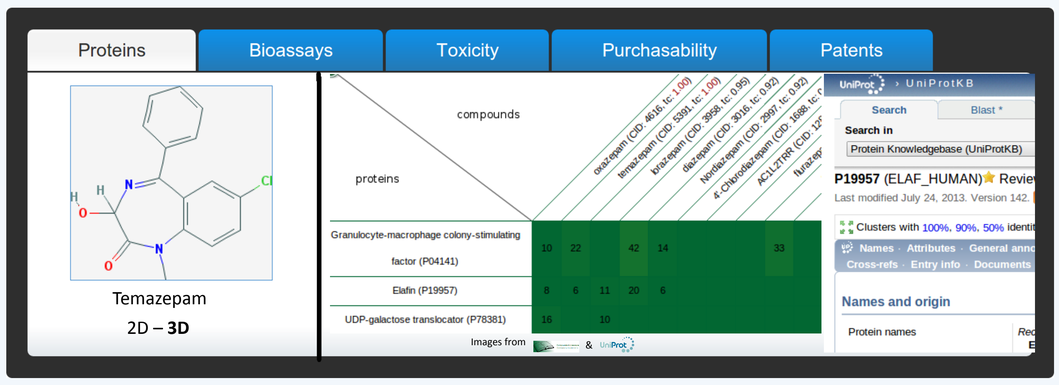
Is the compound purchasable ?

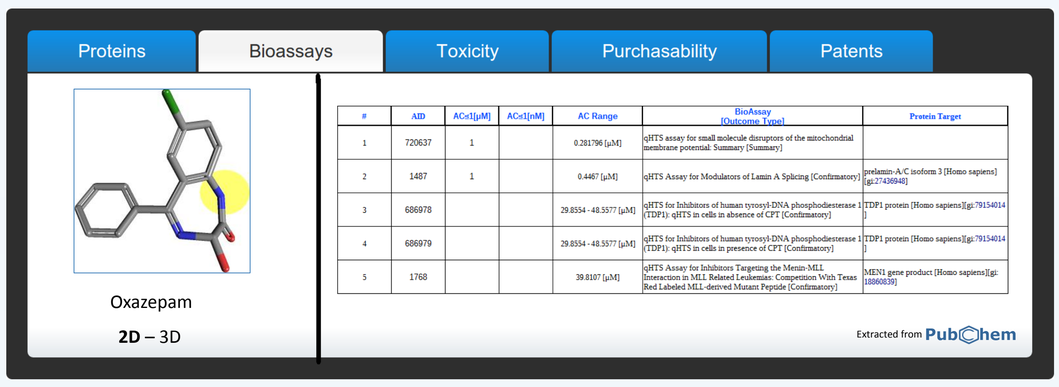
Millions of molecules are commercially available. The **Purchasability** gives information about vendors. These data in CoRS is integrated from ZINC[8] and PBoX[4] sources and displayed in this tab.

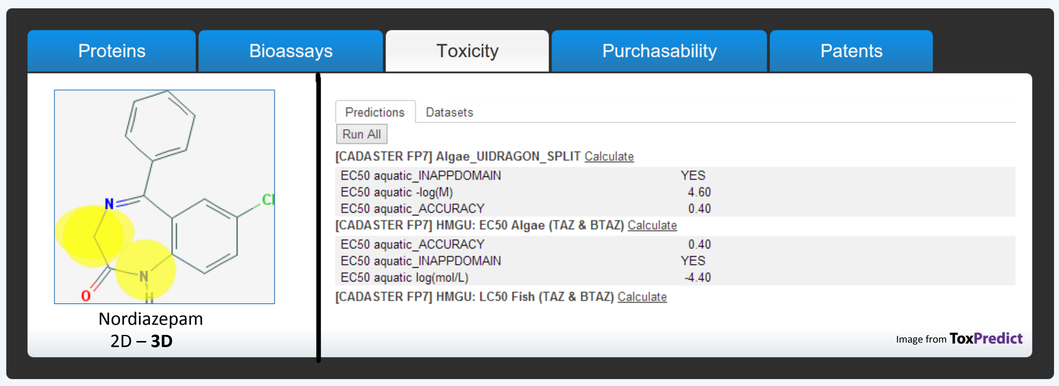
Is the compound patented ?

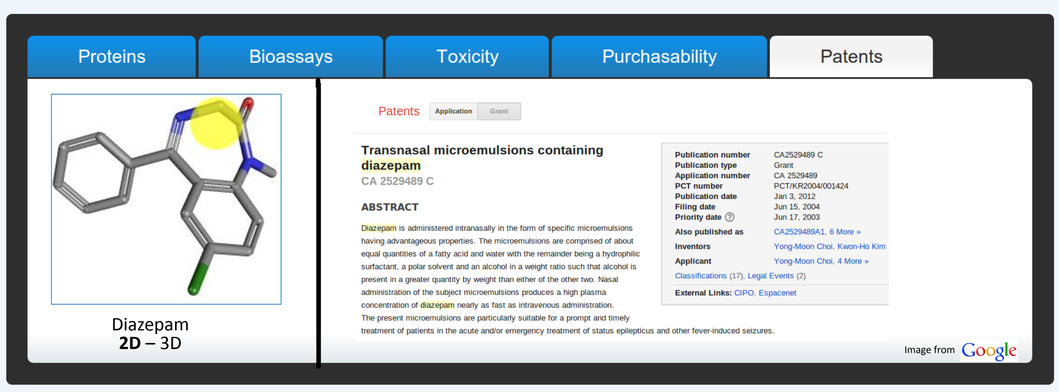
For a newly developed compound, patent information will be crucial. CoRS takes information from Google Patents and displays the **Patents** information. All the documents available through Google Patents originate from US Patent and Trademark Office.

Similar compounds difference is highlighted by yellow color









Which proteins co-occured 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 lit-erature research.

"Information retrieval of small molecules made easy!"

Is the compound toxic?

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

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 predicti-on, purchasability data and patenting information.

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-D66.
[8] Irwin JJ. et al.: **ZINC: A Free Tool to Discover Chemistry for Biology.** *J. Chem Inf Model* 2012, 52(7): 1757-1768.

DFG JUNIOR RESEARCH GROUP PROGRAMME

UNI FREIBURG