SuperNatural: A Database of available natural compounds

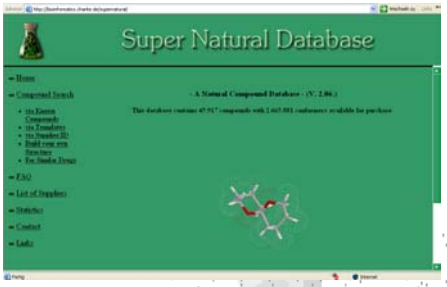
Melanie Füllbeck, Mathias Dunkel and Robert Preissner

3D Data Mining Group, Berlin Center for Genome Based Bioinformatics, Institute of Biochemistry, Charité – Universitätsmedizin Berlin, Monbijoustr. 2, 10117 Berlin

melanie.fuellbeck@charite.de, mathias.dunkel@charite.de, robert.preissner@charite.de

Abstract: Though tremendous effort has been put into synthetic libraries, the majority of marketed drugs are still natural compounds or derivatives thereof. Although there are encyclopaedias on natural compounds, the compounds´ availability is often unclear, and catalogues from numerous suppliers have to be checked. To overcome these problems we have compiled a database of ~ 50,000 natural compounds from different suppliers. Starting point for *in silico* screenings are about 2,500 well-known and classified natural compounds from a compendium that we have added, typical templates or imported molecules. Possible medical applications can be detected via automatic searches for similar drugs in a free conformational drug database containing WHO-indications. Furthermore, we have computed about three million conformers, which are deployed to account for the flexibility of the compounds during usage of the 3D superposition algorithm that we have developed.

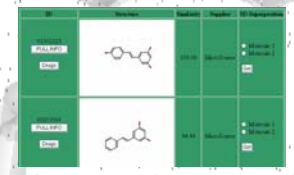
Availability: http://bioinformatics.charite.de/supernatural



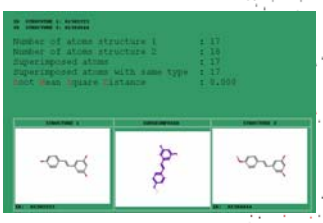
The SuperNatural Database [1] is the first public resource containing 3D structures and conformers of 45,917 natural compounds, derivatives and analogues purchasable from eight different suppliers.

As a starting point for screenings we compiled a searchable compendium of 2,500 well-known natural compounds characterized by CAS-number, which is useful to cross-referencing other databases. This compendium contains systematic names, classification codes, empiric formulae, mixtures and synonyms.

The marvin applet allows the user to build or import a molecular structure and compare it with compounds from the SuperNatural Database.



Similarity searches, which are based on fingerprints and Tanimoto coefficients, are implemented in the SuperNatural Database.



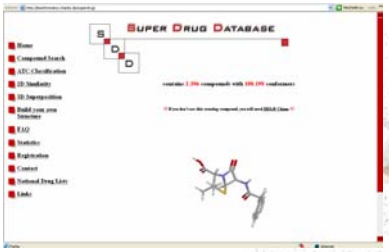
An algorithm, which was developed in our group [2], enables a 3D superposition of two compounds to be made. The algorithm compares all conformers of two compounds to find the best structural alignment.

Conclusions: The chemical diversity and unique properties of natural compounds provide a promising starting-point for developing innovations for scientific, medical and nutritional applications. The SuperNatural Database is a free resource with embedded screening functions for bioactive natural compounds. The extension of the database allows the scientific community simple access to a growing number of available natural compounds.

Natural Compound	2D-Structure	Similar Compounds in SuperNatural (Tanimoto >= 0.85)	Status
Neurological disease area / V03* (All other therapeutic products)			
Morphine		9	Lead compound of Nalorphine
Immunological, inflammatory / L04* (Immunosuppressive agents)			
Tacrolimus		2	Active agent of Tacrolimus (FK -506)
Oncological disease area / L01* (Antineoplastic agents)			
Protopanaxadiol		16	Phase I clinical trials
Triptolide		18	Phase I clinical trials

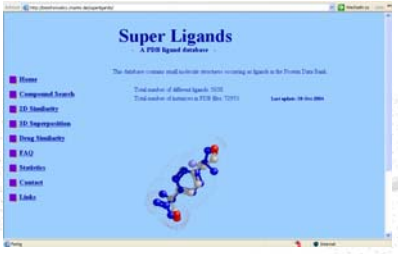
Similarity Screenings in the SuperNatural Database with natural compounds that are known as drugs, from clinical trials or lead compounds for drug development. The database contains a great number of compounds with calculated 2D similarities of ≥ 0.85. The results show that the SuperNatural Database is an excellent source for finding bioactive natural products.

Related Databases:



SuperDrug Database [3]: The database contains ~ 2,500 3D-structures of active ingredients of marketed drugs. An assignment of the Anatomical Therapeutic Chemical (ATC) classification codes to each structure according the WHO scheme is possible.

Availability: http://bioinformatics.charite.de/superdrug



SuperLigands Database [4]: SuperLigands supplements the set of existing resources of information about small molecules bound to PDB structures. Allowing for 3D-comparison of the compounds as a novel feature, this database represents a valuable means of analysis and prediction in the field of biological and medical research.

Availability: http://bioinformatics.charite.de/superligands

References

[1] M Dunkel, M Fullbeck, S Neumann, R Preissner: SuperNatural: a searchable database for available natural compounds. Nucleic Acids Res 2006: Database Issue, in press.
[2] M Thimm, A Goede, S Hougardy, R Preissner: Comparison of 2D similarity and 3D superposition. Application to searching a conformational drug database. J Chem Inf Comput Sci 2004, 44: 1816-22
[3] A Goede, M Dunkel, N Mester, C Frommel, R Preissner: SuperDrug: a conformational drug database. Bioinformatics 2005, 21: 1751-3.
[4] E Michalsky, M Dunkel, A Goede, R Preissner: SuperLigands – a database of ligand structures derived from the Protein Data Bank. BMC Bioinformatics 2005, 6: 122.

This work was done within the context of the NGFN Nationales Genomforschungsnetz and is supported by Projektträger Jülich