

POSTER PRESENTATION

Open Access

Extended graph-based models for enhanced similarity retrieval in Cavbase

Timo Krotzky^{1*}, Thomas Fober^{2*}, Marco Mernberger², Gerhard Klebe¹, Eyke Hüllermeier²

From 8th German Conference on Chemoinformatics: 26 CIC-Workshop
Goslar, Germany. 11-13 November 2012

The problem of estimating the similarity between molecular structures is often tackled by means of graph-based approaches, using graphs for structure representation and measures based on the maximum common subgraph as similarity metrics. In the case of protein binding sites as molecular structures, however, where the graphs can be very large, the computation of these measures may easily become infeasible or at least unacceptably slow.

To this end, Cavbase [1,2] was developed, a database for the automatic detection and storage of putative binding sites on the protein surfaces. Cavbase assigns so-called pseudocenters to the cavity-flanking amino acids, which characterize their physicochemical properties with respect to molecular recognition. On the one side, this representation leads to smaller and more generic representation of a binding site. On the other side, it comes with a loss of information, which is usually compensated by performing further calculations based on additional data. These steps, however, are most often computationally quite demanding, making the whole approach again very slow.

The main drawback of a graph-based model solely based on pseudocenters is the loss of information about the shape of protein surface. In this study, we propose an extended modeling formalism that leads to graphs of the same size, but containing considerably more information. More specifically, additional descriptors of the surface characteristics are extracted from the surface points stored in Cavbase. These properties are included as attributes of the nodes of the graph, which leads to a gain of information and allows for more accurate comparisons between different structures.

* Correspondence: krotzky@uni-marburg.de

¹Institute of Pharmaceutical Chemistry, Philipps-Universität, 35032 Marburg/Lahn, Germany

²Institute of Mathematics and Computer Science, Philipps-Universität, 35032 Marburg/Lahn, Germany

Full list of author information is available at the end of the article

Author details

¹Institute of Pharmaceutical Chemistry, Philipps-Universität, 35032 Marburg/Lahn, Germany. ²Institute of Mathematics and Computer Science, Philipps-Universität, 35032 Marburg/Lahn, Germany.

Published: 22 March 2013

References

1. Schmitt S, Kuhn D, Klebe G: A New Method to Detect Related Function Among Proteins Independent of Sequence and Fold Homology. *J Mol Biol* 2002, **323**:387-406.
2. Hendlich M, Rippmann F, Barnickel G: LIGSITE: Automatic and efficient detection of potential small molecule-binding sites in proteins. *J Mol Graph Model* 1997, **15**:359-363.

doi:10.1186/1758-2946-5-S1-P29

Cite this article as: Krotzky et al.: Extended graph-based models for enhanced similarity retrieval in Cavbase. *Journal of Cheminformatics* 2013 **5**(Suppl 1):P29.

Publish with **ChemistryCentral** and every scientist can read your work free of charge

"Open access provides opportunities to our colleagues in other parts of the globe, by allowing anyone to view the content free of charge."

W. Jeffery Hurst, The Hershey Company.

- available free of charge to the entire scientific community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
<http://www.chemistrycentral.com/manuscript/>


ChemistryCentral