

POSTER PRESENTATION

Open Access

Searching substructures in fragment spaces

H-C Ehrlich*, M Rarey

From 6th German Conference on Chemoinformatics, GCC 2010
Goslar, Germany. 7-9 November 2010

Introduction

Fragment spaces (FSs) are an elegant way to model a large or even infinite number of chemical compounds and their synthetic accessibility. A FS consists of molecular fragments and a set of rules defining how fragments can be combined to products. In virtual screening experiments, FSs might include products with undesired functional groups or inadequate central building blocks. The recognition of such products, especially when they span over multiple fragments, would require their explicit construction from the FS. Due to the generally huge number of possible products in an FS, the complete enumeration is undesired or even impossible. Therefore, algorithms that perform substructure search in FSs must be able to process fragments and joining rules rather than complete molecules. Even though some algorithms that work in FSs exist [1,2], a method that excludes undesired products via substructure definition from a FS is still missing.

Method

We present and compare two algorithms to modify an FS such that no possible product can include a given functional group or substructure. The methods utilize a search procedure based on the Ullmann [3] respectively the VF2 algorithm [4] for subgraph isomorphism. Thereby, we find substructures that are present inside fragments or would be formed by joining two fragments. After the identification of such fragments, they are either removed from the FS or their joining rules are altered in a way that a formation of the substructure becomes impossible.

Results

The algorithms are tested on the BRICS fragment space [1]. We exclude substructures described by SMARTS

patterns that were collected from literature [5]. The experiments show that the VF2 approach is superior in running time.

Published: 19 April 2011

References

1. Degen J, Wegscheid-Gerlach C, Zaliani A, Rarey M: **On the art of compiling and using 'drug-like' chemical fragment spaces.** *ChemMedChem* 2008, **3**(10):1503-1507.
2. Rarey M, Stahl M: **Similarity searching in large combinatorial chemistry spaces.** *J Comput Aided Mol Des* 2001, **15**(6):497-520.
3. Ullmann JR: **An algorithm for subgraph isomorphism.** *J Assoc Comput Mach* 1976, **23**:31-42.
4. Cordella LP, Foggia P, Sansone C, Vento M: **A (sub)graph isomorphism algorithm for matching large graphs.** *IEEE T-PAMI* 2004, **26**(10):1367-1372.
5. Schomburg K, Ehrlich H-C, Stierand K, Rarey M: **From structure diagrams to visual chemical patterns.** *J Chem Inf Model* 2010.

doi:10.1186/1758-2946-3-S1-P11

Cite this article as: Ehrlich and Rarey: Searching substructures in fragment spaces. *Journal of Cheminformatics* 2011 **3**(Suppl 1):P11.

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

* Correspondence: ehrich@zbh.uni-hamburg.de
Center for Bioinformatics, University of Hamburg, 20146 Hamburg, Germany