Chemistry Central Journal

Poster presentation

Incorporating QSPR in the enumeration of fragment space
J Pärn* and M Rarey

Address: Center for Bioinformatics, University of Hamburg, Germany
* Corresponding author

This abstract is available from: http://www.journal.chemistrycentral.com/content/3/S1/P18

The generation of novel bioactive compounds based on existing lead series is a frequently occurring challenge in drug development programs. Bioactivity itself has different aspects which have to be taken into account. While rules based on simple descriptors like Lipinski's rule of five [1] for bioavailability exist, in most cases pharmacologically important properties have to be predicted via statistical models since no direct calculation method is known.

For the lead search and optimization phase, the representation of the search space is of crucial importance. Chemical fragment spaces are a relatively new and a promising approach to model the chemical space in a combinatorial way. A chemical fragment space consists of a set of molecular fragments and a set of rules [2][3]. Each fragment has one or several link atoms, each having a certain type. The set of rules primarily defines which link types are compatible to each other. New chemical entities are generated by connecting fragments using the link atoms according to the compatibility definition.

Based on the idea of recombining fragments of bioactive compounds, we developed a program to enumerate fragment spaces. Since a complete enumeration is in most cases not possible and not desirable our program only enumerates certain parts of a fragment space. The user can define which part of the space should be enumerated by providing min-max ranges for physicochemical constraints which the resulting molecules have to obey [4]. To take properties into consideration which can not be directly derived from simple descriptors, we implemented and incorporated a PLS based QSPR prediction directly into our enumeration engine. The QSPR-model derived properties can either be written out with the molecules or they can be used as filter in the enumeration. As descriptor for the QSPR-model we use a reduced graph representation of a molecule. We reduce the number of different atoms to six pharmacophoric types and count the number of pairs over topological distances.

The integration of QSPR models into the enumeration methodology allows to create molecules from fragment spaces lying within user specified property and QSPR-model ranges. The method can be applied in the lead identification process but might also be useful to study the limitation of QSPR models by creating large, diverse compound sets falling into the same range of predicted QSPR values.

References
1. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ: *Adv Drug Deliv Rev* 2001, **46(1–3):**3-26.
2. Lewell XQ, Judd DB, Watson SP, Hann MM: *J Chem Inf Comput Sci* 1998, **38(3):**511-522.
3. Rarey M, Stahl M: *J Comput Aided Mol Des* 2001, **15(15):**497-520.
4. Paern J, Degen J, Rarey M: *J Comput Aided Mol Des* 2007, **21(6):**327-340.