Chemistry Central Journal

Poster presentation

Assessment of complex SARs using Formal Concept Analysis of fragment combinations

E Lounkine, J Auer and J Bajorath*

Address: Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universtität Bonn, Dahlmannstr. 2, D-53113 Bonn, Germany

* Corresponding author

This abstract is available from: http://www.journal.chemistrycentral.com/content/3/S1/P3

Molecular fragments are widely recognized as powerful descriptors having the advantage of being chemically interpretable [1]. Several methodologies have been introduced to annotate molecular fragments with activity information based on their occurrence and/or abundance in bioactive compounds [2][3]. However, these methods usually do not focus on comparison of several activity classes.

We present FragFCA, a fragment-based approach that utilizes Formal Concept Analysis [4], a data mining technique adapted from information theory. FragFCA makes it possible to generate user-defined complex SAR queries spanning multiple activity classes and incorporating potency information. Using the freely available ToscanaJ software package, we have designed a graphical interface allowing the interactive assembly of queries of varying levels of complexity [5].

Using a publicly available data set of GPCR inhibitors belonging to seven different classes with overlapping activity, we have analyzed the ability of fragment combinations to discriminate between different activities and/or potency ranges. Fragment pairs and triplets, rather than individual fragments, are shown to contain most selectivity information. Moreover, using FragFCA, we have been able to identify fragment combinations that successfully distinguish selective from non-selective cathepsin L inhibitors in HTS data. This suggests that FragFCA can be applied to extract fragment combinations on the basis of training sets for the prediction of compound selectivity.

References

1. Erlanson DA, McDowell RS, O'Brien T: *J Med Chem* 2004, **47:**3463-3482.
2. Schnur DM, Hermsmeier MA, Tebben AJ: *J Med Chem* 2006, **49:**2000-2009.
3. Sutherland JJ, Higgs RE, Watson I, Vieth M: *J Med Chem* 2008, **51:**2689-2700.
4. Priss U: *Ann Rev Inf Sci Technol* 2006, **40:**521-543.
5. Lounkine E, Auer J, Bajorath J: *J Med Chem* 2008 in press.