

# **Novel Algorithms for Molecular Similarity Analysis and Virtual Screening**

Jürgen Bajorath

Department of Life Science Informatics, B-IT, Rheinische Friedrich-Wilhelms-Universität, Dahlmannstr. 2, D-53113 Bonn, [bajorath@bit.uni-bonn.de](mailto:bajorath@bit.uni-bonn.de)

A class of methods termed mapping algorithms is introduced for the evaluation of molecular similarity and application in large-scale virtual screening. Despite algorithmic differences, these methods share similar principles. They are designed to map database compounds to activity-dependent consensus positions in chemical space or to combinations of activity class-specific descriptor value ranges. Mapping algorithms operate in high-dimensional chemical reference spaces that are formed by binary-transformed or original molecular descriptors. During compound mapping, the dimensionality of reference spaces is not held constant but gradually increased in order to separate active and inactive compounds. In virtual screening trials, mapping algorithms produce significant hit and recovery rates and recognize remote molecular similarity relationships. For many activity classes, active molecules are enriched in small selection sets containing 10-50 database compounds.