Abstract

***Chemom. Intell. Lab. Syst.*, 67**, 95-108 (2003)

Scsibrany H., Karlovits M., Demuth W., Müller F., Varmuza K.:

Clustering and similarity of chemical structures represented by binary substructure descriptors.

A set of 1365 substructures has been created for the representation of organic chemical structures by binary substructure descriptors. Software SubMat calculates a matrix of binary fingerprint vectors for given sets of molecular structures and substructures (both in Molfile format). The distribution of the substructures in two spectral databases (IR, MS) has been investigated.

Examples of structure similarity searches based on the Tanimoto index are presented. Cluster analyses of chemical structures by principal component analysis (PCA) have been performed for hitlists from spectral similarity searches.