[P36] Peptide Database Analysis via 3D Shape Fingerprint

Xian Jin, Mahendra Awale, Jean-Louis Reymond

Dept. of Chemistry & Biochemistry, University of Bern, Freiestrasse 3, 3012 Bern, Switzerland

Mining of chemical space for drug discovery, is often faciliated by the various chemoinformatics approaches; popular amongst them are similarity search, scaffold analysis and visualization of compound libraries. To aid in this process, recently we proposed two property spaces 1) Molecular Quantum Numbers (MQN)[1] and 2) SMIfp[2]. Both MQN and SMIfp provided reference feature spaces with capability for visualization of large databases via color coded principle component plane (PC-plane)[3]. However, MQN and SMIfp descriptors do not take into account shape or pharmacophore information from molecules. Furthermore, the use of MQN and SMIfp is so far limited to the relatively small molecules.

In the presented study here, our focus was to develop the property space, which can be applicable to molecules with wide range of size and be able to encode the shape of molecules. Thus, we have designed a simple yet efficient 3D shape fingerprint which was incorparated based on atom pair count within predefined distances in 3D space. The fingerprint was applied to represent the property space of three peptide databases containing molecules with heavy atom count up to 1000. For comparison, compounds from Cambridge Structural Database (CSD) were included in the study.

The 3D shape fingerprint of databases were hashed on to the PC-plane and color coded according to various molecules properties. Another shape analysis, Principal Moment of Inertia (PMI) was carried out to evaluate the 3D shape fingerprint. Clear separation of molecules on the PC-plane according to different databases was observed. This preliminary analysis revealed the shape encoding capability of the designed 3D shape fingerprint. On the other hand, a Java application called Mapplet was applied for interactive visualization of the chemical space of the databases used in the study.

[1] K.T. Nguyen; CemMedChem. 4 (2009) 1803-1805.
[2] J. Schwartz; J. Chem. Inf. Model. 53 (2013) 1979-1989.
[3] M. Awale; J. Chem. Inf. Model. 53 (2013) 509-518.