P72

2-Dimensional substructural fingerprints – a novel method of compound structure representation

Krzysztof Rataj[1], Wojciech Czarnecki[2], Sabina Smusz[1], Andrzej J. Bojarski[1]

[1]Department of Medicinal Chemistry, Institute of Pharmacology, Polish Academy of Sciences, 12 Smętna Street, Kraków, Poland
[2]Faculty of Mathematics and Computer Science, Jagiellonian University, 6 Łojasiewicza Street, Kraków, Poland

In silico methods are becoming more and more popular in drug screening efforts, allowing for a significant increase in the speed of the process, together with a major decrease of its costs. Many of those computational methods are based on the structure of the processed chemical compounds, which must be transformed from a 2-dimensional structural representation into a computer-readable form, preferably a bit string. In case of substructural fingerprints, each bit of such bit string represents the existence of a predefined substructure within the target compound [1]. There are many such representations [2,3], however they all share the same flaw: a complete loss of information about the connectivity between substructures and atoms. This often leads to curious cases, where two structurally diverse compounds are represented with identical bit string, hence rendering their differentiation impossible.

To address this issue, we created a prototype of a novel representation of chemical compounds – a 2-dimensional substructural fingerprint (2DFP). This method does not lose so much needed connectivity data, enabling a more complete depiction of the target compound, while retaining the form needed for computational purposes. The 2DFP methodology has been tested on a small set of 5-HT6R ligands, to determine its efficiency in discrimination between active and inactive compounds. The initial results show a major increase of MCC score of discrimination tests, compared to currently available substructural fingerprints.

[1] Barnard, J. M. & Downs, G. M. *J. Chem. Inf. Model.* 37, 141–142 (1997).
[2] Ewing, T., Baber, J. C. & Feher, M. *J. Chem. Inf. Model.* 46, 2423–2431
[3] Klekota, J. & Roth, F. P. *Bioinformatics* 24, 2518–25 (2008).