[P11] A rotation-translation invariant molecular descriptor and its use in ligand-based virtual screening

Francois Berenger, Kam YJ Zhang

Structural Bioinformatics Team, Division of Structural and Synthetic Biology, Center for Life Science Technologies, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan.

Measures of similarity for chemical molecules have been developed since the dawn of chemoinformatics. Similarity measures are widespread in practice and have proven to be useful in drug discovery. Because of our interest in high throughput ligand-based virtual screening, we sought to exploit the information contained in 3D coordinates and atomic properties of a molecule.
A molecular descriptor is proposed, using the autocorrelation function[1] (figure 1) to encode atomic properties of a whole molecule. An associated scoring function is also proposed to rank-order a database of compounds versus a query molecule. The proposed method can work in one or a combination of the following feature spaces: partial charges, atomic radii and atomic contribution to calculated LogP. It can also perform consensus ranking and automatic query (molecule) optimization. An open source implementation of the method is released. Extensive retrospective ligand-based virtual screening experiments were performed on DUD-E[2], along with comparison against other methods in order to validate the software and protocol.

While it is a simple method, it performs remarkably well in terms of speed, AUC and early enrichment rate[3].

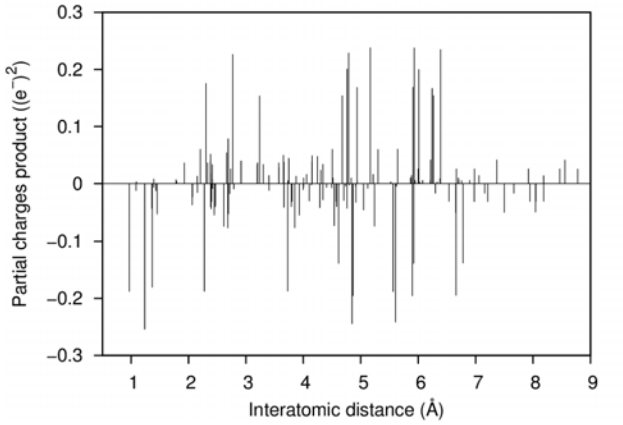

Figure 1: autocorrelogram of a molecule over its partial charges as a set of Kronecker deltas.

[1] Moreau, G., Broto, P. Nouveau Journal de Chimie 4(6) (1980), 359–360.
[2] Mysinger, M.M., Carchia, M., Irwin, J.J., Shoichet, B.K. Journal of Medicinal Chemistry 55(14) (2012), 6582–6594
[3] F. Berenger, A. Voet, XY Lee, K YJ Zhang . Journal of Cheminformatics (2014), submitted manuscript.