SiteAlign: a simple and fuzzy method to measure similarity between binding sites

C. Schalon, J.-S. Surgand, E. Kellenberger, D. Rognan
Bioinformatics of the drug, UMR CNRS-ULP 7175, Faculty of Pharmacy of Strasbourg
74 route du Rhin, BP 24, 67401 Illkirch Cedex, France

A novel method to measure distances between druggable protein cavities is presented [1]. Starting from user-defined ligand binding sites, a set of eight topological and physicochemical properties are projected from cavity-lining protein residues to an 80 triangle-discretised sphere placed at the centre of the binding site, thus defining a cavity fingerprint. Representing binding site properties onto a discretised sphere presents many advantages: (i) a normalised distance between binding sites of different sizes may be easily derived by summing up the normalised differences between the 8 computed descriptors; (ii) a structural alignment of two proteins is simply done by systematically rotating/translating one mobile sphere around one immobile reference; (iii) a certain degree of fuzziness in the comparison is reached by projecting global amino acid properties (e.g., charge, size, functional groups count, distance to the site centre) independently of local rotameric/tautomeric states of cavity-lining residues.

The method was implemented in a new program (SiteAlign) and tested in a number of various scenarios: measuring the distance between 376 related active site pairs, computing the cross-similarity of members of a protein family, predicting the targets of ligands with various promiscuity levels. The proposed method is robust enough to detect local similarity among active sites of different sizes, to discriminate between protein subfamilies and to recover the known targets of promiscuous ligands by virtual screening.

1. Schalon, C.; Surgand, J.-S.; Kellenberger, E.; Rognan, D. A simple and fuzzy method to align and compare druggable ligand-binding sites *Proteins* **2008**, asap