Optimization of a docking-scoring protocol

Alban ARRAULT, Aurélien MONGE, Christophe MAROT, Luc MORIN-ALLORY

Institut de Chimie Organique et Analytique UMR CNRS 6005, Modélisation Moléculaire et Chimiométrie, Université d'Orléans BP6759, 45067 ORLEANS Cedex2, France, alban.arrault@univ-orleans.fr

Docking programs are able to simulate the geometric and electronic complementarities within a ligand-protein complex. The present study has been carried out with two docking programs: FlexX and Fred.

The conformational search is the first step in the protocol. Then, each conformation is docked within the binding site and finally scored.

We have studied the conformer generators of the two docking programs. FlexX has its own "conformer maker", performing an incremental construction of each ligand, under the supervision of FlexX score and in respect to the binding site general features. Fred needs, as input, an independently generated conformers' set. We tested Omega (Openeye) to randomly generate conformations. The two conformer generators were submitted to a comparison to evaluate their importance in the protocol.

Different functions are used for evaluating the energies of each conformer binding to the receptor. Our comparison was based on the cylooxygenase-type 2 target with the same set of compounds in every case. The scoring functions tested are the respectively implemented ones in FlexX (Flexx Score, Dock Score, PMF-Score, Gold Score, Chemscore) and in Fred (PLP score, Screenscore, Shapegauss score, Chemgauss2 and Chemscore).

The performances of each function were inspected under several aspects. A basic analysis is to treat each score as an independent descriptor of the quality of the dock. Another method is to compile a set of selected functions using a consensus method. Our approach was to elaborate a statistical model which promotes the best scoring functions, prunes the worth ones and considers the intermediate functions. Actually, the geometric and electrostatic distributions in every receptor are responsible for the score performance variations and are often at the origin of much confusion during scoring interpretations.