Developing a protein - ligand docking algorithm: Flexligdock

Peter R. Oledzki* and Richard M. Jackson

University of Leeds

A program *Flexligdock* is being developed to provide a flexible ligand docking tool for small molecule docking to proteins. The program is based on Q-fit (Jackson, 2002) a rigid protein-ligand docking algorithm and uses a probabilistic sampling method in conjunction with the GRID (Goodford, 1985) molecular mechanics force field to generate and score solutions.

The method fragments a ligand at each rotatable bond to produce a series of rigid fragments termed *seed fragments* (or anchors). *Flexligdock* then utilises an interaction point methodology to map the ligand fragments onto interaction energy grid maps of the protein target. It then applies an incremental construction algorithm to build the ligand, fragment by fragment, in the protein binding site. This stage employs torsion angle sampling to permit simulation of ligand flexibility. The algorithm has been parameterised on a data set of 46 protein-ligand complexes. The parameterisation data set contains a structurally diverse set of proteins and a variety of ligands that contained between 0-23 torsion angles.

Three main parameters used for filtering out poor solutions during incremental construction were investigated. The filtering parameters investigated were; (1) the intra-molecular ligand collision distance, (2) the percentage threshold for inter-molecular ligand-receptor collisions and (3) the number of solutions passed to each stage of the incremental construction algorithm. With an optimal parameter setting *Flexligdock* docked 46 out of the 55 ligands <2Å RMSD as the top ranked solution with the manual *seed* placement. However, with automatic seed placement only 25 out of the 55 ligands were docked <2Å RMSD as the top ranked solution. This result proves that the incremental construction algorithm produces good ligand conformations with correct seed positioning.

The FlexX validation data set of 200 protein-ligand complexes (Kramer *et al.*, 1999) has been docked with *Flexligdock* to permit comparison against other existing protein-ligand docking algorithms. The FlexX docking algorithm docks 46% of the data set as the top ranked solution. *Flexligdock* (with automatic *seed* placement) docked 47% of the data set as the top ranked solution. Currently, further investigation of parameters and improvements to the docking algorithm are being undertaken to increase the accuracy of *Flexligdock*.

References

Jackson, R.M. (2002). Q-fit: a probabilistic method for docking molecular fragments by sampling low energy conformational space. *J. Comput. Aided. Mol. Des.*, **16(1)**, 43-57.

Goodford, P.J. (1985). A computational procedure for determining energetically favourable binding sites on biologically important macromolecules. *J. Med. Chem*, **28**, 849-857.

Kramer B. *et al.* (1999). Evaluation of the FlexX incremental construction algorithm for protein-ligand docking. *Proteins*, **37**, 228-241.