50. Amino Acid Dependent Weighting Factors for FFT-based Unbound Protein Docking

Philipp Heuser, Dietmar Schomburg

CUBIC - Cologne University Bioinformatics Center, Zuelpicher Str. 47, D-50674 Cologne, Germany, Tel.: +49-221-4707427, Email: philipp.heuser@unikoeln.de

Docking algorithms predict in which orientation two proteins bind under natural conditions. We introduce optimised amino acid specific weighting factors in the FFT-based procedure for the calculation of the geometric correlation. These factors lead to significant enrichment of near-native conformations on the top ranks of the complex prediction.

Protein-protein interactions and complex formation play a central role in a broad range of biological processes, including hormone-receptor binding, protease inhibition, antibody-antigen interaction and signal transduction. To identify how two proteins are interacting will be particularly important for elucidating functions and designing inhibitors. Docking algorithms are developed to predict in which orientation two proteins are likely to bind under natural conditions. They can be split in a sampling step followed by a scoring step.

One of the most widely spread docking methods is based on Fast Fourier Transformations (FFT). The usage of FFT was introduced into docking by Katchalsky-Katzir in 1992. One important aspect of the docking procedure is the representation of the proteins. Most FFT based methods use a grid representation for the proteins. Therefore each protein is mapped on a 3D grid, and different values are assigned to the cells of the grid, representing the surface or the interior of the proteins.

However, each protein-protein interaction depends on the amino acids involved in the interaction. Several attempts to evaluate the importance of the 20 amino acids for protein-protein docking were published in the past. Different properties of the amino acids like hydrophobicity, propensity to be in the interface, electrostatic properties, flexibility and others were tested for their importance in docking.

Since it is nearly impossible to decide which property and which scale is the best one in each single case, we optimised amino acid specific weighting factors for rigid body unbound-unbound protein-protein docking. For the optimisation the nonlinear minimisation function of the R-package based on a Newton-type algorithm was used.

The recalculation of the geometric correlation using the optimised weighting factors resulted in an enrichment of near-native conformations on the top ranks. The optimised parameters were tested on 30 unbound docking test cases from literature. Without the optimised parameter only for 3% of the evaluated complexes a near native conformation can be found within the top 2% of all generated docking solutions, but after the optimisation there is at least one conformation with RMS < 2A within the top 2% for more than 50% of the complexes.

The parameter obtained from the optimisation comply very well with amino acid properties. Very low weighting factors (<1) are assigned to amino acids which have more than 2 freely rotatable bonds in their side-chains (except aromatic residues) and have a very low propensity to be in the interface. These amino acids with long flexible side chains are most likely to be misleading in unbound-unbound protein docking. Values around 1 are assigned to amino acids with no or one freely rotatable bond. High values up to ten are assigned to the aromatic residues which are known to play an important role

in protein interactions. The highest value is assigned to Methionine, which has a rather high propensity to be in a interface but at the same time only about 2% of the interface residues are MET. That means if we do find a Methionine it is quite likely that it lies in the interface region.