Effective Discrimination of Native Protein Structures using Atom-Atom Contacts

McConkey, B.J, Sobolev, V. and Edelman, M.

Department of Plant Sciences, Weizmann Institute of Science

We have developed a method for quantifying inter-atomic contacts within proteins, and have used this measure to conduct a statistical assessment of contacts within known protein structures. The generated scale of atom-atom contact frequencies is useful in the identification of near-native protein structures from within large sets of simulated structures.

The introduced atom-atom contact measure is based on a Voronoi tessellation procedure, which subdivides a protein into cells with a one-to-one correspondence to protein atoms. The contact definition used integrates the solvent accessible surface and atom-atom contacts into a single measure, allowing them to be compared within a statistical framework. The atomˆatom contact measure was used to extract contact preferences from a training set of 648 non-redundant structures, and a contact-based scoring function derived from this data. The accuracy of the scoring function was tested using established protein decoy sets, including decoys from the recent CASP4 competition (http://predictioncenter.llnl.gov/casp4/casp4.html) and those generated by ROSETTA (http://depts.washington.edu/bakerpg). Each decoy set consists of a target native structure plus up to 2000 simulated structures per target. A total of 112 targets and over 48,000 decoy structures were scored. It was found that the scoring function ranked the native protein first within the decoy sets in over 90% of cases, when isolated protein subunits were used as the target structures. If inter-atomic contacts between protein subunits are considered, the accuracy of the method increases to over 97%. This represents a significant improvement in accuracy over currently available methods. Furthermore, the results show that interactions beyond approximately 2 atom diameters (~6.8 Å) are not necessary to determine the fold of a protein for the cases tested, while it is necessary to include interactions between subunits to predict some structures.