

Protein Helical Topology Prediction Using Mixed-Integer Linear Programming

Scott McAllister
Department of Chemical Engineering
Princeton University

The protein folding problem represents one of the most challenging and potentially rewarding problems in computational biology. Distance constraints and topology predictions can be highly useful for the folding problem in reducing the conformational space that must be searched by deterministic algorithms in order to find a protein structure of minimum conformational energy. A novel mixed-integer linear optimization model will be presented for predicting topological contacts and generating interhelical distance restraints between hydrophobic residues in alpha-helical globular proteins. Since the model does not make assumptions about the form of the helices, it is applicable to all alpha-helical proteins, including helices with kinks and irregular helices. This model enhances the ASTRO-FOLD protein folding approach of Klepeis and Floudas (2003), which finds the structure of global minimum conformational energy via a constrained nonlinear optimization problem.

The structure prediction of membrane proteins is comparatively less studied by researchers in the field. However, due to the difficulty in obtaining reliable experimental structures, accurate theoretical prediction of membrane proteins is of paramount importance. A pair of optimization models has been developed to predict the interhelical interactions in alpha-helical membrane proteins. By maximizing the occurrence of highly probable pairwise and three-residue interactions, realistic contacts can be predicted by imposing a number of geometrical constraints. The development of these low distance contacts can provide additional distance restraints for first principles approaches to the tertiary structure prediction problem for membrane proteins.