Local protein structure prediction based on physicochemical properties of amino acids

S.A. Lukshin, Yu. V. Milchevskiy, A.M. Nikitin, V.G. Tumanyan

Abstract

Background: Protein structure prediction methods using local structure information have shown promising improvements. The quality of new fold predictions has risen significantly and in fold recognition incorporation of local structure predictions led to improvements in the accuracy of results.

We developed a local structure prediction method to be integrated into either fold recognition or new fold prediction methods. For local sequence window of a protein sequence the method predicts Cartesian coordinates of local protein backbone.

The first step is to define a set of local structure representatives based on clustering local structures. We have developed an original clustering procedure that uses RMSD as a measure of the distance between the structures. In the second step a stepwise regression model is trained to predict the local structure representative given local sequence information. We have build regression model based on predictors based on physicochemical properties of amino acids

Results: The step of clustering local structures yields an average RMSD quantization error of 0.65Å for 30 structural representatives (for a fragment length of 5 residues).

Conclusion: The described method yields Cartesian coordinates for local protein structure. These local structure predictions can be incorporated either into fold recognition algorithms to improve the overall prediction accuracy or into new fold prediction methods.