



**CALIFORNIA STATE SCIENCE FAIR
2009 PROJECT SUMMARY**

Name(s) Nitish Lakhanpal	Project Number S0414
Project Title Computational Prediction of Beta Structure from Amino Acid Sequence in a Class of Pathologically Relevant Proteins	
<p style="text-align: center;">Abstract</p> <p>Objectives/Goals Because structure dictates the function of proteins - physiological or pathological - protein structure discovery is of great interest to biological science. Though experimental approaches have yielded good results, these efforts have proven ineffective for beta-rich proteins such as amyloids and autotransport proteins - both implicated in pathologies such as Alzheimer's Disease, meningitis, and pertussis - due to difficulties with crystallization and insolubility. Interest has therefore grown in computationally predicting beta structure from primary amino acid sequence. My objective is to develop an algorithm that improves the prediction accuracy achieved by current algorithms. Based on insights into the nature of protein structure, I suggest an algorithm that is governed by both near and distant interactions among the amino acids comprising the protein.</p> <p>Methods/Materials 2.4 GHz Personal computer, 3 GB RAM. Coded in C++: Take a non-redundant sampling of protein structures in the Protein Data Bank. Extract frequencies of residues and pairs of interacting residues involved in beta structures, yielding tables $w(i,j,\theta)$ containing the frequency of pairing between amino acids i and j with orientation θ and $v(i,\theta)$ containing the frequency of observing amino acid i with orientation θ - then compute a "solo" and a "duo" propensity for every subsequence of permissible length. Compute a single score by combining the two. Do this for several threshold values. Compute final globally optimum structure assignment.</p> <p>Results We ran the algorithm on two sets of solved non-redundant proteins: 16 amyloids and 21 autotransporters. Our algorithm outperformed its predecessors in sensitivity to beta strands and in the false positive rate of beta strand discovery, showing approximately a two-and-a-half times improvement.</p> <p>Conclusions/Discussion Our algorithm improved beta structure prediction substantially by considering close as well as distant interactions in a polypeptide chain. We also explored the relationship between prediction sensitivity and false positives and the threshold level used - enabling the algorithm's use for a spectrum of prediction objectives. Though applied to 2 classes of proteins, our algorithm has broad applicability for predicting beta structure.</p>	
Summary Statement We developed an algorithm to predict beta structure in proteins from amino acid sequence based on properties of known beta structures, effectively merging single-amino-acid level analysis with the possibility of long-distance interactions.	
Help Received Discussed idea for project with Dr. Bonnie Berger at MIT.	