



**CALIFORNIA STATE SCIENCE FAIR
2002 PROJECT SUMMARY**

Name(s) Sukhjeet S. Batth	Project Number S0402
Project Title Protein Profile Analysis: A New Application May Streamline Structural Predictions	
<p style="text-align: center;">Abstract</p> <p>Objectives/Goals The ability to predict protein structure from amino acid sequence alone would be invaluable to scientists everywhere. A pharmacologist working on creating a new drug, for example, could analyze a protein sequence with his computer and find its structure in seconds. Computational biology is far more efficient than the long, tedious analyses of structural biologists which include x-ray crystallography and nuclear magnetic resonance. The present study approaches this challenge from a very specific position. A method of protein sequence analysis, Profile Analysis, was applied to create a structural prediction for alpha-helix packing.</p> <p>Methods/Materials Profile Analysis was used to predict initial contact residues (i's) in packing diamonds of 4-3 alpha-helix packing. Position-specific scoring matrices (profiles) were made from a structure correlated scoring matrix and amino acid sequences from either +3 or +4 packing alpha-helices. Different helix lengths were used to create profiles which helped establish the most accurate method of applying Profile Analysis. Reliability was also found with the use of leave-one-out and z-score analyses of predictions in cholecystokinin.</p> <p>Results Profiles made from the helix range consisting of i to i+11 were determined to be the most accurate for the prediction of both +3 and +4 packing alpha-helices among sequences of known structure. In addition, leave-one-out and z-score analyses confirmed that predictions in cholecystokinin by profiles created from the range i to i+11 were within the range of accuracy. These profiles predicted Leucine-13 and Arginine-31 as the initial contact residues of a +3 and +4 helix, respectively, in human cholecystokinin.</p> <p>Conclusions/Discussion Based on accuracy of predictions for known helices and further supported by reliable predictions within the range of accuracy for cholecystokinin, it seems that 4-3 alpha-helix packing relies heavily enough on sequence to be predictable by this method. With further study, this application of Profile Analysis could eventually replace the long, tedious, and costly analyses of structural biology for prediction of structures which rely as heavily on sequence as 4-3 alpha-helix packing.</p>	
Summary Statement The present study attempted to predict a specific protein structure, 4-3 alpha-helix packing, using a computer protein sequence analysis, Profile Analysis.	
Help Received	