

**Problem A3**  
**Predicting and viewing protein structure**  
**Due: February 28, 2006**

For this assignment you are given the amino acid sequence of a chimeric protein (a chimera in this context is defined as a macromolecule composed of pieces from other macromolecules). The sequence is in a FASTA format file **chimera06.fa** on the Homework Problems web page. Hand in requested printouts and written answers (with justifications) to the questions. Label your answers clearly.

**Questions (Total of 40 points)**

1. Use the server at [http://fasta.bioch.virginia.edu/fasta\\_www/chofas.htm](http://fasta.bioch.virginia.edu/fasta_www/chofas.htm) to predict secondary structure for the chimera using the Chou-Fasman method. (Save the results for use in part 4 below.) What amino acids are part of the longest region that is predicted to be alpha helix?
2. Use the BLAST server to determine what proteins were used in creating the chimera, and find 3D structures for each. Make a table containing five columns: (1) the names of the individual proteins whose fragments make up the chimera, (2) the Genpept accession numbers for these proteins (ignore “unnamed protein product” records and list multiple entries if they have the same sequence), (3) the PDB filenames for structures of the individual fragments (see part 3), and (4) the residue numbers of the fragments in the original proteins, and (5) the corresponding residue numbers in the chimera.
3. Obtain the PDB files for the individual pieces. Be careful when picking the PDB files! Make sure that you have a PDB file corresponding to the fragment in the chimera and that you have an actual structure rather than a theoretical model. (a) For each piece, use **JMol** or **RasMol** or **Cn3D** or **Protein Explorer** to print a model in which the entire protein is shown as a black wireframe model and any non-protein atoms are shown as a gray (yellow) spacefill model. Use a white background when printing! (b) Select those residues corresponding to the largest  $\alpha$ -helix actually observed in any of the pieces. For the piece that contains the largest helix, print a model in which the largest helix is shown as a black ribbon model and the remaining residues are shown as a gray (yellow) wireframe model.
4. For each region of alpha helix or beta sheet listed in the PDB files, determine what fraction of that region was correctly predicted by the Chou-Fasman method. Make a table of these results, showing the residue numbers for each region and the fraction correct.

**Extra Credit (5 points each)**

5. Imagine that you submitted the sequence of the chimera and the sequences of the proteins it is composed of to a program that does homology based secondary structure prediction. Do you think that predictions for the various regions of the chimera would be the same or different as the predictions for those regions in the individual proteins? Why or why not?
6. Do Problem 7 in Chapter 10 of the Mount textbook.