CENG 465
Spring 2010-2011

Assignment #4

Programming Assignment about Protein Structures

Finding interacting amino acids between chains of a multi-chain protein

In this assignment, you are going to implement a program to find the amino acids at the interface of different chains of a multi-chain protein. Your program will have two inputs: a protein structure provided as a PDB (Protein Data Bank) file and a distance threshold. The input protein structure will have multiple amino acids chains and your goal will be to find amino acid pairs from different chains whose alpha-carbon atoms are closer than the input distance threshold. In other words each amino acid will be represented as a point which is at the coordinates of the alpha-carbon of that amino acid.

Detailed information about the PDB file format can be found at http://www.wwpdb.org/documentation/format32/v3.2.html . For this assignment, you will only use the ATOM records found at the Coordinate Section. Furthermore, from the ATOM records you will only use the records which are for alpha-carbon atoms. Alpha-carbon atoms are named with the atom name "CA" in the ATOM record. Other records contain additional information about the protein sequences, secondary structure, experimental technique etc. You do not need to parse any other record than the ATOM record in this assignment.

For output, you will report each interacting amino acid pair from two different chains in the following format:

<CHAIN_ID>:<AA CODE>(<AA NUMBER>) interacts with <CHAIN_ID>:<AA CODE>(<AA NUMBER>)

Example Input/Output:

Input: 2XZ8.pdb, distance threshold = 6.0 Angstroms
Output:
A:LYS(255) interacts with B:LEU(353)
A:LEU(256) interacts with B:GLY(354)
A:LEU(353) interacts with B:LYS(255)
A:GLY(354) interacts with B:LYS(255)
A:GLY(354) interacts with B:LEU(256)

Run your program with the test structures at http://www.ceng.metu.edu.tr/~tcan/ceng465_s1011/Assignments/hw4_test_structures.zip using the distance threshold of **7.0 Angstroms**.

Deliverables:
1. The source code of your program. You may use any programming language of your choice.
2. Your results for the test structures.

Submission:
Submit the deliverables as a zip bundle or as a tarball using the COW system.

Late Submission Policy:
Your final assignment grade will be penalized 20 points per late day.

CHECK THE NEWSGROUP REGULARLY FOR POSSIBLE UPDATES ON THE ASSIGNMENT.