

Homework

Due date: April 8, 2003

A printed or hand-written homework (no electronic versions please) is due no later than April 10, 2003. It can be handed in during the class (after the lecture) or you can bring it to my office at any other time.

The **goal of this homework** is to make you familiar with the three-dimensional architecture of proteins by direct observation and elementary analysis of the structures of several representative proteins.

Means of analysis: you will need a computer program that allows visualization of protein structures given atom coordinates of a protein. There is a significant number of software packages designed for these purposes. On the basis of easiness of download and use, I selected protein visualization package RasMol (a Windows version which is easier to use is called RasTop). These programs can be downloaded from the following web sources (see BCHM461 website):

RasMol: <http://www.bernstein-plus-sons.com/software/rasmol/>

RasTop: http://www.bernstein-plus-sons.com/software/RasTop_1.3/

A zip-file ([structures.zip](#)) containing atom coordinates for several proteins has been posted on the Web site and can be downloaded (see Step 2 in the step-by-step instructions below).

Your task is to answer the following questions for **five** of these proteins (you can skip 1ubq.pdb and 1ubq_dry.pdb):

1AAR_dry.pdb

1BAK_6.pdb

1CKA.pdb

gag1.pdb

GB.pdb

1. How many polypeptide chains are in the protein structure? If there are any other molecules besides polypeptides, describe them.
2. Describe how many and what secondary structure elements characterize the fold of this protein.
3. Describe structural motifs (supersecondary elements), if any, that you found in the tertiary structure of the protein.
4. How big is the protein? Specify the length of the polypeptide chain, i.e. the total number of amino acid residues in the protein sequence. What is the size of the protein molecule (in Angstroms)? You may use the molecule analysis tools to measure its size. If the molecule is elongated, indicate its size in the direction of its largest dimension.

Ste-by-step instructions on how to use RasTop.

Step 1. You will have to download and install one of these programs on your computer.

Step 2. Download a set of protein structures from the BCHM461 website. To do this, go to the website, find the assignments table, the line dated March 11, and click on the hypertext "[Protein structures for download](#)". A download panel will appear; its content depends on your web browser. Follow the download instructions, e.g. if you use Internet explorer, select the option to save the file to disk and click the "OK" button. You might be prompted to choose the location for file download on your hard drive.

Step 3. The file you have downloaded is called **structures.zip** --it is a zip-file, so you will have to unzip it. To do this, after you downloaded the file, navigate to the folder where the file is located and double click on it (Windows). This will activate your unzipping program (e.g. WinZip) – follow the instructions for unzipping a file. When unzipped, this file will create a folder "structures", with the following seven files:

1AAR_dry.pdb
1BAK_6.pdb
1CKA.pdb
1ubq.pdb – ignore this one
1ubq_dry.pdb – ignore this one
gag1.pdb
GB.pdb.

All these files are essentially text (ascii) files, and have extension "pdb" (for Protein Data Bank). Their names are somewhat arbitrary – some of them coincide with the structure ID in the Protein Data Bank, but some are different (the structure ID is indicated in the first row in each file). Since these files are text files, they can be read using practically any text editor (e.g. WordPad or Word and, in most cases, Notepad). You might want to make yourself familiar with the data format. This could also be useful, because most of information you will be asked can be retrieved directly from the pdb-files.

Step 4. Now the actual fun begins...

Load/Start RasTop (or RasMol). A black rectangle window will appear. You might want to resize it to fit your visual needs. Click on *File* and then in a pull-down menu panel click on *Load Molecular File...* (or press CTRL-L). A "Select Molecular Coordinate File" panel appears. Navigate to the folder **structures** that contains the downloaded and unzipped protein structures, and select a protein coordinate file that you would like to visualize. If you did everything right, you will see a set of lines representing interatomic bonds in your protein (a so-called wireframe representation). You can use various buttons on the top bar of RasTop screen to choose different representations of the molecule, and to obtain information about its structure (secondary structure elements, Ramachandran map, distances, etc). I briefly demonstrated some of these features in class. The sliders on the bottom could be used to move (rotate or translate) the molecule.

You can load several molecules, one by one, and use the Window button to switch from one loaded molecule to the other.