RasMac 2.6-ucb Tutorial for Chemistry 5.50

Instructions for Looking at Chorismate Mutase (CM) using RasMol The protein structures will be located in a folder that is labeled 5.50.

Within Netscape go to Yahoo and enter RasMol. Click on the first entry describing RasMol. This site is linked to two additional sites that you will find helpful. Initially you should click on Tutorial and go the the site from the U. of Southern Maine. This site leads you step by step through the use of RasMol and describes many of the commands that you will need to know. The second site that is also very helpful, especially as you become more sophisticated in the use of RasMol, is called RasMol Reference Manual. This site has been prepared by Roger Sayle, the developer of this program. You might want to bookmark these two sites.

Open Rasmac 2.6-ubc

File:open Go to the folder labeled 5.50 and open 1ECM.pdb. This protein is one of the chorismate mutases discussed in class. The protein is actually bifunctional, but has been engineered to possess only chorismate mutase activity. The protein has been crystallized with a transition state analog, oxobridged prephenic acid, in the active site. In 1ECM.pdb the inhibitor is labeled hetatm TSA. The names of inhibitors or substrate analogs cocrystallized with the protein are found in the PDB header. The protein is a dimer with chains named A and B.

You will familiarize yourself with the structure of 1ECM and the location of the inhibitor in the active site of this enyzme. Initially you will use the Pull-down menu (File, Edit, Display, Colours, Options etc) on the top of your screen. Look under each heading and familiarize yourself with the transformations that are accessible from this menu. Under Windows: command-line, a second window will open that is white with a RasMol> prompt. Use of this command-line window and commands described below will be essential for allowing you to look the active site of CM.

Display: cartoons

Cartoon displays β sheet strands as arrows pointing toward the C terminal end of the chain. Note you cannot get any information by clicking on the structure, about individual amino acid residues in this display mode. The helical lines, represent helices.

Colours: structures

These commands display your protein with α helices (pink) and β-sheets (yellow). Turns will be in blue if they have been defined by the PDB header., otherwise they will be white. Initially you want to become familiar with the commands that allow you to move your molecule.

I. To rotate the molecule, place the pointer anywhere in the main screen, hold down the mouse button and move the mouse. Dragging up and down rotates the model around the horizontal axis (X) through the molecules center. Dragging left and right rotates around the verticle Y axis. Hold down shift and command keys

while dragging left and right to rotate PurD around the Z axis (perpendicular to the screen). To move CM translationally, hold option and drag. To zoom, hold down the shift key and drag the pointer down the screen to bring the model closer or up to move it farther away.
To return to the original setting go to command-line window
Rasmol>reset <return>

II. Most proteins that we will examine during the semester will have multiple domains. CM, however, is composed of a single helical domain. However, the protein is a dimer. Look at each subunit of CM and using the commands described below, make a plumbing diagram of one subunit. The plumbing diagram should indicate not only the secondary structural features, but also the tertiary features, the spacial relationship of the secondary features to one another. In addition, if the protein has multiple subunits, then the diagram may also indicate the nature of the interaction between the monomers. Often times, the plumbing diagram with multiple subunits becomes too complex to be useful and each subunit is diagramed alone initially.

The following commands within the command-line window will be used to help you look at the CM dimer or each domain separately of a multidomain protein.

Picking and Selecting:
To select residues in CM go through the following exercise. Select is one of the most useful commands.
Rasmol>select 1-109A <return>
This command should allow you to place subunit A in any display mode that you wish. For example display ball and stick will put all of the atoms in subunit A into ball and stick format.
A second useful command is restrict.
Rasmol>restrict 1-109A <return>

If you start with the dimeric form of CM, then restrict A selects all atoms specified and hides all others. The A limits you to the A monomer in this case. This command does not alter the display of the selected atoms. Once you are familiar with the protein and want to simplify the structure to see, for example, the active site, this command is useful.
Display: sticks <return> This command will place the structure in stick form.
Colour: cpk RasMol draws residues 1-109 as a stick model. The cpk command draws the atoms in the appropriate colors: red for oxygen, blue for nitrogen and gray for carbon. This color display format is very useful for identifying the side chains by inspection. If you try to rotate your molecule, it does so around the center of your entire molecule, the AB, dimer. You may want to center your molecule around a specific residue.

Rasmol>center Glu19 A <return>

You can center this domain around any amino acid or specific atom that you like. If for example you want to center around the α carbon (CA) of Glu19

Rasmol>center Glu19A.ca <return>

RasMol is picky about the syntax of expressions. If you have trouble go to the manual.

Draw a plumbing diagrams for CM. Label the aminoa acids at each end of a secondary structure motif. Helix 1 is composed of residues x through y. Often this information is available in the pdb header. In presenting your diagram think about the tertiary as well as the secondary structure. In addition, think about the quarternary structure, the relationship of the two monomers to one another.

An example of a Plumbing Diagram for CM is shown below:

III. Exploring more details about the select and restrict commands.
Rasmol> select * <return>

You have now returned to your entire molecule, the dimer of CM.
Rasmol > reset <return>
Display: sticks

Returns full backbone to the screen and centers the rotation to the entire molecule and original orientation. If the waters associated with the molecule are now apparent, you can remove them by
Rasmol> restrict* and not hoh
Dispay: in any format you desire. Wireframe shows all the side chains and is very useful once you learn how to isolate the region of interest. CPK as noted above is useful as it allows you to identify amino acid side chains by color.

IV. We will look at the active site of CM and the ts analog.
Rasmol> select tsa
Display: ball and stick You will see two ts analogs displayed, one on each subunit. You can see how complex the structure looks. It takes lots of practice to figure out how to simplify the structure so that you can identify the important interactions. We will look at only the A subunit.
Rasmol> restrict 25-88A<return>
Display: stick
Rasmol> select tsa <return>
Display: ball and stick
This process should immediately show you a simplified active site region.

Zoom in on the ball and stick ts analog. Draw what you see. The conformation of the inhibitor is also important to see. Once you have drawn the analog, now try to understand which amino acid side chains in the active site of the enzyme facilitate its binding. First look for charge complementarity. Your inhibitor has two carboxylates. Are there any positively charged amino acid side chains with which these residues interact? Click on arginines and lysines within the active site. The number of the residue should appear in the command window. Write down the residue numbers for future reference. You should have picked out R28, R51 and K39. Note when you look more carefully at the entire structure, you will see that by our original restriction of the monomer to residues 25-88, we actually missed R11 from the second subunit that plays and important role in binding.

Now you will look in more detail in the region around the active site in order to identify the amino acid residues that are responsible for binding the inhibitor and by inference the substrate.
Rasmol> select 1-109A <return>
Rasmol> restrict within (6.0, ___) <return>

The 6.0 refers to the number of angstroms around the active site residue or center of interest. Note you must have a decimal point or the command won't work. The ____ refers to molecules you have just identified as being the center of interest. TSA is your inhibitor and this name should be inserted in place of ___. You can restrict around an inhibitor or specific atoms of amino acid residues in the active site. If you go back and have both subunits present and restrict to within 10.0 Å of the active site you can now find R11 from the second subunit. What is the residue that interacts with the hydroxyl of your inhibitor? (You should have found glu52). Attached is a cartoon of the active site with the inhibitor and all of the amino acid side chains with which it interacts. Make sure you can identify all of these interactions. During the course of the semester you will be looking at the active site of many enzymes. This is the starting point for thinking about function and residues that are involved in catalysis. You must practice restrictions and have confidence that you have identified all the residues sufficiently close to the inhibitor or substrate analog, that you can think about function.

V. Distances: In order for you to know whether an interaction is interesting you must know the distance between the atom of interest in your small molecule and the protein. You must go back and familiarize yourself with H bonding distances, distances associated with single and double bonds, etc.. Let us practice with several distances.
Rasmol>restrict within (8.0, tsa) <return>
Display: stick
Rasmol>center tsaA <return>
Rasmol>set picking monitors <return>

Click on the two atoms of interest, one on tsa and one on an atom within an amino acid side chain of interest within the protein. A dotted line should appear with the distance indicated.
To remove the distances so that new distances can be examined:
Rasmol> set picking ident <return>
Rasmol> monitors off <return>

To familiarize yourselves with looking at protein structures and their active sites, carry out the following exercise. Draw the structure of the inhibitor and side chains as in the above diagram and fill in all of the distances.

V.Saving your work: Scripts

At any time during Rasmol, you can write a script that Rasmol can use to recover your current view. You will want a copy of your active site in your final lab report as a prompt for discussing your kinetic results. The following are instructions as to how to write the Script. Fill the screen with the figure that you wish to preserve.
Rasmol > write script CDRs.spt <return>
After a brief pause, Rasmol writes the Script. This new file will appear in the same location as the Rasmol program icon. Its icon is a little scroll. The file ending .spt (for script) helps distinguish a script from a coordinate file which ends in pdb or ent.

Several additional programs are important for thinking about function. One of these is sequence alignments as residues involved in catalysis are ofter conserved. The web site www.ncbi.nlm.nih.gov/ is very useful. You should try and retrieve the sequences for the E. coli, B. subtilis and S. cerevisiae CMs. Select protein and type in chorismate mutase e. coli etc. You should obtain the following information.

```
CM
>e.coli
mtsenpllalrekisaldekllallaerrelavevgkakllshrpvrdidrerdllerli
tlgkahhldahyitrlfqliiedsvltqqallqqhlnkinphsariaflgpkgsyshlaa
rqyaarhfeqfiesgcakfadifnqvetgqadyavvpientssgaindvydllqhtslsi
vgemtltidhcllvsgttdlstintvyshpqpfqqcskflnryphwkieytestsaamek
vaqaksphvaalgseaggtlyglqvlerieanqrqnftrfvvlarkainvsdqvpakttl
lmatgqqagalveallvlrnhnlimtrlesrpihgnpweemfyldiqanlesaemqkalk
elgeitrsmkvlgcypsenvvpvdpt

>b. subtilis
```

```
mkeetfylvredvlpdamrktlevkklldrkkadsvadavqkvdlsrsafykyrdavfpf
ytmvkeqiitlffhledrsgalsqllqavadsgsnvlsihqtiplqgranvtlsistsam
eedihtlmnklrkfdfvekveilgsga

>s. cerevisiae
mstfgklfrvttygeshcksvgcivdgvppgmslteadiqpqltrrrpgqsklstprdek
drveiqsgtefgktlgtpiammiknedqrphdysdmdkfprpshadftysekygikassg
ggrasaretigrvasgaiaekflaqnsnveivafvtqigeikmnrdsfdpefqhllntit
rekvdsmgpircpdasvaglmvkeiekyrgnkdsiggvvtcvvrnlptglgepcfdklea
mlahamlsipaskgfeigsgfqgvsvpgskhndpfyfeketnrlrtktnnsggvqggisn
geniyfsvpfksvatisqeqktatydgeegilaakgrhdpavtpraipiveamtalvlad
alliqkardfsrsvvh

Note you must use the above format without spaces.  Clustal W is
very picky about format submission.
The ecoli and yeast enzyme are much bigger than the b subtilus
enzyme.  Thus to get good alignments one needs to delete the
section of the gene that does not have CM activity.
```

One you have the sequences , use GOOGLE to go to ClustalW and then paste in the sequences in the appropriate box and submit your search. What is suprising about the results from your search?