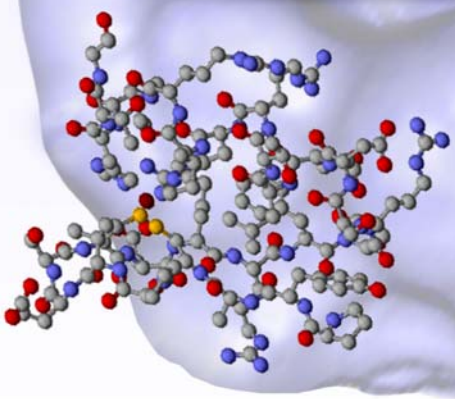
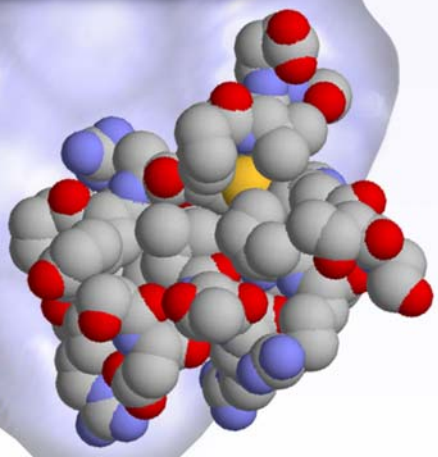
Center for BioMolecular Modeling

RasMol Training Guide

...where teachers come first





Shannon Colton, PhD
Program Director
Center for BioMolecular Modeling
Milwaukee School of Engineering

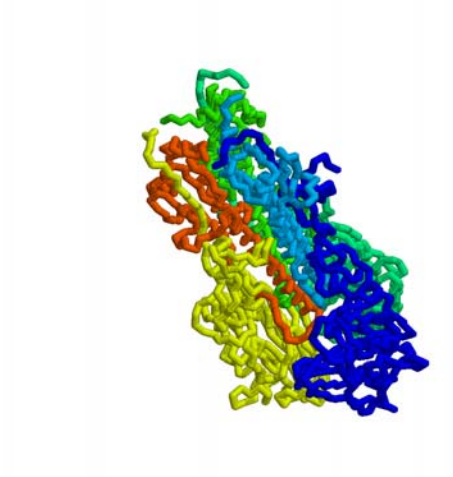
RasMol Training Section I:
Using RasMol as a Computer Visualization Tool

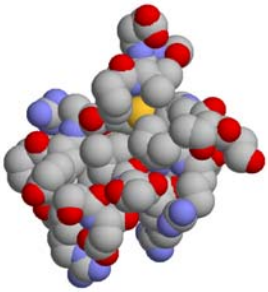
Section I of the MSOE Center for BioMolecular Modeling RasMol Training Guide is designed to assist you as you begin your exploration of RasMol. This program is a simple, yet powerful tool, which enables you to visualize a molecule in "3-Dimensional space". Since you actively manipulate the computer mouse to rotate the molecule in computer space, you develop a sense of the 3-dimensionality of the molecule. Through RasMol, you can also change the display format of the molecule to display different features of the molecule.

One of the goals of the CBM is to combine the use of the physical models with the computer visualization tools in order to allow you the opportunity to further explore the relationship between the structure of a biomolecule and its function. Often times, you will manipulate a CBM-produced model and through your interaction with the model, will generate questions that can be answered by working with RasMol. Through this tutorial, you can become familiar with this program to enable them to further explore the structure/function relationship that is a focal point in biochemistry.

At the end of this section, you should be familiar with how to use RasMol as a computer visualization tool. We will introduce you to the basic features of RasMol, including how to:

- Organize your files
- Start RasMol
- Open a PDB File
- Change the Background Color
- Change Display Format of the Molecule
- Use the Command Line to Change Formats
- Change Colors
- Identify Features within the Molecule
- Select Features within the Molecule





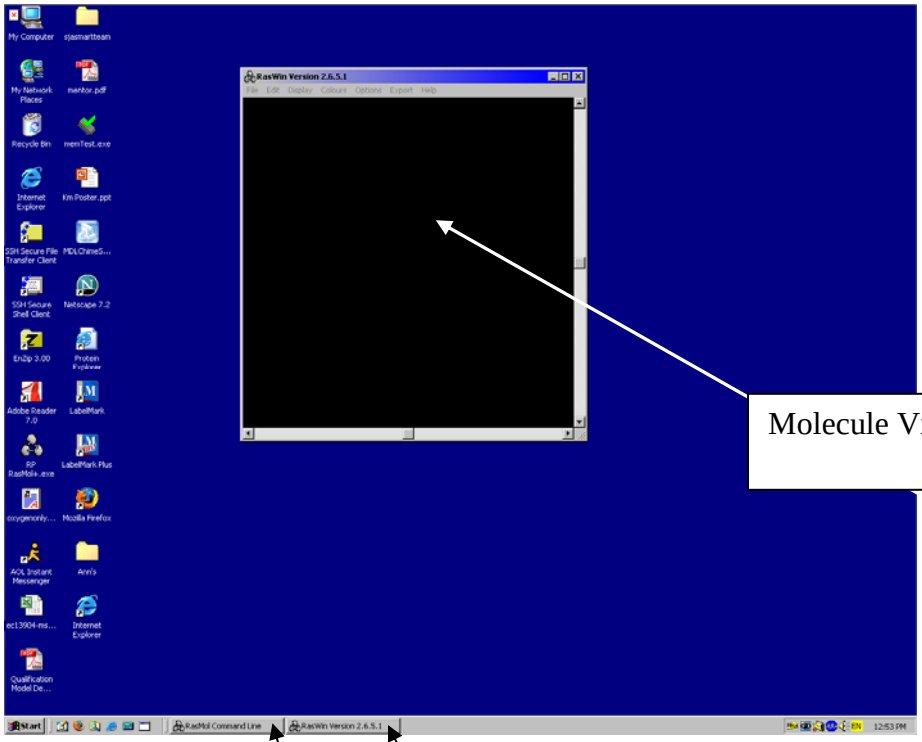
1. **Organize Your Files**
   - We recommend that you maintain separate folders on your computer (desktop or preferred disk drive location) for separate projects.
     - For instance, if you are using RasMol for SMART Teams, you should have a separate folder for each of these projects. Within these "larger folders", keep a separate folder for themes, subjects or projects. This will enable you to keep subjects separated and will become important if you do any design work or generate script files, a way to save your work within RasMol (see RasMol Training Part II).
   - In each folder, place a copy of RasMol and the PDB file with which you are working.
     - As you become more familiar with the program and begin to develop script files, you will probably have several folders with PDB files and script files. RasMol is a small program and it is okay to make multiple copies of it within your computer. We recommend that for each new project that has a new PDB file, you place a copy of RasMol in that folder. This will become especially important as you use RasMol to explore more molecular structures.
     - You can download a copy of RasMol from the CBM website: http://www.rpc.msoe.edu/cbm/smartteams/resources.php

2. **Opening RasMol**
   - To open the RasMol program, double click on the RasMol icon in your folder:


RasMol 2.6.5.1.exe

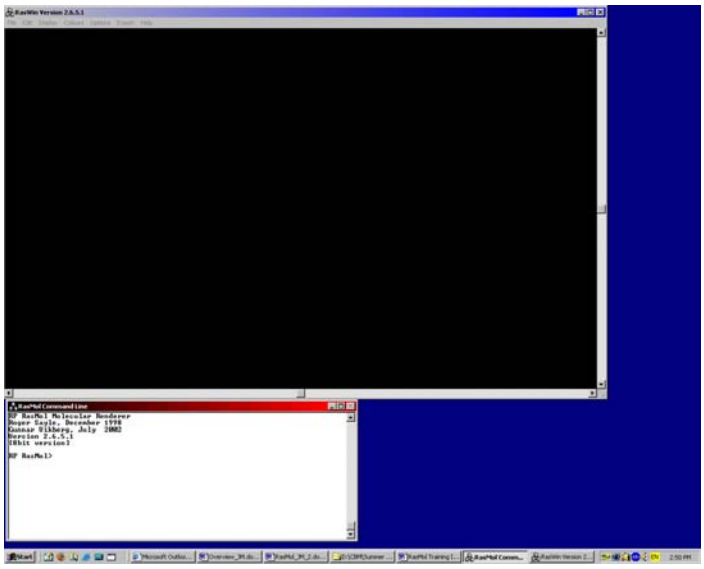
   - **RasMol has <u>Two</u> Screens**
     - You will notice that when you open the RasMol program, a window automatically opens with a black background on your computer monitor. In addition, there will be two buttons that appear on the Task Bar (see figure below). This is because there are two windows associated with RasMol: a Command Window and a Visualization Window.

Note: These directions are specific for PC-users. Macintosh users will utilize RasMac, will have the essential components described herein, but may also have differences. Please see the appendix for Macintosh-specific directions.



When you launch RasMol, this is an image of what will appear on your monitor. Notice the two buttons on the task bar. One button will open the command window (the one on the left) and the other button will open the visualization window (the button on the right; this window will open automatically upon launching of RasMol). Open both of these windows and arrange them on the monitor so that you can see both of them on the screen.

In order to work most effectively with RasMol, we recommend that you organize your screen to display BOTH of these windows. To open the Command Line Window, click on the RasMol Command Line button on the task bar. **Arrange the Molecule Visualization window and the Command Line window so that you can see both windows and they are not overlapping**. You can resize the windows to accommodate your preferences or computer monitor size. You will want the Computer Visualization window to be larger.

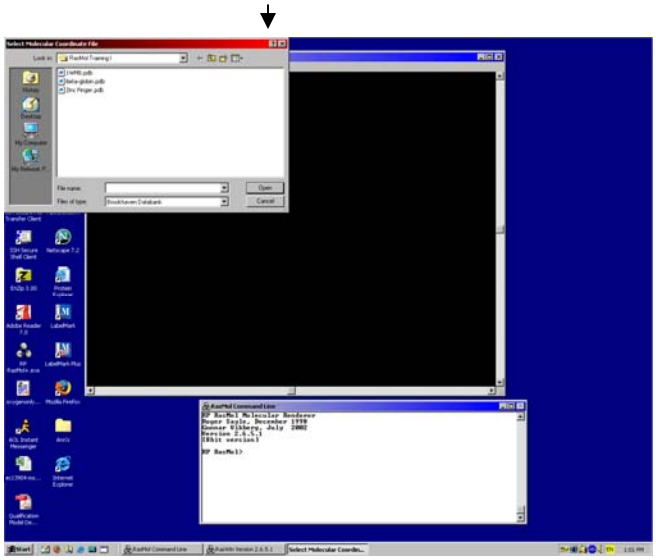


3. PDB Files

- RasMol is a molecular visualization program that enables us to visualize the atomic coordinates of a molecule that has been crystallized.
- These coordinates are stored in a file called a PDB (protein data bank) file.
- PDB files can be located and downloaded for free from the Protein Data Bank website (www.pdb.org).
- More information about saving a PDB file and understanding the information contained within the file will be discussed within RasMol Training Part II.
- For initial exploration of RasMol within Section I of the RasMol Training Guide, example PDB files that are used within this tutorial can also be found on our website at:
    - http://www.rpc.msoe.edu/cbm/smartteams/resources.php

4. Opening a pdb file

- After both windows of RasMol are opened and arranged on the screen, use the Pull-Down menu bar located on the Molecule Visualization window and chose "File."
- Click on "Open" in order to open a PDB file
    - This will bring up a window entitled "Select Molecular Coordinate File" that will allow you to direct the program to open a specific pdb file.



    - Click on the appropriate pdb file that you wish to open.
    - **NOTE:** Organizing your design work will be simplified if you always do the following: Create a RasMol folder with the RasMol program and the PDB file, open the RasMol program by clicking on the RasMol icon inside the folder, use the pull down menu to open the PDB file with the commands >File>Open.

5. Default Display of Molecule

After you open the PDB file, the molecule will appear in wireframe format in the Molecule Visualization window. The wireframe format is the default setting and every time that you open a file, this is the initial format you will see. In the wireframe format,

the thin wire represents the bonds between each of the atoms. We will change the format in a moment.

6. Commands within RasMol

- There are two ways to enter commands and change the display of the molecule in the RasMol program:
    - One way is to use the Pull-down menu and select specific commands (like we just did with opening the PDB file).
    - The other way is to type specific commands into the Command Line window.
        - In the Command Line Window, you will see the prompt "RasMol>". By typing commands at this prompt, RasMol is directed to do certain things. The commands that are entered at this prompt need to be specific. We will discuss this specificity in more detail as we progress through the training guide.
    - For the remainder of this section in the training guide, a combination of commands (both in the command window and the pull-down menu) will be used.

7. Closing a PDB file

- If you have completed work on one PDB file and would like to work on another file, you need to close the PDB file that you are presently working on before you can open a new PDB file. RasMol will not open another PDB file over the original. You must close the current file to open the next one.
- There are two ways to close a PDB file: using the Pull-down menu or using the Command Line window
    - Using the Pull-Down Menu
        - To close a PDB file, use the Pull-Down menu bar located on the Molecule Visualization window and chose "File."
        - Click on "Close" to close the PDB file.
    - Using the Command Line window
        - To close a PDB file, type "zap" at the prompt.
            - RasMol>zap

8. Background Color of Molecule Visualization Window

The default setting for the Molecule Visualization window will be a black background. This can be modified and you can select any color for the background. Dark colors are not recommended if you wish to print an image of your molecule.

- To change the color of the background:
    - RasMol> background white
        - This will change the background of the window to white.
    - RasMol> background green
        - This will change the background to green.
            - This is obviously not a very nice color at which to look for most people, but for some reason, some people like this color.

- RasMol> background blue
  - This will change the background to blue.
  - You get the idea. You can change the color of the screen to any color that you wish. We recommend white, as this has been the color that has worked the best for the most people. It also serves as the best backdrop if you export a picture (*.gif) so that you do not have a black (or green) background for your model. This saves on ink when printing, especially for posters. (See Number 14 in this section of the training guide for how to export an image of your molecule.)
  - Note: We recommend that you avoid the red/green color combination in case you, or others, have problems with color blindness.

9. Rotating the Molecule

- To rotate the molecule within the X-Y axes on the computer screen, place the mouse cursor in the Molecule Visualization window, hold down the **left** mouse button. Moving the mouse will rotate the molecule about the X-Y axes
- To relocate the molecule within the Molecule Visualization window, (for instance, to center the molecule within the window), place the mouse cursor in the Molecule Visualization window, hold down the **right** mouse button and move the molecule to the desired location.
- To zoom in or out, place the mouse cursor in the Molecule Visualization window, hold down the **left** mouse button AND hold the **shift** key. Moving the mouse will then zoom in or out on the molecule.
  - A word of caution: if you zoom in or out too fast, RasMol may close (crash) unexpectedly, as the program has to re-render the image at each step. Zooming too quickly, especially with large molecules, will overwhelm RasMol and the program will crash.
- To rotate the molecule within the Z axis, place the mouse cursor in the Molecule Visualization window, hold down the **right** mouse button and the **shift** key. Moving the mouse will then rotate the molecule about the Z axis.

Table Summarizing Mouse Button/Key Needed for Actions within RasMol

Action	**Mouse Button (and Key) Needed**
Rotate X-Y	Left Mouse Button
Move Molecule	Right Mouse Button
Zoom In/Out	Left Mouse Button & Shift Key
Rotate Z	Right Mouse Button & Shift Key

10. To Change the Format of the Molecule

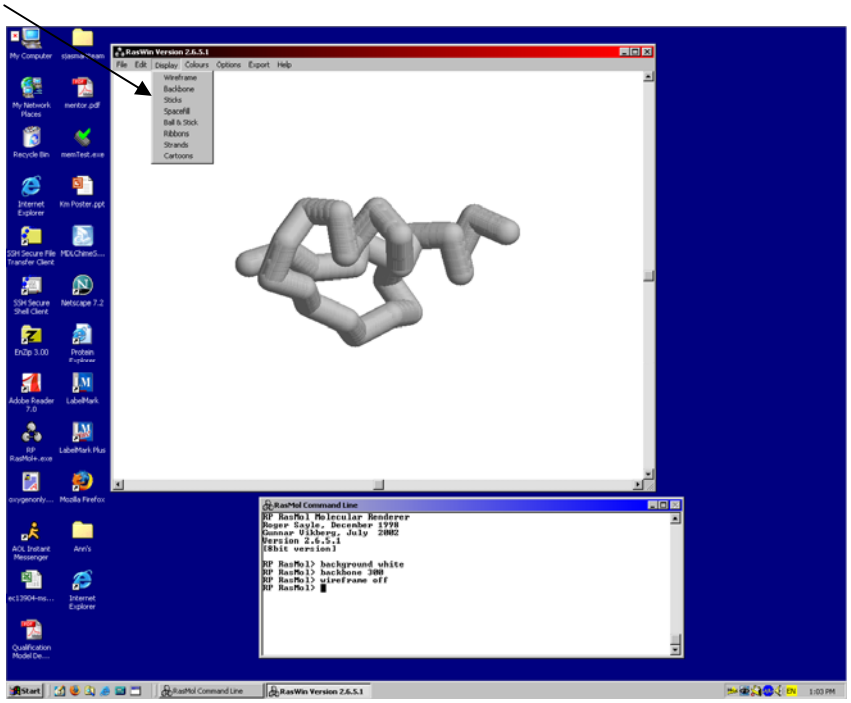
- There are two ways to change the format of the molecule. One is through the pull-down menu and one is through the command line. We will do both so that

you are familiar with both aspects. As you become more comfortable with the program, we do recommend that you use the command line rather than the pull-down menu.

- Using the command line allows you to add or subtract specific features one at a time. If you have created a design using the command line and then use the pull-down menu, the pull-down menu will overwrite all of the molecule features that you have created. This is more important as you develop more sophisticated models with specific features highlighted. In Section III of this RasMol Guide, we will do some exercises that will illustrate this point.

11. Changing format with the Pull-Down Menus

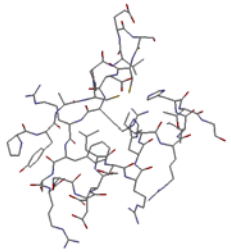
- Under the Pull-down menus in the Molecule Visualization Window, there is a "Display" option (see figure below). Within this pull-down menu, there are several different formats that you can choose to display your model:
  - Wireframe
  - Backbone
  - Sticks
  - Spacefill
  - Ball and Stick
  - Ribbons
  - Strands
  - Cartoon



- Each of these display options will render your molecule in a different format. Some of these options will look new to you and some of them will look very familiar. When we start talking about designing a model for building purposes, we will talk more about which options are available for building, but for now, we are simply talking about using computer visualization programs and therefore we will cover each of the formats.
  - Each option has its advantages and disadvantages; which option you choose to display your molecule depends on which story you wish to tell. The advantage of using RasMol within the classroom is that it enables you to represent the model in a variety of ways, thus illustrating different aspects of the molecule.

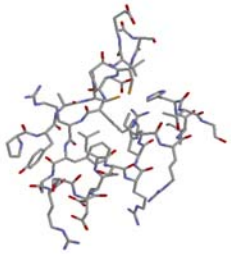
- Let's look at each of these in turn and see what the different formats represent.
    - Please note that the images to the right are of a single zinc finger and are based on amino acids #4-31 of the PDB 1ZAA.pdb.

    - **Wireframe**
        - This is the RasMol default format for displaying molecules initially opened in RasMol. The thin wire represents the bonds between each of the atoms and the ends of the wires represent the atoms. The advantage of the wireframe is that all of the atoms are displayed. It is difficult to distinguish secondary structures in this format.

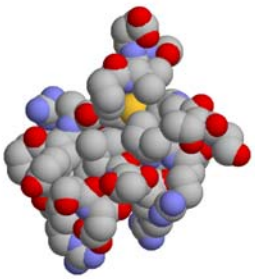


    - **Backbone**
        - The alpha carbon backbone format only displays the position of the alpha carbon in each amino acid by a bend in the backbone. All of the other atoms within the amino acid are not displayed. The advantage of this display format is that it clearly illustrates the secondary structures within a molecule.

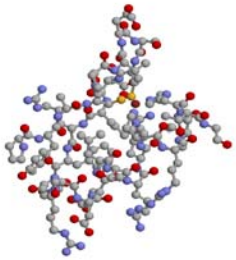
    - **Sticks**
        - This display format is similar to the wireframe format, but rather than a thin wire, the atoms are connected by a stick. As in the wireframe format, all of the atoms in the molecule are displayed.



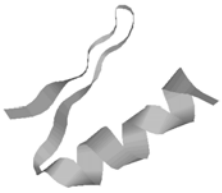
- **Spacefill**
  - In this format, each atom is displayed as a sphere representing the volume that atom occupies in space. For example, hydrogen atoms are smaller than carbon atoms, which are smaller than sulfur atoms. This format is advantageous in that the spacefill model represents the volume and three-dimensional shape of a molecule. Identifying the secondary structures or any internal atoms in this display format is challenging.



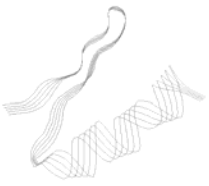
- **Ball and Stick**
  - In this format, each atom is represented by a small sphere (ball) and the bonds connecting each atom are indicated by a stick connecting the spheres. If the molecule is very large, this representation can become too overwhelming and specific information about the molecule can be difficult to determine.



- **Ribbons**
  - This format shows the molecule as if it were traced with a ribbon. This version may be familiar as it is often seen in textbooks. Secondary structures are easily identified within this format with the alpha helix represented as a coiled structure and the beta strand represented as a pleated structure.



- **Strands**
  - This format is similar to Ribbons, except that rather than being a solid ribbon, there are several “strands” or thin lines that make up the model.



- **Cartoon**
  - This format is similar to Ribbons, except that the beta strands are indicated by arrows to point out the directionality of the strands. This version is also widely seen in textbooks.

Format	Advantage	Disadvantage
Wireframe; Sticks; Ball and Stick	Displays all atoms	Difficult to distinguish secondary structures; too busy for large structures
Backbone	Secondary structures clearly seen	Detail of amino acids is lost, unless specifically added
Spacefill	Represents 3D volume and shape	Cannot identify secondary structure or internal atoms
Ribbons; Strands; Cartoon	Secondary structures are easily visualized; cartoon shows directionality of beta strands	No structural details

12. Changing display format with the Command Line

- In the command window, you can change the format of the model by typing the following commands:
  - RasMol>Backbone
  - RasMol>Wireframe
  - RasMol>Spacefill

- To generate the Ball and Stick image, you would use a combination of the wireframe and spacefill commands.

- In contrast with the Pull-down menu in the Molecule Visualization window, in the Command Line window, you can also alter the size. When we get to the model design stage of this training (in section III), this will become especially important. For now, we will just vary the sizes so that you can see what we mean by changing the size of the different formats. Following the format command, you enter a value for the size. Each format as a range of values, usually between 1 and 750. For scaling purposes, 1 RasMol unit equals 1/250 Angstroms.

- Enter the following commands in the Command Line window to see the effects of the commands:
    - RasMol>backbone 300
    - RasMol>backbone 30
    - RasMol>spacefill
    - RasMol>spacefill 100
    - RasMol>wireframe 200
    - RasMol>spacefill 300

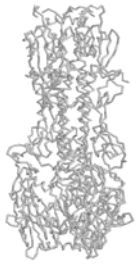
- If you want to turn off a particular format, you would enter the following command
    - RasMol>spacefill off
    - RasMol>wireframe off

13. Changing color with Pull Down Menu

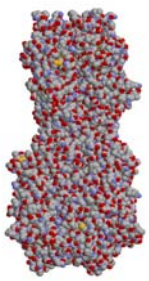
- RasMol has several color display options to make your model look visually pleasing, or to indicate specific features of the molecule.
    - Figures to the right are of hemagglutinin, a protein found on the flu virus membrane and are based on 2HMG.pdb.

The most frequently used color options are:

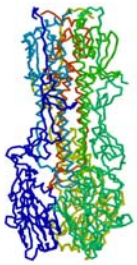
- **Monochrome**
    - This option will color the entire molecule white. This is often a good starting point if you would like the model to have a white backbone with other colors highlighting different features (to be discussed in number 16).



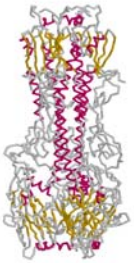
- **CPK**
    - This color scheme was developed by Corey and Pauling and later improved by Kultin in which carbon is gray, hydrogen is white, oxygen is red, nitrogen is blue. This is the standard color code and is utilized in many textbooks.



- **Chain**
  - If a PDB file has multiple polypeptides, as seen in the hemagglutinin protein on the right, in which there are 6 chains, or in the protein hemoglobin, in which there are 4 chains, this color option will color each chain a different color.

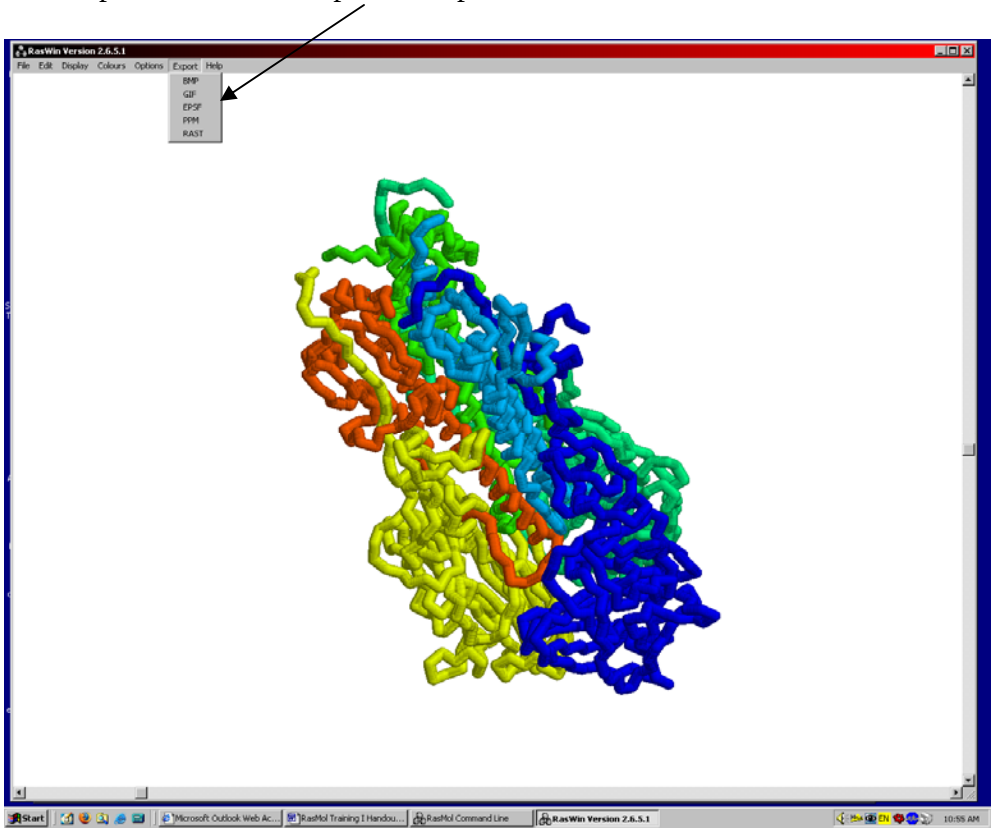


- **Structure**
  - This option will color the secondary structures different colors. Magenta is used to highlight helices, yellow to indicate beta sheets and blue to indicate turns (this feature is not shown in this figure).



14. Exporting an Image from RasMol

- If you would like to create a *.gif image from RasMol, you can do so using the pull-down menu option “Export”



- This menu option will then provide you with several different graphic formats (.gif, .bmp, etc) from which you could choose to save your image. After you select which format you would like to export your image, a save window will appear and you will be asked to name your file and direct the computer to which folder you would like to save the image.
- The exported image that is saved will look *exactly* like the one that is on your screen, which means that if you have a blue background, your image will also have a blue background. This is the reason why we recommend that you change the background to white if you plan to export any images.

15. Identification Feature of RasMol

- One of the features in RasMol is an identification feature that enables you to determine the identity of an atom by placing the mouse cursor over the molecule and pressing the left mouse button. After you click the mouse button, a line will appear in the command window giving identity information about the point on which you clicked.
- The format of the information displayed in the Command Line window will be similar to the following:
    - Atom: CA 140 Group: Asn 19 Chain: B
        - This line gives you several pieces of information:
            - The atom is the alpha carbon (CA) and it is assigned atom number 140.
            - This alpha carbon belongs in the Asparagine amino acid, which is the 19th amino acid in this PDB file [Note: the amino acid number in a PDB file does not always correspond with the amino acid in the protein.]
            - The atom is in chain B of the molecule.
        - This information will be useful in several ways, which we will discuss later in our training.

16. "Select" command

- You can specifically select certain features within your molecule when using RasMol.
- There is a set of predefined terms that are recognized by RasMol (see Quick Reference Guide for full listing) http://www.rpc.msoe.edu/cbm/resources/pdf/refcard.pdf
- Once you select a portion of the molecule, all subsequent commands will apply to that selected portion. RasMol will not know that you meant to select another subset, or that you were done with that set of features, and so you must be very specific with your commands.
    - Examples:
    - **Secondary Structures**
        - RasMol>select helices

- This command will select all of the helices within the molecule.

- After you have selected the helices, you can enter a command to change the format or color, and this command will ONLY be applied to the helices.
    - For Example: RasMol>select helices
      RasMol>color red
      This series of commands will color the helices red.
    - ***NOTE:*** At this point, you have selected the helices. If you enter any other commands, such as a new color, or display format, all commands will be applicable to JUST the helices. You will need to select a new subset of features, or "select all" to get out of the helix-selected mode.

- RasMol>select sheets
    - This command will select all of the beta sheets within the molecule.
    - As we did with the helices, you can color the sheets any color, or change the format in which you have the beta sheets displayed.

- **Heterologous Groups**
    - RasMol> select hetero
        - This command will select all of the hetero atoms that are contained within this file. Remember that these are the atoms that are NOT amino acids. Examples include heme groups, zinc atoms, ADP and water.
        - If you know the name of the group, you can select it directly by its name (such as "Select Zn")

- **Amino Acid Sidechains**
    - Within any molecule, you may wish to display all of the amino acids, or a specific set of amino acids (such as all of the hydrophobic amino acids), or only a select few amino acids based on function (such as the amino acids that make up an active site of an enzyme). Through entering commands in the Command Line window, you can make these selections.

RasMol Exercise

- Let's practice selecting amino acids:
    - Open a PDB file in RasMol
        - Example PDB files can be found http://www.rpc.msoe.edu/cbm/smartteams/resources.php

- Change the background color to white
- Change the display format to backbone

- To select all of the amino acids:
  - RasMol>select sidechain
  - RasMol>spacefill 275
  - RasMol>wireframe 225

- To select a subset of amino acids, you can use predefined terms, such as “polar” or “hydrophobic” to display all of the amino acids that are in these categories.
    - **Note:** There is a list of predefined terms that RasMol recognizes on the RasMol Quick Reference Guide.

  - To select all of the polar amino acids:
    - RasMol>select polar
    - RasMol>color red

  - To select all of the hydrophobic amino acids:
    - RasMol>select hydrophobic
    - RasMol>color yellow

  - You can also select all of particular amino acid, such as all of the histidines in a molecule.
      - **Note**: RasMol recognizes the three letter abbreviations for the amino acids.
    - RasMol>select his
    - RasMol>color magenta
      - This allows you to identify all of the histidines within the molecule. By moving the mouse cursor over these magenta regions, you can click on them to identify which histidine the amino acid is in order to determine the amino acid number so that you can select this amino acid specifically.
  - To select functionally significant amino acids, you could refer to the primary citation associated with the PDB file (more on that in Part Two of the training). If you know that you want to display the sidechains for a particular amino acid, for example

Histidine 63 in the beta-globin protein, you would select that amino acid specifically.

- RasMol> select His63
  - **Note:** Although we typically write His 63 with a space between the letters and the number, there is not a space between the letters and number when writing a command in RasMol.
- RasMol>spacefill 275
- RasMol>wireframe 225
  - These commands will generate a ball and stick representation of all of the atoms in the amino acid. Please **note** that ALL of the atoms within the amino acid are displayed as balls and not just as the sidechain atoms. This is a "bumpy backbone" versus a "clean backbone" and we will discuss this difference in Section III of the RasMol Training Guide.

- RasMol>color CPK
  - This command will color the sidechains CPK.

17. Changing color with Command Line

- To change the color of the model with the command line, there are two options: to use the predefined color list (refer to Quick Reference Guide) or to use the RGB color numbers (again, use the Quick Reference Guide, or any source that indicates RGB [computer color system of Red, Green, Blue] color numbers).
  - RasMol> select helices
  - RasMol> color red **OR** RasMol> color [240,0,0]

Conclusions for Section I

Hopefully at the end of this section, you feel comfortable:

- starting RasMol on your computer
- opening a PDB file
- changing the background color
- changing the display format of the molecule
- changing the colors of the molecule
- selecting certain features of the molecule
- using the pull-down menu
- using the command line

Through these features, we encourage you to use RasMol within the classroom as a visualization tool to enhance your lessons on biochemistry.

Additional Exercise

Exercise I

1. Start RasMol and open Zinc Finger.pdb in RasMol. (can download this file at http://www.rpc.msoe.edu/cbm/smartteams/resources.php)
2. Change the background color to white.
3. Display zinc finger in the alpha carbon backbone format.
4. Display Zinc ion in relation to the zinc finger
5. Display sidechains for the histidine and cysteines coordinating the zinc ion
6. Display the sidechain for Phe16
7. Display the sidechain for Leu22
8. Display the sidechain for Arg18

Exercise I Answers

1. Start RasMol and open Zinc Finger.pdb in RasMol.
2. Change the background color to white.
   RasMol>background white
3. Display zinc finger in the alpha carbon backbone format.
   RasMol>backbone 300
   RasMol>wireframe off
4. Display Zinc ion in relation to the zinc finger
   RasMol>select zn
   RasMol>spacefill
5. Display sidechains for the histidine and cysteines coordinating the zinc ion
   RasMol>select his
   RasMol>spacefill 275
   RasMol>wireframe 225
   RasMol>select cys
   RasMol>spacefill 275
   RasMol>wireframe 225

6. Display the sidechain for Phe16
   RasMol>select phe16
   RasMol>spacefill 275
   RasMol>wireframe 225
7. Display the sidechain for Leu22
   RasMol>select leu22
   RasMol>spacefill 275
   RasMol>wireframe 225
8. Display the sidechain for Arg18
   RasMol>select arg18
   RasMol>spacefill 275
   RasMol>wireframe 225

RasMol Training Section II:
Understanding the Protein Data Bank and More Specific Commands within RasMol

In Section II of the MSOE Center for BioMolecular Modeling RasMol Training Guide, the focus is to learn about the Protein Data Bank, the worldwide repository for the crystal structure files, and to learn additional specific commands within RasMol. Section II will emphasize how to use RasMol in a more sophisticated manner using these specific commands. Mastering the material in this section will prepare you for further understanding of how to use RasMol and using RasMol in the Protein Modeling Challenge in the MSOE Center for BioMolecular Modeling sponsored event in the Science Olympiad (www.rpc.msoe.edu/cbm/scienceolympiad). In this section, we will focus on how to

- Use the Protein Data Bank (www.pdb.org)
- Save your work on RasMol by generating script files
- Open script files in RasMol
- Edit script files
- Center molecules in RasMol
- Select whole fields within RasMol
- Restrict subsets of the molecule within RasMol
- Add and remove hydrogen bonds within the molecule in RasMol
- Add and remove disulfide bonds within the molecule in RasMol

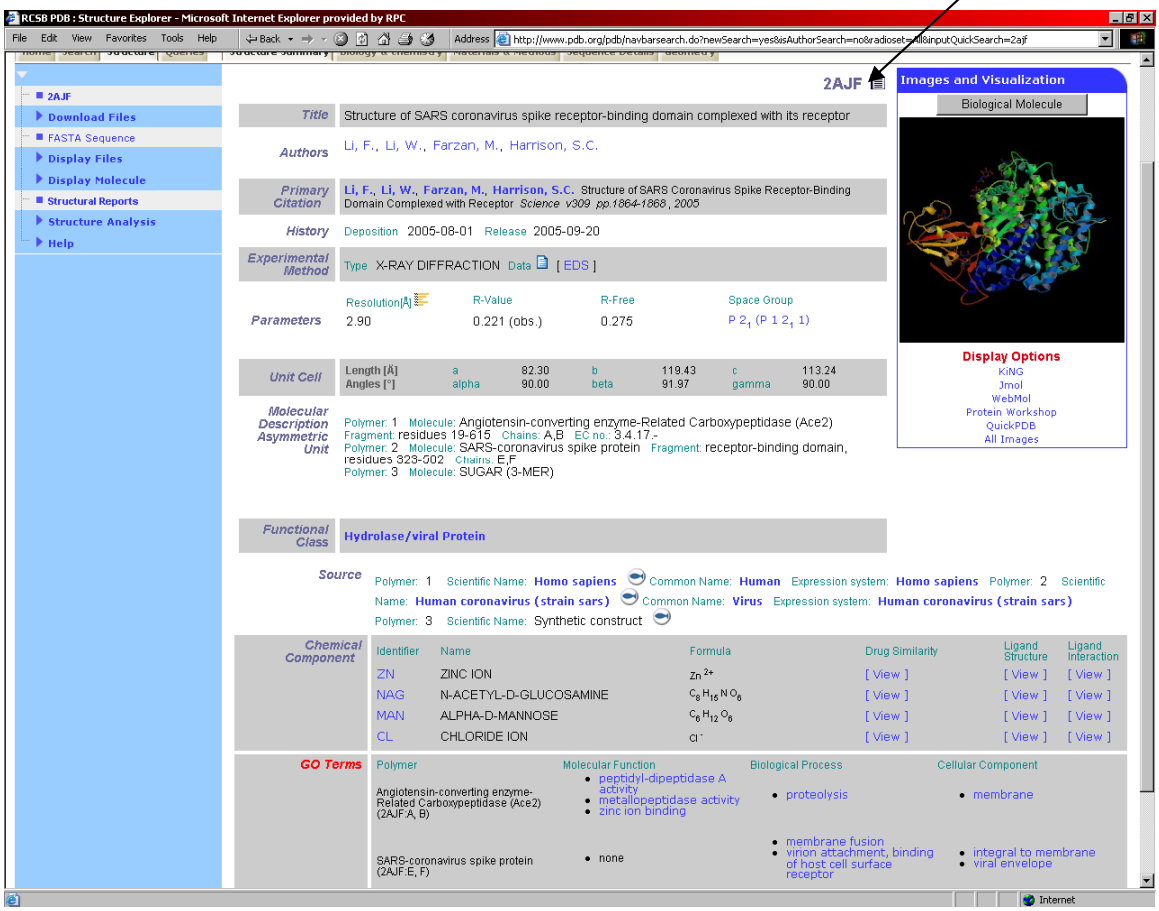
Part I

1. Protein Data Bank

- The PDB (www.pdb.org) is the single worldwide repository for the processing and distribution of 3-D structure data of large molecules of proteins and nucleic acids.

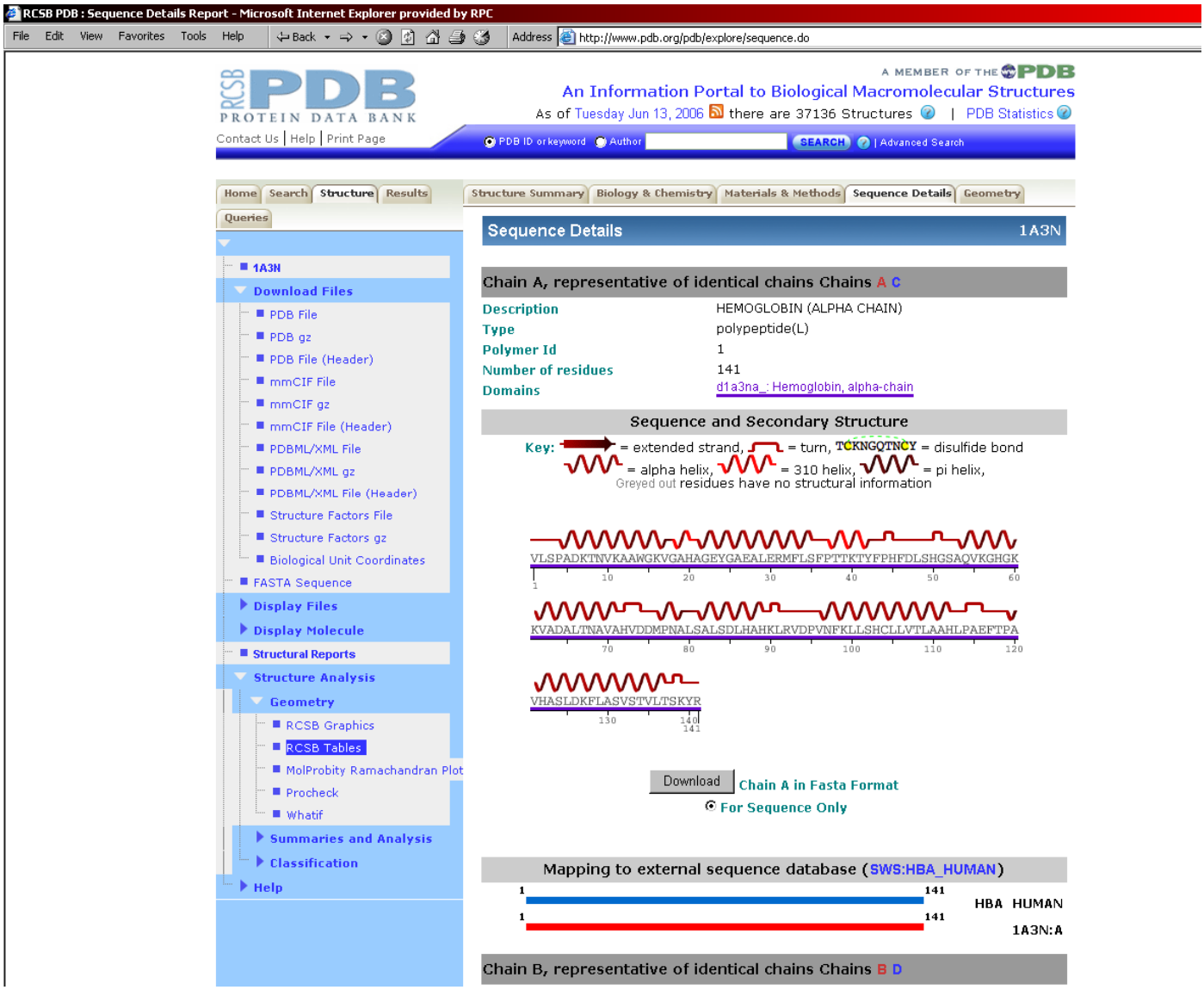
- As of 25 September 2008, there are 53,263 structures within the PDB.
- Resources within the PDB:
  - Molecule of the Month
    http://www.pdb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/index.html
  - PDB Newsletter by Gary Graper describing the Science Olympiad
    http://www.rcsb.org/pdb/general_information/news_publications/newsletters/2005q2/education_corner.html
  - PDB Newsletter (read about the New Jersey Science Olympiad Protein Modeling Challenge Event sponsored by the PDB)
    http://www.pdb.org/pdb/static.do?p=general_information/news_publications/news/news_2006.html#20060404
  - Educational Resources
    http://www.pdb.org/pdbstatic/education_discussion/educational_resources/education_flyer.pdf

- Each entry within the PDB contains several pieces of information:
  - **Structure summary page**
    - PDB Structure ID number
      - This 4 letter/number ID is a unique identifier that is assigned to the crystal data file upon deposition into the database.
    - Title
      - Title of the PDB file



- Authors
    - These are the researchers who were involved with the crystallization of the molecule. Note: The senior author or principal investigator is usually the last author in science publications.
- Primary Citation
    - The journal article that accompanies the PDB file; excellent research resource for understanding the function of the molecule.
- History of deposition and release
    - The date that the molecule was deposited into the PDB and the date the information was released to the public.
- Method of structure determination
    - The method that was used to obtain the structural data (ex: NMR, X-ray diffraction).
- Resolution at which the molecule the structural data was collected
    - How accurate the data is; the smaller the number, the better the data.
- Molecular Description
    - This will tell you the number of chains within the molecule and the chain identity; for example in the hemoglobin file (1A3N.pdb), the chains A and C are the alpha-globin molecules and chains B and D are the beta-globin molecules.
- Functional Class of the molecule
    - What type of molecule is it? (Ex: a toxin, an enzyme)
- Source of the molecule
    - From which species was the molecule isolated? (human, bacterium, virus, mouse)
- Chemical component
    - In this section, you will be able to determine if there are any heterologous groups that were crystallized with the molecule. Not all PDB files will have this section.
    - The 2-3 letter identifier used to designate the chemical components contained within the file listed are recognized by RasMol.
        - For example, if this section stated that there was NAG (N-acetyl-glucosamine) contained within the molecule, RasMol would recognize “NAG” and you could therefore “select NAG” and RasMol would be able to select the atoms within that chemical component of the PDB file.

- **Sequence Details**
  - This section of the PDB file provides specific sequence information as well as secondary structure information about the molecule.
  - In this section of the PDB file, you can identify where within the protein the alpha helix or beta sheets are located, as well as the amino/carboxy termini, which are the first and last amino acids of the protein, respectively (for more information about identifying these amino acids, please refer to number 13 within this section of the training guide).



2. Search the Protein Data Bank

- One of the features of the PDB is the ability to search the database for files. You can search using key words or authors by entering these terms in the search box, highlighted in red in the figure below. Or, if you already know the PDB structure ID number, you can enter that number in the search field. After you have entered the search terms in the field, hit enter or click on the "search" button to the right of the search field.

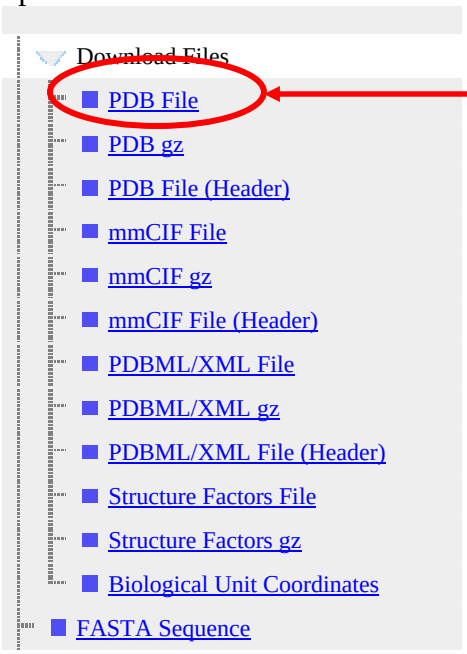
Practice searching:

- In the search field, enter "aquaporin" and click on "search."
  - This search will bring back a search result that gives 18 structure hits.
  - The PDB will list all 18 of these structures.
- If you enter "hemoglobin" into the search field, you will see that there are 362 structure hits.
  - If you do a broad inquiry like this one, you will need to look through the structures to see which one meets your requirements, for example:
    - Human versus mouse
    - Mutant versus "normal"

3. Download a PDB File

To save a PDB file, click on “Download Files”, which is located on the left navigation bar.

This will create a pull-down menu that will give you several options:



Select this option. By clicking on this link, it will bring up a “File Download” window. At this point, you click on the “Save” option and select which folder you would like to save the PDB file on your computer.

Remember our recommendation of organizing your work to include the PDB file and a copy of RasMol in a separate folder for each project on which you are working.

This organization will become important as you save your work (generate script files), which will be discussed in number 8 in this section of the training guide.

4. Searching the PDB and Reading a Structure Summary Page (practice)

- Search the database for "SARS spike receptor binding domain coronavirus protein"
- 8 structures should appear (as of 25 September 2008)
- Look for the one authored by Li *et al* and click on the PDB structure ID number
- You should arrive at the structure summary page for 2AJF.
- Use this structure summary page to answer the following questions about the PDB file (answers are at the end of this section):
    - Who are the authors of the PDB file?
    - In which journal was the primary citation published?
    - On what date was the file deposited into the PDB?
    - How many chains are in this file?
    - Are there any heterologous groups within this PDB file? If so, which ones?
    - From what source was this molecule isolated?

5. Download a File (practice)

- We are now going to work with the potassium channel as our model PDB file.
    - Create a folder labeled "RasMol Training". Within this folder, create a Kchannel folder.
    - Place a copy of RasMol in this folder.
    - Search the PDB for 1J95
    - Download 1J95 to the K Channel folder on your computer
        - This organization will be especially important within this section of the guide.

6. Opening a PDB file (review)

(for more information see Section I)

- Opening PDB files
    - Double-click on the RasMol icon within your folder.
    - Arrange both windows on your screen.
    - Go to the file menu.
    - Click on "Open."
    - Click on the appropriate PDB file (open 1J95).

7. A little review from the previous section

- How do we do the following items?
        - (answers at the end of this section)
    - Change Background to white
    - Change display formats
    - Change colors
    - Highlight the helices as red

8. **Save your work – Creating a Script File**
    - RasMol has a feature that allows for you to save your work. In order to save you work, you will generate a script file.
        - **RasMol> save script filename.spt**
            - Example: RasMol> save script Kchannel.spt
                - We recommend saving multiple versions of your model design as you work.
                    - For example, as you progress through your design, save Kchannel1.spt, Kchannel2.spt, Kchannel3.spt, etc. This allows for you to return to previous files, in case you accidentally save a file that you did not like, or if you make a mistake. **There is not an "undo" command in RasMol.**
                    - Through saving multiple versions of your work, you can return to a previous version if you accidentally do something to your molecule that you do not like, or if you accidentally lose your work.
                    - As with any important document, save your work often!
            - Please **Note** that it is very important to include the "script" and the ".spt" components to this command.
                - If you enter the command with "save filename.spt", you will create a PDB file and NOT a script file.
                - If the save command does not include these items, a script file will NOT be created. If you exit RasMol before a proper script file has been created, all work that you have done will be lost.
        - Once you have entered this command, you have saved your work.
        - Check your folder for your newly created script file.
            - If you have organize your work in such a way that there is a copy of RasMol and the PDB file in the same folder, and have launched RasMol from this folder, the RasMol program will automatically save the script file to this folder.

9. **Opening a script file**
    - To open a script file:
        - First, launch the copy of RasMol in your Kchannel folder.
        - In the command line window, enter the command:
            - RasMol>script filename.spt
                - Example:
                    - RasMol>script kchannel12.spt
        - This should open your version of the molecule in the molecular visualization window.

10. Practice saving/opening a script file

- Open the Kchannel PDB file 1J95 using the pull-down menu.
    - RasMol>select all
    - RasMol>backbone 300
    - RasMol>wireframe off
    - RasMol>color white
    - RasMol>select helices
    - RasMol>color cyan
    - RasMol>select hydrophobic
    - RasMol>color yellow
    - RasMol>save script kchannel.spt

- At this point, we have created a script file, or saved your work. We have generated an image in which the helices are colored cyan, except where there are hydrophobic amino acids.
    - Question: Why are there so many hydrophobic amino acids located on the *outside* of this protein when we have learned that hydrophobic amino acids are typically located buried within the molecule, away from the aqueous external environment? (answer at the end of this section)
- Okay, now that we have generated a script file, let's erase the image from the screen so that we can open the script file
    - RasMol>zap
    - RasMol>script kchannel.spt

- Did your original file open?
    - If so, congratulations, you have created your script file and opened your script file successfully! If not, go back and see where you made your mistake before continuing.
        - Did you have the "save"? Did you have the "script"? Did you have the ".spt"?
        - If the script file did not open, either the command was not entered correctly in the command line window, or the file was saved in a different folder location.

11. Editing your script file

- Any time that you generate a script file, RasMol will place the script file in the folder that RasMol was originally launched. This is the reason that we recommend organizing your work so that each molecule you are exploring has its own folder with a copy of RasMol and the PDB. Each time you generate a script file, it will then be saved in that folder.
- When you save your script file, a specific pathname will be generated and saved within the script file directing RasMol to look within a specific folder for that particular PDB file.
- There may be times in which you work on one computer and wish to transfer your file to another computer. If you simply copy or move the script file to a new

computer or a new folder and try to open the script file using the command described above, RasMol will be unable to locate the PDB file because it will be looking for the PDB through the specific pathname that was generated when the script file was generated.

- This can be circumvented by editing the script file.
- To edit your script file, open your script file using a word processing program such as Notepad or Wordpad. We recommend that you DO NOT use MS Word for this process.

This is what you will see when you open your script file:

```
#!rasmol -script
# File: kchannel.spt
# Creator: RasMol Version 2.6

zap
load pdb "D:\CBM\Summer 2006\RasMol Training 2006\RasMol Training II\1J95.pdb"
background [255,255,255]
set ambient 40
set specular off
```

- The text in **blue** is the pathname and tells RasMol the folder location of the PDB file. If you wish to use this script file on other computers, you will want to edit this line in the script file so that you can open the script file on any computer. In order to open this file on any computer, simply change the pathname to what you see in the text box below in **red**:

```
#!rasmol -script
# File: kchannel.spt
# Creator: RasMol Version 2.6

zap
load pdb ".\1J95.pdb"
background [255,255,255]
set ambient 40
set specular off
```

- By changing the pathname to ".\1J95.pdb", you are making a generic pathname (ie: load the PDB file located in "this" folder). If you have a folder on the new computer with RasMol, 1J95.pdb and the script file, you will be able to open the script file. This generic pathname will work, only if the folder contains RasMol, the PDB file and the script file. You must launch RasMol from this folder for the generic pathname to work.
- After you have saved your script file and you know that you are going to be changing computers, we recommend changing the pathname. This is especially

important if you plan to email the sript file amongst team members. Once you change the pathname, you will not need to do it again.

12. RasMol Commands

- Section I of the MSOE CBM RasMol Training Guide introduced you to the basic commands within the RasMol program. We will now present additional commands to further develop RasMol design skills.

❖ The "*" Command

- The * key within RasMol is wild card for a whole field in which you can designate all atoms within a field.
  - For example, if you wanted to select all of the atoms within Chain D of 1J95.pdb, you would type:
    - RasMol>select *d
      - This command would then select all of the atoms within chain D of the K channel.

❖ Restrict

- This command allows you to restrict your view of the molecule to specific to portions of the molecule dictated by you. You can restrict to specific chains, specific amino acid sections or specific features.
  - For example, if you wish to only look at chain D in the K channel, you can type:
    - RasMol>restrict *d
      - Notice that we have used the "*" command to select all of the atoms in Chain D and we have used the "restrict" command to restrict our viewing of only the atoms in Chain D.
      - Note that with the restrict command, all of the other parts of the molecule have disappeared from view. The other atoms are still in the PDB file, but you have "restricted" your view to the specified region.

❖ Center

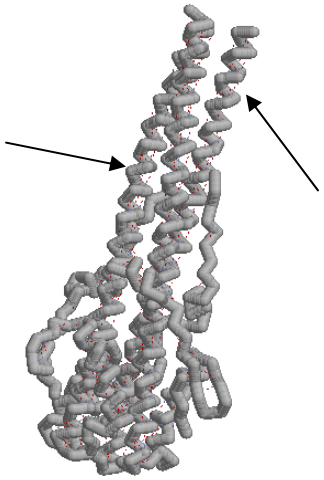
- The "center" command will allow us to center the molecule around a certain portion of the molecule.
  - RasMol will by default center the molecule at the center of the entire molecule. If you restrict your viewing to a certain subset (as we did with the K channel to just Chain D), when we rotate the molecule around in space, the molecule will seem lopsided. This is because although we have restricted our viewing of the molecule to just Chain D, all of the atoms are still present, we just do not see them.
  - Therefore, we can use the "center" command to center the molecule on the restricted region.

        - You will need to enter this command every time that you open the script file.
            - To center the K channel on Chain D, you can type:
                - RasMol>Center *d
                    - Notice that once again, we are using the "*" command to dictate to RasMol that we are centering the molecule around atoms in Chain D.

- At this point, we will open another PDB file so that we can illustrate the remaining items that we will discuss in this section. If you wish to save your script file at this point, you may do so. Remember to zap the existing molecule so that you can open a new PDB file. Or use the file pull-down menu and select "close".
- Please search the Protein Data Bank for 1HTM.pdb and download this file to your folder.
- Open 1HTM.pdb on your computer.
    - Change the background to white.
    - Change the display to backbone.

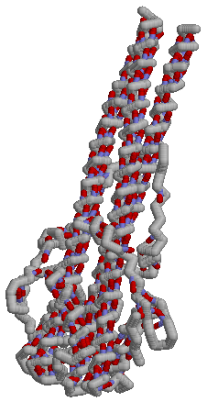
❖ Hydrogen Bonds

- Hydrogen bonds are essential to the stability of secondary structures.
- To add hydrogen bonds to secondary structures within your molecule:



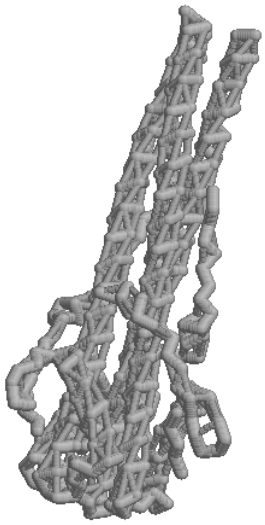
- RasMol>hbonds
    - After you have enter this command, you will notice that there are dotted lines that have appeared. These are the hydrogen bonds.
    - Since these bonds are very difficult to see, let's give them an added dimension.

- RasMol>hbonds 225
    - Notice that the Hydrogen bonds are now thicker, but that they appear to be floating in air. This appearance results from the fact that hydrogen bonds form between the atoms that make up the backbone of the amino acid (the nitrogen and the oxygen atoms), but since we have displayed only the alpha carbon atoms, it appears as if the hydrogen bonds are floating in space. Therefore, we must set the hydrogen bonds to the backbone. Notice that the bonds are blue/red, as they



are based on the CPK color scheme since they connect an oxygen (red) to a nitrogen (blue).

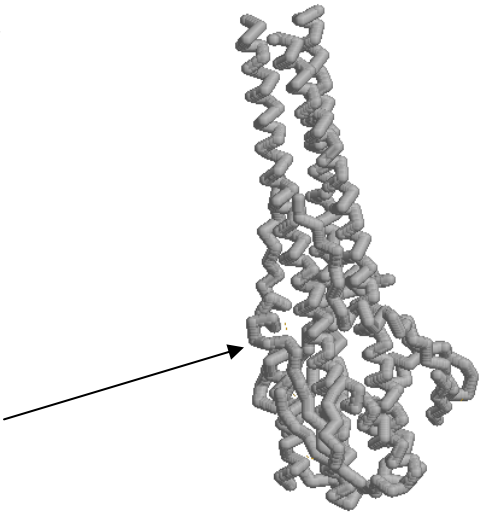
- RasMol>set hbonds backbone
  - Notice that now the hydrogen bonds are attached to the alpha carbon backbone. The bonds are now colored gray since they are connecting the alpha carbon backbone, which is gray (carbon), to another part of the alpha carbon backbone. When bonds connect different sections, they adopt the coloring of the section that they are connecting.



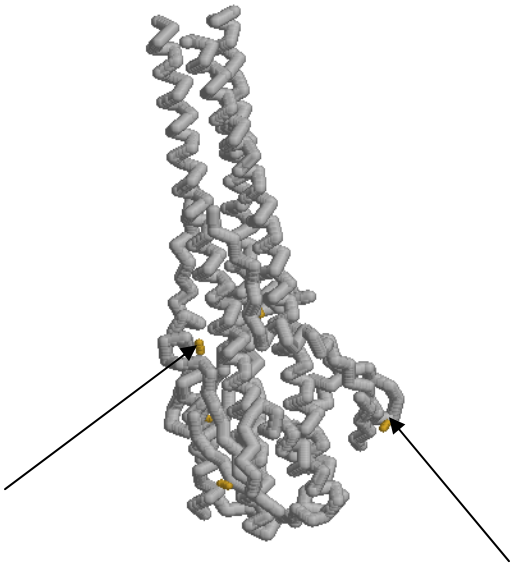
- To turn off the hydrogen bonds:
  - RasMol> hbonds off
    - This command will turn off expression of all of the hydrogen bonds within the molecule, bringing you back to the original state of the molecule in which no hydrogen bonds were displayed.

❖ Disulfide Bonds

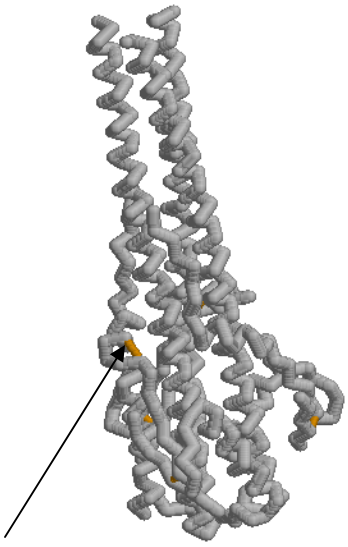
- Some molecules will have disulfide bonds present within the structure. These bonds are between two nearby cysteine amino acids. To visualize these bonds, we can type:
  - RasMol>ssbonds
    - As we saw with the hydrogen bonds, simply typing "ssbonds" will only produce dotted lines. To give these bonds dimension, we must add a value (thickness) to the ssbonds.



- RasMol>ssbonds 225
  - Notice that this command gives the disulfide bonds a thicker dimension, but as we saw with the hydrogen bond, the bond is "floating" in space. This is because the disulfide bond is actually between the sulfur groups of the cysteine sidechains, and not the alpha carbons. To make the disulfide bond connect between the backbone units, we need to set the bonds to the backbone. Note that the disulfide bond is orange (the CPK color for sulfur).



- RasMol>set ssbond backbone
  - Notice that this command sets the disulfide bond to the backbone. Once it does this, the disulfide bond is now gray. RasMol will assign the bond to adopt the color of the backbone that the bonds connect. Since the backbone is gray, now the disulfide bond is gray. Since sulfur is orange in the CPK color scheme, we will enter a command into the command line window to color the bond orange.



- RasMol>color ssbond orange
  - Note that we specified that we wanted to color the ssbond orange. We need to be specific that we want to color the ssbond orange. What happens if you type in "color orange"?

- You may wish to display the disulfide bond connected to the cysteine sidechains, rather than to the backbone of the molecule.
  - In order to set the ssbond between the cysteine residues, you need to selectively display the cysteines.
  - After you have these amino acids displayed, you can the set the disulfide bond to the sidechains, rather than to the backbone.
    - RasMol>set ssbond sidechain
    - This will place the bond in between the sulfur atoms of the cysteines, rather than the alpha carbon backbone, as we saw earlier.
- To remove ssbonds:
  - RasMol>ssbonds off

13. Identifying the Amino and Carboxy Termini

- An important concept in protein structure is that each protein has an amino terminus and a carboxy terminus. Through RasMol, each student can readily identify each of these termini.
  - Amino Terminus
    - The Amino Terminus is the first amino acid in the protein. When a protein is synthesized, it begins with the 5' end of the mRNA and synthesizes in a 5' to 3' fashion. Therefore, the first amino acid in the protein will be the amino acid that is encoded at the 5' end of the mRNA.
    - To determine the amino terminus of the protein in the PDB file, click on the atom at the end of the protein. The atom with lowest amino acid number will be the amino terminus.
    - Alternatively, you may search the PDB sequence information to identify the amino terminus amino acid and use the RasMol command line window to select the specific amino acid. (See Number 1, Sequence Details of the Protein Data Bank File).
  - Carboxy Terminus
    - The Carboxy Terminus, on the other hand, will be the last amino acid in the protein.

- To determine the identity of the carboxy terminus, click on the atom at the end of the protein. The amino acid with the largest number will be the last amino acid in the protein
    - For example, in the 1HTM.pdb protein, in chain, the amino terminus is Ser 40 and the Carboxy terminus is Arg 153.

Conclusions for Section II

At the end of Section II, you should be comfortable with

- Searching the Protein Data Bank
- Reading a structure summary page
- Downloading a PDB File
- Saving script files
- Opening script files
- Editing script files
- The “wild card” (*) command
- Restricting a section of the molecule
- Centering the molecule
- Adding hydrogen bonds
- Removing hydrogen bonds
- Adding disulfide bonds
- Removing disulfide bonds
- Identifying the amino and carboy termini

Answers to Questions Posed within Section II

4. Searching the PDB and Reading a Structure Summary Page (practice)

- Use this structure summary page to answer the following questions about the PDB file (answers are at the end of this section):
  - Who are the authors of the PDB file?
    - *Li F, Li W, Farzan M, Harrison SC*
  - In which journal was the primary citation published?
    - *Science*
  - On what date was the file deposited into the PDB?
    - *1 August 2005*
  - How many chains are in this file?
    - *4 (A, B, E, F)*
  - Are there any heterologous groups within this PDB file? If so, which ones?
    - *Yes*
      - *Zinc, Chloride, Mannose, N-acety-D-glucosamine*
  - From what source was this molecule isolated?
    - *Human*

7. A little review from the previous section

- How do we do the following items?
    - (answers at the end of this section)
  - Change Background to white
    - *RasMol>background white*
  - Change display formats
    - *Use the pull-down menu (display)*
    - *Use command line window*
      - *RasMol>backbone*
      - *RasMol>wireframe*
      - *RasMol>spacefill*
  - Change colors
    - *Use the pull-down menu (color)*
    - *Use the command line window*
      - *RasMol> color red*
  - Highlight the helices as red
    - *RasMol>select helices*
    - *RasMol>color red*

Question: Why are there so many hydrophobic amino acids located on the *outside* of this protein when we learn that hydrophobic amino acids are typically located buried within the molecule, away from the aqueous external environment?

Answer: *The K channel is a membrane embedded protein and therefore will be in contact with a hydrophobic region of the membrane. As such, it is necessary for the part of the K channel that interacts with the hydrophobic portion of the membrane needs to be hydrophobic as well.*

RasMol Training Section III:
Designing a Model to be Built on the Rapid Prototyping Machines

Through this section of the RasMol Training Guide, you will become familiar with the commands needed to design a model that will be built on the rapid prototyping machine. As you become more comfortable using RasMol, this section will enable you to take the next step and be able to become a model designer.

In this section, you will learn how to

- Select an appropriate display format to use in your model design
- Add monitor lines for structural support
- Use Boolean operators
- Add hydrogen bonds within the beta sheet
- Remove "triangle bonds" within beta sheets
- Add sidechains to create a "clean backbone"
- Determine the appropriate sizes for Z corporation printed models
- Select appropriate colors to be used for the Z corporation printer

Model Design

1. How do you choose which type of display format to use in your model design?

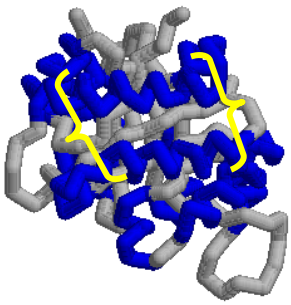
- This is a common question that is asked by model designers. And the answer to this question is another question: "What is the story that you are telling with your model?"
- If your story is focused on a particular active site within the molecule, then perhaps the alpha carbon backbone model displaying key active site amino acid sidechains is the best display format to choose. If your story is focused on how two subunits interface at the surface, then perhaps the spacefill format is the best choice. Ultimately the choice is yours. In Section I, there is a table highlighting the advantages and disadvantages of each display format. This may assist you in deciding which format is the best for telling your story.
- The important point to remember is that no one model will tell every aspect of the story. Using RasMol in combination with the physical model will assist you in telling multiple aspects of your story.
- The CBM will build models in the spacefill format, wireframe and the alpha carbon backbone format. Models in the alpha carbon backbone format can have sidechains and heterologous groups displayed. We currently do not have the ability to build models in cartoon, ribbons or strands format.

2. Monitor Lines: What are they and where to place them

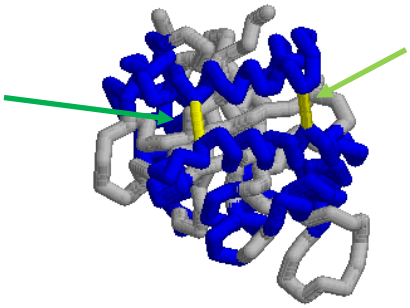
- When a model is built on the rapid prototyping machine, it is done through a layering system. A layer of powder 1/1000th of an inch thick is spread out. This powder is impregnated with droplets of binder with ink (Z corporation printer) or scintered together with a laser (SLS machine). At the end of the production time,

the loose powder not incorporated into the physical model is vacuumed away. During the build time, there needs to be support within the model in order to withstand the additional pressure that accumulates as the powder builds up. Hence, monitor lines are added within the model for support.

- How do you know where to place Monitor Lines?
  - Monitor lines are NOT needed:
    - Within Beta Sheets
      - If a molecule has beta sheets within the structure, the presence of the hydrogen bonds within these sheets will provide very good support. Therefore, you do not need to add monitor lines within beta sheets.
    - Within Alpha Helices
      - Alpha helices are very stable internally, therefore, neither hydrogen bonds nor monitor lines are needed within the helix structure itself.
  - Monitor lines ARE needed:
    - Monitor lines will be needed in regions of the protein that look like they might be able to squeeze together.
      - A good example of this: If you look at the image below on the left, you can see the alpha helices are colored in blue. These helices are stable from top to bottom, as mentioned above, but they do have the tendency to flex from side to side. (See regions indicated by yellow brackets in the figure below, which is based on 1TIM.pdb.) Due to this flexibility from side to side, monitor lines are needed to stabilize the helices.
      - By adding monitor lines in between the helices, as shown in the figure below as yellow bars (and pointed to by the arrows), you will increase the stability of this protein and prevent the flexing of the helices.



Before monitor line addition

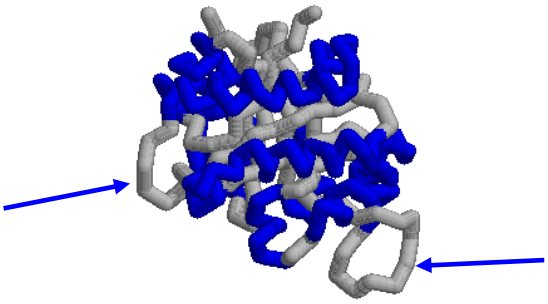


After monitor line addition

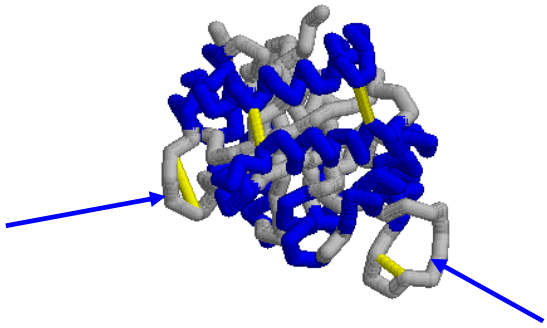
- To determine if the model will need monitor lines, we do the "squeeze test."
    - The "squeeze test" refers to the ability to potentially squeeze a portion of the molecule together. When you look at the image of the molecule on the computer screen, can you see potential regions where you might be able to squeeze the model together, or potential flex points? If so, then you need to add monitor lines to this region to prevent this "squeezing" from occurring.
- We recommend that you add monitor lines at the "top" and "bottom" of the helix in order to anchor the helix. If the helix is exceptionally long, then adding a monitor line in the middle will further stabilize the helix.

- Monitor lines are also needed where there are large loops and turns within a molecule that could potentially be sites that need additional support.

In the image on the left, the blue arrows indicate a loop that would need a monitor line to stabilize the loop. On the image on the left, the blue arrows indicate the positions in which yellow monitor lines have been added to stabilize the loops.



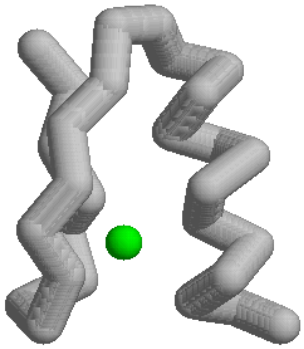
Before monitor line addition



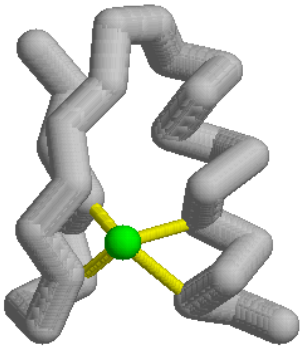
After monitor line addition

- Monitor lines should also be used to connect heterologous groups, such as the zinc ion in a zinc finger, as shown in the figure below. (This figure is based on 1ZAA.pdb, which was used in Section I of the training guide.)

    - If heterologous groups are not connected with monitor lines, they will be built as a separate piece. For example, in the figure to the left there are not currently monitor lines attached to the zinc ion, so the protein component (gray) would be built as one unit and the zinc ion (green) would be a separate piece.

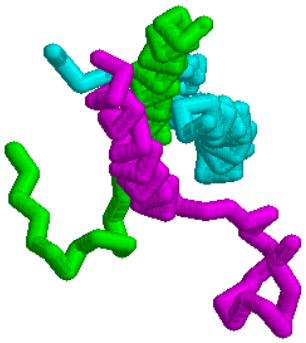


Zinc ion that is **not** attached with monitor lines.

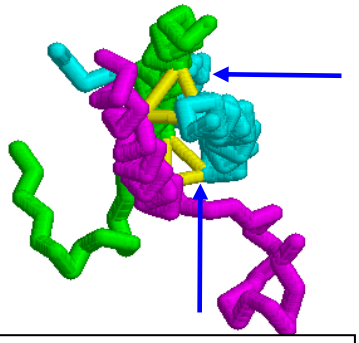


Zinc ion that is attached with monitor lines.

- Monitor lines are also needed to connect subunits if your molecule has multiple subunits. If you do not use monitor lines to connect the subunits, the molecule will build as separate subunits that are not connected.
    - Below are two figures of fibrinogen, based on 1JY2.pdb, which has three chains, each a different color. The figure on the left does not have any monitor lines and if it were built in this fashion, it would be built as three separate chains. In the figure on the right, there are monitor lines (show in yellow and indicated by the arrows) that will hold the three chains together.



Before monitor line addition



After monitor line addition

❖ **Final Note on monitor line placement:**

Monitor lines are used to stabilize the molecule. Anywhere that you think that the molecule looks like it needs additional support, add a monitor line.

3. Adding monitor lines

- Monitor lines are added to the molecule by entering in a command in the Command Line window.
- To add the monitor line, you must first identify the atom numbers for the atoms between which you intend to draw the monitor line. To do this, move your mouse cursor over the atoms and click the **left** mouse button. This action will generate information in your command line window that will provide you with the atom identity, as described in Section I. The first number that is provided in that line of

information is the atom number, and this is the number that you need in order to create a monitor line.

- Once you have identified the two atom numbers between which you would like to have the monitor line exist, the next step is to enter the command to make a monitor line:
  - RasMol>monitor atom number atom number
    - Example: RasMol> monitor 656 8567
      - It is essential to put a space between the numbers.
      - It is also essential to use the **atom number** and **not the amino acid number**.
    - You will notice that monitor lines initially appear as a dotted line, just as the hydrogen bonds and disulfide bonds did (see Section II).
    - In RP-RasMol, to add dimension to monitor lines:
      - RasMol>set monitor 225
        - Please note that this feature is only available in RP-RasMol. You can create monitor lines in RasMol, but you can only add dimension to the monitor lines with RP-RasMol. To obtain a copy of RP-RasMol for academic purposes, please contact Tim Herman at herman@msoe.edu.
  - Once you have given the first monitor line dimension, you will not need to do this again. All future monitor lines will have the same dimension. All monitor lines will have the same dimension.
  - The monitor line command is similar to a toggle switch. If you decide that you do not like the position of the monitor line or if the monitor line is too long, you can simply re-enter the monitor line command (in the above example, RasMol>monitor 656 8567) and it will turn off the monitor line.
  - If you wish to turn ALL of the monitor lines off:
    - RasMol>monitor off
      - Please note that this command will turn off **all** of the monitor lines within the molecule. To turn off specific monitor lines, use the monitor command above.
  - A number will appear next to the monitor line and this number is the length of the monitor line in angstroms. A monitor line should not be longer than 9 angstroms. If the monitor line is longer than 9 angstroms, it is not stable and it defeats the purpose of being present to stabilize the molecule.

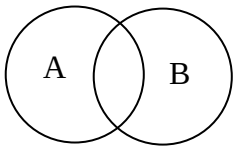
4. Coloring the Monitor lines

- When the monitor line initially appears, it will be the color of the atoms that the monitor line is connecting.
    - If the two atoms are the same color, then the monitor line will be the same color throughout the length of the monitor line.
    - If the two atoms are different colors, then the monitor line will be half one color and half the other color.
- To specifically color the monitor line (as opposed to the default setting described above):
    - RasMol>select all
    - RasMol>color monitor white
- It is essential that you include "monitor" within your command. If you forget to include the "monitor" within the command, you will color your entire molecule whatever color you have selected for your monitor color (in the above example, white).
- We recommend that you choose a light color for your monitor lines, such as white or light gray. Monitor lines are support structures and should not be the emphasis of your model. A bright or dark color will draw the user's eye to that feature and that should not be the focus of your model.
- We at the CBM often color monitor lines and hydrogen bonds the same color (white). Some SMART Teams in the past have opted to color their monitor lines a different color from their hydrogen bonds to differentiate the two types of structural features. This is entirely up to the designers. Our only recommendation is to downplay the color (light colors) choice in order to prevent the focus from being on these structural features.

5. Boolean Operators

- You can link together RasMol commands by using Boolean Operator (And, Or, Not) in order to select very specific things in RasMol.

- Boolean Operators

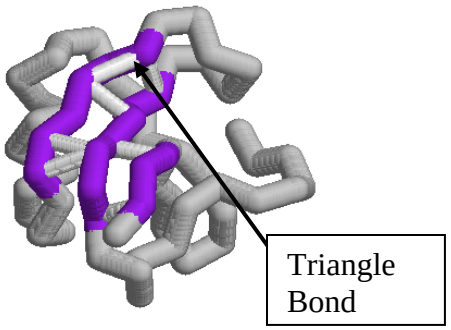


    - OR (RasMol>select A or B)
        - Selects everything in both circles
    - AND (RasMol>select A and B)
        - Selects only that which is in the region that overlaps between the two circles
    - NOT (RasMol>select A not B)
        - Selects the region in A that does not overlap with B

- Practice:
  - Round #1
    - Select men **and** women
    - Stand up
      - NO ONE should have stood up (unless you are a hermaphrodite)
  - Round #2
    - Select men **or** women
    - Stand up
      - EVERYONE should have stood up because this selection process selects everyone who fits the category Men as well as everyone who fits in the category Women
  - Round #3
    - Select women **and** Wisconsin residents
    - Stand up
      - Only women who live in WI should have stood up
  - Round #4
    - Select men **or** those who are wearing black shirts
    - Stand up
      - All men should have stood up, as well as women wearing black shirts
      - **Note** that some may have fulfilled both aspects in that men with black shirts may be present.
- The use of Boolean operators is essential in RasMol in order to select specific portions of the molecule. For example, if you wish to select the sidechain of histidine 63:
  - RasMol>select his63 and sidechain
    - This command selects the atoms that meet both criteria: being a part of histidine 63 and an atom within the sidechain (as opposed to an atom within the backbone).
- As in math, operations that are placed within parentheses are performed first.
  - RasMol>select his63 and (sidechain or alpha)
    - This command will select the sidechain atoms as well as the alpha carbon atom of histidine 63.

6. Hydrogen Bonds

- As discussed in Section II, hydrogen bonds can be added with the command "Hbonds." Since alpha helices are stable structures in terms of buildability on the rapid prototyping machines, please note that we do not recommend placing hydrogen bonds in alpha helices. We do, however, highly recommend placing hydrogen bonds within beta sheets. The placement of these bonds within the beta sheets stabilizes the structure.



- To add hydrogen bonds specifically to the beta sheets:
  - RasMol>select sheets
  - RasMol>hbonds 225
  - RasMol>set hbonds backbone
    - Note: If your molecule has multiple chains, you will need to be more specific with your command. Ie: RasMol>select *a and sheets

- You may notice that some hydrogen bonds may appear to look like "triangle bonds" in that they connect two alpha carbon atoms to create a triangle. The two amino acids will be N and N+2. For example: Amino Acid 6 and Amino Acid 8, or Amino Acid 55 and Amino Acid 57.

- These "triangle bonds" are not "real" and are distracting when we build the alpha carbon backbone models, and they do not add stability to the physical model. We recommend that you remove these bonds.

- To remove these bonds, select the amino acids and remove the bond.
  - RasMol>select 6 or 8
  - RasMol>hbond off
    - Note: if you molecule has multiple chains, you will need to be more specific with your command. Ie: RasMol>select *a and (his6 or arg8)

7. Adding Sidechains with a clean backbone

- If your story requires that you display specific amino acid sidechains to the model, we recommend that you do so in a way that only displays the sidechain atoms, rather than all of the atoms of the amino acid. In previous sections, we have simply selected the amino acid and displayed all of the atoms. In this section, we are going to use the Boolean operators to select just the atoms in the sidechain and display only these atoms.

- To select and display only the atoms of the sidechain of a specific amino acid:
  - RasMol>select his63 and (sidechain or alpha)
    - This command selects the amino acid (histidine 63), but limits the selected atoms to the sidechain atoms and the alpha carbon of that amino acid. It is important to select the alpha carbon atom in addition to the sidechain atoms because we need to attach the sidechain atoms to the alpha carbon. If we do not select the alpha carbon, the sidechain will build as a separate unit from the rest of the molecule

- RasMol>wireframe 225
- RasMol>spacefill 275
    - These two commands together will generate a ball and stick appearance. You can enter just the wireframe command and create a "sticks" appearance, but you cannot enter just the spacefill command. If you enter just the spacefill command, the atoms will be displayed as just little spheres and the spheres will not be connected to one another. It is therefore imperative to add the wireframe command in order to connect the spheres together.

- If you just select the amino acid with command of "select his63", you will select all of the atoms within the amino acid and will generate a "bumpy backbone." Unless the backbone atoms of the amino acid (the amino nitrogen or the carbonyl oxygen) play a specific role, then we generally do not recommend displaying these atoms on the model. It is typically (although not always) the sidechain that has the specific chemical role within the molecule and plays the key role within your story. Therefore, this should be the part that is displayed within your molecule.
- If you need to selectively display the backbone atoms and not the sidechain atoms, see number 10 for additional special design commands.

8. Use of Command Window versus the Pull-Down menu for designing molecules

- We recommend that you use the command line window exclusively when designing your model, rather than using the pull-down menu in combination with the command line window.
- When using the command line window, you can create your model in a step-by-step fashion and you can selectively add or subtract features.
- The pull-down menu is limited in terms of what features can be displayed. For instance, you cannot selectively display sidechains through the pull-down menu. Combining the two command features can potentially cause problems in your design work as the pull-down menu options will could potentially over-write any designs that you have created through the command line window.
- To illustrate this feature, proceed through the follow exercise:
    - Open 1A3N.pdb
    - RasMol>restrict *b
    - RasMol>backbone 300
    - RasMol>wireframe off
    - RasMol>select helices
    - RasMol>color red
    - RasMol>select his63 and (sidechain or alpha) and *b
    - RasMol>wireframe 225
    - RasMol>spacefill 275
    - RasMol>select his92 and (sidechain or alpha) and *b
    - RasMol>select *b
    - RasMol>color cpk
    - Use the pull-down menu to choose Cartoon

        - Notice how the wireframe/backbone combination disappeared and was replaced entirely by the cartoon format.
        - More importantly, notice that the sidechains that you selectively displayed are no longer present. The pull-down menu has overwritten what you have created using the command line window.
    - Return to the command line window
    - RasMol>wireframe 225
        - Notice how you added the wireframe to the cartoon, rather than replacing it.
    - Return to the command line window
    - RasMol>backbone 300
        - Notice that you have added the backbone on top of the pre-existing cartoon and wireframe format.
        - In order to remove the cartoon feature, you will need to turn this off.
    - RasMol>cartoon off
- If you alternate between the pull-down menu and the command line window, you may end up with an odd combination of features, or you may lose work that you have you created within the model. This can be very frustrating, especially if you have spent time designing your model, all to be lost within a few seconds of using a pull-down menu command.
- For this reason, we recommend that you exclusively use the command line window for your design work.

9. Recommended values and colors to use within your model

- During the summer of 2005, the Center for BioMolecular Modeling and 3D Molecular Designs purchased a new Z Corporation printer. This new printer is more accurate with its printing and we have developed a new guideline for recommended values for designing a model. These are the design values that we recommend using if your molecule is 200-1000 amino acids in length. If your molecule is smaller or larger, please contact us for other recommended values.
    - Backbone 300
    - Spacefill 275
    - Wireframe 225
    - Monitor lines 225
    - Hydrogen bonds 225
    - Disulfide bonds 225
- We also have recommendations with respect to colors:
    - The printer is a CMY printer (cyan, magenta, yellow)
        - Therefore, cyan, magenta and yellow are considered to be the "pure" colors and will be the brightest.
        - All other colors will be a mixture of these three colors.

        - Secondary colors created by mixing any two of the primary colors are also a good color choice. Ie: cyan and yellow will give you green.
        - We will **NOT** print black.
    - Colors should highlight features. In your model, the colors that "jump out at you" should be on the features that you want to emphasize in your model to assist in telling a specific story.
        - For example, you would not want your monitor lines to be colored magenta and everything else colored a light color, since the monitor lines will become the focal point of your model and the monitor lines should be the least important feature of your model. You should reserve the bright colors, like magenta, for the important features, such as the helices or the sidechains of the active or binding site.

10. Additional Design Commands

- Selecting a specific atom
    - If you would like to specifically select an atom, you would use your mouse to identify the atom number, (see Section 1 for further information on how to do this) and then you would use the "select" command and the "atomno=" command, which is needed to identify the atom number, rather than the amino acid number.
        - If a number is entered into the prompt line, the default assumption is that the number is the Group Number (such as the amino acid number). If you wish to specifically select the atom number, you need to designate the number as the atom number.
            - RasMol>select atomno=376
                - This command will select the atom within the molecule that is numbered 376.

- Selecting a region within the molecule
    - You can select a region within the molecule by selecting the range of amino acids or a range of atoms
        - Range of Amino Acids
            - If you would like to select amino acids within a range, you can do so by entering the following command:
                - RasMol>select #-$
                    - The # indicates the number of the first amino acid of your range and the $ indicates number of the last amino acid of your range.
                - For example:
                    - If you wished to select amino acids 4-31 of the molecule, you would enter the following command:
                        - RasMol>select 4-31

- Note: If you need to select 4-31 of a particular chain, you need to include that within your selection criteria.
    - RasMol>select *a and (4-31)

- Range of Atom Numbers
    - If you would specifically like to select a range of atom numbers, you can do so by entering the following command:
        - RasMol>select (atomno>=___) and (atomno<=___)
            - the underline region is where you would insert the specific atom number of your range
        - For example:
            - If you would like to select all of the atoms between 436 and 862:
                - RasMol>select (atomno>=436) and (atomno<=862)

- Turning the monitor labels off
    - With each monitor line added to the molecule, there will appear a number next to the line that designates the length of the monitor line in angstroms. If you are exporting this image, the numbers may clutter the image. Or, if there are several monitor lines, the numbers may hinder your viewing of the molecule. You may then wish to remove the labels from the molecule.
    - To turn off the labels, enter the following command:
        - RasMol>set monitor off
    - Note: Make sure that you have the word "set" included in the command. If you enter the command "monitor off", all of the monitor lines that you have added to the molecule will be turned off.
- Repeating a command
    - By pressing the up arrow key on the keyboard, the previous commands that you have entered into the command line window will be repeated.
- Selecting specific types of atoms
    - If you wish to select a subset of atoms, such as all of the alpha carbons, you can do so by using the "*" command and the 1 or 2 letter codes for the different atoms within the molecule
    - For example, if you wished to select all of the alpha carbons:
        - RasMol>select *.ca
            - The "ca" refers to the alpha carbons

Atom in Backbone	**RasMol Designation**
Nitrogen	n
Alpha Carbon	ca
Carbonyl carbon	c
Carbonyl oxygen	o

Conclusions

At the end of this section, you should feel comfortable designing a molecule that could successfully be built on the Z Corporation machine. To this end, you should be comfortable with the following items:

- Adding hydrogen bonds
- Removing triangle bonds
- Adding monitor lines
- Adding sidechains with a clean backbone
- Using “good” colors and “good” values

Any comments/suggestions should be sent to Shannon Colton at **colton@msoe.edu**

Acknowledgements

I would like to thank the teachers who have attended the workshops that the MSOE Center for BioMolecular Modeling has presented over the years for their input into creating this document. Their inspiration and requests for something that would assist them in working with RasMol has allowed me to put this Training Guide together.

I would like to thank Margaret Franzen for her valuable insight while I was writing.

I also thank Jennifer Morris for all of her contributions to this document. Without her training initially, I would not know how to use RasMol, and would not be able to write this document in the first place. Jennifer also provided many comments on rough drafts of the Training Guide and her editing of the document was very much appreciated.

I also to need to acknowledge the support of Tim Herman, Director of the Center for BioMolecular Modeling, as he continues to encourage me to "do something remarkable."

This publication was made possible by Grant # 1 R25 RR022749-01 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH).

Appendix to RasMol Training Guide for Macintosh Users

1. Mouse Control keys

As those of you who use Macintosh computers are aware, there are differences with the mouse and keyboard keys between PCs and Macs. The table below provides a comparison between the mouse/key combinations needed to perform different functions within RasMol between a PC and a Mac.

Action	Windows	Macintosh
Rotate X, Y	Left	Unmodified
Translate X, Y	Right	Command*
Rotate Z	Shift-Right	Shift-Command*
Zoom	Shift-Left	Shift
Slab Plane	Ctrl-Left	Ctrl

*On some Macs, the Option (Alt) key has the same effect on RasMol as the Command key.

2. Changing the PDB extension

When you download a PDB file from the Protein Data Bank, the file is saved with a **.pdb** extension. Macintosh users should change the **.pdb** extension to **.txt**.

For example, **1a3n.pdb** becomes **1a3n.txt**.

In some cases, we have had users need to add the **.txt** to the **.pdb**, rather than replacing the **.pdb**. In this case, **1a3n.pdb** becomes **1a3n.pdb.txt**.

3. Editing Script files

As with script files generated on a PC, script files that will be used between Mac computers will need to have the path name modified in order to be opened when moved from one computer to another (or from one computer to another). The modification of the script file is more specific with script files on a Mac than on a PC.

On the PC script file, this is how it appears:

```
#!rasmol -script
# File: kchannel.spt
# Creator: RasMol Version 2.6

zap
```

```
load pdb "D:\CBM\Summer 2006\RasMol Training 2006\RasMol Training
II\1J95.pdb"
background [255,255,255]
set ambient 40
set specular off
```

On a script file that was generated on a Mac, this is what will appear:

```
#!rasmol -script
# File: kchannel.spt
# Creator: RasMol Version 2.6

zap
load pdb "MacIntosh HD:Users:SColton:Desktop:Kchannel:1J95.txt"
background [255,255,255]
set ambient 40
set specular off
```

Notice that there are colons, rather than back slashes.
Notice that the computer username is listed after users.
Notice that the location of the file (desktop in this case) is listed.
Notice that the folder in which the PDB is located is listed (Kchannel in this case).
Notice that there is a .txt extension, rather than a .pdb extension.

Please NOTE: If you move the script file from a Macintosh computer to a PC computer, it will be essential to change the load command line to the PC format (with the back slashes, rather than the colons). Likewise, if you move a script file from a PC to the Macintosh computer, it will important to change the load command line to the Macintosh format.

4. Operating Systems

If you have OS X on your computer, in order to run RasMol, you have to have Classic Environment (emulator) installed on your computer. If you need to install this on your computer, you can visit http://docs.info.apple.com/article.html?artnum=151871.