

**Wisconsin Science Olympiad Protein Folding Challenge**  
**A Guide to Using RasMol for Exploring Protein Structure**

Prepared by MSOE Center for BioMolecular Modeling  
Milwaukee, WI

Shannon Colton, Ph.D.

Timothy Herman, Ph.D.  
Director of Center for BioMolecular Modeling

Jennifer Morris, Ph.D.

January 2005

# RasMol

## A Guide for Designing Molecular Models

RasMol is a popular molecular visualization software written by Roger Sayle in the early 1990's, while he was a graduate student at the University of Edinburgh. This program generates a computer image of a molecule based on x,y,z coordinate data contained in a specific format of a PDB file. Through a combination of depth-cueing and shadowing, the two-dimensional image on the computer screen takes on a three-dimensional quality, especially as the molecule is "rotated" in real time.

The Science Olympiad Team will use RasMol to generate a computer image of a particular protein. The teams will use this image to build a model of the protein using toobers.

To learn more about RasMol and its history, check out the following web sites:

the RasMol HomePage

<http://www.umass.edu/microbio/rasmol/index2.htm>

the RasMol Manual (recommended)

<http://www.umass.edu/microbio/rasmol/distrib/rasman.htm>

Roger Sayle's personal account of the history of RasMol

<http://www.umass.edu/microbio/rasmol/pershist.txt>

**This Guide is intended to provide Science Olympiad Teams with the basic information of RasMol necessary to compete successfully in the Wisconsin Science Olympiad Protein Modeling Challenge.**

## Table of Contents

### Guide to Using RasMol for Exploring Protein Structure

- A. Introduction
  - 1. Overview
  - 2. Molecule of the month
  - 3. Get the PDB file
  - 4. Identify key features of the molecule
- B. RasMol Tutorial and Design Guide
  - 1. Organizing your work
  - 2. Installing RasMol
  - 3. Downloading a PDB file
  - 4. Protein Databankd and RasMol
    - a. RasMol Windows
    - b. Background Color
    - c. Open PDB File
    - d. Mouse Controls
    - e. Molecule Display Window
    - f. Colors
  - 5. RasMol Commands
    - a. Display Formats
    - b. Select
    - c. Color
    - d. Restrict
    - e. Predefined Sets
    - f. Center
    - g. SS Bonds
    - h. Display Sidechains
    - i. Atom ID
    - j. Slab
  - 6. Saving Your Work as a Script File
  - 7. Boolean Operators
  - 8. Additional Tutorials

## Guide for Physical Model Design Using RasMol

### A. Introduction

1. Overview. This guide is provided as an outline for the Science Olympiad Teams and is not intended to be all inclusive. We limit the guide to presenting the RasMol commands you are likely to use in the Protein Modeling Challenge. Additional commands can be found at the RasMol website given below. The following discussion includes RasMol commands that are important to display features of protein structure.

For more detailed information regarding all RasMol commands, consult the RasMol Manual that is available at <http://www.umass.edu/microbio/rasmol/distrib/rasman.htm>

Another useful tutorial can be found at

<http://www.bio.cmu.edu/Courses/BiochemMols/RasMolTutorial/RasTut.html>

2. Molecule of the Month. The molecule that will be modeled at the state competition will be taken from David Goodsell's Molecule of the Month, which can be found on-line at <http://pdbeta.rcsb.org/pdb/>. Students will be responsible for learning about the Molecules of the Month from April 2004 through March 2005. We suggest that students read the Molecule of the Month descriptions, download the PDB files, and explore the protein structures in RasMol. It is important for the students to read through the information that David Goodsell has provided in the Molecule of the Month.

**Questions about the molecule at the State Competition will be taken from this material, and from material contained in the "header" of the pdb file.** A copy of the PDB file and the Molecule of the Month of the protein selected for the on-site challenge will be provided at the State Competition. Students can prepare by becoming familiar with the material for the on-site challenge.

3. Get the PDB file. After selecting a molecule, the appropriate PDB file is downloaded from the Protein Data Bank at <http://pdbeta.rcsb.org/pdb> (see page 5-6 for the procedure.) The PDB file contains the data that specifies the three-dimensional location of each of the atoms in the molecule

4. Identify key features of your molecule. The material provided in the Molecule of the Month will describe the structure and the key features of the molecule. These features and function may include:

- Number of amino acids
- Number of chains in the molecule (many structures will consist of a complex of several proteins)
- Alpha-helices or beta-sheets
- Substrate or co-factor binding sites
- Enzyme active site(s)
- Binding site(s) to other molecules
- Hydrophobic or hydrophilic regions
- Functional or other structural domains
- Disulfide bonds

After identifying the key features of the molecule, the Science Olympiad team will decide which features will be included in the model. A physical model of the protein region will be built using the toobers and materials provided at the Science Olympiad competition.

## **B. RasMol Tutorial and Design Guide**

1. Organizing your work. (VERY IMPORTANT!!) Create a folder named “RasMol” on your computer. In this folder you will place a copy of (i) the molecule PDB file and (ii) the RasMol program.

### 2. Installing RasMol

- Go to the RasMol homepage <http://www.umass.edu/microbio/rasmol/>
- Select RasMol from the top line.
- Select “Getting RasMol or Chime and their Manuals” from half way down the page.
- Select the first option, “Getting and installing RasMol”.
- Choose the version you need (“Download RasWin for PC/Windows”).
- Select “32-bit RasWin”.
- Under “Save as”, click on your “RasMol” folder and click SAVE.



### 3. Protein Data Bank and Downloading a PDB File

Let's start by looking the oxygen transporting protein hemoglobin (file 1A3N.pdb)

- Start an internet browser program (Netscape, Internet Explorer, Mozilla)
- Go to the Protein Data Bank at <http://pdbeta.rcsb.org/pdb>
- In the box labeled: “Enter a PDB ID or keyword” at the top of the page, type “1A3N” and click the button to the right “SEARCH” or press enter.
- This will bring up the “Structure Summary Page” for the 1A3N pdb file. This page will provide information about the protein structure including.
  - Title: A description of the molecule
  - Authors: the scientists who crystallized and solved the structure of the protein
  - Primary Citation: journal article in which this information was published
  - History: when the data was deposited into the databank and released for public viewing
  - Experimental Method: method used for collection of data points
  - Parameters: resolution at which the structure was crystallized (a smaller number means better resolution)
  - Molecular Description: identifies the number of chains in the molecule, whether it is protein or nucleotide
  - Functional Class: defines what role the molecule plays
  - Source: identifies from which organism the molecule was isolated

Chemical Component: identifies any heterologous groups (any ions or other molecules that are a part of the crystal structure)

- To download the file, click on the “File” link on the navigation bar above the “Structure Summary Page” title.
- Select the first option, “download a pdb file.”
- Select the “uncompressed (.pdb)” option. These files are small, so there is no need to download a compressed or zipped file.
- A file download box will appear. Click Save.
- You will be given the option to save your file into a specific folder. Make sure that you save this file to the same folder containing the RasMol program.

4. Rasmol.

A. RasMol Windows. Now that there is a pdb file in your folder, open the



RasMol2.6.exe

RasMol program by double-clicking on the RasMol icon.

RasMol has two windows that you need to simultaneously use. The first window is the “molecule display window” with a black background. This window automatically opens with the program. The second window is called the “RasMol Command Line” window and is opened by clicking on “RasMol Command Line” box located at the bottom of the screen.

Organize your computer screen so that both windows are visible and not overlapping.

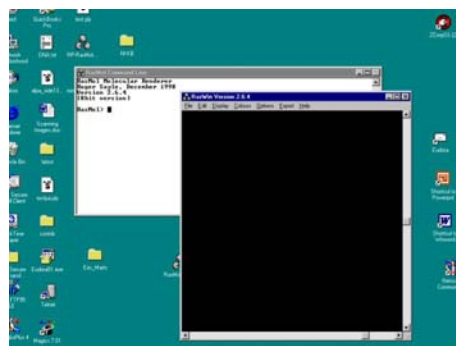
- Click on the structure window and drag the top of the window to nearly the top of the screen.
  - Open the RasMol Command Line by clicking on in the command line at the bottom of the screen.
  - Resize and reposition the RasMol command line window so the window is in the upper left hand corner of the computer screen and does not overlap the “structure” window.
- B. Background color. The background color of the molecule display window can be changed. We suggest changing the background color to white. Type the following command in the RasMol command window:
- RasMol>Background white

C. Open PDB file.

- In the molecule display window, select FILE from the menu bar and select OPEN.
- Click on 1A3N.pdb and select OPEN

D. Mouse controls. You can manipulate the image in the molecule display window using a combination of the keyboard and mouse.

- Left-Click and drag the mouse on the background, and the molecule will



- rotate about the X- and Y-axis running through its center-of-mass.
- Right-click and drag the mouse on the background, and the molecule will be translated up, down, left, and right.
- Shift+Left-Click and drag the mouse on the background, and the molecule will zoom in (up) and out (down). Don't zoom in too fast or the program will bomb and you will have to restart it.
- Shift+Right-click and drag the mouse on the background, and the molecule will rotate about the Z-axis.

Later, in section 5 on page 10, we will have you demonstrate that Ctrl+Left, when the Slab Plane command is activated, moves through the molecule image.

- E. Molecule display format. RasMol can display the atoms of the protein in different formats on the computer screen. The default display for a molecule when the file is first opened is a wireframe structure colored in cpk (a standard color scheme used by chemists). We will be constructing a toober model of the challenge protein based on a RasMol image of the protein displayed in the backbone (or alpha-carbon backbone) format. In this format, each amino acid is represented by only the alpha carbon. None of the other atoms of the amino acid are shown.

You may modify the image of the molecule in either the molecule display window or the command line window. The molecule display window has a bar across the top of the window with pull down menus. In the command line window, you enter commands by typing specific RasMol commands following the Rasmol prompt (see section 5).

To display hemoglobin in the alpha carbon backbone format, use the "Display" pull down menu in the molecule display window. Select "backbone".

To increase the thickness of the backbone, go to the RasMol "command" window and type:

- o RasMol>Backbone 250

- F. Colors. Unless other colors are selected, RasMol will use the CPK method to assign colors to the molecule. CPK is a standard chemistry method of identifying specific atoms by color. Here are the colors of the cpk system:

Carbon	light gray	Chlorine	green
Oxygen	red	Bromine, Zinc	brown
Hydrogen	white	Sodium	blue
Nitrogen	light blue	Iron	orange
Sulfur	yellow	Calcium, Metals	dark gray
Phosphorus	orange		

You may also assign other colors to selected regions of a protein. The predefined colors recognized by RasMol are shown in the following chart along with their RGB definition:

Blue	[0,0,255]	black	[0,0,0]
------	-----------	-------	---------

Cyan	[0,255,255]	green	[0,255,0]
Green blue	[46,139,87]	magenta	[255,0,255]
Orange	[255,165,0]	purple	[160,32,240]
Red	[255,0,0]	red orange	[255,69,0]
Violet	[238,130,238]	white	[255,255,255]
Yellow	[255,255,0]		

5. RasMol Commands. Open RasMol and arrange the molecule display window and the command line window on your desktop. Use the File pull down menu to open the 1A3N pdb file for hemoglobin. Change the background to white.

A. Display formats: **Backbone, Spacefill, Ball and Stick, Wireframe.**

- PDB files initially open in RasMol as a cpk colored wireframe display. The toober models will be based on the alpha-carbon backbone format. Change the display to the “backbone” display format. In the molecule display window, select the “Display” pull down menu and select “**Backbone**”.
- You can change the thickness of the backbone in the command line window. You can use values between 1-750. In the command line window, type:
  - RasMol>Backbone 250
- **Spacefill** format represents the volume of the atoms. You may wish to use spacefill for heterologous atoms or amino acid sidechains. To change the display to spacefill, use the molecule display window and the pull down “Display” menu. Select “spacefill”
  - The volume of the spacefill can be specified with values between 1 and 750. To change the spacefill size, type in the command line window:
    - RasMol>Spacefill 250
- **Wireframe** format represents the atoms as a bent wire and cpk color. This format is useful for displaying amino acid sidechains. You select “wireframe” display by using the pull down Display menu in the molecule display window. You may change the thickness of the wireframe by typing a specific value in the command line window.
  - RasMol>Wireframe 200
- **Ball and stick** format has the individual atoms represented by a “ball” and covalent bonds represented with a “stick”. You can select the “ball and stick” format in the molecule window pull down “Display” menu. To change the size of the ball and sticks, use the command line window and type the following:
  - RasMol>Spacefill 250

- RasMol>Wireframe 200
- B. Select. Sometimes you may wish to modify the display of only a portion of the molecule that you exploring. To select a portion of the molecule you type the command “select” in the command line window. **Note: when you select a part of the molecule, any subsequent commands only apply to the selected region and do not apply to the whole molecule.** To select amino acids 1-30, type the following in the command line window:
- RasMol>Select 1-30
- C. Color. After you select a region of the molecule you can change the color by entering typing the command “color” in the command line window. RasMol has predefined colors that are listed on the condensed quick reference card to RasMol. **The “\*” can be used as a “wild-card”. “\*A” would select all of the atoms in chain A.** To select the atoms of chain A in a PDB file and change the color to red, type the following in the command line window:
- RasMol>Select \*a
  - RasMol>Color red
- D. Restrict. The PDB file contains the x,y,z coordinates for the entire protein and may contain more than one protein chain. If you want to only display a single chain or a small portion of the protein, use the command line window and the “restrict” command. When you use the restrict command, the rest of the protein will be “invisible” on the screen.
- RasMol>Restrict \*A
- E. Predefined sets. The RasMol program has predefined sets which use single words to represent portions of a molecule. Examples of predefined sets that may be of interest to Science Olympiad Teams are: **Alpha** (amino acid alpha carbons), **Backbone** (amino acid atoms in the peptide backbone), **Charged** (amino acids containing a sidechain that is charged), **Hydrophobic** (amino acids containing a sidechain that is hydrophobic), **Hetero** (for heterologous, any atom that is in the pdb file that is not an amino acid), **Sheet** (the secondary protein structure, beta sheet), **Helix or Helices** (the secondary protein structure, alpha helix), **Sidechain** (the atoms contained in the amino acid “R” group or sidechain), **Water** (H<sub>2</sub>O)
- F. Center. When you restrict the RasMol display to only a portion of the protein, you can only see the “restricted” portion. However, the remaining portion of the molecule is still present, but not shown. If you use the mouse control to move the image, the image will rotate around the center of the entire molecule and this can be awkward if you are only displaying a small part of a larger molecule. You can use the “center” command to change the center point of the displayed molecule.

To change the center of rotation of a molecule to the atoms in chain A, type the following in the command line window:

- RasMol>Center \*A
- G. SS bonds. The sidechain of the amino acid cysteine can form a covalent bond with a nearby cysteine sidechain to form a disulfide bond, also known as an SS bond. Frequently, disulfide bonds are important to stabilize protein structure. Using the command line window you can display disulfide bonds, change their thickness and color.
- RasMol>SSbonds 150
  - RasMol>Set ssbonds backbone
  - RasMol>Color ssbonds red
- H. Display sidechains. When proteins are displayed in the backbone format, none of the atoms of the amino acid sidechains are shown. Sometimes specific amino acid sidechains have an important role in the protein structure or function and should be included in the model. You may add a specific amino acid sidechain, ie: his 62, to the display with the following commands entered in the command line window:
- RasMol>Select his62 and sidechain
  - RasMol>wireframe 200
- I. Atom ID. You can use the mouse to identify an amino acid displayed in the molecule display window. Place the mouse cursor over the atom of interest and right-click the mouse button. The identity of the atom will be displayed in the command line window. A typical identity line will be:
- O 2165 Group: ASN 139 Chain: B
  - The atom is Oxygen with the assigned atom number of 2165 and the oxygen is part of the amino acid asparagine 139 on protein chain B.
- J. Slab. Sometimes it is useful to look at a molecule by “slicing” through the molecule. To move through a molecule type the following command in the command line window:
- RasMol>Slab 50

The Ctrl+Left mouse button will move the image through the molecule.

#### 6. Saving your work as a script file.

In the command window type

RasMol>SAVE SCRIPT FILENAME1.SPT

Make sure the script file is saved into your RasMol folder.

To open the script file, type:

RasMol>SCRIPT FILENAME1.SPT

This will save the file in the RasMol folder. A hard lesson to learn is that you must “save early, save often”. RasMol has very limited editing resources. Any change made to the display of the molecule is permanent. In this tutorial whenever you make a change to a

molecule you should type the SAVE SCRIPT command but change the number to the next consecutive number. In that way you will be saving previous versions of the script file that you could use. You can always delete the early versions of the script files when you are done with your project.

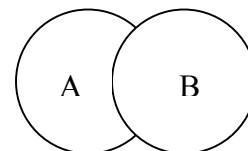
You can go to an earlier command with the UP ARROW and hit RETURN but it does the same as if you typed it in on the last command line. (Note also that you cannot CUT and PASTE).

7. Boolean operators. RasMol allows you to construct more complex commands out of simple commands by using the logical operators **AND**, **OR**, **NOT**.

The operator **OR** selects everything contained in both circles A and B.

The operator **AND** selects only the region of circles A and B that overlaps.

The operator **NOT** selects the regions of circle A that do not overlap with circle B (or vice versa). Another level of complexity can be added by grouping some of the commands in parentheses. Any commands grouped within a set of parentheses are executed first.



Example:

```
RasMol>SELECT (GLU6 OR HIS63 OR HIS92) AND (SIDECHAIN OR ALPHA)
```

This RasMol command selects both the sidechain and alpha carbons for each of the amino acids glu 6, his 63, and his 92.

10. Additional tutorials. Go to <http://www.umass.edu/microbio/rasmol/rastut.htm>  
For a RasMol manual go to <http://www.umass.edu/microbio/rasmol/distrib/rasman.htm>.