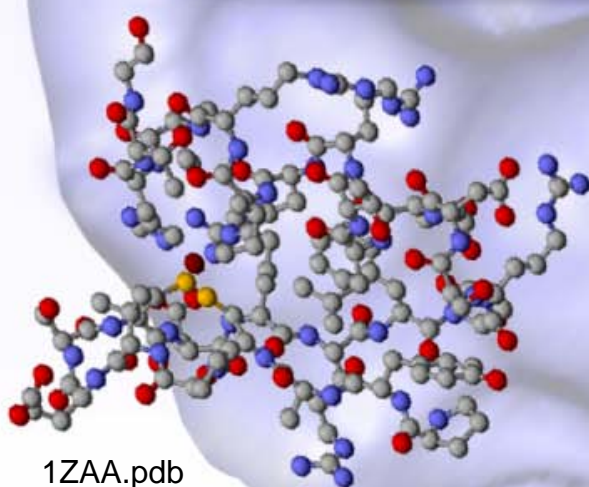
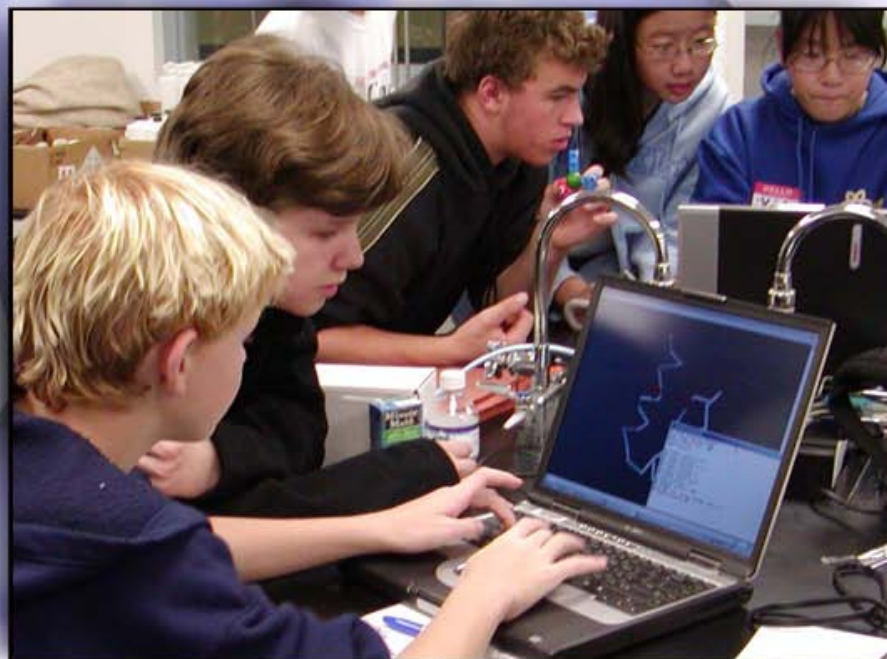
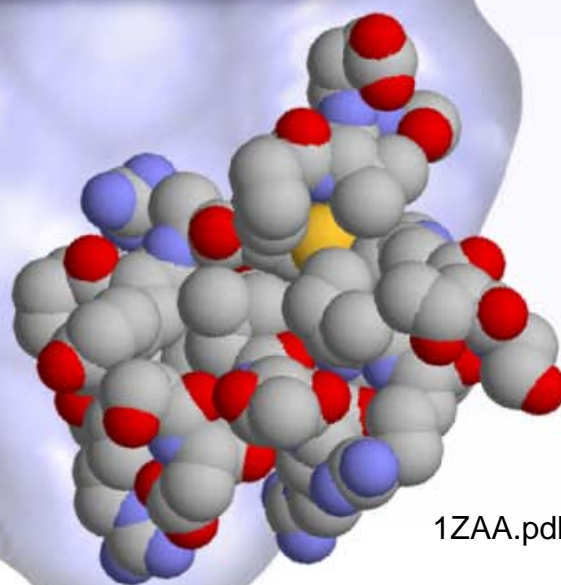


Jmol Training Guide

version 1.0



1ZAA.pdb



1ZAA.pdb



Center for
BioMolecular
Modeling

<http://cbm.msoe.edu>

Copyright 2010

Contents

Acknowledgements

Section 1: Using Jmol as a Computer Visualization Tool

- 1-1 - Downloading Jmol
 - 1-2 - Protein Data Bank (PDB) Files
 - 1-3 - Launching Jmol and Opening a PDB File
 - 1-4 - The Command Line and Common Display Formats and Display Colors
 - 1-5 - The Select Command
 - 1-6 - Boolean Operators
 - 1-7 - Exporting Images
 - 1-8 - Saving Your Work and Organizing Files
-

Section 2: Understanding the Protein Data Bank

- 2-1 - Using the RCSB Protein Data Bank
 - 2-2 - Searching the RCSB Protein Data Bank
 - 2-3 - Inside a PDB File
 - 2-4 - Referencing the RCSB Protein Data Bank
-

Section 3: Designing Models for Rapid Prototyping

- 3-1 - What is 3D Printing?
- 3-2 - Selecting Appropriate Display Formats and Colors
- 3-3 - Adding and Removing Hydrogen Bonds and Disulfide Bonds
- 3-4 - Adding and Removing Support Struts
- 3-5 - Coloring Hydrogen Bonds, SS Bonds, and Struts
- 3-6 - Adding Sidechains with a "Clean Backbone"
- 3-7 - Additional Miscellaneous Jmol Commands

Acknowledgments

The MSOE Center for BioMolecular Modeling would like to thank the countless teachers who have attended workshops over the years for their input in creating this document. Their inspiration and requests for something that would assist them in working with molecular visualization and Jmol has allowed the CBM to put this Training Guide together.

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

Section I – Using Jmol as a Computer Visualization Tool

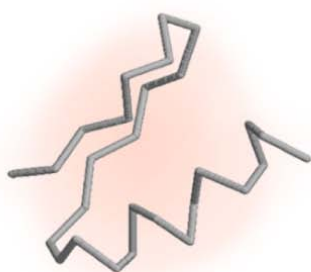
Jmol is a free open source molecular visualization program used by students, educators and researchers to explore protein structure. Section I of this Jmol Training Guide is designed to introduce you to this software. This program is a simple, yet powerful tool, which enables you to visualize a molecule in 3-dimensional space. As you actively manipulate the computer mouse to rotate the molecule, you will develop a sense of the 3-dimensional nature of a protein. Through the Jmol command line, you can also change the display format of the molecule to display different features of the molecule.

At the MSOE Center for BioMolecular Modeling (CBM), we believe that physical models of proteins are powerful tools that can be used synergistically with computer visualization tools to explore protein structure and function. Recent advances in 3-D printing technology make it possible to generate a physical model of any virtual image of a protein generated in the computer environment of Jmol. Therefore, as you learn to generate different images of proteins in Jmol, you will also be designing a physical model of that protein. Welcome to the amazing world of proteins --- and the fascinating stories that each one can tell!

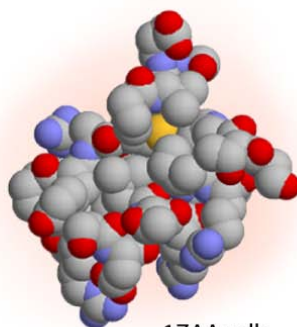
At the end of this section, you should be familiar with how to use Jmol as a computer visualization tool. We will introduce you to the basic features of Jmol, including:

Section 1: Using Jmol as a Computer Visualization Tool

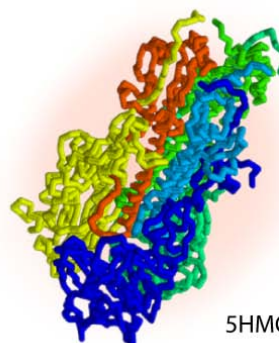
- 1-1 - Downloading Jmol
- 1-2 - Protein Data Bank (PDB) Files
- 1-3 - Launching Jmol and Opening a PDB File
- 1-4 - The Command Line and Common Display Formats and Display Colors
- 1-5 - The Select Command
- 1-6 - Boolean Operators
- 1-7 - Exporting Images
- 1-8 - Saving Your Work and Organizing Files



1ZAA.pdb



1ZAA.pdb



5HMG.pdb

1-1 - Downloading Jmol

- **About Jmol** – Jmol runs on a Java platform and therefore functions equally well in a PC or Mac environment. Because Jmol is freeware and an open source program, it can be obtained from many locations. For total functionality in designing a structure to be built on a 3-D printer, you must use the version 12.0.14 that is available on our website.
- **Downloading Jmol** - We suggest that you download Jmol directly from the CBM website to ensure you are using the correct version for creating physical models.
 - <http://cbm.msoc.edu/stupro/smart/resources.html>
- **Unzipping Jmol** - The program Jmol should download as one zipped folder that will need to be unzipped. Most computers running Microsoft Windows will automatically do this using the program WinZip. If your computer does not have WinZip, you may download the free unzipping software **CAM UnZip**, available at:
 - <http://www.camunzip.com/>
- **Unzipping for Mac** - If you are using a Mac that does not have an unzipping program, you may download the free unzipping software MacZip, available at:
 - http://download.cnet.com/MacZip/3000-2250_4-10025248.html
- **Jmol File Location** – The Jmol folder contains all the files your computer needs to run Jmol, and they need to be kept together in the same folder. You may save this folder in any convenient location on your computer. Jmol does not need to be "installed", so once these files are in place on your computer, the program will be ready to use.

★ When first launching Jmol, you will need to click the "Jmol.jar" file (this will be discussed in more detail in section 1-3). For convenience, it is a good idea to create a shortcut to this "Jmol.jar" file on your desktop as well as in each of your working folders so you can easily launch Jmol. In Windows, right click on the Jmol.jar file and select 'create shortcut'. Drag and drop the shortcut to your desktop.

★ On some PCs, the Jmol.jar file is seen by the computer as a zipped folder. If your computer shows the Jmol.jar file as a zipped file, you can open Jmol by clicking on the Jmol.bat file. Similarly, you can create a shortcut to the Jmol.bat file and save it on your desktop.

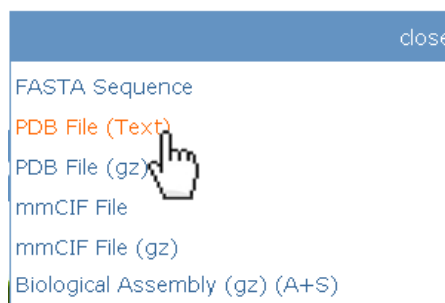
1-2 - Protein Data Bank (PDB) Files

- Molecular structures can be determined through the use of X-ray Crystallography, Nuclear Magnetic Resonance (NMR), or computer algorithm-based calculations.
- Once a structure has been determined, each atom in the structure is assigned an (X,Y,Z) coordinate to mark its location in 3-dimensional space. These coordinates are then stored in a file called a Protein Data Bank (PDB) file.
- Molecular visualization software, such as Jmol, can then use the coordinates stored in a PDB file to create an interactive 3-dimensional visualization of a molecular structure.
- **Downloading PDB Files** - The most common source of free PDB files of macromolecules is the RCSB Protein Data Bank website at <http://www.pdb.org>. For the first section of this Training Guide, we will be using Jmol to view the PDB file 1qys.pdb which contains the atomic coordinates for a single molecule of the protein "Top 7". To download this file from the Protein Data Bank:

- Go to the website <http://www.pdb.org>.
- In the top right corner of the website there is a search bar similar to the image below. Type in the four number/letter file name "1qys" and click the "Search" button.



- This should bring you to the page for "1qys.pdb - Top 7". Just below the search box on the right should be a list of four options similar to the image shown below left. Click "Download Files" and you will see an expanded menu similar to the image shown on the right below.



- Click "**PDB File (Text)**" to begin the download of the PDB file containing the coordinates for Top 7. This file, named "1qys.pdb", can be saved to the location of your choosing on your computer.

★ Note: It is a good idea to create a new folder for each molecule you work on to organize all PDB files, images, and other related work. (See section 1.8 for more information.)

- More detailed information about the Protein Data Bank and PDB files will be discussed within the Jmol Training Guide Part II.

1-3 - Launching Jmol and Opening a PDB File

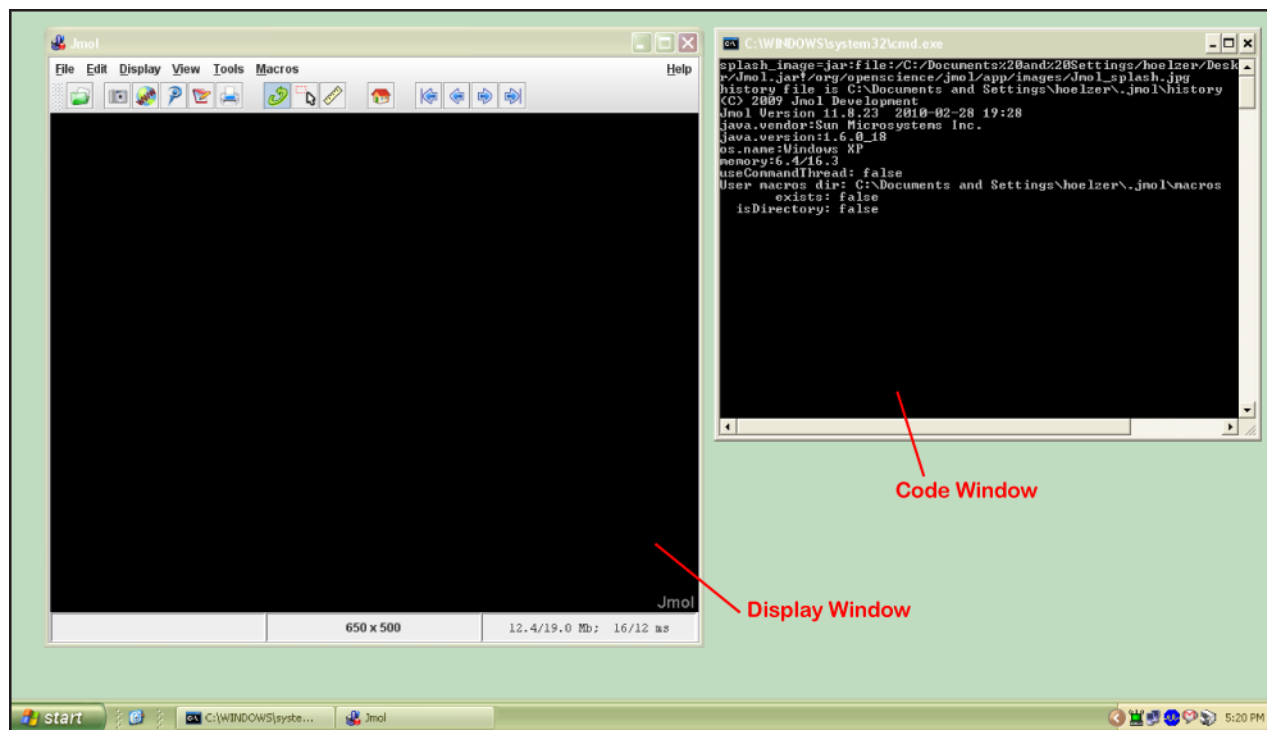
- **Launching Jmol** - To launch Jmol, open the Jmol folder you downloaded and unzipped earlier. In that folder, select the file "Jmol.jar". Note: it should have an icon similar to the image just to the right for PC and far to the right for Mac.



★ When first launching Jmol, you will need to click the "Jmol.jar" file. For convenience, it is a good idea to create a shortcut to this "Jmol.jar" file on your desktop as well as in each of your working folders so you can easily launch Jmol. In Windows, right click on the Jmol.jar file and select 'create shortcut'. Drag and drop the shortcut to your desktop.

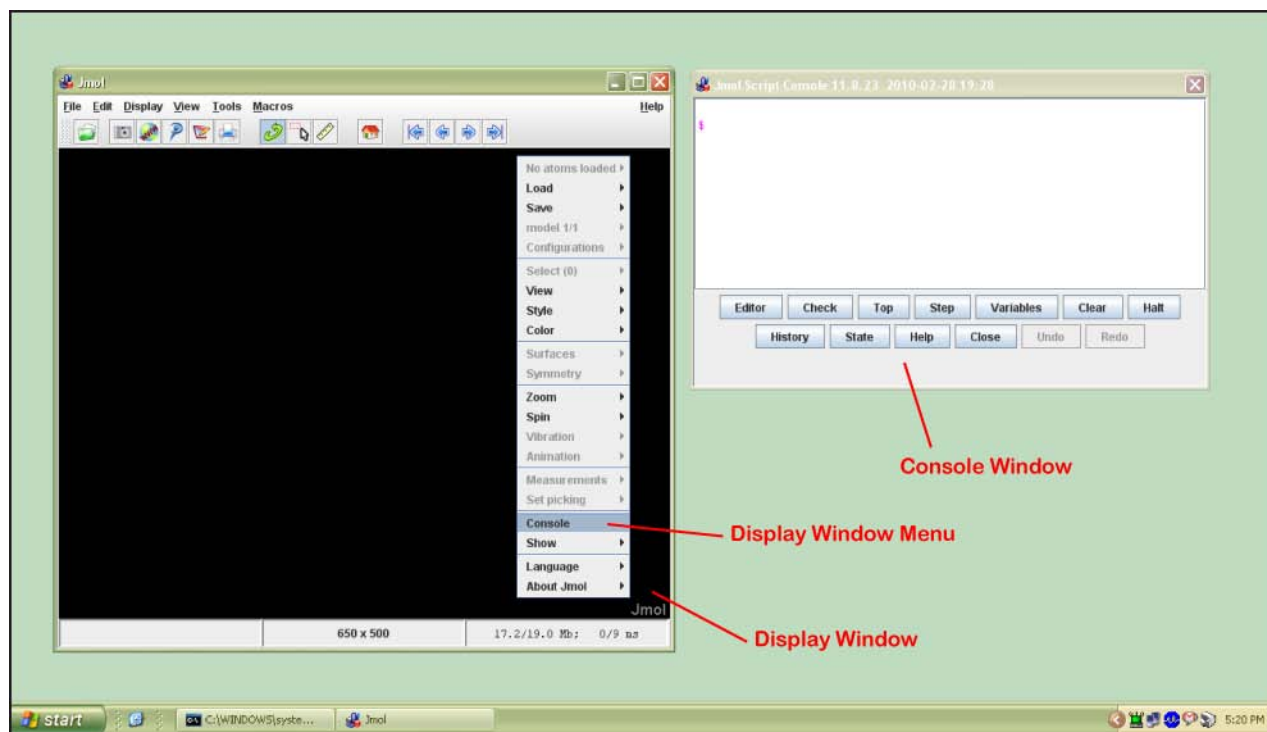
★ On some PCs, the Jmol.jar file is seen by the computer as a zipped folder. If your computer shows the Jmol.jar file as a zipped file, you can open Jmol by clicking on the Jmol.bat file. Similarly, you can create a shortcut to the Jmol.bat file and save it on your desktop.

- **Two Windows** - Once launched, Jmol will display two windows. The Display Window (below, left) is used to view your 3-dimensional images. The Code Window (below right) will not be used in this section of the Training Guide and can be minimized.



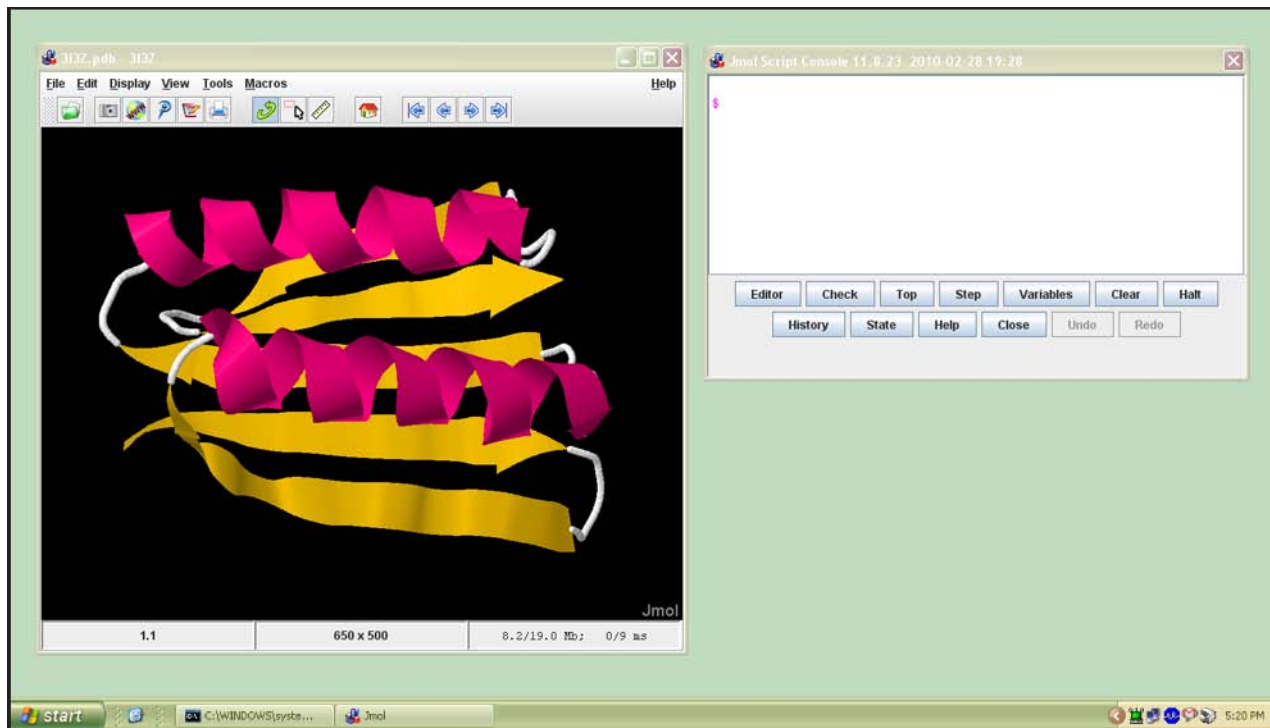
- **The Console** – Right click anywhere in the Display Window to bring up the Display Window Menu. In this menu, click on the button "Console" which will open a new window called the Console Window.

★ Because the Console and Display Windows will almost always be used in unison, arrange the windows so that they do not overlap and both fit comfortably on your screen.

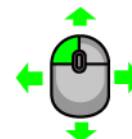


- **Opening a PDB File** - Click on the "File" button located at the top left of the Display Window. From the File Menu, click on the "Open" button.
 - The Jmol Open Menu will appear. For the remainder of this section, the Training Guide will use the protein Top 7 based on the PDB file 1qys.pdb that was downloaded in Section 1.2.
 - Locate the file 1qys.pdb, select it in the Jmol Open Menu, and click the "Open" button in the bottom right corner of the Jmol Open Menu.
 - The default Display Format for a protein opened in Jmol is cartoon format colored by secondary structure. Ligands are displayed with a spacefill and a wireframe while ions or individual water molecules are displayed in full spacefill format.

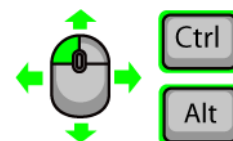
- **Interacting With The Display Window** - Your Jmol Display Window should now display the molecule similar to the image below. You can now rotate, relocate or zoom in on the molecule in the Display Window.



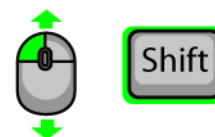
- **Rotating the Molecule** - To rotate the molecule within the X-Y axes on the computer screen, place the mouse cursor in the Display Window, hold down the left mouse button, and move the mouse.



- **Relocate the Molecule** - To relocate the molecule, place the mouse cursor in the Display Window, hold the left mouse button along with CTRL and ALT, and move the mouse.



- **Zoom** - To zoom in or out, place the mouse cursor in the Display Window, hold the left mouse button along with SHIFT, and move the mouse up or down. If the mouse you are using has a center rolling wheel, you can also zoom in and out using this wheel.



- **Barrel Roll** - To barrel roll (rotate the molecule on the Z axis), hold the left mouse button along with SHIFT as if you were going to zoom, but move your mouse left or right instead of up or down.



1-4 - The Command Line and Common Display Formats and Display Colors

- **The Command Line** - If you click in the white area of the Console Window, you will enter the Command Line. This area, highlighted by a magenta dollar sign (\$), is where text commands are entered. After typing a command and hitting enter, the 3-dimensional image of your molecule in the Display Window will change.

★ If you have typed a command incorrectly or the command is not recognized by Jmol, the text in the command line that is normally black will turn red.

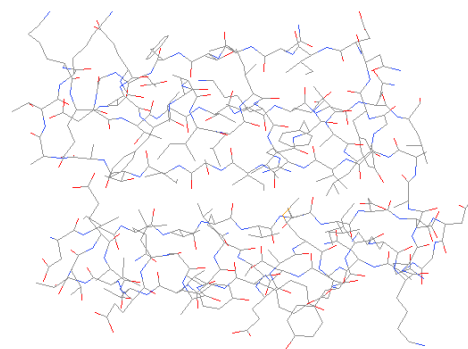
- **Different Display Formats** - The first text commands we will explore are the Display Formats. These key terms, when typed into the command line, will let you change the way your molecule is displayed. Below is a description of each of the common Display Formats along with an image of how the Top 7 from 1qys.pdb looks when this command is entered.

★ To reset the image so your display will match the images below, enter the following commands:

- `color cpk`
- `cartoon off`
- `spacefill off`

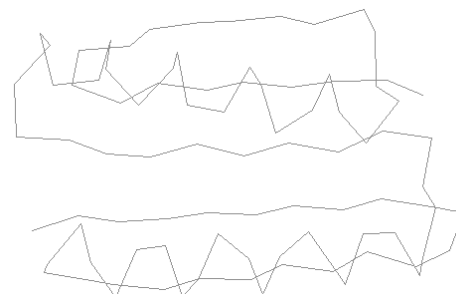
At this point, you should have a blank display window.

- **Wireframe** - The thin wire represents the bonds between atoms and the ends of the wires represent the locations of the atoms. The advantage of the wireframe is that all of the atoms are displayed. However, it is difficult to distinguish secondary structures in this format.

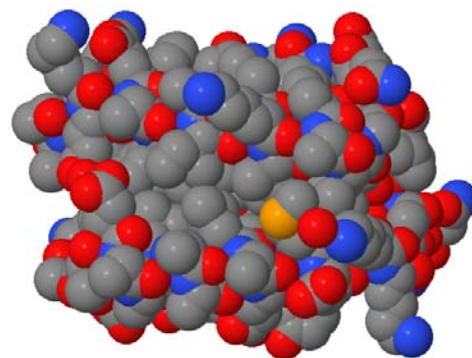


- To turn this format on, type `wireframe on`
- To turn this format off, type `wireframe off`

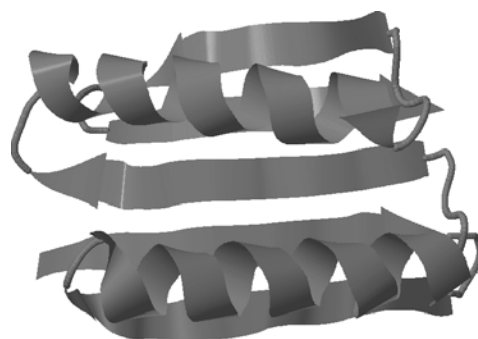
- **Backbone** - The alpha carbon backbone format only displays the position of the alpha carbon in each amino acid by a bend in the backbone. All of the other atoms within the amino acid are not displayed. The advantage of this display format is that it clearly illustrates the secondary structures within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



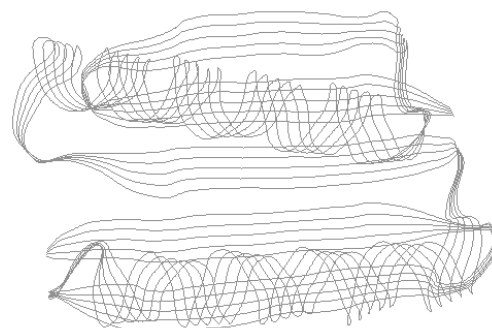
- To turn this format on, type **backbone on**
 - To turn this format off, type **backbone off**
- **Spacefill** - In this format, each atom is displayed as a sphere representing the volume that atom occupies in space. For example, hydrogen atoms are smaller than carbon atoms, which are smaller than sulfur atoms. This format best represents the volume and three-dimensional shape of a molecule. However, identifying the secondary structures or any internal atoms in this display format is challenging.



- To turn this format on, type **spacefill on**
 - To turn this format off, type **spacefill off**
- **Cartoon** - 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. The direction of the protein chain can also be identified by arrows at the ends of beta strands.



- To turn this format on, type **cartoon on**
 - To turn this format off, type **cartoon off**
- **Strands** - This format is similar to Cartoon, except that rather than displaying a solid ribbon, there are several "strands" or thin lines that make up the model.



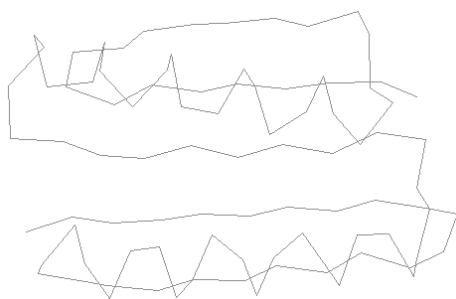
- To turn this format on, type **strands on**
- To turn this format off, type **strands off**

- **Advantages and Disadvantages** - 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. Different formats are used for different purposes. Which option you choose to display your molecule depends on the story you wish to tell. The advantage of using Jmol within the classroom is that it enables you to represent the model in a variety of ways, thus illustrating different aspects of the molecule. A summary of the advantages and disadvantages of each format is provided below.

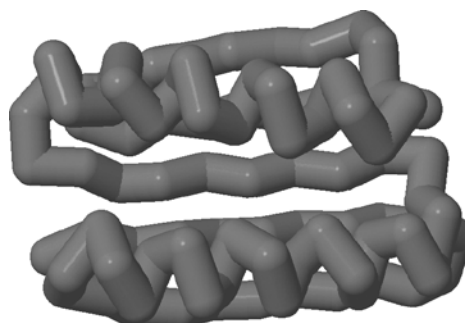
Format	Advantage	Disadvantage
Wireframe	Displays all atoms	Difficult to see secondary structures
Backbone	Secondary structures clearly seen	Detail of amino acids is lost
Spacefill	Represent 3D volume and shape	Cannot identify secondary structures
Ribbons and Strands	Secondary structures are easily visualized	No structural details

★ Note: Ribbons and strands are not used in designing models to be built using 3-D printing. Typically, backbone, wireframe and spacefill formats are used in combination to design a pleasing and informative model for 3-D printing.

- **Altering the Thickness of a Display Format** - Many of the display formats can be further altered by changing the thickness. To do this, simply add a number after any display format typed into the Command Line and then hit ENTER.



backbone on



backbone 1.5

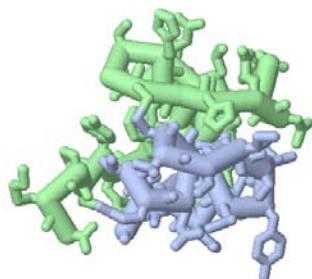
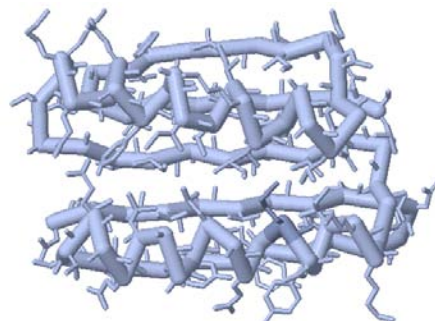
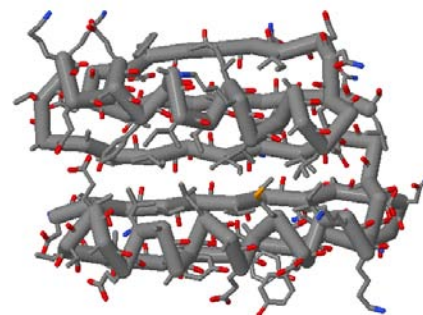
★ Note: The number added after the Display Format must include a decimal. For example, you will need to enter **backbone 1.5** and not simply **backbone 1**. This number represents the thickness you want in Ångströms (1×10^{-10} meters). So if you enter **backbone 1.0**, you will be showing a Display Format with a one Ångström diameter backbone.

★ Note: The standard scale for designing models for 3D printing is **backbone 1.5**.

- **The Different Display Colors** - Jmol has several common Display Color options to make your model look visually pleasing or to indicate specific features of the molecule. Below is a description of each of the Display Colors along with an image of how the Top 7 protein from 1qys.pdb will look with that color scheme.

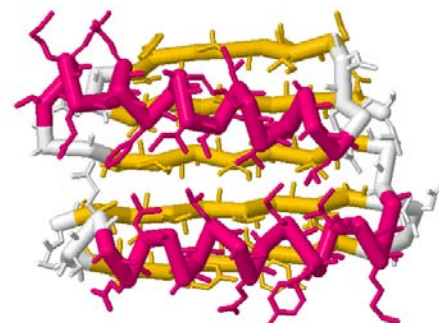
★ To reset the image so your display will match the images below, enter the following commands:

- **backbone 1.0**
- **wireframe 1.0**
- **CPK** - This is the default color scheme used in Jmol and was developed by **C**orey and **P**auling and later improved by **K**ultin. In CPK Color, carbon is gray, hydrogen is white, oxygen is red, nitrogen is blue and sulfur is yellow. This is the standard color code utilized in many textbooks.
 - To turn this color on, type **color CPK**
- **Chain** - If a PDB file has multiple chains (also known as polypeptides), this color option will color each chain a different color.
 - To turn this color on, type **color chain**



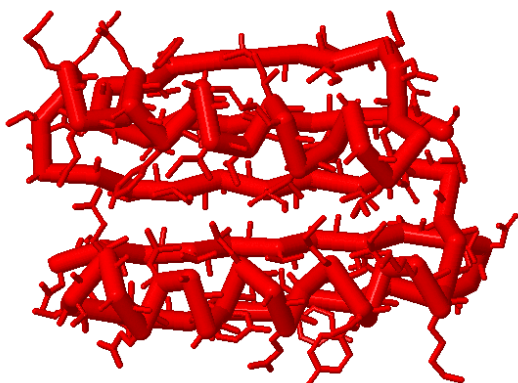
★ Note: The Top 7 protein has only one polypeptide chain and therefore is colored one color when the "color chain" command is used. Many proteins have more than one polypeptide chain and will therefore have more than one color. For example, a molecule of human insulin from the file 3i3z.pdb has two chains and therefore would have two colors.

- **Structure** - This option will color the secondary structures different colors. Magenta is used to highlight helices, yellow to indicate beta sheets, and white for non-structured regions.

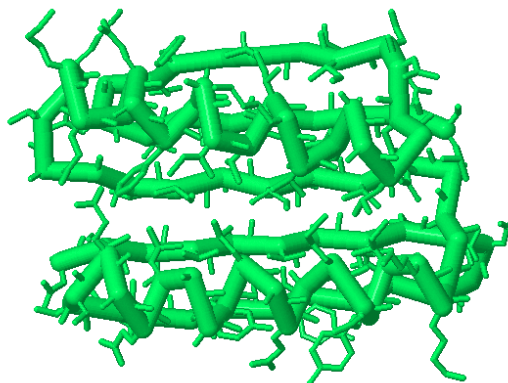


- To turn this color on, type `color structure`

- **Any Color** - This option allows you to color your molecule any solid color. Jmol recognizes a wide range of colors (for specific colors, visit: <http://Jmol.sourceforge.net/jscolors/>), or you can use RGB color codes entered between brackets and separated by commas [R,G,B] to specify a more specific color.

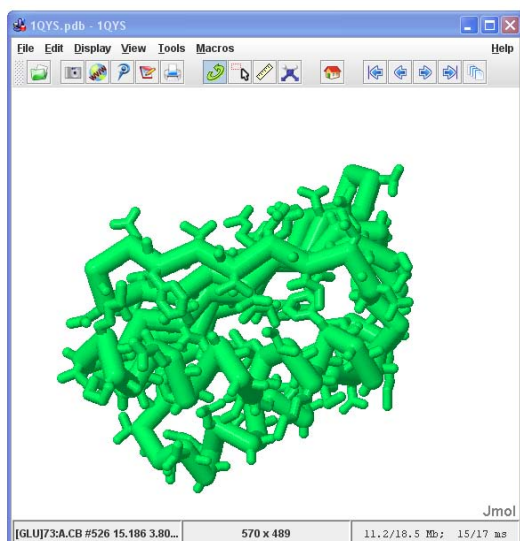


`color red`

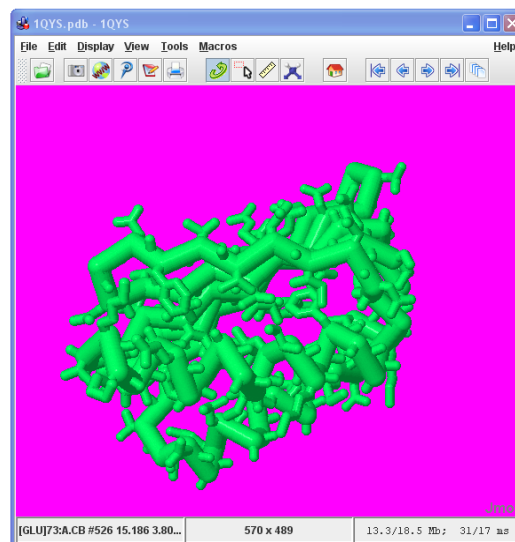


`color [0,255,100]`

- **Background Color of the Display Window** - The default color for the background of the Display Window will be black. This can be modified by using the `color background` command followed by the color you wish to change the background to.



color background white

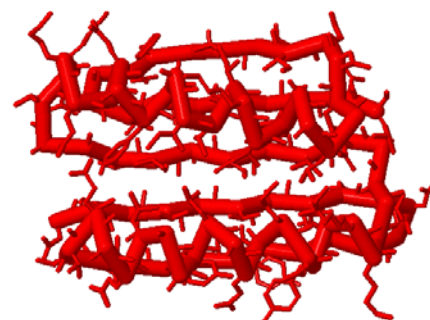


color background magenta

★ Note: We recommend that you avoid the red/green color combination in case you, or others, have problems with color blindness. Bright background colors from a computer monitor are also often hard on the eyes and therefore make it more difficult to focus on the protein. When creating an image for use on a poster, it is best to use a white background.

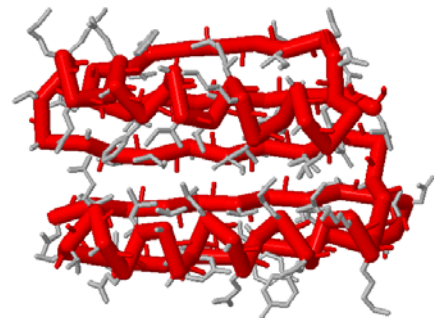
1-5 - The Select Command

- **The Select Command** - Thus far, every command we have typed into the Command Line has affected the entire protein. However, we can specify specific parts of a protein or molecule that we want to have future commands affect by using the Select command. Below is a description of each of the Select Commands along with an image of how the Top 7 from 1qys.pdb will look with the selected area colored red.
 - **All** - Select all is the default Jmol selection. With select all, every atom contained in a PDB file is selected and will be affected by future commands.
 - To select every atom, type `select all`
 - To color this selection green, type `color red`



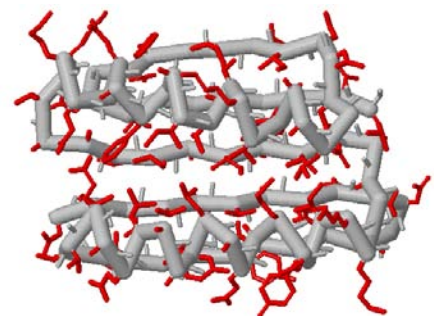
★ Note: The images of Top 7 shown below have been reset for each set of commands. To do this, use the `select all` command followed by `color gray`.

- **Backbone** - With select backbone, only the backbone atoms will be affected by future commands.
 - To select the backbone, type `select backbone`
 - To color this selection red, type `color red`

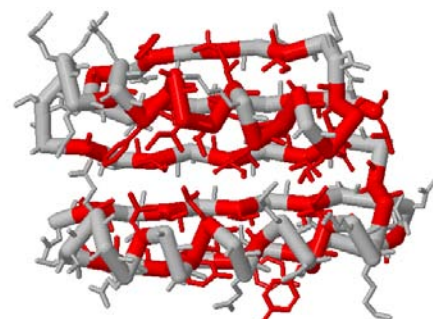


★ Note: The sidechains are not colored red.

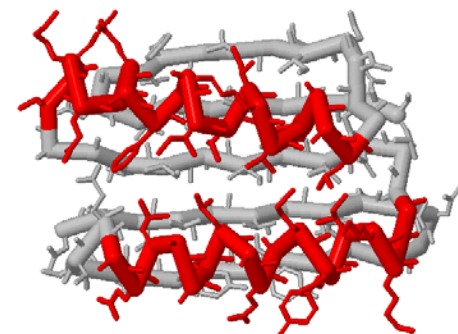
- **Sidechain**- With select sidechain, only the atoms that are part of sidechains in a protein will be affected by future commands.
 - To select the sidechains, type `select sidechain`
 - To color this selection red, type `color red`



- **Hydrophobic** - With select hydrophobic, only the hydrophobic amino acids in a protein will be affected by future commands.
 - To select hydrophobic, type `select hydrophobic`
 - To color this selection red, type `color red`



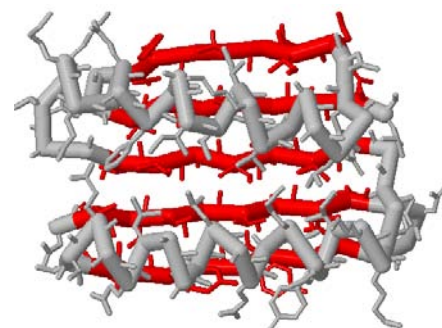
- **Helix** - With select helix, all atoms that are part of alpha helices will be affected by future commands.
 - To select helices, type `select helix`
 - To color this selection red, type `color red`



★ Note: Jmol does not recognize "helices" (plural) but only "helix" (singular).

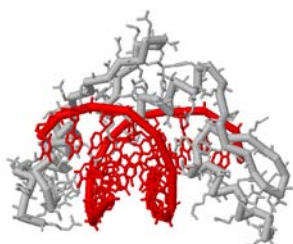
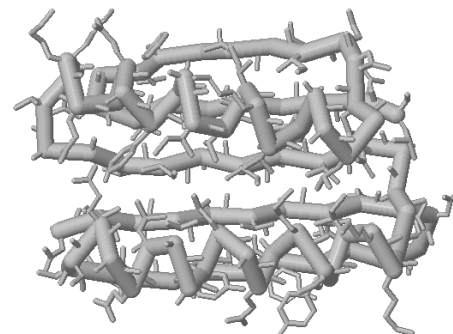
- **Sheet** - With select sheet, all atoms that are part of beta pleated sheets will be affected by future commands.

- To select beta pleated sheets, type `select sheet`
- To color this selection red, type `color red`



- **Nucleic**- With select nucleic, all atoms that are part of nucleic acids (DNA, RNA) will be affected by future commands.

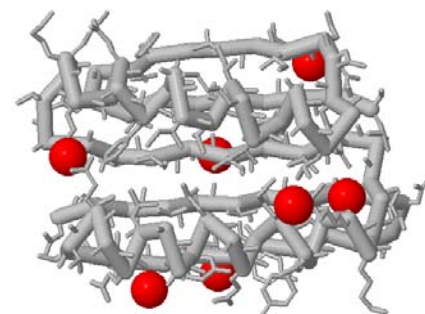
- To select nucleic acids, type `select nucleic`
- To color this selection red, type `color red`



★ Note: The Top 7 protein does not contain any nucleic acids and therefore nothing is colored when the "select nucleic" command is used. A zinc finger protein based on the file 1zaa.pdb does contain nucleic acids and therefore will have areas colored red.

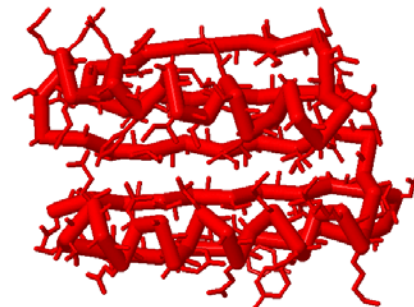
- **Water**- With select water, all water molecules will be affected by future commands.

- To select water molecules, type `select water`
- To display water molecules, type `spacefill`
- To color this selection red, type `color red`

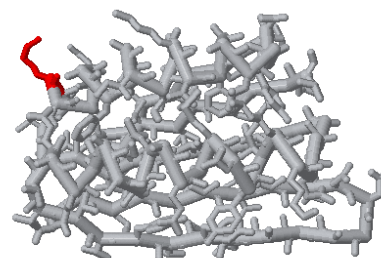


★ Note: Hydrogen atoms are not displayed in pdb files that are obtained via X-ray crystallography. Therefore water molecules appear as a single oxygen atom. Since the oxygen is not visually bonded to another atom, the only way water molecules can be displayed is in spacefill format.

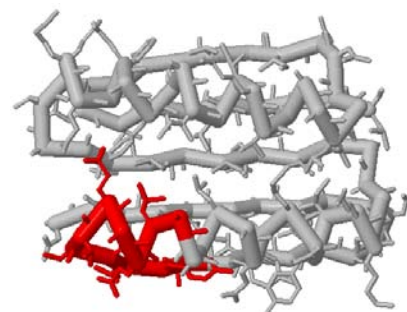
- **Chain-** With select chain * followed by the letter of the chain you would like to select, only the chain you select will be affected by future commands.
 - To select chain A, type `select *a`
 - To color this selection red, type `color red`



- **Residue Number-** With the Residue Number select command, you can select a specific amino acid or a range of amino acids that all future commands will affect.
 - To select residue number 58, type `select 58`
 - To color this selection red, type `color red`

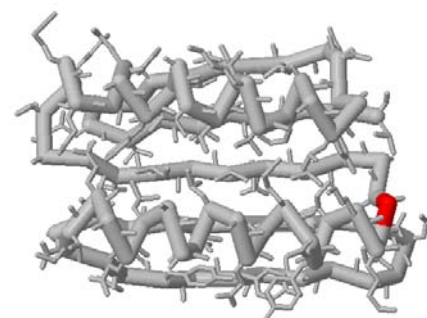


- To select residues 20 through 30, type `select 10-20`
- To color this selection red, type `color red`

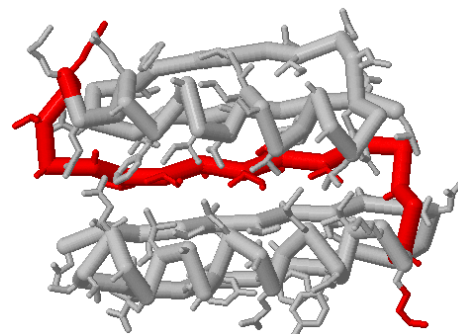


★ Note: You can also select an amino acid by its three letter abbreviation followed by the Residue Number. For example, instead of `select 58` as shown above, you could also use `select lys58` because the amino acid at Residue Number 58 in the Top 7 protein is a Lysine.

- **Atom Number-** With the Atom Number select command, you can select a specific atom number or range of atoms that all future commands will affect.
 - To select atom number 317, type `select atomno=317`
 - To color this selection green, type `color red`



- **Atom Number Range**- With an alteration to the Atom Number select command, you can select a range of atoms that all future commands will affect.
 - To select atom number 303-410, type `select atomno>= 303 and atomno<= 410`
 - To color this selection green, type `color red`



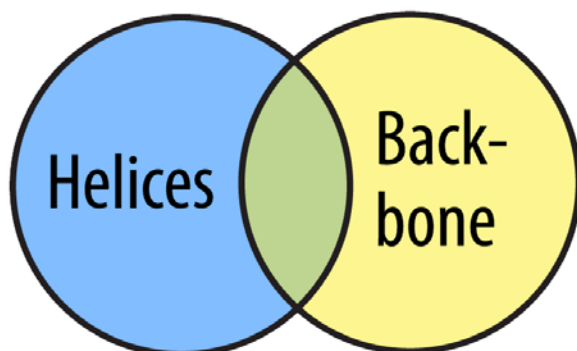
★ Note: If you do not know the Residue Number, Chain, Amino Acid Type, or the Atom Number you want to select, you can click on the residue in the Display Window. Clicking on an atom will result in information about that atom appearing in the Console Window. The image below explains this information in more detail:

[TYR]14:A:CA#92 - 3.55 - 21.94 - 5.889

- [TYR]**: This is the three letter abbreviation for the amino acid residue clicked - in this case, TYR for Tyrosine.
- 14**: This is the residue number that will be used with the select command - in this case, 14.
- A**: This letter identifies what chain the clicked residue is on - in this case, A.
- CA**: These letters identify what atom type has been clicked - in this case, CA for "Carbon Alpha".
- #92**: This number identifies the atom number that has been clicked - in this case, 92.
- 3.55 - 21.94 - 5.889**: These remaining three numbers identify the X,Y,Z coordinates for the atom that has been clicked.

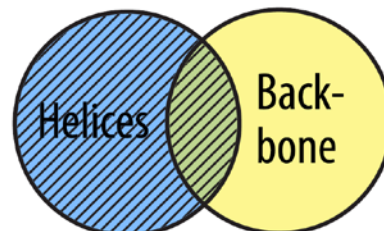
1-6 - Boolean Operators

- Once you are comfortable using the Select Command, you can begin to link together selections by using Boolean Operators. Imagine you are working with two predefined sets, for example, **helix** and **backbone** which are illustrated in the Venn diagram below.



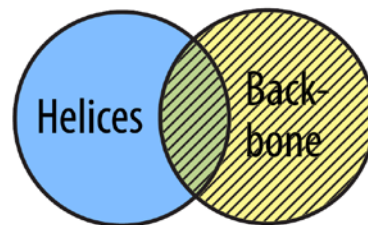
- You can select just the helices. This set includes all the alpha helices (including any backbone atoms that happen to be in alpha helices). The area with the hatched lines represents what is selected.

- `select helix`



- You can also select backbone. This set includes all the atoms that are part of the backbone (including any backbone atoms that also happen to be located in alpha helices).

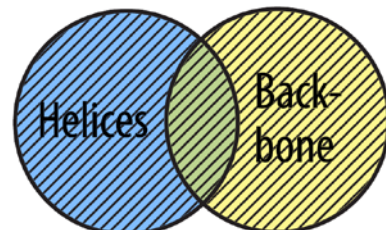
- `select backbone`



- You can use the **or** Boolean Operator to select all the amino acids that are either part of alpha helices or the atoms that are part of the backbone.

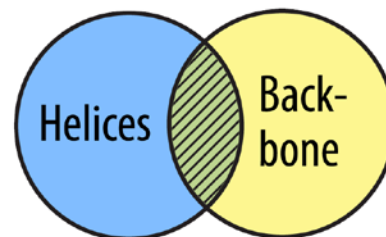
- `select helix or backbone`

- The phrase 'select helix or backbone' means 'select all atoms that are either a helix or in the backbone (or both)'.



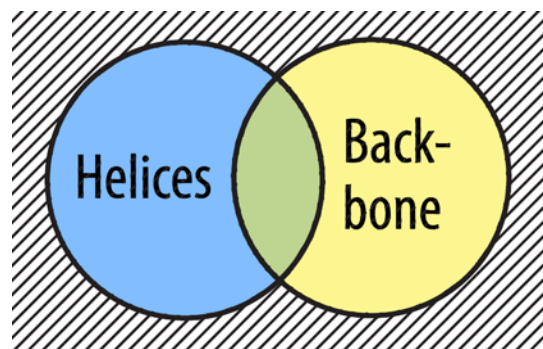
- You can use the **and** Boolean Operator to select all the amino acids that are both backbone amino acids as well as amino acids in alpha helices.

- select helix and backbone**
- The phrase 'select helix and backbone' means 'select all atoms that are both part of a helix AND part of the backbone'.



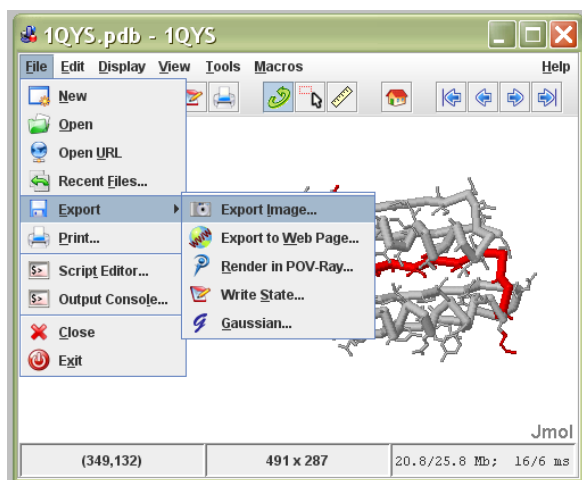
- You can use the **not** Boolean Operator to select only the atoms that are not part of helices and not part of the backbone.

- select not helices and not backbone**
- The phrase 'select not helices and not backbone' means 'select all atoms that are neither a helix nor part of the backbone'.



1-7 - Exporting Images

- Exporting an Image from Jmol** - If you want to create a JPEG image of your molecule from Jmol (to use in a poster, for example), click the "File" menu located on the top left of the Display Window. In the file menu, select "Export" and then "Export Image...".



it

★ Note: It is often a good idea to change your default black background to white to save cost in printing and for a better appearance.

★ Note: Your JPEG image will appear exactly as is displayed in the Jmol Display Window on your screen. The larger the size of your Display Window on the screen, the larger will be the dimensions of the JPEG that is exported.

1-8 - Saving Your Work and Organizing Your Files

- **Saving Your Work** - If you would like to save your work while working in Jmol, you simply export a JPEG file following the exact same directions shown in section 1-7. These JPEG files, even though they appear to be regular images, also contain information on what PDB file you have used and what commands you have entered. We recommend that you save your work frequently using sequential numbers
 - For example: Filename1.jpg
Filename2.jpg
Filename3.jpg

- **Organizing Files** - We recommend that you create a folder for each molecule you are designing. In this folder, place the following files:
 - Shortcut to Jmol
 - PDB file being used
 - All exported JPEG files

★ Note: We recommend the folder organization shown above because, in order for your saved work from your JPEG to load properly, you must have the PDB file and the JPEG file located in the same folder on your computer. When exporting a JPEG to save your work, save your file to this same folder.

- **Loading Your Saved Work** - To load your saved work from the JPEG that you have created, simply drag the JPG from the folder onto the Display Window. This will bring your saved work back up in Jmol.
- **Sending Your Saved Work** - When you are ready to send your saved model design for review or 3D Printing, attach the exported JPEG file to an e-mail. If the PDB file that was used to generate the model design and JPEG is not a standard PDB file available from www.pdb.org, the PDB file will need to be sent as an e-mail attachment as well.

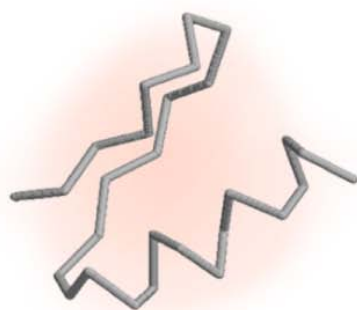
Section II – Understanding the Protein Data Bank

The focus of Section II of the MSOE Center for BioMolecular Modeling Jmol Training Guide is to learn about the Protein Data Bank, the worldwide repository for large crystal structure files. Section II will emphasize using Jmol in a more sophisticated manner. Mastering the material in this section will prepare you for both designing Rapid Prototyped Models and competing in the Protein Modeling Event developed by the CBM for Science Olympiad.

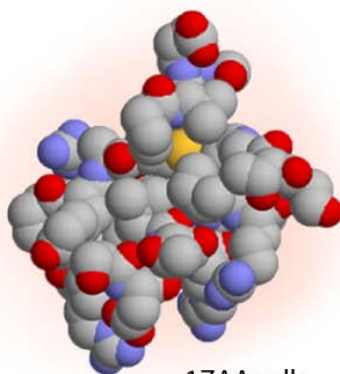
At the end of this section, you should be familiar with the RCSB Protein Data Bank and PDB files, including:

Section 2: Understanding the Protein Data Bank

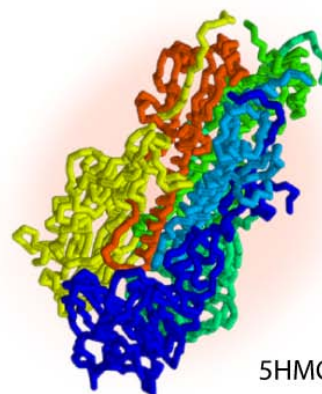
- 2-1 - Using the RCSB Protein Data Bank
- 2-2 - Searching the Protein Data Bank
- 2-3 - Inside a PDB File
- 2-4 - Referencing the RCSD Protein Data Bank



1ZAA.pdb



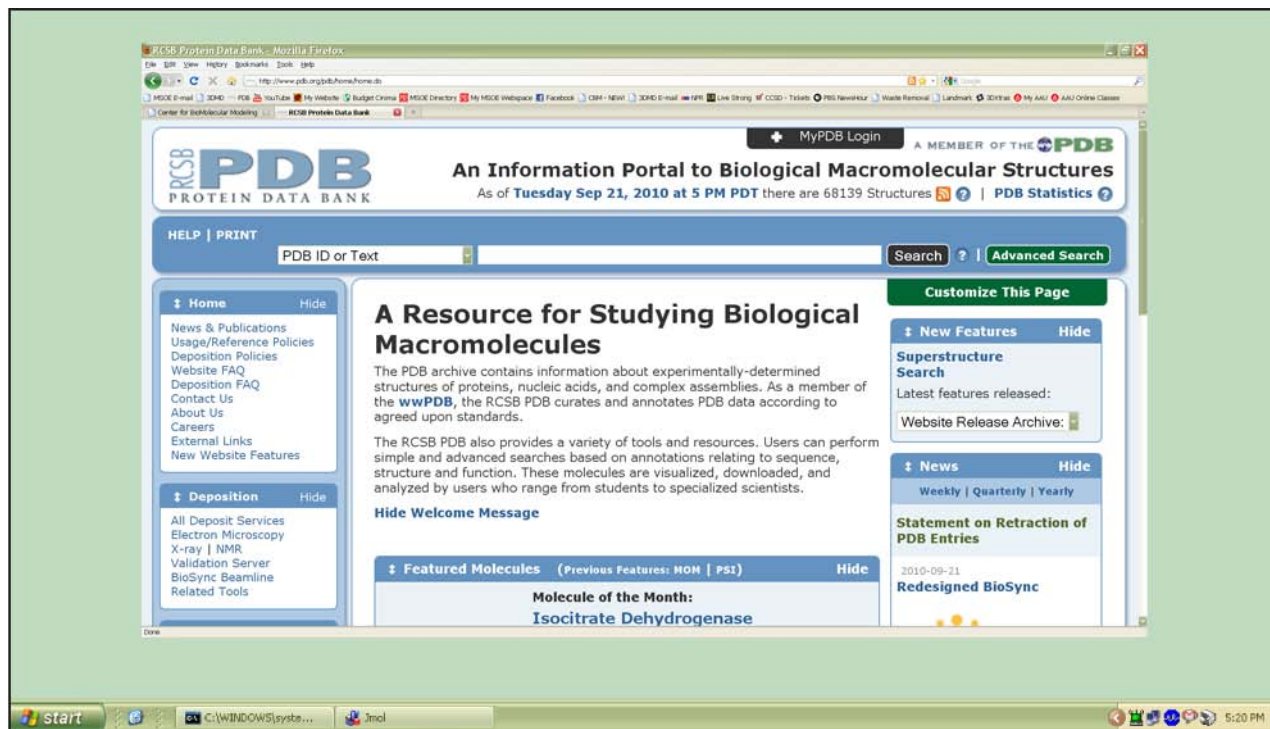
1ZAA.pdb



5HMG.pdb

2-1 - Using the RCSB Protein Data Bank

- The RCSB PDB (<http://www.pdb.org>) is the largest worldwide repository for the processing and distribution of 3-D structure data of large molecules of proteins and nucleic acids.



- PDB.org has several regularly updated features, as well as some interesting interviews and newsletters. All of these resources can be accessed from the PDB homepage.
 - Molecule of the Month** - David Goodsell explores the structure/function relationship of a new protein each month.
http://www.pdb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/current_month.html
 - PDB Newsletter** – This monthly newsletter includes updates from the PDB as well as articles about how people are utilizing the PDB for various projects. There have been several references to the Science Olympiad Protein Modeling Event in recent newsletters.
http://www.pdb.org/pdb/static.do?p=general_information/news_publications/index.html
 - Educational Resources** – Background information on using the PDB, as well as lesson plans, tutorials, and current events are posted here.
http://www.pdb.org/pdb/static.do?p=education_discussion/educational_resources/index.html

- **Structure Summary Page** - Every PDB file in the Protein Data Bank has its own page with information about that PDB file. These pages include:

The image shows a screenshot of the PDB Structure Summary Page for Human insulin (PDB ID: 3I1Z). The page is titled "Human insulin" and includes the following information:

- PDB Structure ID Number:** 3I1Z
- Title:** X-ray investigation of gene-engineered human insulin crystallized from a solution containing polyisialic acid.
- Authors:** Timofeev, V.L., Chapiro-Netchay, R.N., Samigina, V.R., Bezuglov, V.V., Miroshnikov, K.A., Kuronova, L.P.
- Primary Citation:** Journal: 2010 Acta Crystallogr. Sect.F 66: 259-263. PubMed: 20789155. PMID: 20789155. DOI: 10.1107/S174091110000461.
- Molecular Description:**
 - Classifications: Homology
 - Structure Weight: 5737.65
 - Molecule: Insulin A chain (Length: 21)
 - Polymer: 1 (Type: polypeptide(L))
 - Chain: A
 - Molecule: Insulin B chain (Length: 30)
 - Polymer: 2 (Type: polypeptide(L))
 - Chain: B
- Source of the Molecule:**
 - Source: Homo sapiens (Common Name: Human, Expression System: Escherichia coli)
 - Polymer: 1
 - Source: Homo sapiens (Common Name: Human, Expression System: Escherichia coli)
 - Polymer: 2
- Functional Class:** Insulin A chain, Insulin B chain
- Chemical Component:** Insulin A chain, Insulin B chain
- History:**
 - Deposition Summary: 2010-07-01
 - Release: 2010-01-05
 - Last Modified (REVDAT): 2010-05-05
- Method of Structure Determination:** X-RAY DIFFRACTION
- Resolution:** 1.50 Å

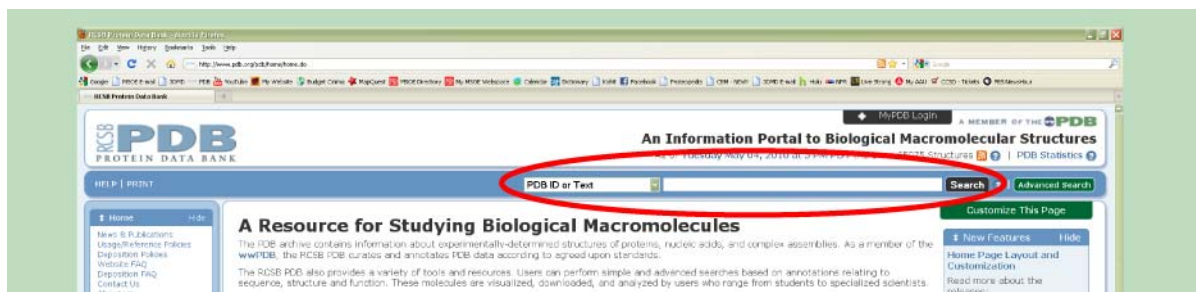
- **PDB Structure ID Number** - This 4 letter/number ID is a unique identifier that is assigned to the crystal data file upon deposition into the database.
- **Title** - Title of the PDB file
- **Authors** - These are the researchers who were involved with the crystallization of the molecule.
- ★ The senior author or principal investigator is usually the last author in science publications. The first author is the one who did the most research related to the article.
- **Primary Citation** - The journal article that accompanies the PDB file; excellent research resource for understanding the structure 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** – This indicates how accurately the position of each atom was determined. The smaller the number, the better the data.
- **Molecular Description** - This indicates the number of chains within the molecule and the chain identity; for example in the hemoglobin file (1A3N.pdb), chains A and C are alpha-globin molecules and chains B and D are beta-globin molecules.
- **Functional Class** - What type of molecule is it? (Ex: a toxin, an enzyme)
- **Source of the Molecule** - From which species was the molecule isolated? (Ex: human, bacterium, virus, mouse)
- **Chemical Component** - Heterologous groups that were co-crystallized with the molecule are listed here. 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 Jmol.
 - For example, if NAG (N-acetyl-glucosamine) is contained within the crystal, Jmol recognizes “NAG” for this file. If you enter the command “select NAG” Jmol would recognize the atoms within that chemical component of the PDB file.
- **Sequence Details** - Just above the Title there is a series of tabs - the third of which is "Sequence". This section of the PDB file page provides specific sequence information as well as secondary structure information for the molecule. You can identify the alpha helices or beta sheets as well as the amino/carboxy termini, which are the first and last amino acids of the protein, respectively.



2-2 - Searching the Protein Data Bank

- **PDB Search** - One of the key features of the RCSB Protein Data Bank is the ability to search the database for structures. You can search using key words or authors by entering these terms in the search box, located near the top center of every <http://www.pdb.org> page. 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.



- Often more than one PDB file exists for a specific type of protein - for example, there are over four hundred PDB file entries for "hemoglobin". It is often a good idea to use this structure summary page to answer the following questions about a searched PDB file so that you can select the best possible file.
 - 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?
 - What is the resolution of the structure?
- **Jmol Preview Window** - Every PDB file can be previewed using the online Jmol Preview Window. To view this preview, simply click the "View in Jmol" button that is located directly below the molecule image on each Structure Summary Page.



★ Note: You will not be able to design a model for printing using this preview version of Jmol.

2-3 - Inside a PDB File

- A PDB file is the file format for the X,Y,Z coordinates for each atom in the molecular structure. Molecular visualization programs like Jmol translate these X,Y,Z coordinates into

interactive 3-dimensional images, but you can also examine the file directly by opening it in any number of text editing programs such as Microsoft Word or WordPad.

- Most PDB Files will begin with a "Header", containing information about the structure, the lab that determined it, and the techniques used in the lab. Any comments like this will always start with the word "REMARK" and are ignored by Jmol. The information used to create the Structure Summary page of the RCSB PDB for each PDB file is included in the Header.
- The main portion of a PDB file will be the X,Y,Z coordinates. Although there are several similar versions of the nomenclature used to record these coordinates, the basic information remains the same. Below is a sample set of coordinates from a PDB file that has been opened with text editing software.

```
ATOM 1132 NH1 ARG A 149 31.814 -31.597 16.995
ATOM 1133 NH2 ARG A 149 32.203 -32.934 18.816
ATOM 1134 N ASN A 150 29.346 -24.359 18.812
ATOM 1135 CA ASN A 150 28.480 -23.190 18.933
ATOM 1136 C ASN A 150 28.606 -22.168 17.808
ATOM 1137 O ASN A 150 27.803 -21.276 17.678
ATOM 1138 CB ASN A 150 28.732 -22.524 20.282
ATOM 1139 CG ASN A 150 28.284 -23.389 21.447
ATOM 1140 OD1 ASN A 150 27.205 -23.981 21.430
ATOM 1141 ND2 ASN A 150 29.110 -23.463 22.466
ATOM 1142 N LEU A 151 29.629 -22.313 16.996
ATOM 1143 CA LEU A 151 29.868 -21.415 15.894
ATOM 1144 C LEU A 151 29.953 -22.205 14.597
ATOM 1145 O LEU A 151 30.149 -23.422 14.614
ATOM 1146 CB LEU A 151 31.208 -20.735 16.100
ATOM 1147 CG LEU A 151 31.436 -19.884 17.337
ATOM 1148 CD1 LEU A 151 32.846 -19.333 17.256
```



2-4 - Referencing the RCSB Protein Data Bank

PDB Search - When you use a PDB file to create an image or a model, you must cite the PDB file in your reference list. A PDB structure entry should be cited with its PDB ID and primary reference. For example:

PDB ID: 1QYS. Kuhlman, B., Dantas, G. Ireton, G.C., Varani, G., Stoddard, B.C., Baker, D. (2003). Design of a Novel Globular Protein Fold with Atomic Level Accuracy. *Science* 302: 1364-1368.

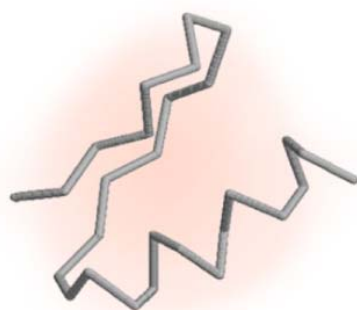
Section III - Designing Models for 3D Printing

In this section of the Jmol Training Guide, you will become familiar with the commands needed to design a model that will be built on a 3D Printer. As you become more comfortable using Jmol, this section will enable you to take the next step and design physical protein models.

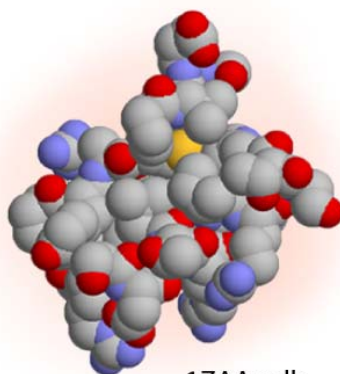
At the end of this section, you should be familiar with how to use Jmol to create a physical protein model on a rapid prototyping machine, including:

Section 3: Designing Models for 3D Printing

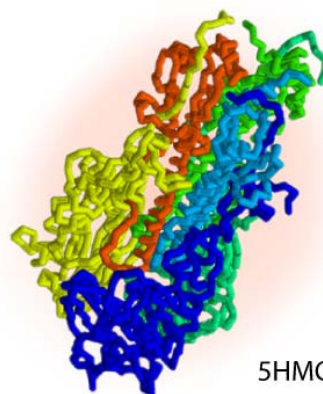
- 3-1 - What is 3D Printing?
- 3-2 - Selecting Appropriate Display Formats and Colors
- 3-3 - Adding and Removing Hydrogen Bonds and Disulfide Bonds
- 3-4 - Adding and Removing Support Struts
- 3-5 - Coloring Hydrogen Bonds, SS Bonds, and Struts
- 3-6 - Adding Sidechains with a "Clean Backbone"
- 3-7 - Additional Miscellaneous Jmol Commands



1ZAA.pdb



1ZAA.pdb



5HMG.pdb

3-1 - What is 3D Printing?

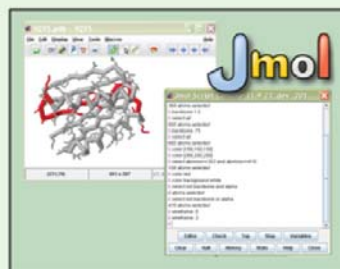
- 3D Printing is a cutting-edge technology that can use stacked serial planes to build physical 3-dimensional models from any computer generated 3D file. Every physical protein model design goes through four key steps:

1



X,Y,Z coordinates in the form of a PDB file are the starting point of every protein model.

2



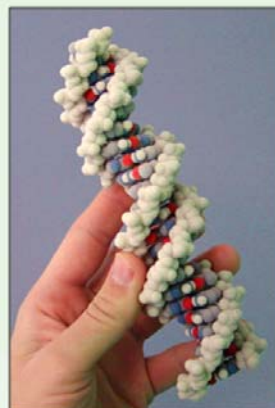
Design software is then used to edit the features and format of the protein model

3



The model is then exported as a file that is recognized by a 3D printer

4

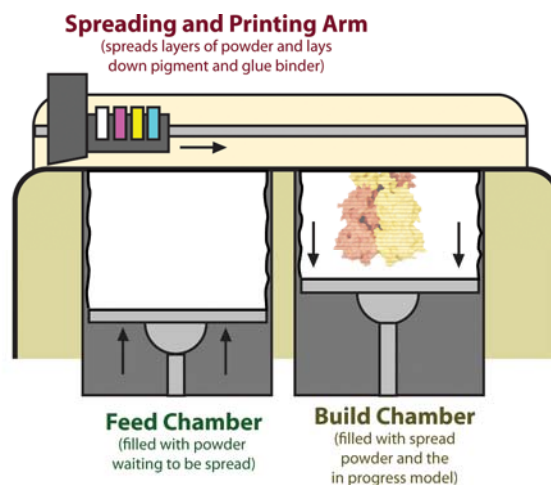
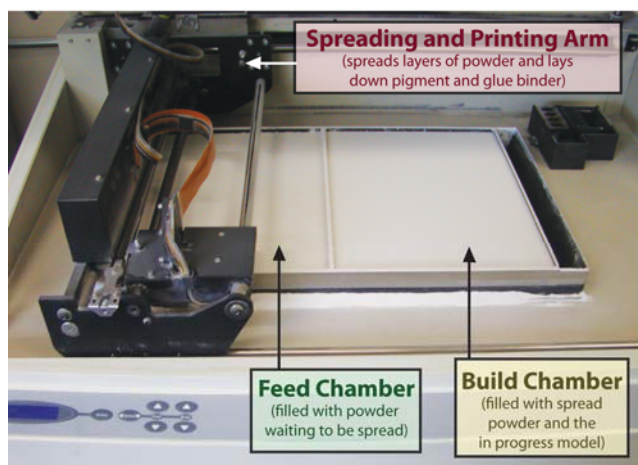


Finally, the physical model is removed from the printer and completed in a post production process.

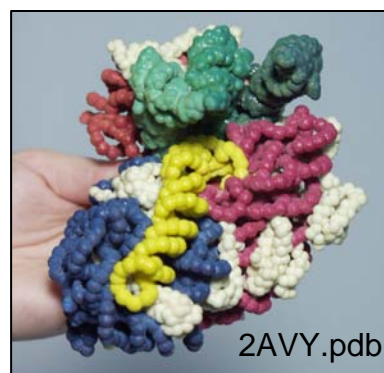
- Z-Corporation 510 and 650** - Although many varieties of 3D Printing and Rapid Prototyping machines exist, the MSOE Center for BioMolecular Modeling relies mainly on two Z-Corporation machines - the Z510 and the Z650.



- **How a 3D Printer Works** - The inside of a 3D Printer has three main parts: the feed chamber that stores the plaster powder to be used, the build chamber where the printed model will be built, and the spreading and printing arm which applies a glue and pigment solution to the growing model in the build chamber. While printing, the spreading and printing arm moves from left to right pushing a thin layer of plaster powder into the build chamber. The arm then makes a second pass printing a glue pigment solution onto the new layer in the build chamber. The build chamber is then lowered as the feed chamber is raised so a new layer of plaster can be spread. This process continues until all layers of the model have been completed.



- **The End Result** - A designed and printed protein model will be an exact replica of the model created in the 3D visualization program.

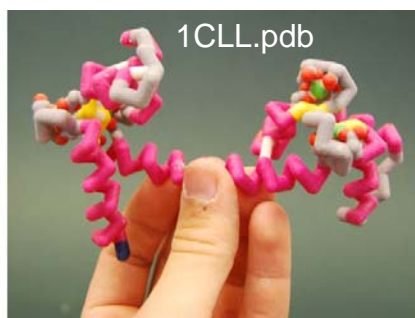


3-2 - Selecting Appropriate Display Formats and Colors

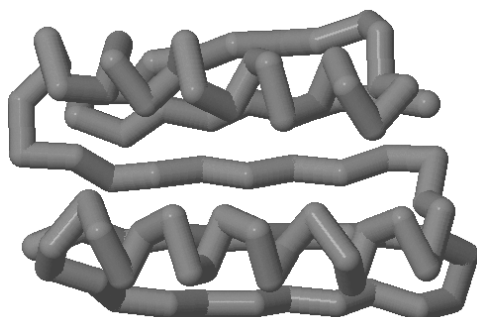
- **What Display Should I Use?** - This is a common question asked by model designers. And the answer to this question is another question: "What is the story that you are telling with your model?"
 - If your story is focused on a particular active site within the molecule, then perhaps the Alpha Carbon Backbone Format displaying key active site amino acid sidechains in Ball and Stick (Wireframe and Spacefill) Format would be best. 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 will assist you in deciding which format is the best for telling your story.

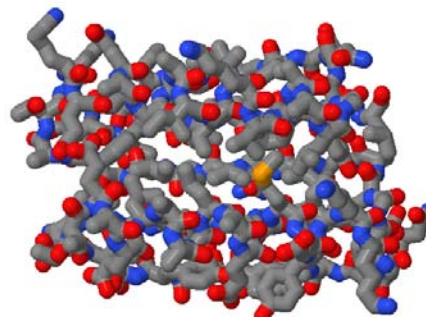
- The important point to remember is that no one model will tell every aspect of the story. Using Jmol in combination with a physical model will assist you in telling multiple aspects of your story.
- The CBM will build models in the Spacefill Format and the Alpha Carbon Backbone Format. Models in the Alpha Carbon Backbone Format can have sidechains and heterologous groups displayed. We currently do not build models in Ribbons or Strands Format.



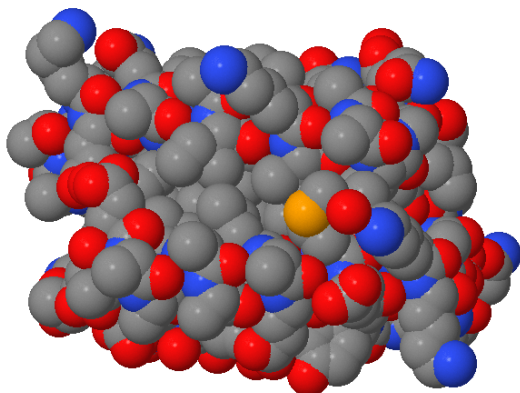
- **Format Sizes** - Every display format in Jmol is able to be displayed with various thicknesses (see section 1.4 for more information on Display Formats). If a display format is too thin, a protein model will be more likely to break during the 3D printing process. If a format is too thick, the model will use excessive amounts of material and details will be less visible. Below is a list of suggested sizes for the different formats:



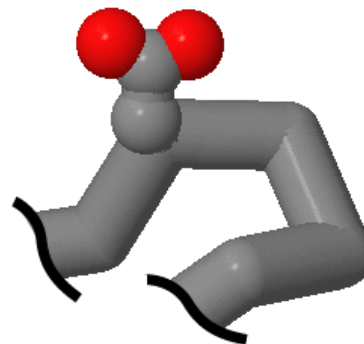
Backbone 1.5



Wireframe 1.25



Spacefill



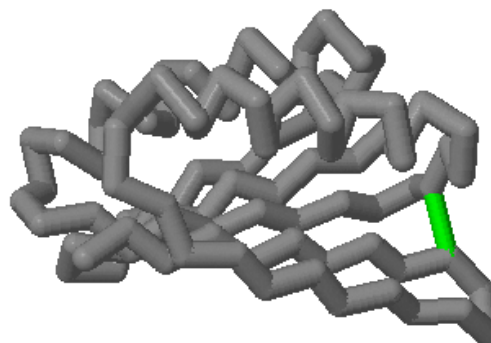
(for sidechains only)

★ Note: The default spacefill does not include a number. This is known as Full Spacefill.

Wireframe 1.25
Spacefill 1.5

For all bonds and struts

hbonds 1.0
ssbonds 1.0
struts 1.0



- **Which Colors Should I Use?** - Which colors you use, like which Display Format you choose, is largely influenced by the story you would like your model to show.
 - If the overall shape and fold of your protein is important to your story, you may want to color the secondary structures (alpha-helices and beta-sheets) a bright eye-catching color.
 - If there is a specific sidechain or active site area that is important to the function of your protein, you may want to highlight that area with a unique or bright color.
 - You can color a protein model in an almost unlimited variety of ways so it is very important to consider what you want to show and what colors would be best to highlight the key points to your protein's story.

- **Colors Not to Use** - There are a few colors or color combinations that you should avoid while designing a protein model to be build using 3D printing technologies.

- Do not use any black or extremely dark colors. The 3D printers do not print well with dark colors and often leave stripes or unevenness in the final physical model.



- Do not use colors that are too similar to highlight different features. These will often not be distinguishable in the final physical model.



- Do not use too many extremely bright colors together on the same model. When used rarely, very bright colors are extremely good at drawing the eye to important areas of the protein model. Overuse of bright colors can distract from the model.

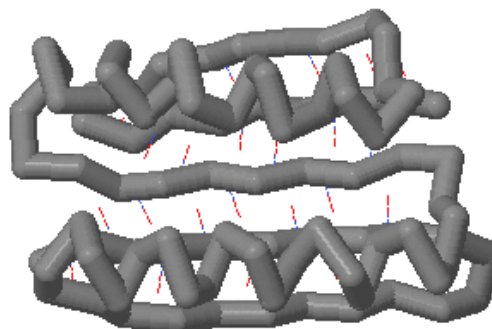


3.3 - Adding and Removing Hydrogen Bonds and Disulfide Bonds

- **Hydrogen Bonds** - Hydrogen bonds (referred to as "hbonds" in Jmol) are essential to the stability of secondary structures in a protein. They form between the backbone oxygen of one amino acid and the backbone nitrogen in a second amino acid. To add hydrogen bonds to secondary structures within your model, use the `calculate hbonds` command. We typically do not add hydrogen bonds to alpha-helices, since they do not add stability to the model and actually clutter the view of the structure. Adding hydrogen bonds to beta-sheets provides additional support for the final model and is recommended. Use the following commands:

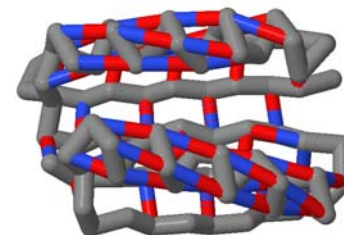
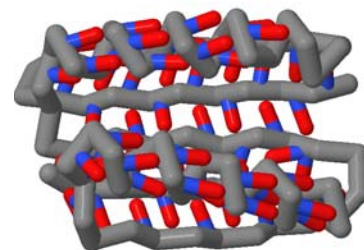
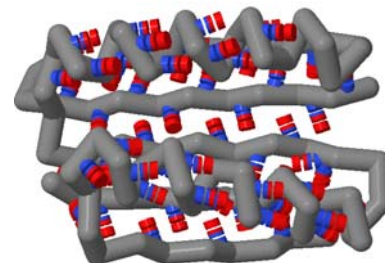
- `select sheet`
- `calculate hbonds`

- After you have entered this command, you will notice that there are dotted lines that have appeared.



These are the hydrogen bonds.

- The default size when you first add hydrogen bonds is the same as for wireframe which is very thin.
- **Thickness** - Hydrogen bonds, like wireframe, backbone, and spacefill, can be thickened by placing a number after the `hbonds` command.
 - `hbonds 1.0`
- **Solid hbonds** - The default display for hydrogen bonds is a dashed line. You will need to change this into a solid cylinder for building a physical protein model using 3D printing. You can do this using the command:
 - `set hbonds solid`
- **Setting hbonds to the Backbone** - Notice that the hydrogen bonds are now thicker and solid, 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 in Backbone Format, it appears as if the hydrogen bonds are floating in space. Therefore, we must set the hydrogen bonds to the backbone using the command:
 - `set hbonds backbone`
- **Removing all hbonds** - To turn off all the hydrogen bonds in a selected area, use the `hbonds off` command.
 - `hbonds off`
- **Summary of Commands to Add hbonds:**
 - `select sheet`
 - `calculate hbonds`
 - `hbonds 1.0`
 - `set hbonds solid`
 - `set hbonds backbone`
- **Adding or Removing Individual hbonds** - At times you may want to add or remove a single hydrogen bond. First, you need to know the two **amino acid residue numbers** that the hydrogen bond connects. To do this, click on the two residues in the Display Window (see section 1.5 for more information on identifying atom numbers). Once you know the two



residue numbers, you must select only these two amino acids. Finally, use the `hbonds 1.0` command to add an hbond or the `hbonds off` command to remove the bond between the two selected amino acids.

- For example: `select 167 or 73`
`hbonds 1.0`

For example: `select 321 or 334`
`hbonds off`

- **Disulfide Bonds** - Some molecules will have disulfide bonds (referred to as "ssbonds" in Jmol) present within the structure. These bonds form between two cysteine amino acids that lie close to each other in 3D space. Disulfide bonds are added using the command `ssbonds`.

- `ssbonds`

- **Thickness** - As we saw with the hydrogen bonds, simply typing `ssbonds` by itself will only produce thin lines. To give these bonds dimension, we must add a thickness (number) to the ssbonds.

- `ssbonds 1.0`

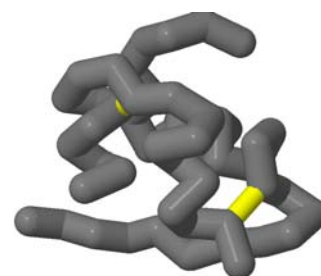
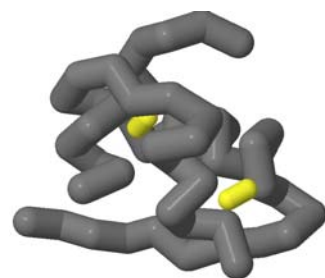
- **Setting ssbonds to the Backbone** - Notice that this command gives the disulfide bonds a thicker dimension, but as we saw with the hydrogen bond, the bonds are "floating" in space. This is because the disulfide bond is actually between the sulfur groups of the cysteine sidechains, and not the alpha 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).

- `set ssbonds backbone`

★ Note: You may wish to display the sidechains of the cysteines involved in the disulfide bond. If that is the case, then you will not need to set ssbonds to the backbone.

★ Note: Not every PDB file has ssbonds. The Top 7 protein PDB file used in Section 1 does not have a disulfide bond. To practice adding disulfide bonds, try the PDB file 2hiu.pdb which contains the coordinates for one molecule of human insulin.

- **Adding or Removing Individual ssbonds** - At times you may want to add or remove a single disulfide bond. To do this, you first need to know the two **amino acid residue numbers** that the bond connects. To do this, click on the two residues in the Display Window (see section 1.5 for more information on identifying atom numbers). Once you know the two



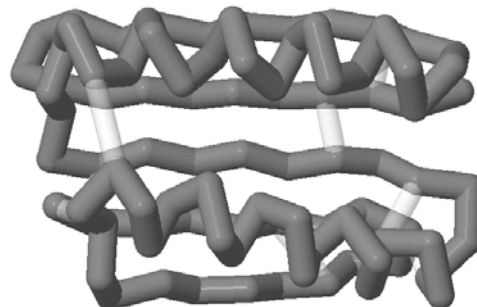
residue numbers, you must select only these two amino acids. Finally, use the `ssbonds 1.0` command to add an ssbond or the `ssbonds off` command to remove the bond between the two selected amino acids.

- For example: `select 167 or 73`
`ssbonds 1.0`

For example: `select 321 or 334`
`ssbonds off`

3.4 - Adding and Removing Support Struts

- **Support Struts** - When a model is built on a 3D printer, it needs to be supported in order to withstand the additional pressure as powder builds up and as the model is removed and cured. Hence, support struts are added within the model for support. To do this, select your entire model and use the calculate struts command.



- `Calculate struts`

- **Thickness** - Struts, like hydrogen bonds and ssbonds, can be set to different thicknesses by adding a number after the strut command.

- `Strut 1.0`

- **Adding or Removing Individual Struts** - At times you may want to remove a single strut (for example, if it goes through a sidechain you have represented). To do this, you first need to know the **atom numbers** of the two atoms the strut connects. To do this, click on the two atoms in the Display Window (see section 1.5 for more information on identifying atom numbers). Once you know the two atom numbers, you must select only those two atoms using the `select atomno=` command. Finally, use the `struts 1.0` command to add a strut or the `strut off` command to remove the bond between the two selected amino acids.

- For example: `select atomno=167 or atomno=73`
`strut 1.0`

- For example: `select atomno=321 or atomno=334`
`strut off`

★ Note: The calculate strut command may not always provide enough structural support. Additional struts may be necessary. Struts will need to be added manually to attach any ligands (substrate, inhibitor, cofactors) to the protein.

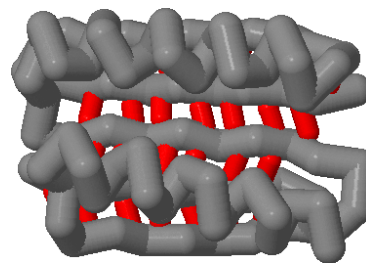
3.5 - Coloring Hydrogen Bonds, SS Bonds, and Struts

- **Default Bond Colors** - When bonds or struts initially appear, they will be the default Jmol colors (a translucent white for struts and the colors of the atoms that they connect for hbonds and ssbonds).
 - If the two atoms the bond is connecting are the same color, then the hbond or ssbond will be the same color throughout the entire length of the monitor line.
 - If the two atoms are different colors, then the hbond or ssbond will be half one color and half the other color.
- **Changing All Bonds** - To color all of a certain type of bond or strut, select all and then use the color command followed by the type of bond or strut you want to color and then the color you would like to color them.
 - For example:

- `select all`
- `color hbond red`

or:

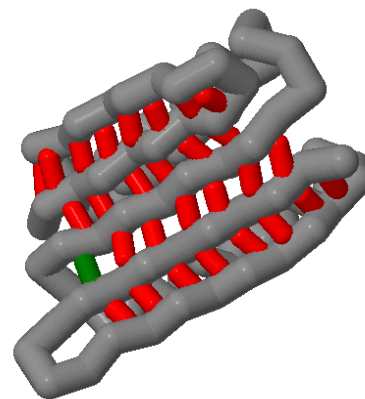
- `select all`
- `color strut red`



★ Note: We recommend that you choose a light color for your struts, such as white or light gray. Struts are support structures and should not be the emphasis of your model. A vibrant or dark color will draw the user's eye to the struts, and that should not be the focus of your model.

- **Changing A Single Bond Color** - To change the color of an individual bond or strut, follow the same steps for removing an individual bond or strut. Identify the two atom numbers that the bond or strut connects. Then, select only those two atoms. Finally, use your color command to apply the new color.
 - For example:

- `select atomno=333 or atomno=65`
- `color hbond green`



3.6 - Adding Sidechains to Create a "Clean Backbone"

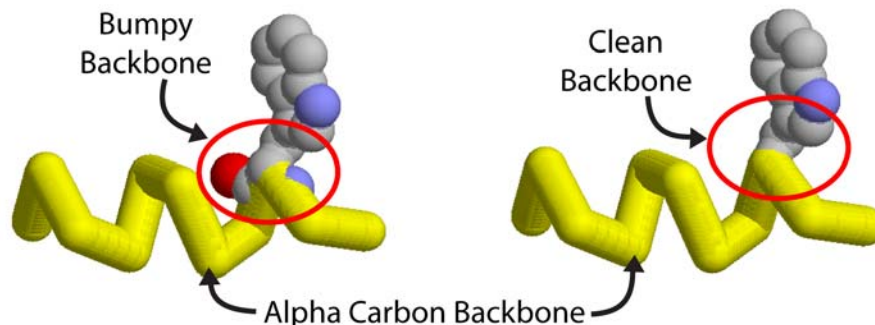
- If your molecular story requires that you display specific amino acid sidechains in the model, we recommend that you do only display the sidechain atoms, rather than all of the atoms of the amino acid. This gives the model a more aesthetically pleasing appearance. In previous sections, we have simply selected the amino acid by its residue number and displayed all of the atoms in that amino acid. In this section, we are going to use 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, you want to use the `select` command followed by the amino acid name/number and end with the `and (sidechain or alpha)` text.
 - `select cys30 and (sidechain or alpha)`
 - This command selects the amino acid at residue 30, but limits the selected atoms to the sidechain atoms and the alpha carbon of that amino acid.

★ Note: It is important to select the alpha carbon atom in addition to the sidechain atoms because this will attach the sidechain atoms to the alpha carbon. If you do not select the alpha carbon, the sidechain will build as a separate unit from the rest of the molecule. Once your sidechain is connected, give it wireframe and spacefill dimensions:

- `wireframe 1.0`
- `spacefill 1.25`

★ Note: 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 little spheres and the spheres may not be securely connected to one another. The image MAY look fine on the screen, but the model will come out of the 3D printer in pieces that can't be put back together! It is therefore imperative to add the wireframe command to connect the spheres together.

- If you do not use the `and (sidechain or alpha)` command when displaying sidechains, you will generate a "bumpy backbone." Unless the backbone atoms of the amino acid (the amino nitrogen or the carbonyl oxygen) play a specific role in the story you are telling, we 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.



3.7 - Additional Miscellaneous Jmol Commands

- In Section I of the Jmol Training Guide, you were introduced to the basic commands of Jmol. We will now present additional commands to further develop your Jmol design skills.
 - **The “Center” Command** - This command allows you to center the rotation of the molecule on a certain portion of the molecule.
 - - **center**
 - Jmol will by default center the molecule at the center of the entire molecule. If you restrict your viewing to a certain subset, when you rotate the molecule around in space, the molecule will seem lopsided. This is because although you have restricted your viewing to just a part of the molecule, all of the atoms are still present and being used to determine the center of rotation; you just do not see them.
 - Therefore, you need to use the **center** command to center the molecule on the restricted region.
 - For example: **center *c**

★ Note: Once again, we are using the ***** command to dictate to Jmol that we are centering the molecule on the all the atoms in Chain C.

- **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 Jmol, you 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 each end of the protein. The atom with the 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 Jmol command line window to select the specific amino acid.
- **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 each end of the protein. The amino acid with the largest number will be the last amino acid in the protein.

★ Note: Occasionally a PDB file will have gaps or missing sections in a chain because of incomplete experimental data in the file. This can easily be confused for the Carboxy Terminus. Be sure that you are really looking at the last amino acid in the chain and not just the beginning of a gap.

- **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 displayed, in reverse order. You can edit these commands before hitting enter.
- **Undo Button** - Near the bottom right of the Console Window is an undo button. This button can be used to undo recent commands and take your protein design back to a previous state. Be sure to save your work frequently! Use judiciously; sometimes you need to undo multiple steps to get to the command you want to undo.
- **Selecting specific types of atoms** - If you wish to select a subset of atoms, such as all of the carbon atoms, you can do so by using the `select` command followed by the type of atom you would like to select.
 - For example: `select carbon`

Quick Reference Standard Sizes for SMART Team Models

Standard Display Format Sizes

```
backbone 1.5  
wireframe 1.25  
spacefill
```

Standard Bond Sizes

```
hbond 1.0  
strut 1.0  
ssbond 1.0
```

Standard Clean Sidechain Sizes

```
spacefill 1.25  
wireframe 1.0
```