

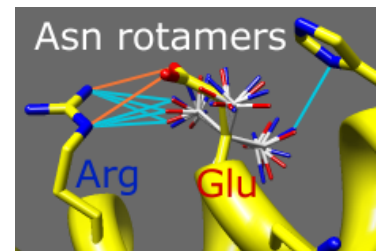
# Molecular Visualization with UCSF Chimera: Getting Started

- Presentation (~30 mins):
  - welcome and introductions
  - UCSF Chimera, the big picture
- Hands-on tutorial:
  - **Getting Started - Menu Version** (~35 minutes)
- Optional tutorial (available from **Help** menu):
  - **Getting Started - Command Version** (~25 minutes)

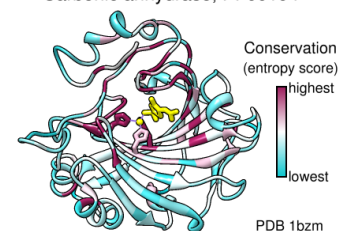
---

General molecular visualization and analysis (sessions 1 & 2):

- **many display styles and levels of abstraction** (from atomic detail to ribbons to surfaces)
- chemical perception, hydrogen addition, charge assignment
- high-quality H-bond, contact, and clash detection
- measurements: distances, angles, surface areas, RMSDs
- amino acid rotamers, Ramachandran plot
- bond rotation, structure building, **multimer assembly**
- structure superposition and morphing
- showing properties with colors and “worms”
- molecular dynamics trajectory playback and analysis
- interface for screening docking results
- web fetch from numerous databases, including PDB, EDS, EMDB, ModBase, CASTp, Pub3D, VIPERdb, and UniProt



Carbonic anhydrase, PF00194

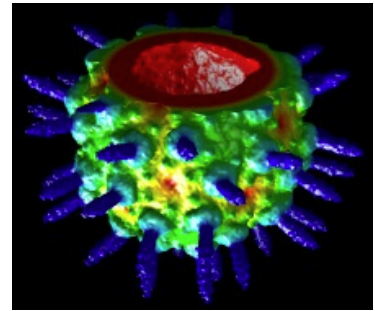
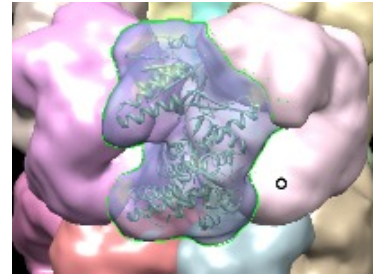


Sequence-structure tools (session 3):

- display of individual sequences and multiple sequence alignments
- sequences automatically associate with structures
- superimposing structures with or without pre-existing alignments
- generating structure-based sequence alignments from superpositions
- calculation of sequence and structure conservation per alignment column
- protein BLAST search via web service
- Modeller comparative modeling via web service or local installation

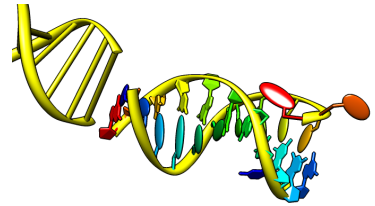
## Electron density maps and other "volume data" (session 5):

- many formats, display as mesh or **solid isosurfaces**
- interactive threshold adjustment
- map zoning around atoms of interest or pseudoatomic "markers"
- many options for smoothing, filtering, segmenting, and editing maps
- fitting of atomic coordinates to maps and maps to maps
- measurements: isosurface area, enclosed volume, map symmetry
- volume planes, time series playback, morphing



## Images and movies for presentation and publication (sessions 6 & 7):

- visual effects including fog, silhouettes, shadows, and special backgrounds
- convenient presets (combinations of display settings)
- 2D labels: text, symbols, arrows, and color keys
- structure-derived geometric objects: axes, planes, centroids, ellipsoids
- special nucleotide representations
- morphing and trajectory playback integrated with movie recording



## Getting Started Tutorial - Menu Version

Many tasks in Chimera can be accomplished in multiple ways. For example, colors and [display styles](#) can be changed with the [Actions menu](#) or by entering [commands](#). In general, commands are more concise and powerful, but menus allow easy access to features without knowledge of commands and their syntax.

In this tutorial, many of the same tasks performed with commands in the [Getting Started Tutorial - Command Version](#) are carried out using the menus instead.

To follow along, first [download](#) the PDB files included with this tutorial to a convenient location on your computer:

- [1zik.pdb](#) - leucine zipper
- [1d86.pdb](#) - DNA and netropsin

---

### Menus, Part 1 - Manipulation, Selection, and Chains

On **Windows/Mac**, click the **chimera** icon; on **UNIX**, start Chimera from the system prompt:

```
unix: chimera
```

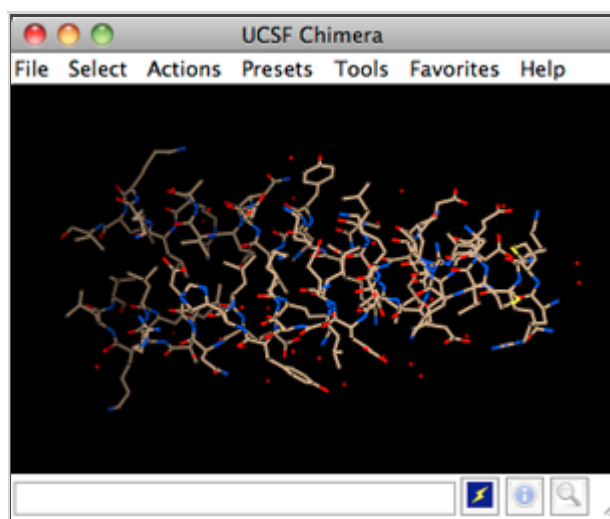
A splash screen will appear, to be replaced in a few seconds by the main Chimera [graphics window](#) or [Rapid Access](#) interface (it does not matter which, the following instructions will work with either). If you like, resize the Chimera window by dragging its lower right corner.

Now open a structure. Choose **File... Open** from the menu and use the resulting [file browser](#) to locate and open the previously [downloaded](#) file [1zik.pdb](#). The structure is a leucine zipper formed by two peptides.

A [preset](#) is a predefined combination of display settings. Apply interactive preset #2:

#### **Presets... Interactive 2 (all atoms)**

This displays all atoms and color-codes atoms other than carbon [by element](#) (oxygen red, nitrogen blue, etc.); carbons are left in the initial model color, in this case tan.



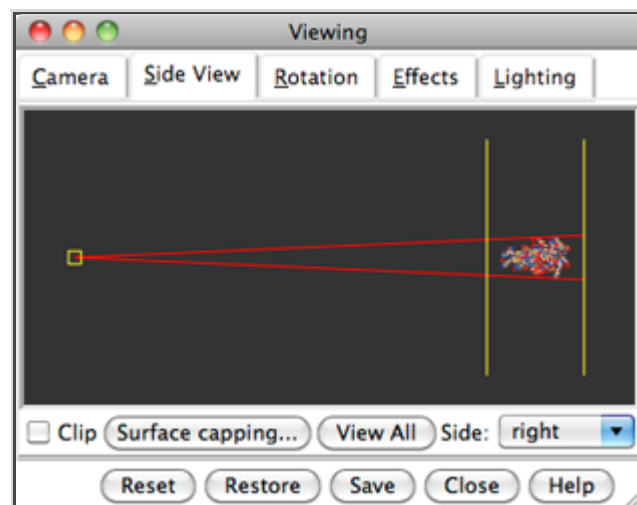
Try [moving the structure with the mouse](#) in the main graphics window. By default:

- the left mouse button controls rotation

- the middle mouse button controls XY translation
- the right mouse button performs scaling (zooming)

If you are using a touchpad or single-button mouse, modifier keys allow emulating the middle and right mouse buttons. These are **option** and **command** (⌘) on Mac keyboards.

Use the **Favorites** menu to show the **Side View** for interactive scaling and clipping. It shows a tiny version of the structure. Within the **Side View**, try moving the eye position (the small square) and the clipping planes (vertical lines) with the left mouse button. The **Side View** will renormalize itself after movements, so that the eye or clipping plane positions may appear to “bounce back,” but your adjustments have been applied.




Continue moving and scaling the structure with the mouse in the graphics window and **Side View** as desired throughout the tutorial. When the mouse focus is in the graphics window (you may need to click into it if you have been interacting with a different window), hovering the mouse cursor over an atom or bond (without clicking any buttons) will show identifying information in a pop-up “balloon.” The balloon will disappear when the cursor is moved away.

In Chimera, **selection** specifies atoms, bonds, residues, *etc.* for subsequent operations with the **Actions menu**. Ways to make a **selection** include using the **Select menu** or **picking** from the screen. The **Actions menu** applies to whatever is selected, but when nothing is selected, the **Actions menu** applies to everything.

Select the water (red dots):

**Select... Structure... solvent**

Another way to make the same selection is with **Select... Residue... HOH**. The selection is highlighted in green, and the magnifying glass icon near the bottom right of the window is also green: , indicating that something is selected. Hide the selected atoms:

**Actions... Atoms/Bonds... hide**

Even though the water atoms are hidden, they are still selected. Clear the selection, thicken the lines, and display only the chain trace:

**Select... Clear Selection**

**Actions... Atoms/Bonds... wire width... 3**

**Actions... Atoms/Bonds... backbone only... chain trace**

The chain trace includes just the  $\alpha$ -carbons (atoms named CA), connected in the same way

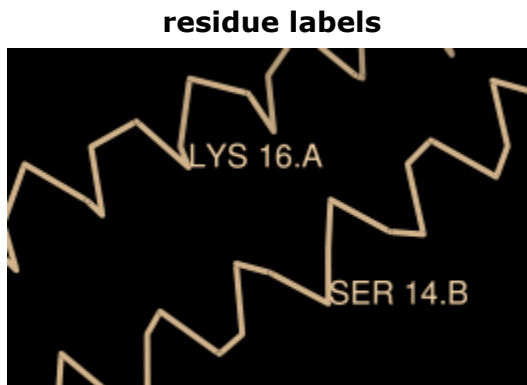
that the residues are connected.

By default, [picking](#) from the screen is done by clicking on the atom or bond of interest with the left mouse button while pressing the **Ctrl** key. To add to an existing selection, also press **Shift**. Try picking two atoms, one from each peptide chain (Ctrl-click the first, Shift-Ctrl-click the second).

Label the atoms you have selected, first by atom name and then by residue name and number:

**Actions... Label... name**  
**Actions... Label... off**  
**Actions... Label... residue... name + specifier**

The specifier includes residue number and chain ID. One peptide is chain A and the other is chain B. Use the **Favorites** menu to show the [Preferences](#), change to **Category: Background**, and adjust the **Label font** and size as you wish. Click **Save** before closing the preferences if you want the settings to apply to later uses of Chimera.



Clear the selection by Ctrl-clicking in empty space, as if [picking](#) "nothing." Turn off the labels:

**Actions... Label... residue... off**

Color the two chains different colors:

**Select... Chain... A**  
**Actions... Color... cyan**

Repeat the process to color chain B yellow.

Select chain A by [picking](#) any atom or bond in the chain, then pressing the up arrow key twice, once to [expand the selection](#) to the entire residue and another time to expand it to the entire chain. Display its full backbone:

**Actions... Atoms/Bonds... backbone only... full**

Display all atoms of chain A only (which is still selected):

**Actions... Atoms/Bonds... show only**

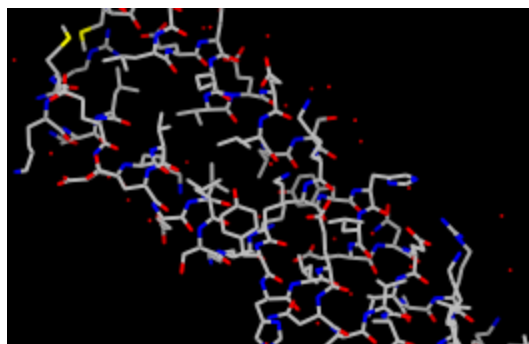
Display all atoms and color them by element:

**coloring by element**

**Select... Clear Selection**  
**Actions... Atoms/Bonds... show**  
**Actions... Color... by element**

The **by element** coloring is the same as **by**

**heteroatom** except it also color-codes carbons (gray). Heteroatom-only coloring is useful for keeping different structures distinguishable by their different carbon colors.



Generally, each file of coordinates opened in Chimera becomes a model with an associated model ID number. Models are assigned successive numbers starting with 0. Models are listed in the left side of the **Model Panel** (**Tools... General Controls... Model Panel**). A checkbox in the **A**(ctive) column of the **Model Panel** shows that the model is [activated for motion](#); unchecking the box makes it impossible to move the model. Checking the box again restores the movable state. Make sure **1zik.pdb** is highlighted on the left side of the **Model Panel** (if not, click on it) and then click **close** in the list of functions on the right side. Next, use the **Close** button at the bottom to dismiss the **Model Panel**.

Go on to **Part 2** below, **OR** exit from Chimera with **File... Quit**.

---

## Menus, Part 2 - Molecular Representations and Surfaces

With Chimera started as described at the beginning of [Part 1](#), choose the menu item **File... Open**. Use the resulting [file browser](#) to locate and open the previously [downloaded](#) file [1d86.pdb](#). It contains the molecule netropsin bound to double-helical DNA.

[Move and scale](#) the structure with the mouse in the graphics window and **Side View** as desired throughout the tutorial.

Apply the "all atoms" preset, which will show the DNA as wire and netropsin as spheres:

**Presets... Interactive 2 (all atoms)**


Color carbons white, then undisplay the water:

**Select... Chemistry... element... C**

**Actions... Color... white**

**Select... Structure... solvent**

**Actions... Atoms/Bonds... hide**

Remember that hiding atoms does not deselect them; they remain selected, as indicated by the green magnifying glass icon  near the bottom right of the window, until the selection is cleared or replaced with a new selection.

Color the different nucleotides different colors. For example, color the adenine deoxynucleotides blue:

**Select... Residue... DA**

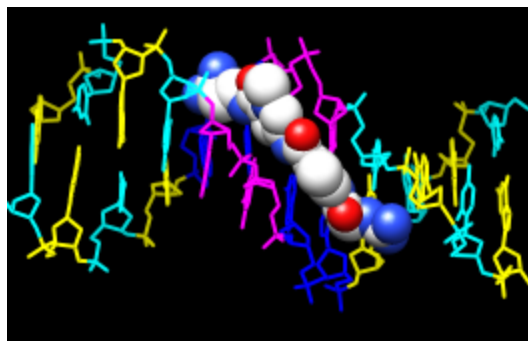
**Actions... Color... blue**

Analogously, color DC residues cyan, DG residues yellow, and DT residues magenta. Clear

the selection with **Select... Clear Selection** or by [picking](#) (Ctrl-click) in empty space.

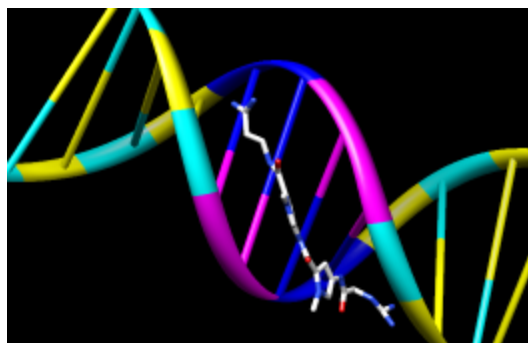
Next, try some different [display styles](#), or representations.

**Actions... Atoms/Bonds... sphere**  
**Select... Chain... A**  
**Actions... Atoms/Bonds... ball & stick**  
**Select... Clear Selection**  
**Actions... Atoms/Bonds... stick**



Showing ribbon automatically hides the [mainchain](#) (backbone) atoms.

**Actions... Ribbon... show**  
**Actions... Ribbon... edged**  
**Actions... Ribbon... rounded**



DNA can be shown with special [nucleotide objects](#). We will show "lollipops," boxes, and a ladder.

**Actions... Atoms/Bonds... nucleotide objects... settings...**

In the resulting **Nucleotides** dialog:

1. set **Show side (sugar/base) as** to **tube/slab**
2. set **Show base orientation** to **false**
3. click **Slab Style** tab, set slab style to **skinny**
4. click **Slab Options** tab, set **Slab object** to **ellipsoid**
5. click **Apply**; these are the "lollipops"

Nucleotide settings can be applied to just the selected residues (not necessarily all of the DNA). One way to select specific residues is in the [Sequence](#) tool:

**Favorites... Sequence**

Show the sequence of chain A and select one or a few residues in the sequence window with the mouse; this selects the corresponding part of the structure. **Quit** from the sequence window. In the **Nucleotides** dialog (also under **Tools... Depiction** in the menu):

1. set **Show base orientation** to **true**
2. set **Slab object** to **box**
3. click **Apply**; base orientations are shown with "bumps"

Clear the selection (**Select... Clear Selection** or Ctrl-click in empty space) and use **Nucleotides** to show the DNA as a ladder:

1. set **Show side (sugar/base) as** to **ladder**



2. in the **Ladder Options**, set **Rung radius** to **0.3 Å**
3. click **OK** (which will also dismiss the dialog)

To return to more general display styles, turn off the nucleotide objects:

**Actions... Atoms/Bonds... nucleotide objects... off**

Hide the ribbons and show everything as ball-and-stick:

**Actions... Ribbon... hide**

**Actions... Atoms/Bonds... ball & stick**

Finally, have some fun with surfaces. There are [built-in categories](#) within structures such as **main** and **ligand**; when nothing is [selected](#), **Actions... Surface... show** displays the surface of **main**.

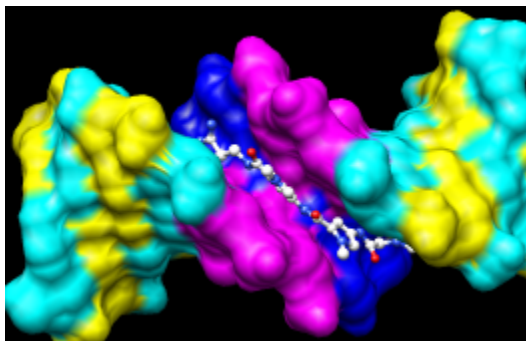
**Actions... Surface... show**

**Actions... Surface... hide**

**Select... Structure... ligand**

**Actions... Surface... show**

**Actions... Surface... mesh**



Surface color can be specified separately from the colors of the underlying atoms. The ligand surface is tan and white because the original model color (tan) is used for surfaces of atoms not explicitly recolored by the user, and above, only the carbon atoms were changed to white. With the ligand still selected, choose **Actions... Color... all options...** to open the [Color Actions](#) dialog. In that dialog,

1. change the **Coloring applies to** (target) setting to **surfaces**
2. click **red**
3. click **Close** (which will automatically reset the coloring target back to **all of the above**)

Clear the selection, change back to a solid surface, and then undisplay the surface:

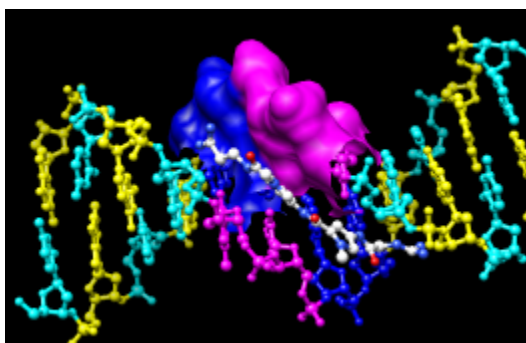
**Select... Clear Selection**

**Actions... Surface... solid**

**Actions... Surface... hide**

As an example of a more complicated selection process, show the surface of the adenine and thymine deoxynucleotides in chain B only:

1. change the selection mode: **Select... Selection Mode... append**
2. **Select... Residue... DA**
3. **Select... Residue... DT**
4. change the selection mode: **Select... Selection Mode... intersect**





5. **Select... Chain... B**
6. **Actions... Surface... show**

To prepare for any subsequent operations, restore the selection mode and clear the selection:

**Select... Selection Mode... replace**  
**Select... Clear Selection** (or Ctrl-click in empty space)

The [command](#) equivalent is much more concise, but requires some knowledge of the [atom specification syntax](#):

*Command:* **[surf](#) :da.b,dt.b**

Sometimes it is helpful to make a solid surface transparent:

**Actions... Surface... transparency... 50%**

Choose **File... Quit** from the menu to terminate the Chimera session.

---

*meng-at-cgl.ucsf.edu / September 2012*

## Getting Started Tutorial - Command Version

Many tasks in Chimera can be accomplished in multiple ways. For example, colors and [display styles](#) can be changed with the [Actions menu](#) or by entering [commands](#). In general, commands are more concise and powerful, but menus allow easy access to features without knowledge of commands and their syntax.

In this tutorial, many of the same tasks performed with menus in the [Getting Started Tutorial - Menu Version](#) are carried out using commands instead.

To follow along, first [download](#) the PDB files included with this tutorial to a convenient location on your computer:

- [1zik.pdb](#) - leucine zipper
- [1d86.pdb](#) - DNA and netropsin

---

### Commands, Part 1 - Manipulation, Selection, and Chains

On **Windows/Mac**, click the **chimera** icon; on **UNIX**, start Chimera from the system prompt:

```
unix: chimera
```

A splash screen will appear, to be replaced in a few seconds by the main Chimera [graphics window](#) or [Rapid Access](#) interface (it does not matter which, the following instructions will work with either). If you like, resize the Chimera window by dragging its lower right corner.

Use the **Favorites** menu to show the [Command Line](#). Now open a structure. Enter the command:

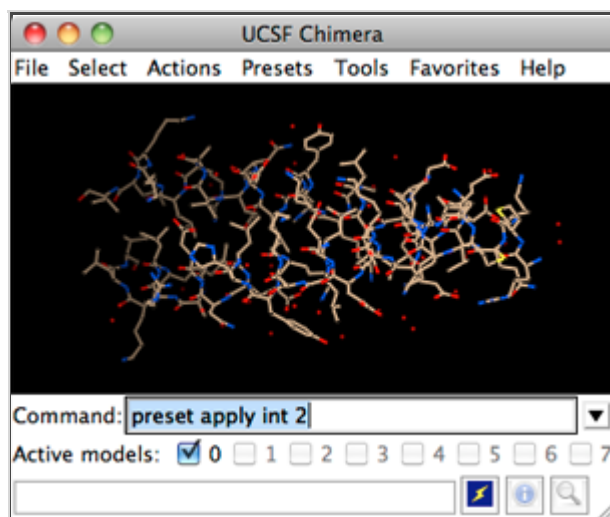
```
Command: open
```

and use the resulting [file browser](#) to locate and open the previously [downloaded](#) file [1zik.pdb](#). The structure is a leucine zipper formed by two peptides.

A [preset](#) is a predefined combination of display settings. Apply interactive preset #2:

```
Command: preset apply int 2
```

This displays all atoms and color-codes atoms other than carbon [by element](#) (oxygen red, nitrogen blue, etc.); carbons are left in the initial model color, in this case tan.



Try [moving the structure with the mouse](#) in the main graphics window. By default:

- the left mouse button controls rotation
- the middle mouse button controls XY translation

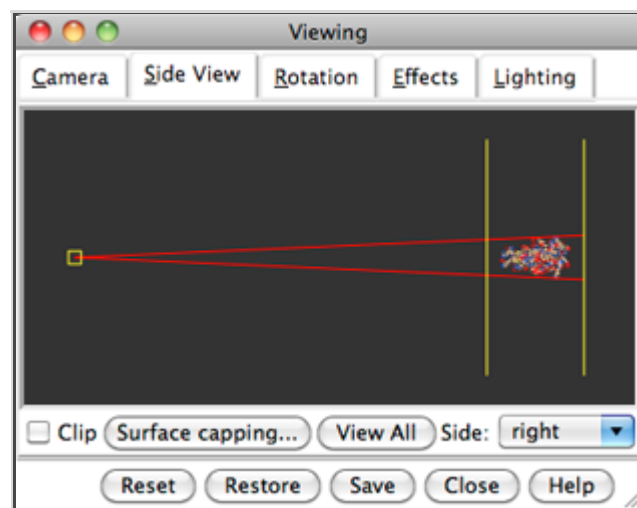
- the right mouse button performs scaling (zooming)

If you are using a touchpad or single-button mouse, modifier keys allow emulating the middle and right mouse buttons. These are **option** and **command** (⌘) on Mac keyboards.

Show the [Side View](#) for interactive scaling and clipping:

Command: [start Side View](#)

It is also listed in the **Favorites** menu by default. It shows a tiny version of the structure. Within the **Side View**, try moving the eye position (the small square) and the clipping planes (vertical lines) with the left mouse button. The **Side View** will renormalize itself after movements, so that the eye or clipping plane positions may appear to “bounce back,” but your adjustments have been applied.



Continue moving and scaling the structure with the mouse in the graphics window and **Side**

**View** as desired throughout the tutorial. When the mouse focus is in the graphics window (you may need to click into it if you have been interacting with a different window), hovering the mouse cursor over an atom or bond (without clicking any buttons) will show identifying information in a pop-up “balloon.” The balloon will disappear when the cursor is moved away.

A Chimera command may include **arguments** and/or an [atom specification](#). For example, in the following,

Command: [color hot pink :lys](#)

the [color name hot pink](#) is an argument of the command [color](#), and [:lys](#) specifies all residues named LYS.

A blank specification is interpreted as all applicable items. For example,

Command: [color hot pink](#)

makes all atoms (and their labels, surfaces, etc.) hot pink.

Many commands have “~” versions that perform the opposite function. The following will change a structure back to its default color:

Command: [~color](#)

The command [help](#) can be used to show the manual page for any command. For example,

Command: [help color](#)


Atom Specification Symbols		
Symbol	Meaning	Usage
#	model	# <i>model</i> (model ID number)
:	residue	: <i>residue</i> (residue name or number)
::	chain	:: <i>chain</i> (chain ID)
@	atom	@ <i>atom</i> (atom name)
=	partial wildcard	matches partial atom or residue name, e.g., @C= specifies all atoms with names beginning with C
?	single-character wildcard	matches single character in atom or residue name, e.g., :G?? specifies all residues with three-letter names beginning with G

shows the manual page for the command [color](#). The [Chimera Quick Reference Guide \(PDF\)](#) lists most [commands](#) and gives some examples of command-line [atom specification](#).

Thicken the lines and display only the atoms named CA ( $\alpha$ -carbons):

```
Command: linewidth 3  
Command: show @ca  
-OR- (equivalently)  
Command: ~display ~ @ca
```

The latter command undisplay atoms *not* named CA.

By default, [picking](#) from the screen (a type of [selection](#)) is done by clicking on the atom or bond of interest with the left mouse button while pressing the **Ctrl** key. To add to an existing selection, also press **Shift**. Try picking two  $\alpha$ -carbons, one from each peptide chain (Ctrl-click the first, Shift-Ctrl-click the second). The selection is highlighted in green, and the magnifying glass icon near the bottom right of the window is also green: , indicating that something is selected.

The word **selected**, **sel**, or **picked** can be used in commands to specify the current selection. Label the atoms you have selected:

```
Command: label sel
```

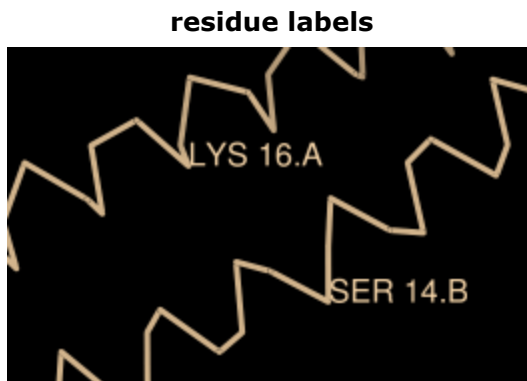
The [label](#) command shows atom information (atom name, by default). Undisplay the atom labels, then show labels for the *residues* containing the selected atoms:

```
Command: ~label  
Command: rlabel sel
```

Each residue label is of the form:

**res\_name res\_number.chain**

One peptide is chain A and the other is chain B. Use the **Favorites** menu to show the [Preferences](#), change to **Category: Background**, and adjust the **Label font** and size as you wish. Click **Save** before closing the preferences if you want the settings to apply to later uses of Chimera.



Clear the selection by Ctrl-clicking in empty space, as if [picking](#) "nothing." Turn off the labels:

```
Command: ~rlabel
```

Color the two chains different colors; note that commands can be truncated to unique strings:

```
Command: color cyan :.a  
Command: col yellow :.b
```

Residues and atoms can also be specified, along with or independent of chain:

```
Command: col orange :5-9.a,12.a,8.b  
Command: col magenta :14-18  
Command: disp :leu.b
```

Command: [col green :leu.b@cb](#)

The structure also includes water, which can be shown with:

Command: [disp solvent](#)

-OR- (equivalently)

Command: [disp :hoh](#)

Display the full backbone of chain A:

Command: [disp :a@n,ca,c,o](#)

Display all atoms in chain A only:

Command: [show :a](#)

Display all atoms and color them by element:

Command: [disp](#)

Command: [col byelement](#)

Coloring **byelement** is the same as **byhet** except it also color-codes carbons (gray). Heteroatom-only coloring is useful for keeping different structures distinguishable by their different carbon colors.

Generally, each file of coordinates opened in Chimera becomes a model with an associated model ID number. Models are assigned successive numbers starting with 0. The **Active models** line right under the **Command**

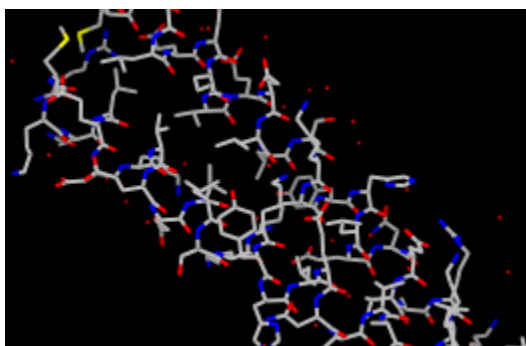
**Line** shows which models are [activated for motion](#). The checkbox for **0** is activated. Unchecking the box makes it impossible to move model 0. Checking the box again restores the movable state.

Command: [close 0](#)

closes the model. Go on to **Part 2** below, **OR** exit from Chimera with the following command:

Command: [stop](#)

coloring by element



---

## Commands, Part 2 - Molecular Representations and Surfaces

With Chimera started and the [Command Line](#) opened as described at the beginning of [Part 1](#), choose the menu item **File... Open**. Use the resulting [file browser](#) to locate and open the previously [downloaded](#) file [1d86.pdb](#). It contains the molecule netropsin bound to double-helical DNA.

[Move and scale](#) the structure with the mouse in the graphics window and **Side View** as desired throughout the tutorial.

Apply the "all atoms" preset, which will show the DNA as wire and netropsin as spheres:

Command: [preset apply int 2](#)

Color carbons white, then undisplay the water:

Command: **color white C**

Command: **~disp solvent**

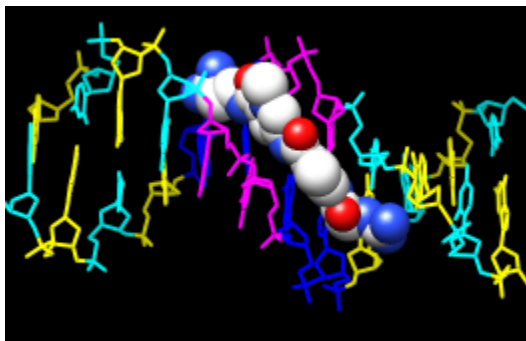
Residue names can be identified by looking in the **Select... Residue** menu or by hovering the cursor over an atom or bond to see information in a pop-up "balloon." Color the different nucleotides different colors, specifying them by residue name:

Command: **color blue :da**

Command: **color cyan :dc**

Command: **color yellow :dg**

Command: **color magenta :dt**



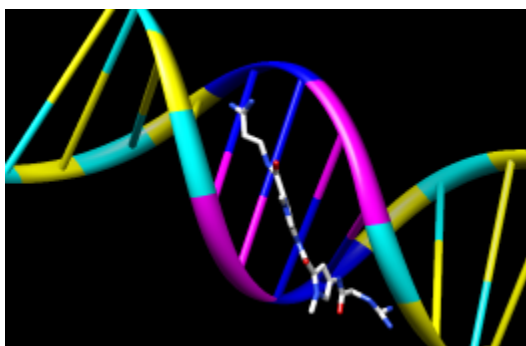
Next, try some different [display styles](#), or representations.

Command: **represent sphere**

Command: **repr bs :.a**

Command: **rep stick**

Notice that commands (but not necessarily their keyword arguments) can be truncated to unique strings. For example, the command **represent** can be shortened to **repr** or **rep** but not **re** (because other commands also start with **re**), whereas its keywords **stick**, **sphere**, etc. cannot be truncated. If the truncation is not unique, one of the corresponding commands will be executed, but it may not be the one intended.



Showing ribbon automatically hides the [mainchain](#) (backbone) atoms.

Command: **ribbon**

Command: **ribrep edged**

Command: **ribr rounded**

DNA can be shown with special [nucleotide objects](#). We will show "lollipops," boxes with orientation bumps, and then a ladder. You can copy and paste into the **Command Line**. The command-line contents can be edited, and past commands can be accessed using the up and down arrow keys or **Ctrl-p** (previous) and **Ctrl-n** (next).

Command: **nuc side tube/slab shape ellipsoid orient false style skinny**

Command: **nuc side tube/slab shape box orient true style skinny :8-10.a**

Command: **nuc side ladder radius 0.3**

To return to more general display styles, turn off the nucleotide objects:

Command: **~nuc**

Hide the ribbons and show everything as ball-and-stick:

Command: **~ribbon**

Command: [rep bs](#)

Finally, have some fun with the **surface** command. There are [built-in categories](#) within structures such as **main** and **ligand**; when nothing is specified, [surface](#) shows the surface of **main**.

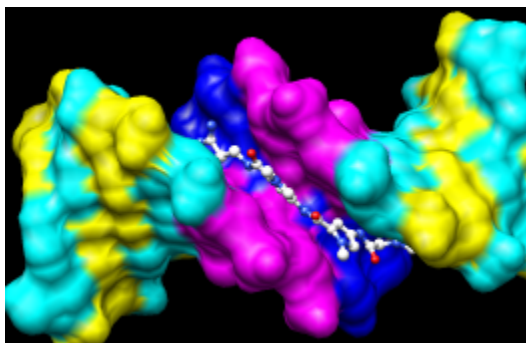
Command: [surface](#)

Command: [~surf](#)

Command: [surf ligand](#)

-OR- (equivalent)

Command: [surf :nt](#)



Surface color can be specified separately from the colors of the underlying atoms. The ligand surface is tan and white because the original model color (tan) is used for surfaces of atoms not explicitly recolored by the user, and above, only the carbon atoms were changed to white. Show the ligand surface as red mesh:

Command: [surfrep mesh](#)

Command: [color red,s ligand](#)

Command: [surfrep solid](#)

Parts of a surface can be shown:

Command: [~surf](#)

Command: [surf :da,dt](#)

Command: [~surf](#)

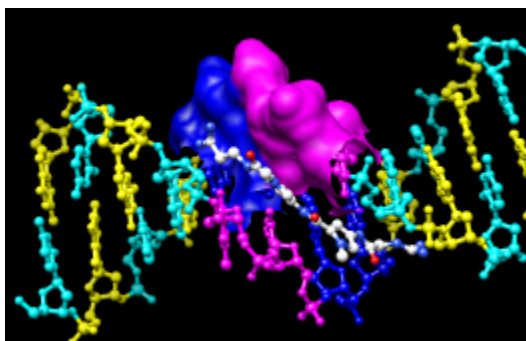
Command: [surf :da.b,dt.b](#)

Sometimes it is helpful to make a solid surface transparent:

Command: [surftrans 50](#)

When finished, exit from Chimera:

Command: [stop now](#)





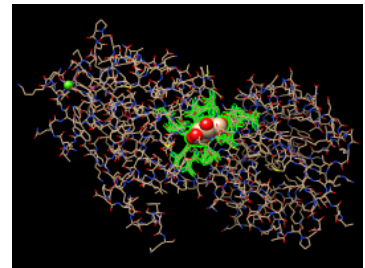


# Analyzing and Comparing Structures with UCSF Chimera

- Demonstration (~20 mins):
    - general concepts and tips for using Chimera
    - preview of some of the tools you will be using
    - OK (preferable?) to watch rather than trying to follow along; if you do try things, close session or quit/restart before starting the tutorial
  - Hands-on tutorial:
    - **Structure Analysis and Comparison** (~45 minutes)
  - Optional tutorials (available from **Help** menu):
    - **Attributes** (~20 minutes)
    - **Trajectory and Ensemble Analysis** (~35 minutes)
    - **The Model Panel and Ensembles** (~20 minutes)
- 

## Turning Data into Information

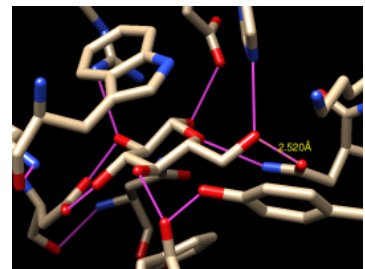
- molecular graphics as an analysis tool
- visual presentation of information carried in atomic coordinates
- mapping non-coordinate information onto graphics



## Demonstration Outline

General concepts, basic structure analysis:

- customizing tool startup and icons with **Preferences**
- zone selection
- measuring a distance
- identifying hydrogen bonds (**FindHBond**)
- commands and atom specification symbols
- command help, **Help** buttons, **Help** menu
- documentation is local (included in download) and searchable



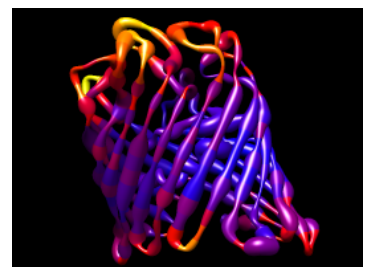
Structure comparison:

- superimposing structures with **MatchMaker**
- morphing between different conformations
- viewing and recording a morph trajectory with **MD Movie**



Attributes (atom and residue property values):

- showing B-factor with colors using **Render by Attribute**
- showing amino acid hydrophobicity with "worms"
- attributes can be used in commands
- users can define custom attributes in a simple text format
- various operations in Chimera also create attributes (surface area, conservation in an alignment, etc.)





## Structure Analysis and Comparison Tutorial

This tutorial includes binding site analysis and comparison of related structures by superposition and morphing. Internet connectivity is required to fetch the structures **2zcp** and **2zco**.

- [Background and Setup](#)
- [Distances, H-Bonds, Contacts](#)
- [Angles, Rotamers, Clashes](#)
- [Surfaces and Attributes](#)
- [Superposition and Morphing](#)

### ← Background and Setup

The pathogenic organism *Staphylococcus aureus* makes a pigment called staphyloxanthin. The pigment imparts a golden color (hence *aureus*), but more importantly, contributes to virulence by protecting the bacteria from killing by the host immune system. The *S. aureus* enzyme CrtM may be a good drug target because it catalyzes a key step in staphyloxanthin synthesis:

[A cholesterol biosynthesis inhibitor blocks \*Staphylococcus aureus\* virulence.](#) Liu CI, Liu GY, Song Y, Yin F, Hensler ME, Jeng WY, Nizet V, Wang AH, Oldfield E. *Science*. 2008 Mar 7;319(5868):1391-4.

We will view and compare different structures of this enzyme.

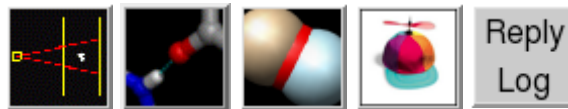
On **Windows/Mac**, click the **chimera** icon; on **UNIX**, start Chimera from the system prompt:

```
unix: chimera
```

A splash screen will appear, to be replaced in a few seconds by the main Chimera [graphics window](#) or [Rapid Access](#) interface (it does not matter which, the following instructions will work with either). If you like, resize the Chimera window by dragging its lower right corner.

Show the [Command Line](#) (**Tools... General Controls... Command Line**). Choose **Favorites... Add to Favorites/Toolbar** to place some icons on the toolbar. This opens the **Preferences**, set to **Category: Tools**. In the **On Toolbar** column, check the boxes for:

- [Side View](#) (under Viewing Controls)
- [FindHBond](#) (Structure Analysis)
- [Find Clashes/Contacts](#) (Structure Analysis)
- [Rotamers](#) (Structure Editing)
- [Reply Log](#) (Utilities, the last section)



If you want these icons to appear in later uses of Chimera, click **Save** before closing the preferences.

Fetch a structure from the [Protein Data Bank](#) and use the ribbons [preset](#):

Command: **open 2zcp**

Menu: **Presets... Interactive 1 (ribbons)**

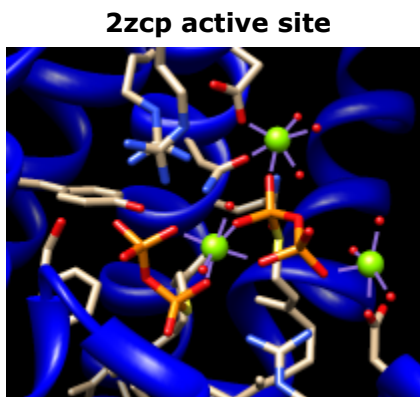
This preset displays the two protein chains (A and B) as red and blue ribbons; bound molecules and

nearby sidechains are shown as sticks. The two chains are two copies of the enzyme. Delete chain B:

Command: **[delete](#) ::b**

Move and scale structures [using the mouse](#) and [Side View](#) as desired throughout the tutorial.

The enzyme combines two 15-carbon molecules of farnesyl pyrophosphate to form a 30-carbon lipid. This structure contains farnesyl **thi**opyrophosphate, which differs from the substrate by having a sulfur in the place of one oxygen. Sulfur atoms are shown in yellow, phosphorus orange, oxygen red, and nitrogen blue. Label the ligand residues:



Command: **[rlabel](#) ligand**

In this structure, the farnesyl thiopyrophosphate molecules are named FPS and are residues 657 and 658 of chain A. Three metal ions help to offset the negative charges on the phosphates. These are shown as greenish spheres; hovering the mouse cursor over each reveals (assuming default [balloon preferences](#)) that they are Mg ions. Metal coordination bonds from FPS, water, and the protein are shown with dashed purple lines. Lines drawn to indicate interactions other than covalent bonds are called **pseudobonds**. Hovering the cursor over a pseudobond or bond shows balloon information about its end atoms and length.

Delete the water, label residues with displayed atoms, and place the labels near  $\alpha$ -carbons instead of residue centroids:

Command: **[del](#) solvent**

Command: **[rlabel](#) @/display**

Command: **[setattr](#) m residueLabelPos 2**

Water could be included in the various analyses, but is removed here to simplify the tutorial.

## ← Distances, H-Bonds, Contacts

It looks like several sidechains could be donating hydrogen bonds to phosphate oxygens. (Although the structure does not include hydrogens, we know they are there!)


One of the displayed residues is Ser 21. To measure a distance:

1. Ctrl-click to pick the sidechain oxygen of Ser 21
2. Shift-Ctrl-doubleclick on the nearest phosphate oxygen
3. click **Show Distance** in the resulting menu

Similarly, measure the distance between the sidechain oxygen of Tyr 248 and the same phosphate oxygen.

The distances seem consistent with hydrogen bonds. However, rather than measuring many distances and trying to remember the appropriate hydrogen-bonding distances for different types of atoms, just use [FindHBond](#). We will limit the search to H-bonds involving the FPS residues:

1. Select the FPS residues (for example, with **Select... Residue... FPS** in the menu).

2. Start [FindHBond](#) by clicking its [icon](#): 
3. In that dialog, turn on the options:
  - o **Only find H-bonds with at least one end selected**
  - o **Write information to reply log**
4. Set **Line width** to 3.
5. Click **OK**.


The H-bonds are shown as pseudobonds of the specified color and line width. Both measured distances were found to be consistent with hydrogen bonding. Details can be viewed in the **Reply Log**, which can be opened by clicking its [icon](#). (In your own work, you may prefer to write the information to a file instead.) Remove the "hydrogen bond" pseudobonds:

*Command:* [~hbond](#)

[Find Clashes/Contacts](#) has some similarities to **FindHBond**, but it can also identify nonpolar interactions:

- **clashes** - unfavorable interactions where atoms are too close together; close contacts
- **contacts** - all kinds of direct interactions: polar and nonpolar, favorable and unfavorable (including clashes)

We will identify contacts of the FPS residues:


1. Start [Find Clashes/Contacts](#) by clicking its [icon](#): 
2. With the FPS residues still selected, click **Designate**. 48 atoms should be designated for checking against **all other atoms**.
3. Set the **Clash/Contact Parameters** to the default **contact** criteria (by clicking the button marked **contact**).
4. Set the **Treatment of Clash/Contact Atoms** to:
  - o **Select**
  - o **Write information to reply log**
  - o (turn off any other treatment options)
5. Click **Apply**.

In the **Reply Log**, the atom-atom contacts are listed in order of decreasing VDW **overlap**, the amount of intersection of atomic van der Waals spheres: 0 if just touching, positive if interpenetrating, negative if separated by space. Slightly separated spheres are still considered contacts by default. The distances between atomic centers are also given.

Even though all of the contacting atoms have been selected, not all of them are displayed. Click into the graphics window, press the keyboard up arrow key to promote the selection from atoms to whole residues, and then display the selection (**Actions... Atoms/Bonds... show**).

One might simply want a list of the interacting residues rather than the details of each atomic contact. A list of the selected residues can be saved. First, deselect the ligand residues and ions, leaving the protein residues selected:

*Command:* [~sel ligand | ions](#)

Open the [Selection Inspector](#) by clicking the magnifying glass icon  near the bottom right corner of the main window. It reports that 26 residues are selected. Click its **Write List...** button (or choose **Actions... Write List** from the main menu). In the resulting dialog, indicate that

**selected residues** should be written. Click **Log** to write the list to the **Reply Log** instead of to a file.

## ← Angles, Rotamers, Clashes

Amino acid torsion angle values can be viewed in the [Selection Inspector](#). Focus on Tyr 248:

*Command:* **focus :248**


Ctrl-click to pick any atom or bond in Tyr 248. This will replace (clear) any pre-existing selection. Open the [Selection Inspector](#) again if it is not already open. Inspecting **Residue** (instead of **Atom**) shows that the sidechain angles chi1 and chi2 in this residue are approximately  $-107^\circ$  and  $-142^\circ$ , respectively. The backbone angles phi and psi are also given. **Close** the **Selection Inspector**.

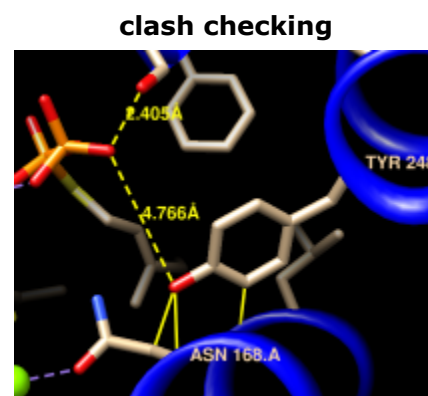
We will set up clash checking and rotate the sidechain interactively.

1. Display residues within 4 Å of Tyr 248:

*Command:* **disp :248 z<4**

2. Click into the main graphics window, then press the up arrow key on the keyboard to promote the atom or bond selection within Tyr 248 to the whole residue.

3. Start [Find Clashes/Contacts](#) by clicking its [icon](#): 
4. In that dialog, click **Designate**. The 12 atoms of Tyr 248 should be designated for checking against **all other atoms**.



5. Set the **Clash/Contact Parameters** to the default **clash** criteria (by clicking the button marked **clash**).
6. Set the **Treatment of Clash/Contact Atoms** to **Draw pseudobonds**; turn off any other treatment options.
7. Set the **Frequency of Checking** to **after relative motions**.

No clashes are found in the current position. To rotate the sidechain of Tyr 248 interactively:

1. Ctrl-doubleclick its CA-CB bond (the stick segment right next to the ribbon)
2. click **Rotate Bond** in the resulting menu

The rotatable bond will be listed in the [Adjust Torsions](#) dialog. In that dialog, if the **Near** atom is set to **N**, the value reported is the chi1 angle. There are several ways to rotate a bond, including dragging the pointer on the dial, clicking the black arrowheads, and editing the angle value directly. If the dialog becomes obscured by other windows, it can be raised again from the menu (**Tools... Structure Editing... Adjust Torsions**).

Use whichever method you prefer to rotate the bond. As the sidechain moves, yellow pseudobonds show any clashes and the distance monitor updates automatically. The sidechain is fairly constrained; only chi1 angles of approximately  $-(100-120)^\circ$  avoid clashes if only that bond is rotated. The sidechain can be frozen in a new position, but for tutorial purposes, simply restore it to its original position: in the **Adjust Torsions** dialog, click the entry under **Bond** to show a menu. In that menu, choose **Revert** (move back to original position) and then **Deactivate** (make the




bond no longer rotatable). **Close** the torsion dialog.

Click the **Find Clashes/Contacts** [icon](#) to bring the dialog to the front and **Close** it to halt clash checking.

Next, we will compare the conformation of Tyr 248 in the structure to tyrosine rotamers from a library. This will indicate whether the sidechain is in a frequently observed conformation and whether other conformations might also fit into the structure.

With part of Tyr 248 still selected (if not, Ctrl-click to to select any of its atoms or bonds), start

**Rotamers** by clicking its [icon](#) . Click **OK** to show **TYR** rotamers from the [Dunbrack backbone-dependent library](#). The rotamers are shown in the wire representation and listed in another dialog.

Clicking a line in the rotamer dialog displays just that rotamer. Clicking the lines one by one shows that none of the rotamers match the conformation in the structure.

The rotamer dialog reports chi1 and chi2 angles and probabilities from the library based on the residue's backbone conformation. Other than the backbone angles, the probabilities do not take into account any interactions specific to this structure. However, the rotamer dialog is integrated with clash and H-bond detection. Choose **Columns... Add... Clashes** and click **OK** to use the default settings; repeat with **Columns... Add... H-Bonds**.

The new columns show that each rotamer forms several clashes but no hydrogen bonds. We showed above that Tyr 248 H-bonds with the ligand and avoids clashes, which may compensate for its nonrotameric (presumably strained) conformation.

A sidechain can be replaced with a chosen rotamer. With a single rotamer shown, either click **OK** in the rotamer dialog to replace the sidechain with that rotamer and remove the others, or simply **Close** the dialog to remove the rotamers without replacing the sidechain. Although not done in this case, rotamers can be shown for a different amino acid type than in the structure, allowing for virtual mutations.

Unlike Tyr 248, Tyr 41 closely resembles the highest-probability tyrosine rotamer given its backbone angles. If you like, focus on Tyr 41, select one of its atoms or bonds, and use [Rotamers](#) to show and evaluate rotamers of tyrosine or some other amino acid at that position.

Clear the selection, unlabel the residues, zoom back out, and remove the distance pseudobonds:

Command: [~sel](#)

Command: [~rlab](#)

Command: [focus](#)

Command: [~dist](#)

## ← Surfaces and Attributes

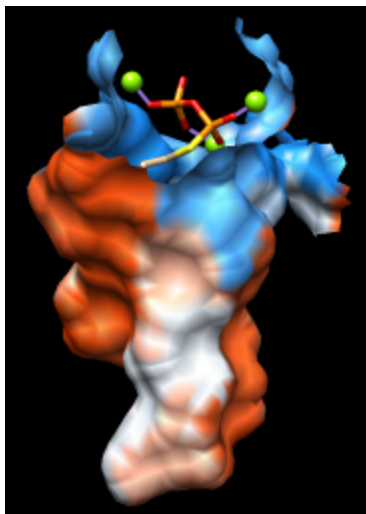
FPS and the natural substrate both have a highly polar/charged "head" and a long hydrocarbon "tail." The enzyme pocket is very large and deep, as needed to accommodate two of these molecules.

A surface representation is best for showing the shape of the pocket. Choose **Presets... Interactive 3 (hydrophobicity surface)** from the menu to display a molecular surface color-coded by amino acid hydrophobicity. This [preset](#) colors the surface ranging from dodger blue for the most polar residues to orange red for the most hydrophobic, with white in between.

As one might expect, the mouth of the pocket near the phosphates and ions is mostly polar, while the rest of the pocket is largely hydrophobic. However, the pocket is so deep that it is hard to see when the whole surface is shown. Restricting the surface display to the pocket region allows it to be viewed from the outside, as shown in the [figure](#).

Hide the ribbon and the protein atoms, then show only surface near the ligand:

### 2zcp pocket surface



Command: `~ribbon`

Command: `~disp protein`

Command: `sop zone #0 ligand 5.0 max 1`

The `sop` ("surface operations") command limits the display of surface #0 to a 5.0-Å zone around the ligand and shows only the largest resulting fragment. Without `max 1`, a few additional disconnected shards of surface would be shown. If you like, try different cutoff values; the [figure](#) shows results for a cutoff of 4.7 Å.

Amino acid hydrophobicity is one example of an **attribute**. Another attribute commonly used in structure analysis is atomic B-factor. B-factor values indicate which parts of the structure are more or less flexible or disordered.

Display all atoms by choosing **Presets... Interactive 2 (all atoms)** from the menu. Now the ligand atoms are shown as spheres and the protein as wires. Show the protein as sticks:

Command: `repr stick protein`

The **Render by Attribute** tool shows attribute values with colors and other visual cues. Choose **Tools... Depiction... Render by Attribute** from the menu to start this tool. The dialog is set to **atoms** when it appears. Choose **bfactor** from the attribute menu. The B-factor values are shown in a histogram, and the vertical bars on the histogram define a color mapping. These bars or **thresholds** can be dragged along the histogram with the mouse, and their colors can be changed. Clicking the square [color well](#) allows changing the **Color** of the threshold that was most recently clicked or moved. The colors can also be changed collectively by choosing a different **Palette** or clicking the button to **Reverse threshold colors**. (The [figure](#) shows the dialog after a user has added more thresholds, chosen the **Rainbow** palette, and then reversed the colors.) Adjust the color mapping as desired and **Apply** it to the structure. The lowest B-factors are in the protein core, the highest in a loop over the active site and the C-terminus on the opposite side. If any sidechain has been replaced with a rotamer, the new atoms will not have B-factor values, but they can be assigned a **No-value color**.

In the **Render** dialog, switch from **atoms** to **residues** and show the histogram for the attribute **kdHydrophobicity**, the [Kyte-Doolittle hydrophobicity](#) of amino acids. This is the same attribute as used by the surface [preset](#), but different coloring schemes can be applied. Larger positive values correspond to more hydrophobic residues, negative values to hydrophilic residues. The **No-value color** pertains to residues without Kyte-Doolittle hydrophobicities, namely any that are not amino acids (such as the ligands in this structure). The colors can be changed as described above. Apply coloring by hydrophobicity if you like, then **Close** the dialog.

## ← Superposition and Morphing

Comparing different structures of a protein is another way to evaluate flexibility. We have been

viewing a structure bound to substrate analogs, **2zcp**. A structure of the same enzyme without ligands is also available.

Fetch the "empty" structure and apply the ribbons preset:

Command: **open 2zco**

Command: **preset apply int 1**

Now the first structure is tan and the new structure is sky blue, as shown in the **Model Panel (Tools... General Controls... Model Panel)**.

The structures are in completely different positions, so the next step is to superimpose them. This is easily accomplished with **MatchMaker** or its [command equivalent](#):

Command: **mm #0 #1**

**MatchMaker** generates a sequence alignment using residue types and secondary structure (tries to align helix with helix and strand with strand) and then fits CA-CA pairs in the same columns of the sequence alignment. By default, the fit is iterated so that far-apart pairs are not included in the final match. Match statistics are reported in the **Reply Log**. The sequence alignment is not shown by default, but can be displayed by adding **show true** to the above command.

Alternatively, the following produces exactly the same superposition:

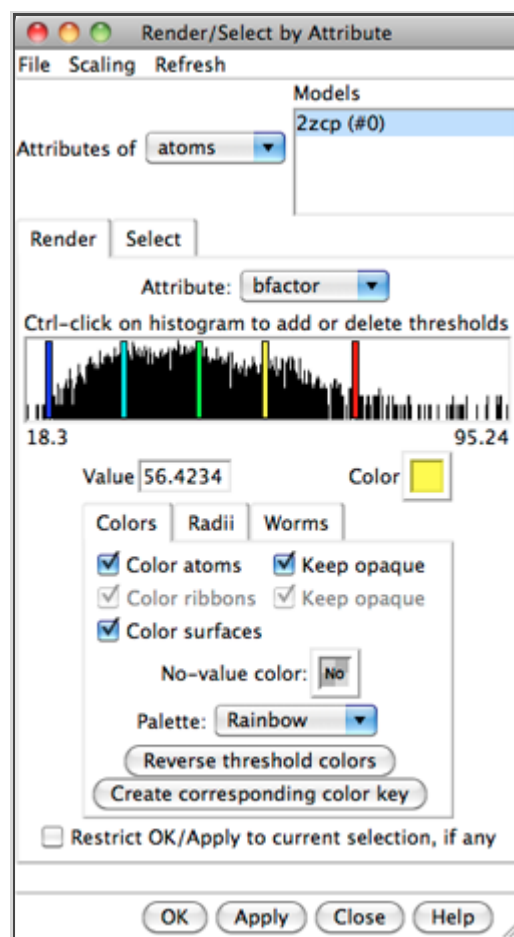
Command: **match iterate 2.0 #1@ca #0@ca**

This command works because the structures have equal numbers of CA atoms, and the first CA atom in one structure correctly pairs with the first in the other, the second with the second, and so on. The **match** command does not generate a sequence alignment, so it is faster than **MatchMaker**. Another potential advantage is that it can use any atoms, not just CA. Its main disadvantage is that in most cases, one must figure out and specify exactly which atoms in one structure should pair with which atoms in the other. While not used by default, iteration can be turned on, as shown above.

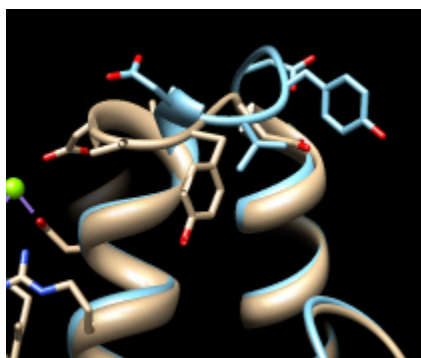
Either way, iteration is recommended so that similar parts will superimpose well and dissimilar parts will stand out better. These two structures are very similar except for the loop at approximately residues 51-56 (you can see residue numbers by hovering the cursor over the ribbon):

Command: **disp :51-56**

Morphing involves calculating a series of intermediate, interpolated structures between the original



**different loop conformations**



input structures. In Chimera, the series of structures is treated as a trajectory that can be replayed, saved to a coordinate file, or saved as a movie using [MD Movie](#).

Start the [morphing tool](#):

*Command:* **[start Morph Conformations](#)**

Click **Add...** and in the resulting list of models, doubleclick to choose #0, #1, and #0 again, corresponding to a morph trajectory from the ligand-bound structure to the empty structure and back. **Close** the model list. In the main **Morph Conformations** dialog, set the **Action on Create** to **hide Conformations**, and then click **Create**.

The progress of the calculation is reported in the [status line](#). When all the intermediate structures have been calculated, the input structures are hidden, the trajectory is opened as a third model (#2), and the [MD Movie](#) tool appears. The trajectory can be played continuously or one step at a time by clicking buttons on the tool, or a frame number can be entered directly. If the **MD Movie** playback dialog becomes obscured by other windows, it can be raised using its [instance in the Tools menu](#) (near the bottom of the menu, below the horizontal line).

The initial display of the morph trajectory is as ribbons rather than all atoms, but all atoms that are present in both input structures are also present in the morph. Their display can be controlled just as in any other structure, for example:

*Command:* **[disp #2:51-56](#)**

Ligand residues were only present in one of the structures, so they are not included in the morph trajectory. However, part or all of the original structures can be displayed along with the trajectory. Display of the other models can be re-enabled by checking the **S(hown)** boxes in the [Model Panel \(Tools... General Controls... Model Panel\)](#) or with a command, for example:

*Command:* **[modeldisp #0](#)**

Then, to show only the ligand residues and ions from that model:

*Command:* **[~ribbon #0](#)**

*Command:* **[show :fps,mg](#)**

In this case, the secondary structure does not change much between the input structures. When conformational differences are larger and ribbons will be displayed, one may want to re-evaluate the secondary structure at each step of the trajectory. That can be done by putting the command [ksdssp](#) in a [per-frame Chimera script](#) in [MD Movie](#).

The [Morph Conformations](#) page lists a few additional [example systems](#).

When finished enjoying the morph trajectory, choose **File... Quit** from the menu to exit from Chimera.

# Exploring Sequence-Structure Relationships with UCSF Chimera

- Demonstration (~20 mins):
  - preview of some of the tools you will be using
  - OK (preferable?) to watch rather than trying to follow along; if you do try things, close session or quit/restart before starting the tutorial
- Hands-on tutorials:
  - **Sequences and Structures** (~25 minutes)
  - **Superpositions and Alignments** (~40 minutes)

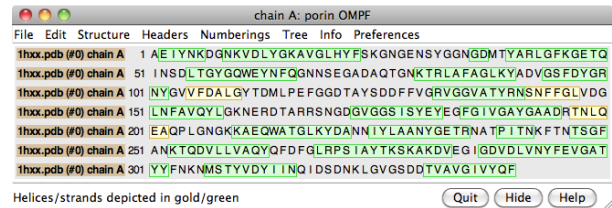
## Combining Structure with Other Data

- Structure provides spatial information
- Sequence adds non-spatial information
- "The whole is greater than the sum of the parts"

## Demonstration Outline

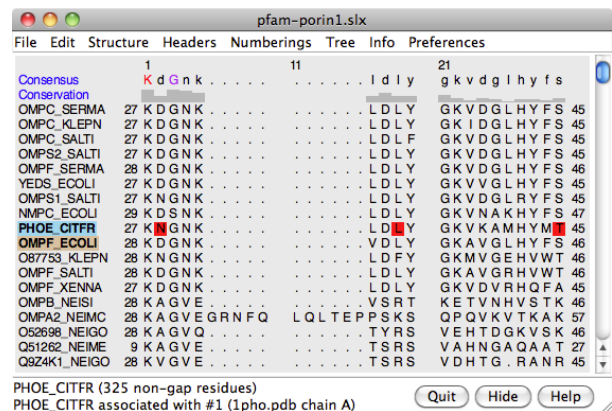
Basic sequence-structure crosstalk:

- **Sequence** shows the individual sequences of structure chains
- selecting in sequence selects structure and vice versa
- pastel boxes show helix and strand assignments
- mouse over sequence reports structure residue name/number



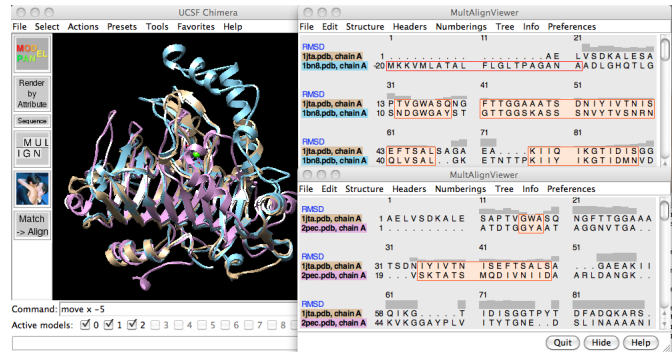
Multiple sequence alignment from a file:

- [external sources](#) include many other programs and databases; example is from PFAM
- shown with **Multalign Viewer** (actually same tool as **Sequence**)
- sequence-structure association is automatic, tolerates some mismatches
- many appearance settings (**Multalign Viewer** preferences)
- lines of information above the sequences are "headers"
- numerical headers are automatically assigned as residue attributes of associated structures
- many options for calculating **Conservation** header (**Multalign Viewer** preferences)



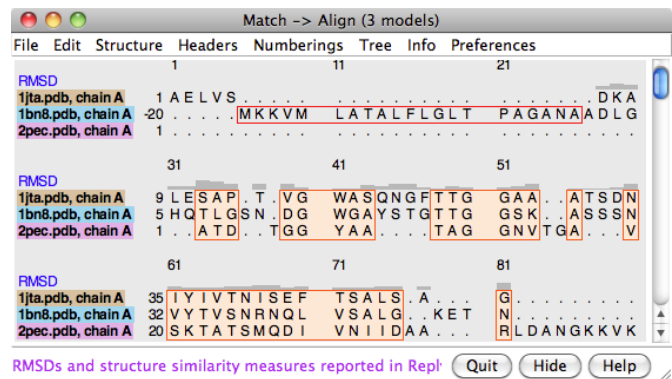
When multiple structures are associated with an alignment:

- structures can be superimposed based on the sequence alignment
- **RMSD** header shows “spatial conservation”



Creating sequence alignments in Chimera:

- **MatchMaker** creates pairwise sequence alignments which are then used to superimpose the structures
- given a multiple structure superposition (regardless of how it was made), **Match**→**Align** creates a multiple sequence alignment
- sequence alignments can be saved in many formats, those that can be read plus EPS





## Sequences and Structures Tutorial

This tutorial is an introduction to working with sequence alignments and associated structures using [Multalign Viewer](#). Data are taken from the enolase superfamily:

[The enolase superfamily: a general strategy for enzyme-catalyzed abstraction of the alpha-protons of carboxylic acids](#). Babbitt PC, Hasson MS, Wedekind JE, Palmer DR, Barrett WC, Reed GH, Rayment I, Ringe D, Kenyon GL, Gerlt JA. *Biochemistry*. 1996 Dec 24;35(51):16489-501.

To follow along, first [download](#) the sequence alignment file [super8.msf](#) to a convenient location. This sequence alignment contains the barrel domains of eight enolase superfamily members.

On **Windows/Mac**, click the **chimera** icon; on **UNIX**, start Chimera from the system prompt:

```
unix: chimera
```

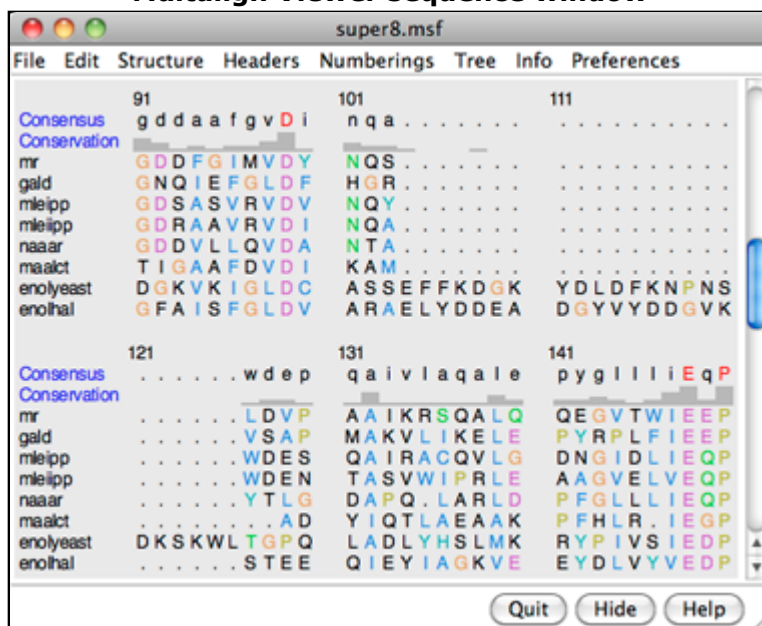
A splash screen will appear, to be replaced in a few seconds by the main Chimera [graphics window](#) or [Rapid Access](#) interface (it does not matter which, the following instructions will work with either).

Choose **File... Open**, then locate and open [super8.msf](#). Opening a [sequence alignment file](#) automatically starts **Multalign Viewer** and uses it to display the alignment.

**Multalign Viewer** has its own set of preferences, including display options. From the alignment window menu, choose **Preferences... Appearance** and adjust settings as desired for **Multiple alignments**, then **Close** the preferences dialog.

Size and place the sequence window and main Chimera window so that both are visible. If the sequence window later becomes obscured, it can be raised by choosing **Tools... MAV - super8.msf... Raise** from the menu, or by clicking **super8.msf** in the lefthand **Active Dialogs** section of [Rapid Access](#) (shown/hidden by clicking the lightning-bolt icon at the bottom of the Chimera window).

Multalign Viewer sequence window



A [Consensus sequence](#) and [Conservation histogram](#) are shown above the sequences. They can be hidden/shown using the **Headers** menu. Red capital letters in the consensus sequence indicate a few positions in which all of the sequences have the same type of residue.

The first sequence in the alignment, named **mr**, corresponds to the barrel domain of mandelate racemase from *Pseudomonas putida*. The next-to-the last sequence in the alignment, named



**enolyeast**, corresponds to the barrel domain of enolase from *Saccharomyces cerevisiae*. There are multiple structures in the Protein Data Bank for each of these sequences; **2mnr** (mandelate racemase) and **4enl** (enolase) are used in this tutorial.

If you have internet connectivity, structures can be obtained directly from the [Protein Data Bank](#). Use **Fetch by ID** (**File... Fetch by ID** in the Chimera menu) to get **2mnr** and then **4enl** from the **PDB** database. If you do not have internet connectivity, instead [download](#) the files [2mnr.pdb](#) and [4enl.pdb](#) included with this tutorial and open them in that order with **File... Open**.

The view is initially centered on the first structure opened, mandelate racemase. Off to the side and possibly out of view is the second structure, enolase. Use the menu to bring everything within view:

### Actions... Focus

Apply the ribbons [preset](#) (which may or may not change the appearance, depending on your preference settings):

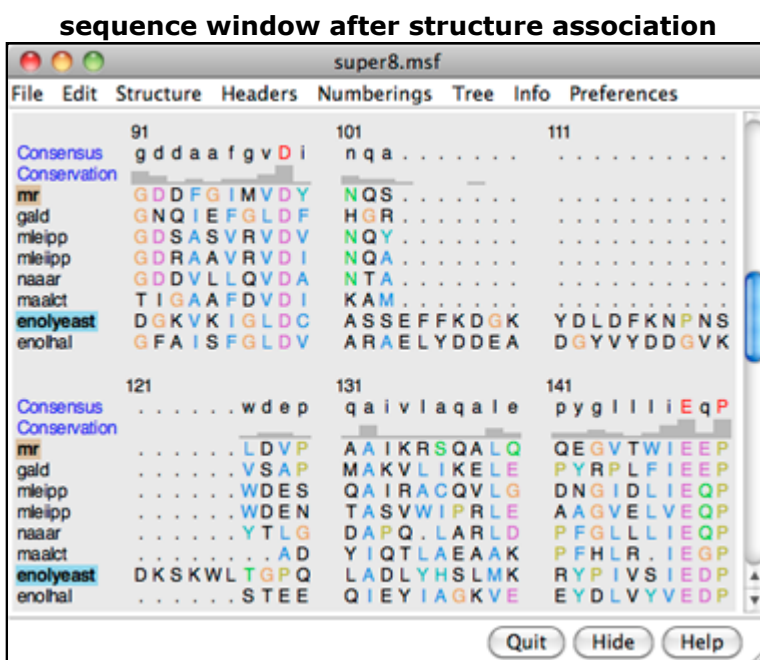
### Presets... Interactive 1 (ribbons)

[Move and scale](#) the structures as desired throughout the tutorial.

Notice that in the sequence window, the sequence names **mr** and **enolyeast** are now shown in bold within colored boxes. Each structure was [automatically associated](#) with its best-matching sequence, and the colors indicate the pairing.

Association enables several types of crosstalk, which will be explored in more detail below:

- selecting in the sequence window selects residues in the structure(s) and *vice versa*
- information from the sequences can be shown on the structures and *vice versa*
- the structures can be superimposed using the sequence alignment



Association tolerates some mismatches, which is important for working with mutated or partial sequences and/or structures. As a practical matter, sequence data from common sources are rarely perfect matches to the sequences of structures, even if both are available for the same species. Further, structures may not be available for the exact proteins in the sequence alignment, only their homologs.

First, we will use the sequence alignment to guide a structural match. From the **Multalign Viewer** menu, choose **Structure... Match** and click **Apply** without checking any boxes. This superimposes the structures using the alpha-carbons of all pairs of residues aligned in the sequence alignment.

Readjust the view to focus on the active sites, in this case, the metal ions and surrounding atoms:

**Select... Structure... ions**  
**Select... Zone...** (try an 8-Å cutoff)  
**Actions... Focus**

Clear the selection by Ctrl-clicking in an empty part of the graphics window.

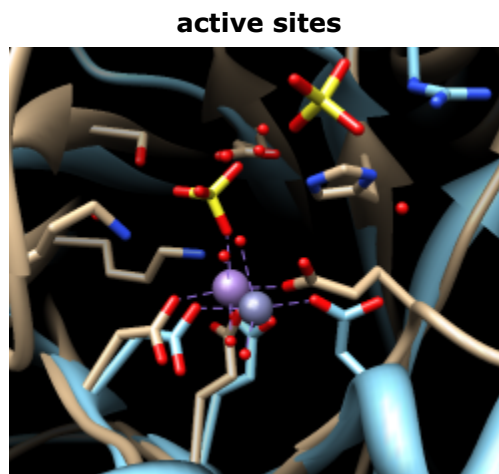
Match statistics are reported briefly in the status line at the bottom of the Chimera window and written to the [Reply Log](#) (**Favorites... Reply Log**). The match is fairly rough, giving an RMSD of 8.4 Å, and the active sites are somewhat offset. Apparently, not all 175 pairs of residues in the same columns of the sequence alignment superimpose well in space. Try matching again, but this time, turn on the option to **Iterate by pruning...** and edit the angstrom value to **1.0** before clicking **OK**. This improves the superposition of the active sites by omitting dissimilar parts from the fit.

Zoom back out to view the overall structures:

**Actions... Focus**

The structures have N-terminal and barrel domains, but only the barrel domains are included in the sequence alignment. With the mouse, drag a box to highlight the entire sequence alignment. That will [select](#) the corresponding parts of the proteins. Invert the selection to the parts of the proteins *not* included in the sequence alignment, then hide the ribbon for those parts:

**Select... Invert (all models)**  
**Actions... Ribbon... hide**  
**Select... Clear Selection**  
**Actions... Focus**



Boxes drawn on the alignment are called [regions](#). A single region may consist of several disconnected boxes. The active region is shown with a dashed outline. Delete the active region by clicking in some blank area of the sequence window (not on the sequences) to make sure this window has the mouse focus, then pressing the keyboard **delete** key. If the region is not active (the border is not dashed), first click on the region to activate it, then press **delete**.

Communication also goes in the opposite direction: a [selection](#) in the structures is shown on the sequence alignment as a region. For example, select all aromatic residues in the structures:

**Select... Residue... amino acid category... aromatic**

The green boxes in the sequence alignment are the selection region. Clear the selection by Ctrl-clicking in an empty part of the graphics window.

Placing the cursor over any structure-associated residue in the sequence alignment shows the corresponding structure residue number near the bottom of the sequence window. For example, clicking into the sequence window and then placing the cursor over the first residue in the **mr** sequence shows that it is associated with Val 134 in chain A of model #0 (**2mnr**, mandelate racemase).

Next, see where some of the conserved residues are within the structures. In the sequence

window, find the first completely conserved residue in the alignment. It is an aspartic acid, D, at alignment position 99. Drag with the mouse to create a region containing just that column of the alignment. The corresponding structure residues will be [selected](#).

These aspartic acid residues are already displayed because the ribbons [preset](#) (used above) automatically displays sidechains near any bound ions or ligand molecules.

At alignment position 150, there is a completely conserved proline (P). Create another region for this column (this time, press **Ctrl** along with the mouse button to start a *new* region instead of replacing the first). Only ribbons are shown for these residues; display the sidechain atoms too:

#### **Actions... Atoms/Bonds... show**

Alignment regions can also be created automatically. From the **Multalign Viewer** menu, choose **Structure... Secondary Structure... show actual**. This creates regions in the structure-associated sequences named **structure helices** (light yellow with gold outline) and **structure strands** (light green with green outline).

The region names are listed in the [Region Browser](#) (**Info... Region Browser** in the **Multalign Viewer** menu). Clicking the **Active** checkbox for a region will [select](#) the corresponding residues in any associated structures. Only one region can be active at a time. **Close** the **Region Browser**.

Choose **Tools... General Controls... Command Line** from the main Chimera menu and use a command to close the enolase structure:

*Command:* [close 1](#)

Clear the selection (if any) by Ctrl-clicking in an empty part of the graphics window.

A structure can be colored according to the conservation in an associated sequence alignment. Choose **Structure... Render by Conservation** from the **Multalign Viewer** menu. The resulting [Render by Attribute](#) tool shows a histogram of the [residue attribute](#) named **mavConservation**. The "mav" part of the name is shorthand for **Multalign Viewer**, and "Conservation" indicates the values correspond to what is shown in the **Conservation** line of the [sequence window](#). By default, the value for a column is the fraction of sequences with the most common residue type at that position, for example: 1.0 where all 8 sequences have the same type of residue, and  $2/8 = 0.25$  where only 2 (but any 2) share a type.

The **Render by Attribute** histogram contains vertical bars (or [thresholds](#)) for mapping the attribute values to colors, radii, or worms. Use the **Colors** section. You can define your own color mapping by dragging the thresholds along the histogram and/or changing their colors. The **Value** and **Color** are shown for the most recently clicked or moved threshold. The **Value** changes when the threshold is moved, or the position can be changed by entering a value and pressing return. The **Color** can be changed by clicking the square [color well](#) and using the [Color Editor](#). Adjust the color mapping as desired before clicking **Apply**. Coloring the structure shows the high conservation of the metal-binding residues and the low conservation of most residues around the outside of the barrel.

**conservation shown with color**



Attributes can also be used in the command line. For example, show only the residues with **mavConservation** values greater than 0.7:

*Command:* [show](#) **:/mavConservation>.7**

The conservation can be calculated in different ways (entropy, variability, *etc.*) using the AL2CO program included with Chimera. Choose [Preferences... Headers](#) from the **Multalign Viewer** menu, set **Conservation style** to **AL2CO**, and adjust additional AL2CO parameters as desired. The **Conservation** line in the sequence window will update automatically. In **Render by Attribute**, it may be necessary to use **Refresh... Values** to update the histogram before recoloring the structure.

[Multalign Viewer](#) includes many features, only a few of which are sampled here. Users are encouraged to explore the menus and [documentation](#) for more information.

End the Chimera session:

*Command:* [stop now](#)

## Superpositions and Alignments Tutorial

In this tutorial, [MatchMaker](#) is used to align protein structures (create a superposition), [Match -> Align](#) is used to generate a multiple sequence alignment from the structural superposition, and [Morph Conformations](#) is used to morph between related structures.

Sequence alignments are displayed in [Multalign Viewer](#), which is covered in more detail in the [Sequences and Structures tutorial](#), and the morphing trajectory is displayed in [MD Movie](#), which is covered in more detail in the [Trajectory and Ensemble Analysis tutorial](#).

Internet connectivity is required to fetch the structures used in this tutorial: **1tad, 121p, 1r2q, 1j2j, 1puj, 1tnd, 1tag**

- [Background and setup](#)
- [Different but related proteins](#)
  - [Superposition](#)
  - [Structure-based sequence alignment](#)
- [Different conformations of the same protein](#)
  - [Morphing](#)

### ← Background and Setup

Protein structures are classified within databases such as [SCOP](#), [CATH](#), and [HOMSTRAD](#). Classifications range from groups of highly similar and closely related proteins to larger, more diverse sets. For analysis and comparison, it is often useful to superimpose related structures. Although it is not always clear whether proteins with the same fold are evolutionarily related (homologous), they should still be superimposable. In general, more closely related proteins are easier to superimpose.

**G proteins** (guanine nucleotide-binding proteins) are used as examples. G proteins are important in signal transduction. They act as molecular switches, changing conformation and interaction partners depending on whether GTP or GDP is bound. Many diverse structures are known. The two main subsets are the small monomeric G proteins, such as Ras, and the larger heterotrimeric G proteins, which act immediately downstream of G-protein-coupled receptors. The  $\alpha$  subunits of heterotrimeric G proteins are homologous to the small G proteins.

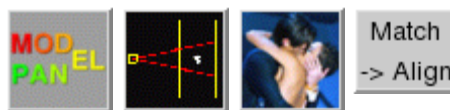
On **Windows/Mac**, click the **chimera** icon; on **UNIX**, start Chimera from the system prompt:

```
unix: chimera
```

A splash screen will appear, to be replaced in a few seconds by the main Chimera [graphics window](#) or [Rapid Access](#) interface (it does not matter which, the following instructions will work with either). If you like, resize the Chimera window by dragging its lower right corner.

Show the [Command Line](#) (**Tools... General Controls... Command Line**). Choose **Favorites... Add to Favorites/Toolbar** to place some icons on the toolbar. This opens the [Tools section](#) of the preferences, which lists the same tools as Chimera's **Tools** menu. In the **On Toolbar** column, check the boxes for:

- [Model Panel](#) (under General Controls)
- [Side View](#) (Viewing Controls)
- [MatchMaker](#) (Structure Comparison)
- [Match -> Align](#) (Structure Comparison)



You can also specify tools such as the [Command Line](#) to **Auto Start** (start when Chimera is started). If you want these settings to apply to subsequent uses of Chimera, click **Save** before closing the preferences.

Fetch a structure from the [Protein Data Bank](#):

*Command:* [open 1tad](#)

The structure contains three copies of the  $\alpha$  subunit of transducin, a heterotrimeric G protein. Delete solvent and two of the copies, chains B and C:

*Command:* [del solvent](#)

*Command:* [del :.b-c](#)

Move and scale the structures [using the mouse](#) and [Side View](#) as desired throughout the tutorial.

## ← Different but Related Proteins

We will [superimpose structures](#) of a sample of G proteins, then [create a sequence alignment](#) from the superposition.

The  $\alpha$  subunit of the heterotrimeric G protein transducin was already opened in the [setup](#). Fetch structures for the monomeric G proteins H-Ras, Rab5a, and ADP-ribosylation factor 1, respectively:

*Command:* [open 121p](#)

*Command:* [open 1r2q](#)

*Command:* [open 1j2j](#)

Use the ribbons [preset](#) (which may or may not change the appearance, depending on your preference settings):


*Menu:* **Presets... Interactive 1 (ribbons)**

This preset displays ribbons plus ions, ligands, and nearby sidechains.

## ← Superposition

### superimposed G proteins

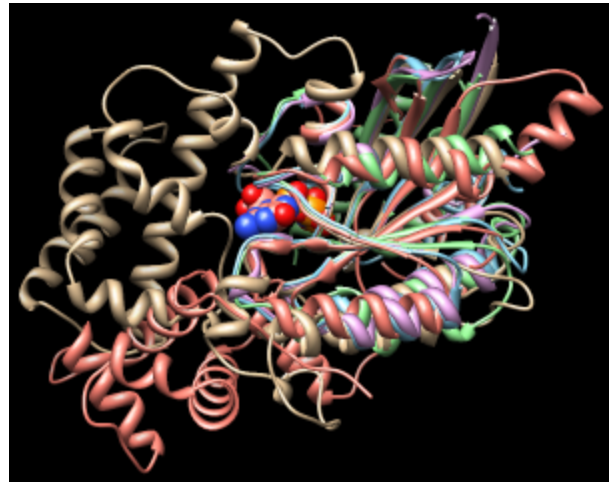
The structures need to be superimposed so that they can be compared. Start [MatchMaker](#) by

clicking its [icon](#): 

[MatchMaker](#) superimposes structures pairwise by first aligning their sequences and then fitting the  $\alpha$ -carbons of residues in the same columns of the sequence alignment. Usually the fit is iterated so that residue pairs aligned in sequence but far apart in space are not used in the final 3D match.

Several parameters control the sequence alignment step:

- algorithm - Needleman-Wunsch (global; default) or Smith-Waterman (local)
- scoring function -
  - residue similarity matrix (default BLOSUM-62)
  - whether secondary structure information should be used (default yes)
  - weighting of the secondary structure and residue similarity terms (default 30% and 70%, respectively)
  - gap penalties



Click **Reset to defaults** (near the bottom of the dialog) to ensure that the default settings will be used. All of the structures should already be chosen as the **Structure(s) to match**; keep that the same, but choose **1tad** as the **Reference** and click **Apply** to match all the others to it.

The number of  $\alpha$ -carbon pairs and RMSD in the final iteration of each pairwise fit are reported in the **Reply Log** (in the menu under **Favorites**). However, simple visual inspection of the overall structures is often the most useful indicator of success.

Another visual indicator is how well similar ligands superimpose. Show only residues classified as [ligand](#), and label them:

Command: [show ligand](#)

Command: [rlab ligand](#)

Each of these structures includes GTP or an analog of GTP in the binding site. However, some other ligands were simply present in the crystallization solution and are not biologically relevant. GOL is glycerol and can be removed:

Command: [del :gol](#)

Command: [~rlab](#)

Try using different reference structures in **MatchMaker** (click a line in the **Reference structure** list, click **Apply**). With the default alignment parameters, the superposition is similar and basically correct no matter which structure is used as the reference. Detailed examination of the match statistics and guanine nucleotide positions suggests results may be slightly better with **1r2q** as the reference.

Next, try a structure that is harder to superimpose, and display its ligand in the sphere representation:

Command: [open 1puj](#)

Menu: **Presets... Interactive 1 (ribbons)**

Command: [show ligand](#)

Command: [repr sphere ligand & #4](#)

Besides lacking sequence similarity, this protein is **circularly permuted** compared to the others: its N-terminal part structurally matches the C-terminal part of other G proteins and vice versa.







resurrected by choosing **MAV - alignment-name... Raise** from near the bottom of the **Tools** menu. In **Multalign Viewer**:

- choose **Preferences... Appearance** and adjust settings for **Multiple alignments** as desired
- use **Info... Percent Identity** to compare **all sequences** with **all sequences**, confirming that the pairwise identities are <30% for these examples
- use **Edit... Delete Sequences/Gaps** to delete the sequence named **2 x 1puj, chain A** and any resulting all-gap columns

Now the alignment clearly shows the large insertion in  $\alpha$ -transducin (**1tad**) relative to the small monomeric G proteins. Select and display residues that are completely conserved in the sequence alignment:

Command: **sel** **:/mavPercentConserved=100**

Command: **disp sel**

Some of the conserved residues are Gly (no sidechain). Clear the selection by Ctrl-clicking in an empty area of the graphics window.

### ← Different Conformations of the Same Protein

(To jump to this section right after performing the [setup](#), open the sequence alignment file [4gees.afa](#) included with this tutorial.)

Now we will compare **1tad** with different structures of the same protein, transducin- $\alpha$ :

Command: **open 1tnd**

Command: **open 1tag**

Delete solvent and chains B-C (extra copies in **1tag**):

Command: **del solvent**

Command: **del ::b-c**

If **Multalign Viewer** (the sequence alignment window) is hidden, bring it to the front by choosing **MAV - alignment-name... Raise** from near the bottom of the **Tools** menu.

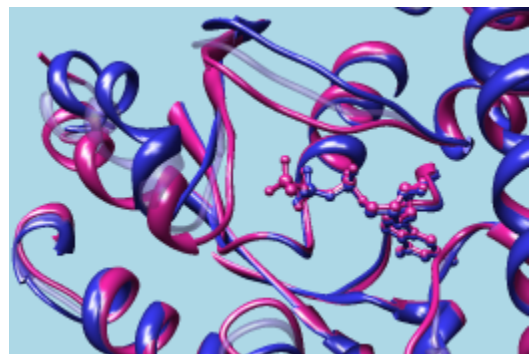
**Multalign Viewer** displays lines of information called **headers** above the sequences in the alignment. Use the **Headers** menu to hide **Consensus** and **Conservation** and to show **RMSD**, if not already shown. The sequence name **1tad, chain A** has a dashed green line around it, indicating that the sequence is associated with multiple structures. The **RMSD** header shows the spatial variability of residues associated with each column ( $\alpha$ -carbon root-mean-square distance); currently, it contains high values everywhere because the structures are not all superimposed.

To superimpose the structures using the sequence alignment, choose **Structure... Match** from the **Multalign Viewer** menu. One structure (it does not matter which) should be designated as the reference, and all three can be designated as the structures to match. Check the option to **Iterate by pruning...** using a 2.0-Å cutoff and click **OK**. The **RMSD** header is automatically recomputed, showing much lower values.

Superposition of proteins with the same or nearly the same sequence is generally trivial. We used

#### GTP-binding switch

(**1tagA**, **1tndA**, morph intermediate)



**Multalign Viewer** since we already had a sequence alignment, but **MatchMaker** (or its [command equivalent](#)) or the command **match** could have been used instead. These other methods are used and discussed in the [Structure Analysis and Comparison tutorial](#).

Use the ribbons preset (which may or may not change the appearance, depending on your preference settings) and focus on the ligand residues:

Command: **preset apply int 1**

Command: **focus ligand**

Command: **r1ab ligand**

Open the **Model Panel** and use the **S(hown)** checkboxes to view the structures individually.

The **1tad** structure (tan) represents the activated form of a G protein; even though it includes GDP, the GDP and ALF (AlF<sub>4</sub><sup>-</sup>) residues together mimic the transition state of GTP hydrolysis. **1tnd** (light blue) contains the GTP analog GSP and also represents the activated form. The third structure, **1tag** (purplish pink), includes GDP and represents the nonactivated form.

Use the **Model Panel** checkboxes to show all three structures together. Remove the labels and focus on the overall structures:

Command: **~r1ab**

Command: **focus**

Although the structures are mostly similar, the nonactive conformation (pink) differs from the activated ones (tan and light blue) in specific areas, termed **switch regions**.

In the sequence alignment window, the three most prominent "humps" in the **RMSD** header correspond to the known G protein switch regions at approximately residues 173-183, 195-215, and 227-238 of transducin- $\alpha$ . The third switch region is unique to heterotrimeric G proteins; it is an insertion relative to the monomeric G proteins. Placing the cursor over a position in the **1tad** sequence lists the associated structure residues near the bottom of the sequence window, and drawing a box around residues in the sequence alignment (click to start, drag to expand) [selects](#) the associated parts of the structures.

Close **1tad**:

Command: **close 0**

The RMSD histogram looks much the same; now it simply shows the CA-CA distances between the two remaining structures, **1tnd** representing the activated form and **1tag** representing the nonactivated form.

## ← Morphing

Finally, morph between the two structures. Morphing involves calculating a series of intermediate structures. In Chimera, the series of structures is treated as a trajectory that can be replayed, saved to a coordinate file, or saved as a movie using [MD Movie](#).

Start the [morphing tool](#):

Command: **start Morph Conformations**

Click **Add...** and in the resulting list of models, doubleclick to choose #2, #1, and #2 again,

corresponding to a morph trajectory from the nonactivated structure to the activated and back. **Close** the model list. In the main **Morph Conformations** dialog, set the **Action on Create** to **hide Conformations**, and then click **Create**.

The progress of the calculation is reported in the [status line](#). When all the intermediate structures have been calculated, the input structures are hidden, the trajectory is opened as model #0, and the [MD Movie](#) tool appears.

The trajectory can be played continuously or one step at a time using the buttons on the tool. If the player dialog becomes obscured by other windows, it can be resurrected by choosing **MD Movie - trajectory-name... Raise** from near the bottom of the **Tools** menu. If you want to see the original structures again, use the **S(hown)** checkboxes in the [Model Panel](#).

When you have finished viewing the morph trajectory, choose **File... Quit** from the menu to exit from Chimera.

# Molecular Visualization with UCSF Chimera: Question & Answer Session

- New features in Chimera 1.7:
  - Improved graphics
    - performance increase for newer video cards
    - higher default frame rate for smoother motions
    - glossy lighting
    - interactive shadows
  - New tools
    - PDB2PQR - assignment of atomic partial charges and radii
    - APBS - continuum electrostatics calculation
    - AutoDock Vina - molecular docking
    - Model/Refine Loops - interface to Modeller
    - Notepad - save user-entered descriptive text with sessions
  - New image/movie features
    - export in HTML/WebGL format for interactive web viewing
    - Animation tool for creating simple movies
- Future directions:
  - Chimera 2.0 - web- and desktop-based application
    - Web-based application using modern web standards (HTML5, JavaScript, WebGL)
    - Desktop-based application using modern cross-platform toolkits (Qt, OpenGL)
    - Performance from parallelization (threads, GPU)
    - Integration with web services (databases, computation, storage)
  - interface to cloud computing infrastructure such as Amazon Web Services
    - Access to extensive computational resources
    - Minimize cost for hardware setup and maintenance
    - Pre-made system images for common tasks (molecular dynamics, virtual screening)
- What is a feature you would like to see in Chimera?
  - Tell us now
  - Mail suggestions to [chimera-users@cgl.ucsf.edu](mailto:chimera-users@cgl.ucsf.edu) later