Practice Session 1:
3D structure visualization and high quality imaging using UCSF Chimera

Antoine Daina, Ute Röhrig, Vincent Zoete – 2023
Contact : vincent.zoete@sib.swiss
antoine.daina@sib.swiss

Presentation of UCSF Chimera

This introductory practice about 3D structure visualization and high quality imaging for publication will make use of the free program UCSF Chimera. The official website of Chimera can be found at the following address: https://www.cgl.ucsf.edu/chimera

Here is a brief description of Chimera and its features taken from the website:

" UCSF Chimera is a highly extensible program for interactive visualization and analysis of molecular structures and related data, including density maps, supramolecular assemblies, sequence alignments, docking results, trajectories, and conformational ensembles. High-quality images and animations can be generated. Chimera includes complete documentation and several tutorials, and can be downloaded free of charge for academic, government, non-profit, and personal use. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics, funded by the National Institutes of Health (NIGMS P41-GM103311). "

The program can be downloaded at the following address:
https://www.cgl.ucsf.edu/chimera/download.html

When using UCSF Chimera, one should cite this reference:
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. "UCSF Chimera - A Visualization System for Exploratory Research and Analysis." *J. Comput. Chem.* **25**(13):1605-1612 (2004).

The following exercises will be used as tutorials to introduce some basic commands of Chimera..

You are highly encouraged to test the different options that you will encounter in the different menus.

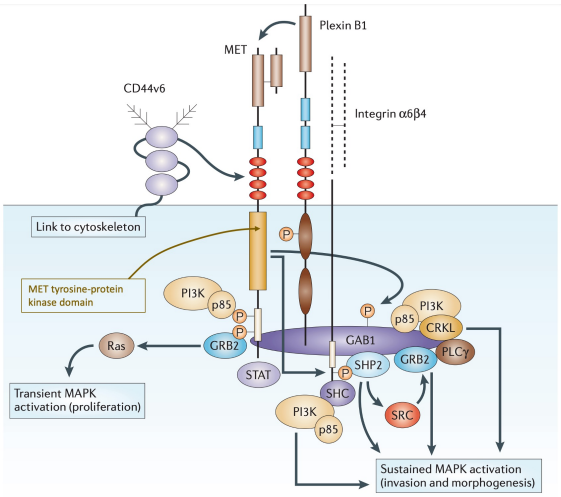
Grey boxes contain commands or examples to be tested during the practice

Pink boxes contain very important notes

Yellow boxes contain notes that allow users to explore possibilities beyond those exposed in the main practice.

Exercise 1 – First steps with UCSF Chimera

The 3D structure of the tyrosine-protein kinase domain of the Hepatocyte Growth Factor Receptor, a.k.a. HGFR or c-MET, in complex with the FDA-approved drug crizotinib (2GWJ in PDB), obtained by X-ray crystallography, will be used as an example.
MET is essential for embryonic development, organogenesis and wound healing. However, abnormal activation of MET or the MET pathway plays an important role in the development of cancer through activation of downstream key oncogenic pathways (e.g. Ras or PI3K pathways). Consequently, MET constitutes an attractive oncology target for therapeutic intervention. Several MET inhibitors were designed, including crizotinib. The latter was approved by the FDA in 2011.

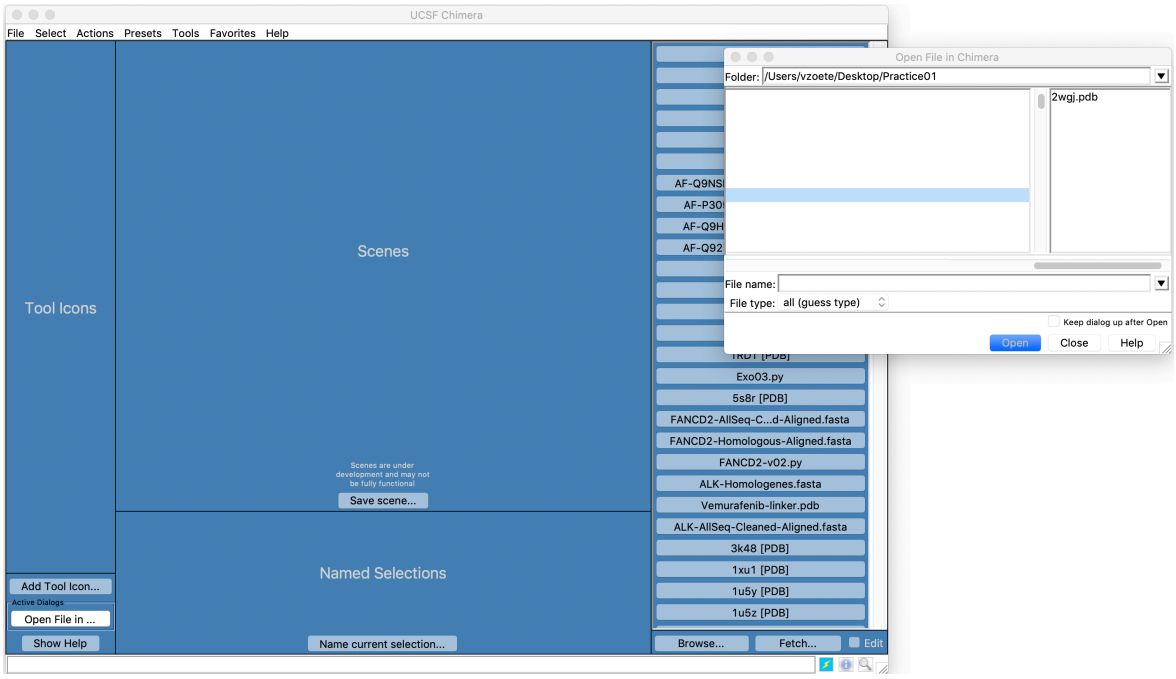


MET pathway. Adapted from *Nature Reviews Cancer* volume 6, pages 637–645 (2006)

Loading a structure into Chimera

Two methods can be used to load a structure into Chimera.

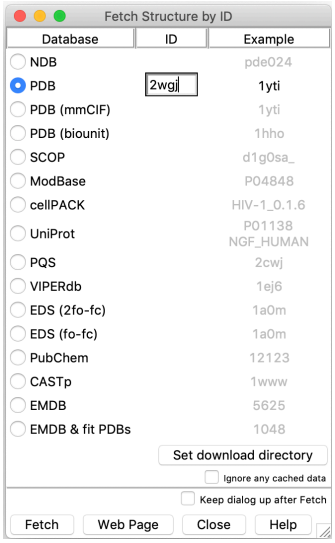
1) If the structure file is present in the user's computer, choose the menu item "**File/Open**". Then select "2wgj.pdb" and click "**Open**".



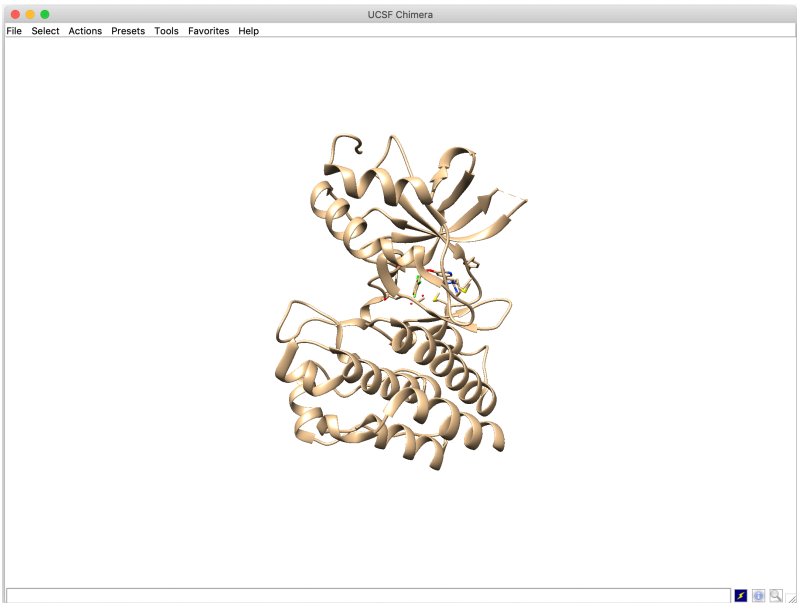
Note: *Under Linux or Mac OSX, the structure (for instance 2wgj.pdb) can also be loaded using the following commands:*

```
(standard Linux)   > chimera 2wgj.pdb
or
(standard macOS)   > /Applications/Chimera.app/Contents/MacOS/chimera 2wgj.pdb
```

2) The structure can also be fetched directly from a database over the internet, when available. Choose the menu item “**File>Fetch by ID…**”. Choose the PDB databank and type the PDB code “2wgj”. Then click “**Fetch**”



Once the structure has been loaded, the complex between MET and crizotinib should appear in a representation automatically chosen by UCSF Chimera:



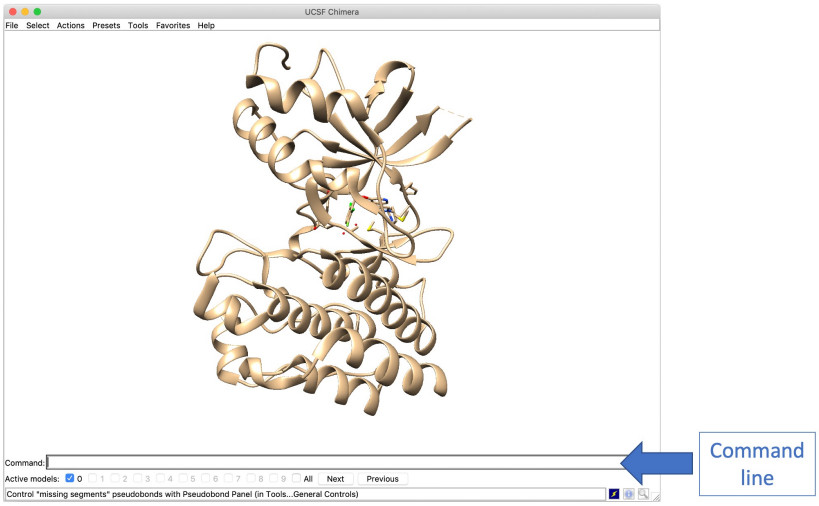
Moving / zooming

The macromolecule can be rotated by clicking the left button of the mouse (or the corresponding gesture on the trackpad) and dragging the cursor over Chimera's window. The translation is obtained by using the central mouse button. Finally, one can zoom in and out using the right mouse button.

Rotate: press left button and drag
Translate: press central button and drag
Zoom: press right button and drag

Command line...

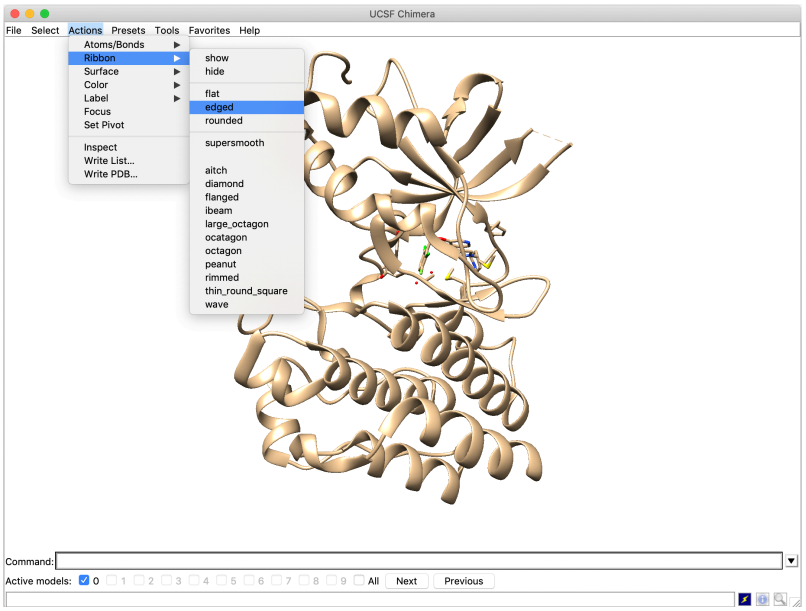
Most of Chimera functionalities can be accessed by typing a given command in the "**Command line**", which is situated at the bottom of the Chimera window. If the command line is not present, display it using "**Favorites>Command Line**". A line preceded by "Command:" will appear on the lower part of the main window. Examples of how to use the command line will be given later.



Ribbon representation

The secondary structure elements can be shown using the ribbon representation. This can be accessed through the "**Actions>Ribbon"** sub-menu.

The ribbon representation of the secondary structure elements can be switched on or off by choosing "**show**" or "**hide**" in the "**Actions>Ribbon**" menu, respectively. Three main variants are available: "**flat**", "**edged**" and "**rounded**". You can try them all by clicking the corresponding menu item. Select the representation you prefer.

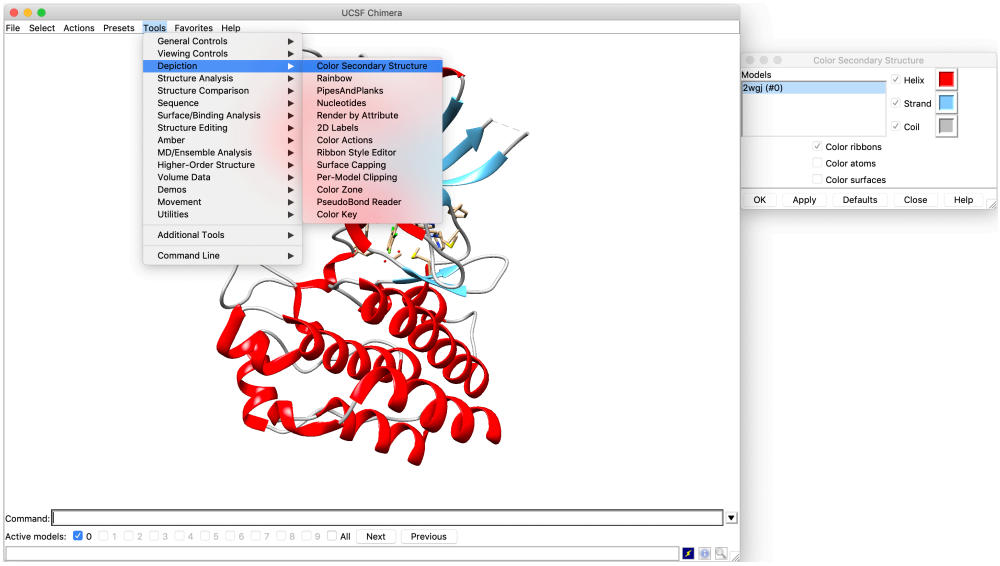


The following images will be generated using the "**rounded**" option.

Note: when necessary, you can type "**ksdssp**" in the command line to recalculate the secondary structure elements of the protein. This uses an implementation of the **K**absch and **S**ander algorithm for **d**efining the **s**econdary **s**tructure of **p**roteins. Biopolymers 22:2577 (1983).

Coloring the secondary structure elements.

The ribbon can be colored according to the secondary structure element, i.e. strand, helix or loop. Select "**Tools>Depiction>Color Secondary Structure**". This will open the "**Color Secondary Structure**" window. Check the "**Helix**", "**Strand**" and "**Coil**" boxes.



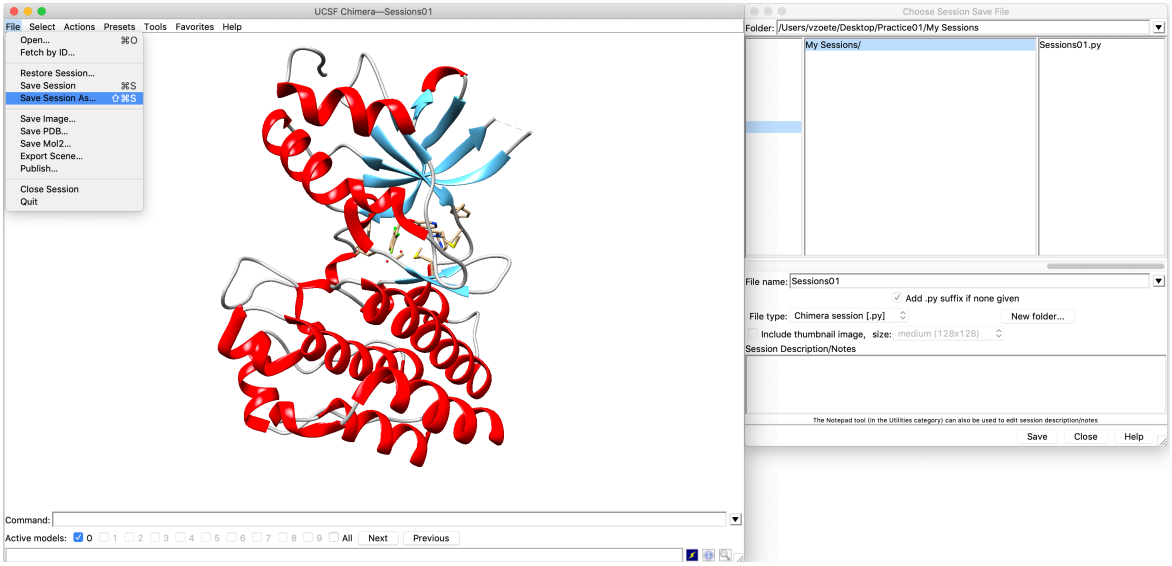
For each secondary structure element, it is possible to modify the default color by clicking on the corresponding colored square to open the "**Color Editor**" window and change the RGB cursors positions. Click on the "**Apply**" button of the "**Color Secondary Structure**" window to apply the coloring.

You can try several color combinations and keep the one you prefer. Finally, you can close the "**Color Secondary Structure**" and "**Color Editor**" windows by clicking the "**Close**" button, or the top "**X**" icon of the window itself.

Note: The RGB Color Editor is available to choose the color of many structural elements, text labels, etc. It is generally accessible when a color is displayed within a square, like in the example above. In this case, just click within the square to change the color.

Saving the session status

The Chimera session (the actual state and representation) can be saved for future use or modifications, or to share it with a collaborator. To do so, use the "**File>Save Session As...**" menu. Select a location on your hard drive, type or select a file name, and click "**Save**". The saved file is actually a python script and will have the ".py" extension.



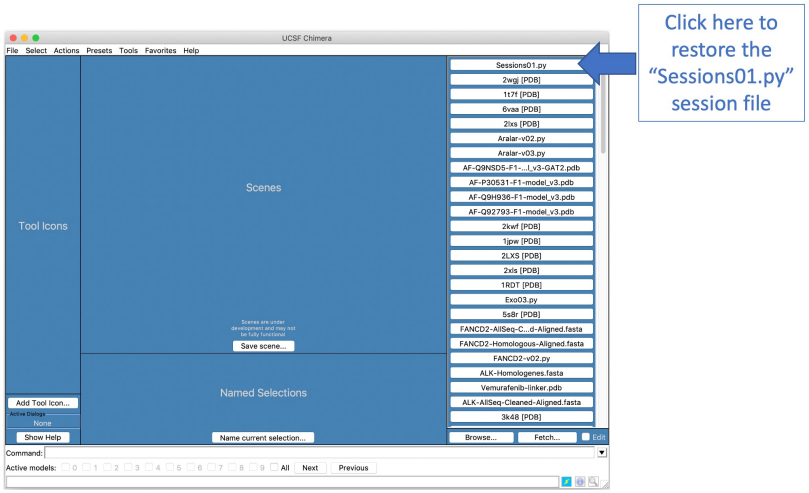
Note: There is no undo button in UCSF Chimera. **The best way to mimic an undo action is to save a session file regularly**, or before every action that is potentially deleterious to the representation (atom deletion, actions on massive selections, etc.). In case of error, the session file can be loaded to revert the presentation to the previous state (see below).

Closing the session.

The session can be closed using the "**File>Close Session**" menu item.

Restoring a previous session

A previously saved session can be restored using the "**File>Restore Session**" menu item. Select the file (with a .py or .pyc extension) and click "Open". The session will be restored exactly in the state it was at the last "Save session" action.
As an alternative, it is also possible to restore a recent session file by clicking on its name, in the right column of an empty chimera window.



Note: **Session files are system independent.** It is therefore possible to share a session file with a collaborator who has a different machine, with a different Operating System (Mac OSX, Windows, Linux…). The collaborator will see exactly the presentation you prepared on his/her own screen.
However, these session files are dependent on the version of Chimera used: a session file created with a recent version of Chimera might not be correctly restored by an older version of the software.

Changing the display of bonds and atoms. Selections

The objective of this part is to display the global structure of MET, as well as crizotinib and some of the surrounding residues of MET. Other residues will be hidden to simplify the representation. This will require to select parts of the structure and to apply different actions and representation schemes to them.

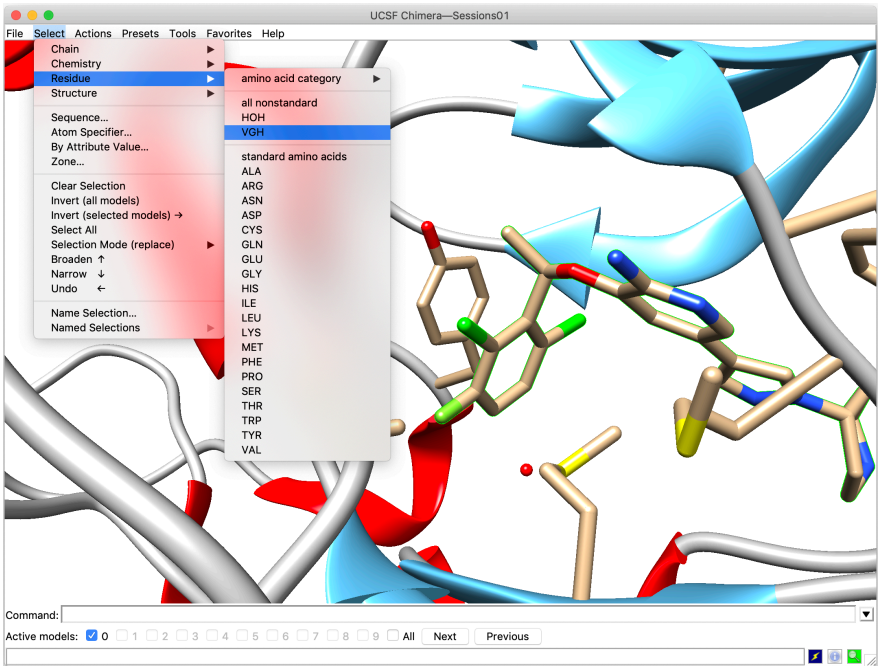
Note: UCSF Chimera allows to display differently each atom, bond, secondary structure element, etc. according to the user's needs to obtain the final desired representation. To achieve this, it is possible to select each of these elements independently to apply a different representation scheme to it. Item selection is therefore an essential function in UCSF Chimera, and several alternatives exist.

There are three ways of selecting atoms with Chimera: using the "**Select**" menu, using the "**Command Line**", or using **the mouse** to select atoms from the screen.

Selecting using the select menu

Open the "**Select**" menu and verify that the "**Selection Mode**" is "**replace**". Otherwise, choose it.

The first four menu items allow to select part of the structure according to the "**Chain**" (i.e. a protein chain), the "**Chemistry**" (chemical nature of the atoms or functional groups), the "**Residue**" (residue name or type) and the "**Structure**" (ligand, water, ions, backbone, side chains, secondary structure, etc...). You can try to select different parts of the complex. For instance, you can select crizotinib by choosing the "**Select>Residue>VGH**" menu item. You will see that the selected parts of the structures are surrounded by thin green lines.



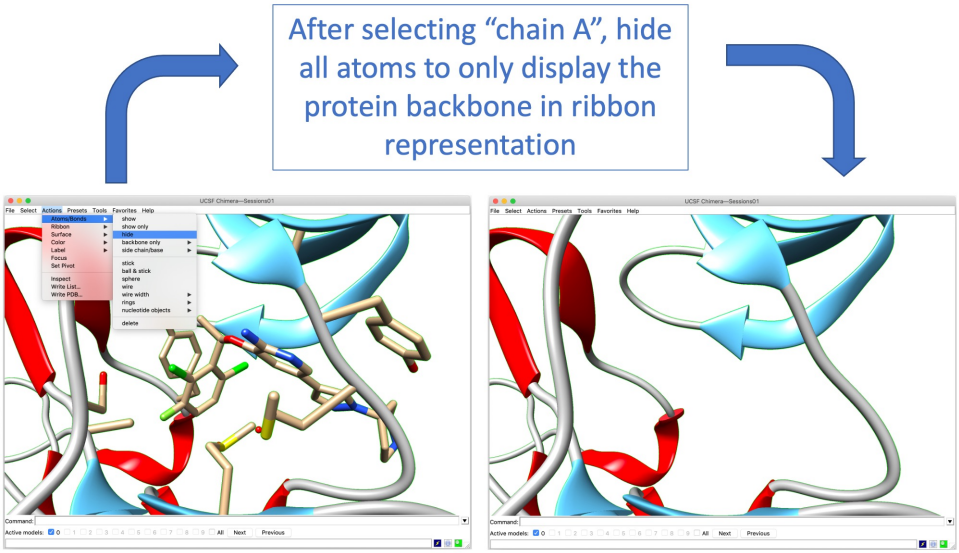
You can try selecting different parts of the complex: the water molecule, chain A of the system (i.e. in this case, the protein, the ligand and the water molecules), the aspartate residues, the strands, etc...

Finally, you can clear all selections using "**Select>Clear Selection**".

Note: Most of the options of the "**Select**" menu, like the list of available protein chains or the list of residues present in the structure, are dynamically created by UCSF Chimera upon analysis of the content of the loaded PDB file. Therefore, they will differ between PDB files.

Deleting and hiding atoms

Select chain A (which groups the entire complex in this particular case) by selecting "**Select>Chain>A**". Then, hide all the selected atoms using "**Actions>Atoms/Bonds>hide**".

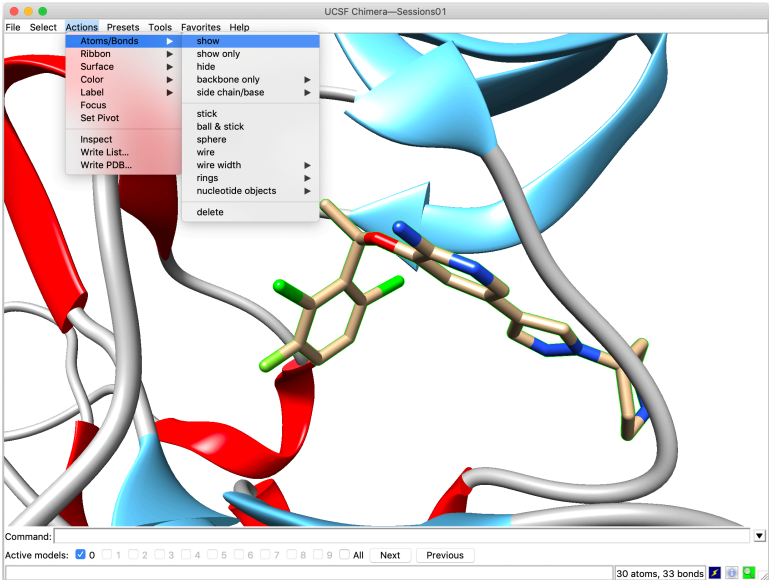


Note: There is a fundamental difference between deleting and hiding atoms in Chimera.
- **"delete" definitively suppresses the atoms from the system**. It will not be possible to show them again, unless the user re-loads their coordinates into Chimera
- **"hide" undisplays the atoms from the current representation but keeps them in the system**. Consequently, these atoms can still be selected and shown again.

Note: When no structural element is selected, commands are applied to all atoms.

Now that the representation has been cleared and simplified, we can display crizotinib. For this, select the residue VGH using "**Select>Residue>VGH**" and display it using "**Actions>Atoms/Bonds>show**".

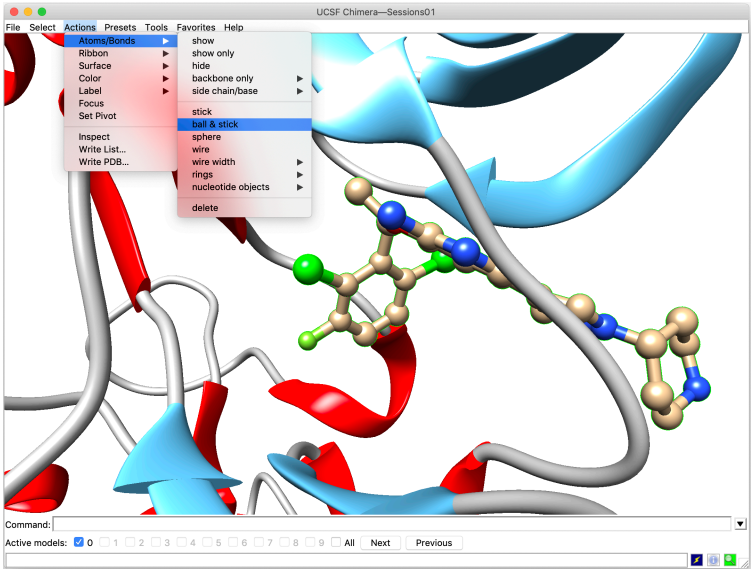
Indeed, VGH is the three-character code of crizotinib in the PDB, as can be seen in https://www.ebi.ac.uk/pdbe/entry/pdb/2wgj and https://www.ebi.ac.uk/pdbe-srv/pdbechem/chemicalCompound/show/VGH.



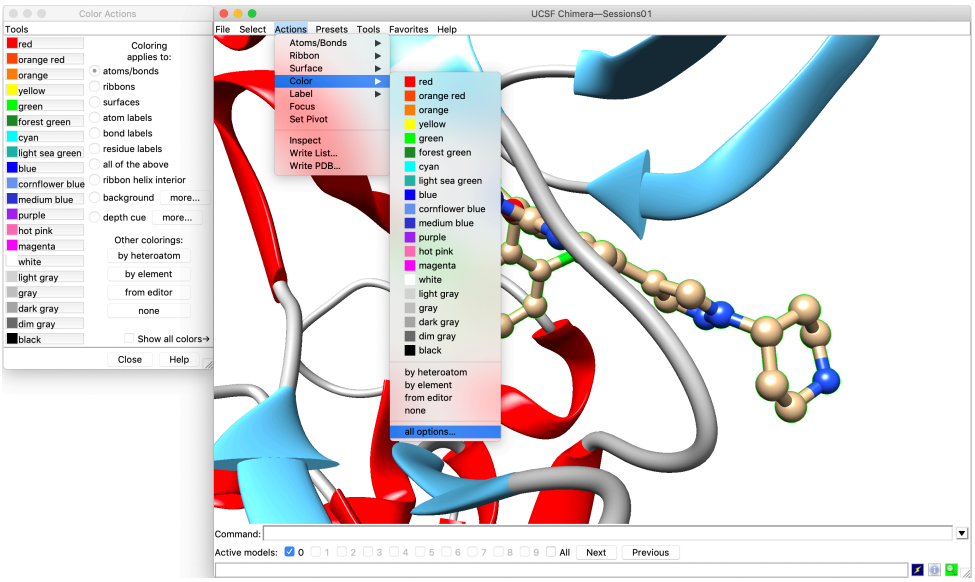
Changing bond display and color. Changing background color.

Verify that crizotinib (residue VGH) is still selected. If so, it should be surrounded by a thin green line. If needed, select it again using "**Select>Residue>VGH**".

In the "**Actions>Atoms/Bonds**" menu, choose successively the "**stick**", **ball & stick**", "**sphere**" and "**wire**" menu items and see how the ligand is displayed in each case. Finally, choose the "**ball & stick**" representation.

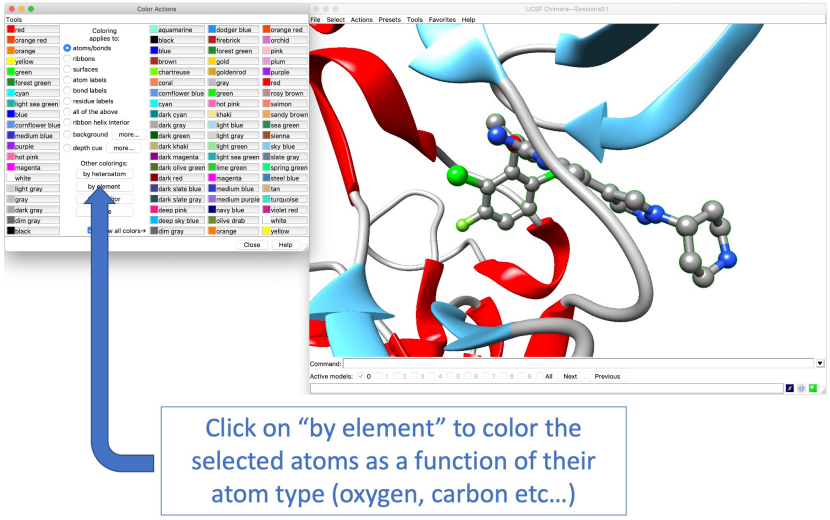


Open the detachable color action menu using "**Actions>Color>all options…**". Check the "**atoms/bonds**" button so that the color changing will be applied only to atoms and bonds.



Then, you can select a color from the left column that will be used for all ligand atoms. More colors can be obtained after checking the "**Show all colors**" box. In addition, users can define their own colors by clicking on "**from editor**", which opens the RGB Color Editor.

Once you have tested different possibilities, color all atoms according to their atom types. This is obtained by clicking "**by element**".



You can also check the "background" button in the "Color Actions" menu to change the color of the background only. Generally, we use white or black for better clarity and contrast.

You can close the "Color Actions" menu.

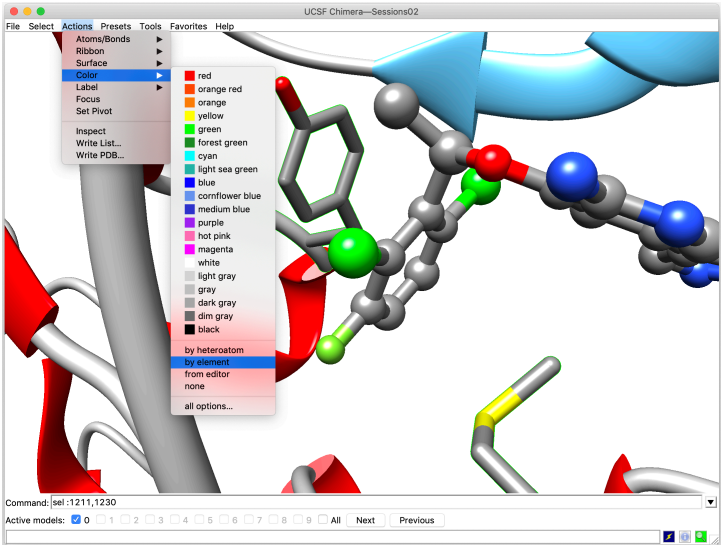
Selecting using the command line

If it's not already done, open the command line by choosing "**Favorites>Command line**". A line preceded by "Command:" will appear on the lower part of the main window. This line can be used to type and execute commands relative to selection and display, for instance. To execute a command, one needs to type it in the command line and **press Return**.
The union (logical 'OR') and intersection (logical 'AND') selection keywords are noted "|" and "&", respectively. The negation symbol is "~".

A detailed description of the different selection possibilities using the command line are provided in the UCSF Chimera quick reference guide (https://www.cgl.ucsf.edu/chimera/current/docs/UsersGuide/quickref.pdf). Here is a limited list showing some possible selections based on our particular structure that you could try:

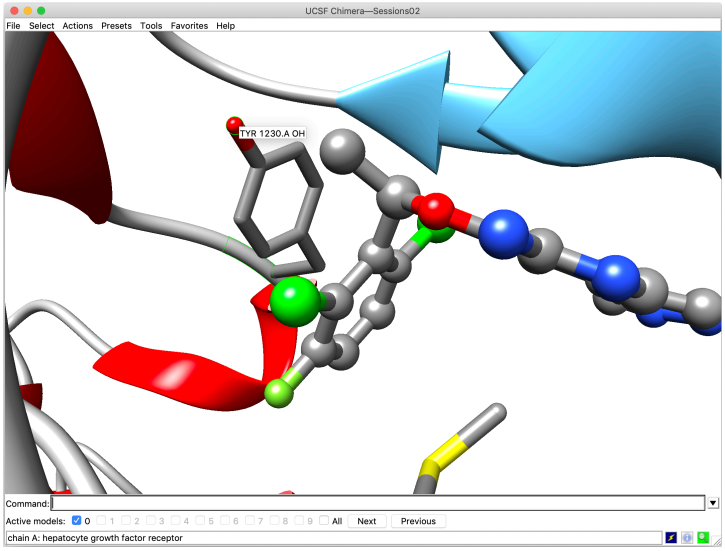
"select" : select everything.
"select :.A" : select chain A of the protein.
"select :.A,.B" : select chains A and B of the protein (if a chain B is present)
"**select ligand**" : select the ligand.
"**select :VGH**": select the residue named VGH. Another way to select the ligand.
"**select :HOH**" : select the residues named HOH, i.e. the water molecules.
"**select :HOH l ligand**" : select both the water molecules and the ligand.
"**select :1230**" : select all residues that are numbered 1230 in the PDB.
"**select :1230@CA**": select atom Cα of residues 1230.
"**select :1230.A**" : select residue 1230 of chain A only (useful if several chains are present)
"**select :1211,1230**" : select all residues that are numbered 1211 and 1230 in the PDB.
"**select strand**" : select strands.
"**select :VGH z<5**" select all atoms that are within 5 Å from residue VGH.

Select residues 1211 and 1230 by typing "**select :1211,1230**" in the command line, and show them in stick representation, colored according to the atom type: "**Actions>Atoms/Bonds>show**", "**Actions>Atoms/Bonds>stick**", "**Actions>Color>by element**".



Selecting using the mouse and the keyboard

To select one atom that is currently displayed on the screen, it is possible to press the "**Control**" key on the keyboard while performing a **left-click** with the mouse on the atom of interest. Use this technique to select the oxygen atom of the Tyr1230 side chain of MET.



This oxygen atom (named "TYR 1230.A OH") is now surrounded by a thin green line.
Now that this atom has been selected, you can perform the following commands to obtain different selection schemes:

- Pressing the "**arrow up**" key on the keyboard will select the entire residue Tyr1230
- Pressing again "**arrow up**" key will select the secondary structure element to which Tyr1230 belongs (a loop in this case)
- Pressing again "**arrow up**" key will select the entire protein chain to which Tyr1230 belongs (in this case chain A, excluding ligand)
- Pressing again "**arrow up**" key will select the entire system (including proteins, ligands, etc…)
- Pressing the "**arrow down**" key will change the selection down to the protein chain to which Tyr1230 belongs
- Pressing "**arrow down**" key again will change the selection down to the secondary structure element to which Tyr1230 belongs
- Pressing the "**arrow down**" key again will select down residue Tyr1230
- Finally, pressing the "**arrow down**" key again will select down the atom OH of Tyr1230

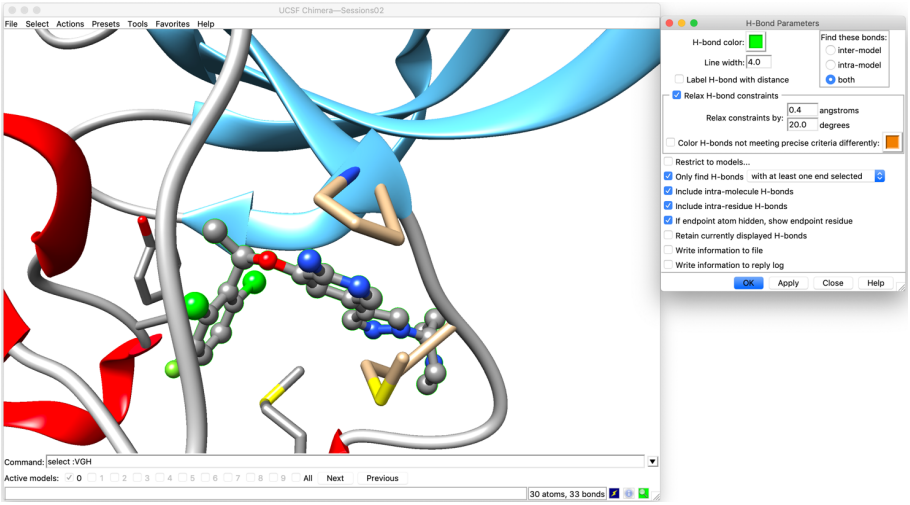
Note: It is possible to select several atoms using the keyboard and mouse. Select the first atom using the "**Control**" key and the **left mouse click** as described above, then press **simultaneously the "Control" and "Shift" keys** of the keyboard and perform a **left click with the mouse** on the second atom. You can select as many atoms as you wish with this procedure.

It is also possible to clear the selection by using the "Control" key and clicking with the left mouse button in a void space (in the background).

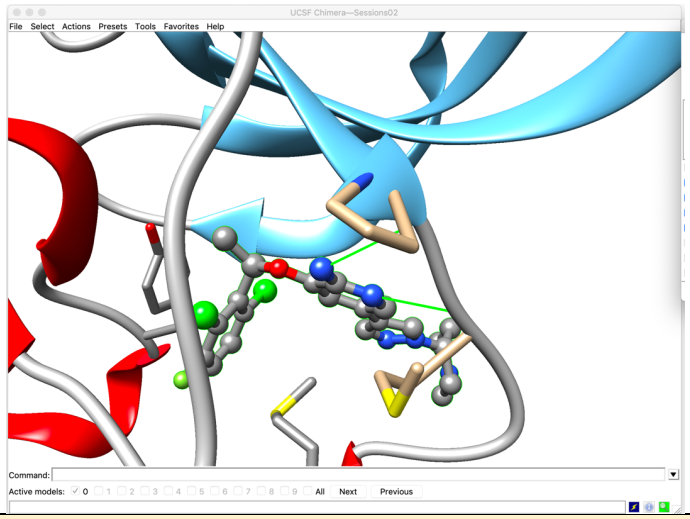
Note: Pressing the "**Control**" key on the keyboard while holding the **left button of the mouse and dragging it** on the Chimera window will select all atoms within the rectangle defined by the dragging of the pointer. Note that the selection is done in 3D.

Calculating and showing hydrogen bonds

Chimera can calculate and display hydrogen bonds between given atoms.
To show the hydrogen bond network between the ligand and MET, the following sequence of actions can be performed. Select crizotinib using the command line, the selection menu or the keyboard/mouse technique. Open the "**Tools>Surface/Binding Analysis>FindHBond**" window. Check the "**Only find H-bonds with at least one end selected**" box, so that it will calculate only hydrogen bonds involving crizotinib. Also check the box "**If endpoint atom hidden, show endpoint residue**" to display all residues that are making a hydrogen bond with crizotinib, even if they are currently hidden (this is very useful not to miss any hydrogen bond).



Finally, click on the "**Apply**" button. Hydrogen bonds will be displayed by thick green lines.

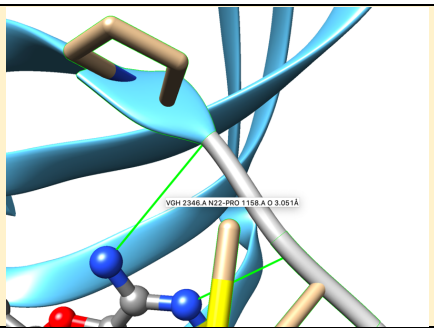


Note: You can change the hydrogen bond color and line width from the "**H-Bond Parameters**" window and click "**Apply**" to apply them.

Hiding and displaying ribbon representation for selected residues

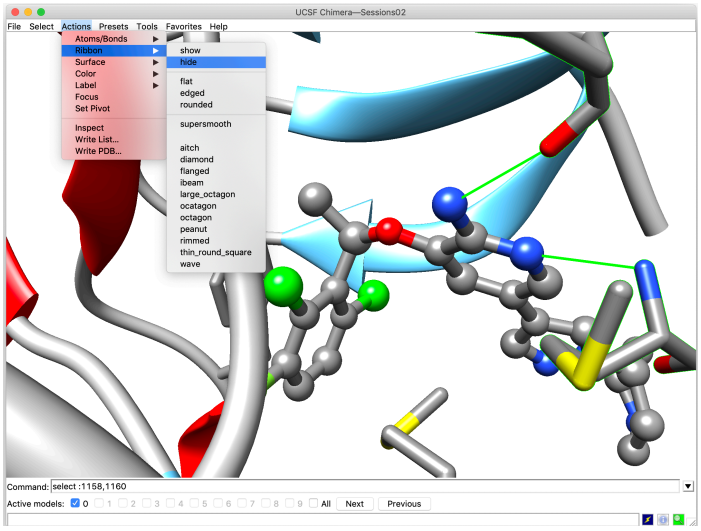
The "**FindHBond**" function of Chimera has found hydrogen bonds between crizotinib and Pro1158 of MET as well as between crizotinib and Met1160 of MET. However, since the ribbon representation is applied by default to all protein residues, the backbone atoms are hidden (only the side chain atoms of these two residues are displayed). To better see these hydrogen bonds, it is therefore necessary to switch off the ribbon representation for these two residues.

Note: If you leave the pointer of the mouse over the thin green line of a hydrogen bond, information regarding the atoms involved will be displayed:



For showing backbone atoms, select these two residues. This can be done using the mouse as seen above, or by typing "**select :1158,1160**" in the command line. Once the residues are

selected, hide their ribbon representation by selecting "**Actions>Ribbon>hide**". Keep these two residues selected, and color them according to their elements.



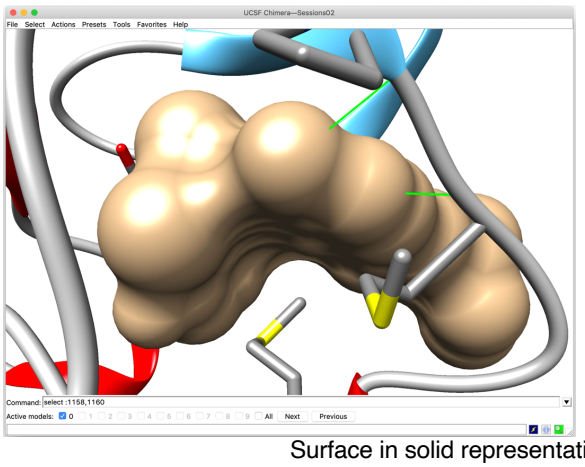
To show again the ribbon representation for these two residues, select them (if not already done) and click on "**Actions>Ribbon>show**".

Showing the ligand surface

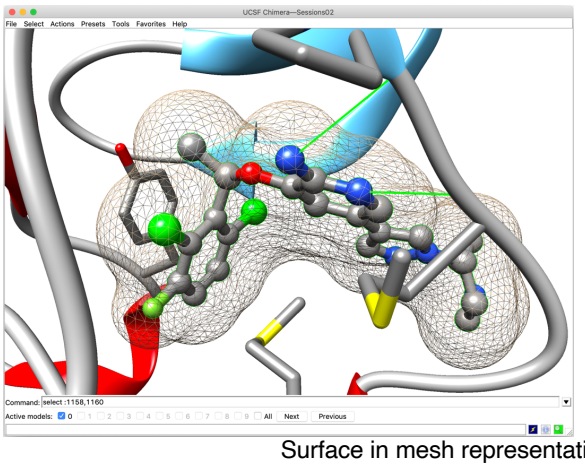
Select crizotinib, and then click "**Actions>Surface>show**".

Note: It is possible to change the color of the surface. Open the "**Actions>Color>all option…**" menu and check the "surfaces" button so that the color changing will be applied only to surfaces. Surfaces can be colored using a selected color, or by elements.

Through the **Actions>Surface** menu, you can try to modify the surface representation to "**mesh**" and "**dot**". Then, go back to the "**solid**" representation.

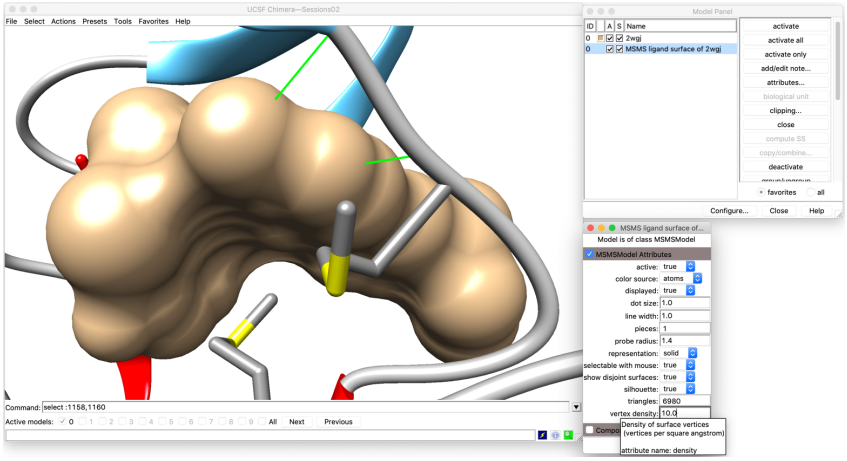


Surface in solid representation

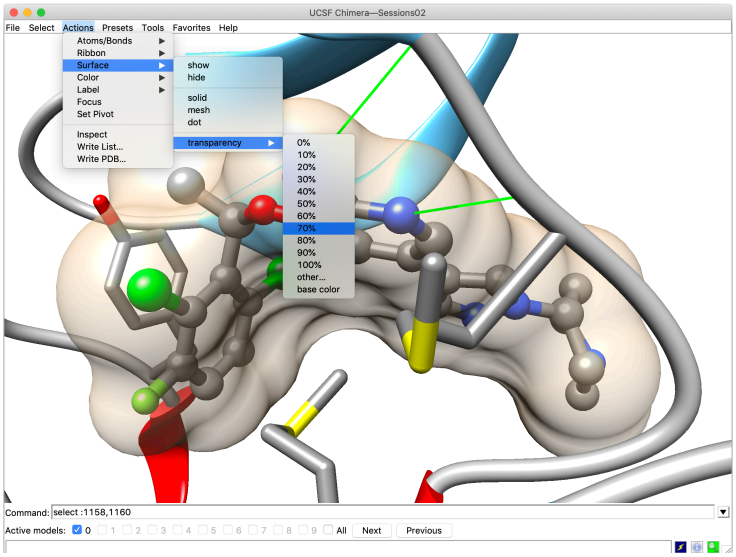


Surface in mesh representation

The surface smoothness may be increased by changing the corresponding attribute. In the "**Favorites**>**Model Panel**" window, select the surface by clicking on the second row in the list "MSMS ligand surface of 2wgj" (this one is highlighted in blue). Then click "**attributes…**" on the right panel and change the "**vertex density**" attribute to 10.0. Press Enter to apply the change and "**Close**".

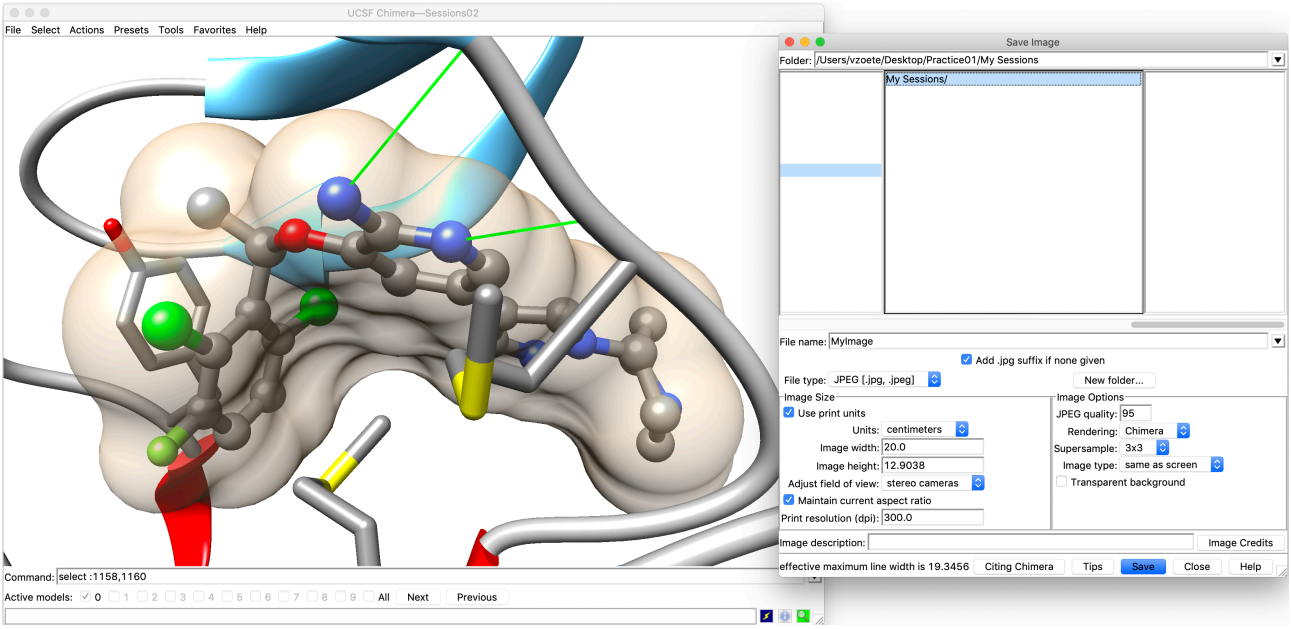


This surface gives a good idea of the volume occupied by the ligand. However, it also hides the molecule of interest. To correct this, the solid surface can be made transparent. In the "**Action>Surface>transparency**" menu, select 70%.



Saving images

Clear all selections using "**Select>Clear Selection**". Choose an orientation and a zoom that provides a satisfying point of view. Then, select the "**File>Save Image…**" menu item. In the new window, choose an image resolution of 300 dots per unit (dpi). Select "**Maintain current aspect ratio**" and enter an "**Image width**" of 20 centimeters. Select the format of the file that will be saved in the "**File Type**" menu. Select JPEG in this case. Finally, choose a "**File name**" and click "**Save**".



Images are saved as RGB figures.

Note: Generally, scientific journals require JPEG, EPS or TIFF images, with a 300 dots per inch resolution or higher, and with RGB colors. Such options are provided by UCSF Chimera.

Note: On Windows PCs it might happen that no image is generated. In this case, replace the unit of the "**Image width**" from centimeters to pixels, and choose an image width of about 1000 pixels.

Saving the session status

Save the last state of this exercise as a Chimera Session (as seen previously). This session will be used in exercise 2.

Closing the session. Quitting Chimera

The session can be closed using the "**File>Close Session**" menu item. One can quit Chimera with "**File/Quit**".

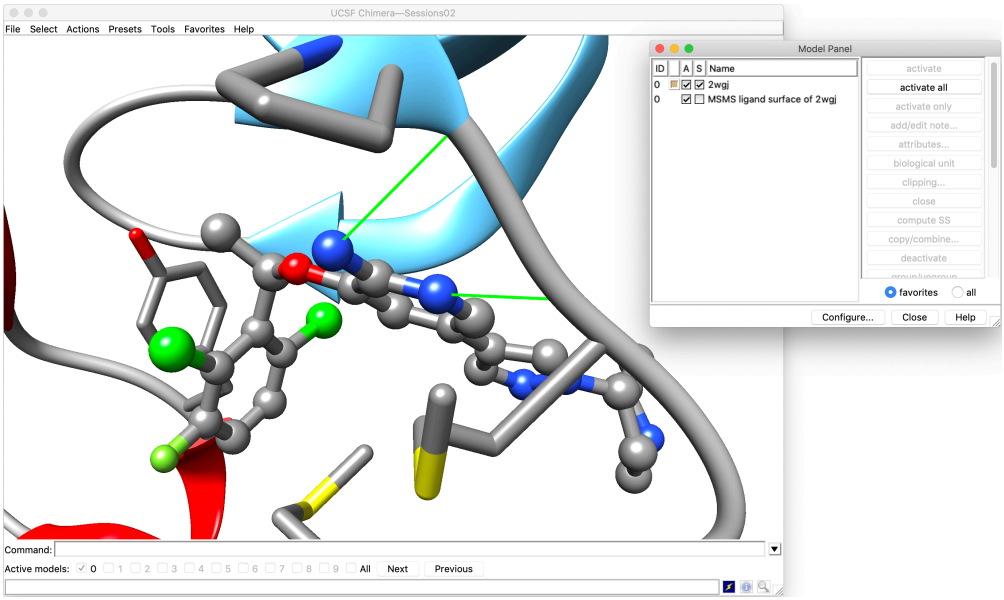
Exercise 2 – Advanced surface options

Restoring the last session of exercise 1

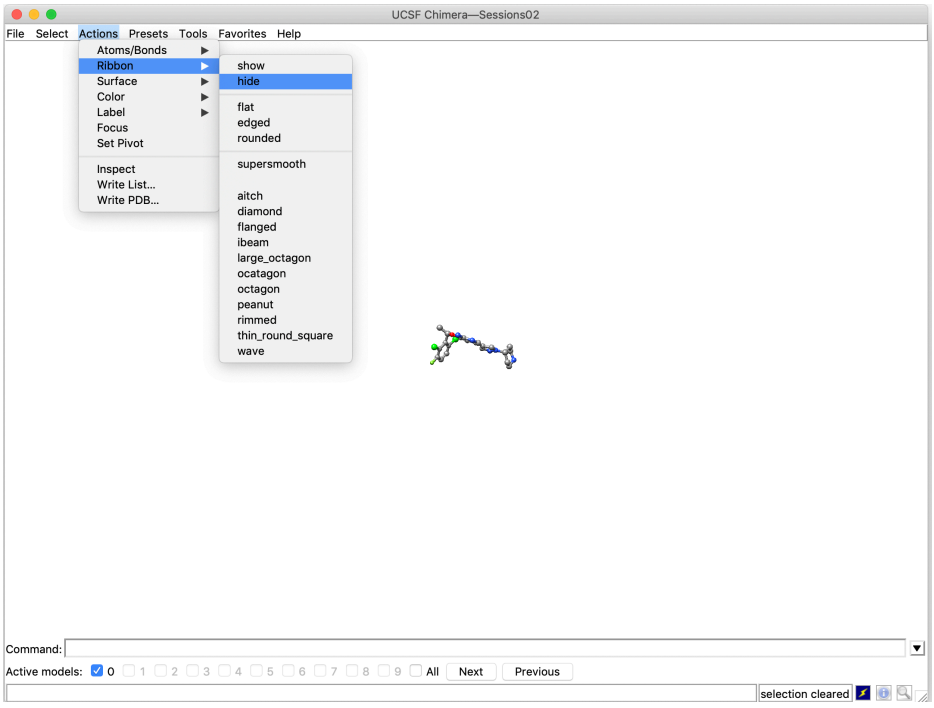
Restore the session saved at the end of exercise 1 using the "**File>Restore Session**" menu item

Showing the molecular surface of the protein

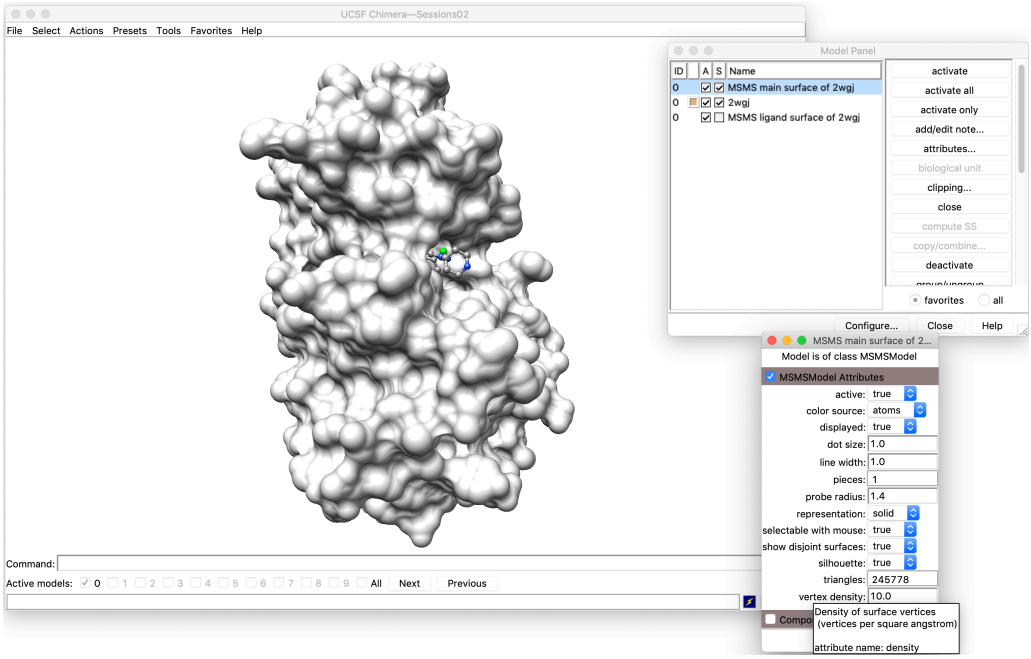
First, hide the surface of crizotinib. For this, open the Model Panel, using "**Favorites/Model Panel**". Then, uncheck the box S (for 'shown') for **"MSMS ligand surface of 2wgj"**.



Second, hide the bonds, atoms and ribbon of MET and keep only crizotinib visible. To do so, select the protein, for example by choosing "**Select>Residue>standard amino acids**". Then choose "**Actions>Atoms/Bonds>hide**" and "**Actions>Ribbon>hide**".

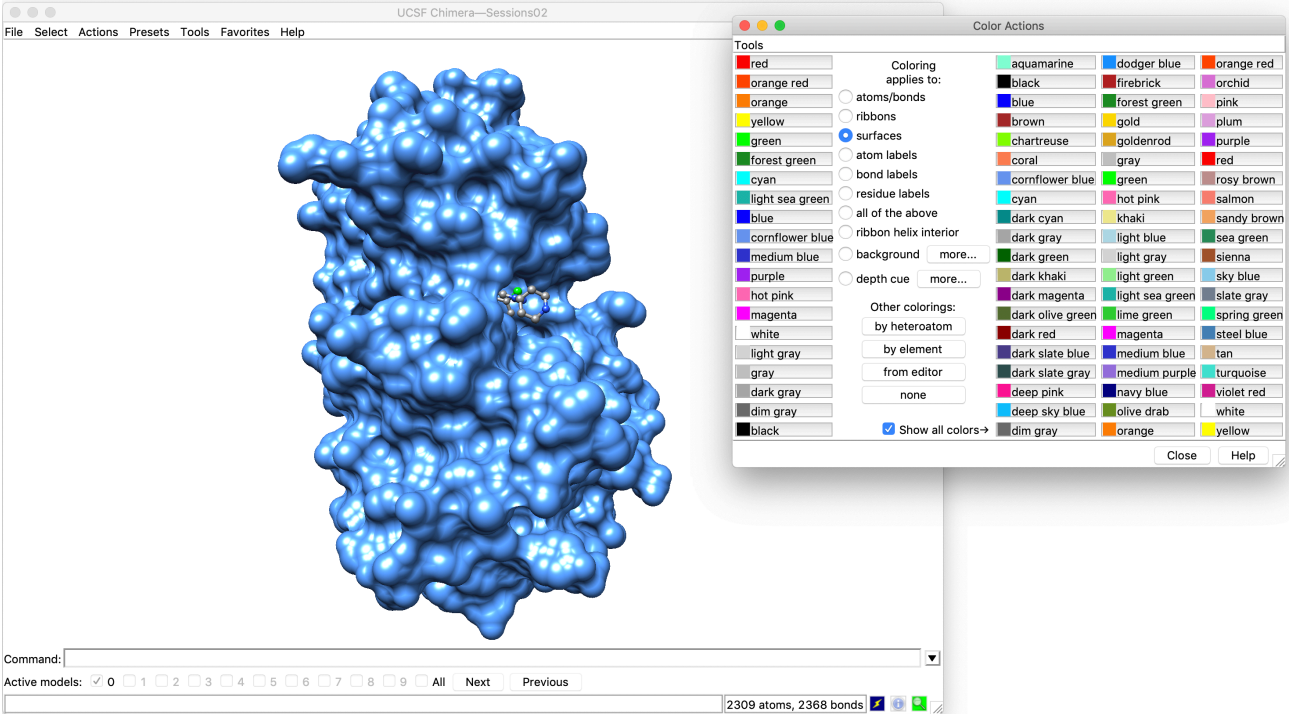


Be sure that the protein is still selected. If necessary, or in case of doubt, select it again as described above. Then, choose “**Action>Surface>show**”. Open the “**Favorites>Model panel**” menu, select the “**MSMS main surface of 2wgj**” in the left list (click to highlight in blue), then click “**attributes…**” and change the “**vertex density**” to 10. If your computer is too slow, it might be necessary to reduce the vertex density to 5. Click on “**Close**”



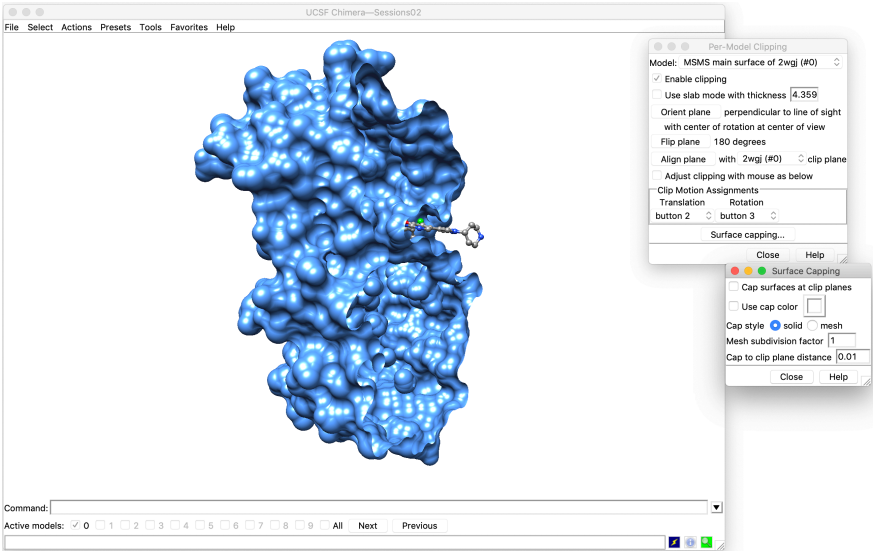
Changing the surface color

Be sure that the protein is still selected. Open and detach the “**Actions>Colors>all options…**” menu. Check the “**surfaces**” button and click on the color you would like to apply to the protein surface. The following image has been obtained using the cornflower blue color. Click on “**Close**”.



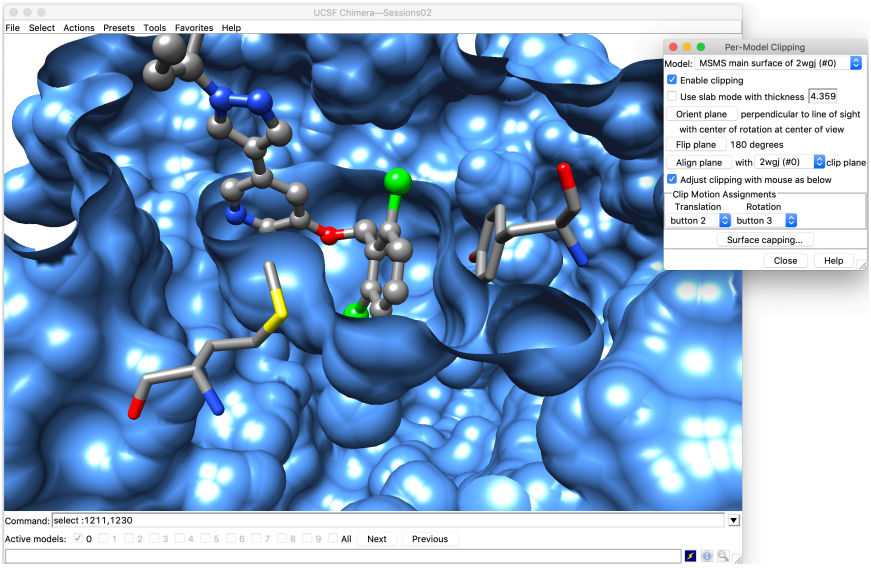
Clipping the protein

It is possible to cut the protein surface to get a better view of the ligand binding mode inside the binding site. To do so, open the "**Tools>Depiction>Per-Model Clipping**" menu. In "**Model**", select the "**MSMS main surface of 2wgj**" and check the "**Enable clipping**" box.
To begin with, we will remove the capping of the surface. For this, click on the **"Surface capping…"** button, and uncheck the **"Cap surface at clip planes"** box in the new **"Surface capping"** window. We will reactivate the capping in the next section.
Rotate the structure to get a better view of the clipping effect.

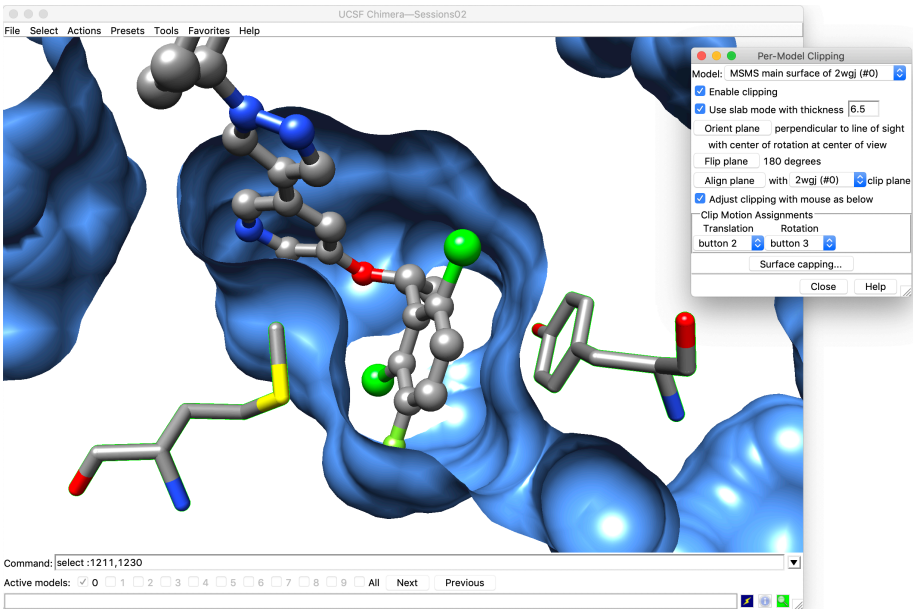


It is possible to manipulate the clipping position using the mouse. Check the "**Adjust clipping with mouse as below**" box. A click/drag of the central mouse button will control the translation of the clipping plane. A click/drag of the right mouse button will modify its orientation. Note that this cancels any possibility of protein translation or zoom using the mouse. You can get back to the normal behavior of the mouse by unchecking the "**Adjust clipping with mouse**" box.

You can select and display residues 1211 and 1230 by typing "**select :1211,1230**" in the command line (followed by **Return**), and then display them in the stick representation. Change the position and orientation of the clipping plane to have a good view of the interactions between these residues and the ligand.



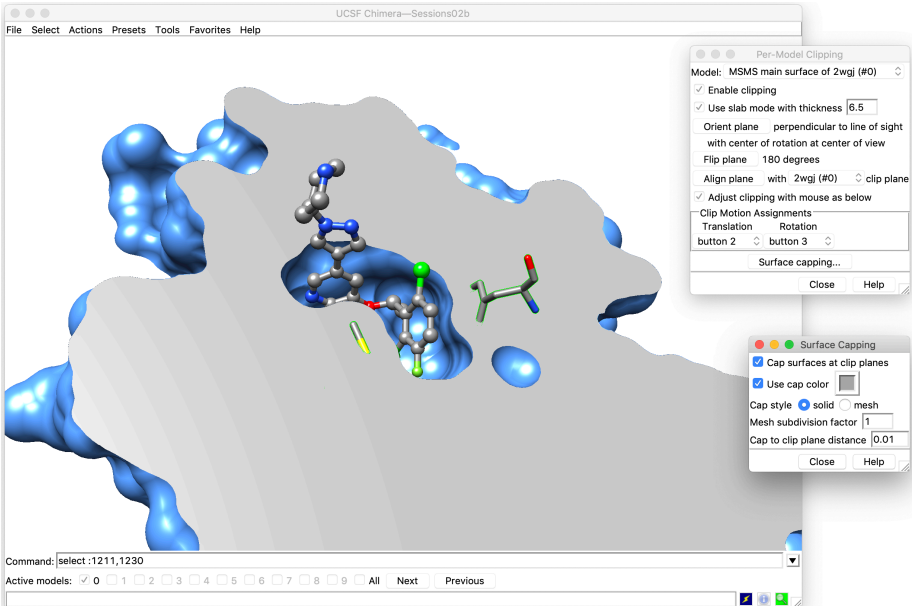
You can also get a good view of the shape of a buried binding site using the slab mode. In the "**Per-Model Clipping**" menu, check the "**Use slab mode with thickness**" and choose 6.5 for the thickness. This creates a protein slab that you can manipulate using the mouse by checking the "**Adjust clipping with mouse**" box.



Uncheck the "Use slab mode with thickness" box.

Surface capping

A clipped surface may be capped. To do so, choose "**Tools>Depiction>Surface Capping**" or click on "Surface capping…" in the "**Per-Model Clipping**" window**.** In the **"Surface Capping**" window, check the "**Cap surface at clip planes**" box. You can change the color of the capping plane by checking the "**Use cap color**" box and choose the color by clicking the colored square next to it.



This capping can also be applied with the slab mode, as in the figure above.

Save an image as described in the previous exercise, then save the session and close it.

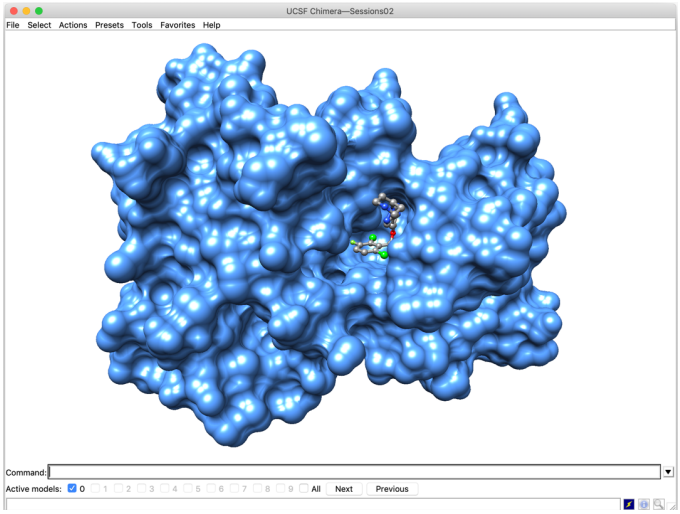
Exercise 3 – Lighting, cueing and shadow effects

Opening a previous session

Restore the session saved at the end of exercise 1 using the “**File>Restore Session**” menu item

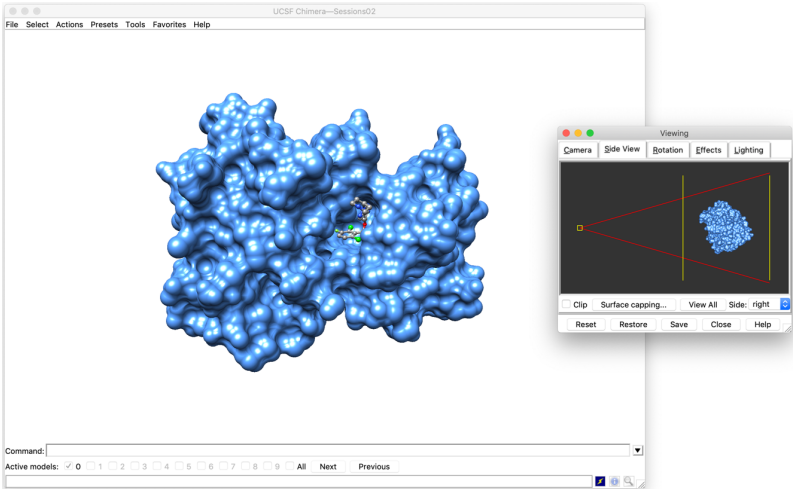
Showing the molecular surface of the protein

Follow the instructions given on page 19 to display the surface of the protein this way:



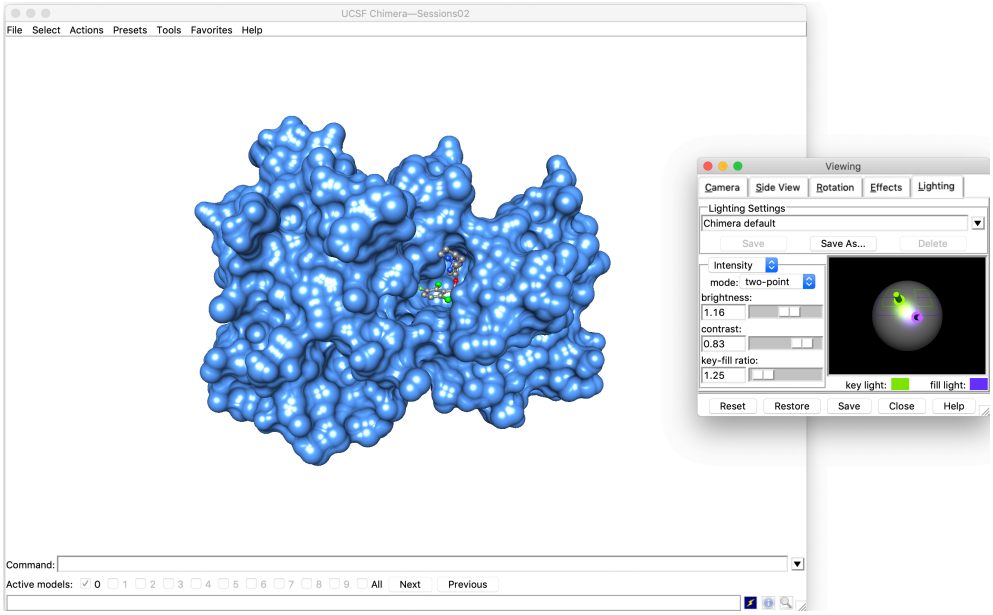
Using the side view

Open the “**Viewing**” menu by choosing the “**Favorites>Side view**” menu item. The “**Side View**” tab of the “**Viewing**” window should be active. Otherwise, click on the “**Side View**” tab. You will see a reduced view of the structure appearing in the “**Viewing**” window. The vertical lines show the clipping planes that define the region of space displayed on the screen. The square gives the viewer's eye position. The red lines show the field of vision.
Clicking the “**View All**” button adjusts the scale and clipping plane positions so that the view will include everything that is displayed. It is possible to move the position of the viewer's eye and clipping planes using the mouse.



Changing the lighting

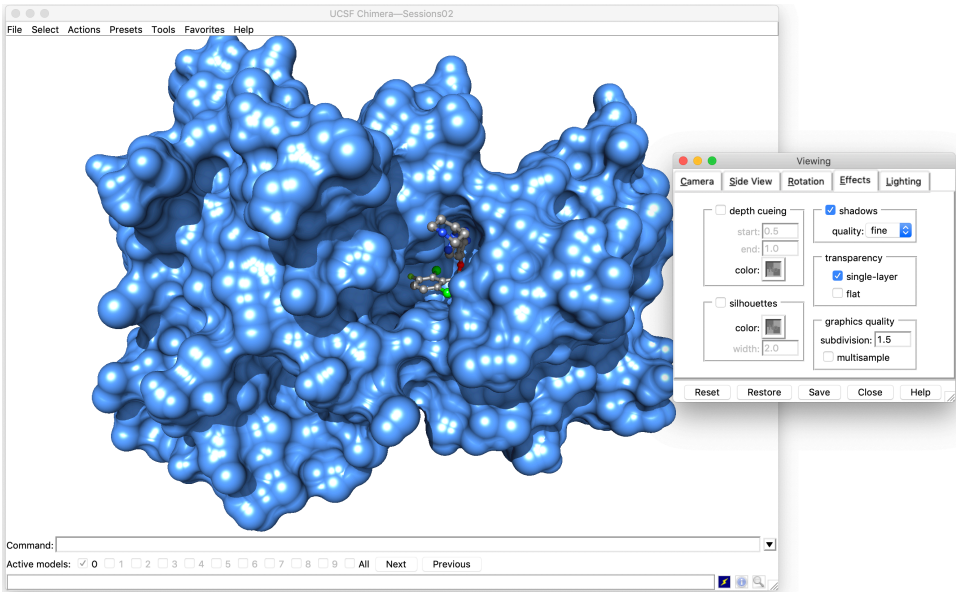
It might be useful to change the lighting (intensity and direction) to get a better view of the system. Select the “**Lighting**” tab in the “**Viewing**” window. This tab displays the light sources and parameters. The key light is the dominant brighter source of light. The fill light gives a secondary source. The solid arrows in the right view allow manipulating the lighting directions with the mouse. You can try different lighting directions to see their effect.



Adding 3D effects

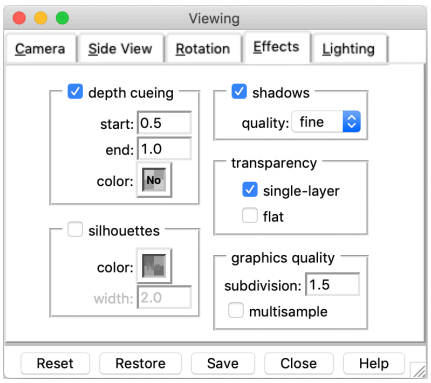
Shadows

Based on the lighting seen above, shadows can be calculated and displayed using UCSF Chimera. For this, select the “**Effects**” tab in the “**Viewing**” window, and check the box “shadows”. You can see the effect of the lighting on the shadows by modifying the position of the lights as seen in the previous paragraph.

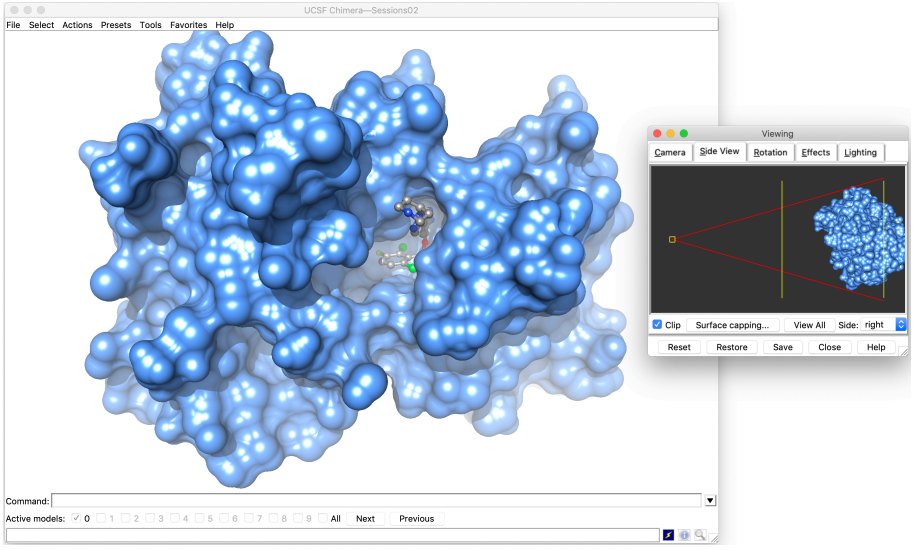


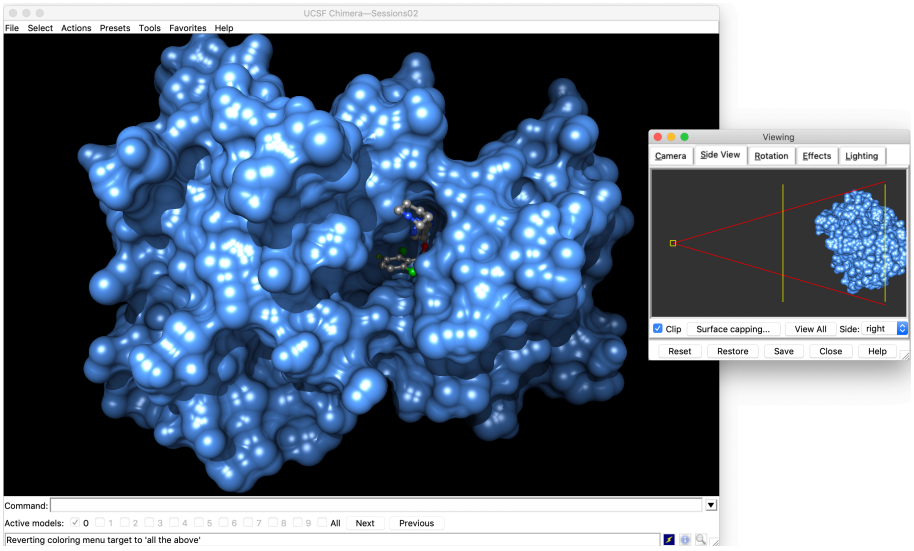
Depth Cueing

Depth cueing can be started by checking the corresponding box in the “**Effects**” tab of the “**Viewing**” window.



Depth cueing makes the rear part of the system disappear in the shadow (if the background color is black) or in the fog (if the background color is white). This is extremely useful to provide a 3D effect on a 2D image. The depth cueing effect can be controlled by dragging the rear and front clipping planes closer to or further from the protein





You can save and close the session.

Exercise 4 - Comparing structures

Loading macromolecular structures

The objective of this exercise is to learn how to compare two structures. For this, we will compare the kinase domain of MET, which we used during the previous exercises, and the experimental structure of the kinase domain of B-Raf, another therapeutic target in oncology.

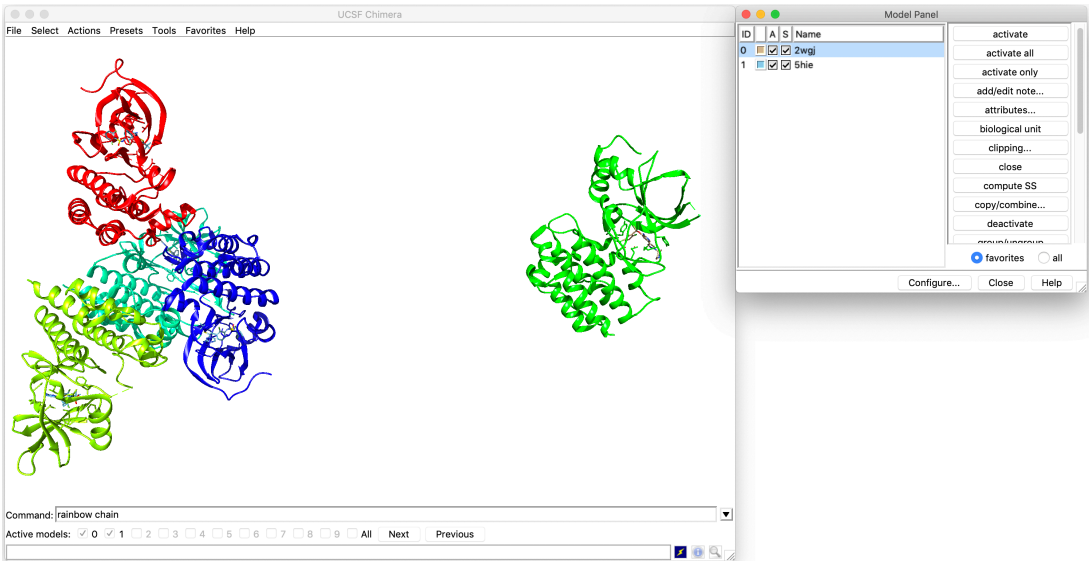
Fetch or load successively the PDB files:

- 2wgj, which contains the MET kinase domain binding crizotinib that we already used
- 5hie, which contains the B-Raf experimental structure, in complex with the FDA-approved drug dabrafenib

There are now two active models shown in the lower line of the main Chimera window and in the Model Panel. Model 0 is MET (2wgj), whereas model 1 is B-Raf (5hie). Zoom out with the mouse to see both proteins in the 3D window.

Type "**rainbow chain**" in the command line to automatically color each chain differently.

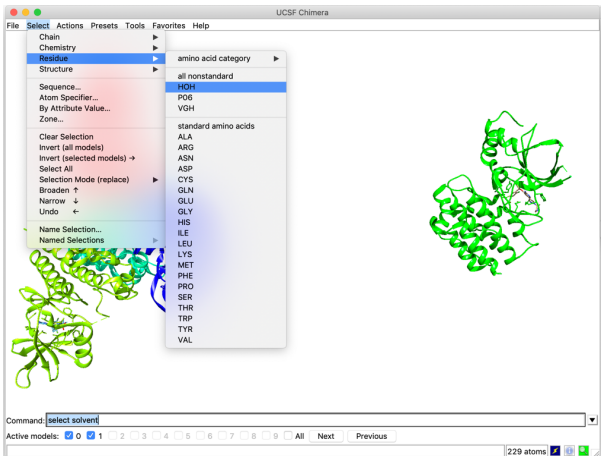
Finally, open the "**Favorites>Model Panel**", which should show the two systems: 2wgj and 5hie.

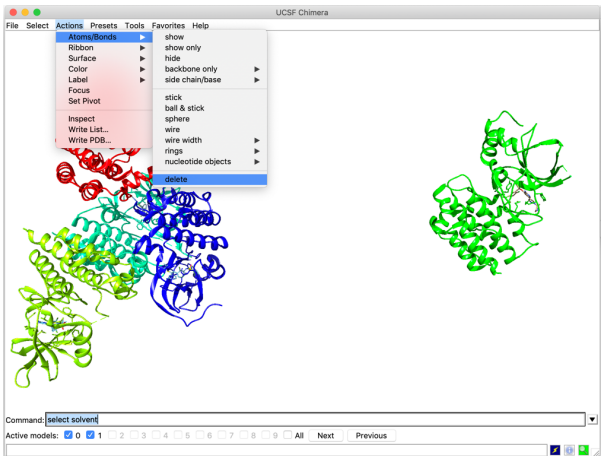


Delete atoms

To simplify the analysis, we will delete the atoms that are not of interest here.

For this, you can select all water molecules, by typing "**select solvent**" in the command line or by clicking "**Select>Residue>HOH**". Then, click "**Actions>Atoms/Bonds>delete**". Now, the system contains only natural amino-acids (i.e. the proteins) and some copies of the ligands.



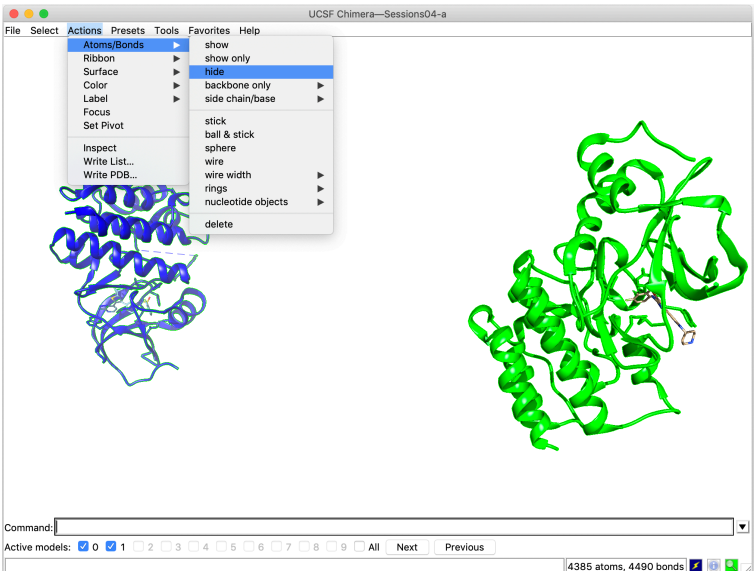


Four copies of B-Raf, corresponding to chains A, B, C and D are present in the 5hie structure. Since we only need one copy, we will delete chains B, C and D. For this, you can select one chain at a time in the selection menu, e.g. "**Select>Chain>B**" and then delete the corresponding atoms using "**Actions>Atoms/bonds>delete**". Alternatively, you can type "**select #1:.B,.C,.D**" in the command line to select all the three chains at the same time, and then delete the atoms, again using "**Actions>Atoms/bonds>delete**".

Note: The same result could have been obtained in one action, by typing "**delete #1:.B,.C,.D**" in the command line

Now, the system contains only one chain of each protein.

For clarity, we will hide all atoms and bonds, except those belonging to the ligands. For this, select the natural amino acids with "**Select>Residue>standard amino acids**" and hide them with "**Actions>Atoms/bonds>hide**".



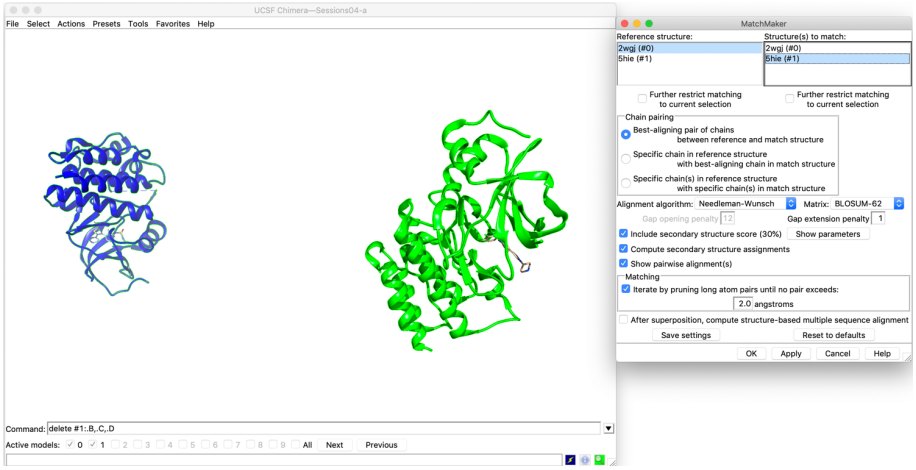
Match two proteins. Obtain a structural alignment

Open the structural alignment tool using "**Tools>Structure Comparison>MatchMaker**". Select 2wgj as "**Reference structure**" and 5hie as "**Structure(s) to match**". Click on "**Best aligning pair**

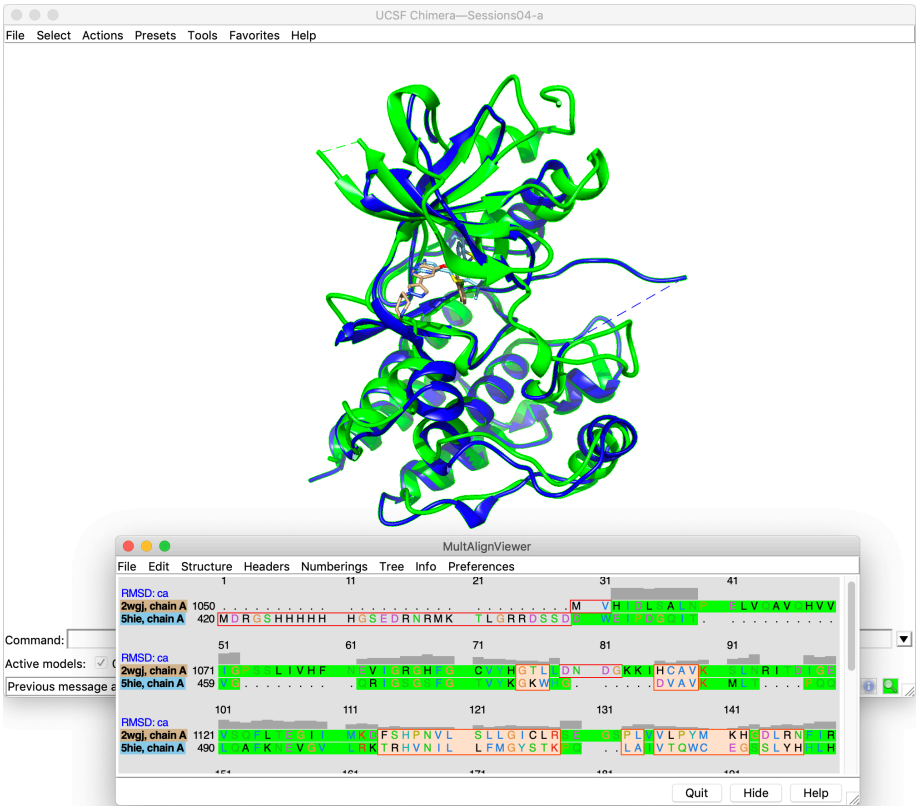
of chains […]", so Chimera will try to superimpose each possible chain of the first protein to each possible one of the second protein.

Note: Here, there is actually no need to select one particular option, since each protein has only one chain. However, these 3 "**Chain pairing**" options can be useful for instance to superimpose systems composed of different chains, by allowing to specify which chains to use as reference for the structural alignment.

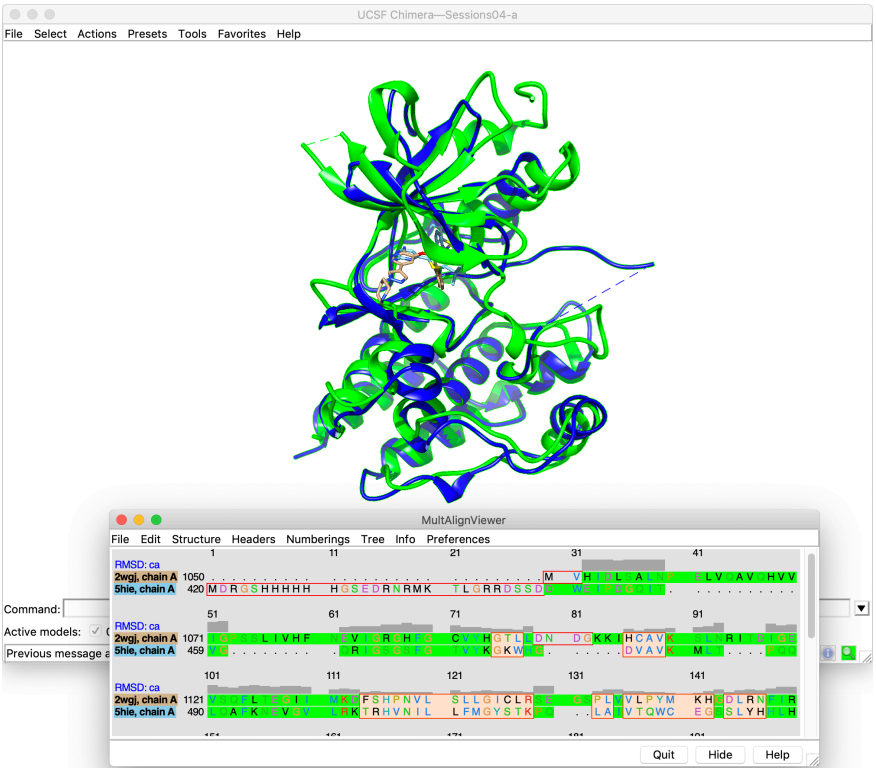
Verify that the "**Show pairwise alignment(s)**" box is selected. Finally, select the "**Needleman-Wunsch**" algorithm and click "**OK**". This will perform a matching of the two molecules, based on their sequence and 3D structure similarities.



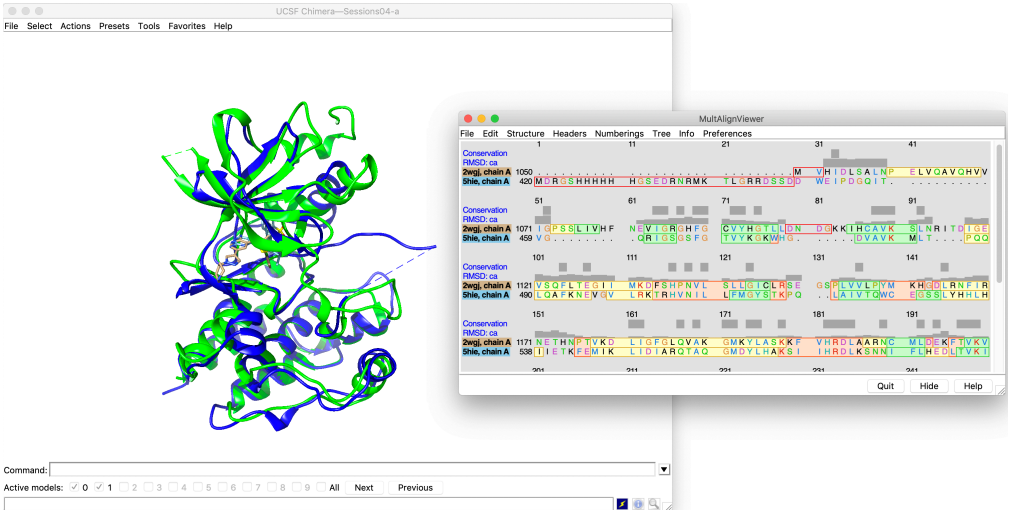
After a short moment, the two molecules should be superimposed. A new window, called "**MultiAlignViewer**" will also appear, showing the corresponding sequence alignments.



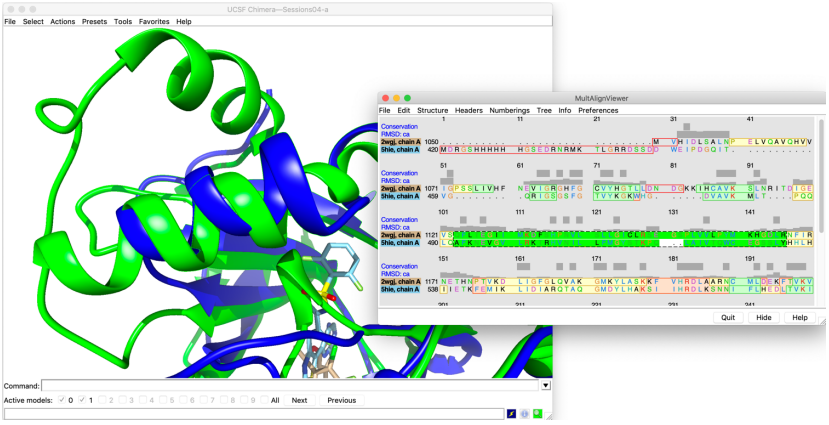
The percentage of sequence identity between the two molecules can be obtained using the menu "**Info>Percent identity...**" of the MultiAlignViewer window (here, 26.82%). It is possible to show the conservations between the two sequences by selecting "**Headers>Conservation**" in the MultiAlignViewer window.



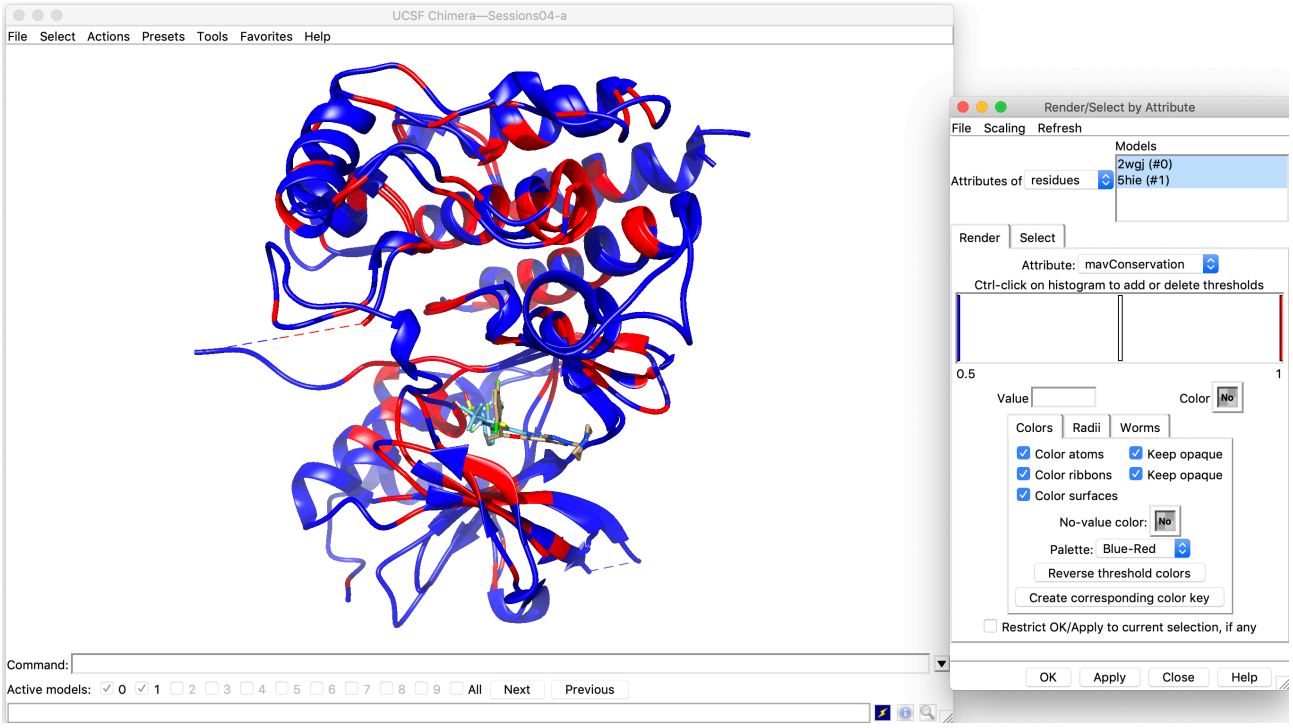
The secondary structure elements can be highlighted in the sequence using the MultiAlignViewer command: "**Structure>Secondary Structure>show actual**". Secondary structure elements appear in green for β-strands and yellow for α-helices.



Isolated residues or groups of residues can be selected by clicking over the sequence or in the "**Conservation**" line of the MultiAlignViewer, allowing the analysis of the structural differences or analogies between the proteins.



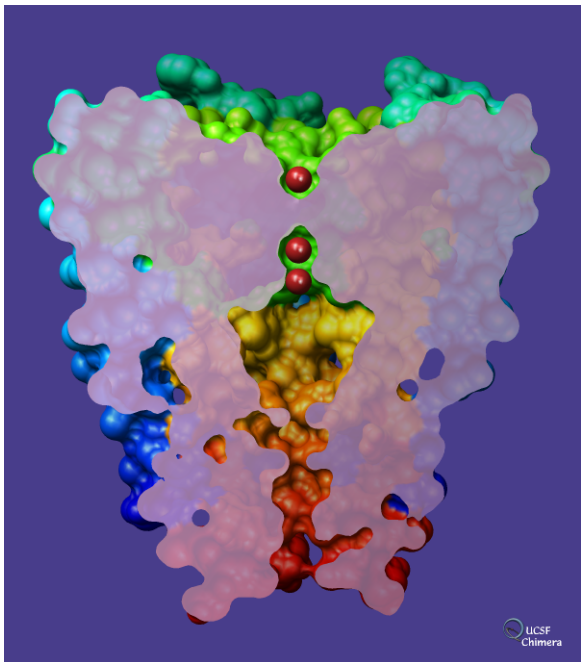
It is possible to color the structures according to the percentage of conservation in the sequence alignment. For this, open the "**Structure>Render by Conservation**" menu of the MultiAlignViewer, then click "**Apply**" using the default values for the different options. In the present case, the most conserved regions are colored in red and the less conserved ones in blue.



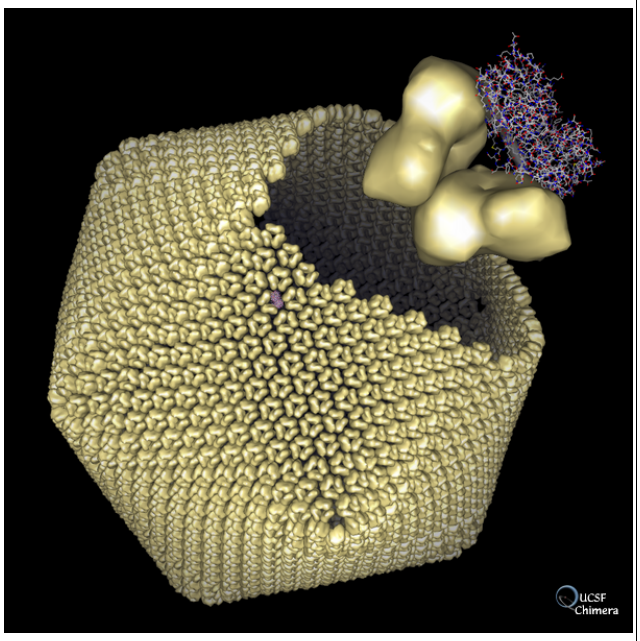
Going further with UCSF Chimera

These exercises give only a very limited overview of what Chimera is capable of. You can find a detailed documentation, as well as some tutorials, at the following address:
http://www.cgl.ucsf.edu/chimera/docindex.html
Here are some examples of images produced using Chimera that were taken from the official Website.

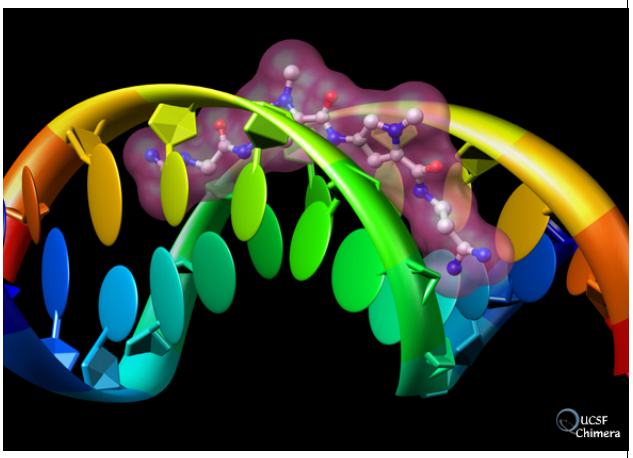
Sliced potassium channel



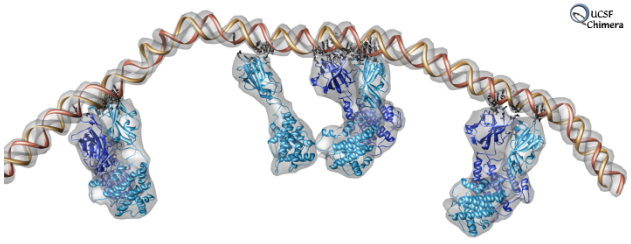
Paramecium Bursaria Chlorella Virus



DNA and Netropsin



Bluetongue Virus and Viral RNA



Practice Session 2:
Ligand-protein docking

Antoine Daina, Ute Röhrig, Vincent Zoete – 2023
Contact : vincent.zoete@sib.swiss or antoine.daina@sib.swiss

Exercise 5. Docking of anti-inflammatory drug Celecoxib into protein COX2.

❖ *Don't forget to save **Chimera sessions** from time to time for easy recovering in case of mistake!*

First save: **>File >Save Session As…** Give a name to the .py file and click the **"Save"** button**.**

Save frequently: **>File >Save Session.**

To get the last session back: **>File >Restore Session…**

❖ *For this exercise, you should have access to **Command Line**.*

>Favorites >Command Line

- *Download the crystal structure of COX2 in complex with inhibitor Naproxen from the Protein Databank into Chimera*

  >File >Fetch by ID…

  In the pop-up window, **select PDB** and type "**3q7d**" in the **text field**
  click the "Fetch" button.

  It may take a few seconds for the system to be downloaded and displayed.

- *Remove the chains B of the protein*
  >Select >Chain >B
  >Actions >Atoms/Bonds >delete

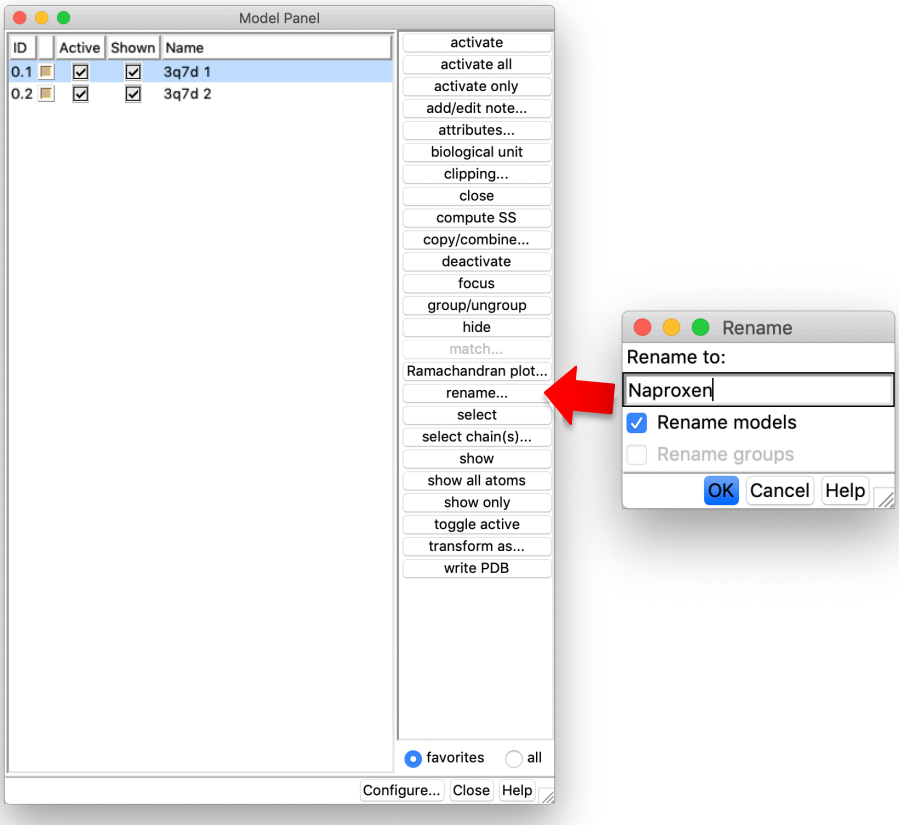
- *Spot on non-protein parts of the structure*
  >Select >Residue >all nonstandard
  Fly-over all highlighted items to identify the name of residues

- *Remove glycans and adjuvants*
  In the Command Line, type **sel :BOG,NAG**
  >Actions >Atoms/Bonds >delete

- *Color, display and extract the molecule NPX (Naproxen) from the cocrystal.*
  In the Command Line, type: **sel :NPX**
  >Actions >Color > by element
  >Actions >Atoms/Bonds > ball & stick
  In the Command Line, type: **split atoms :NPX**

- *rename the protein (COX2) and the ligand (Naproxen).*
  In the **Model Panel (>Favorites >Model Panel),** select on the first line **(0.1)** and click on the **"rename…"** button on the right**.**
  In the Rename window**,** type **"Naproxen"** and then click **"OK".**

  In the **Model Panel** select on the second line **(0.2),** click on the **"rename…"** button and type **"COX2"** and **"OK".**



- *Generate the protein surface*
  >Select >Residue >standard amino acids
  >Actions >Surface >Show
  >Select >Clear Selection

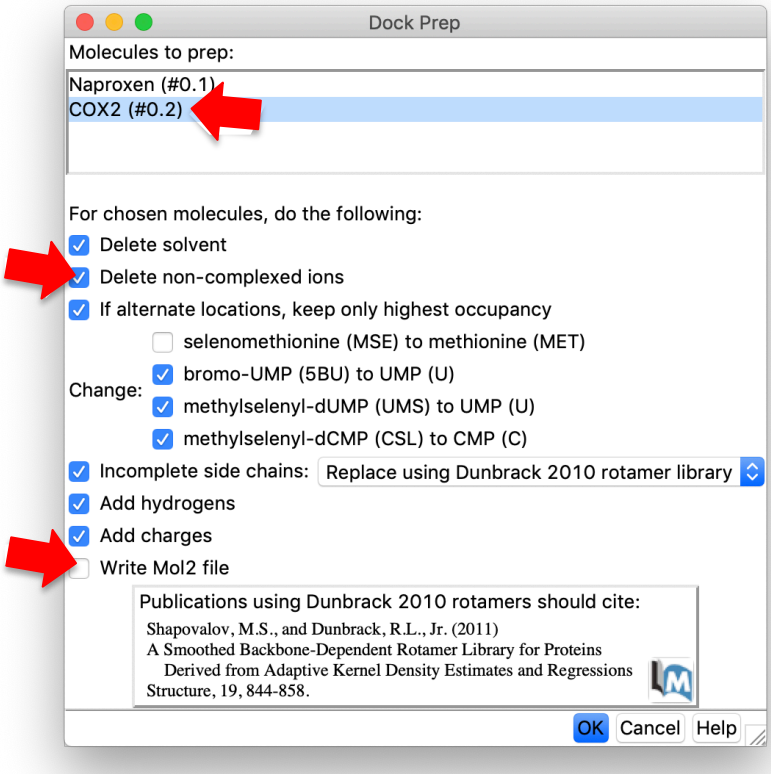
  ❖ At this point, one should have the protein COX2 in entry **0.2**, its surface (MSMS) in **0.2** (in another line in the Model Panel) and the cocrystallized inhibitor Naproxen in entry **0.1**.

  Now, you can visualize easily the binding pocket of the heme.
  Try rotating the structure to visualize the other distant, more buried binding site of the inhibitors, by locating the Naproxen.

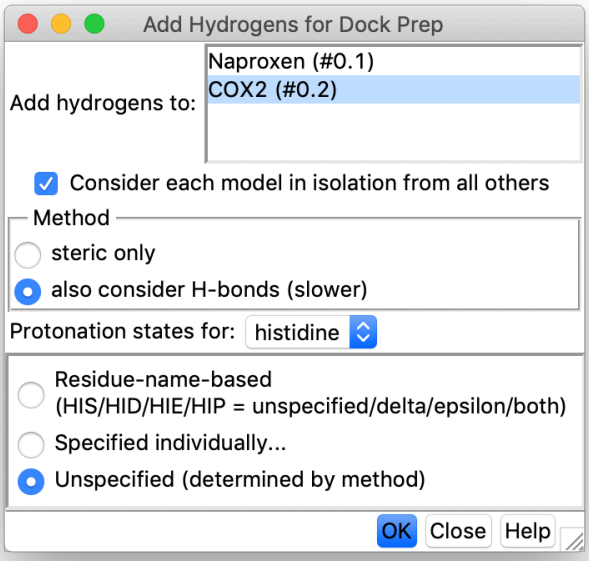
- *Prepare the COX2 structure for docking*
  >Tools >Surface/Binding Analysis >Dock Prep

  ○ In the **first panel "Dock Prep":**
  Select **only** the protein entry **COX2 (#0.2)** to be prepared in the **"Molecules to prep" list.**
  Tick all boxes except "Write Mol2 file".
  Make sure that the **Incomplete side chains will be replaced** using the **"Dunbrack 2010 rotamer library".**

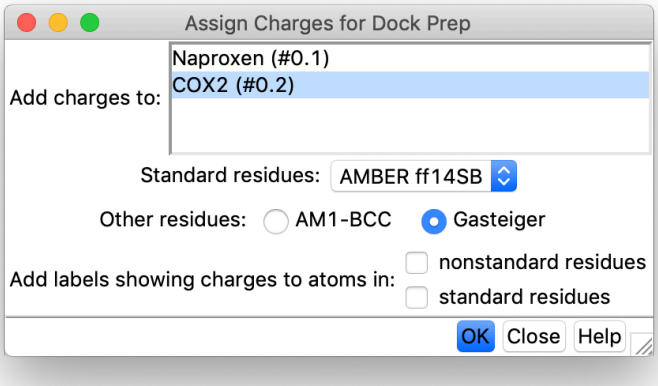
  Click **"OK"**



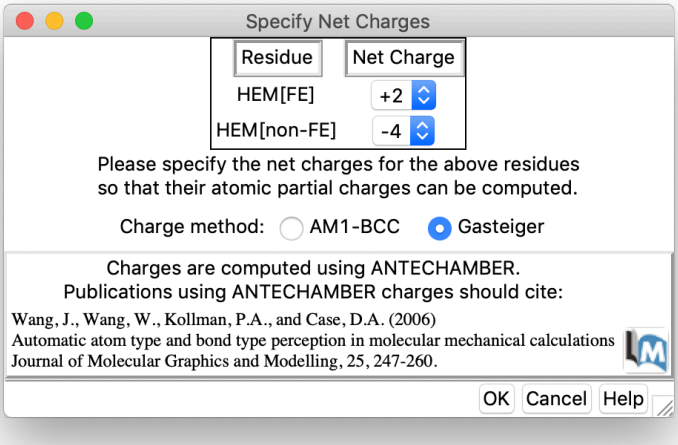
- In the **second panel "Add Hydrogens for Dock Prep":**
  Select the protein to be prepared in the **"Add hydrogens to" list**
  Add hydrogen atoms **also considering H-bonds (slower)**
  Treat histidine **Unspecified (determined by method)**
  Click **"OK".**



- In the **third panel "Assign Charges for Dock Prep":**
  Select the protein to be prepared in the **"Add charges to" list**
  Use **AMBER ff14SB** charges for standard residues
  Use **Gasteiger** charges for non-standard residues
  Let both label boxes unchecked.
  Click **"OK".**



- In the **fourth panel "Specify Net Charges":**
  Assign a charge of **+2 to HEM[Fe]**
  Assign a charge of **-4 to HEM[non-Fe]**
  Use **Gasteiger** method of calculation.
  Click **"OK".**



- *Download Celecoxib from the PubChem database*

One needs a reliable structure of Celecoxib with a realistic 3D geometry in a manageable file format. This can be found in the **PubChem database**.

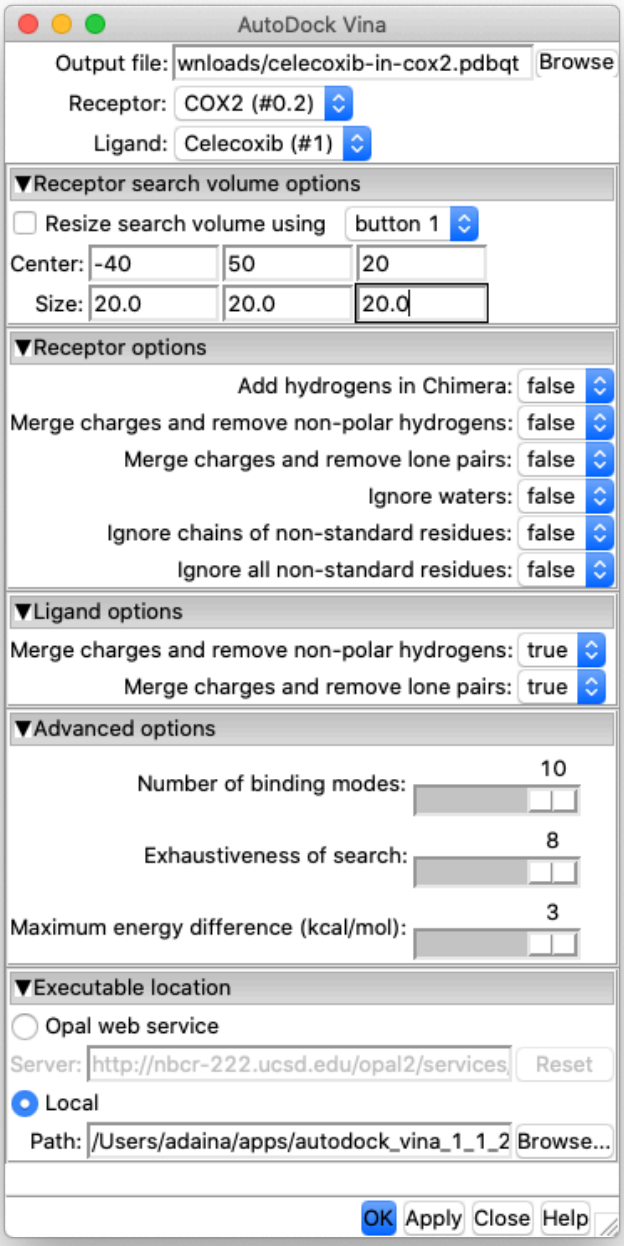
>File >Fetch by ID…
In the pop-up window, **select PubChem** and type Celecoxib ID code "**2662**" in the text field
click the "Fetch" button.
In the **Model Panel (>Favorites >Model Panel),** select on the line **(1)** and click on the **"rename…"** button on the right**.**
In the Rename window**,** type **"Celecoxib"** and then click **"OK".**

❖ At this point, one should have the cocrystallized ligand Naproxen in **#0.1**, the target COX2 prepared for docking and its MSMS in **#0.2**, and the Celecoxib to be docked in **#1**.

- *Setup and run docking (with VINA)*
  >Tools >Surface/Binding Analysis >AutoDock Vina

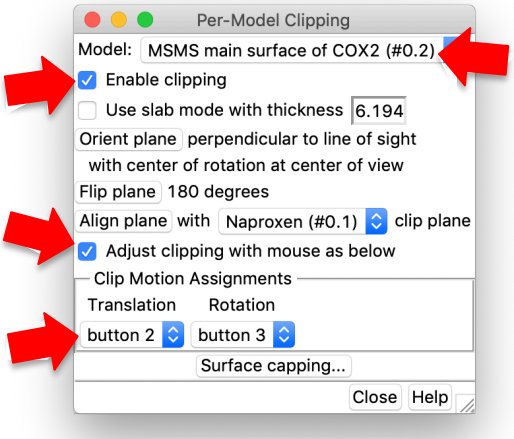
  - In the AutoDock Vina window, click the **"Browse"** button, navigate to a convenient location on your hard drive (you have the possibility to create a new folder), enter **"celecoxib-in-cox2.pdbqt"** in the **file name: text box** and click on **"Set Output Location"**.

  - In the AutoDock Vina window, select **COX2 (#0.2)** as **Receptor** and **Celecoxib (#1)** as **Ligand.**



- Expand the **Receptor search volume options** panel to define a box encompassing the binding site, where to focus the docking search.
  Enter random numbers in the **Center** boxes (x,y,z coordinates, eg. **1 1 1**).
  Enter **20 20 20** the in the **Size** boxes (side length in Å);
  You should be able to visualize a green box. If not, zoom out.

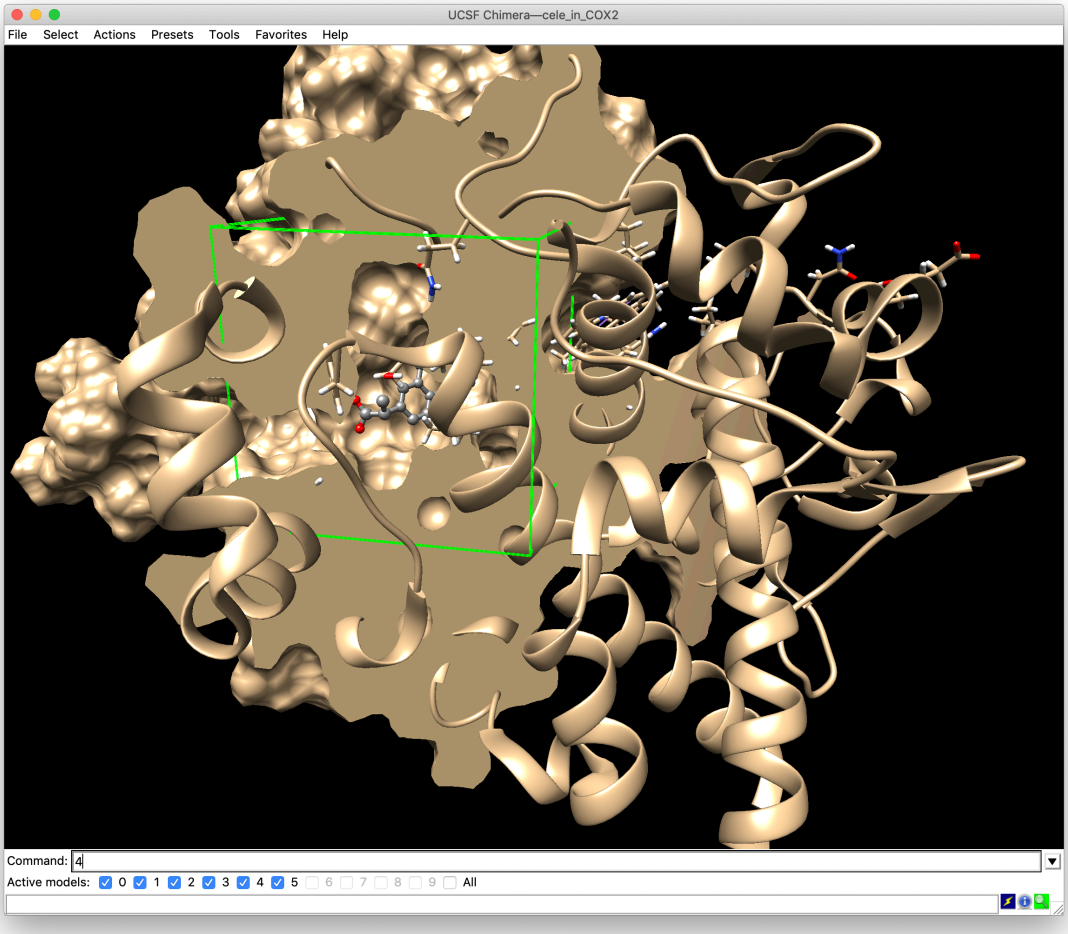
  Visualization of the binding pocket is easier by clipping the surface of COX2:
  Tools >Depiction >Per Model Clipping.
  Select the protein surface,
  tick "Enable clipping",
  tick "Adjust clipping with mouse as below"
  Select "**button 2**" or another.



- Play with mouse with e.g. **button 2 pressed and hold** to find the best cutting plan to visualize the binding site.
  Untick "Adjust clipping with mouse as below".

- Go back to the **AutoDock Vina** Panel.
  Tick the box **"Resize search volume using** (eg.) **button 2**" in.
  By **clicking and holding** mouse button 2, one can **translate** the green box in the space. Place the cubic search region so that it includes the volume of the guessed binding pocket entirely

- This is the trickiest part; you have to play with the rotation/zoom of the system and the translation of the box. Something like:



- Develop the **Receptor options** panel
  Set **all options** to "**false**" (because the receptor was already prepared through the Dock Prep procedure).
- Develop the **Ligand options** panel
  Set **both options** to "**true**"

- Develop the **Advanced options** panel
  Verify that the Number of binding modes is 9, the Exhaustiveness of search is 8 and the Maximum energy difference is 3 kcal/mol.

- Develop the **Executable location** panel
  click on **Local** (the Opal server is out of service, contrarily to what is in the video)
  Give the **path where you have installed Autodock VINA** on your computer.
  You can also navigate to the **executable file** by clicking on "**Browse**".

  On Mac and Linux, your path is where you have uncompressed the TAR archive. For instance:
  `/Users/adaina/apps/autodock_vina_1_1_2_mac_catalina_64bit/bin/vina`

  On Windows, something like:
  `\Program Files (x86)\The Scripps Research Institute\Vina\vina.exe`

In case, please refer to Vina documentation:
https://vina.scripps.edu/manual/#installation

- Launch your docking! Click on **"Apply"**.
  The message "Running" appears at the bottom of the main Chimera Windows.
  The Vina run duration is normally less than one minute. Upon completion a window "ViewDock" including the docking results pops-up. ***Go to page 42 to analyze****.*

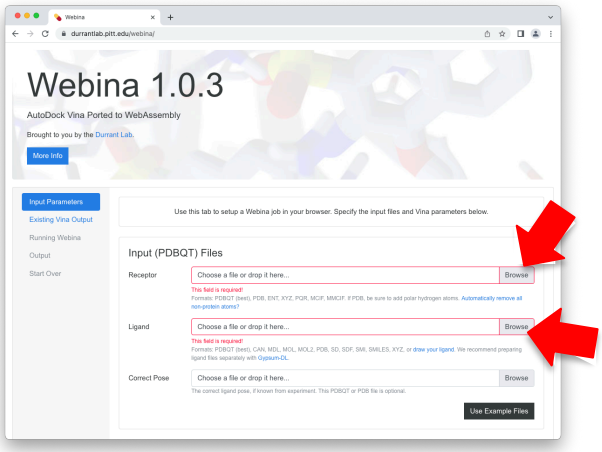
Workarounds (if the VINA program not installed locally)

In case it has not been possible to install AutoDock VINA locally on your computer, here are two workarounds: 1. Use the Webina interface or 2. Download pre-calculated VINA docking results.

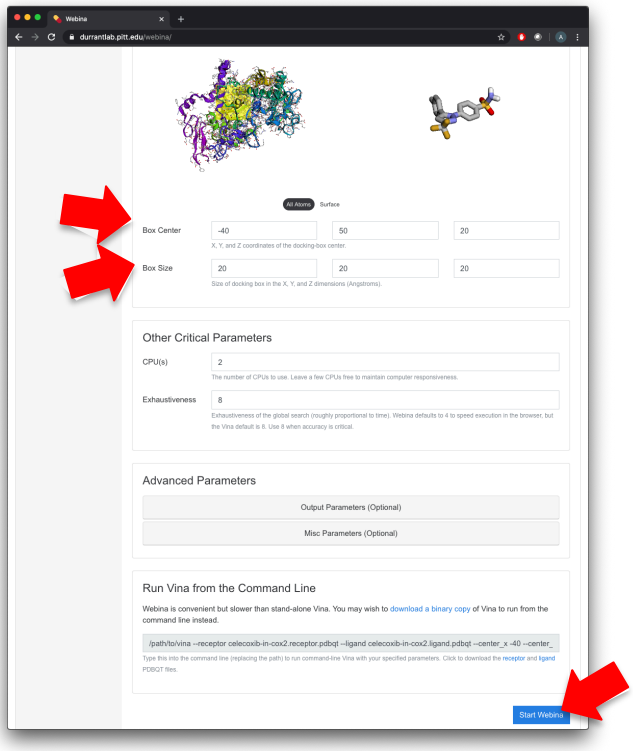
- *Workaround 1: Run docking through the Webina interface.*

Use **Webina**, a Web interface that performs local Vina docking without installation.
Go to: https://durrantlab.pitt.edu/webina/ (best in Chrome or Firefox browsers)

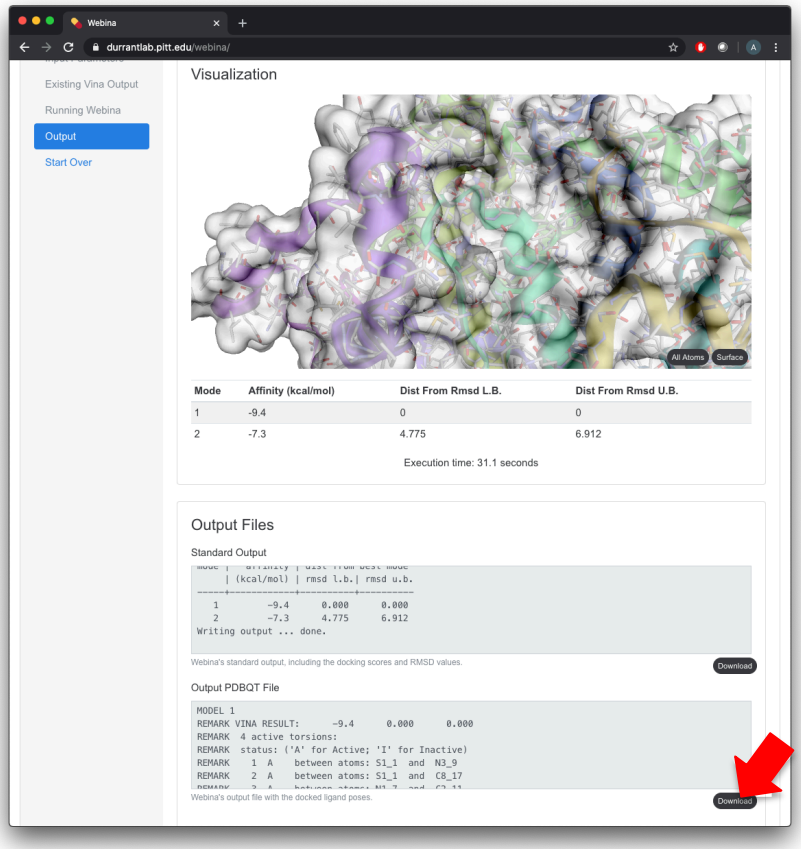
- You have to **input both prepared receptor and ligand as PDBQT** files. Both files were created on your hard drive, where you have **"Set Output Location"** in Chimera (p. 37)



- After upload (protein can take a few seconds and unimportant
- warnings can be displayed), define the Search Volume, as Box
- center and Box Size, the same way as set before in Chimera (pp.
- 37-38). Let all other parameters unchanged.



- Click "**Start Webina**"
- After one minute or so (depending on your computer), you can **download the Output PDBQT file** on your hard-drive (something like **webina_out.pdbqt**).
- **Go to page 42** for analyzing the docking results.



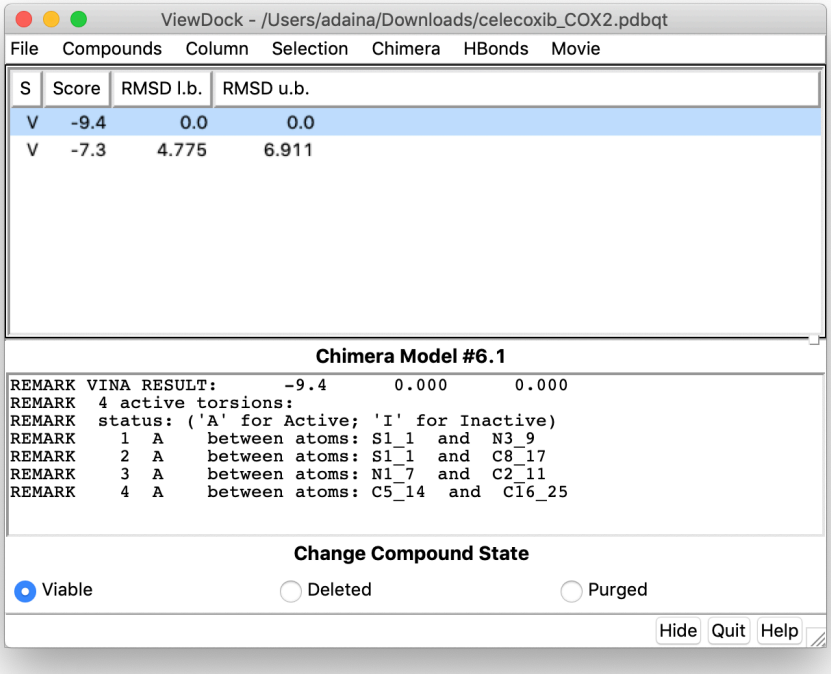
Workaround 2: Don't run docking yourself and download pre-calculated results.

Download an **already calculated** VINA docking at
http://drug-design-teaching.ch/download/docking-celecoxib-in-cox2.pdbqt

Depending on your Web browser, the download may not start automatically, and the full text is displayed. If so, save the file on your hard-drive and keep the .pdbqt extension (not .txt).
Go to next paragraph for analyzing the docking results.

- *Analyze docking results*

If you have performed a Vina docking locally using the Chimera interface, the **ViewDock** window should pop up automatically upon docking completion. If you have used the Webina interface (workaround 1) or have downloaded the pre-calculated results (workaround 2), you can open the saved output file in Chimera with **>Tools >Surface/Binding Analysis >ViewDock** (possibly named **"webina_out.pdbqt"** or **"celecoxib-in-cox2.pdbqt")**



- In the **ViewDock** window, the lines starting with a "V" are individual docking poses. Go through every docking solution by **clicking the first line** and pressing the 🡻 **key**. It is also possible to display multiple poses by holding the **ctrl key.**

    - *How many docking poses are proposed?*
    - *Are all poses located inside the binding site as defined?*

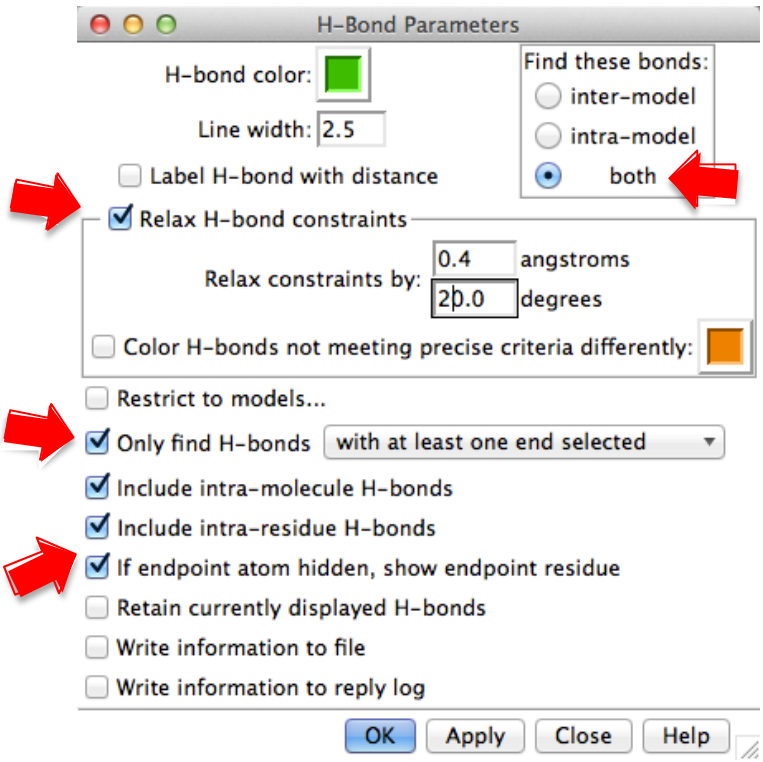
- If cocrystallized Naproxen is hidden, display it:
  In the **Model Panel (>Favories >Model Panel**):
  Tick the "**S**" box in line Naproxen (#0.1)
  Untick the "**S**" box in line MSMS main surface of COX2 (#0.2)
  Untick the "**S**" box AutoDock Vina Search Volume

    - *Qualitatively, is one Celecoxib docking pose better superimposed over the crystallographic pose of Naproxen?*
    - *Is it the solution evaluated as the most favorable by the scoring function? Note the score. In what range is the predicted K_i of Celecoxib inhibiting COX2?*

- Evaluate the intermolecular **hydrogen-bonds** explaining (at least partly) the recognition of the Celecoxib by COX2.

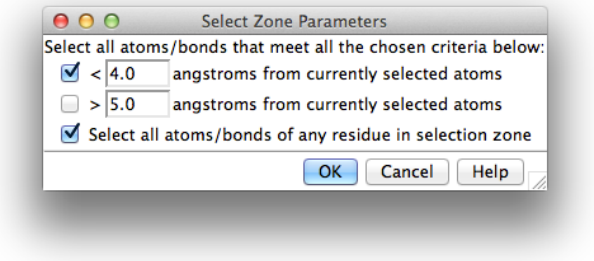
  Hide the protein surface and the Naproxen, and display COX2 (from the **Model Panel**).

  Display and select what you consider the best docking pose (for instance, **select #3.1**), change the rendering: **Actions >Atoms/bonds > ball & stick**.
  and then **>Tools >Surface/Binding Analysis >FindHbond.**
  In the **H-Bond Parameters** window, make sure that "**both**" (intra- and intermolecular hydrogen-bonds will be searched) and that **Relax H-bond constraints** will be applied.
  Select "**Only find H-bonds with at least one end selected**", and **Include intra-molecule H-bonds** and **Include intra-residue H-bonds**, and **If enpoint atom hidden, show endpoint residue.** Finally click **"Apply".**



- *How many possible intermolecular hydrogen-bonds were found?*
- *Which amino acid(s) is (are) involved? Which atoms are involved? Are they part of backbone or side chain?*

Display the amino acids close to the docked Celecoxib: Select the best pose (for instance **select #3.1),** then **Select > Zone…** In the Select Zone Parameters window, **Select all atoms/bonds of any residue in 4.0 Å** of the binding solution. Click **"OK"**.



Then **Actions >Atoms/Bonds >Show only.**

For clarity, hide non-polar hydrogens: **>Select >Chemistry >IDATM type >HC** and **>Actions >Atoms/Bonds >hide**.

- *Give two examples of hydrophobic interactions.*
- *Can you spot aromatic interactions?*
- *Looking at the protonation of amino acids (made automatically), would it be worth*
- *to go back and re-run docking with alternative states/tautomers?*
- *Propose one targeted mutation that could validate the predicted binding mode of Celecoxib in COX2.*

- Make the best possible JPEG image that illustrates how Celecoxib is predicted to bind to COX2. Emphasize, select and label (once selected **>Actions >Label >residue >name + specifier**) the amino acids involved in intermolecular hydrogen-bonds.
  >File >Save Image…

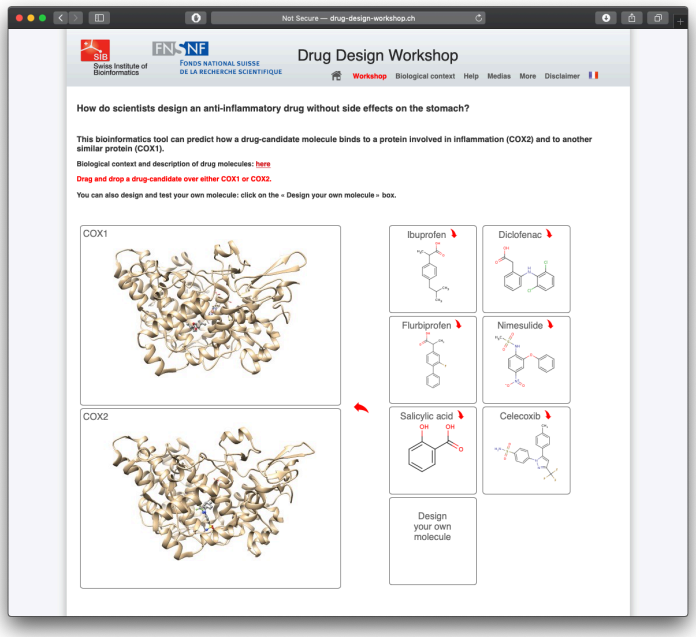
  Please note that in some bugs were reported for the generation of images with the Windows version of UCSF Chimera. If the picture generated is all black, please try to set "print unit" to "points" and adapt the "Image width" to 225. You may also try with another output format (PNG, for instance).

- *Perform structure-based optimization yourself! Are you able to design, build and dock a chemically modified Celecoxib? Or a totally different putative COX2 inhibitor? If not, you can use the simplified (educational) web-based procedure "Drug-Design-Workshop" described in the next section.*

Exercise 6. Structure-based optimization of COX2 inhibitors

- *Retrieve precalculated docking pose of Celecoxib*

  In a web browser, go to the website **http://www.drug-design-workshop.ch/cox.php**
  Drag-and-drop Celecoxib structure over the picture of “COX2”.



  Almost immediately, the **pre-calculated docking** results appear in the web page.

  - *Can you recognize the binding mode you predicted in the part A of the session?*

- *Dock interactively your own molecules into COX2*

  Click on the “**Design your own molecule**” box.
  A **sketcher** appears at the bottom of the page.
  Draw the molecule to dock and click on “**Done**”. The molecule structure appears in the box.

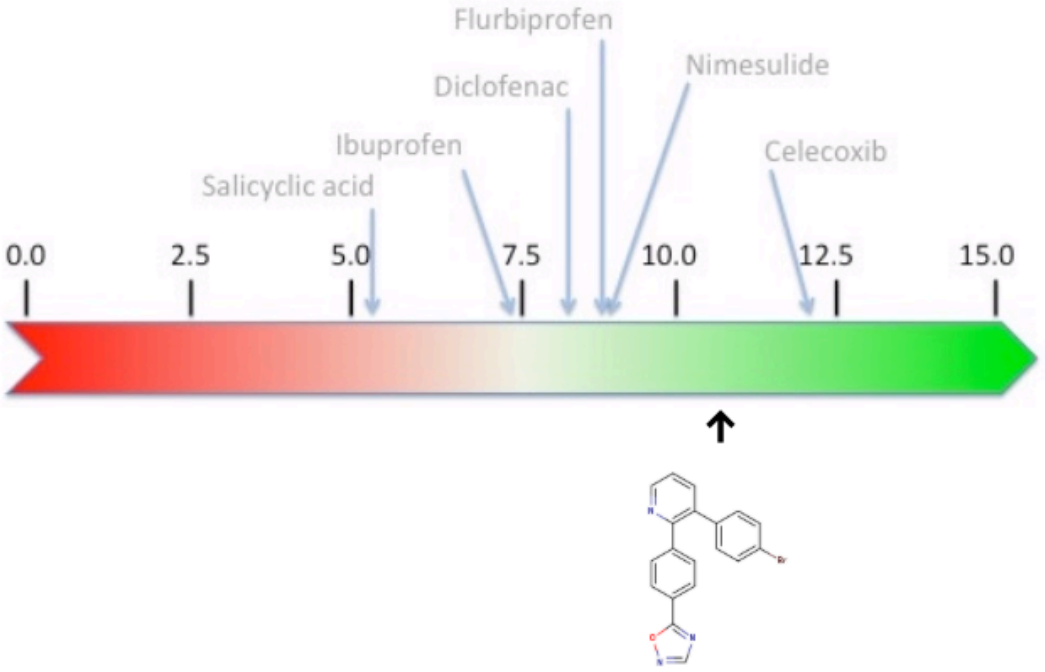
  Drag-and-drop it over the “COX2”. The docking is prepared and starts automatically, when a processor on the server is free.

  Upon completion (less than a minute, you can monitor your docking on a progression bar), the message “Docking completed!” appears.
  Click on the link: “**The results are available here**”.

  The highest ranked docking pose is displayed together with its score on a scale.

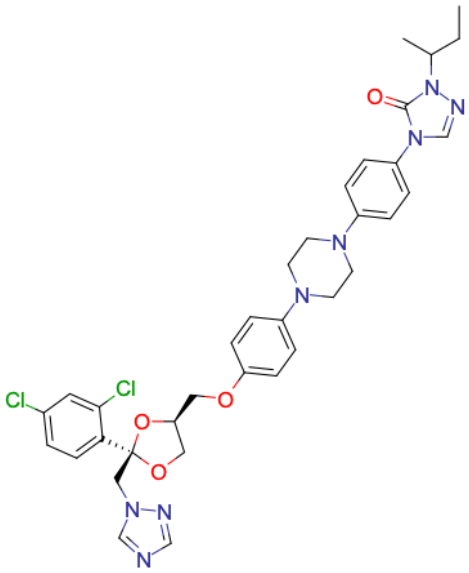
  - *With this information will you be able to optimize rationally high affinity molecules iteratively?* (Return to the **first tab** of your browser, click on **“Design your own molecule” box** and optimize the structure or restart from scratch!)

Your molecule has a score of: 10.7



Practice Session 3. Ligand-based virtual screening with SwissSimilarity

Before the actual workshop, we will perform a the simple preliminary virtual screening of a commercially available antifungal drug, **Itraconazole** (Sporanox®), which inhibits fungal 14α-demethylase, important for cell wall synthesis. We will search for similar molecules inside the library of all drugs approved by the FDA. This will illustrate the concept of ligand-based virtual screening, and show how to calculate enrichment. Here are the 2D structure and the SMILES of Itraconazole:

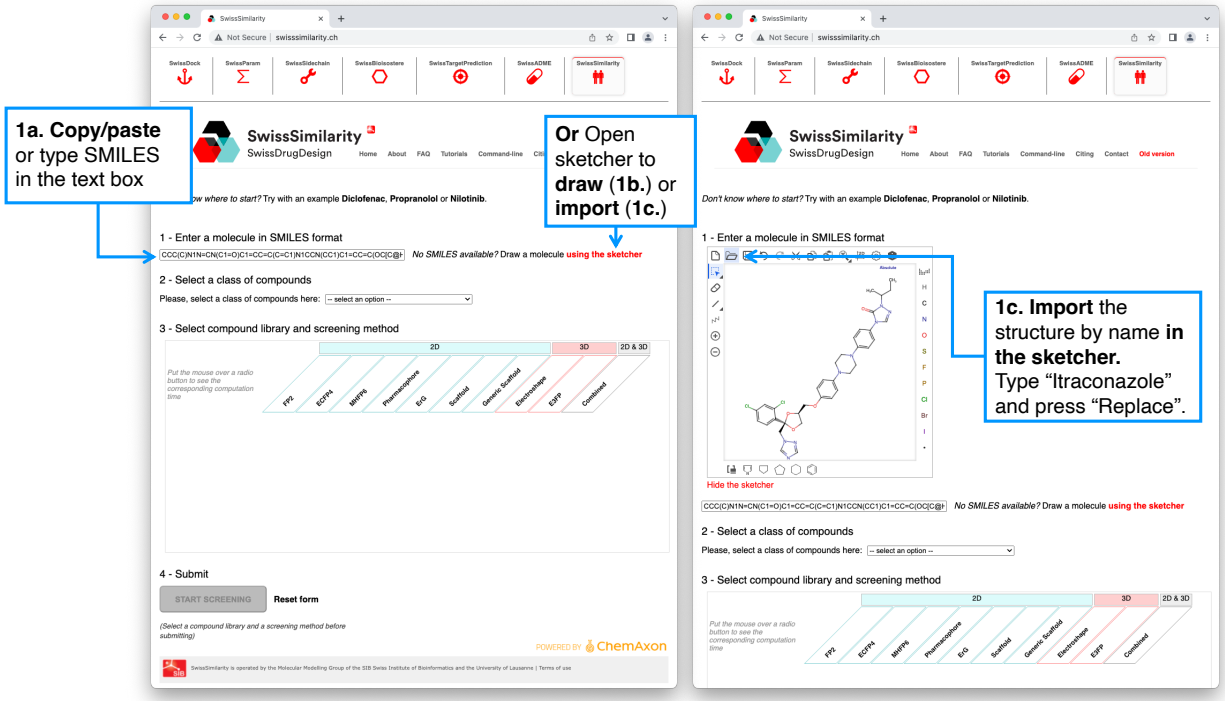


SMILES: `CCC(C)N1N=CN(C1=O)C1=CC=C(C=C1)N1CCN(CC1)C1=CC=C(OC[C@H]2CO[C@@](CN3C=NC=N3)(O2)C2=CC=C(Cl)C=C2Cl)C=C1`

- Access the web interface of SwissSimilarity in a web browser by typing the following URL http://www.SwissSimilarity.ch (preferably with *Google Chrome* or *Mozilla Firefox*).

1. Use one of the following options to **input the molecule** to be screened (i.e. Itraconazole):
   a. copy/paste (or type) the **SMILES** in the text box,
      or click on "using the sketcher" to :
   b. **draw the structure** in the sketcher, **or**
   c. **import the structure** (by name) in the sketcher.

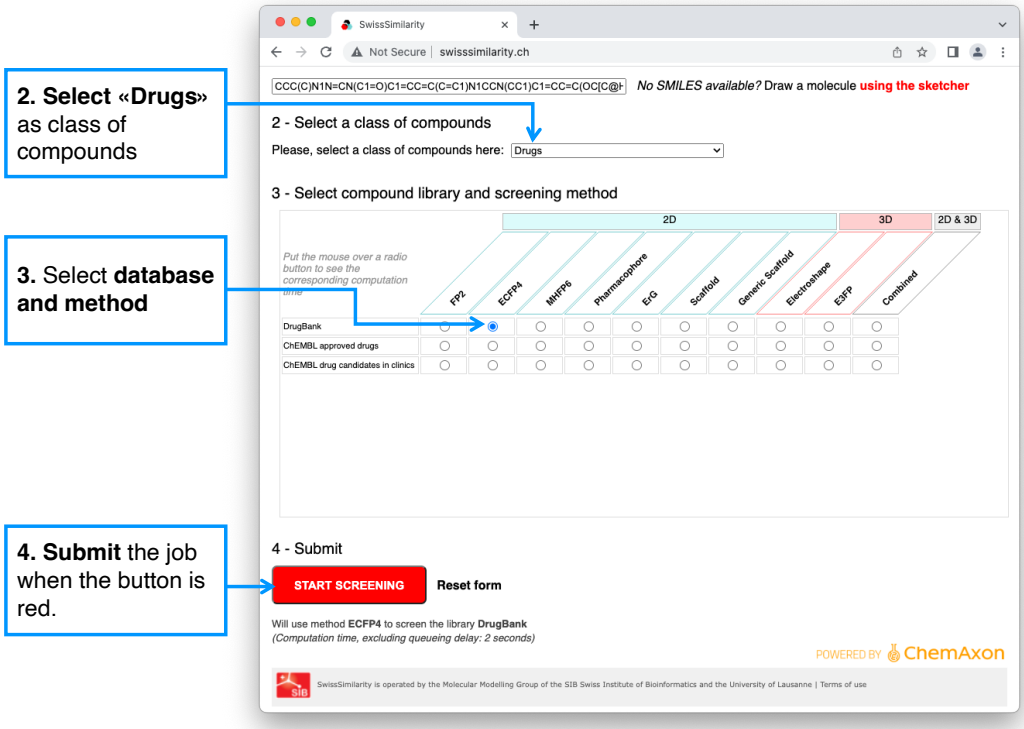
   Please note that the SMILES in the text field and the 2D structure are synchronized.



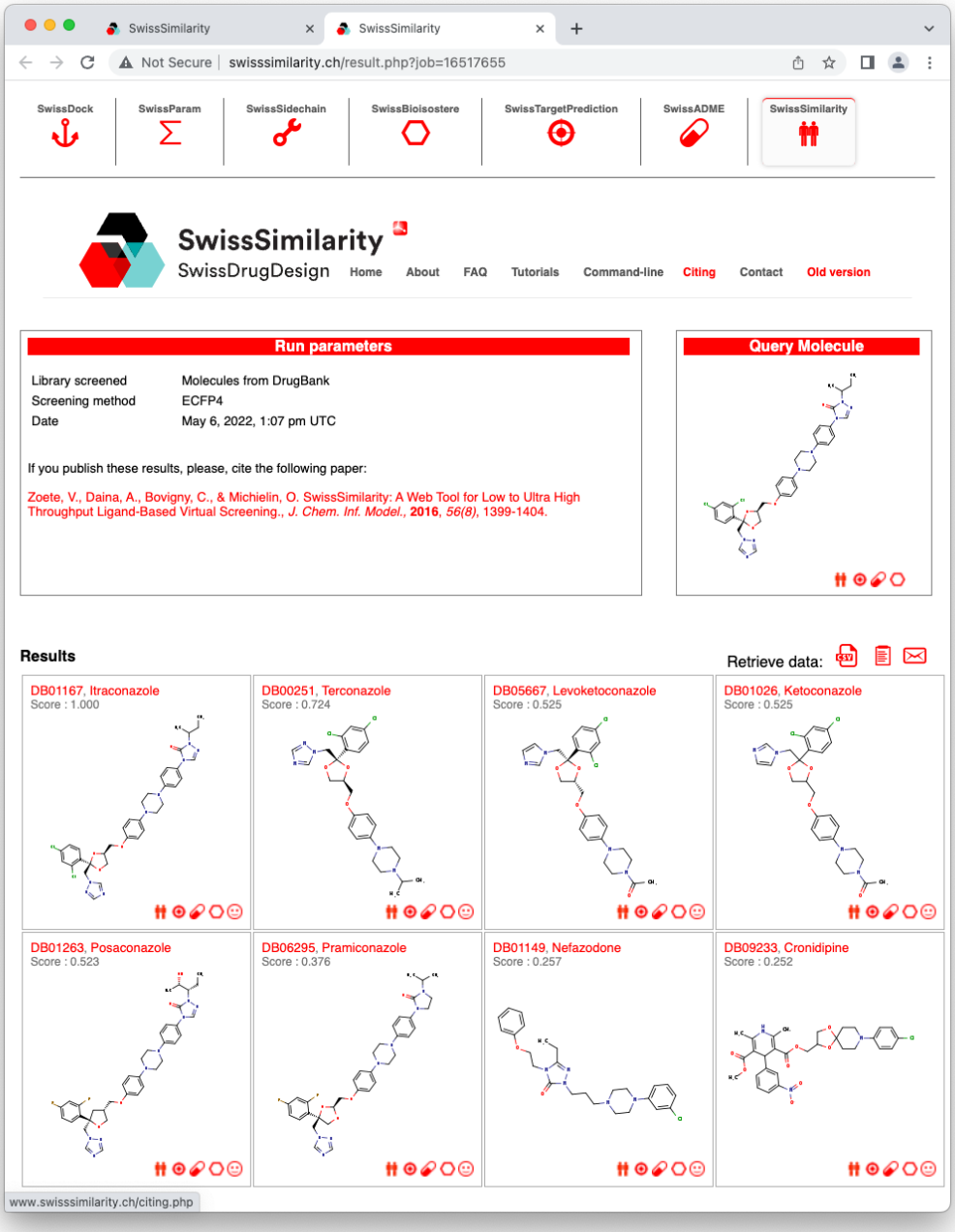
2. Select “Drugs” as the class of compounds in the drop-down menu.

3. In the dynamic table of databases and screening methods, click on the radio button to **select “DrugBank”** to search for similar compounds to Itraconazole using “**ECFP4**” fingerprints.

4. Submit the computation by clicking the **“START SCREENING” button**, which turns red only upon the input of molecule and selection of database / method. A text describing the job to be launched appears below the button as well.



After a few seconds, the DrugBank compounds most similar to Itraconazole, from a chemical point of view, are displayed in a new tab.



- Looking at the output page, let's answer these questions:
  - *How many DrugBank drugs have been found?*
  - *Are these structures chemically similar to the one of Itraconazole?*

- Keep this page open for the next exercise.

Exercise 7 (preliminary). Enrichment of antifungal drugs molecules in FDA-approved drugs.

Calculate the **enrichment factor** (EF) obtained by screening the DrugBank database for Itraconazole, using ECFP4 fingerprints.

There are **2726 FDA-approved drugs** in the DrugBank database, among which **28 are antimycotic azole derivatives** targeting fungal 14α-demethylase**.** All their names **end with** suffixes **-conazole or -imazole**.

The rate of antimycotic azoles in the whole DrugBank database is $\mathbf{r_{db} = 28 / 2726 = 0.01.}$

*Calculate the enrichment factor (**EF**) at **top 8** (i.e. at 0.3%)*

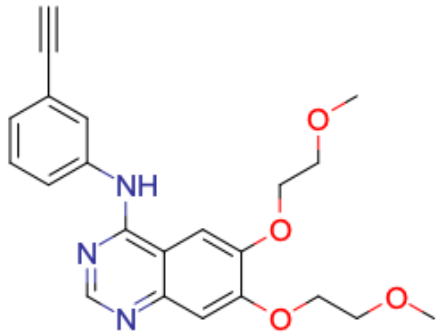
- *How many of these* antimycotic azoles *can be found in the top 8 of your screening* **(n_{screen})** ?
- *The rate at top 8:* $\mathbf{r_{screen} = n_{screen} / 8}$

$$EF = \frac{r_{screen}}{r_{db}}$$

This first preliminary exercise is finished. You can close all windows and tabs.

Let's start with the workshop!

Erlotinib (Tarceva®) is an inhibitor of the epidermal growth factor receptor (EGFR, erbB1, uniprotID P00533) tyrosine kinase. It has been approved by the FDA as an anti-cancer drug to treat non-small cell lung cancer, pancreatic cancer and several other tumor types.
The 2D structure and SMILES of Erlotinib are displayed below:



SMILES: `COCCOC1=CC2=C(C=C1OCCOC)C(NC1=CC=CC(=C1)C#C)=NC=N2`

The objective of this exercise is to use the tool **SwissSimilarity** to screen the collection of active molecules from the ChEMBL database in order to retrieve compounds similar to Erlotinib in terms of chemical structures (2D) and in terms of shape (3D).

- Access the web interface of SwissSimilarity in a web browser by typing the following URL http://www.SwissSimilarity.ch.
  (Please note, a detailed video tutorial is available at:
  http://www.swisssimilarity.ch/tutorials.php)

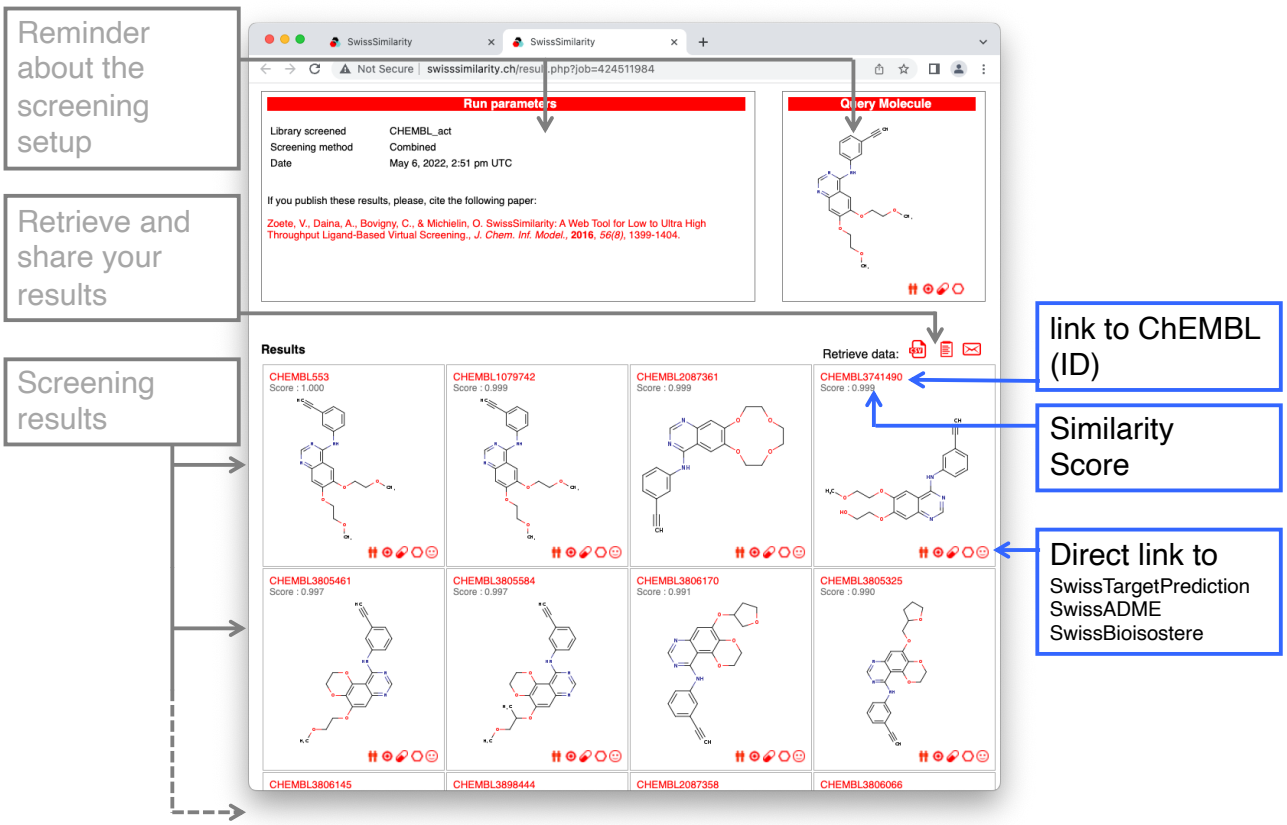
- Use one of the following options to **input the molecule** to be screened against (i.e. Erlotinib):
  a. copy/paste or type the SMILES in the text box, or
  b. draw the structure in the sketcher, or
  c. import the structure (by name) in the sketcher.

- Select "**Bioactive**" as the class of compounds in the drop-down menu.

- Click on the radio button in the table to **select both the screening method and the library** to screen. Here it is proposed to evaluate the library "**ChEMBL (actives only)**" for similarity with Erlotinib using a "**Combined**" method. The latter makes the process screen the library with both 2D-fingerprint (FP2) and 3D-electroshape (ElectroShape) to return a consensus value, corresponding to a combined score obtained by logistic regression of both individual (2D and 3D) similarity measurements.

- Start the screening by clicking the **"Submit" button**, which turns red only upon the input of a reference molecule and the selection of a database/method. A text describing the job to be launched appears below the button as well.

- After less than one minute, your **screening results** should appear in a new tab of your web browser. In the meantime, you can follow the progression of the calculations thanks to the blue bar.



📣 Given those results let's try to answer the following questions about compound CHEMBL461792:

- *What is the similarity score and ranking of compound CHEMBL461792?*
- *What are the two structural differences between CHEMBL461792 and Erlotinib?*
- *Which of these chemical modifications makes CHEMBL461792 more rigid than Erlotinib?*
- *Any clue about the potential benefit to test a more rigid ligand?*

Please, KEEP this SwissSimilarity result page OPEN as it is the starting point for the following sessions.

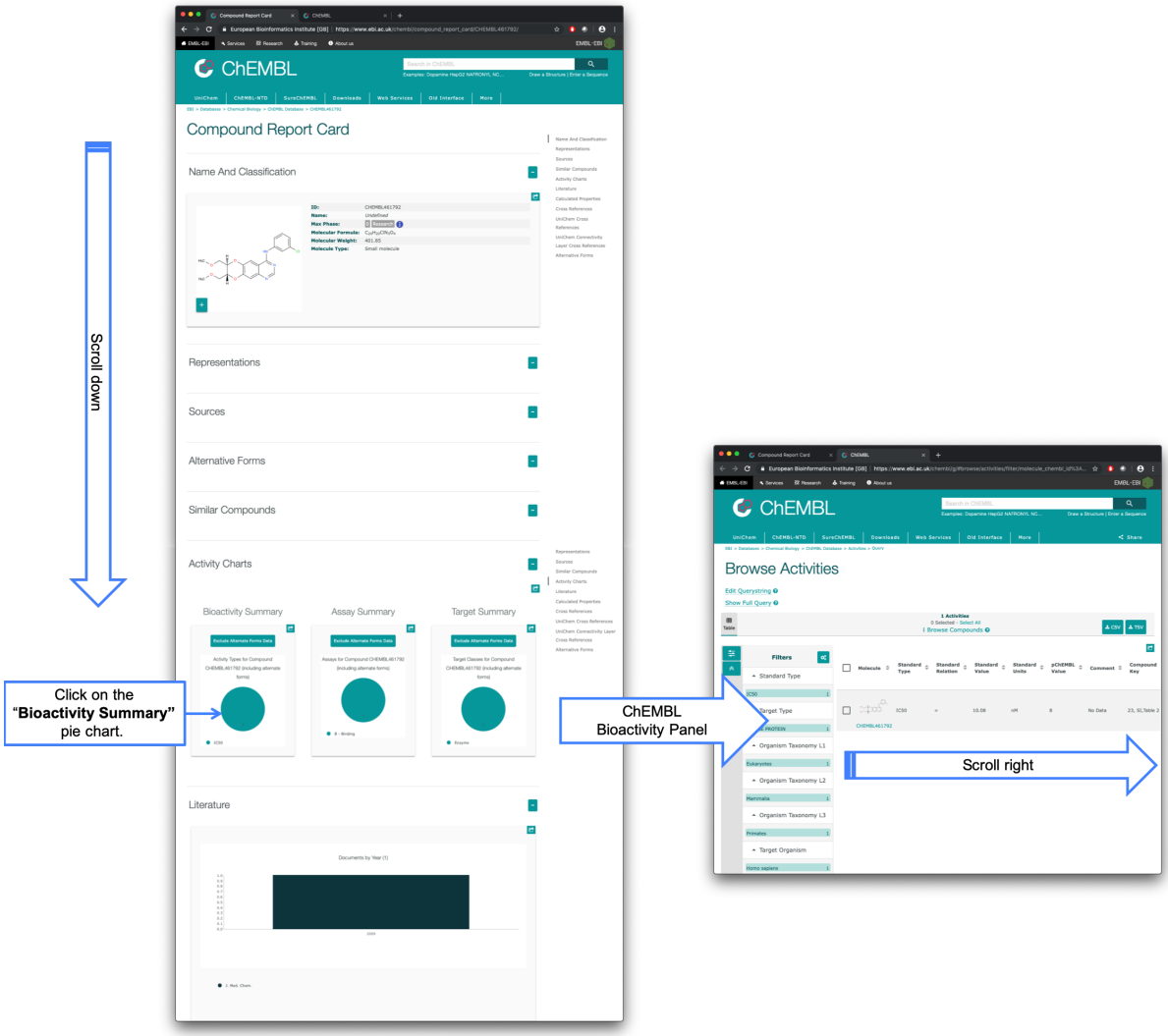
You can copy, bookmark or email the URL of your SwissSimilarity result page.

Practice Session 4. Reverse screening with SwissTargetPrediction

The purpose of this session is to analyze the biological activity of compound CHEMBL461792, which shows significant similarity with Erlotinib. The first actions illustrate the interoperability of the different SwissDrugDesign tools and the link with external resources.

1. Let's go back to the **SwissSimilarity result page** in your web browser and access to ChEMBL database for entry CHEMBL461792 (by **clicking the ID link**).

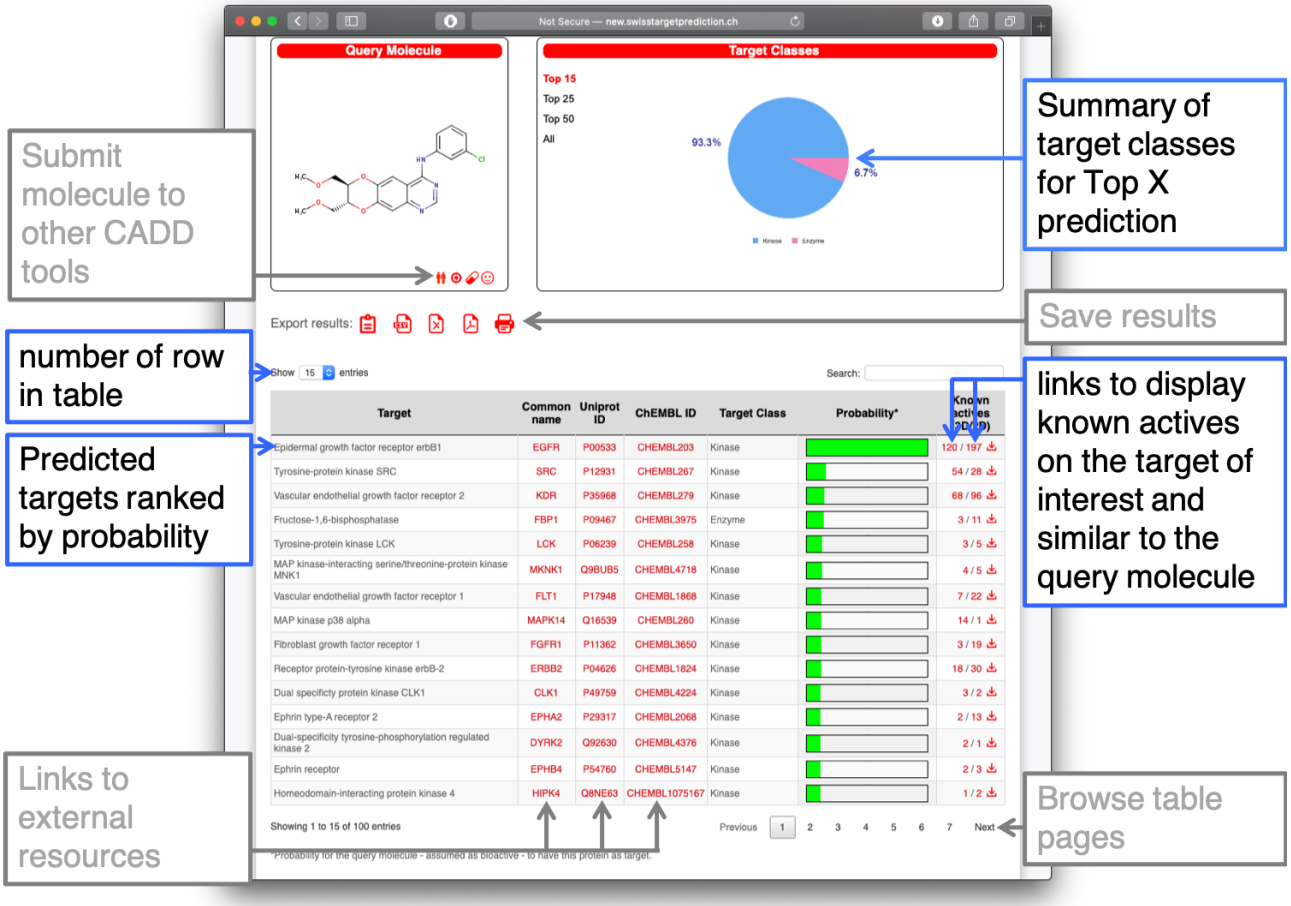
2. A first **ChEMBL** panel corresponding to entry CHEMBL461792 opens. Scroll down to the "Activity Charts" section and click on the "**Bioactivity Summary"** pie chart.
   This brings you to a second "ChEMBL Bioactivity" panel.



Confirm that compound CHEMBL461792 has been tested on EGFR and note the IC_{50}.

3. Go back to the SwissSimilarity result page in your web browser. Submit compound CHEMBL461792 to **SwissTargetPrediction for reverse screening to predict protein targets**, either by
   a. clicking the corresponding "**target" icon** from the SwissSimilarity results page, or
   b. by going directly to http://www.swisstargetprediction.ch in a new tab or window. In this case, you will need to draw the chemical structure in the sketcher or to copy/paste the SMILES in the dedicated text box, and finally click on the "Submit" button.

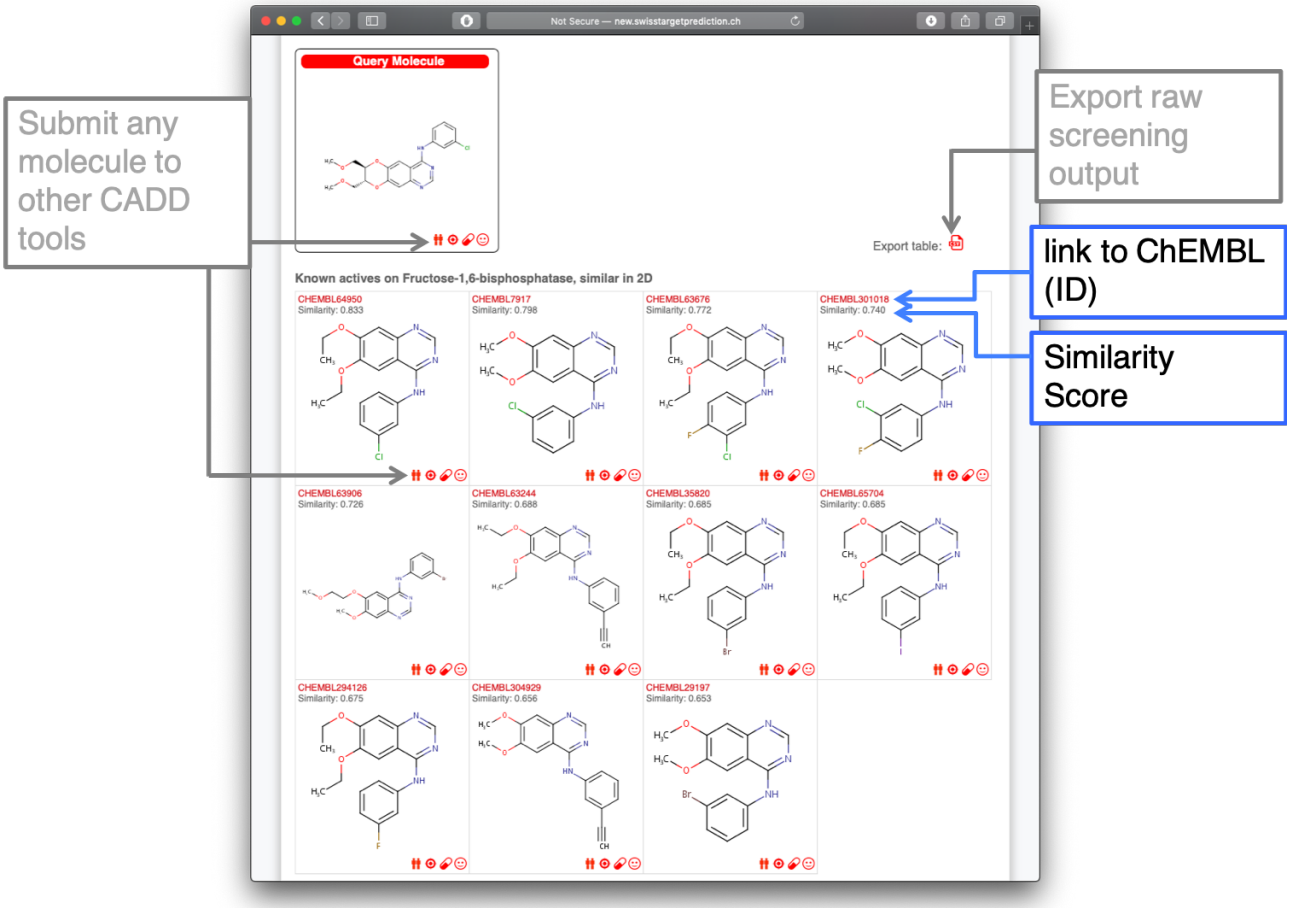
- After less than one minute, the **target prediction results** should appear in your web browser.



📣 With those results let's try to answer the following questions about compound CHEMBL461792:

- *How many protein targets are predicted in total (select* Show "**All**" entries *on the menu above the table). What is the proportion of kinases (click on "***All***" left to the pie-chart)?*
- *What is the most probable protein target for this compound? Can we consider this result as an actual prediction? Why?*
- *Looking at the ranking, what is the most probable non-kinase target?*
- *How many known actives of this non-kinase protein are similar to molecule CHEMBL461792 based on 2D chemical structure similarity? Same question for 3D shape similarity?*

4. By **clicking the number of similar molecules** for a given target (either from 2D or 3D screenings), a second window opens with a complete description of the outcome of the reverse screening. The structure of similar molecules active on the target of interest (those having driven the prediction) is displayed.



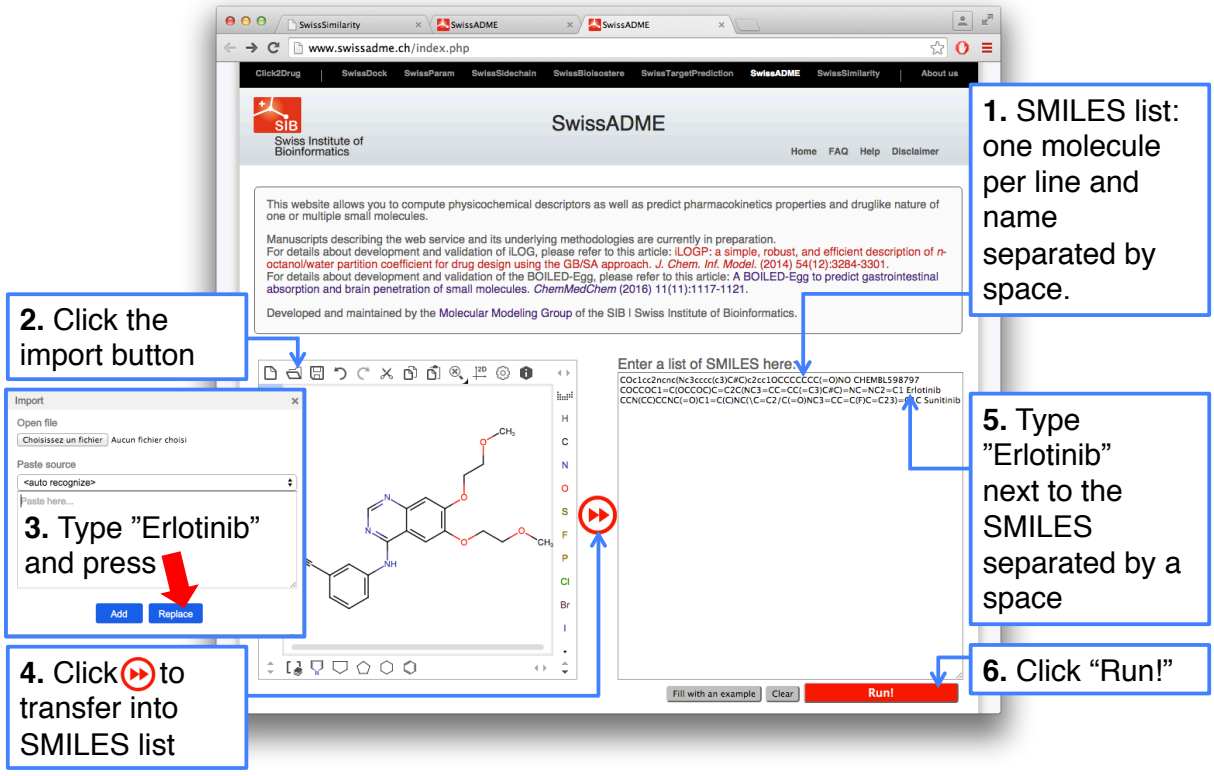
- *Are the actives most similar in 2D and in 3D to CHEMBL461792 the same compounds?*

Practice Session 5. Pharmacokinetics evaluation with SwissADME

During this session, it is proposed to estimate the druglikeness and two important ADME parameters regarding absorption and side effects: the gastrointestinal passive absorption and the brain penetration, respectively. To this end, the SwissADME web tool will be used to analyze Erlotinib and an analogue molecule (CHEMBL598797) together with another kinase inhibitor: the anti-cancer drug Sunitinib.

- Let's go back to the **SwissSimilarity result page** in your web browser. Locate compound CHEMBL598797 (score: 0.961) and click on the corresponding "**pill" icon** to directly launch a SwissADME calculation.

- A new tab opens with SwissADME calculating for about 3 to 10 seconds. Upon completion the output panel related to CHEMBL598797 is displayed in the same page. Now let's add two anti-cancer drugs Erlotinib and Sunitinib to the list and let's rerun the calculation.

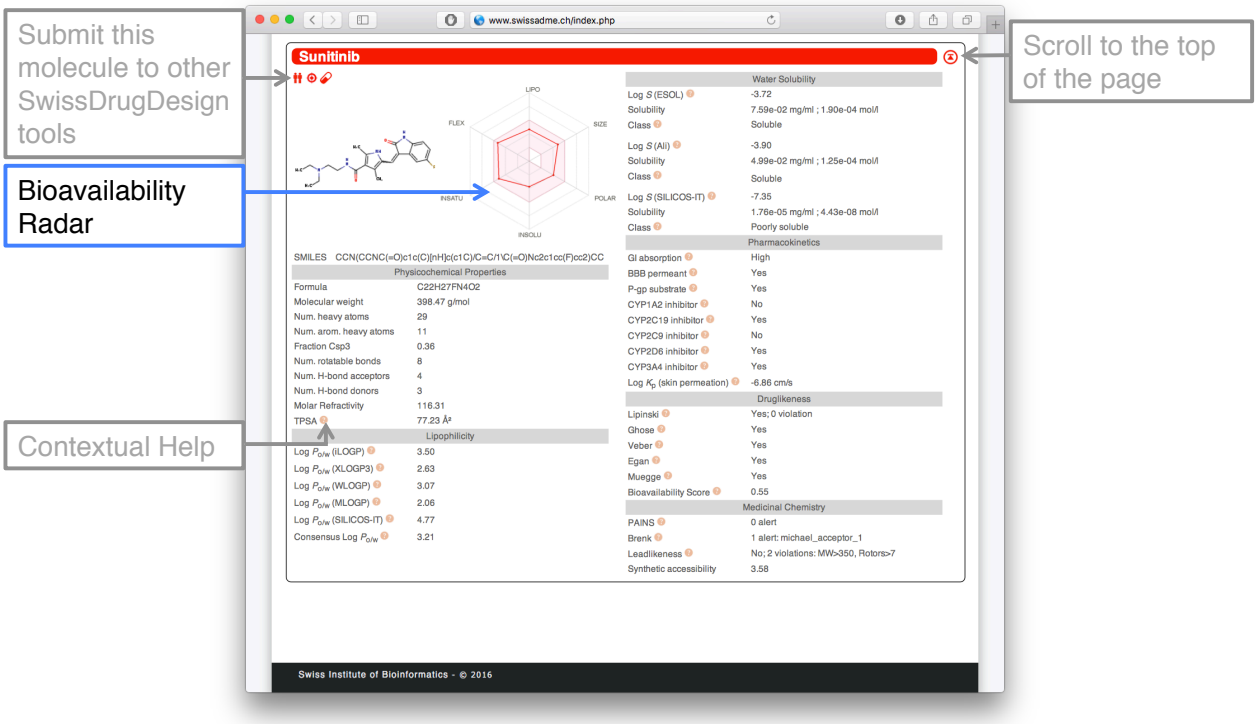


1. Scroll to the top of the SwissADME page and in the SMILES list box, type the name of the compound "CHEMBL598797" next to the SMILES, separated by a **space**. Then press **enter** to go to the next line.
2. Click on the "**Import**" button of the sketcher (second button top-left).
3. Type "Erlotinib" in the pop-up "Paste source" box and click the "**Replace**" blue button.
4. The structure of Erlotinib appears in the sketcher. Click on the **double-arrow button** in between the sketcher and the SMILES list to transfer the molecule into the list.
5. Type "Erlotinib" next to the SMILES on the second line, separated by a **space**.

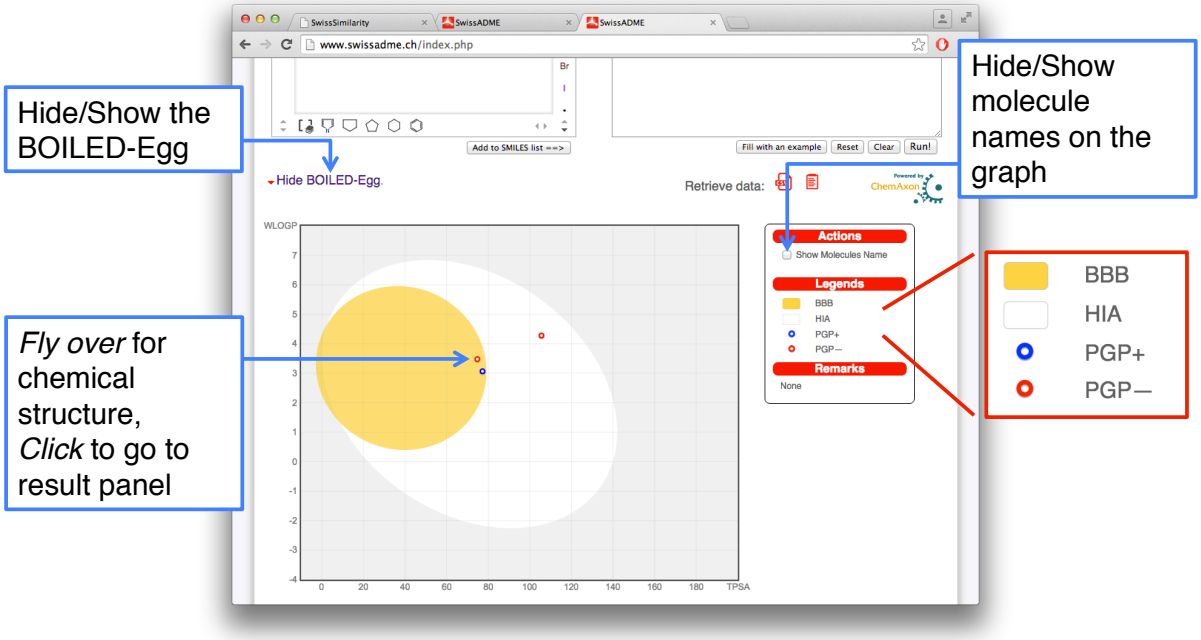
- Repeat steps 2, 3, 4 and 5 for Sunitinib (*no matter if there are blank lines in the text box*)

6. Click on the "**Run!**" button at the bottom of the list.

- Upon calculation completion (3 to 10 seconds per molecules), the results appear in the same page; three panels per molecule, each summarizing the values for a given compound.



- By clicking on the "**Show BOILED-Egg**" red button (below the sketcher), the graphical output is displayed on the same page



- With those results let's try to answer the following questions about the ADMET of those three molecules:
    - *One of these compounds is predicted toxic, can you point out which one and the alert related to this prediction?*
    - *Between both marketed drugs, which one of Sunitinib or Erlotinib is more prone to create drug-drug interactions linked with metabolism?*
    - *Which of these molecules is the less druglike? What is the molecular property responsible for that?*
    - *Are all three compounds predicted as well-absorbed by the gastrointestinal tract when administered orally?*
    - *Qualitatively, what is the propensity for each compound to passively cross the blood-brain barrier?*
    - *Which is the physicochemical property mostly explaining the difference in passive brain permeation behavior?*
    - *Which compound(s) is (are) predicted to be actively pumped out from the central nervous system? Why?*
    - *Finally, which compound has the highest probability to be in significant concentration in the brain?*

Exercise 8. Pharmacokinetics optimization of EGFR inhibitor.

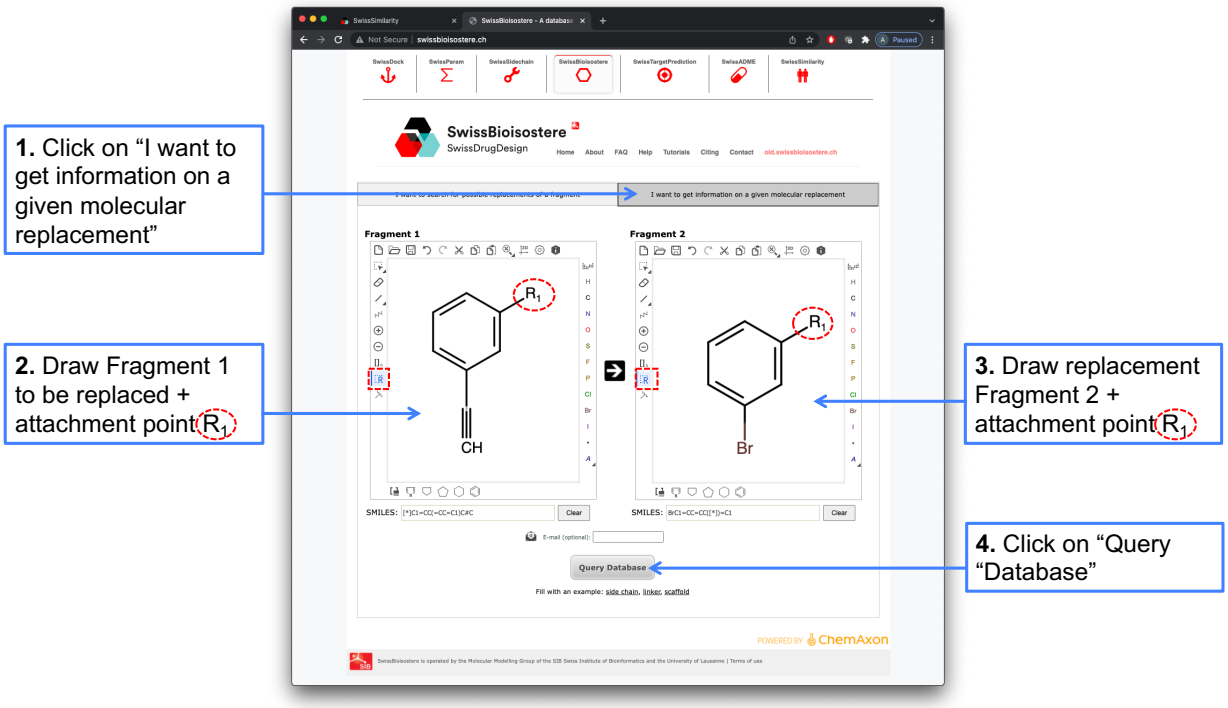
- Imagine that your endeavor consists in optimizing the properties of CHEMBL598797, which has to inhibit a kinase located in the central nervous system (CNS). Try some small chemical modifications (e.g. copy/paste SMILES in the sketcher, apply chemical modifications and transfer multiple entry lines to the SMILES list). You have so initiated an iterative optimization process. Once you are happy with the ADMET properties, click on the target icon to submit your optimized molecule to SwissTargetPrediction.
    - *Describe your optimization strategy.*
    - *What are your conclusions regarding pharmacokinetics and pharmacodynamics?*

Practice Session 6. Bioisosteric design with SwissBioisostere

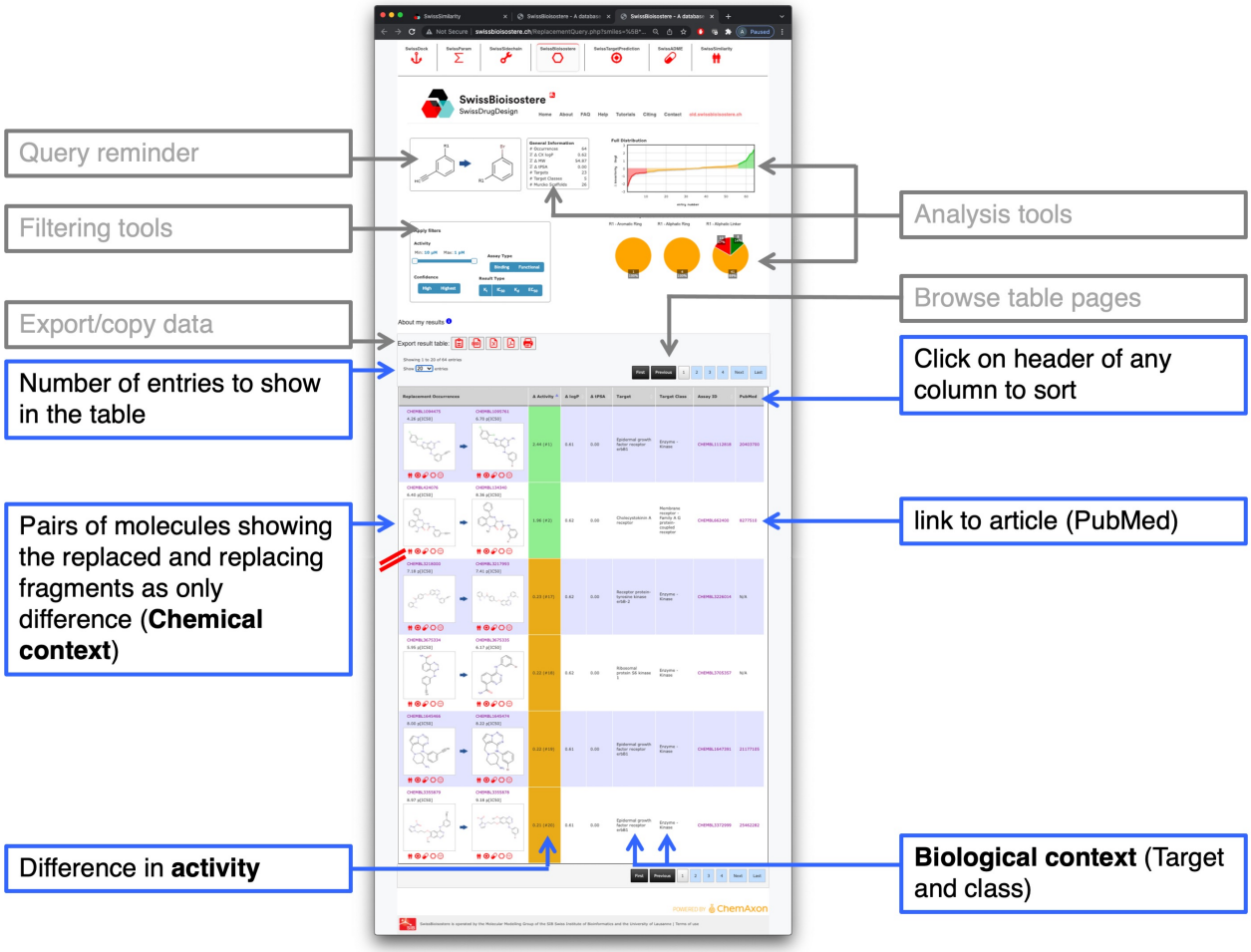
The purpose of this session is twofold. First it is proposed to analyze a specific replacement found in the virtual screening. Then we will suggest other possible bioisosteric replacements for ligand design.
Please note, detailed video tutorials are available at: http://swissbioisostere.ch/tutorials.html

- Go back to the **SwissSimilarity result page** in your web browser, locate compound CHEMBL2087361 (rank 3) and compound CHEMBL2087355 (rank 13). *Can you point out the chemical difference between those two molecules?*
- Let's investigate this specific molecular replacement. Type the following URL in a new tab or window of your web browser **http://www.SwissBioisostere.ch** to access the submission page of SwissBioisostere. Alternatively, you can click on the link in the black toolbar at the top of any SwissDrugDesign Website.

1. Click on the "**I want to get information on a given molecular replacement**" grey tab. A second sketcher "Fragment 2" appears on the right.
2. Draw in the *left* sketcher the fragment in CHEMBL2087361 that is replaced. Add the attachment point (R_1) with "smart R-group" in the left tool bar (numbering is automatic).
3. Draw the replacing fragment in CHEMBL2087355 in the *right* sketcher. Pay attention to add the same attachment point (R_1).
4. Start the search by clicking the button "**Query Database**" at the bottom of the page.



- After a few seconds the SwissBiosisostere output page is displayed in a new tab, compiling all examples found in the literature about the replacement of *m*-ethynylbenzene fragment by *m*-bromobenzene.

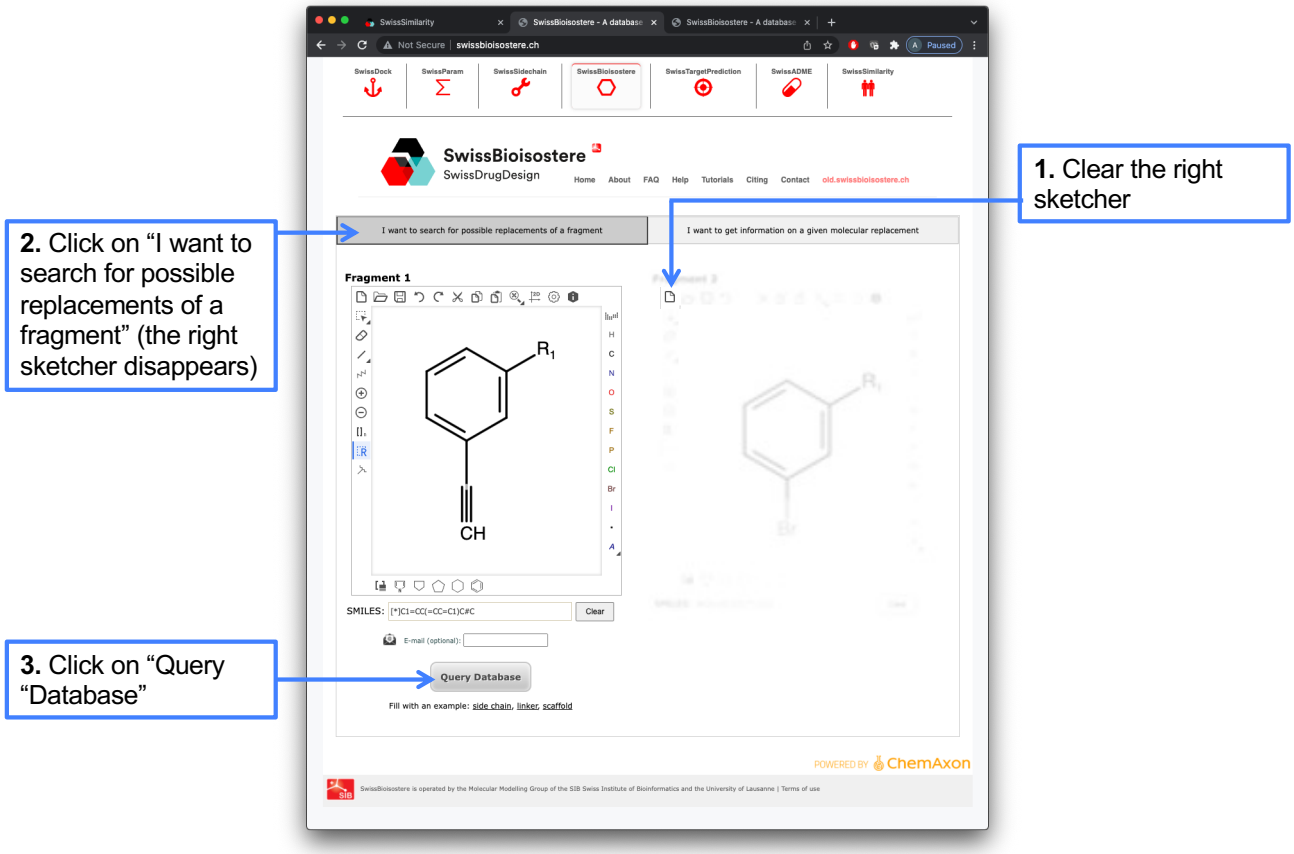


📣 Display all lines by setting "**All**" in the number of **entries to show** on the upper left corner of the result table. Let's try to answer the following questions:

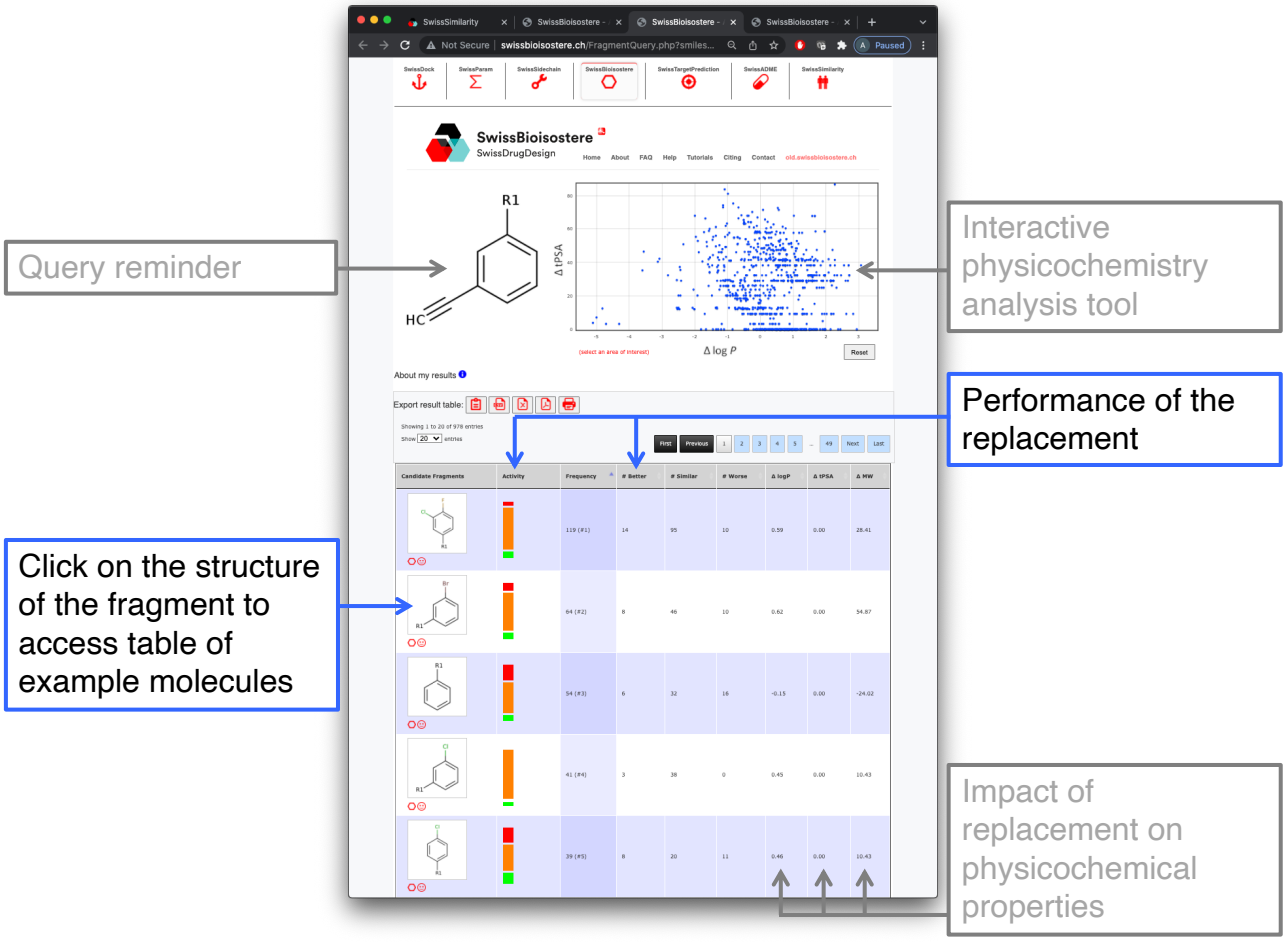
- *Overall, how many times this specific replacement was found in the literature?*
- *Generally speaking, what is the trend: increasing or decreasing or similar biological activity?*
- *In what biological context this replacement was mainly tried?*
- *How many times this replacement was found for compounds tested on our target of interest (Epidermal growth factor receptor erbB1)?*
- *What is the trend for activity on this specific protein?*
- *Can you find the entry corresponding to our case (CHEMBL2087361 to CHEMBL2087355)?*
- *Which molecule is the most potent on erbB1 among CHEMBL2087361 and CHEMBL2087355? Note that the activity is given as pIC50.*
- *Any idea why is the entry seems duplicated? Click on the PubMed link to get more info from the abstract.*

- Let's try to find **other** relevant **replacements** for *m*-ethynylbenzene. **Go back to the tab** of your browser where you have made the last SwissBioisostere request.

1. Clear the *right* sketcher by clicking on its most upper-left button.
2. Click on the "**I want to search for possible replacements of a fragment**" grey tab. The right sketcher disappears.
3. Verify that the *m*-ethynylbenzene is correctly drawn in the left sketcher and click on "**Query Database**" for SwissBiosiostere to search for all possible molecular replacements.



- After a few seconds the SwissBiosisostere output page listing all possible replacements of *m*-ethynylbenzene found in the literature is displayed in a new tab.



📣 With these results let's try to answer the following questions:

- *Can you find the replacement that we studied in detail in the previous section (CHEMBL2087361 to CHEMBL2087355)?*
- *Propose two other sensible replacements for* m*-ethynylbenzene in our biological/chemical contexts and explain how you found them.*

The workshop is finished. Hope you've enjoyed it!