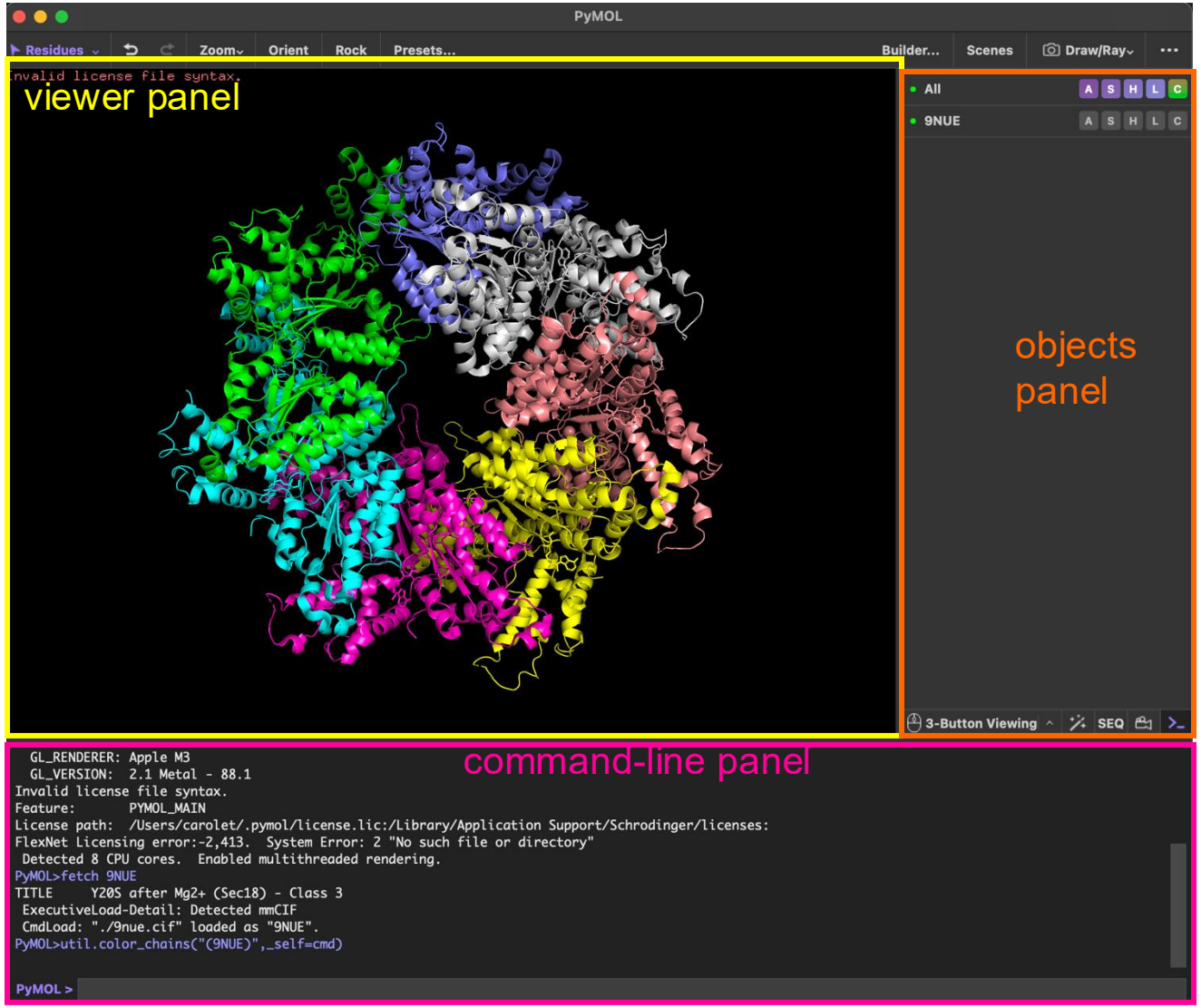

DTU

Introduction to Bioinformatics - 22111

PyMol tutorial


PyMOL
Residues
Zoom
Orient
Rock
Presets...
Builder...
Scenes
Draw/Ray
Invalid license file syntax.
viewer panel
All
9NUE
objects panel
3-Button Viewing
SEQ
command-line panel
GL_RENDERER: Apple M3
GL_VERSION: 2.1 Metal - 88.1
Invalid license file syntax.
Feature: PYMOL_MAIN
License path: /Users/carolet/.pymol/license.lic:/Library/Application Support/Schrodinger/licenses:
FlexNet Licensing error:-2,413. System Error: 2 "No such file or directory"
Detected 8 CPU cores. Enabled multithreaded rendering.
PyMOL>fetch 9NUE
TITLE Y20S after Mg2+ (Sec18) - Class 3
ExecutiveLoad-Detail: Detected mmCIF
CmdLoad: "./9nue.cif" loaded as "9NUE".
PyMOL>util.color_chains("(9NUE)",_self=cmd)
PyMOL >

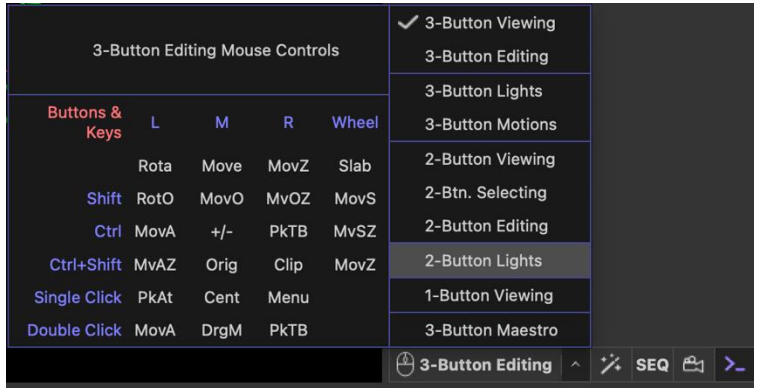
Commands and actions included in this tutorial

- fetch (load proteins)
- remove (delete objects, can be chain, residues, etc)
- select (select and name objects)
- color (color the full protein, separate chains, residues, color by secondary structure)
- align (align two chains or two structures)

Tasks – learn to configure your mouse!

- Load in the 1GCI structure

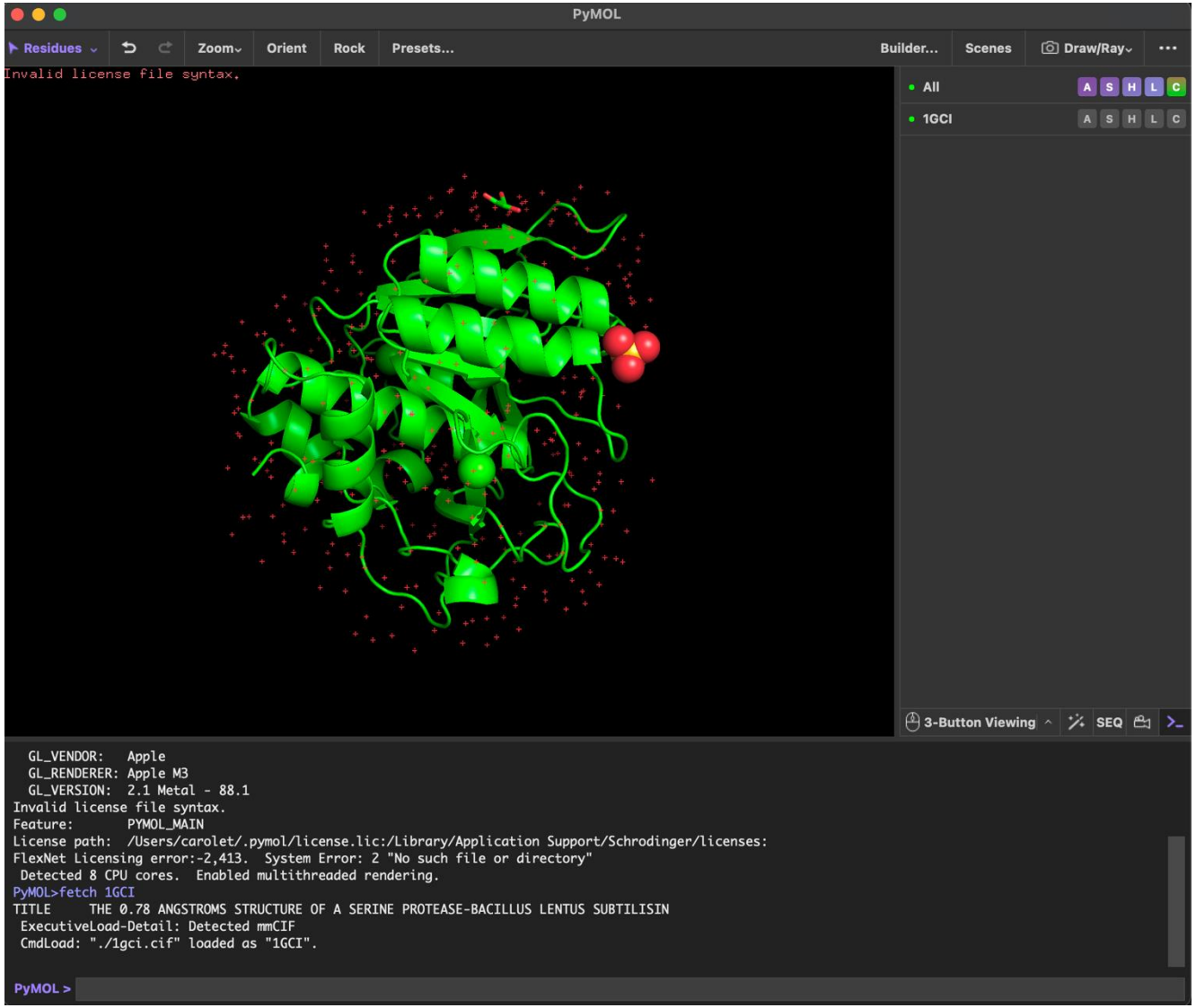
- Play around with the interface – learn how to rotate and zoom the structure:
  - Rotate: Click and hold left mouse button and move around
  - Zoom:
    - 1) Right click + move up/down (a bit slow)
    - 2) Shift + control + scroll wheel
    - 3) Shift + control + two finger drag on mouse pad



- Background info:
  - The structure is of the Novozymes peptidase “Savinase” that we have worked with before

  - PDB link: http://www.rcsb.org/pdb/explore/explore.do?pdbId=1GCI
  - UniProt link: http://www.uniprot.org/uniprot/P29600

How to load a protein structure?



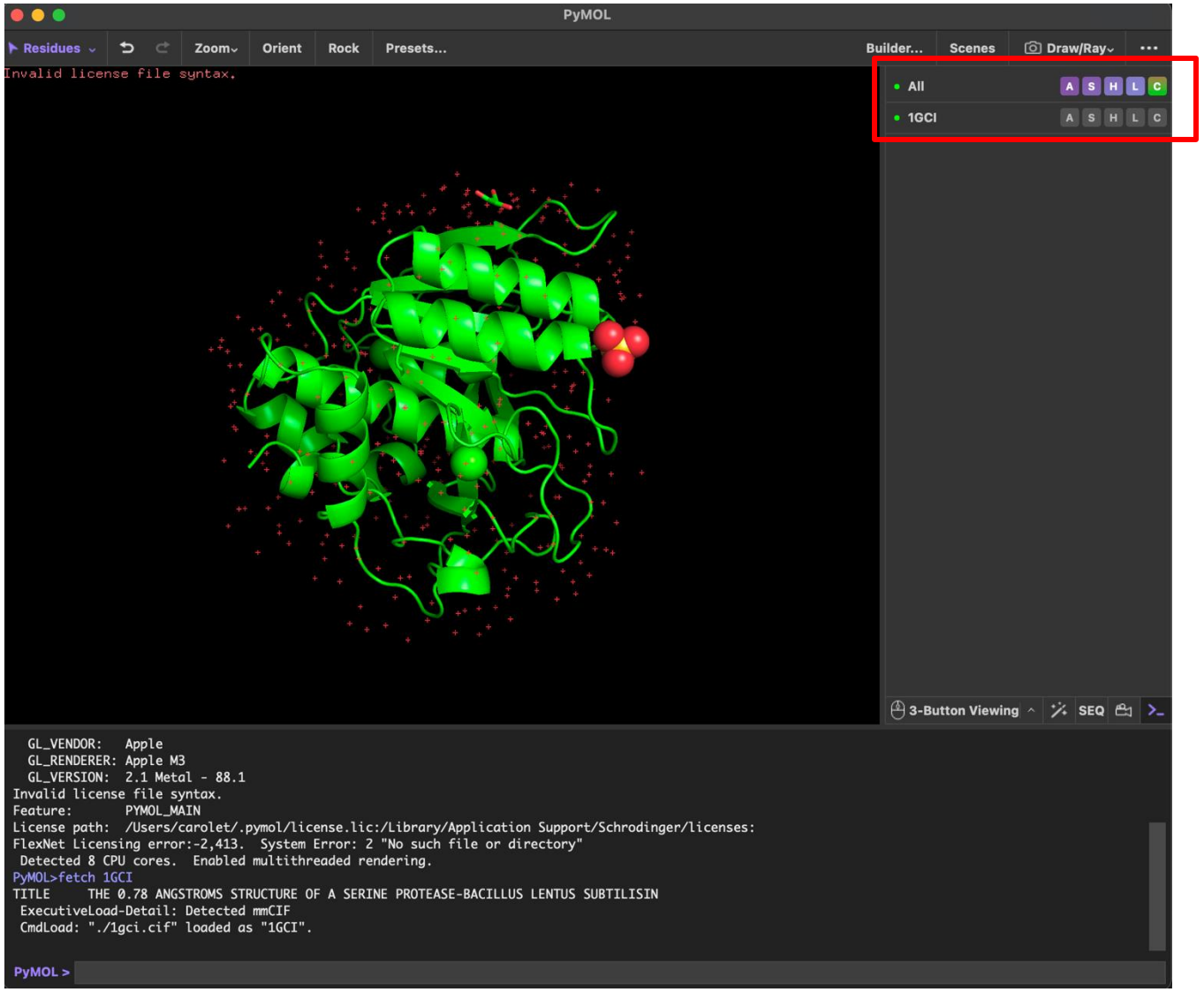
Structure: 1GCI (PDB)

0.78Å structure of Savinase

Two ways:

- Write "**fetch 1GCI**" in the command field
- Download .PDB file from pdb.org and use the File->Open menu

Working with the structure



A = Action

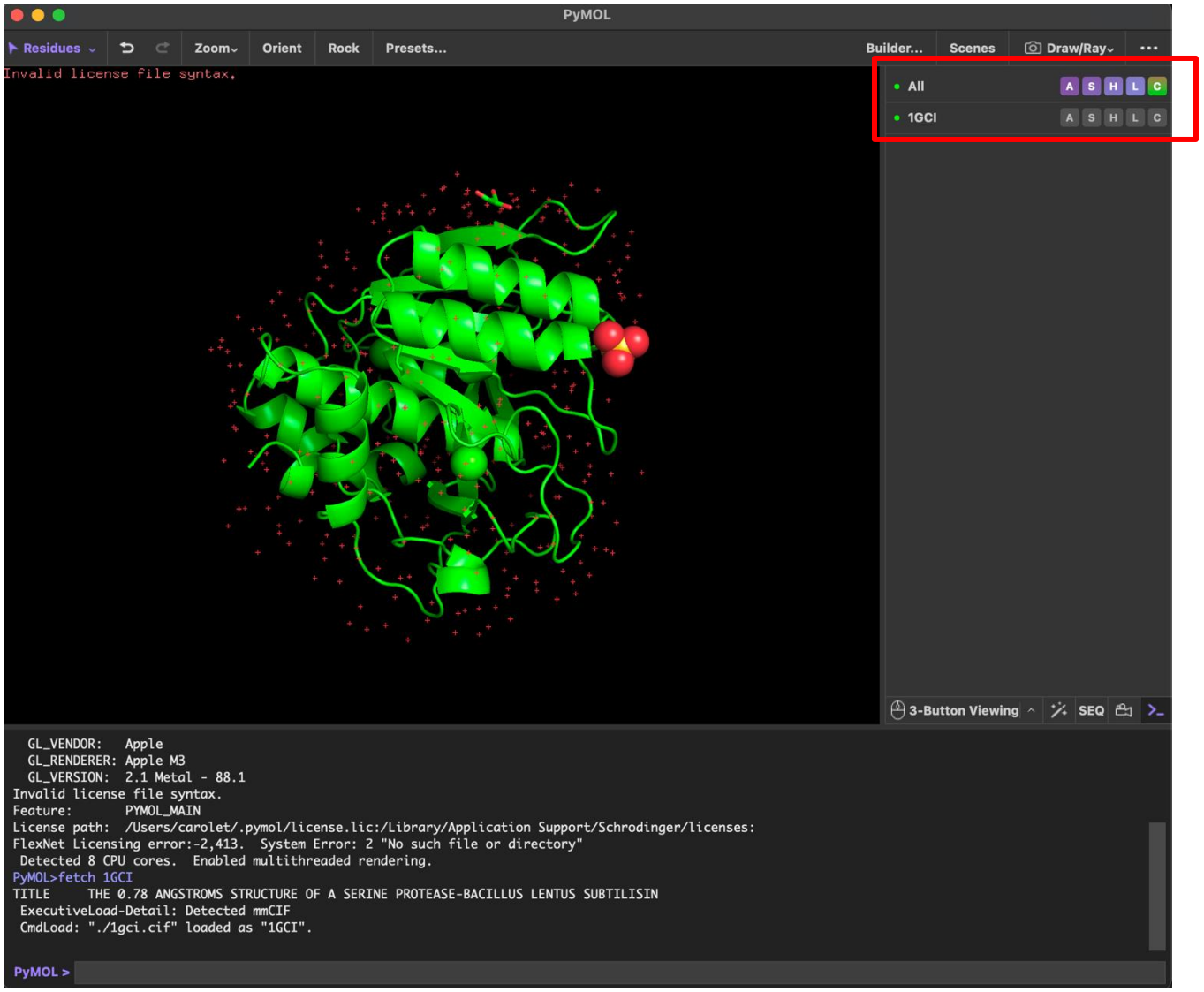
S = Show

H = Hide

L = Label

C = Color

Working with the structure



List of things to manipulate

"all": shortcut to ALL objects

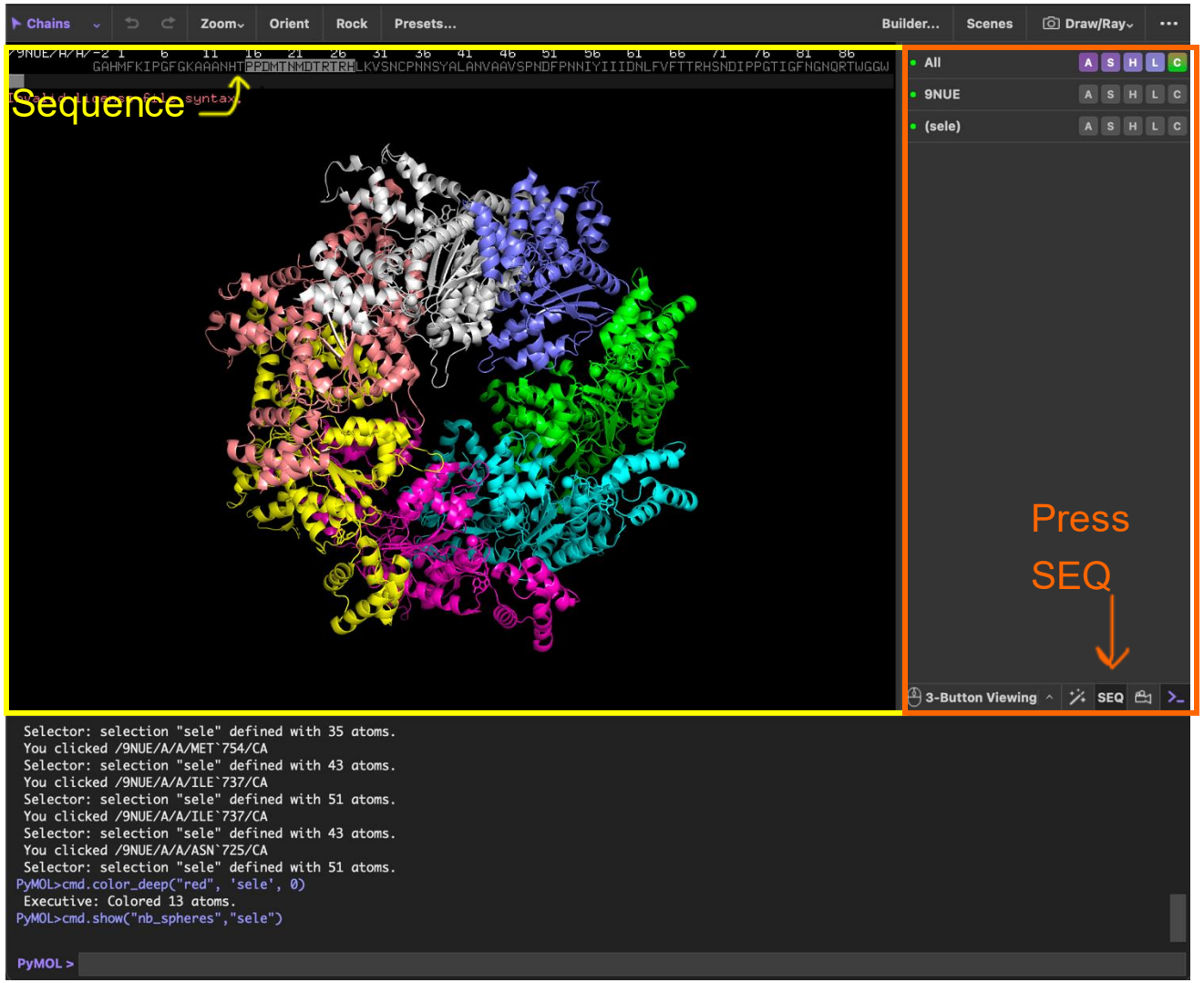
1GCI: the structure we have loaded

Multiple structures, peptide chains and selections can be in play, as we shall see later.

TASKS

- Play around with the coloring menu and figure out how to change the color of the entire structure (red, green, blue etc.)
- Next, figure out how to color according to the secondary structure, and select a scheme that will high-light 1) alpha helices 2) beta-strand 3) turns
- Rotate the structure to make it easier to see the different kinds of secondary structure
- Align different chains of the molecule
- *… we'll return to coloring, after we have learned how to select subsets of the structure …*

Show the sequence and color an element

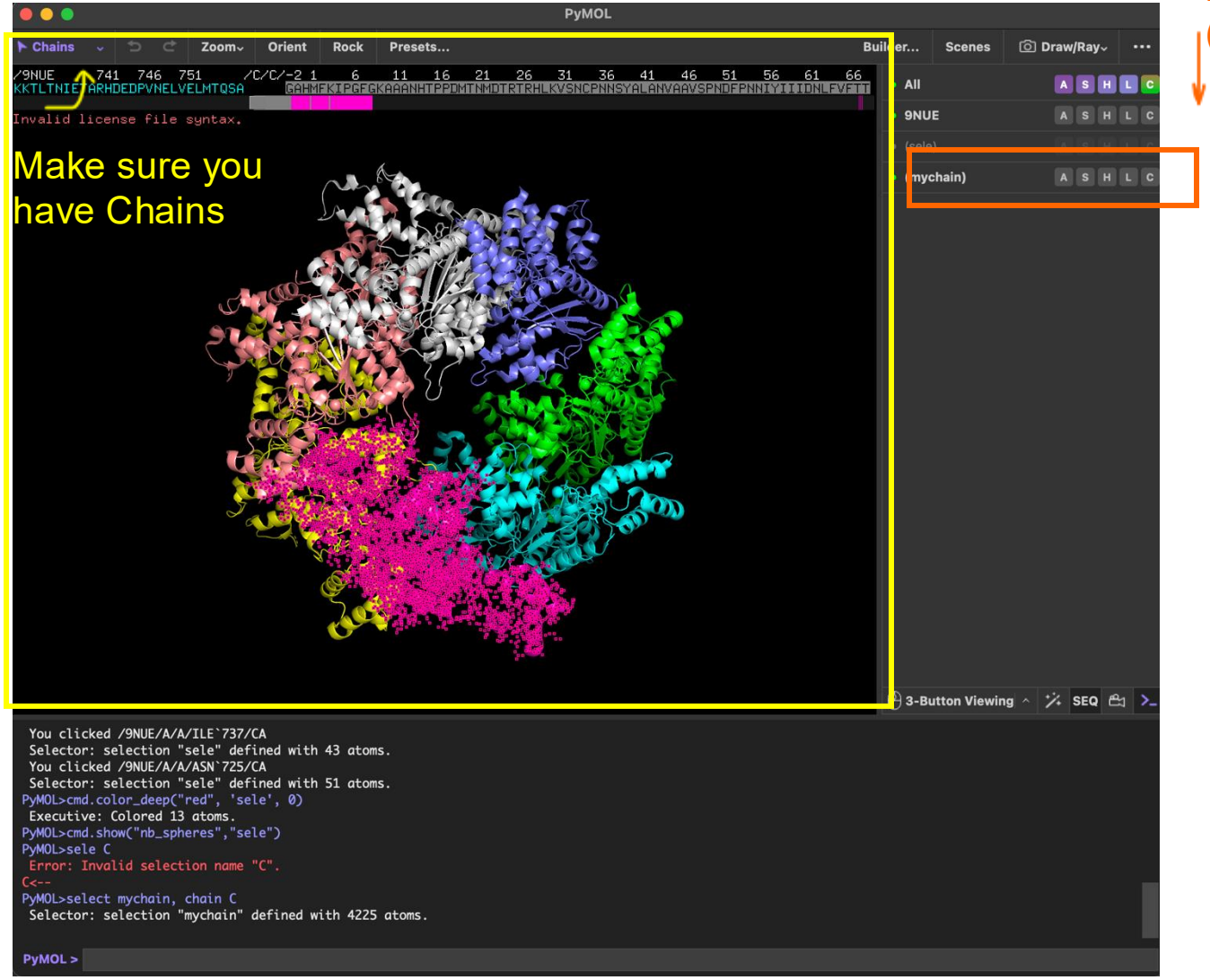


Structure: 9NUE (PDB)

Translocase from S.cereviseae

- Press SEQ and the sequence will appear in the viewer panel
- scroll to the left and check the name of the structure and chains.
- How many chains are in the structure?

Show the sequence and color an element



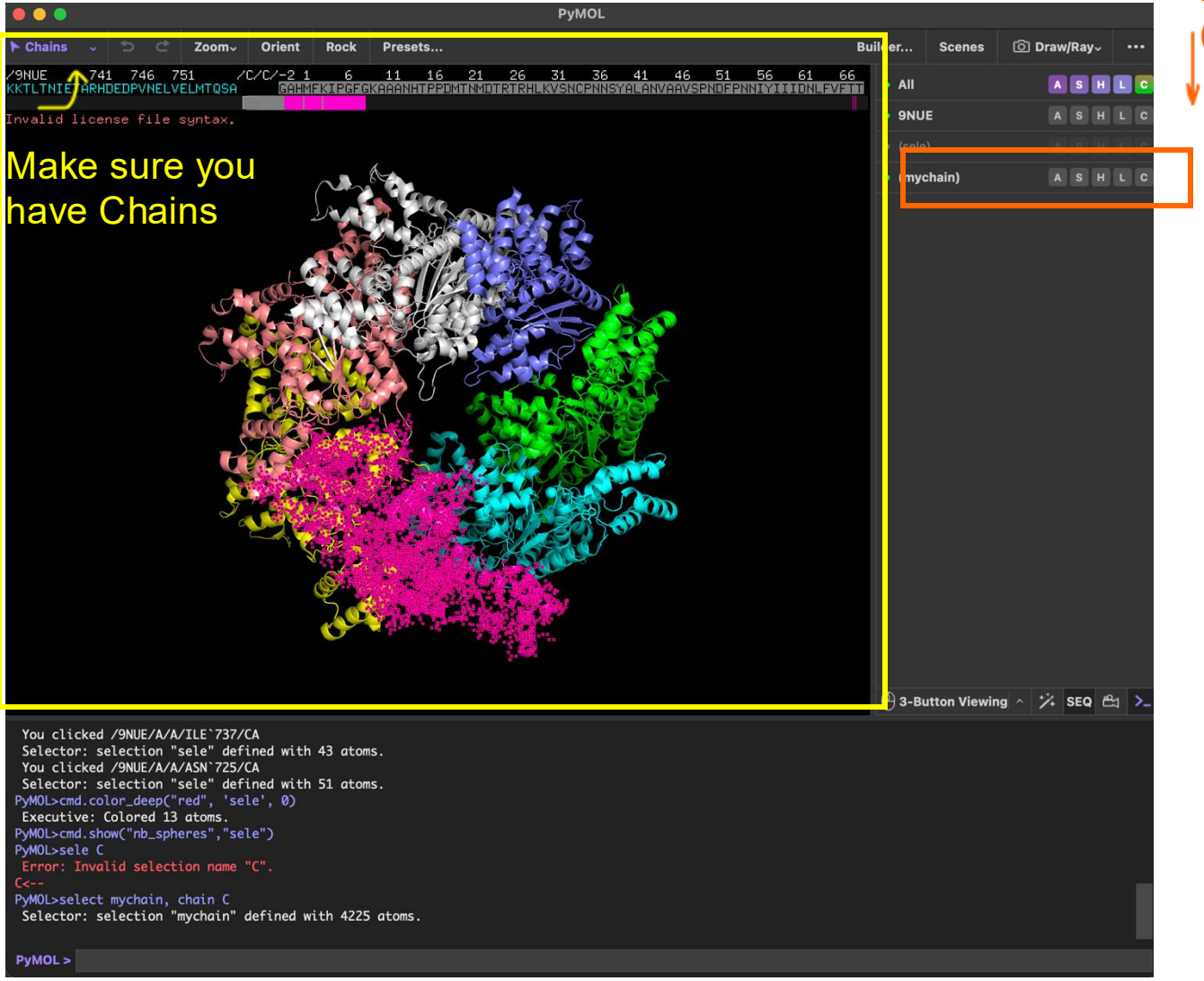
in the selection (sele) Press C

Structure: 9NUE (PDB)

Select chain C and color red

- make sure you have chains as selecting element (yellow)
- you can scroll on the sequence to find chain C
- then go to the objects panel and click C (color) and select red

Show the sequence and color an element



in the selection (sele) Press C

Structure: 9NUE (PDB)

Select chain C and color red

Type in the Command line panel

OPTION 1

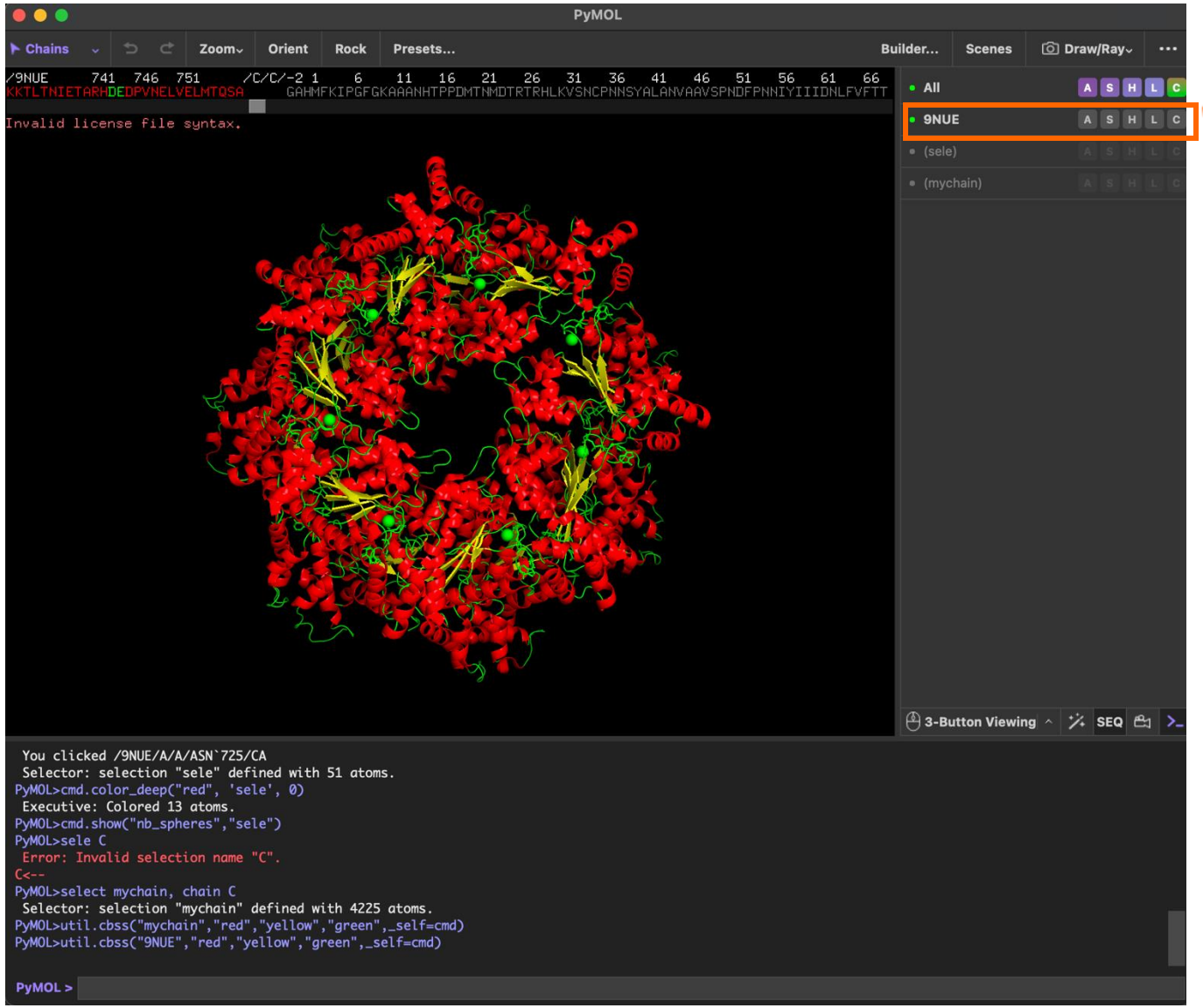
- select mychain, chain C
- color red, mychain

or OPTION 2

- color red, chain C

What are the differences?

Color by secondary structure

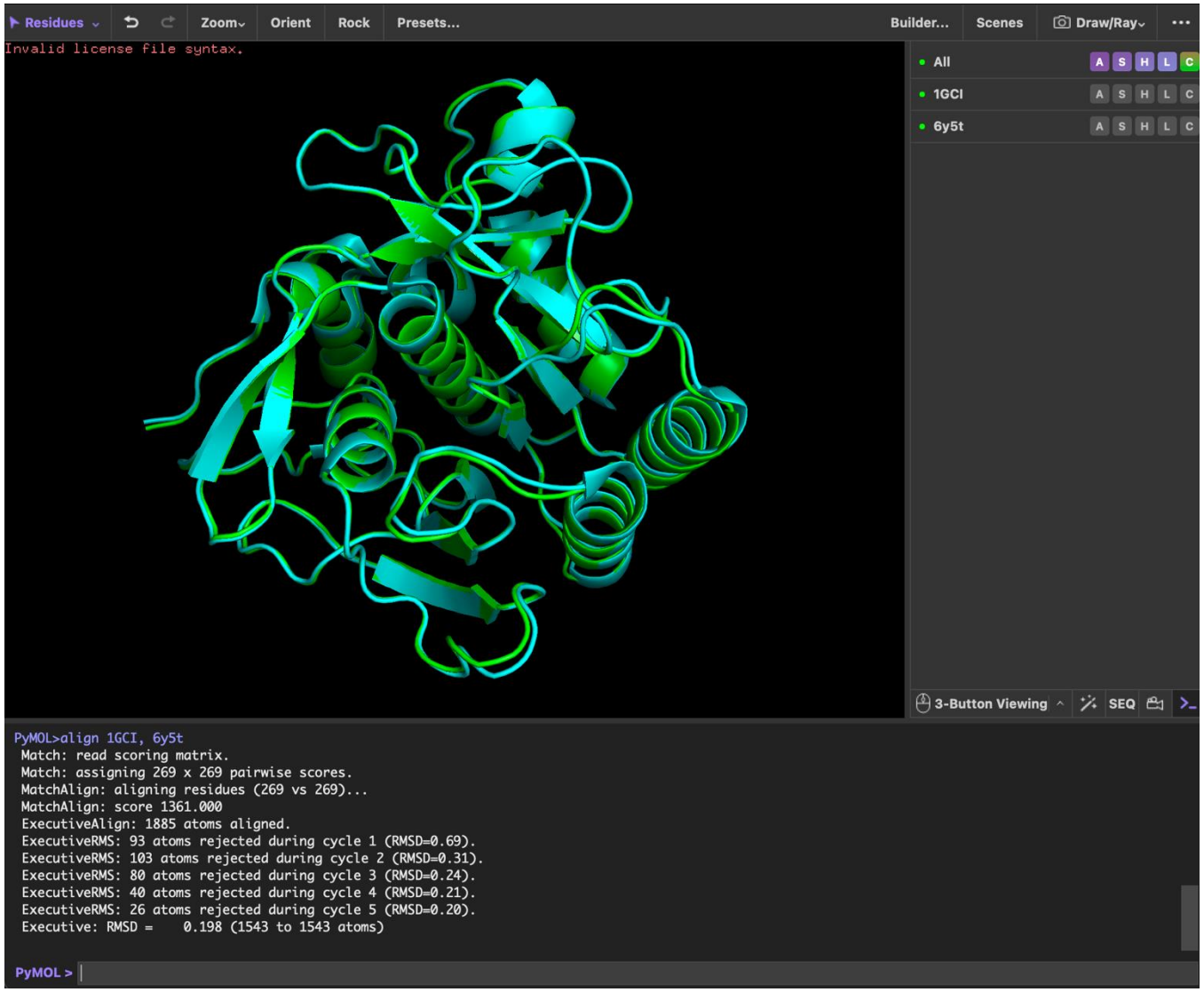


in the structure name press C

Structure: 9NUE (PDB)

color by secondary structure (ss)

Align two protein structures

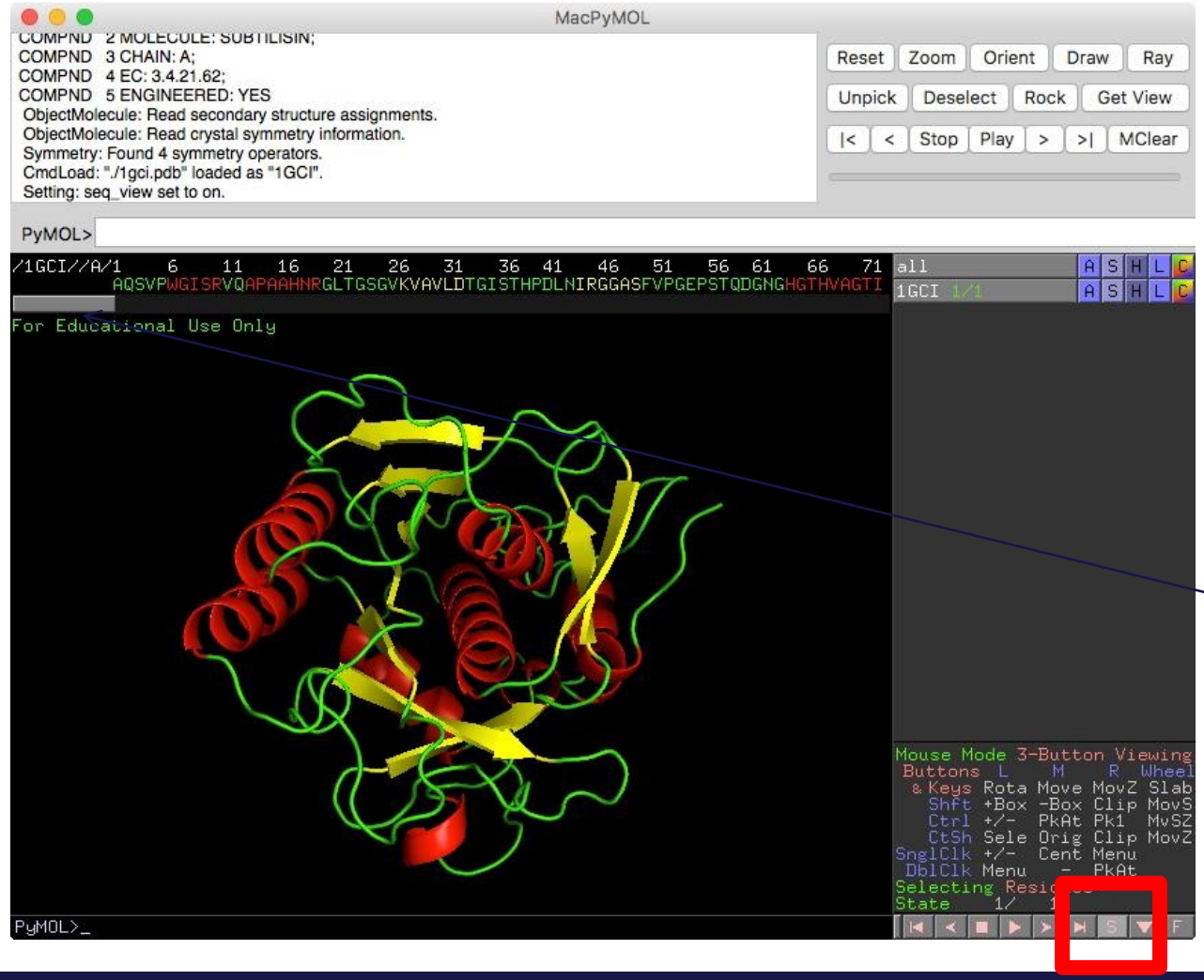


Download another crystal structure of the Savinase and align them.

For example: 6Y5T

- PyMOL>fetch 1GCI
- PyMOL>fetch 6y5t
- PyMOL>remove hetatm
- align 1GCI, 6y5t

Amino acid sequence



View sequence by:

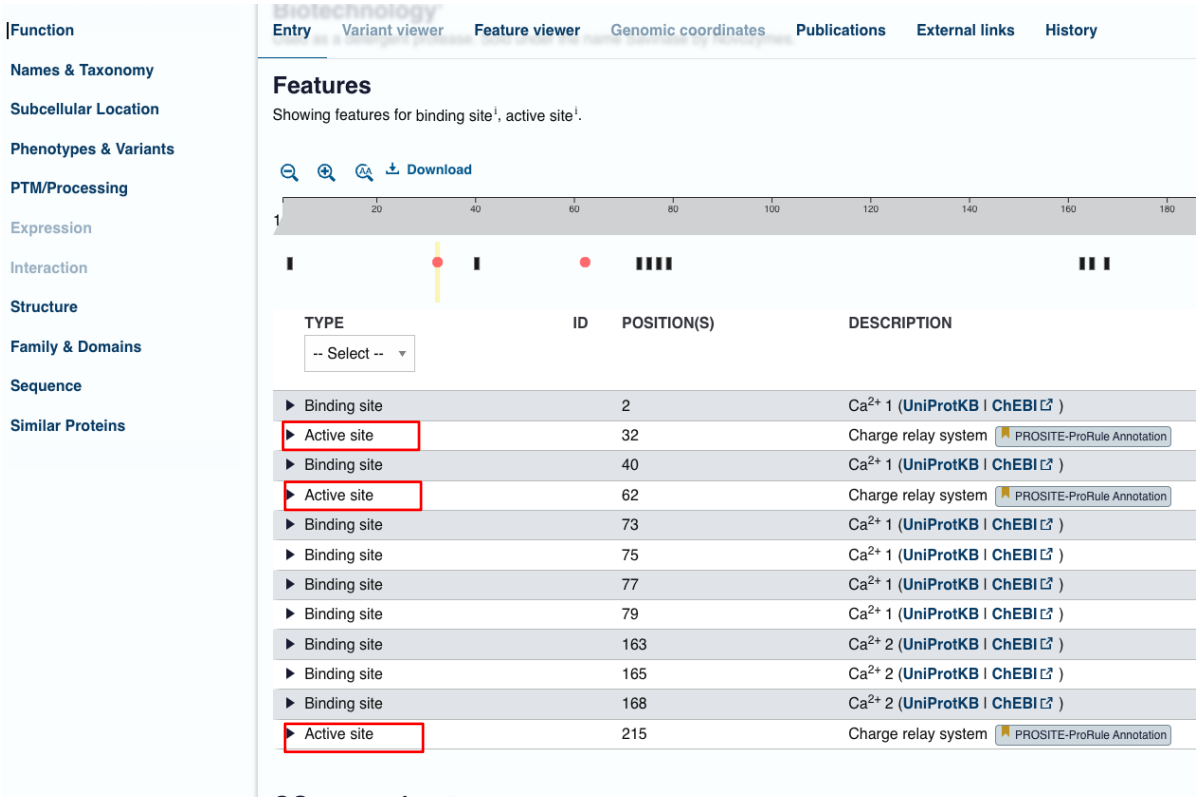
1) Pressing the “S” button in the lower right hand corner

2) Menu: Display -> Sequence On

Scroll by dragging the bar

Select amino acids by clicking them

Active site

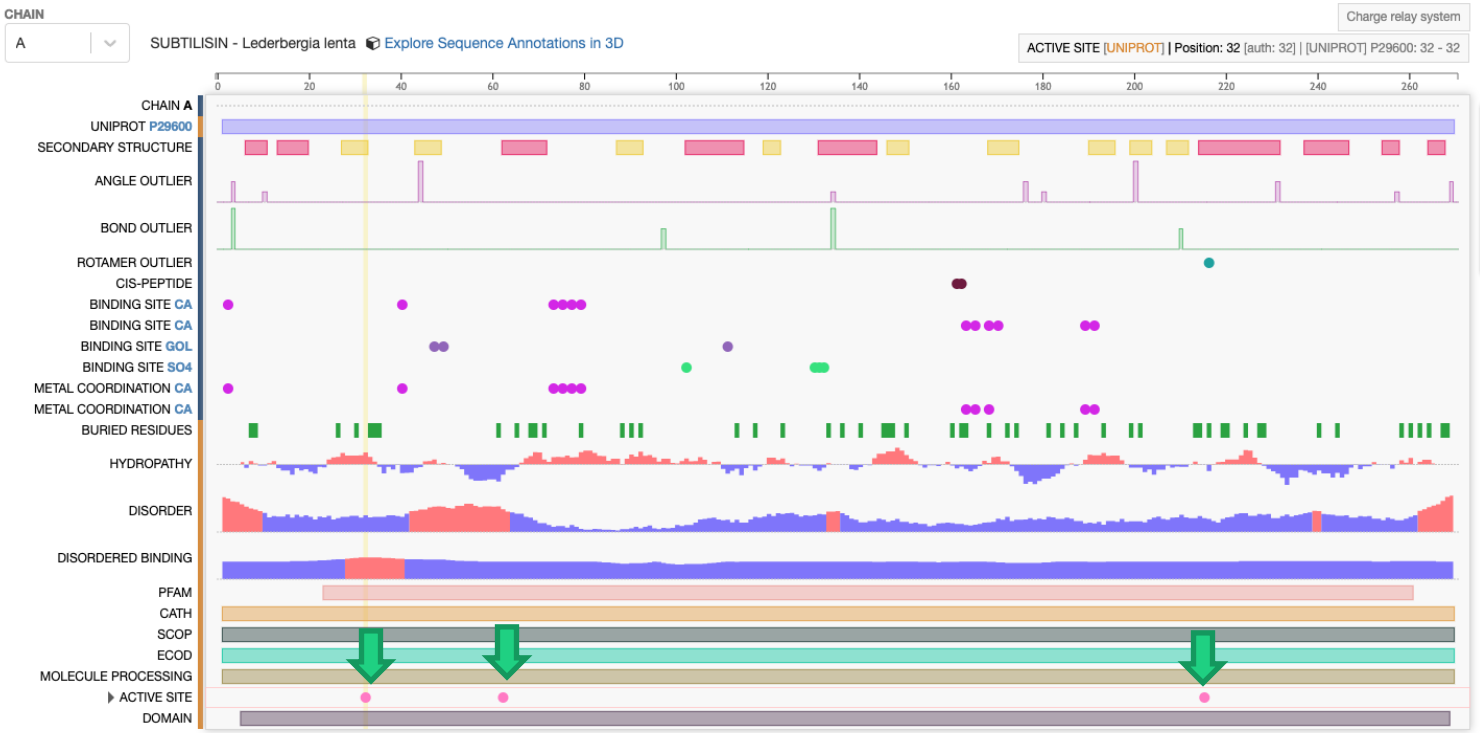


Active site consists of three amino acids

Easy to look up in UniProt

Active site according to UniProt: D32, H62, S215

PDB vs. UniProt numbering



PDB structures may choose to follow a different sequence numbering scheme than UniProt, **even if the sequence is identical!**

This is the case for Savinase, as seen in the figure, and we need to do **coordinate-mapping** of the active site information.

The three amino acids in the active site have been high-lighted.

ACTIVE SITE [UNIPROT] | Position: 32 [auth: 32] | [UNIPROT] P29600: 32 - 32

ACTIVE SITE [UNIPROT] | Position: 62 [auth: 64] | [UNIPROT] P29600: 62 - 62

ACTIVE SITE [UNIPROT] | Position: 215 [auth: 221] | [UNIPROT] P29600: 215 - 215

TASKS

- Turn on sequence mode, fetch the protein Structure: 1GCI (PDB)

- The ACTIVE SITE of the protein consists of (after coordinate mapping):
  - (D) Asp-32
  - (H) His-64
  - (S) Ser-221

- Play around with the sequence bar and figure out how to select these three amino acids (and only those)

Selection

Amino acid residues you click on (in both the sequence and in the actual structure) ends up in a special object named "(sele)".

As with any other object you can apply styles, colors etc. to this object.

That way it's possible to apply a different visualization to a subset of the structure.

TASKS

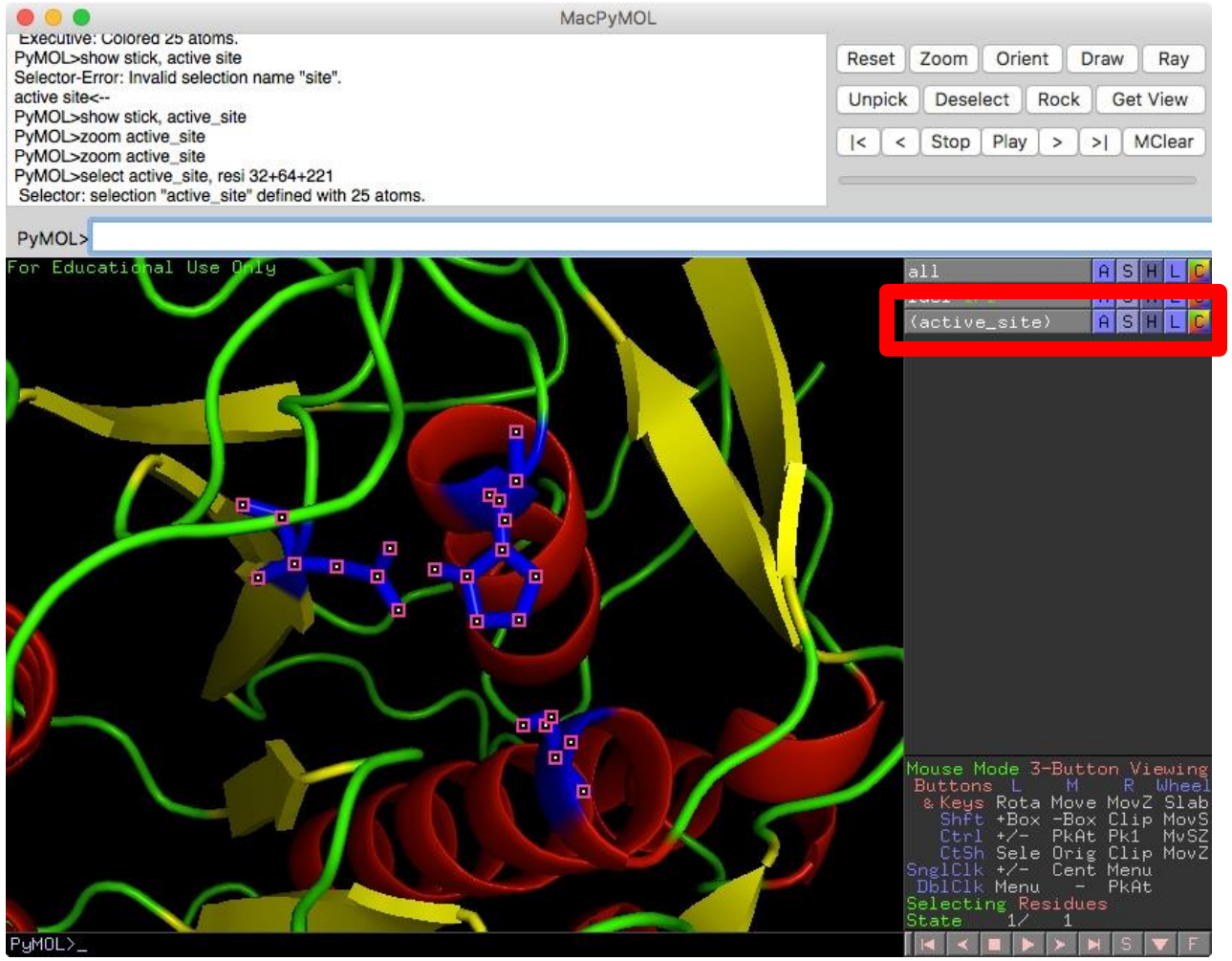
- Work with your selection to show the amino acids in the active site as:
  - "Sticks"
  - With a different color

- Figure out a way to maintain the "Cartoon" style of the backbone and have the sticks show as an additional feature

- If you mess up the visualization you can reset to where we were before by applying the following to the "**1GCI**" object:
  - Show -> As -> Cartoon
  - Color -> By ss -> (pick the first color scheme)

- Finally zoom in a bit so it's easier to see the three amino acids in the catalytic triad
  - Either use the mouse
  - Click "A" -> zoom
  - Or write "**zoom sele**" in the command field

Selection – commands, renaming



Selections can be renamed into something more useful by using the Action (A) button, and thus "saved for later use".

PyMol also makes it possible to specify selection ranges (and name) directly in the command field, as detailed on the next slide

TASK

- Play around with the selection command

- The general syntax for selecting individual amino acids is:
  - select resi 1 (Select only aa #1)
  - select resi 1-5 (Select the range 1-5)
  - select resi 1+5+10 (Select aa#1 and aa#5 and aa#10)

- Select the catalytic triad (D32, H64, S221)

- Rename your selection to something useful for late use:
  - Click “A” -> Rename selection

- You can also specify a name directly in the selection command:
  - select my_name, resi xx+yy+zz

In conclusion

- In this tutorial you have learned how to:
  - Load a structure into PyMol
  - Apply specific styles and colors
  - How to see the amino acid sequence behind the structure
  - How to select specific amino acids ranges in the structure
    - By clicking
    - By using commands
  - How to give those a different visual style + color
  - How to name selections for later use

- PyMol can do a lot of other things, and commands exists for automating the entire process of loading structures, selecting styles, colors, orientation, zooming and exporting the result as images.

- The tutorial has on purpose been kept simple and has only focused on working with amino acid selection – PyMol can do a lot of advanced stuff with atom level selection as well.

- Link to command overview:
  - http://pymol.sourceforge.net/newman/user/S0220commands.html