

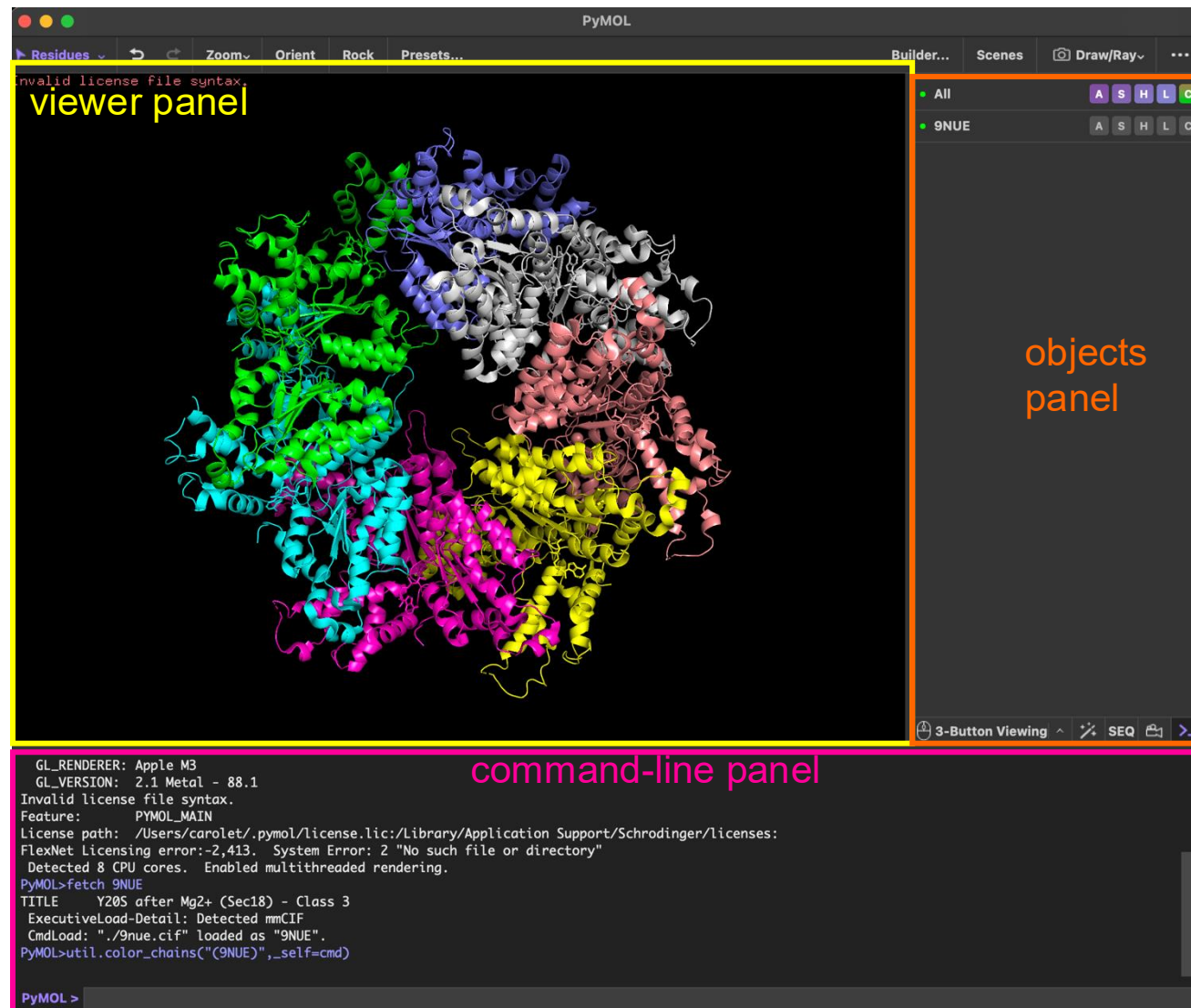
DTU



Introduction to Bioinformatics - 22111

PyMol tutorial

Created by Carolina Barra Quaglia

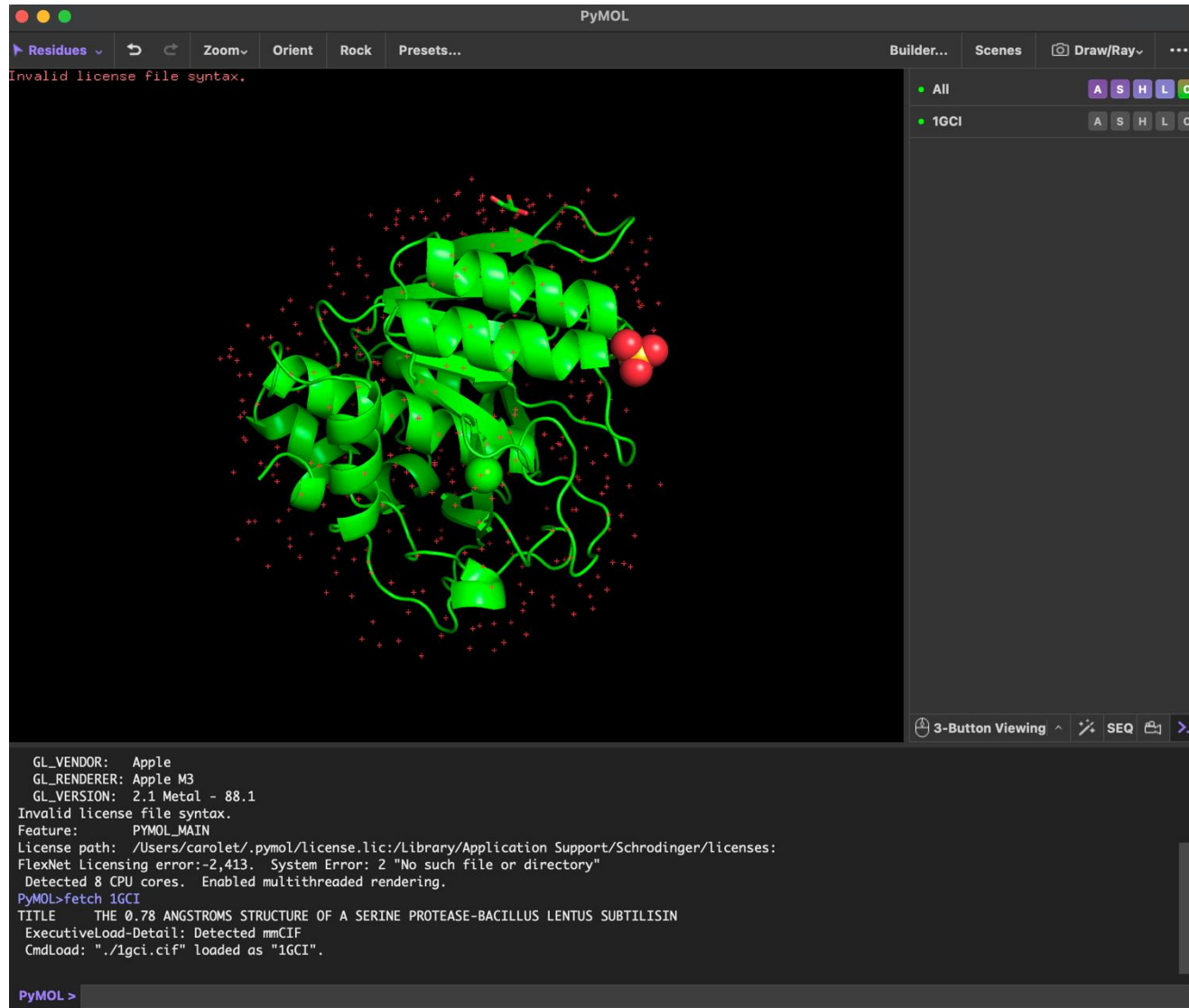


Commands and actions included in this tutorial

- 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)

Note: The word ***residue*** means **amino acid** and is used all along the tutorial. Do not get confused with ***atom***, which means one single chemical **element**.

How to load a protein structure?



Structure: 1GCI
(PDB)

0.78Å structure of
Savinase

Two ways:

- Write **fetch 1GCI** in the command line panel
- Download .PDB file from [pdb.org](https://www.rcsb.org/pdb) and use the File->Open menu

Tasks – learn to configure your mouse!

- Click on the *3-Button editing* on the objects panel, a new panel will open allowing you to configure your mouse

- Play around with the interface – learn how to rotate and zoom the structure:

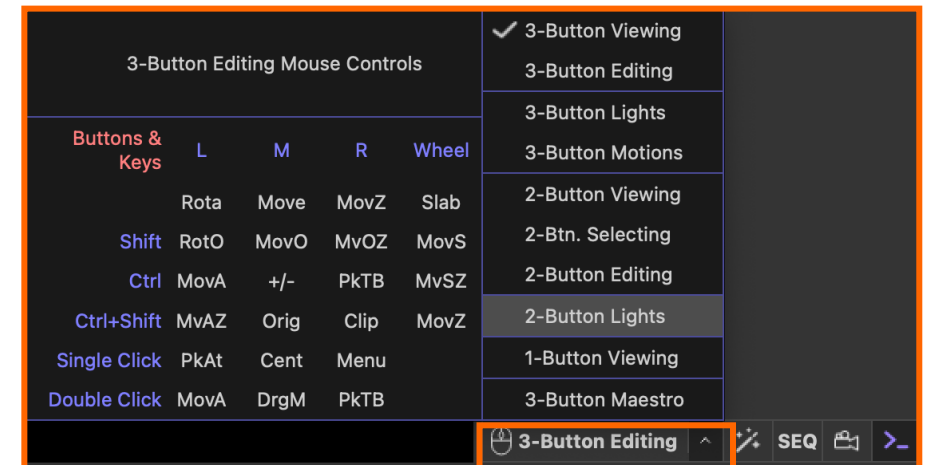
- Rotate: Click and hold left mouse button and move around
- Zoom:
 - 1) Right click + move up/down (a bit slow)
 - 2) Shift + control + scroll wheel
 - 3) Shift + control + two finger drag on mouse pad

- Background info:

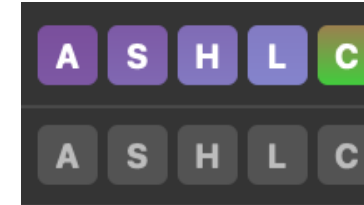
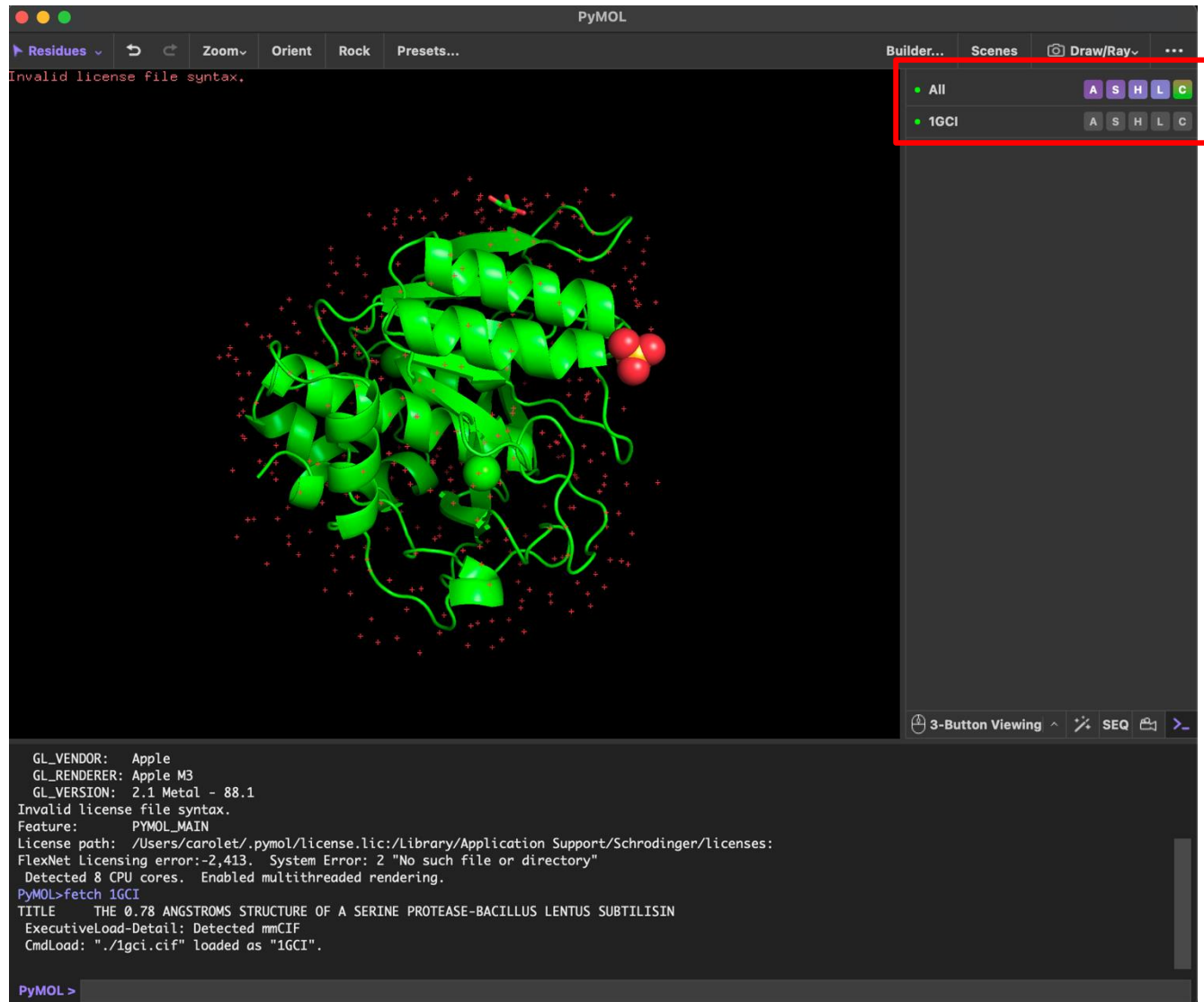
- The structure is of the Novozymes peptidase “Savinase” that we have worked with before

- PDB link: <http://www.rcsb.org/pdb/explore/explore.do?pdbId=1GCI>

- UniProt link: <http://www.uniprot.org/uniprot/P29600>



Working with the structure



A = Action

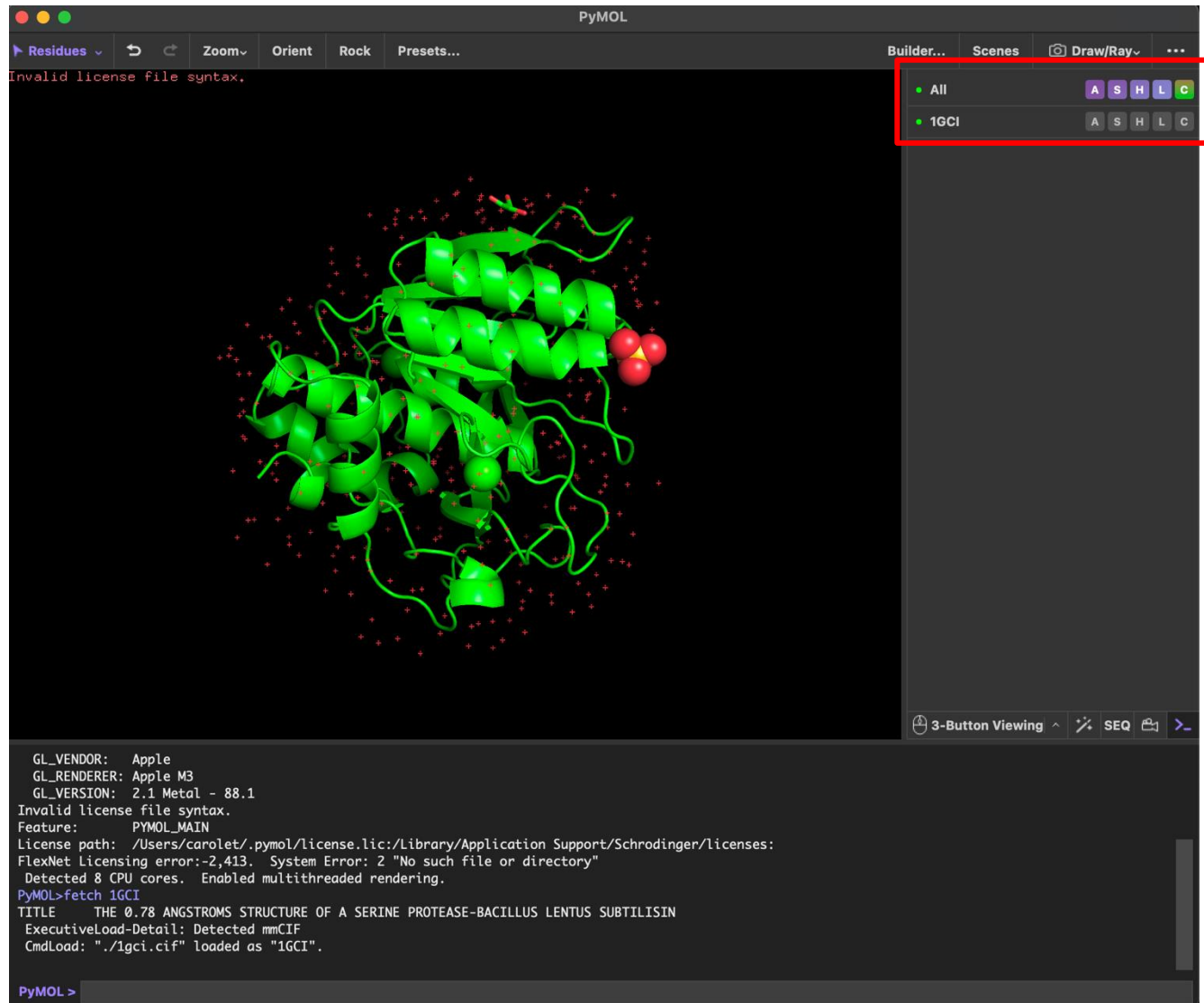
S = Show

H = Hide

L = Label

C = Color

Working with the structure



List of objects to
manipulate

Pressing *All* selects
ALL objects in the
panel

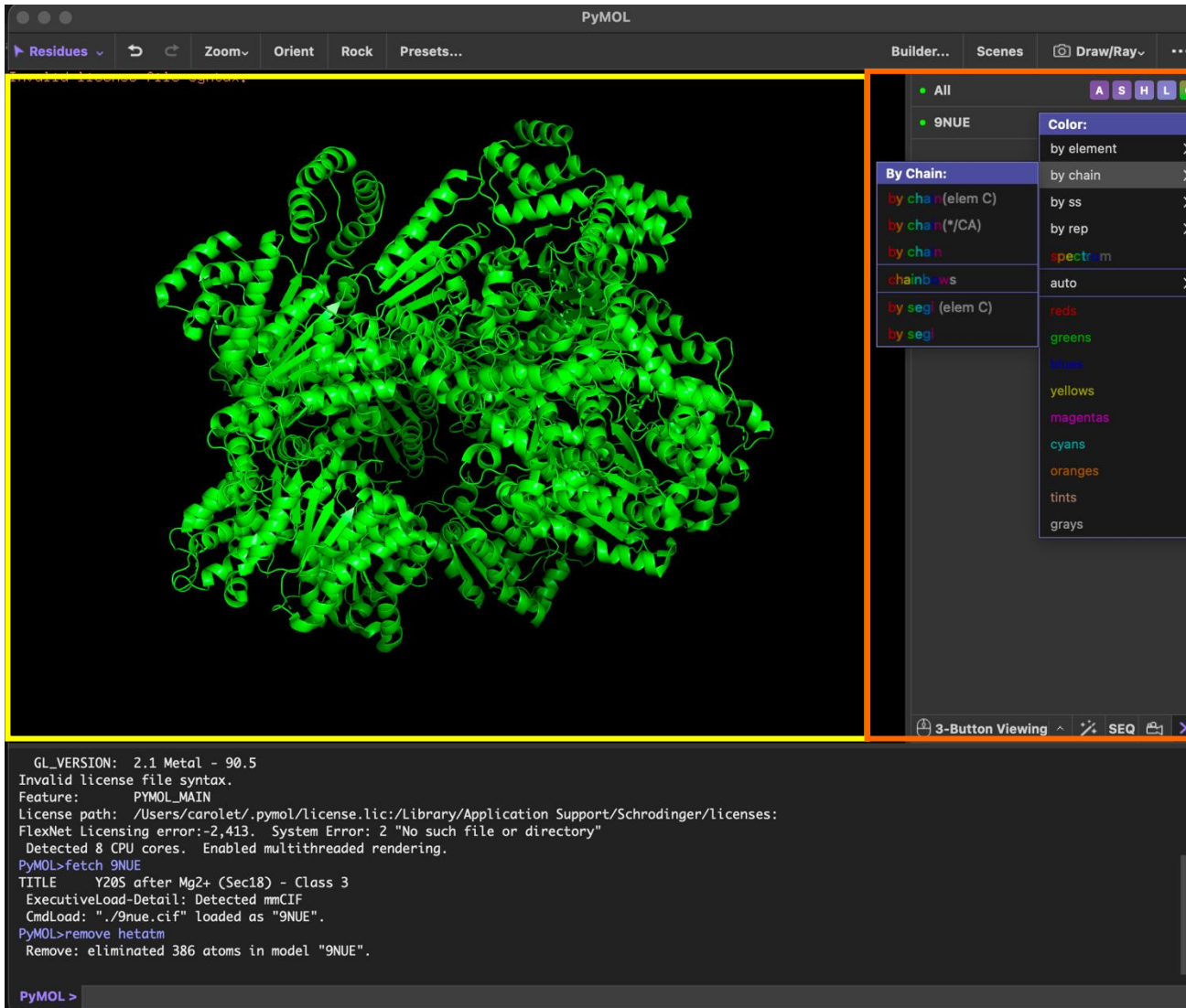
1GCI: the structure
we have loaded

*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 highlight 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



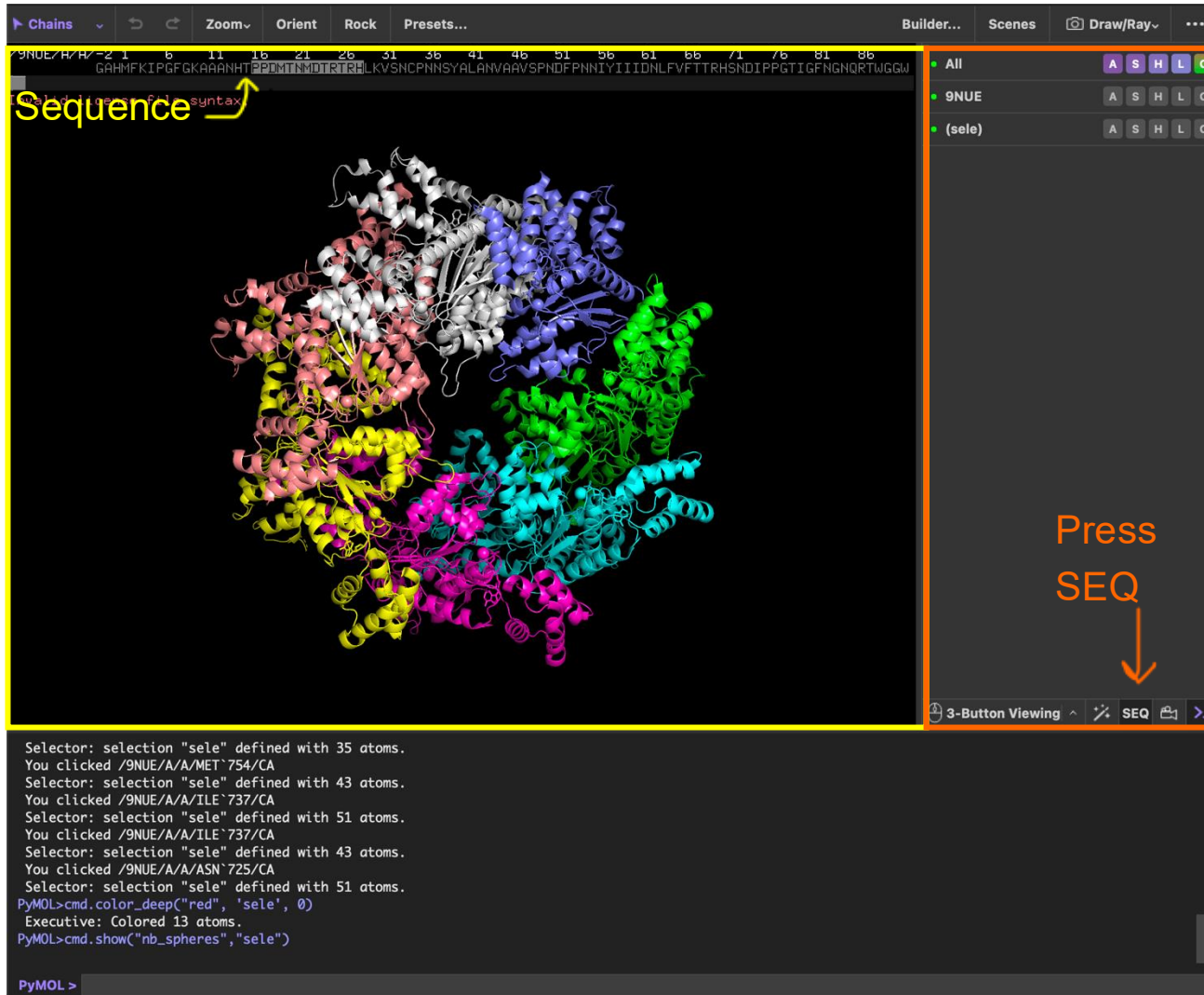
fetch 9NUE

Translocase from
S.cerevisiae

press **C (color)** in 9NUE

- Color by chain->
- Select (elem C)
- How many chains are in the structure?

Show the sequence and color an element



fetch 9NUE

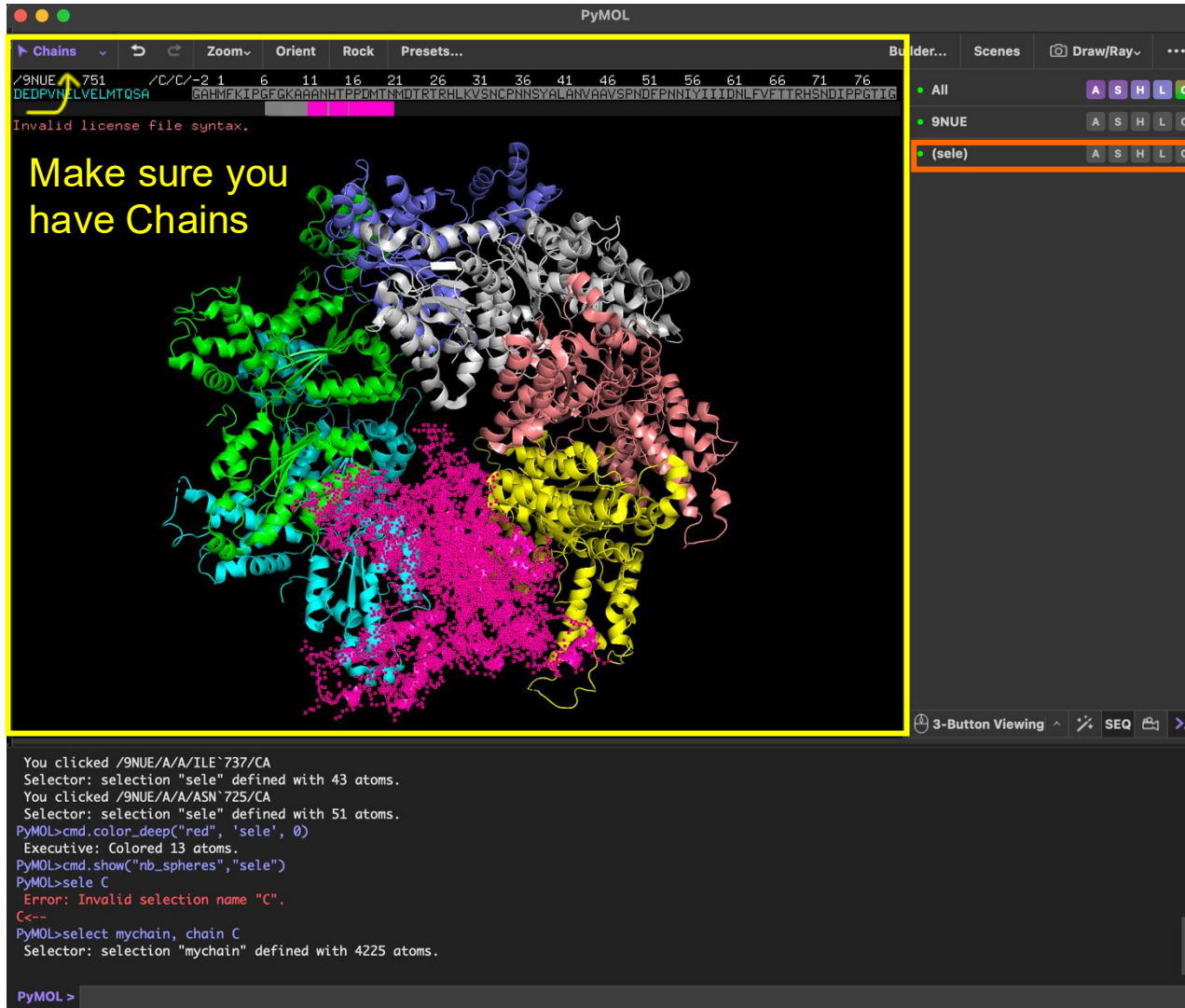
- Press *SEQ* in the object panel and the sequence will appear in the viewer panel
- scroll to the left and check the name of the structure and chains

Chain Name



- How many chains are in the structure?

Show the sequence and color an element



Make sure you have Chains

In the new object called (sele) press C (color)

PyMOL

Chains

/9NUE 751 /C/C/-2 1 6 11 16 21 26 31 36 41 46 51 56 61 66 71 76
 DEDPVNLLVELMTQSA GAHMFKIPGGKRAAHHTPPNTIMDTTRHLKVSNCPPNNSYALANVRAVSPNDFPNNITITDNLVFTTRHSNDIPPGTIG

Invalid license file syntax.

You clicked /9NUE/A/A/ILE`737/CA
 Selector: selection "sele" defined with 43 atoms.
 You clicked /9NUE/A/A/ASN`725/CA
 Selector: selection "sele" defined with 51 atoms.
 PyMOL>cmd.color_deep("red", "sele", 0)
 Executive: Colored 13 atoms.
 PyMOL>cmd.show("nb_spheres", "sele")
 PyMOL>sele C
 Error: Invalid selection name "C".
 C<--
 PyMOL>select mychain, chain C
 Selector: selection "mychain" defined with 4225 atoms.

PyMOL >

In the new object called (sele) press C (color)

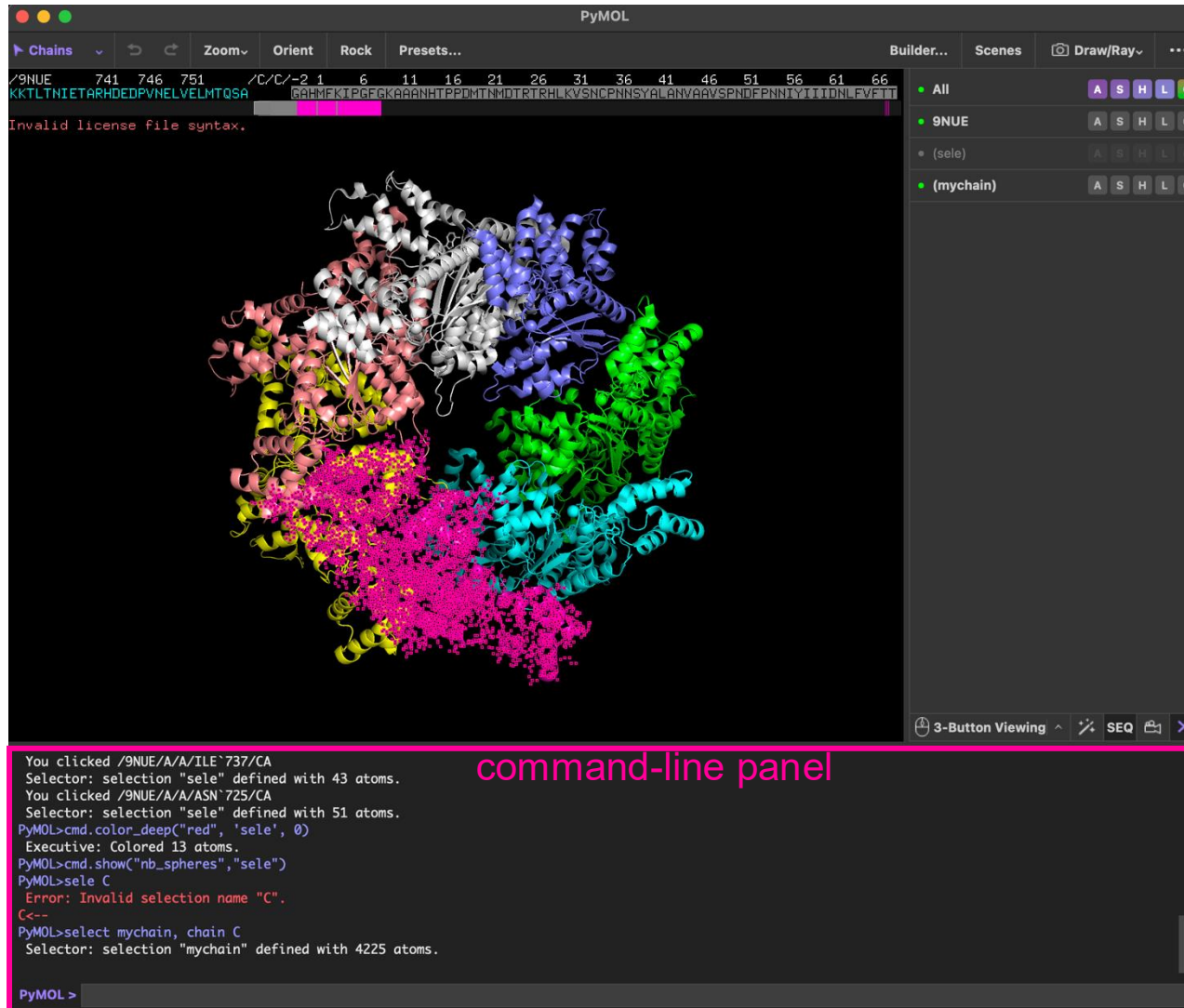
Select chain C and color red

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

If instead of *Chain* you have **residues** you can click and select individual amino acids (both the sequence and in the actual structure) After you click a special object named "(sele)" will appear

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

Show the sequence and color an element



Alternatively in the command line panel..

Type:

OPTION 1

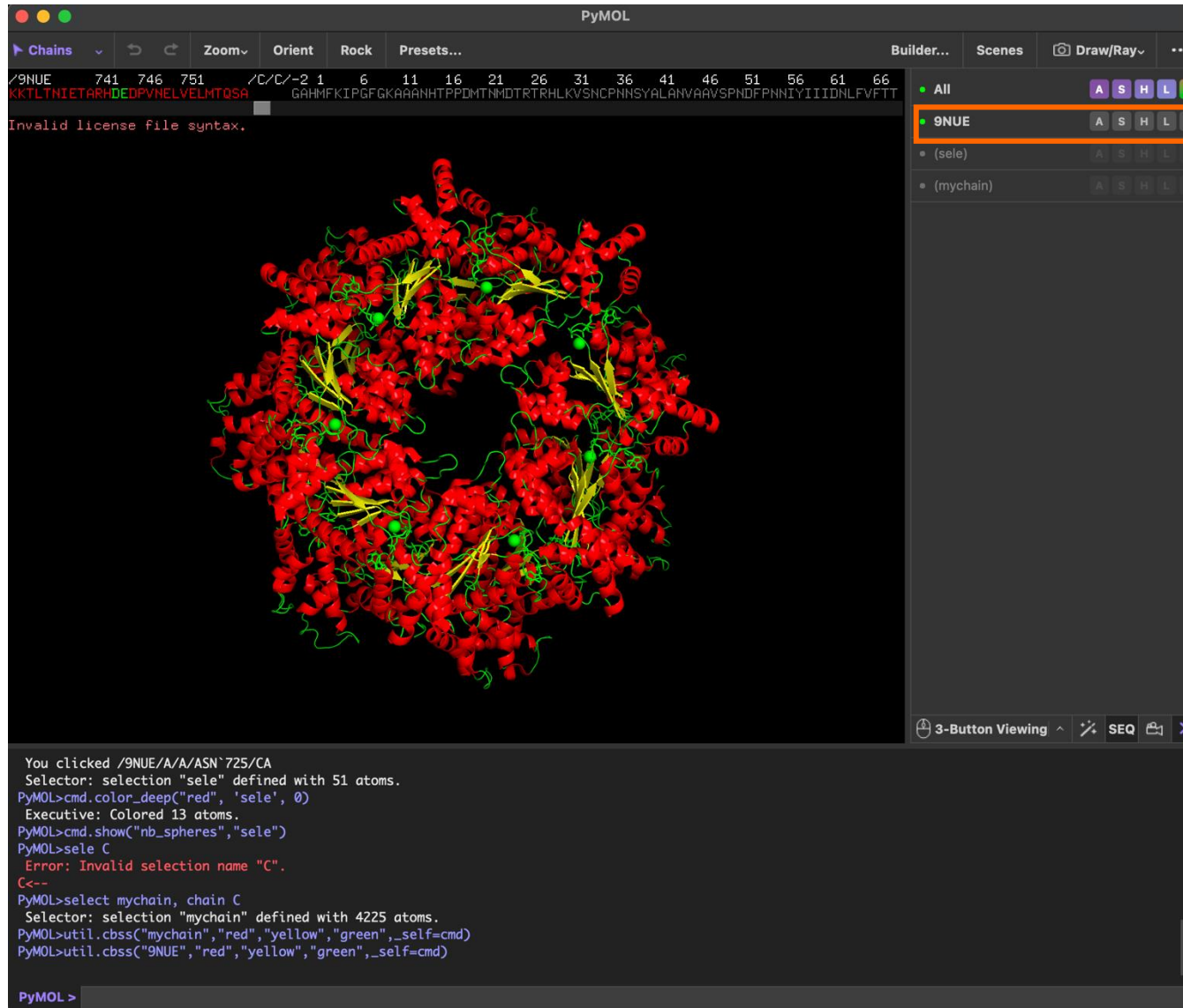
select mychain, chain C
color red, mychain

OPTION 2

color red, chain C

- What are the differences?

Color by secondary structure

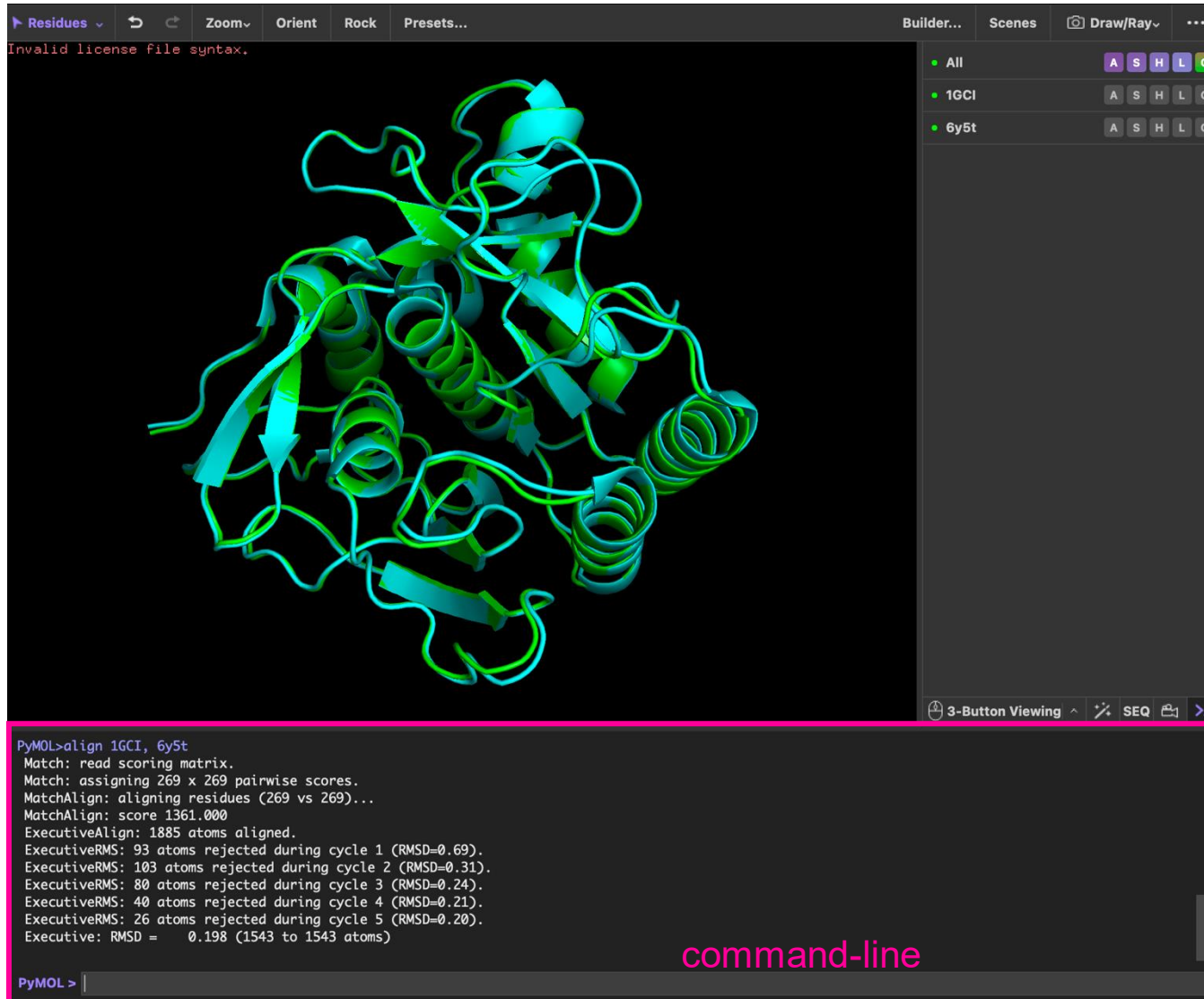


in the structure name press C

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

On the command line type:

```
PyMOL>fetch 1GCI
```

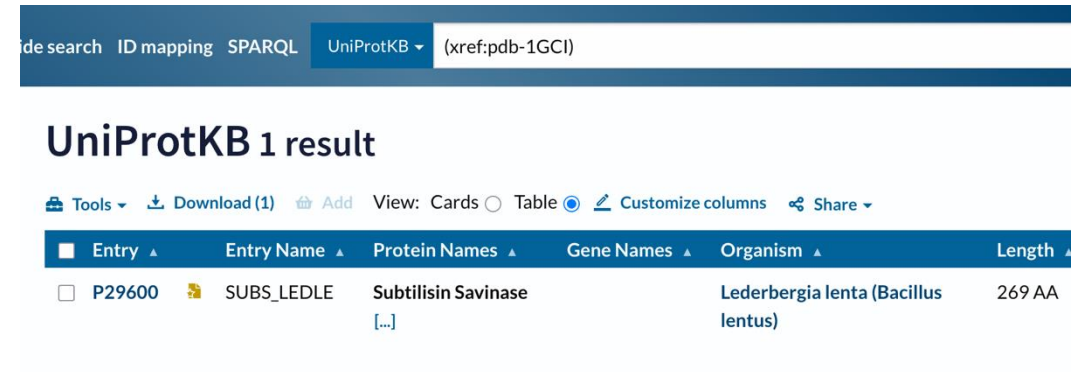
```
PyMOL>fetch 6y5t
```

```
PyMOL>remove hetatm
```

```
PyMOL>align 1GCI, 6y5t
```

TASKS

- Turn on sequence mode, fetch the protein Structure: 1GCI (PDB)
- The ACTIVE SITE of the enzyme consists of (We can get this information from UniProt):
 - (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)
- Are the amino acids of the active site in the same positions in the PDB structure?



UniProtKB 1 result

Tools Download (1) Add View: Cards Table Customize columns Share

Entry	Entry Name	Protein Names	Gene Names	Organism	Length
☐ P29600	SUBS_LEDLE	Subtilisin Savinase [...]		Lederbergia lenta (Bacillus lentus)	269 AA

Active site

UniProt BLAST Align Peptide search ID mapping SPARQL UniProtKB

Function
Names & Taxonomy
Subcellular Location
Phenotypes & Variants
PTM/Processing
Expression
Interaction
Structure
Family & Domains
Sequence
Similar Proteins
Homologs

P29600 · SUBS_LEDLE | Subtilisin Savinase

Tools Download

Entry Variant viewer Feature viewer Genomic coordinates Publications External links History

Features

Showing features for binding site¹, active site¹.

Download

0 20 40 60 80 100 120 140 160 180 200

Q S V P W G I S R V Q A P A A H N R G L T G S G

	TYPE	ID	POSITION(S)	DESCRIPTION
+	Binding site		2	Ca ²⁺ 1 (UniProtKB ChEBI)
+	Active site		32	Charge relay system PROSITE-ProRule Annotation
+	Binding site		40	Ca ²⁺ 1 (UniProtKB ChEBI)
+	Active site		62	Charge relay system PROSITE-ProRule Annotation
+	Binding site		73	Ca ²⁺ 1 (UniProtKB ChEBI)
+	Binding site		75	Ca ²⁺ 1 (UniProtKB ChEBI)
+	Binding site		77	Ca ²⁺ 1 (UniProtKB ChEBI)
+	Binding site		79	Ca ²⁺ 1 (UniProtKB ChEBI)
+	Binding site		163	Ca ²⁺ 2 (UniProtKB ChEBI)
+	Binding site		165	Ca ²⁺ 2 (UniProtKB ChEBI)
+	Binding site		168	Ca ²⁺ 2 (UniProtKB ChEBI)
+	Active site		215	Charge relay system PROSITE-ProRule Annotation

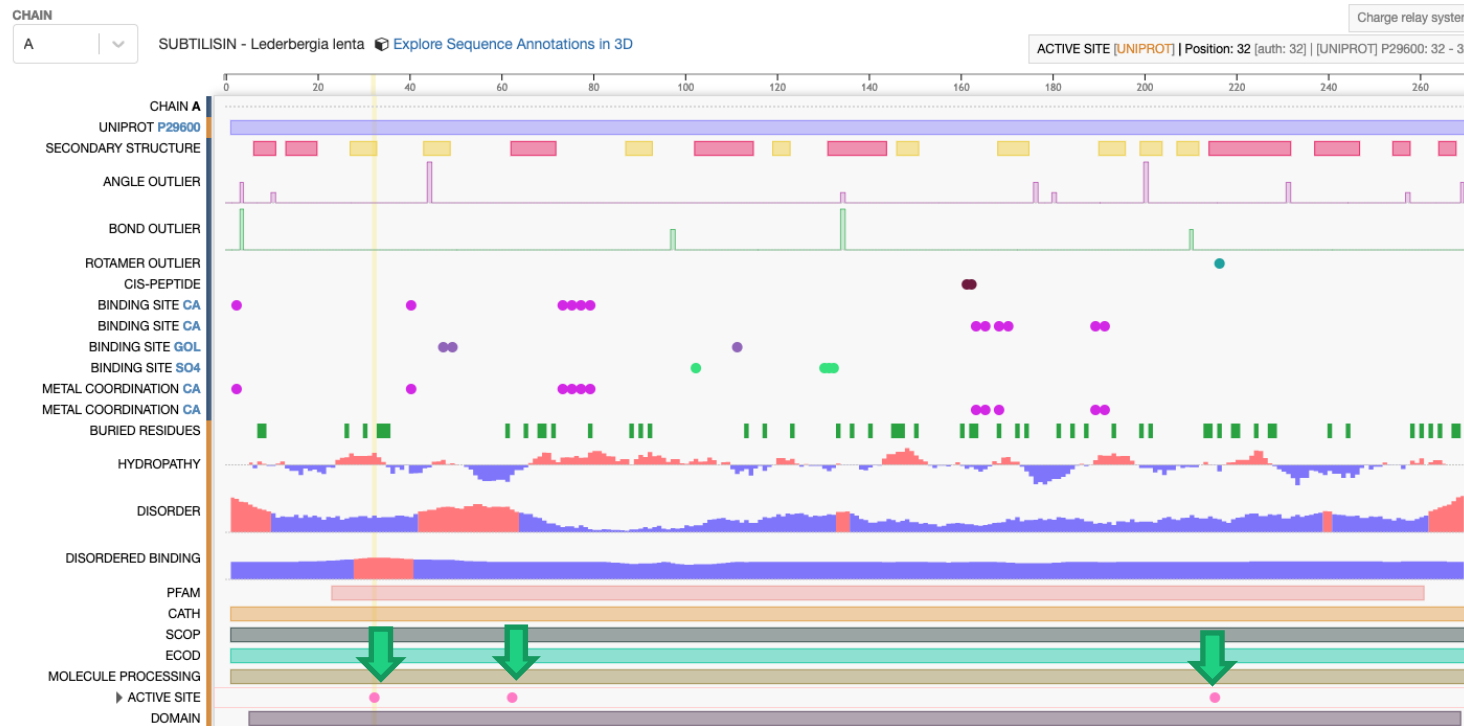
Active site consists of three amino acids

Easy to look up in UniProt

(Function -> Features)

Active site according to UniProt: D32, H62, S215

PDB vs. UniProt numbering



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

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

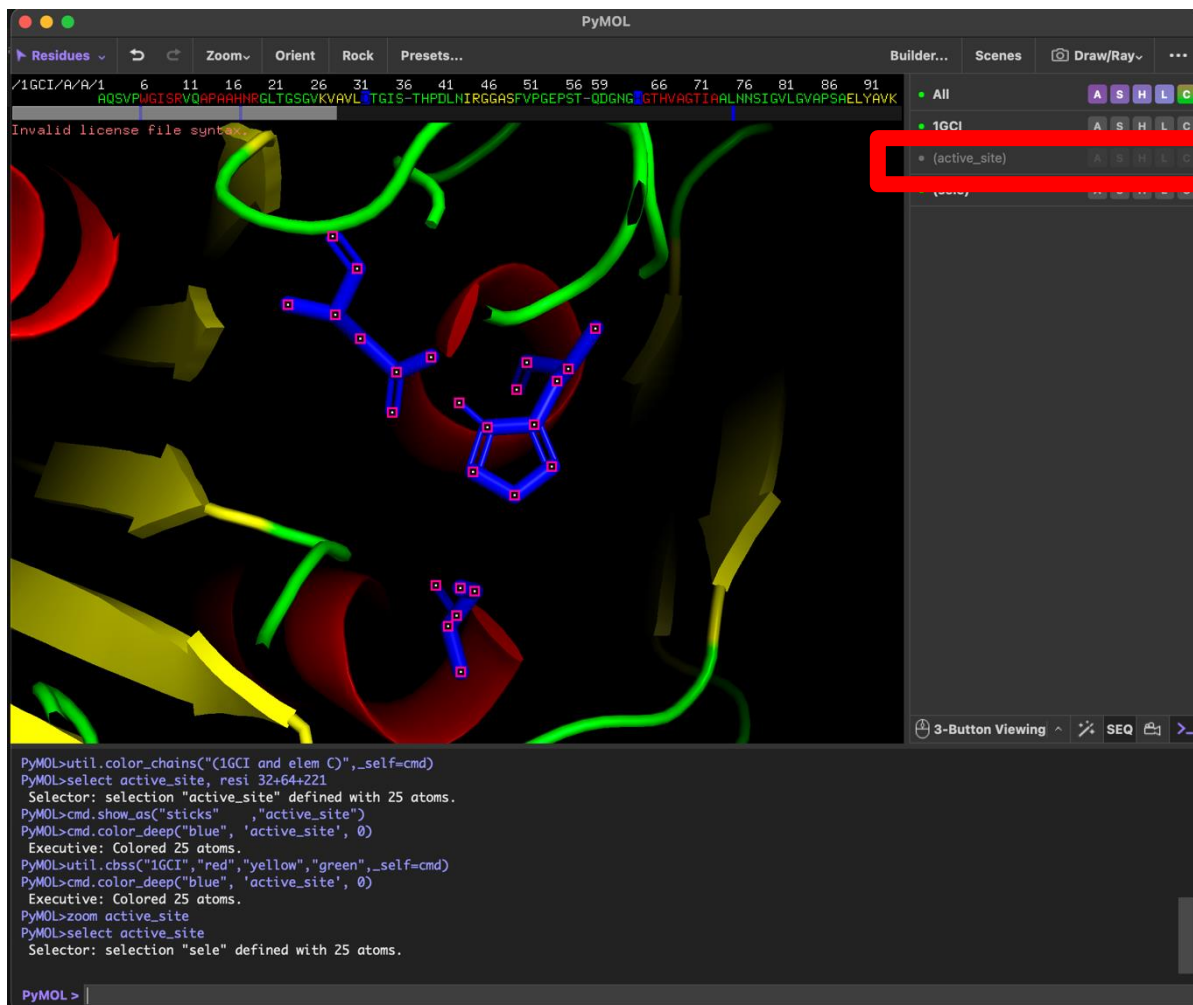
This is the case for Savinase, as seen in the figure, and we need to do **coordinate-mapping** of the active site information.

The three amino acids in the active site have been highlighted.

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"

You can also do that directly in the command line:

select active_site, resi 32+64+221

you can also change color, display and zoom

color blue, active_site

show sticks, active_site

zoom active_site

In summary

- In this tutorial you have learned how to:
 - Load structures into PyMol
 - Apply specific styles and colors both in the command line and in the interactive windows
 - Display the amino acid sequence behind the structure
 - 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 or cool videos for your presentations
- 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>