

# PREPARAZIONE DI IMMAGINI DI STRUTTURE PROTEICHE UTILIZZANDO IL SOFTWARE PYMOL.

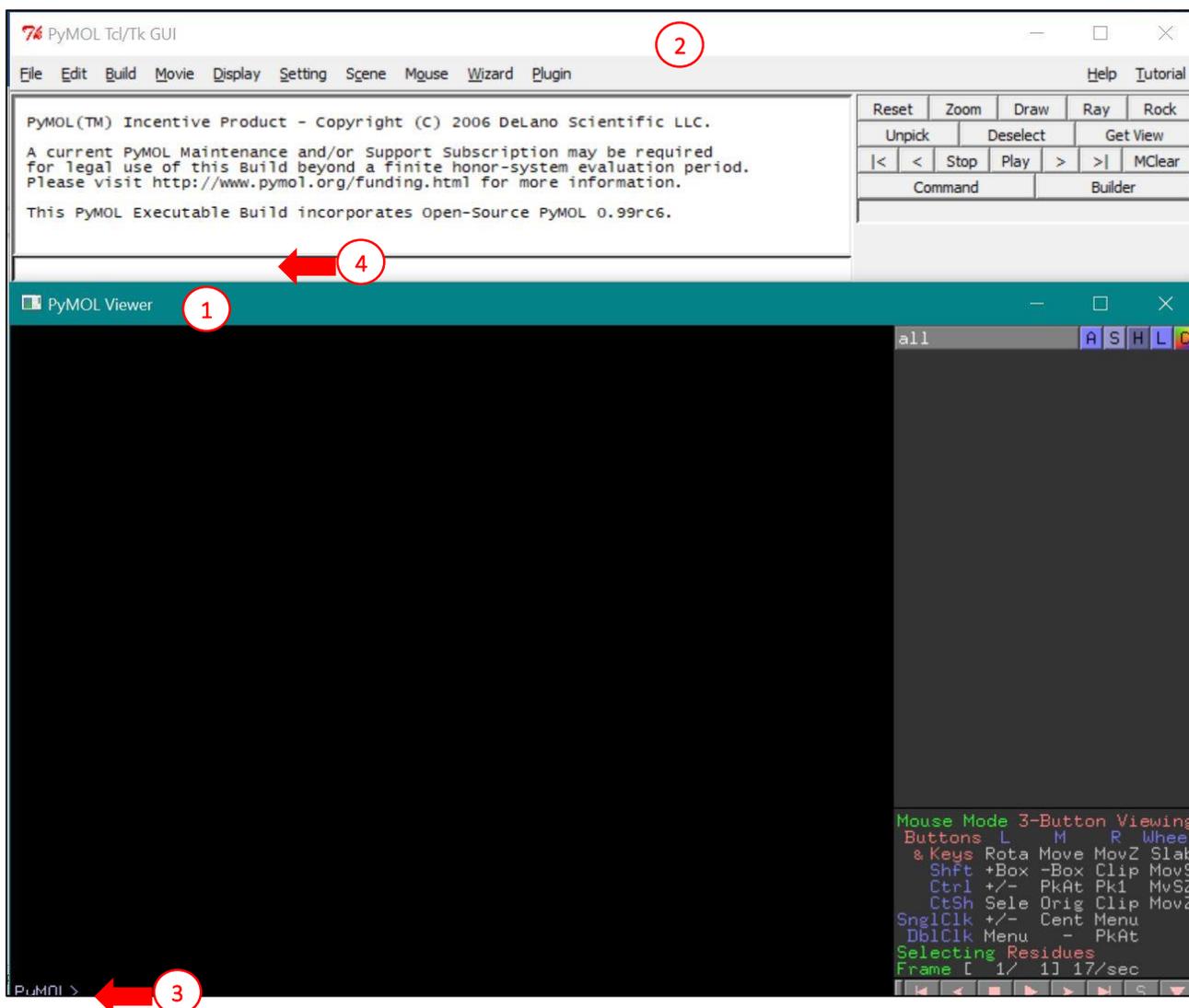
Il software PyMOL [1] è uno strumento che consente di visualizzare strutture proteiche, di confrontarne le conformazioni e di ottenere immagini di buona qualità sia della struttura complessiva sia dei dettagli locali. Come input il software richiede un file di tipo pdb, che può essere ottenuto dall'affinamento portato a termine precedentemente, oppure dalla banca dati Protein Data Bank [2]. Esiste una guida all'utilizzo di questo software, consultabile facilmente come wiki [3], che contiene anche tutorial ed esempi.

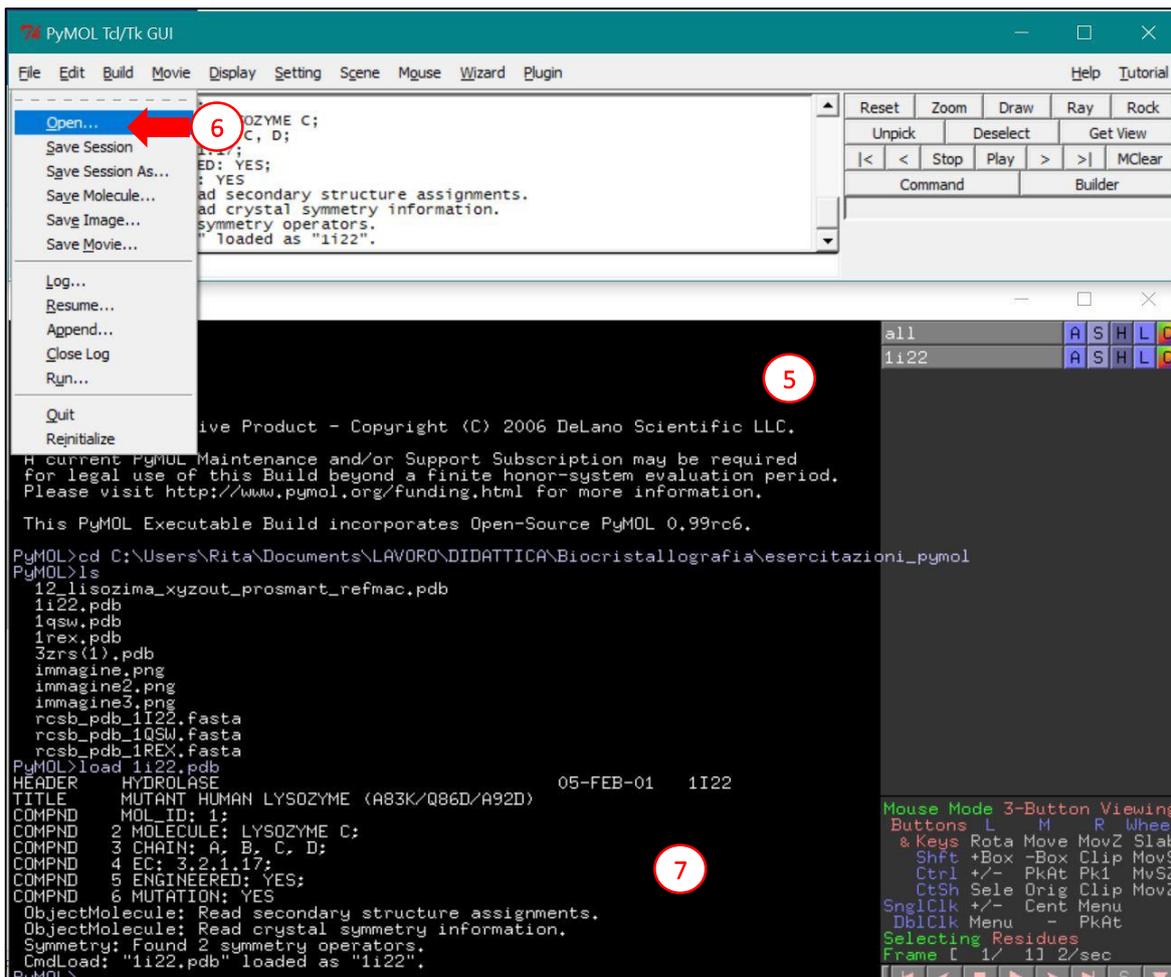
## Visualizzazione di un file pdb

Quando il programma viene avviato, si aprono 2 diverse finestre, una finestra grafica (1) dove l'utente visualizza la struttura ed una finestra di comando (2). Per lavorare con PyMOL è possibile sia scegliere le diverse opzioni dai menù della finestra di comando, sia digitare comandi nella finestra grafica (3) oppure nella finestra di comando (4). Premendo il tasto Esc sulla tastiera mentre si lavora nella finestra grafica è possibile visualizzare la sequenza dei comandi (5).

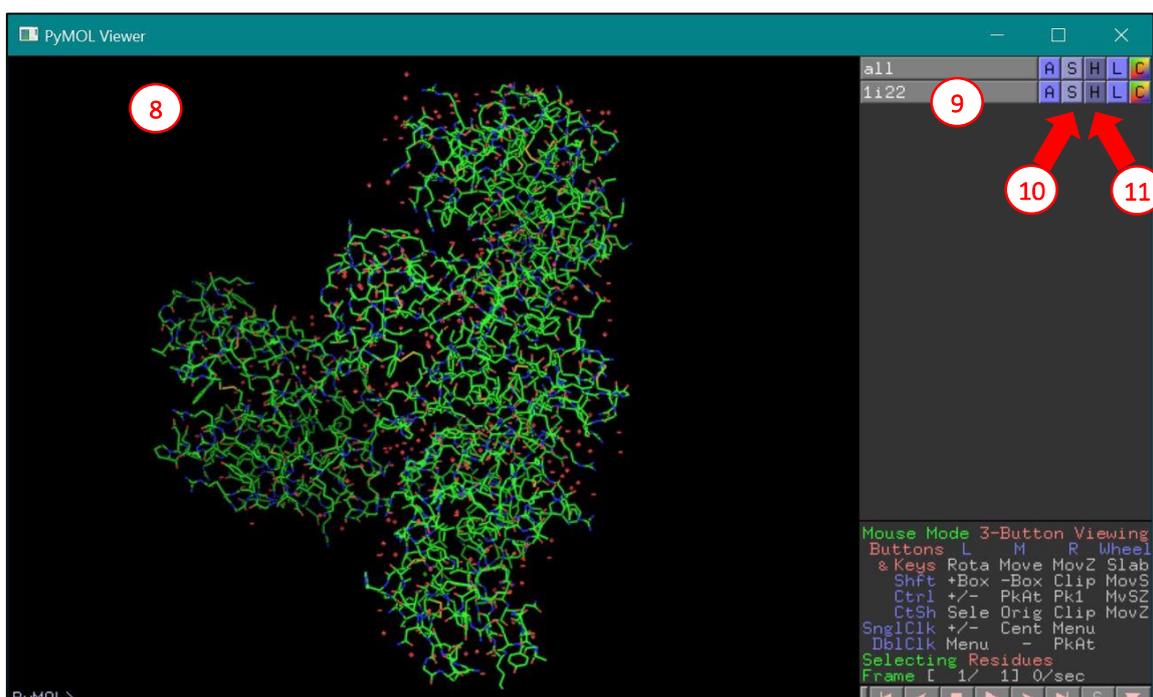
Per aprire un file pdb contenente le coordinate di una struttura proteica, si può selezionare dal menù in alto nella finestra di comando *File > Open* (6), oppure si può caricare il file attraverso la riga di comando:

```
> cd C:\<working directory>  
> load <nomefile>.pdb, <nome dell'oggetto creato (opzionale)>
```





Quando la struttura viene caricata, il software legge automaticamente anche la simmetria del cristallo (7), presente nella riga CRYST1 del file pdb, assieme ai parametri della cella elementare. La molecola caricata è rappresentata nella forma di linee (8). E' possibile caricare più di una struttura alla volta, ma visualizzarne solo una e deselectionare le altre premendo sul nome della struttura nel menù a destra nella finestra grafica (9).



## Opzioni di visualizzazione della struttura proteica

Il software permette diverse opzioni di visualizzazione, tra cui:

- *lines*: i legami sono rappresentati come semplici linee
- *sticks*: i legami sono rappresentati come bastoncini
- *cartoon*: la struttura secondaria è rappresentata come eliche o frecce (per i  $\beta$ -sheet), senza le catene laterali dei residui
- *ribbon*: solo la struttura di carboni  $\alpha$  viene rappresentata
- *spheres*: le sfere di van der Waals di ciascun atomo sono rappresentate.

E' possibile selezionare la diversa visualizzazione dal menù sulla destra della finestra grafica, sotto la lettera **S** (show, **10**), oppure digitare il comando:

```
> show <lines/sticks/cartoon/ribbon/spheres>, <nome della molecola>
```

I metodi di visualizzazione della molecola che non sono desiderati, invece, possono essere nascosti dal menù a destra, sotto la lettera **H** (hide, **11**), oppure digitando:

```
> hide <lines/sticks/cartoon/ribbon/spheres>, <nome della molecola>
```

## Comandi e sintassi di selezione

Al posto del nome della molecola, è possibile anche digitare una specifica selezione di atomi, utilizzando i seguenti identificatori:

- *object*: identifica lo specifico oggetto (l'identificatore può essere anche omissso, digitando direttamente il nome della struttura che compare sulla destra)
- *chain*: lettera identificativa della catena presente nell'oggetto indicato
- *resi*: selezione di specifici residui utilizzando i loro numeri (esempio: *resi 4-14* indica l'intervallo tra i residui 4 e 14, inclusi)
- *resn*: identifica il tipo di residui (esempio: *resn trp* seleziona tutti i triptofani presenti nella struttura)
- *name*: identifica il nome di un atomo (esempio: *name CA* seleziona tutti i carboni il cui nome è CA presenti nella struttura)

La selezione utilizza anche gli operatori logici, tra cui in particolare *and* e *or*.

Ad esempio, se si vuole selezionare i residui da 4 a 14 della catena A dell'oggetto 1i22 e anche tutti i residui di prolina presenti nella stessa catena A, il comando che si digita è:

```
> sele object 1i22 and chain A and (resi 4-14 or resn pro)
```

Se invece siamo interessati a selezionare i residui da 4 a 14 della catena A dell'oggetto 1i22, ma tutti i residui di prolina presenti nell'oggetto, il comando diventa:

```
> sele object 1i22 and ((chain A and resi 4-14) or resn pro)
```

Questi residui possono essere rappresentati come *stick* utilizzando il comando:

```
> show sticks, sele
```

Oppure, direttamente, senza digitare il comando di selezione

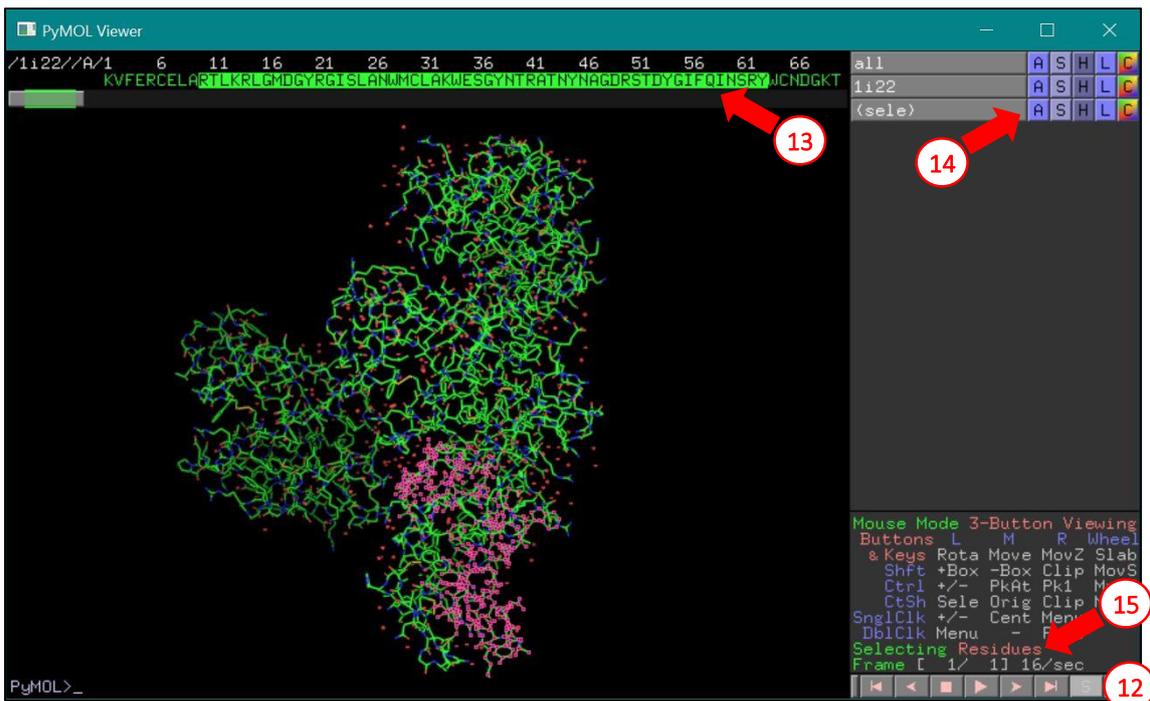
```
> show sticks, object 1i22 and ((chain A and resi 4-14) or resn pro)
```

La stessa selezione può essere effettuata utilizzando il mouse e selezionando un residuo alla volta, oppure visualizzando la sequenza dei residui mediante il tasto **S** in basso a destra (**12**) e selezionando, sempre con il mouse, i residui di interesse nella sequenza che compare in alto (**13**). A questo punto, le opzioni di visualizzazione per questa sequenza possono essere modificate utilizzando i menù **S** e **H** dell'oggetto (*sele*) comparso in alto a destra in seguito alla selezione (**14**).

Oltre a selezionare singoli residui, è anche possibile selezionare interi oggetti, o catene, o singoli atomi. Per modificare il metodo di selezione, è necessario clickare sull'opzione *Selecting* in basso a destra (**15**).

La selezione di atomi può essere anche colorata opportunamente utilizzando il menù in alto a destra alla lettera **C** (color, **16**), oppure con il comando:

```
> color <nome del colore>, <selezione>
```



### Ulteriori opzioni di visualizzazione

Esistono molte ulteriori opzioni per la visualizzazione delle strutture, per le quali si fa riferimento alla guida wiki presente online [3], e alla lista che può essere visualizzata dal menù *Settings > Edit All* della finestra di comando. Le stesse variabili presenti nel menù possono essere cambiate anche dalla riga di comando, utilizzando il comando `set`:

```
> set <variabile>, <valore>
```

Alcune delle opzioni presenti in questo menù:

- *sphere\_scale*: cambia la rappresentazione degli atomi come sfere, riducendone o aumentandone il raggio rispetto al raggio di van der Waals dell'atomo. Il valore iniziale è 1, corrispondente ad un raggio di van der Waals. Accetta come valore un numero.
- *stick\_radius*: cambia la rappresentazione dei legami come *stick* facendone variare lo spessore. Accetta come valore un numero.
- *ortho*: permette di passare dalla visione prospettica (quella iniziale all'apertura del programma) alla visione ortoscopica. Accetta come valore 1/0 (o anche on/off).
- *depth\_cue*: il software rappresenta la profondità utilizzando una sfumatura per le molecole (o le regioni di una molecola) più lontane dall'osservatore. Questa opzione può essere rimossa fissando questa variabile ad un valore di 0. Accetta come valore 1/0 (o anche on/off).
- *ray\_shadow*: nella rappresentazione normale di PyMOL, quando l'immagine viene preparata per essere salvata, il programma disegna le ombre delle catene, rispetto alla/e fonte/i di illuminazione. Accetta come valore 1/0 (o anche on/off).
- *ray\_opaque\_background*: normalmente lo sfondo della figura è riportato nell'immagine con un colore unico (nero di default, ma può essere modificato, si veda sotto). Con questo comando lo sfondo può essere reso trasparente, opzione molto utile per la preparazione di immagini per presentazioni. Accetta come valore 1/0 (o anche on/off).
- *bg\_rgb*: modifica la colorazione dello sfondo della figura. Accetta 3 valori, che devono essere inseriti separati da uno spazio, corrispondenti rispettivamente alla componente rossa, verde e blu del nuovo colore di sfondo. Ad esempio, il colore bianco si ottiene con [1 1 1], il nero è [0 0 0]. Lo stesso risultato può essere ottenuto con il comando:

```
> bg_color <colore>
```

oppure dal menù *Display > Background*.

## Visualizzazione di molecole relazionate per simmetria

Quando una struttura cristallografica viene caricata, il programma rappresenta solo il contenuto del file pdb, corrispondente all'unità asimmetrica. Tuttavia, è possibile (e in molti casi è utile) rappresentare anche le molecole simmetriche. Per farlo, è necessaria l'informazione relativa a cella elementare e simmetria riportata alla riga CRYST1 del file pdb (<oggetto>). Se questa informazione è presente, il programma può rappresentare le molecole simmetriche attorno ad un centro (<selezione>) ed entro un raggio (<raggio>) definiti dall'utente. Il comando per la rappresentazione delle molecole simmetriche è:

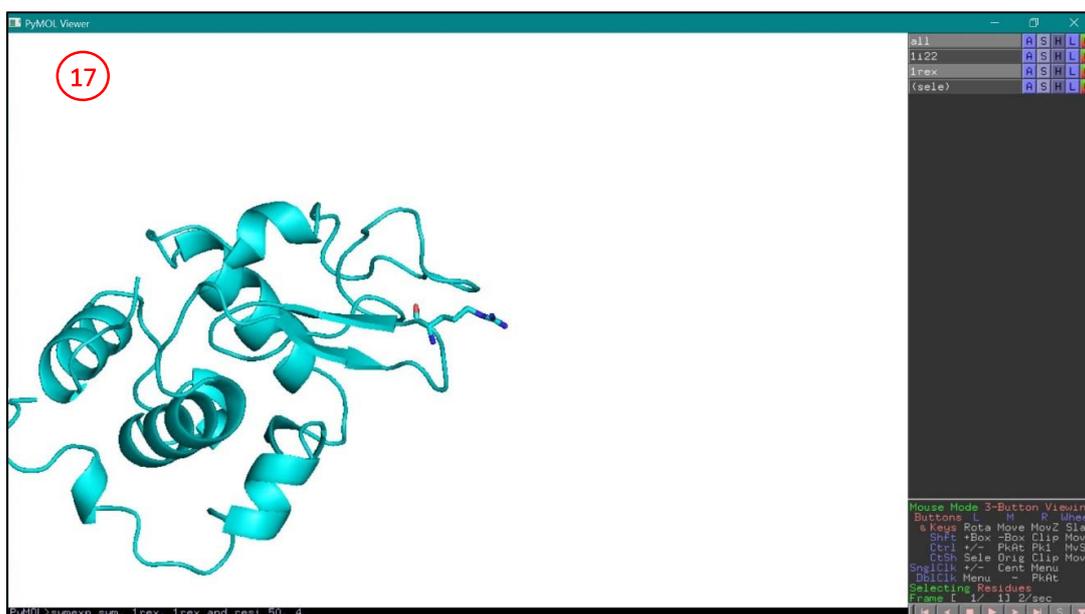
```
> symexp <prefisso>, <oggetto>, <selezione>, <raggio>
```

Il programma crea altri oggetti contenenti le molecole simmetriche e il cui nome contiene il prefisso selezionato.

Nell'esempio qui sotto (16-18), la molecola viene spostata in modo che il centro dell'immagine si trovi sul residuo 50 e il residuo viene rappresentato come stick:

```
> center lrex and resi 50
```

```
> show stick, lrex and resi 50
```



Il comando successivo serve a creare gli oggetti simmetrici (in questo caso solo uno, 19) che verranno chiamati con il prefisso indicato:

```
> symexp sym, 1rex, 1rex and resi 50, 4
```



Questo comando di PyMOL è molto utile per analizzare i contatti proteina-proteina nel cristallo e valutare come questi possono influenzare la conformazione delle catene laterali. Inoltre, nel caso in cui la proteina in esame sia oligomerica e si trovi su un elemento di simmetria cristallografica, sarà necessario mostrare anche le catene proteiche relazionate per simmetria per osservare la proteina intera e funzionale.

### Misura di distanze, angoli e angoli torsionali

Con PyMOL è anche possibile misurare e rappresentare distanze, angoli e angoli torsionali. Per farlo, è necessario selezionare gli atomi che definiscono il legame (2 atomi), l'angolo (3 atomi) o l'angolo torsionale (4 atomi). Gli atomi devono essere selezionati con un metodo di selezione del mouse diverso da quello usuale. Per cambiare metodo di selezione, fare click in basso a destra, sull'indicazione *Mouse mode* (20) che verrà cambiato da *Editing* a *Viewing*. A questo punto gli atomi vengono selezionati (21). A seconda del parametro geometrico da misurare, si seleziona l'adeguato numero di atomi. Infine, per la misura si utilizzando i comandi *distance*, *angle* e *dihedral*. Ad esempio, per misurare una distanza (22):

```
> distance d1, pk1, pk2
> color black, d1
```

Il programma ha creato un nuovo oggetto, la distanza stessa, le cui proprietà possono essere modificate anche dal menù sulla destra. In questo caso il nome che viene dato all'oggetto distanza è *d1*. Le variabili successive, *pk1* e *pk2*, sono gli atomi selezionati (presenti anche nell'elenco sulla destra, 23). (Il secondo comando, opzionale, permette una miglior visualizzazione della distanza che inizialmente appare colorata in giallo).

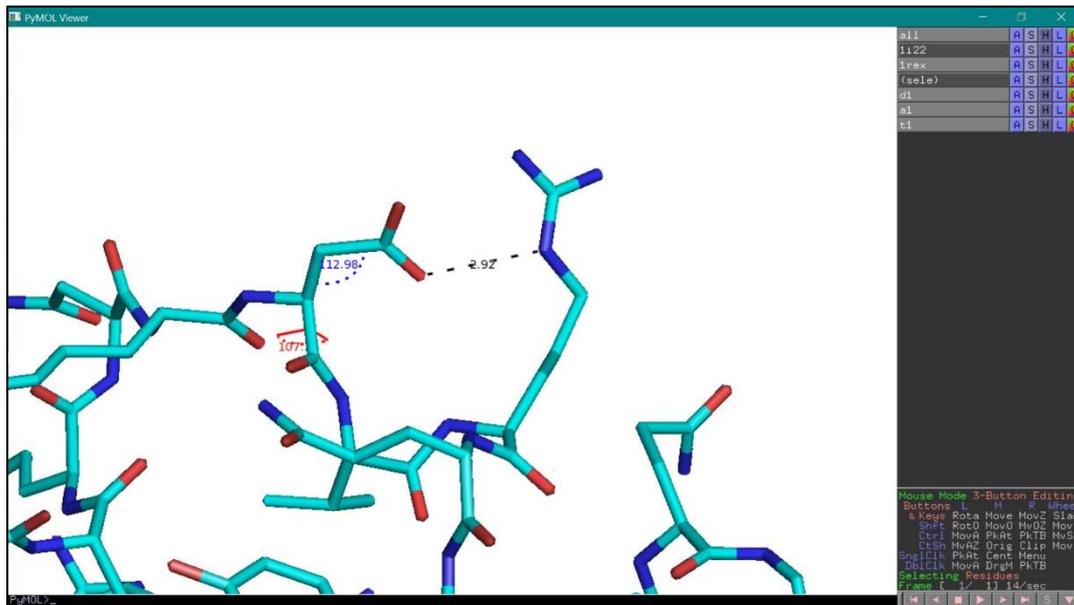
Nel caso in cui si voglia misurare un angolo, si selezioneranno 3 atomi (24) e si utilizzeranno i comandi:

```
> angle a1, pk1, pk2, pk3
> color blue, a1
```

La misura di un angolo diedro può essere fatta selezionando 4 atomi (25):

```
> dihedral t1, pk1, pk2, pk3, pk4
> color red, t1
```





### Sovrapposizione di proteine omologhe per confrontarne la struttura

Quando si confrontano strutture diverse di proteine della stessa famiglia può essere molto utile visualizzarle in modo che siano allineate, per verificare le zone uguali e diverse nelle strutture omologhe. Per farlo, PyMOL ha un comando specifico, il comando `align`, che sposta una molecola (*selezione 1*) per farla coincidere con l'altra (*selezione 2*), in modo da minimizzare la distanza tra gli atomi corrispondenti. La sintassi del comando è la seguente:

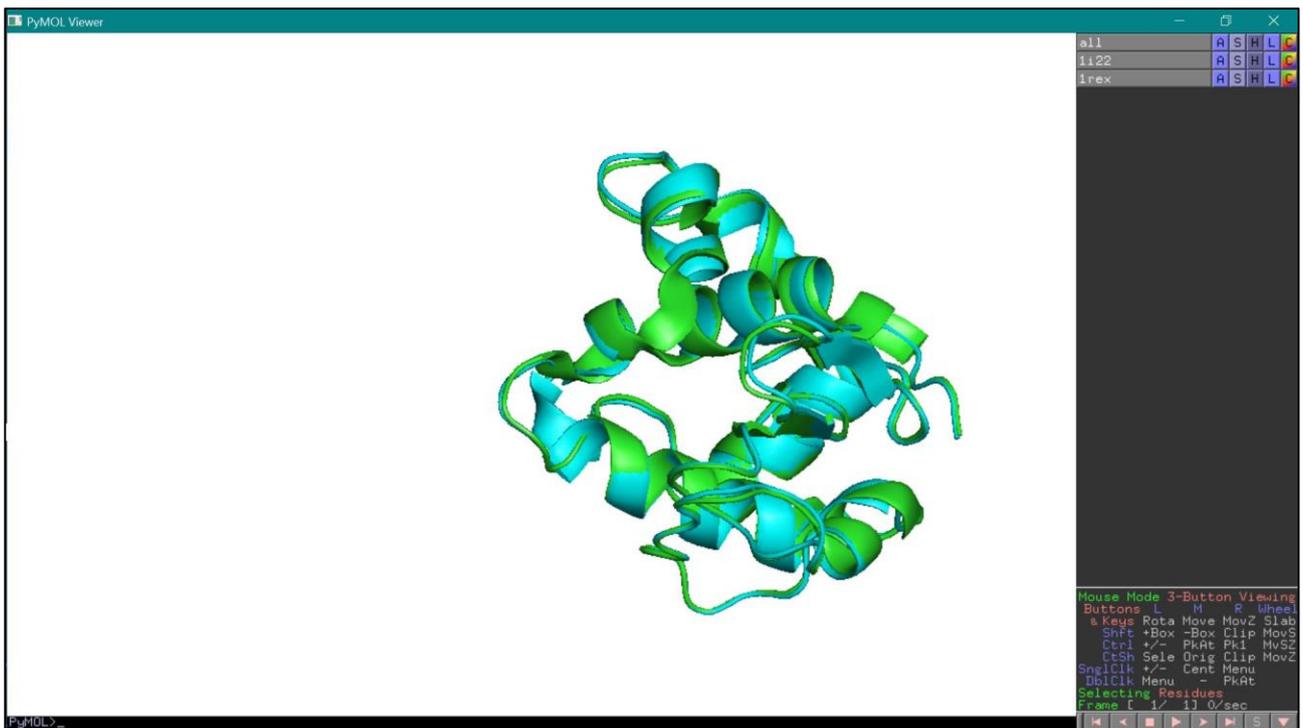
```
> align <selezione 1>, <selezione 2>
```

Perché l'allineamento abbia successo, le due selezioni devono appartenere a oggetti diversi.

Ad esempio, per allineare le strutture della lisozima corrispondenti ai codici PDB 1i22 (struttura in verde) e 1rex (struttura in ciano):

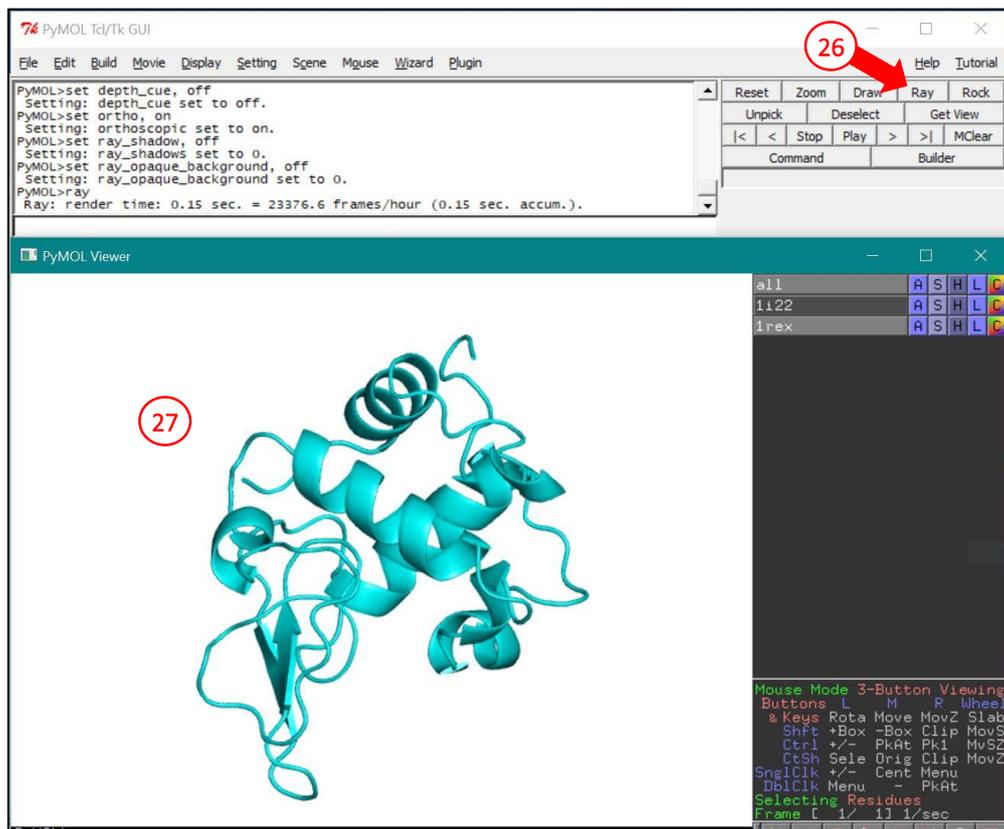


```
> align 1rex and chain A, 1i22 and chain A
```



### Preparazione di immagini della struttura proteica

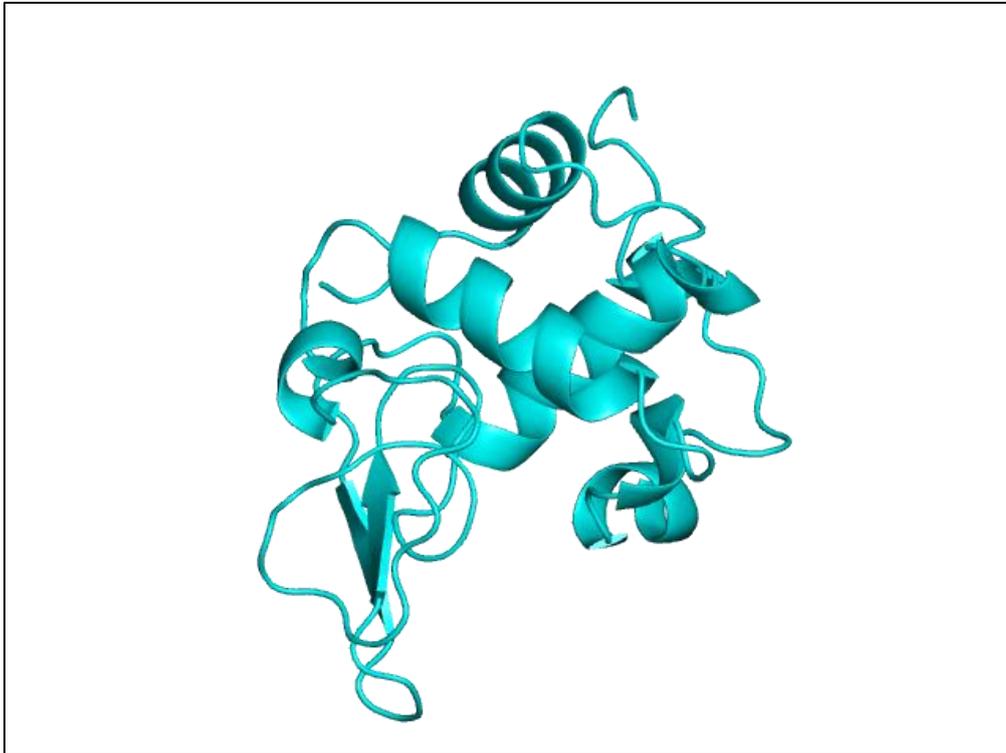
Un'ultima funzione importante del programma PyMOL è quella che consente di preparare immagini di buona qualità della struttura in esame. Per migliorare la qualità della visualizzazione prima di preparare l'immagine, si utilizza il comando *ray*, che può essere digitato nella barra di comando oppure selezionato dalla finestra di comando (26). La qualità dell'immagine che ne risulta è più definita e l'osservatore percepisce la profondità grazie agli effetti della luce sulla molecola stessa (27).



A questo punto la struttura non dovrà essere modificata o ruotata, altrimenti l'ottimizzazione della visualizzazione andrà persa (ma può essere facilmente ripristinata con il comando *ray*). Il comando successivo, che serve per salvare l'immagine in formato png, è semplicemente:

```
> png <nome file>.png
```

Attenzione: il file verrà salvato nella cartella di lavoro. Se il file pdb iniziale è stato aperto dall'icona del file stesso, PyMOL salverà i file ottenuti in una cartella apposita all'interno della directory del programma. E' possibile visualizzare quale sia questa directory, digitando il comando *pwd*.



#### Riferimenti.

- [1] The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.
- [2] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. "The Protein Data Bank" **Nucleic Acids Research**. 2000; 28, 235-242 (<https://www.rcsb.org/>)
- [3] Community-run support site for the PyMOL molecular viewer (<https://pymolwiki.org/>)