

PREPARAZIONE DI IMMAGINI E FILMATI DI STRUTTURE PROTEICHE UTILIZZANDO IL SOFTWARE UCSF CHIMERAX.

Esercizio: Analisi della struttura della proteasi di HIV nelle conformazioni aperta e chiusa e del suo legame con il substrato

La storia dello studio strutturale della proteasi del virus HIV comincia molto lontano, subito dopo la scoperta del virus (1983), ed è una delle storie di maggior successo dell'approccio *structure-based drug design*. All'inizio, non essendo ancora disponibili gli strumenti di biologia molecolare che oggi rendono relativamente semplice il lavoro di espressione e purificazione di proteine ricombinanti, il campione di proteina utilizzato per gli studi strutturali era stato ottenuto per sintesi chimica [1]. Sulla base delle prime strutture cristallografiche inibitori sempre più potenti sono stati sintetizzati e messi sul mercato come farmaci antiretrovirali (ART, *anti-retroviral therapy*) [2].

Purtroppo, uno dei problemi maggiori dell'epidemia di HIV è l'insorgere di mutazioni nel virus, inclusa la sequenza proteica della proteasi, che rendono le nuove varianti resistenti ai farmaci ART. Lo sviluppo di varianti resistenti è facilitato da regimi farmacologici discontinui, una situazione che purtroppo si incontra spesso nelle regioni in cui si sviluppano nuovi focolai epidemici, spesso in paesi economicamente svantaggiati. Nonostante la strada fatta, quindi, c'è ancora bisogno di nuovi farmaci che abbiano attività inibitoria non solo sulla proteina *wild-type* ma anche sulle sue varianti cliniche.

La proteasi di HIV è uno dei 3 enzimi presenti nel genoma virale (gli altri sono la trascrittasi inversa e l'integrasi). E' una proteasi aspartica omodimerica in cui i residui catalitici, aspartato 25, sono presenti su entrambi i monomeri, ma il sito attivo si forma solo con la dimerizzazione. Questa proteasi è specifica per la sequenza Phe/Thr-Pro, presente nella proteina Gag-Pol ottenuta per traduzione della sequenza genica del virus. Perché il virus possa essere replicato all'interno della cellula ospite, è indispensabile l'idrolisi della poliproteina Gag-Pol catalizzata dalla proteasi, che produce tutte le proteine virali funzionali, sia proteine strutturali che enzimi.

L'obiettivo di questo esercizio è di esplorare in un video la struttura della proteasi da HIV nelle sue conformazioni aperta (struttura *apo*, senza substrati, codice PDB: 2PCO, [1]) e chiusa (con alcuni substrati, codici PDB: 1KJ4, 1KJ7, 1KJG, 1F7A, 1KJF, 1KJH [2]). Nel caso delle strutture con substrato, la proteina utilizzata non è la forma attiva della proteasi, ma una proteina mutata in cui i residui di aspartato sono stati sostituiti con residui di asparagina.

Per prima cosa è necessario ottenere da PDB i file con le coordinate dei modelli e aprirli. Questo può essere fatto con:

```
> open 2pc0 fromDatabase pdb
> open 1kj4 fromDatabase pdb
> open 1kj7 fromDatabase pdb
> open 1kjg fromDatabase pdb
> open 1f7a fromDatabase pdb
> open 1kjf fromDatabase pdb
> open 1kjh fromDatabase pdb
```

Nel pannello *Log* sono descritte le principali caratteristiche delle strutture cristallografiche. Dall'esame di questo pannello dopo aver caricato i file delle strutture si osserva che:

- (1) la struttura 2PCO ha solo una catena polipeptidica di 99 residui nell'unità asimmetrica, oltre ad uno ione magnesio e una molecola di glicole propilenico (probabilmente un additivo utilizzato nella cristallizzazione); per ottenere la molecola intera sarà necessario applicare gli operatori di simmetria.

- (2) la struttura 1KJ4, invece, ha 4 catene polipeptidiche nell'unità asimmetrica, che formano 2 dimeri di proteasi; inoltre, sono presenti 2 corte catene peptidiche (denominate P ed S), di 10 residui ciascuna, che costituiscono il substrato; sono presenti anche ioni acetato derivati dal buffer di cristallizzazione.
- (3) tutte le altre strutture hanno un dimero di proteasi nell'unità asimmetrica, assieme ad un peptide substrato (catena P, 10 residui per tutti i modelli); anche in queste strutture sono presenti ioni acetato.

Per prima cosa uniformiamo la visualizzazione e selezioniamo un background bianco:

```
> hide atoms
> show ~/P~/S cartoon
> set bgColor white
```

Per vedere le differenze tra i due dimeri di proteina presenti nella struttura 1KJ4, è necessario per prima cosa creare due nuovi oggetti che contengono ciascuno dei dimeri. Per farlo, prima di tutto si divide l'oggetto iniziale nelle sue catene (comando **split**) e poi si creano due nuovi oggetti con le catene opportune (comando **combine**), includendo anche i substrati presenti all'interno del sito attivo di ciascun dimero:

```
> hide ~#2 cartoon
> split #2
> combine #2.1 #2.2 #2.5
> combine #2.3 #2.4 #2.6
> color #9 red
> matchmaker #9 to #8
```

Da qui osserviamo che ci sono solo piccole differenze per ciascuna catena, soprattutto sui loop esterni, probabilmente dovute all'impaccamento cristallino. Per comodità, considereremo solo il primo dimero (oggetto #8).

Passiamo alla visualizzazione dell'unità funzionale della prima struttura, 2PCO.

```
> hide #2,8,9 cartoon
> show #1 cartoon
> crystalcontacts #1 distance 2
> view
```

Osserviamo che sono state create altre 2 catene polipeptidiche simmetriche (oggetto #10), una delle quali (#10.1) è la catena polipeptidica che forma il dimero con quella presente nell'unità asimmetrica della struttura. Combiniamo le catene che formano un dimero:

```
> combine #1 #10.1
> hide #1,10 cartoon
```

A questo punto sovrapponiamo le strutture dei dimeri caricati:

```
> matchmaker #11,8,4,5,6,7,8 to #3
> show #3,4,5,6,7,8,11 cartoon
```

La prima cosa che è interessante osservare dal punto di vista strutturale è la posizione dei loop che sono posizionati al di sopra del sito attivo, generalmente chiamati *flaps*. Per farlo visualizziamo solo le strutture della proteina *apo* (dimero, oggetto #11) e di uno dei complessi con il substrato (ad esempio #3) e facciamo una transizione dell'uno all'altro:

```
> hide ~#3,11 cartoons
> split #3
> combine #3.1 #3.2
> morph #11 #12
```

Il comando **morph** crea una traiettoria per interpolazione dall'una all'altra struttura. Questo comando, però, non funziona su oggetti che hanno un numero di catene diverso, quindi è necessario anche creare un ulteriore oggetto (#12) che include le catene polipeptidiche del dimero presente nella struttura del complesso con il

substrato (oggetto #3). Il comando `morph` ha creato un ulteriore oggetto (#13) che contiene le traiettorie. Si osserva che c'è una chiusura dei *flaps* contemporanea ad una rotazione relativa dei due monomeri che porta alla chiusura del sito attivo della proteina quando il substrato è presente.

Osserviamo a questo punto le posizioni dei residui attivi delle proteine in esame (Asp 25 di entrambi i monomeri nella proteina *apo*, mentre Asn 25 di entrambi i monomeri per i complessi con i diversi substrati) e dei substrati presenti:

```
> name catalytic_residues (#3,4,5,6,7,8:25) | #11:1025
> name substrates /P#3,4,5,6,7,8
> show catalytic_residues atoms
> show substrates atoms
> hide substrates cartoon
> style stick
> hide :HOH atoms
```

Coloriamo le strutture con diversi colori, in modo da poterle più facilmente visualizzare (i colori sono stati scelti da [3]):

```
> show #3,4,5,6,7,8,11 cartoon
> color #3/A,B #c51b8a; color #3/P #fbb4b9
> color #4/A,B #31a354; color #4/P #c2e699
> color #5/A,B #8856a7; color #5/P #b3cde3
> color #6/A,B #f03b20; color #6/P #fecc5c
> color #7/A,B #2b8cbc; color #7/P #bdc9e1
> color #8/A,B #636363; color #8/P #cccccc
> color #11/A,B blue
> color byhet [questo comando consente di colorare con i colori usuali gli atomi che non sono carbonio per ciascuna struttura: azoto blu, ossigeno rosso, solfo giallo, idrogeno bianco, fosforo arancione]
> close #13 [questo comando serve a chiudere l'oggetto generato dal comando morph; se l'oggetto non viene chiuso, può dare problemi successivamente, quando vogliamo registrare il filmato].
```

A questo punto siamo pronti per registrare un piccolo filmato che mostra l'analisi che abbiamo condotto. Prima di tutto la "sceneggiatura":

- 1) all'inizio il filmato mostrerà la proteina *apo* nella sua conformazione aperta, sia nella struttura complessiva (posizione della camera: *overall*), sia nel dettaglio del sito attivo (posizione della camera: *active_site*);
- 2) utilizzeremo il *morphing* per mostrare come è cambiata la conformazione della forma aperta per ingresso del substrato nel sito attivo;
- 3) a questo punto osserveremo alcune delle conformazioni dei substrati nelle diverse strutture dal lato (posizione della camera: *sub1*), dall'alto rispetto al sito attivo (posizione della camera: *sub2*) e nel particolare del sito attivo (posizione della camera: *sub3*);
- 4) infine, torneremo alla conformazione *apo* della proteina.

Salviamo le posizioni della camera che vogliamo utilizzare per il filmato:

```
> view name overall1 [posizione che mostra tutta la proteina apo, con l'asse di simmetria C2 verticale e parallelo alla camera]
> view name overall2 [posizione che mostra tutta la proteina #3, con l'asse di simmetria C2 verticale e parallelo alla camera, con lo stesso zoom della posizione precedente; è necessario correggere la posizione perchè la sovrapposizione delle strutture apo e con il substrato non è perfetta e l'asse (o pseudoasse) non coincide perfettamente]
> view name active_site [posizione in cui si osservano i residui catalitici e la tasca del sito attivo]
```

- > **view name sub1** [dalla posizione overall, un ingrandimento mostra la posizione del substrato; l'asse C2 non è presente nelle strutture che contengono il substrato perchè il substrato stesso rompe la simmetria della proteina; però possiamo considerare un pseudo asse C2 e posizionare la camera in modo che lo pseudo asse sia parallelo alla camera e verticale]
- > **view name sub2** [posizione a 90° rispetto alla precedente, con lo pseudoasse perpendicolare alla camera]
- > **view name sub3** [posizione simile alla precedente, ma con maggiore zoom e rimuovendo con un comando `clip` la parte superiore della proteina (*iflaps*)].

Prima di iniziare la registrazione del filmato, salviamo la sessione di ChimeraX (file `cxs`), con il comando:

- > **save C:\[path]\HIVprotease.cxs**

Comandi per la preparazione del filmato:

- > **hide ~#11 cartoons, atoms**
- > **show #11:1025 atoms**
- > **color byhet**
- > **view overall1**
- > **cofr centerOfView** [questo comando fissa l'origine del sistema di assi utilizzato per la rotazione al centro dell'immagine visualizzata]
- > **movie record**
- > **roll y 2 180**
- > **fly overall1 30 active_site**
- > **crossfade 10; show (#11:1025 :< 8) atoms** [questo comando mostra tutti gli atomi dei residui ad una distanza minore di 8 Å, appartenenti al modello #11]
- > **rock y 30 60 cycle 30** [ruota la proteina di 30° attorno all'asse y per 60 frame, con cicli di 30 frame]
- > **wait 10**
- > **fly active_site 30 overall1; crossfade 30; hide atoms**
- > **morph #11,12; crossfade 50; show #3/P atoms; color #13 #c51b8a; view overall2; wait 3**
- > **show #!11 models; hide #13 cartoon; hide #11 cartoon, atoms**
[questo comando serve a nascondere il *morphing* (#11) ed attivare di nuovo il modello #11 che era stato nascosto durante il *morphing*; equivale a riattivare la visualizzazione del modello #11 dal pannello *Models*]
- > **show #3/A,B cartoon; show (#3/P :< 8) atoms; wait 3**
- > **fly overall2 30 sub1 30 sub2**
- > **crossfade 10; hide ~(#3:25|#3/P) atoms**
- > **fly sub2 30 sub3**
- > **rock y 30 60 cycle 30**
- > **wait 10**
- > **crossfade 20; hide #3 cartoon, atoms; show #4/A,B cartoon; show #4:25|#4/P atoms; cofr centerOfView**
- > **fly sub3 20 sub2; clip off; cofr centerOfView** [il comando `clip off` serve per riportare alla posizione iniziale il primo piano e lo sfondo]
- > **fly sub2 20 sub1**
- > **crossfade 20; hide #4 cartoon, atoms; show #5/A,B cartoon; show (#5/P :< 8) atoms; wait 20**
- > **fly sub1 20 sub2; wait 20**
- > **crossfade 20; hide ~(#5:25|#5/P) atoms; view sub2; wait 10**
- > **crossfade 20; clip off; hide #5 cartoon, atoms; show #6/A,B cartoon; show #6:25|#6/P atoms; wait 10**
- > **fly sub2 20 sub1; wait 10**

```
> crossfade 10; hide #6 cartoon, atoms; show #7/A,B cartoon; show
#7:25|#7/P atoms; wait 10
> fly sub1 20 sub2; wait 10
> crossfade 20; hide #7 cartoon, atoms; show #8/A,B cartoon; show
#8:25|#8/P atoms; wait 10
> fly sub2 20 sub1 20 overall2
> crossfade 20; hide #8 cartoon, atoms; show #11/A,B cartoon; view
overall1
> wait 10
> movie encode
```

Riferimenti.

- [1] Heaslet H, Rosenfeld R, Giffin M, Lin YC, Tam K, Torbett BE, Elder JH, McRee DE, Stout CD. "Conformational flexibility in the flap domains of ligand-free HIV protease." *Acta Crystallogr D Biol Crystallogr*. 2007. 63(Pt 8):866-75. doi: 10.1107/S0907444907029125.
- [2] Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. "Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes." *Structure*. 2002. 10(3):369-81. doi: 10.1016/s0969-2126(02)00720-7.
- [3] <https://www.cgl.ucsf.edu/chimerax/docs/user/commands/palettes.html>