

# TRABAJO PRÁCTICO N°7

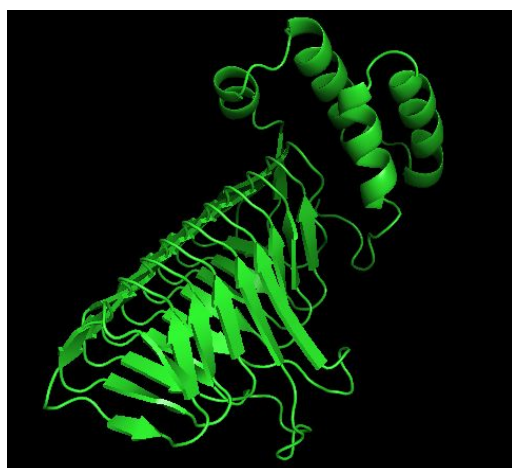
---

## 1. ¿Cómo describís la estructura de esta proteína?

Las estructuras de las proteínas pueden ser muy diversas. Se las puede clasificar por su contenido en estructuras secundarias y según su cantidad de dominios. ¿Cómo podrías describir la estructura de esta proteína? Realiza la misma descripción pero para la proteína 1THJ, 30GB.

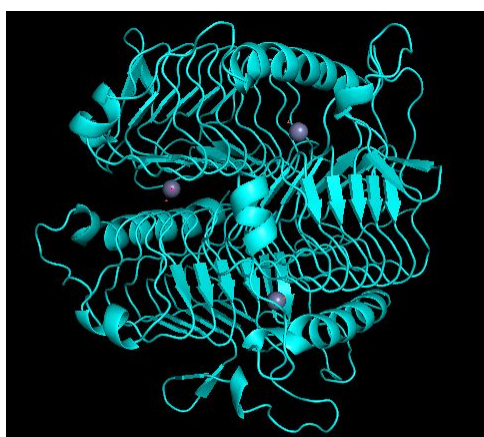
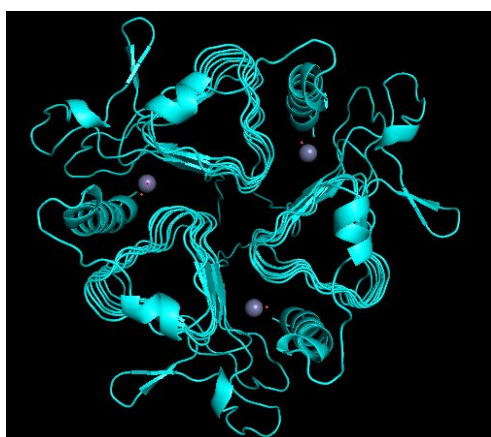
### Proteína 1LXA

Antes de hablar del dominio de una proteína debemos definir qué es. Un



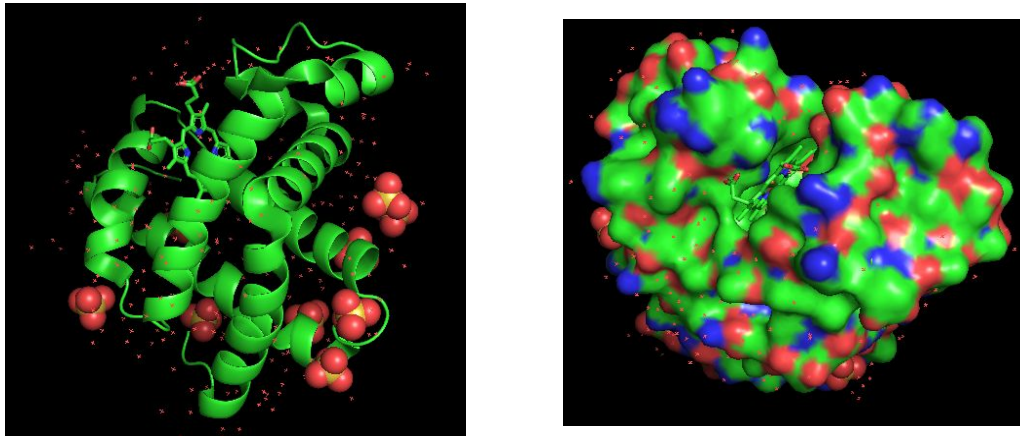
dominio proteico es la zona de la proteína donde se halla mayor densidad, es decir, donde hay más plegamientos. Una cadena polipeptídica puede tener uno o más dominios. Si giramos un poco la proteína 1LXA podemos ver que tiene una zona con varias hojas beta plegadas (formando una especie de triángulo) y otra zona que cuenta con varias alfa hélices. Ambas zonas están unidas por giros beta.

### Proteína 1THJ



Esta proteína tiene una estructura un poco más compleja que la anterior. Podemos ver que cuenta con 3 regiones “principales” de estructuras en formas de triángulo constituidas por varias hojas beta plegadas cada una. A su vez, cada uno de estos triángulos tiene una alfa hélice en cada extremo. También hay alfa hélices separando cada uno de los triángulos.

### Proteína 30GB



Esta proteína cuenta con varias alfa hélices y no tiene hojas beta plegadas. Además cuenta con un grupo HEM que podemos apreciar como forma de “bolsillo” en la segunda foto.

## 2. Proteínas con estructura cuaternaria

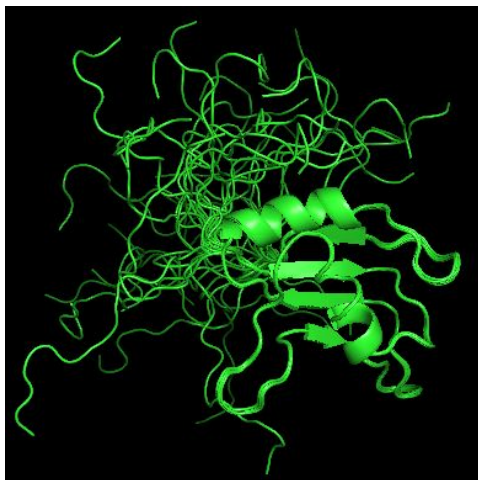
Utilizando la estructura 1THJ vamos a usar un comando para ver las distintas cadenas que la componen. En la línea de comandos, escribí “split\_chains”. Inmediatamente verás en el panel que la estructura 1THJ se ha subdividido en sus cadenas que la componen (A, B y C usualmente). Este



comando también es útil para separar las cadenas cuando la estructura derive de un estudio NMR.

#### 4. Un mar de conformaciones...

Estudiá la proteína 2CPE. ¿Qué tipo de proteína es? ¿Cómo la describirías? Utilizá el comando “set all\_states, on” para ver todos los estados conformacionales estimados para esa estructura. Compará el espacio conformacional de la 2CPE con la de la estructura de la mioglobina (1MYF). Utilizando el modo secuencia, seleccioná la HIS 64. Mostrala en formato stick o lines. ¿Qué función podría llegar a tener? ¿Y la HIS 93?



La proteína 2CPE figura en PDB como “Solution structure of the RNA recognition motif of Ewing Sarcoma(EWS) protein”. A pesar de que tiene un par de hojas beta plegadas, dos alfa hélices y muchos giros beta (o loops), podemos decir que es una proteína desordenada ya que no cumple con ninguno de los patrones de las “estructuras conocidas”.



La mioglobina posee un grupo HEM y la 2CPE no. Tiene una estructura un poco más “ordenada”. El histon HIS-64 además de unir ADN, participa en la regulación positiva de la transcripción por la ARN polimerasa II. Con respecto al HIS-93, tiene participación en la unión de la mioglobina con la hemo.

**RETO I:** Estas estructuras difieren de las estructuras sobre las que venimos trabajando en su determinación. Como habrás notado estas fueron obtenidas mediante la técnica de MNR, ¿Pero en qué consiste esta técnica?

La técnica de MNR (Nuclear Magnetic Resonance) funciona de la siguiente manera: Al someter a una molécula a un campo electromagnético sus átomos

empiezan a actuar como imanes. Luego, se somete a la molécula a un espectro de radiofrecuencia lo que va a hacer que cada núcleo de cada átomo empiece a resonar en su propia frecuencia (esto varía según el tipo de átomo). Esta información es traducida a un espectro MNR. Cuanto mas resuene un átomo, mayor será el pico que veamos en el gráfico del espectro. Gracias a estas vibraciones de los átomos los científicos son capaces de determinar una estructura tridimensional aproximada de una molécula, sus movimientos y su comportamiento.

**RETO II:** Investigá en qué consisten las interacciones puentes de hidrógeno,  $\pi$ - $\pi$  y  $\pi$ -catión y qué aminoácidos podrían intervenir en dichas interacciones.

Puentes de hidrógeno: Son enlaces que se producen a partir de la atracción existente entre un átomo de hidrógeno y un átomo de oxígeno, flúor o nitrógeno con carga negativa.

Enlaces  $\pi$ - $\pi$ : Los enlaces pi son enlaces químicos covalentes donde dos lóbulos de un orbital en un átomo se superponen con dos lóbulos de un orbital en otro átomo y esta superposición se produce lateralmente.

Enlaces  $\pi$ -catión: La interacción catión- $\pi$  es una interacción molecular no covalente entre la cara de un sistema pi rico en electrones (vg. benceno, etileno) con un catión adyacente.

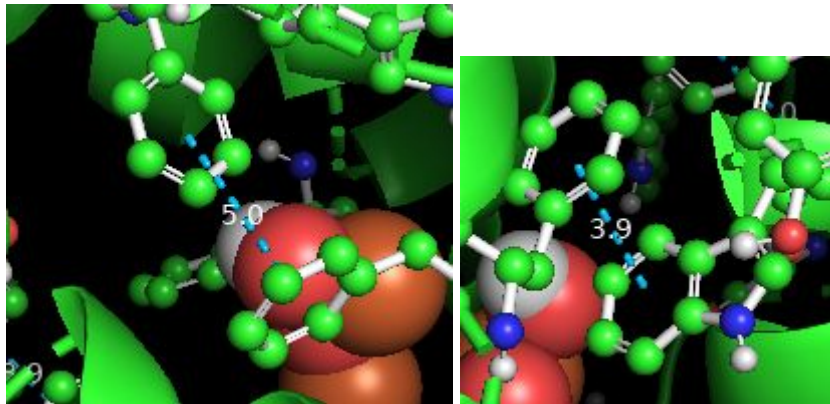
## 6. Identificación de interacciones $\pi$ - $\pi$ y $\pi$ -catión

Estudiaremos más de cerca estas interacciones, utilizando la proteína miohemeritina (hacé fetch del PDB 1A7E). Una vez que la proteína se cargó:

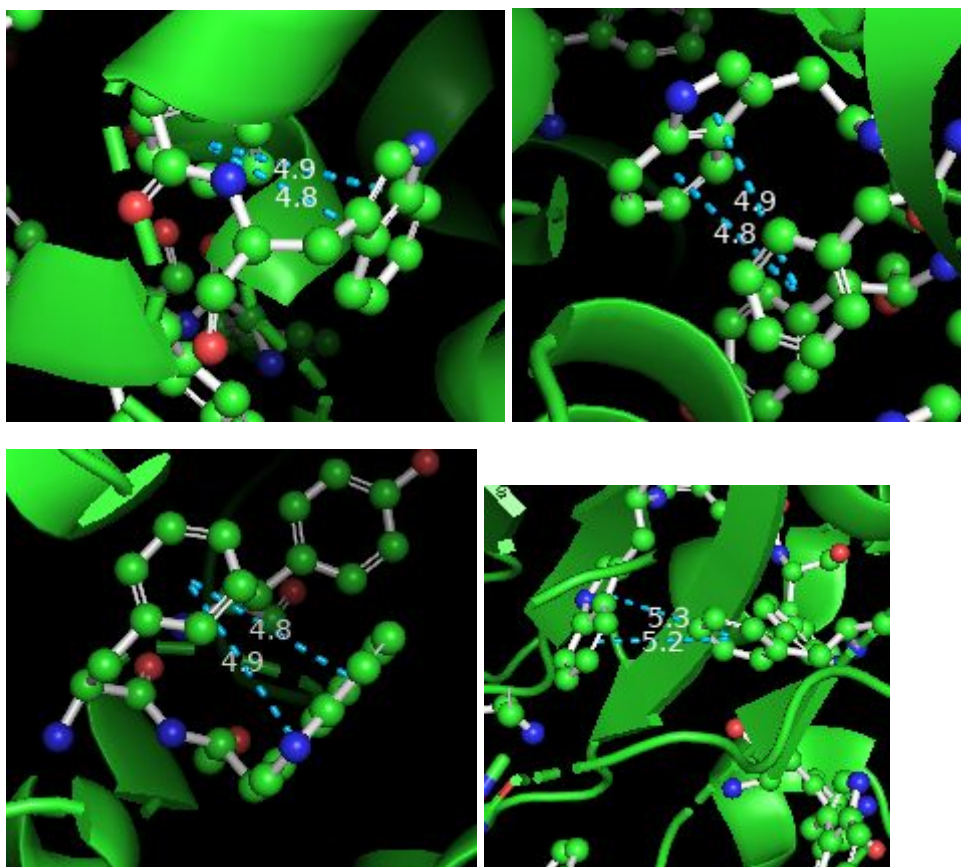
- Ocultá las moléculas de agua (H  $\rightarrow$  waters) ,
- Seleccioná los residuos aromáticos (PHE/F, TYR/Y, TRP/W) manualmente en la barra de secuencia o mediante el siguiente comando: `select aro, resn PHE+TYR+TRP`. El comando `select` se puede utilizar para seleccionar cualquier átomo, residuo, cadena, etc. Para más información sobre el comando `select` visite `Select help`. Más detalles en el ítem `Selection_Algebra`. [`select selection_name, type XXX+XXX+...`]
- Una vez seleccionados los residuos aromáticos, mostralos como bolas y palitos (A  $\rightarrow$  preset  $\rightarrow$  balls and sticks) y localizá anillos aromáticos que participen de interacciones  $\pi$ - $\pi$ . Una vez identificadas medí la distancia entre los anillos, para ello, en el panel de medidas (measurements)



seleccioné Measurement mode → distance to rings. Usando las estructuras 1A7E y la 1EJ1 y los protocolos descritos en la parte anterior, exploré la proteína en busca de interacciones  $\pi$ -catión. Para ello seleccioné los aminoácidos con carga positiva y mostralos como bolas y palitos. Medí la distancia entre los grupos intervinientes.



1EJ1:

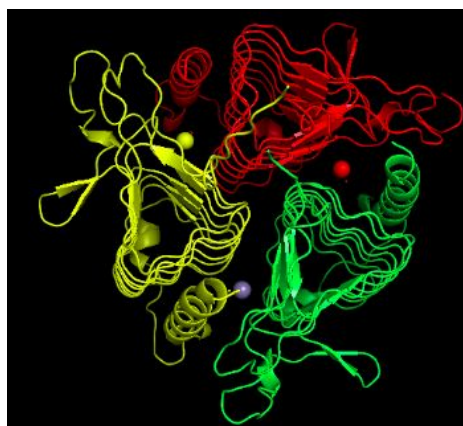


## 7. Sitio activo de la anhidrasa carbónica

Utilizado la estructura de la anhidrasa carbónica 1THJ y la opción Select en el Menú Edit:

a. ¿Cuántas subunidades tiene la estructura nativa?

Podemos notar 3 subunidades principales



b. ¿Tiene heteroátomos esta molécula? ¿Cuáles?

Si. Figuran abajo de todo en el PDB. Son el Zinc y el Oxígeno

HETATM	4834	ZN	ZN	A	214	37.283	50.955	133.578	1.00	41.02	ZN
HETATM	4835	ZN	ZN	B	214	28.238	68.130	148.898	1.00	42.04	ZN
HETATM	4836	ZN	ZN	C	214	15.406	47.566	144.477	1.00	44.05	ZN
HETATM	4837	O	HOH	A	215	29.571	68.877	147.658	1.00	25.06	O
HETATM	4838	O	HOH	B	215	14.158	48.942	145.115	1.00	28.77	O
HETATM	4839	O	HOH	C	215	36.130	49.609	132.673	1.00	27.90	O

c. Identificá los residuos que unen el Zinc. Para ello utilizá los siguientes comandos:

- select zincs, metals
- select nearzincs, zincs around 6

También se puede usar el comando: select all within 6 of metals. Cambiá la visualización de la selección a bolas y palillos.

d. ¿Los residuos que unen el Zinc pertenecen a la misma subunidad?

No. Podemos ver que cada zinc está unido a dos subunidades al mismo tiempo. Es decir, actúa como ligando.

e. Determiná los contactos proteína-proteína en la estructura cuaternaria de la anhidrasa carbónica. En este punto usaremos distintos programas para establecer qué aminoácidos participan en contactos proteína-proteína. Para ello accedé al servidor Protein-Protein interaction server InterProSurf

(<http://curie.utmb.edu/pdbcomplex.html>) y cargá el código pdb 1THJ. Registrá

el número de los residuos que mantiene contactos entre subunidades. Visualizá los mismos en Pymol. Utilizá la barra de secuencia o los siguientes comandos:

- select interactingA, chain A & resi x+x+x+x... donde x es el número de los residuos que se desea seleccionar (Obtenidos de InterProSurf). Realizá la misma operación para las cadenas B y C.

Se puede ver en las imágenes los residuos que están en contacto en la interfaz para la secuencia A seleccionando los residuos que se pueden ver en el comando que sigue.

```
PyMOL>select interactingA, chain A & resi 0+1+2+10+13+15+16+17+18+19+35+36+37+39+201+204+208+209+212
```

Selector: selection "interactingA" defined with 155 atoms.



```
PyMOL>select      interactingB,      chain      B      &      resi
0+1+2+3+10+12+13+15+16+35+57+81+99+117+208+209+212

Selector: selection "interactingB" defined with 142 atoms.
```



```
PyMOL>select      interactingC,      chain      C      &      resi
0+1+2+3+10+12+13+15+16+35+37+39+122+135+136+205+208+209

Selector: selection "interactingC" defined with 150 atoms.
```



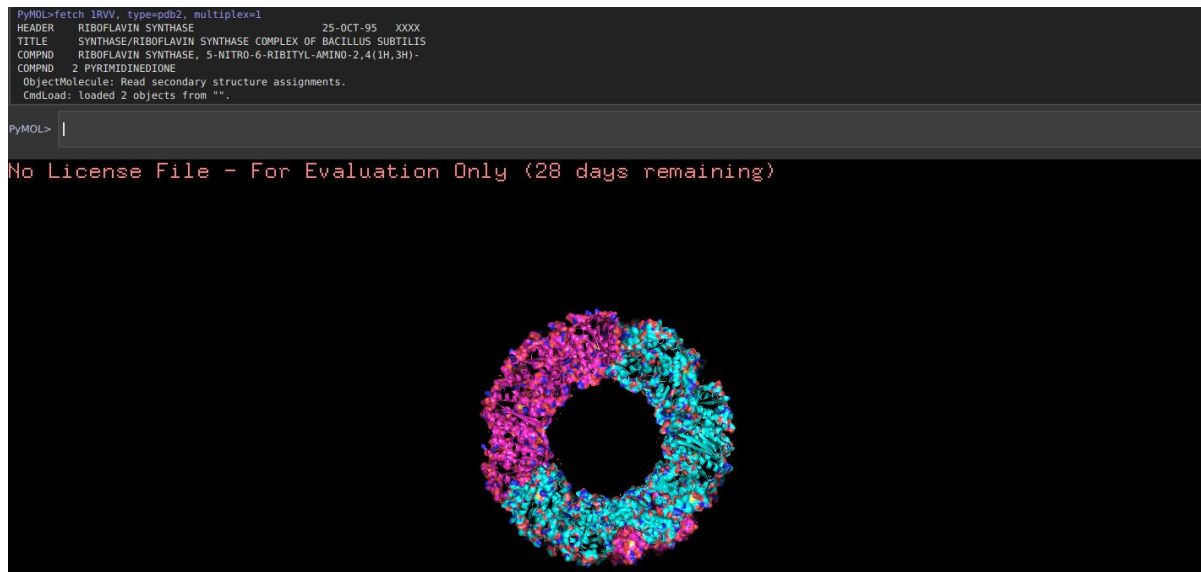


## 8. Uso de superficies y slab model en pymol

Cargá la estructura 1RVV biological assembly 2 a partir del archivo 1RVV\_full.pdb, o utilizando el comando:

```
fetch 1RVV, type=pdb2, multiplex=1
```

Esta proteína es la lumazina sintasa y está formada por 62 monómeros formando una esfera. Utilizando **S** → Surface determiná la superficie de la proteína. Ahora usando la rueda del mouse utilizá el slab mode para recorrer en profundidad la esfera que forma la proteína.



## 9. Estudio de la porina de *N. gonorrhoeae*

Utilizando el programa Pymol estudiá la estructura terciaria y cuaternaria de la proteína 4AUI. Determiná la superficie, localizá ligandos e identifícalos. Si querés obtener más información visitá el sitio <https://www.rcsb.org/> y buscá por el ID: 4AUI.

Para seleccionar los ligandos podés usar el comando:

```
select ligandos, organic
```

Para seleccionar grupos inorgánicos podés usar el comando:

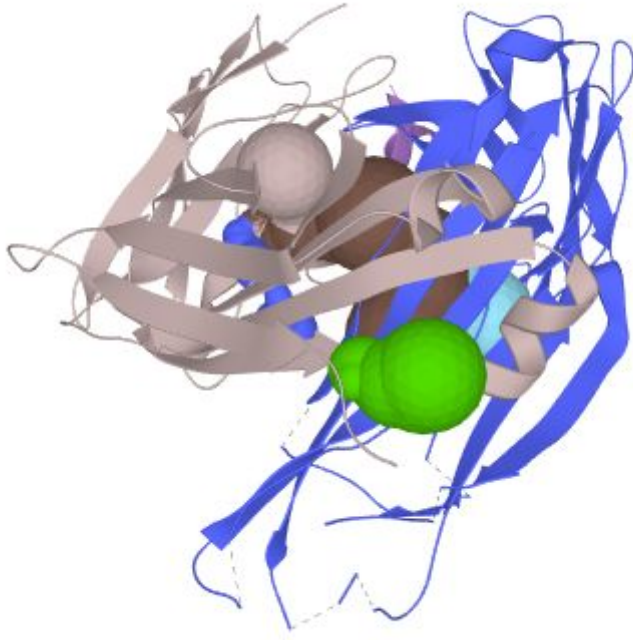
```
select inorganicos, inorganic
```

## 10. Estudio de túneles

Utilizando el programa (o servidor <https://mole.upol.cz/>) MOLE estudiá los túneles asociados al sitio activo de las estructuras 1THJ y 1F90. Este programa te pedirá que indiques a partir de qué regiones de la proteína debe comenzar la exploración en búsqueda de túneles. El punto de inicio

puede ser a partir de cavidades, átomos o residuos seleccionados manualmente, o a partir de la base de datos CSA (Catalytic site atlas).

1F90:



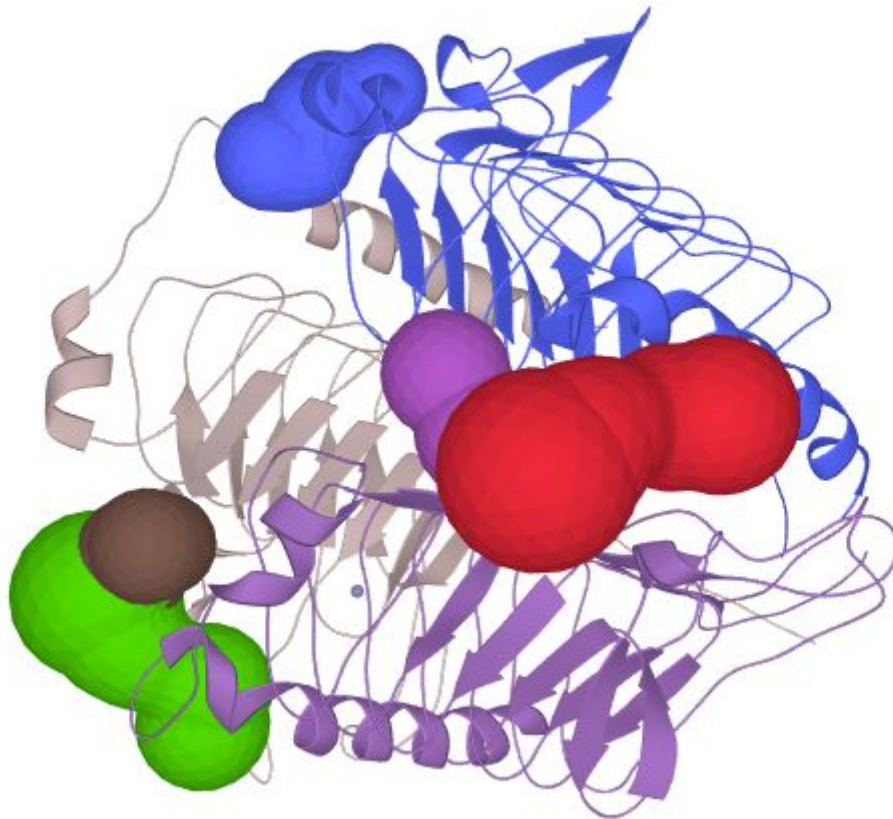
Podemos ver que en mole nos muestra 5 túneles y 13 cavidades.

1THJ:



Podemos visualizar un solo túnel con mole. Nos dice que tiene 12 cavidades.

Ajustando parámetros de las cavidades y los túneles encontramos esto:



Acá vemos 3 cavidades y 5 túneles.

### 11. Estudio de pockets en EGFR

El EGFR es uno de los principales marcadores de cáncer de pulmón. Utilizando la sesión de Pymol estudié los pockets. 1M14 es un conformero activo y 3W32 uno inactivo. Comparé el sitio activo de ambos conformeros así como también los tamaños de los pockets. La determinación de las cavidades se puede realizar mediante el servidor fpocket: <https://mobyli.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::fpocket> Desactivar la opción de demo: "Test the service with server sample data (input parameters will be discarded)".

### 12. Túneles y diseño de fármacos

El estudio estructural de proteínas nos proporciona múltiples campos de aplicación, uno de los más explorados en la actualidad es el diseño racional de fármacos. Si se conoce la base biológica de una enfermedad, es decir se conocen las moléculas implicadas, es posible diseñar medicamentos

que interactúen con la molécula responsable, de tal forma que la modifique y se modifique el cuadro patológico. En otras palabras, el diseño racional de fármacos consiste en la aplicación del conocimiento biológico y estructural de los receptores (proteínas involucradas en una dada enfermedad) para diseñar moléculas que interactúen sólo con estos... dentro de lo posible! Veamos el ejemplo de la proteína Acetil-CoA sintetasa de humano, involucrada en cáncer de mamas. Como bien sabemos para que una droga pueda unirse a la proteína y modificar su función debe tener dónde hacerlo. Tomando como base la estructura PDB 3GPC, estudié sus túneles utilizando el Fpocket

¿Encontrás algún/os túnel/es en la proteína donde pueda unirse algún fármaco?

¿Qué residuos se encuentran en dicha cavidad?

Ahora analicemos los resultados de los investigadores de la UBA y UNQ (<https://doi.org/10.1007/s00018-020-03679-5>), que mediante una técnica de acoplamiento molecular (docking) diseñaron un inhibidor para esta proteína, capaz de unirse a la misma y reducir entonces la proliferación de células tumorales.

**RETO III:** Investigá en qué consiste el docking, en qué ideas basa su funcionamiento.

El docking o acoplamiento molecular es una técnica de mecánica molecular usada para predecir la orientación del enlace de una pequeña molécula candidata a fármacos, con la proteína que ejercerá acción, con lo que se podrá predecir la afinidad y la actividad de la molécula pequeña. Por tal motivo este método tiene un rol muy importante en el diseño racional de fármacos. Para tal fin entran en juego una serie de procesos químicos gobernados por leyes físicas, entre ellas las que tienen que ver con la energía que se consume o libera en tal proceso.

Descargá la estructura de la proteína y la del inhibidor y visualizalas en Pymol (abrí ambos archivos en la misma sesión) ¿Dónde se une el inhibidor? ¿Coincide con el túnel que propusiste anteriormente? ¿Dada esta inspección



ocular cómo creés que actúa el inhibidor? Usando el play de la botonera, observá las distintas conformaciones del ligando o inhibidor.

### **13. Asignación de estructura secundaria basado DSSP**

Para inferir estructura secundaria utilizaremos el servidor 2StrucCompare con la Endolisina de bacteriófago en dos variantes (3F8V y 4LMZ).

1) Primero determiná la estructura secundaria de ambas conformaciones por separado usando el servidor <https://2struccompare.cryst.bbk.ac.uk/index.php>

a) Corré el análisis por separado de cada estructura (recuadro derecho), cargando su código PDB y haciendo click en submit

b) Observá las distintas regiones de estructura secundaria, ¿todas son de igual cantidad de residuos? ¿Se requieren más o menos residuos para formar una alfa hélice que un loop? ¿Por qué?

2) Luego podemos comparar ambas conformaciones utilizando el mismo servidor <https://2struccompare.cryst.bbk.ac.uk/index.php>. Esta vez cargaremos las estructuras a la par (recuadro izquierdo) y el servidor nos mostrará las diferencias

a) Utilizando el botón Colour Scheme especificá las zonas de mayor movimiento de la proteína