

Introducción

El acoplamiento molecular es un método bioinformático que permite predecir y calcular computacionalmente la posición más favorable de interacción entre un ligando y un blanco usualmente proteico a partir de sus representaciones tridimensionales (Figura 1). Mientras más estable, específica y favorable sea la unión entre un ligando (fármaco) y su blanco proteico (diana terapéutica), mayor será su actividad biológica (fármaco más efectivo). Por tanto, este método bioinformático juega un rol importante en el descubrimiento y desarrollo de nuevos principios activos o fármacos.

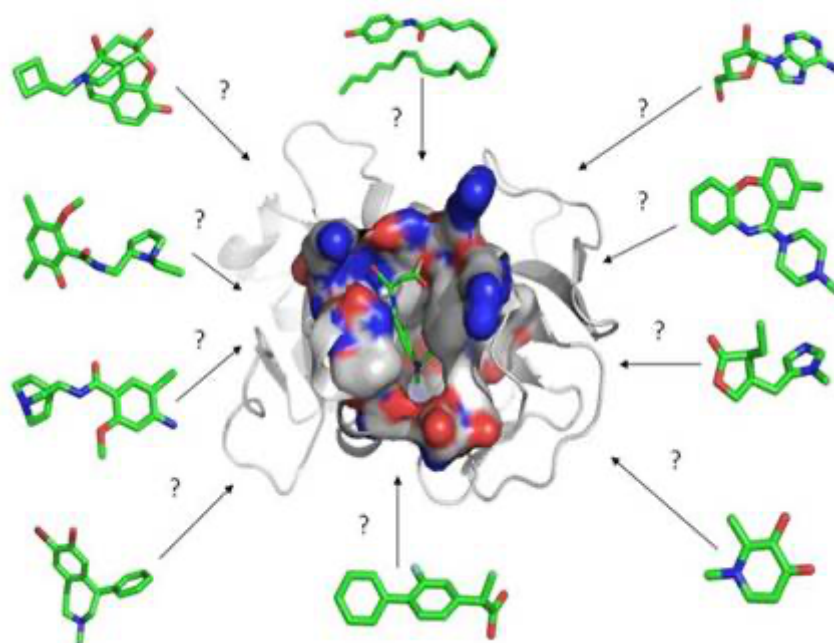


Figura 1. Representación de diferentes ligandos dispuestos alrededor del sitio activo de un blanco proteico. Los ligandos serán evaluados por acoplamiento molecular y seleccionados virtualmente de acuerdo a un algoritmo de muestreo y de puntaje (Extraído de Beckham, 2014).

DOCK es uno de los primeros programas de acoplamiento desarrollados. Su versión actual DOCK6, introduce nuevas funciones de puntuación y métodos de análisis. Como se mencionó anteriormente, este es uno de los pocos programas con implementación de GPU para la puntuación AMBER y el cálculo PBSA/GBSA para complejos proteína-ligando. Como está diseñado para funcionar con DockPrep, se requiere la descarga y el uso de USCF Chimera, sin embargo, esto permite la introducción de otra herramienta muy útil y eficiente. Las desventajas relativas de DOCK son que solo se implementa en Linux, requiere cierta experiencia en compilación de software y un rendimiento relativamente inferior de su función de puntuación predeterminada en comparación con otros programas.

Métodos y resultados

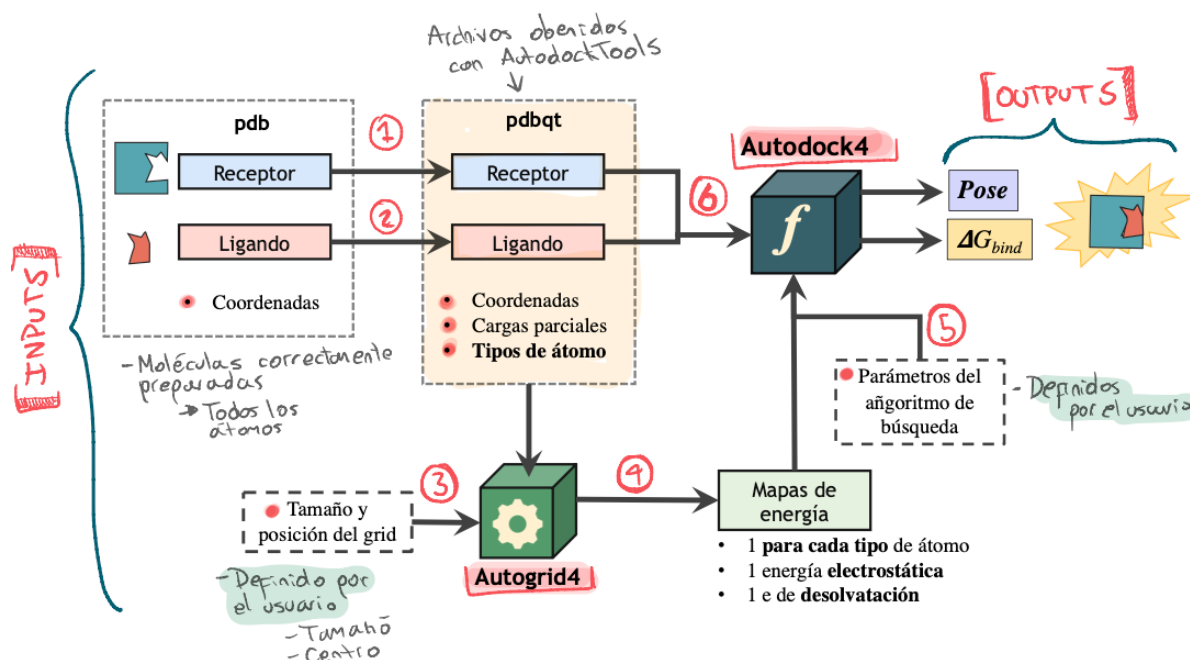


Figura 2. Acoplamiento Molecular con AD4. Flujo de ejecución de Autodock 4.2 (J. Ricci L., 2021)

Obtención de la proteína

Se trabajó con la proteína CDK2 cuya información se encuentra en el siguiente enlace: <https://bit.ly/30FUxav>; una quinasa involucrada en el control del ciclo celular; esencial para la meiosis, pero prescindible para la mitosis. Actúa en la etapa G1-S para promover el programa transcripcional E2F y el inicio de la síntesis de DNA, y modula la progresión de G2; controla el momento de entrada en mitosis/meiosis controlando la activación posterior de la ciclina B/CDK1 por fosforilación, y coordina la activación de la ciclina, papel crucial en la orquestación de un fino equilibrio entre

la proliferación celular, la muerte celular y la reparación del ADN en células madre embrionarias humanas (hESC).

Posteriormente se descargó la proteína y se abrió con UCSF Chimera. Seguidamente guardamos en un archivo .pdb los átomos de la proteína CDK2. Con el ambiente de conda dock activado, preparamos la estructura de la proteína con PDB2PQR dentro de wd_dk, creando un nuevo archivo prot.pdb con la nomenclatura correcta en Amber y con los estados de protonación a pH 7 de los residuos ionizables.

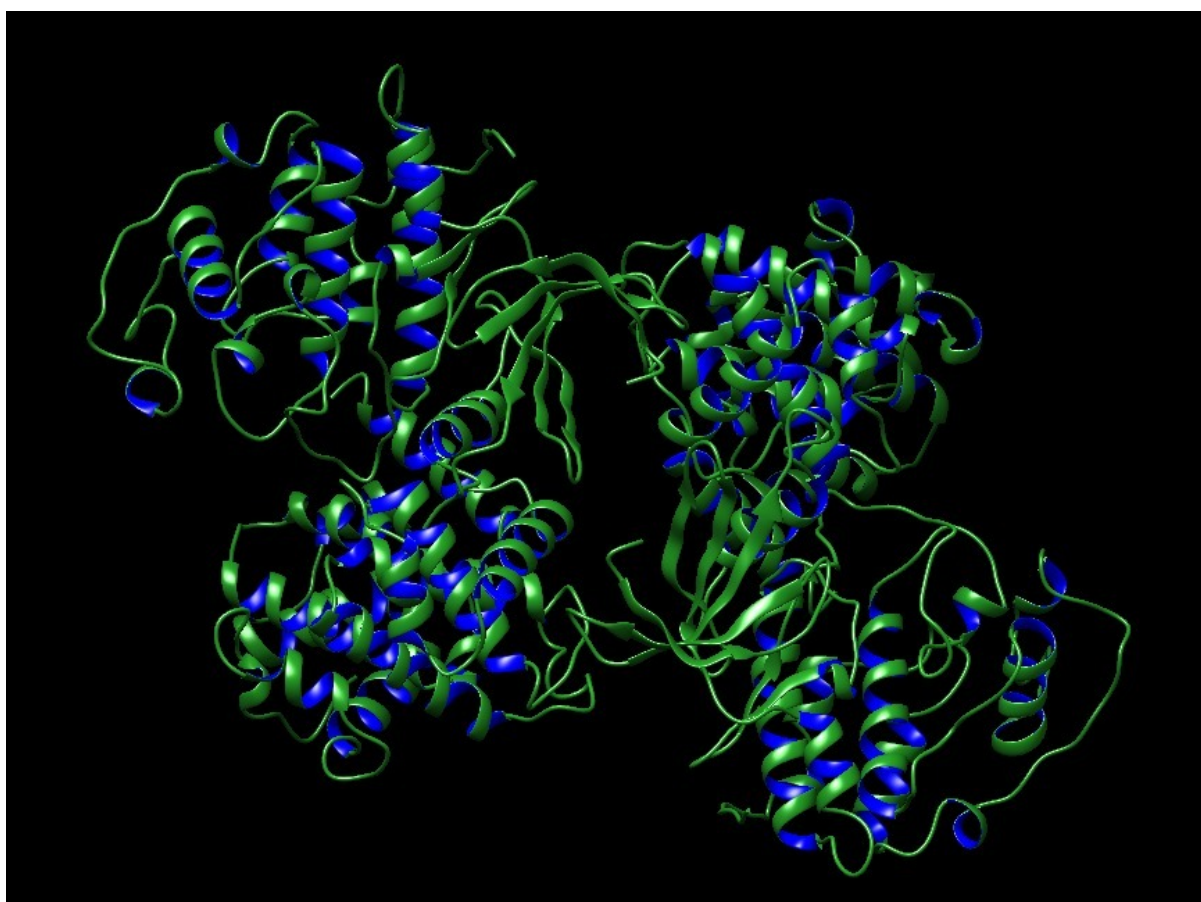


Figura 3. Visualización de la proteína (prot.pdb) en UCSF Chimera.

Empleamos la base de datos PubChem para descargar la “estructura” de la molécula de ATP, que usaremos como ligando. Guardamos el SMILES de esta molécula en un archivo de texto en la carpeta de trabajo como smiles_atp.smi.

Para obtener y preparar las moléculas, primero creamos el ambiente de trabajo para guardar los archivos que se irán generando. Convertimos el SMILES a formato mol2 utilizando Open Babel para generar las coordenadas 3d, añadir hidrógenos a pH 7, y añadir cargas parciales de Gasteiger.

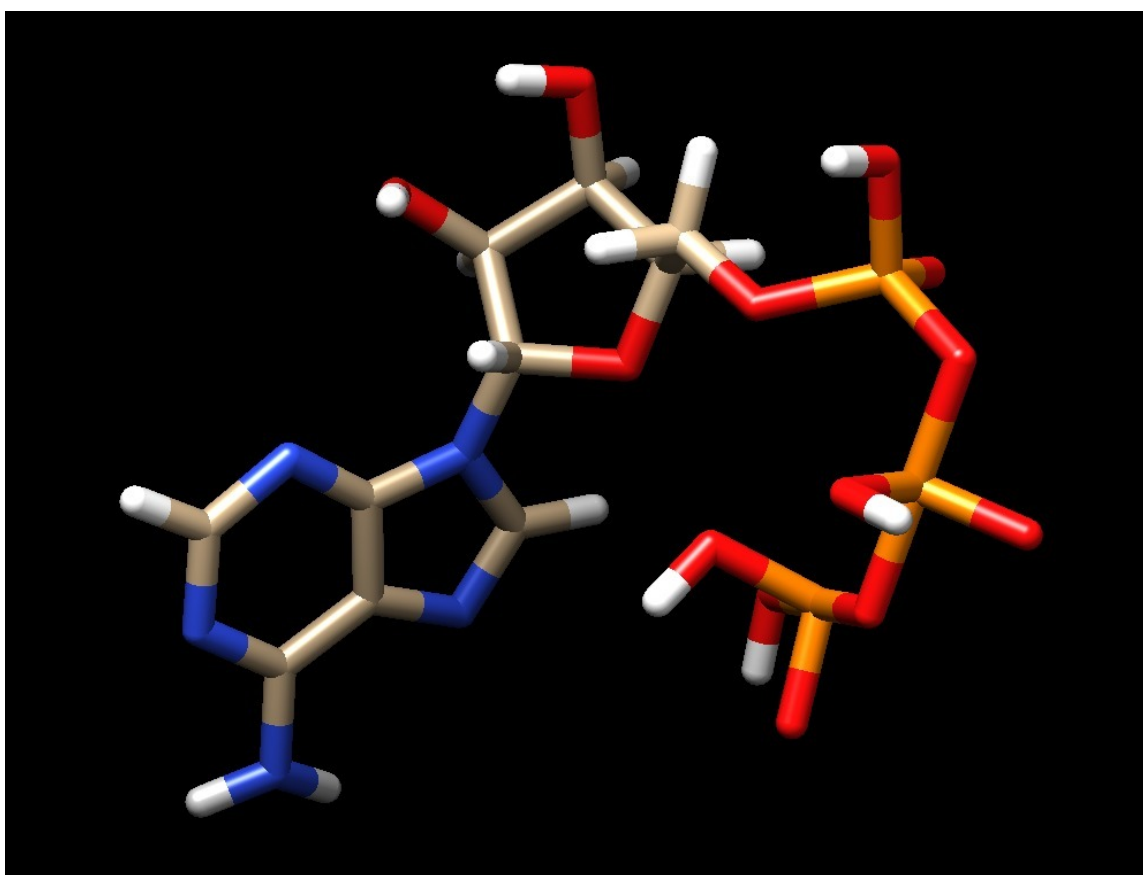


Figura 4. Visualización de la molécula de ATP (ATP.mol2) en UCSF Chimera.

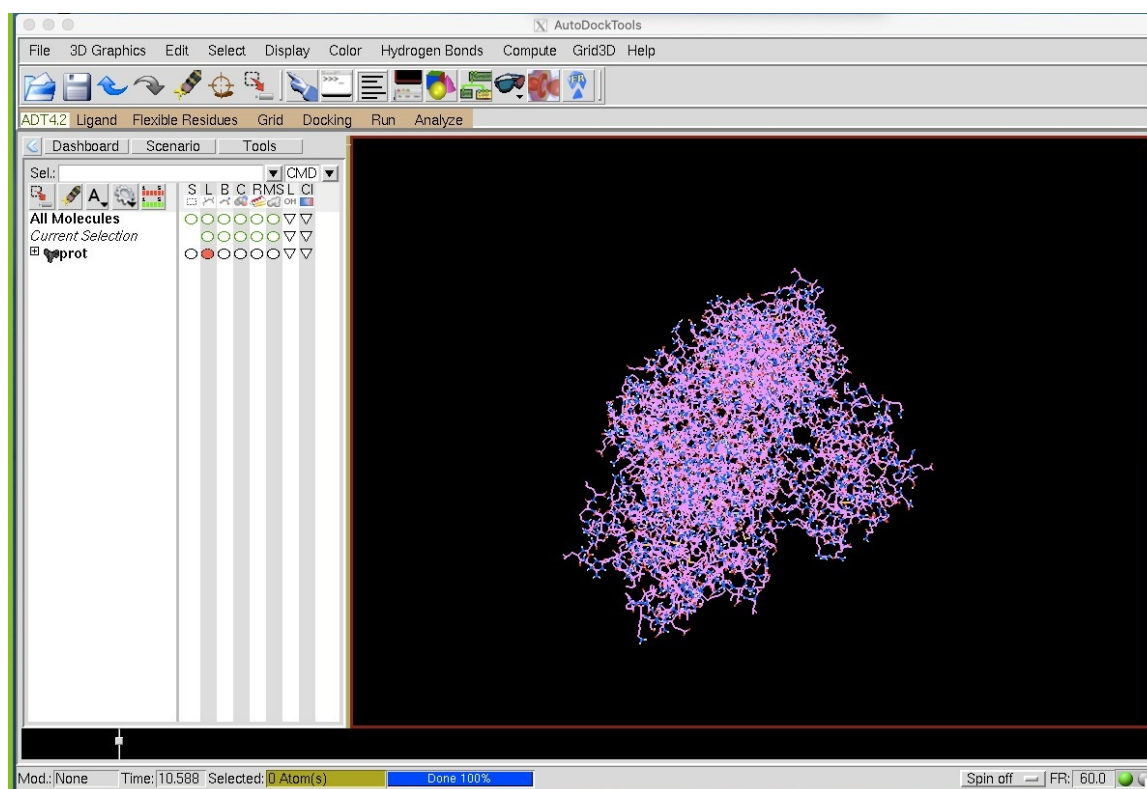


Figura 5. Proteína sin coordenadas ni parámetros añadidos.

Autodock 4.2

Preparamos los archivos de la proteína (prot.pdb) y el ligando (ATP.mol2) en formato PDBQT, los cuales son necesarios para la ejecución de AD4.

Para preparar el ligando, utilizamos las coordenadas del archivo ATP.mol2 para generar un archivo PDBQT que incluye la siguiente información:

- Q: Cargas parciales de Gasteiger.
- T: Tipo de Átomo según el campo de fuerza usado por AD4.

Creamos un archivo de parámetros para el ejecutable autogrid4: el archivo .gpf (grid parameter file). Este gpf fue creado a partir de información del receptor y la(s) molécula ligando, y especifica lo siguiente:

- El espacio de búsqueda con respecto al receptor.
- Centro y tamaño de la rejilla
- La distancia entre los nodos de la rejilla (resolución).
- Los tipos de átomos a usar en los ligandos y el receptor.
- 1.2.2.1.1 Determinar el centro y tamaño de la rejilla

Abrimos AutoDockTools y cargamos el archivo pdbqt de la proteína. Modificamos el tamaño y posición de la caja que aparece en el visualizador. Cambiamos los parámetros hasta posicionar la caja sobre el sitio activo de la proteína:

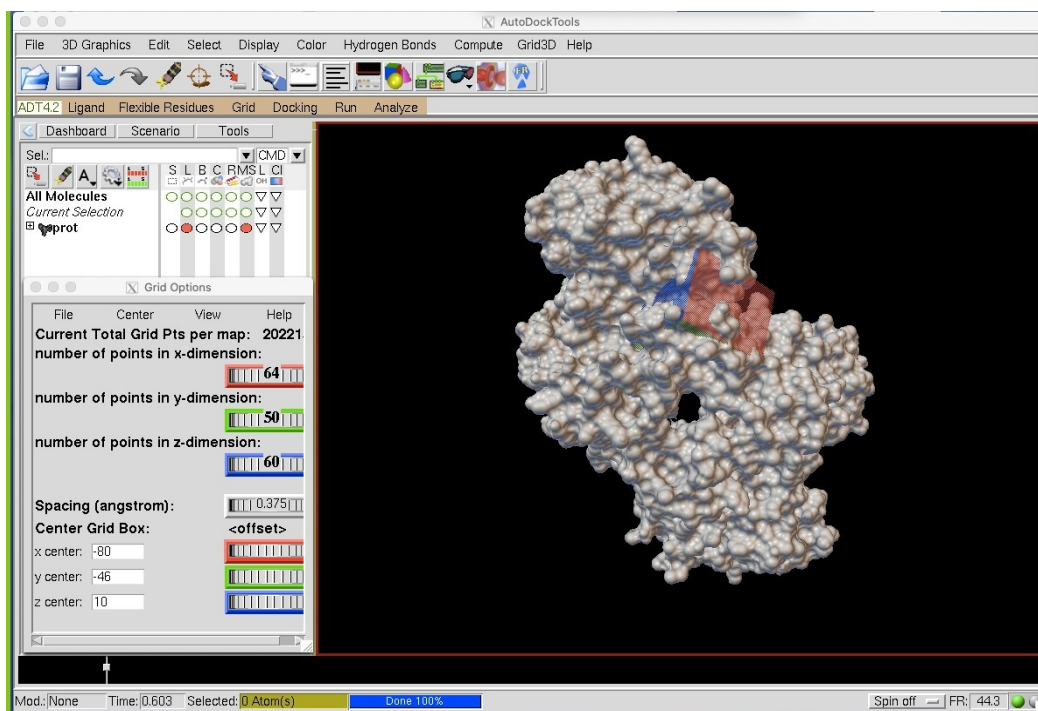


Figura 6. Determinación del centro y tamaño de la rejilla en UCSF Chimera.

¿Qué tipos de átomos contiene la molécula de ATP?

- P, O, C, H, N.

¿Cuántos enlaces rotables tiene?

- 15

¿Hay algún átomo que no posea carga parcial?

- No

Creamos e inspeccionamos el archivo GPF.gpf; ejecutamos AutoGrid y como resultado obtuvimos el archivo prot.maps.fld y los archivos .map, correspondientes a los grid maps de cada tipo de átomo.

Creamos el archivo de parámetros para el ejecutable autodock4: el archivo .gpf (grid parameter file), empleado para llevar a cabo el acoplamiento entre el ligando y el receptor, incluyendo aquellos que usará para el algoritmo de búsqueda. Especificamos además los archivos .map a utilizar para facilitar el cálculo de la función de puntaje.

¿Qué tipo de cargas parciales fueron asignadas a la proteína? ¿Gasteiger o Kollmann?

- Gasteiger.

¿Por qué usar una carga u otra?

- Las cargas de Gasteiger pueden ser analizadas en AutoDockTools, se determinan sobre la base del equilibrio de electronegatividad.

Obtuvimos el archivo de salida DLG.dlg (docking log), archivo que contiene las poses de acoplamiento tras el resultado de la ejecución de docking.

¿Se usará un algoritmo genético o simulated annealing?

- Algoritmo genético

¿Cuál es el número máximo de evaluaciones de la función de scoring?

- 1,000,000

¿Cuántas corridas/ensayos independientes de docking se realizarán?

- 3

¿Qué significa el parámetro rmstol?

- Define cuán diferentes pueden ser dos resultados para estar incluidos en el mismo grupo.

Análisis de los resultados

Inspeccionamos el archivo DLG.dlg y localizamos la información.

¿Qué significa el término torsional degrees of freedom? ¿Cómo afecta a la energía final de interacción?

R = Son los grados de libertad que tiene una molécula para rotar en sus ángulos diedros.

¿Cuántos MODEL hay y a qué corresponde cada uno?

- 3, son repeticiones utilizando diferentes seeds.

¿Cuál es la energía libre de interacción total estimada del primer MODEL?

- +4.48 kcal/mol

¿Es el primer MODEL la pose con la mejor energía?

- Es el que posee la mayor energía.

¿Cuáles son los cuatro términos que autodock usa para estimar dicha energía?

- Energía intermolecular final (vdW + Hbond + desolv energy),, energía interna final total, energía libre de torsión, energía del sistema sin enlace.

¿Cuál es la contribución de las interacciones electrostáticas?

- Confieren polaridad a la molécula.

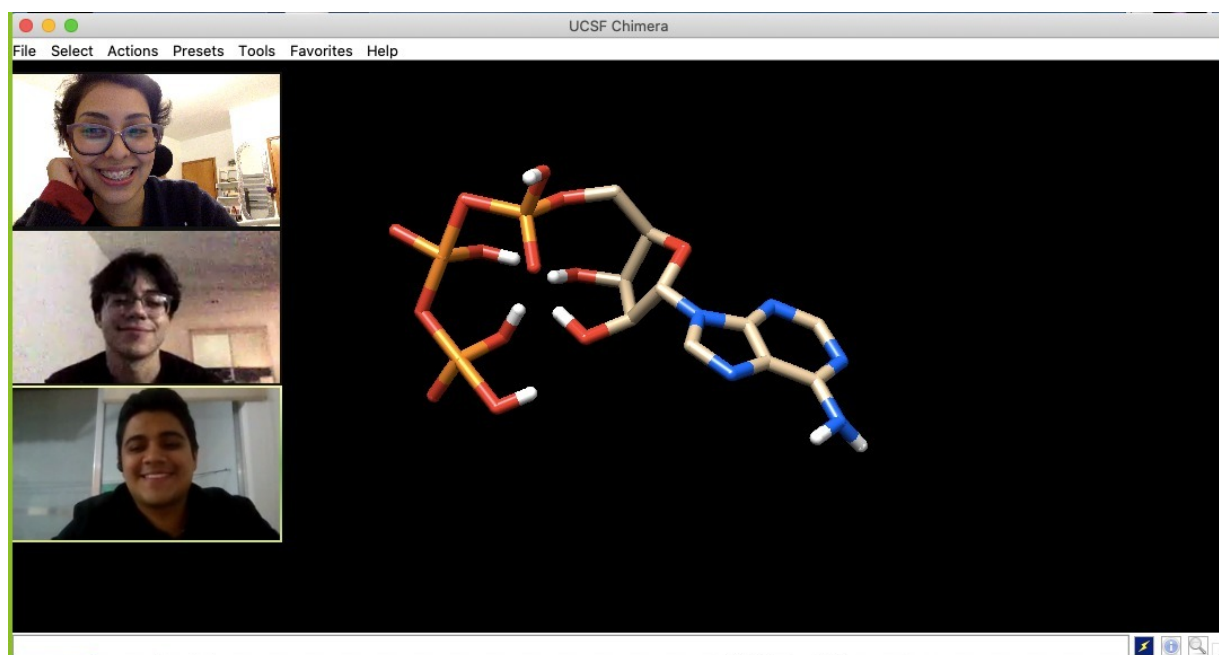


Figura 7. Acoplamiento molecular de ATP.mol2 y equipo de trabajo.

Conclusiones

Desde su primera aparición a mediados de la década de 1970, el Docking Molecular ha representado una herramienta *in silico* sin precedentes para ayudar al diseño y descubrimiento de fármacos. Sin embargo, más allá de las aplicaciones para las que se desarrolló originalmente, el Docking ahora también se emplea ampliamente para ayudar a una variedad de otras tareas de descubrimiento de fármacos, como la identificación de andamios químicos novedosos dentro de grandes bibliotecas de compuestos, para realizar descubrimientos y perfilados *in silico*, reposicionamiento de fármacos, predicción de efectos adversos, entre otras. Al ser una herramienta versátil, el Docking sin duda encontrará aplicación también en otros campos.

Aunque la reciente introducción de técnicas permitió mejorar los resultados de las pruebas de acoplamiento, su aplicabilidad depende en gran medida de, como vimos, la disponibilidad de datos estructurales, químicos o de bioactividad. Por lo tanto, estos enfoques pueden no representar la opción óptima para mejorar el rendimiento de la predicción de acoplamiento cuando se trata de dianas terapéuticas recientemente identificadas o que aún no se han estudiado a fondo.

Basándonos en los resultados de este proyecto, el acoplamiento representa un enfoque valioso también para predecir nuevos horizontes terapéuticos para medicamentos ya aprobados, compuestos naturales y ligandos ya sintetizados, especialmente cuando se usa en conjunto con otros métodos computacionales, como los enfoques de similitud basados en ligandos.

Referencias

- Ricci-López, J. 2021. Taller de simulación molecular: Docking y dinámica molecular. Recuperado el 18/11/2021:
https://jriccil.github.io/Taller_Simulacion_Molecular/index.html.