

TP2: Métodos para la predicción de complejos proteína-droga: docking molecular

Docentes a cargo:

Dr. Juan Pablo Bustamante

Dra. María Cecilia Gómez

Diseño & Descubrimiento de Drogas

1- Introducción

El Docking molecular busca predecir el modo de asociación entre dos moléculas, a nivel atómico. Esta metodología involucra también determinar la posición estructural de los ligandos en relación con su correspondiente receptor, comprender el proceso de reconocimiento para establecer relaciones estructura/actividad/propiedad y generar un sentido crítico que permita hipotetizar, dentro de un conjunto de ligandos, cuál es el que genera las interacciones más fuertes con el receptor.

2- Procedimiento

NOTA 1: Cree una carpeta correspondiente al TP2 y trabaje siempre desde allí.

2.A- Preparación de archivos de docking

2.A.a - Preparación del receptor: utilizaremos el archivo SIRT_close_dock.pdbqt generado en el TP1. Realice una copia de dicho archivo en la carpeta del TP actual. En un editor de texto borre las líneas correspondientes a los átomos de Na agregados para la dinámica molecular ya que Autodock no tiene parametrizados esos iones y no son relevantes para la región donde vamos a hacer el docking.

2.A.b - Generación de los ligandos para estudio: partiremos de los 7 compuestos que obtuvieron el mejor *score* en el cribado final del TP1 y realizaremos modificaciones sobre sus estructuras por medio de técnicas de modelado molecular, buscando maximizar las interacciones entre los compuestos y el receptor.

Con las herramientas que ya maneja (ADT, VMD, Ligplus, etc), analice las diferentes interacciones entre los compuestos seleccionados y la proteína, así como también las interacciones que estabilizan al ligando de la estructura 4I5I. **Vaya anotando en el informe sus conclusiones respecto a las interacciones observadas, incluya además imágenes para explicar dichas conclusiones.** A partir de esto, ¿qué elementos de la proteína puede llegar a aprovechar para estabilizar mejor a los compuestos?. Tómese su tiempo y describa para algunos (2 o 3) compuestos al menos 3 interacciones que, a su criterio, permitan diseñar un buen candidato muy afín (pero no tanto) a su receptor, procure elegir distintos tipos de interacciones. **Una vez concluida esta actividad, y habiendo dejado por escrito este último análisis, continúe con el TP.**

Genere una tabla con los 7 compuestos, siguiendo el modelo a continuación, que le permita comparar si cumplen las reglas de selección para líder. Deberá recurrir a la página de Zinc para obtener los datos solicitados o por búsqueda en internet.

Nombre	Peso molecular	Lipofilicidad (log P)	Donantes HB	Aceptores HB	Enlaces Rotables	Nº de Anillos	Área de Sup. polar
--------	----------------	-----------------------	-------------	--------------	------------------	---------------	--------------------

Descargue también las imágenes en 2D de cada uno de los compuestos y analícelas para detectar al menos un grupo de compuestos similares, si le sirve ayúdese con las interacciones detectadas para los complejos receptor-ligando. **Esquematice al menos el farmacóforo de uno de los grupos detectados.**

A partir del análisis de las estructuras e interacciones de los grupos identificados y de las propiedades fisicoquímicas de cada ligando, seleccione un compuesto representativo y proponga modificaciones de modelado molecular con el objetivo de maximizar las interacciones en un solo compuesto, buscando siempre cumplir de la mejor manera con las reglas de selección para un compuesto líder.

Luego de describir y fundamentar en su informe las modificaciones propuestas, indicando cuáles interacciones esperaba que se produzcan y la naturaleza de las mismas, genere 2 de los compuestos sugeridos utilizando como punto de partida el archivo .pdb del compuesto seleccionado como representativo, modificando en dicha estructura lo que corresponda.

NOTA 2: Cree una carpeta independiente para cada uno de los ligandos.

NOTA 3: Es posible que Gaussview presente inconvenientes con el formato pdb que produjo o babel, si ocurriera esto basta con abrir la estructura en VMD y guardarla reemplazándola sobre el mismo archivo.

Si recuerda la etapa de la preparación de los archivos, se había pedido que se agregaran los hidrógenos polares. En este paso vamos a pedirle a Gaussview que agregue todos los H para poder calcular correctamente la estructura optimizada y las cargas de los átomos. Para tal fin vaya al menú *Edit* → *Atom List* y dentro de la ventana que se le abra haga click en *Edit* → *Add Hydrogens for All Atoms*. En este paso, es vital revisar que la cantidad de hidrógenos y las posiciones en las que se encuentran se correspondan con la estructura que queremos generar. **¿Recuerda por qué motivo esto es muy importante? Déjelo asentado en su informe.** Verifique y ajuste los átomos de H y también los tipos de enlaces entre los átomos (dobles, triples, etc). Debido a que las estructuras obtenidas fueron generadas por nosotros, no tenemos certeza de que las posiciones de los átomos resultantes sean las de menor energía para el compuesto. Por tal motivo, se requiere realizar una serie de minimizaciones de energía (u optimizaciones de geometrías) y verificaciones que adecuen la molécula a una estructura de mínima energía. Para contar con una estructura razonable haga click en *Edit* → *Clean* y verá cómo se van ajustando las posiciones de los átomos. Esta optimización no es un mínimo de energía pero sí una aproximación realizada por el programa Gaussview teniendo en cuenta algunos parámetros de los campos de fuerzas. Cuando finalice leerá en la parte inferior de la pantalla "*Clean completed*". Guardar las estructuras de los nuevos ligandos como compuesto1_AAA.pdb y compuesto2_AAA.pdb, siendo AAA las iniciales de su nombre.

Estudio por mecánica cuántica

Debido a que precisamos obtener luego la descripción electrónica de nuestros compuestos, **¿por qué le parece que esto es así? (déjelo asentado en el informe)**, el siguiente paso consistirá en realizar una optimización energética por métodos *ab initio*.

Haga click derecho sobre el área de trabajo y seleccione *Calculate* → *Gaussian Calculation Setup*. En *Job Type* indique que se va a realizar una optimización. En la pestaña *Method* indique las cargas y multiplicidad que correspondan para sus compuestos. En el método seleccione *DFT* con el funcional *b3lyp* y la base *6-311G*. En la pestaña *Link 0* seleccione la cantidad de memoria que desee ocupar y procesadores y destilde la opción *Full Path* en la línea correspondiente a archivo *checkpoint file*. Por último, en la pestaña *General* tilde *Ignore Symmetry* y destilde *Write Connectivity*, deje los demás parámetros de esa pestaña por defecto. Luego haga clic en *Edit* y *Save*. Seleccione que desea generar un nuevo archivo, póngale de nombre compuestoN.gjf (reemplazar N con el número de compuesto propuesto comenzando por 1) y luego cierre la ventana que se le abrirá. Ejecute desde la consola cada una de las optimizaciones con el siguiente comando:

```
>g09 compuestoN.com compuestoN.log
```

Una vez finalizado el cálculo obtenga la estructura final y realice un cálculo *single point* que calcule las cargas RESP de cada átomo. La línea de comando de Gaussian debe contener lo siguiente:

```
#p b3lyp/6-311g ginput iop(6/7=3) density guess=Read iop(6/50=1) iop(6/33=2) iop(6/42=6)  
pop=mk test units=(ang,deg)
```

Luego del listado de los átomos deberá dejar una línea en blanco (o dos si el archivo trae información del PDB) y colocar un nombre de archivo con extensión .gesp. Corra el archivo con la misma sintaxis que el paso anterior pero con nombres distintos sobre cada compuesto, tanto para la salida de Gaussian, como para el .chk y el archivo .gesp. Cuando termine el cálculo tendrá el archivo .log del cálculo que contendrá una descripción completa de la distribución electrónica de cada uno de los átomos junto con el cálculo del potencial electrostático obtenido por métodos de mecánica cuántica.

Emplearemos el módulo Antechamber para obtener, de la salida de Gaussian, las estructuras y las descripciones electrónicas que precisamos para nuestros programas de docking. Para esto, ejecute el siguiente comando:

```
>antechamber -i compuesto_N.log -fi gout -o compuesto_N.mol2 -fo mol2 -c resp
```

Pasamos ahora a crear los archivos de entrada para el docking. Abra el programa Autodock Tools. Dentro del programa verifique que la opción del menú *File* → *Preferences* → *Set* → *Autodock Tools*: "Automerge NPHS" esté desactivada. Luego vaya al menú *Ligand* → *Input* → *Open* y seleccione uno

de los archivos .mol2. Entienda el cartel que le aparece. Luego vaya nuevamente al menú *Ligand* → *Torsion Tree* y seleccione *Detect root* y verá como se observa un punto verde en el centro de masa de la molécula. Posteriormente vaya al menú *Ligand* → *Torsion Tree* → *Choose Torsions* y verifique que los enlaces marcados son correctos. ¿Se le ocurre en qué caso trivial un enlace rotatable podría ser dejado fijo sin que perdamos conformaciones diferentes? **Comente sus conclusiones en el informe.** Una vez que tenga los enlaces correctos vaya al menú *Ligand* → *Output* → *Save as PDBQT*

2.A.c - Configuración de la grilla

En esta oportunidad sí vamos a precisar crear las grillas de afinidad para cada elemento químico presente en nuestros compuestos. Para tal fin, seguiremos usando el paquete de programas Autodock Tools, con el que configuraremos la grilla y luego con Autogrid las construiremos.

Vaya al programa Autodock Tools y allí al menú *Grid* → *Macromolecule* → *Open* y seleccione el archivo *SIRT_close_dock.pdbqt* del TP1. Preserve las cargas que ya contenía el archivo.

Para configurar la grilla vaya ahora al menú *Grid* → *Grid Box* y en la ventana emergente deberá introducir los parámetros para confeccionar la grilla. Para el centro de la grilla usaremos las coordenadas utilizadas en el docking final del TP1. Elija un espaciamiento de 0.20Å con 100 puntos en cada dirección. Una vez configurado estos parámetros dentro de la ventana *Grid Options* vaya al menú *File* → *Close Saving current*. Teniendo en cuenta lo que fue explicado en la teoría sobre algoritmo, **infiera qué significa cada uno de los parámetros elegidos. Calcule e informe el volumen de la caja utilizada.**

Por último, tenemos que indicarle qué átomos tienen nuestros ligandos para que el programa considere cuáles grillas deberá construir. Haga clic en *Grid* → *Set Map Types* → *Open ligand* (o *Choice Ligand* si actualmente está cargado el ligando en el programa) y seleccione el archivo .pdbqt del compuesto correspondiente.

Una vez configurados todos los elementos que precisa la grilla la guardaremos yendo al menú *Grid* → *Output* → *Save GPF*, nómbrelo como *Grilla.gpf*. Verifique en la carpeta que guardó el archivo que se ha creado el mismo exitosamente. Abra una terminal en esa ubicación y con el siguiente comando generaremos las grillas de cada tipo de átomo empleando el programa autogrid4:

```
>autogrid4 -p Grilla.gpf -l Grilla.glg
```

Luego de unos minutos tendrá en la carpeta los mapas generados. Visualícelos con VMD, para esto, abra cada grilla como un nuevo objeto, seleccionando *Autodock Grid Map* en “*Determine file type*” y luego, en *Graphics* → *Representations*, seleccione “*Isosurface*” y en la opción “*Draw*” seleccione “*Solid Surface*” para visualizarla como una superficie uniforme. **Interprete qué significan; cargue a la vez más de un mapa, e intente extraer conclusiones sobre lo que ve.** Debe ver algo como lo de la Fig. 1.

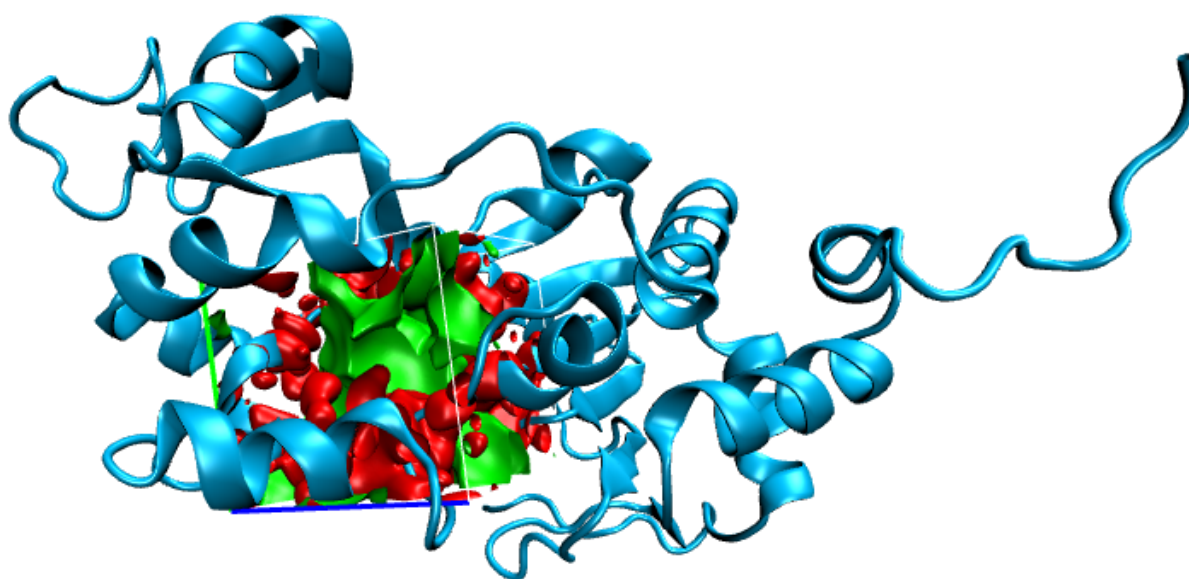


Fig. 1 - Estructura tridimensional de la proteína, en color celeste, y 2 mapas de afinidad, en verde para átomos aromáticos y en rojo para el potencial electroestático.

Debido a que el receptor a utilizar será el mismo para todos los compuestos, puede reutilizar las grillas copiando los archivos generados en las carpetas donde va a realizar los dockings. Tenga presente que si alguno de los compuestos tiene algún elemento químico diferente de los presentes en el compuesto que usó para generar los archivos, va a tener que generar la grilla para ese elemento.

2.B - Ejecución del docking y análisis asociado

Una vez que tenemos los archivos pdbqt del receptor, ligandos y las grillas calculadas, solo nos falta configurar los parámetros del docking en sí.

Abra el programa Autodock Tools. Haga click en el menú *Docking* → *Macromolecule* → *Set Rigid Filename* y abra el archivo pdbqt del receptor SIRT_close_dock.pdbqt. Luego cargaremos el ligando desde el menú *Docking* → *Ligand* → *Open*, se le abrirá una ventana emergente para que seleccione el archivo pdbqt generado anteriormente para cada ligando. En la ventana siguiente deje los valores de los parámetros por defecto.

Ya tenemos cargados nuestros archivos de entrada, ahora pasaremos a configurar los parámetros del experimento.

Realizaremos un estudio de búsqueda basado en algoritmos genéticos. Desde el menú *Docking* → *Search Parameters* → *Genetic Algorithm* podrá visualizar las diferentes opciones que pueden configurarse para este tipo de búsqueda. **Busque y discuta con los docentes y compañeros qué significan cada uno de ellos.** Luego seleccione un tamaño de población de 1500 individuos y en *Maximum Numbers of top individuals that automatically survive* en 10, el resto de los parámetros déjelos por defecto.

Además de los parámetros en el algoritmo de búsqueda hay otras configuraciones que pueden hacerse sobre el proceso de docking, como los movimientos de rotación y traslación de

los átomos, o los criterios para evaluar y presentar los resultados finales, entre otros. Explore las opciones desde el menú *Docking* → *Docking Parameters*.

Para generar el archivo que precisamos para correr el experimento vaya al menú *Docking* → *Output* → *Lamarckian GA (4.2)*. Guárdelo como *Dock_compuesto1.dpf*.

Correremos el docking con el programa autodock mediante el siguiente comando:

```
> autodock4 -p Dock_compuesto1.dpf -l Dock_compuesto1.dlg
```

Repita el procedimiento para todos los compuestos propuestos. Analice la salida obtenida primeramente con un editor de texto, identificando las distintas partes que forman el archivo. **Encuentre dentro de dichos archivos los histogramas y clusters de confórmeros. ¿Obtuvo algún resultado de energía de unión positiva?** De ser así anótelos para luego analizarlos con más detenimiento.

Abra nuevamente Autodock Tools y haga clic en el menú *Analyze* → *Dockings* → *Open*. Seleccione allí alguno de los archivos de salida del docking (extensión *dlg*). Luego cargue la estructura del receptor dentro del menú *Analyze* → *Macromolecule* → *Open* y por último, para ver las diferentes conformaciones, desplácese en los menús hacia *Analyze* → *Conformations* → *Play*. Pruebe cambiar en el listado de moléculas de la izquierda la representación para visualizar el ligando y luego utilizando la ventana emergente con los botones de control, vaya observando las diferentes conformaciones que adoptó su ligando en la cavidad. Puede encontrar más opciones



de visualización dentro de las opciones accediendo al ícono . Pida al programa que muestre los puentes de hidrógeno.

¿Identifica alguna relación entre los puentes de hidrógeno que reconoció el programa y los valores energéticos hallados para cada complejo receptor-ligando?

Guarde cada complejo en un archivo independiente. **Visualice las interacciones con LigPlus y extraiga conclusiones.**

¿Cómo fueron los resultados de las interacciones entre los grupos añadidos a los compuestos y el receptor? ¿Se cumplieron sus predicciones? ¿Recuerda el objetivo de proponer estos nuevos compuestos, cree que se cumplió al menos para la mayoría de los compuestos? ¿Cada uno de los nuevos ligandos parece tener mejor o peor afinidad por el receptor respecto al ligando original del cual partió en cada caso? ¿Qué conclusiones puede sacar de las posiciones que adoptaron los confórmeros con energía de unión positiva, si es que obtuvo alguno?. Por último realice el docking utilizando autodock vina y comente las similitudes/diferencias entre los resultados de autodock y vina.

Hasta aquí llegan los estudios de docking sobre una estructura del receptor completamente rígida. Pero... sabemos que esto es solo una aproximación, y que en realidad las proteínas se mueven, y presentan regiones con distintos grados de flexibilidad y, que además, en muchos casos “se acomodan” al ligando, reforzando interacciones, creando y eliminando algunas otras. Sabemos también, del cuatrimestre pasado, que las moléculas de agua juegan un rol preponderante en muchos complejos receptor-ligandos. Con todo esto, se imaginarán que el paso siguiente es hacer un estudio de dinámica molecular de los complejos, para evaluar este tipo de cuestiones y analizarlos desde un punto de vista dinámico. Dejemos eso para el TP3