

Taller 1: Campos de fuerza y Schrödinger

Sebastián Franco Ulloa y Gian Pietro Miscione

Módulo de Mecánica Molecular

Septiembre 9, 2017

Este taller tiene como objetivo familiarizar al estudiante con las herramientas más importantes distribuidas por Schrödinger para realizar cálculos de *docking*, *single-point energy* y búsquedas conformacionales. Entre estas herramientas se encuentran *PrepWizard*, *LigPrep*, *Glide* y *Macromodel*. Más aun, se pretende esclarecer las nociones de cargas parciales, energía potencial interna, términos energéticos enlazantes y no-enlazantes, y coordenadas internas. Los estudiantes deberán manipular varias de las opciones ofrecidas por los programas y así entender a fondo su funcionamiento. A lo largo del taller, los estudiantes estudiarán un sistema biológico de una ADN girasa la cual está involucrada en el proceso de replicación del ADN bacteriano, por lo que su inhibición puede resultar en la muerte del microorganismo. Por ende, es un receptor llamativo para el diseño de medicamentos.

1 Visualización y modificación de moléculas en Maestro

Lo primero que haremos será entender la lógica detrás de la interfaz gráfica de Schrödinger también conocida como Maestro. Este programa se puede descargar de forma gratuita para cualquier sistema operativo desde el sitio web de Schrödinger, Inc. En esta sección daremos un recorrido sobre algunos aspectos importantes de Maestro que nos permiten visualizar y modificar moléculas.

- Descargue del repositorio del curso el archivo *aspirina.mol2*.
- Abra una ventana en la terminal (*Ctrl + Alt + T*) y abra el archivo descargado con su editor de texto de preferencia (e.g. *gedit aspirina.mol2*) ¿A qué corresponde cada columna de este archivo?
- Abra Maestro. Si está trabajando en una máquina con Linux puede escribir en la terminal “Maestro” o “\$SCHRODINGER/maestro”. Si está trabajando en una máquina con Windows o Macintosh, dele doble click al ícono de Maestro que se generó tras su instalación.
- A través de *File → Import Structures*, importe el archivo descargado.
- Identifique la carga total del sistema (se muestra en el panel inferior de Maestro).
- El sistema en cuestión corresponde a la forma protonada de la aspirina pero nos interesa generar el estado metilado y etilado de la misma.
- Hay varias formas de modificar estructuras en Maestro. Por ejemplo, para generar la estructura metilada, debe duplicar la estructura protonada (*Click derecho (sobre el nombre de la estructura en el panel izquierdo) → Duplicate → In Place*). Oprima la tecla *A* para asegurarse que la selección esté en modo “átomo” y seleccione el protón que se desea modificar por un grupo metilo. Ahora abra el *3D-Builder* de Maestro (*Edit → 3D Builder*) y sustituya el átomo de hidrógeno por uno de carbono. Maestro saturará automáticamente este sustituyente.
- Para generar la estructura etilada duplique nuevamente la forma protonada, vuelva a seleccionar el protón y a abrir el *3D-Builder*. Esta vez seleccione la opción *Add Fragments → ... → Diverse Fragments* y seleccione *ethyl*.

Estas son solo dos de las formas con las que se pueden agregar o remover secciones de una estructura. Ahora veremos rápidamente cómo se pueden cambiar las coordenadas internas de una molécula para, por ejemplo, generar una geometría inicial para un cálculo de cualquier índole.

- Por ahora trabajaremos con la estructura protonada de la aspirina.
- Presione la tecla *A* para seleccionar únicamente átomos.
- Para variar el largo de un enlace: *Selecione dos átomos enlazados (para seleccionar múltiples átomos mantenga presionada la tecla Ctrl) → Click derecho → Adjust Distance.*
- Para variar un ángulo: *Selecione 3 átomos → Click derecho → Adjust Angle.*
- Para variar un ángulo dihedro: *Selecione 2 átomos → Click derecho → Rotate Dihedral.*
- Otra forma de acceder a estos menús es abriendo el *3D-Builder* y seleccionar la flecha a la derecha de la opción *Move*.

Respecto a la visualización de las moléculas, todo lo que concierne a la representación, textura y color de los átomos se puede manejar desde el panel **Style** localizado en la parte superior de Maestro. La mejor forma de representar moléculas es muy personal y depende de lo que cada persona quiera ver o esté buscando. Desde este panel se pueden hacer cosas como cambiar el color de los átomos, cambiar su modo de representación (van der Waals, *ball-and-stick*, etc.), y mostrar u ocultar átomos, superficies y estructuras secundarias de moléculas biológicas.

2 Preparación del receptor

Ya que se ha adquirido un poco más de confianza con Maestro, podemos proceder a usarlo para hacer cálculos más complicados como uno de *docking*. Primero elimine las entradas que se crearon en la sección 1, es decir, la aspirina y su forma metilada y etilada. Para esto haga click derecho en los nombres de las moléculas en el panel izquierdo de Maestro y seleccione “Delete”.

El primer paso para un cálculo de *docking* es la preparación del receptor, el cual, al venir de la base de datos PDB, contiene varios problemas discutidos en clase con anterioridad.

- Para descargar el archivo **.pdb* correspondiente al receptor de interés, se debe ir a la página principal de la base de datos PDB ([link aquí](#)) y en el motor de búsqueda escribir “DNA gyrase”.
- Para filtrar más los resultados, en el costado izquierdo puede seleccionar en la sección de **ORGANISM**, la bacteria *Staphylococcus Aureus*.
- Buscar y seleccionar la entrada con código **4PLB**. En esta página puede encontrar toda la información pertinente del cristal.
- En el recuadro **Download Files**, seleccione **PDB Format** para descargar el archivo *4PLB.pdb*.
- Abrir con *gedit* el archivo *4PLB.pdb* e identifique los residuos y átomos faltantes al igual que algunos de los residuos con más de una conformación (Ayuda: Vaya a la línea 4832 para ver cómo se exhiben las conformaciones alternas de un residuo).
- Ir a Maestro y, al igual que en la sección anterior, importar la estructura que se descargó. Aunque depende de la configuración y la versión de Maestro, en verde se debería estar mostrando nuestro ligando de interés.
- Jugar con la visualización de la proteína con las opciones del recuadro **Style** localizado en la parte superior de Maestro. Debe identificar cuántas cadenas, moléculas, residuos, átomos y cargas tiene el sistema. También debe localizar la doble hebra de ADN y la posición del ligando.

- Seleccionar en la esquina superior izquierda la opción **Protein Preparation Wizard**. Esta es la interfaz para la preparación del receptor.
- Nótese que en el recuadro **Import structure into Workspace** se puede descargar el archivo **.pdb* escribiendo el código en el recuadro de texto. En este caso esto no es necesario.
- En el recuadro **Preprocess the Workspace structure** deje todas las opciones predeterminadas y además seleccione las opciones “Fill in missing side chains using Prime”, “Fill in missing loops using Prime” y en “Delete waters beyond # Å from het groups” escriba 0.0. En este taller no se considerarán aguas por simplicidad. Haga click en **Preprocess**.
- Luego puede ver, en **View Problems**, los problemas que tuvo Maestro al preparar el receptor. Por ejemplo, átomos faltantes, átomos a los que no se pudo asignar un *atom type* y residuos con conformaciones alternas (¿hay algún residuo problemático cerca al ligando? ¿Las conformaciones diferentes corresponden a interacciones químicas diferentes?).
- En el recuadro **Review and Modify**, tras hacer click en la opción **Analyze Workspace**, puede ver las cadenas, moléculas no convencionales (*het*), aguas, etc,...
- En el recuadro **Refine**, se minimizan los hidrógenos los cuales se pudieron haber puesto mal en los pasos anteriores.
- Seleccione la opción “Minimize Hydrogens of Altered Species” y luego haga click en **Optimize**.
- Podemos ignorar la sección **Remove waters**, pues ya las eliminamos todas.
- En la última sección **Restraigned minimization**, coloque un valor de RMSD igual a 0.30 y haga click en **Minimize**. Esta minimización va a relajar el sistema evitando que hayan átomos muy cercanos entre sí.
- Note que en cada uno de los pasos descritos anteriormente se crea una nueva entrada que se importa automáticamente a Maestro, entonces, la última entrada corresponde al receptor preparado. Para no confundirlas es útil cambiarles el nombre o eliminar las entradas generadas en pasos intermedios.
- Una buena idea para tener una noción de las interacciones presentes entre el ligando nativo y la proteína es hacer un mapa de interacciones. Maestro es capaz de hacer esto por nosotros. Seleccione la estructura preparada y seleccione en la esquina superior izquierda la opción **Ligand Interaction Diagram**.
- Para *Glide* en particular no es necesario eliminar el ligando nativo del cristal para hacer el *docking*. El mismo entiende quién es el ligando nativo y lo ignora.
- Como estamos ignorando el problema de las conformaciones alternas, Maestro, por defecto, va a usar las conformaciones ‘A’.

3 Generación de la cuadrícula

Procederemos con la generación de la cuadrícula la cual se hace con la herramienta *Glide* (la misma que hará el *docking*). La cuadrícula es un único archivo que contendrá toda la información del receptor necesaria para hacer el *docking*; esto incluye geometría, conectividad, potenciales, restricciones, etc.

- Selecciona e incluya la estructura del receptor preparado. OJO: Maestro hace un gran énfasis en la diferencia entre “Selección” e “Inclusión”. La entrada Incluida es aquella cuya geometría se muestra en el área de trabajo, mientras que la entrada Seleccionada es la que se pone azul cuando se le hace click en el costado izquierdo de Maestro. Una entrada puede estar incluida pero no seleccionada y viceversa.
- Busque en *tasks* la opción “Receptor Grid Generation” y seleccione la herramienta que hace mención a *Glide*.

- Para excluir el ligando nativo del docking se le debe decir a *Glide*. Para esto, en el recuadro **receptor**, seleccione “Pick to identify the ligand” y seleccione un átomo del ligando nativo.
- Seleccione la opción “Use input partial charges” para que use las cargas de los *atom types* que se asignaron durante la preparación del receptor. Si no se hiciera esto, *Glide* volvería a asignar *atom types*.
- En la parte inferior del recuadro de **Advanced settings** se puede permitir que los átomos del receptor formen puentes H con hidrógenos aromáticos, halógenos como receptores y halógenos como aceptores. En este caso ninguna de las opciones es particularmente importante.
- En el recuadro **Site** se deben especificar los parámetros de la caja interna y externa. Ya que tenemos un ligando cocrystalizado la mejor opción es centrar las cajas en el “Centroid of Workspace ligand”. En el recuadro de **Advanced settings** se pueden personalizar las dimensiones de las cajas pero mantendremos las predeterminadas.
- En el recuadro de **Constraints** vamos a darle la posibilidad a *Glide* de fijar algunas restricciones (esto no lo obligará en el paso del *docking*). Vaya a la pestaña de “H-bond/Metal” y luego seleccione los 2 átomos de oxígeno (del grupo carboxilato) presentes en los 2 residuos Asp1083 (Ayuda: están interactuando con el ligando nativo).
- En este ejemplo no vamos a darle libre rotación a ninguno de los dihedros que nos ofrece *Glide* ni excluirémos ningún volumen. No obstante esto se puede hacer en las pestañas de **Rotatable Groups** y **Excluded Volumes**. OJO: No olvide que imponer restricciones al *docking* puede resultar en el sesgo de los resultados. En este caso es posible imponer la formación de este puente H ya que se está dando entre dos grupos cargados (un nitrógeno con carga +1 en el ligando y los oxígenos del aspartato). Debido a la forma en la que se diseñó *Glide*, el programa considera como más favorables los puentes H entre grupos neutros que entre grupos cargados, lo cual va en contra de la teoría de la química orgánica. Imponer este puente H sesga nuestros resultados pero restaura el sentido físico del sistema.
- Haga click en **Run**. Este cálculo debe tomar aproximadamente 3 minutos y resulta en un único archivo *.zip.

4 Preparación de un ligando

Hasta este punto terminamos de lidiar con el receptor, pero aun falta preparar el ligando, lo cual se hará con la herramienta *LigPrep* de Schrödinger. En este caso haremos el *docking* del ligando nativo y compararemos la pose predicha con la original del cristal.

- Seleccione e incluya en el espacio de trabajo la estructura “4PLB” preparada. Haga un duplicado de la estructura haciendo *Click derecho* → *Duplicate* → *In place*.
- Ahora seleccione solo el ligando haciendo click en la letra **L** presente en la parte superior izquierda de Maestro en el recuadro “Quick Select”. Luego seleccione la opción “Invert” y suprima para borrar todo excepto el ligando.
- Busque en *Tasks* la herramienta *LigPrep*. Esto debería abrir una ventana con una única pestaña.
- Lo primero que se debe especificar es el (o los) ligando a preparar. El ligando a preparar puede estar en un archivo que no se haya importado a Maestro, o puede estar presente en el proyecto actual. Por simplicidad, haremos lo segundo.
- Asegurese de que el ligando esté incluido en el área de trabajo y en la opción “Use structures from:” seleccione “Workspace (# included entries)”.
- Al hacer *docking* de un solo compuesto, no es necesario hacer un filtro, por lo que nos saltaremos esta opción.

- En este ejercicio generaremos todos los posibles estados de ionización en un pH de 7.0 ± 2.0 usando Epik (otra distribución de Schödinger).
- En este ejemplo no es necesario pero es buena costumbre siempre seleccionar la opción de “Desalt”.
- Seleccione la opción de “Generate Tautomers”.
- Con respecto a la estereoquímica, generaremos todas las posibles combinaciones (lo cual no es común), estableciendo el máximo de estructuras a 32.
- Este cálculo generará un único archivo con todas las moléculas ya listas para hacerles el *docking*. Este archivo puede ser *.maegz o *.sdf. Ambos formatos pueden almacenar la misma información pero el primero es un formato binario que solo puede leer Maestro mientras que el segundo corresponde a un archivo de texto que se puede abrir con la mayoría de programas en química computacional. Por estas razones se acostumbra a exportarlo como *.sdf.
- Haga click en **Run**. El resultado, aunque corresponda a un único compuesto, contiene varias moléculas. Analice las diferencias.

5 Docking

Una vez se han preparado el receptor y el ligando, podemos proceder a hacer la inserción de uno en el otro. Esto se hará con la misma herramienta con la que se generó la cuadrícula: *Glide*.

- Para hacer el *docking* debemos abrir la ventana correspondiente para lo cual debe buscar en *Tasks* la herramienta *Ligand Docking*. Al hacer click debe abrir una ventana con 6 pestañas.
- Lo primero que se debe especificar es la cuadrícula que se usará para hacer el *docking*. Esto se hace en la parte superior de la ventana donde dice **Receptor Grid**. Busque en el navegador el archivo *.zip generado en la sección 3.
- En la pestaña **Ligands** se deben elegir las moléculas que se insertarán en el receptor. Estos ligandos se pueden especificar como un archivo, como entradas seleccionadas y como entradas incluidas. Seleccione la opción de su preferencia pero sea consistente. e.g. si elige “Workspace (Included Entries)” asegúrese de que todas las moléculas se muestren en el espacio de trabajo.
- Dependiendo de los recursos computacionales disponibles, es posible que se quiera no evaluar todos los ligandos disponibles sino solo los primeros *N*. Seleccionar la casilla *End* indica que se evaluarán todas las moléculas del archivo.
- Seleccione la casilla *Use input partial charges* por razones discutidas anteriormente. Mantenga el resto de opciones en esta pestaña con sus valores predeterminados.
- En la pestaña de **Setting** se debe especificar todo lo concerniente al *docking*. La casilla *Precision* es donde se especifica la función de *scoring* a usar. En este caso usaremos *SP*.
- En la casilla *Ligand Sampling* se especifica si se desea hacer *docking* con el ligando flexible o rígido. En este caso lo dejaremos flexible.
- El resto de opciones las dejaremos en sus valores predeterminados pero asegúrese de marcar con un chulo las opciones “Add Epik state penalties to docking score”, “Reward intramolecular H-bonds” y “Enhance planarity of conjugated pi groups”
- En la parte inferior del recuadro de **Advanced Settings** se puede especificar si se desea que los átomos de los ligandos puedan participar en puentes H con hidrógenos aromáticos, halógenos como aceptores o halógenos como donores.

- En la pestaña de **Constraints** se mostrarán las restricciones que se definieron durante la generación de la cuadrícula. Usted deberá hacer el cálculo de *docking* 2 veces. La primera vez sin ninguna restricción y la segunda vez exigiendo un puente de hidrógeno con al menos uno de los 4 átomos de oxígeno seleccionados en la sección 3. Compare los resultados.
- En la pestaña **Output** se especifican los aspectos técnicos de lo que retornará el cálculo. En este caso mantendremos casi todas las opciones con sus valores predeterminados.
- En la opción “Write at most: # poses per ligand”. Cambie el número a 3. Esto hará que el programa imprima las 3 mejores poses del ligando en el receptor (para cada molécula).
- Seleccione la opción “Write per-residue interaction scores” y pídalo para residuos hasta 5Å del centro de la cuadrícula. Con esto *Glide* imprimirá la energía de interacción con cada uno de los residuos que satisfaga esta condición.
- Haga click en *Run*. El archivo más importante que se genera con este tipo de cálculos es un archivo **_pv.maegz*.
- Importe (si no se hizo automáticamente) el archivo mencionado anteriormente. La primera entrada corresponde al receptor y las demás son las moléculas en sus respectivas poses.
- Para mantener el receptor fijo dele *click derecho* → *Show in Workspace* → *Fix within project*. Luego para moverse entre las moléculas, solo inclúyalas en el espacio de trabajo. Así podrá ver más rápidamente todos los complejos predichos.
- Para ver propiedades como las interacciones por residuos, puede abrir la tabla del proyecto (símbolo # en la parte superior derecha de Maestro). Ahí se mostrarán todas las propiedades de todas las entradas del proyecto. Dentro de estas propiedades se encuentran dos tituladas “Glide Score” y “Docking score”. Estas dos columnas corresponden al valor que toma la función de *scoring* cuando se evalúa en esa molécula para esa pose particular. La diferencia entre estas dos columnas es que “Docking score” incluye un término calculado por *Epik* (“Epik State Penalties”) que da cuetas de qué tan probable es encontrar la molécula en cuestión a condiciones fisiológicas. En la tabla también encontrará las contribuciones energéticas de cada uno de los residuos especificados.

6 *Single-Point energy* y búsqueda conformacional

Es posible que una molécula tenga una muy buena energía libre de enlace (ΔG_{bind}) pero que no sea un buen inhibidor. Esto puede ocurrir cuando el ligando requiere de demasiada energía para llegar a la pose en la que interactúa propiamente con el receptor. En otras palabras, si la energía del mínimo global es mucho más baja que la energía de la molécula en la pose de inhibición, su actividad puede ser menor que la predicha. Para verificar esto haremos dos cálculos en los que hallaremos la energía potencial interna de una molécula. En uno de los cálculos lo haremos para la pose predicha y el otro para la conformación de su mínimo global.

- Identifique la molécula con mejor energía libre de unión (*Docking score* más bajo) e inclúyala en el área de trabajo. Primero calcularemos la energía potencial interna de esta molécula en esta pose.
- Para abrir el calculador de energía MM: *Tasks* → *Current Energy*. Esta interfaz es una dependencia de *Macromodel*.
- Seleccione *None* como modelo de solvente y mantenga una constante dieléctrica de 1.0.
- En la casilla *ECalc* seleccione la opción de *Complete*. No es necesario modificar las demás opciones para este cálculo en particular.
- Asigne un nombre al cálculo y seleccione **Run**. El cálculo no debe tardar más de 1 minuto.

- Se debieron generar 6 archivos. Los 3 a continuación corresponden a archivos que Maestro escribe para poder correr el cálculo pero no tienen información nueva: *.com con los parámetros del cálculo, *.mae con las estructuras y propiedades del sistema y *.sbc con los grados de libertad del sistema. Además, cuando el cálculo termina, debió generar 3 nuevos archivos: *.log con la información de cómo corrió el cálculo, *.maegz con la estructura (coordenadas, conectividad y propiedades) de la geometría final (en este caso la misma inicial) y .mmo con la energía potencial interna de la molécula descompuesta en los diferentes términos del campo de fuerza.
- Abra el archivo *.maegz si no se abrió ya automáticamente.
- En la esquina superior derecha seleccione el símbolo # para abrir la tabla del proyecto. Ahí encontrará todas las propiedades del sistema incluyendo la columna *Potential Energy-OPLS-2005* la cual contiene la energía potencial interna de la molécula en *kJ/mol*. El valor absoluto de esta energía no es muy diciente pues depende en gran parte del número de átomos, enlaces, ángulos, etc. por eso necesitamos otra energía con la cual compararla.

Ahora compararemos esta energía con la energía del mínimo global. Para esto vaya a *Tasks* → *Conformational Search (Macromodel)*.

- Deje casi todas las opciones con sus valores predeterminados.
- En la pestaña de **CSearch** se especifica el algoritmo con el que se hará la búsqueda conformacional. En la opción *Method* seleccione “Torsional Sampling (MCM)”.
- Asegúrese de seleccionar e incluir únicamente la molécula cuya búsqueda se desea hacer.
- En la opción *Number of structures to save for each search*, el 0 significa guardar todas las estructuras no redundantes. Esto puede resultar en cientos de geometrías para una misma molécula. Cambie este número a 5 para que el programa solo imprima las 5 geometrías con menor energía. Aquella con la menor energía corresponde al mínimo global.
- Haga click en **Run**.
- Compare las geometrías y los valores de energía potencial interna cuando la molécula se encuentra en su estado ligado versus cuando está en su mínimo global. ¿Cuál es la diferencia de energía entre la pose ligada y el mínimo global?

7 Virtual Screening

En los cálculos anteriores se estudió una única molécula. Si tuvieramos interés en encontrar nuevos compuestos potencialmente activos sería necesario estudiar una mayor variedad de compuestos. En esta sección usaremos la cuadrícula generada en el paso 3 pero haremos el *docking* de compuestos con estructuras disponibles en línea pero cuyas actividades desconocemos.

- Los ligandos los descargaremos de la base de datos ZINC (<http://zinc.docking.org/>). Vaya a la base de datos y haga click en las siguientes opciones:

Subsets → *Catalog* → *Biosynth* → *Downloads* → *All (SDF, Mid pH 6-8)*

Así debe descargarse un archivo *.sdf.gz.

- Importe el archivo que acaba de descargar en Maestro (Maestro admite formatos comprimidos). En total debe haber abierto 1609 compuestos.
- Por motivos de tiempo, para disminuir el número de ligandos, quédese con los primeros 200 que se importaron y elimine los demás.

- Prepare los ligandos similar a como se hizo en la sección 4. Esta vez filtre los compuestos para quedarse únicamente con aquellos que tengan un peso molecular entre 250 y 500g/mol y que tengan una carga total entre $\{-2, 2\}$. Además seleccione la opción “Retain specified chiralities (vary other chiral centers)”.
- Haga el *docking* de la misma manera en la que se hizo en la seccion 5 sin imponer restricciones.
- ¿Cuál es el compuesto más tentativo para decir que va a ser un buen inhibidor? ¿Qué le dice su instinto químico? ¿Se ven las interacciones que se ven con el ligando nativo?