

# **Manual de acoplamiento molecular (*docking*): teoría y ejercicios**

**Constanza Duran Becerra**

**Juan Francisco Cortés Benítez**

**Juan Francisco Palacios Espinosa**



## Índice

1	Modelado molecular .....	5
1.1	Cálculos de campos de fuerza .....	5
1.2	Métodos ab initio .....	6
1.3	Métodos semiempíricos .....	6
2	¿Qué es el acoplamiento molecular ( <i>docking</i> )?.....	6
3	Antecedentes del <i>docking</i> .....	7
4	Algunas aplicaciones .....	7
5	Flexibilidad en el <i>docking</i> .....	8
5.1	Docking rígido.....	8
5.2	Docking flexible .....	9
6	Algoritmos.....	9
6.1	Algoritmos sistemáticos .....	10
6.2	Algoritmos aleatorios o estocásticos .....	10
6.3	Métodos de simulación o determinísticos.....	12
7	Funciones de evaluación o <i>score</i> .....	12
8	Metodología del <i>docking</i> .....	13
9	Softwares .....	14
10	Antes de comenzar... .....	16
11	Cómo preparar un ligando.....	17
11.1	Construirlo desde cero .....	17
11.2	Cómo minimizar un ligando.....	21
11.3	Con una base de datos .....	24
12	Cómo preparar una proteína .....	30
12.1	Con RCSB Database .....	30
12.2	Con PDB Loader Service de PyMOL. ....	31
12.3	Con comandos en PyMOL.....	32
12.4	Cómo minimizar una proteína .....	34
13	<i>Docking</i> en Autodock4.2.....	42
13.1	¿Qué necesitamos? .....	42
13.2	Visualización de aminoácidos del sitio de unión.....	42
13.3	Preparación en ADT.....	43

13.4	Generar el archivo *.gpf .....	46
13.5	Generar el archivo *.dpf .....	51
13.6	Para ejecutar el cálculo .....	55
13.7	Generar el archivo *.glg.....	56
13.8	Generar el archivo *.dlg .....	57
13.9	Monitorear avance del docking .....	58
13.10	Visualización de resultados .....	60
13.11	Guardar mejor conformación .....	62
14	<i>Docking</i> en Autodock VINA.....	65
14.1	¿Qué necesitamos? .....	65
14.2	*.pdbqt del ligando.....	65
14.3	*.pdbqt de la proteína .....	65
14.4	Conf.txt.....	65
14.5	Para ejecutar el cálculo en VINA.....	67
14.6	Visualización de resultados .....	71
14.7	Guardar mejor conformación .....	73
15	<i>Docking</i> en GOLD.....	74
15.1	Scores de GOLD .....	74
15.2	¿Qué necesitamos? .....	74
15.3	Verificar ligandos.....	75
15.4	Nombrar ligandos.....	77
15.5	Para ejecutar el cálculo .....	77
15.6	Visualización de resultados .....	87
15.7	Guardar mejor conformación .....	89
16	Cómo calcular RMSD .....	90
16.1	Teoría .....	90
16.2	Ejemplo .....	90
17	Ejercicios .....	97
17.1	Ejercicio 1 .....	98
17.2	Ejercicio 2 .....	100
17.3	Ejercicio 3 .....	102
17.4	Ejercicio 4 .....	105
17.5	Ejercicio 5 .....	108

18	Referencias .....	113
19	Respuestas .....	115
19.1	Ejercicio 1 .....	115
19.2	Ejercicio 2 .....	116
19.3	Ejercicio 3 .....	116
19.4	Ejercicio 4 .....	118
19.5	Ejercicio 5 .....	119
20	Anexo 1. Instalación de programas .....	122
20.1	Avogadro .....	122
20.2	Autodock Tools (MGL Tools).....	123
20.3	YASARA View .....	125
20.4	PyMOL .....	126
20.5	BIOVIA Discovery studio visualizer.....	128
20.6	Ejecutables de Autodock 4.2 .....	130
20.7	VINA .....	131
20.8	GOLD .....	132
21	Anexo 2. Comandos y sitios web de interés .....	134
21.1	Símbolo del sistema.....	134
21.2	AD4.2.....	134
21.3	VINA .....	134
21.4	Páginas web.....	134

## 1 Modelado molecular

El modelado molecular es un conjunto de métodos computacionales y teóricos que se utilizan para la representación aproximada del comportamiento de moléculas, iones y partículas, considerando sus características tridimensionales.<sup>1,2</sup> Los métodos computacionales son una gran herramienta dentro del modelado molecular y se utilizan para calcular propiedades moleculares o simular resultados experimentales. Estos métodos computacionales se pueden clasificar en las siguientes categorías:

### 1.1 *Cálculos de campos de fuerza*

Los campos de fuerza estiman la energía potencial de un sistema bajo estudio y calculan las fuerzas de interacción que actúan sobre los átomos involucrados. Por lo tanto, son elementos clave para el modelado de interacciones interatómicas de sistemas moleculares y se usan como componentes base de un gran número de métodos de simulación.<sup>3</sup> Los campos de fuerza utilizan varios tipos de cálculos y generalmente se verifican con valores experimentales.

- **Campo de fuerza universal (UFF, por sus siglas en inglés):**

Este método es un campo de fuerza que tiene parametrizaciones para cualquier átomo de la tabla periódica con un número atómico menor a 103. Debido a su gran flexibilidad, se puede aplicar a distintos tipos de sistemas, como moléculas orgánicas y complejos metálicos. UFF asume que la topología del sistema simulado es invariable, por lo tanto, se considera un campo de fuerza no reactivo. Los campos de fuerza no reactivos permiten cambios geométricos pequeños del sistema, pero prohíben cualquier rearrreglo grande como generación o el rompimiento de enlaces covalentes, así como cambios en los estados de hibridación de los átomos. Una desventaja es que, al ser universal, tiene una menor precisión que otros campos de fuerza.<sup>3</sup>

- **Campos de fuerza clásicos (CFF, por sus siglas en inglés):**

Este método se lleva a cabo haciendo uso del algoritmo conjunto de reacción de Monte Carlo (REMC por sus siglas en inglés), el cual otorga una solución numérica exacta al problema. Este algoritmo es una traducción directa del enfoque termodinámico macroscópico al ambiente molecular correspondiente. Requiere una lista específica de especies e incorpora de manera implícita las porciones no ideales (residuales) de los potenciales químicos por medio de simulaciones moleculares relacionadas a los campos de fuerza de la especie.<sup>4</sup>

- **Campo de fuerza reactiva (ReaxFF):**

Este método utiliza campos de fuerza para modelar reacciones químicas mediante la formación y disociación de enlaces químicos, interacciones de van der Waals y de Coulomb. Estos modelos son de especial importancia en estudios de dinámica molecular ya que pueden modelar la cinética de reacción.<sup>4</sup>

## **1.2 Métodos ab initio**

Estos métodos se basan en la química cuántica. En ellos, las estructuras moleculares se pueden calcular con base en la ecuación de Schrödinger, los valores constantes fundamentales y el número atómico de los átomos presentes.

- **Mecánica cuántica (QM, por sus siglas en inglés):**

Se realiza una simulación de mecánica cuántica de los átomos en el sistema y un análisis empírico basado en *clusters* de la estructura atómica equilibrada utilizando la identificación molecular como criterio. No se requieren datos estándar de potencial químico. Tiene una demanda computacional muy alta y mientras más compleja sea la molécula, también lo será este método.<sup>4</sup>

## **1.3 Métodos semiempíricos**

Los métodos semiempíricos utilizan aproximaciones de datos experimentales para proveer el *input* a los modelos matemáticos.

- **Mecánica cuántica/mecánica molecular (QM/MM):**

Separan el problema en dos partes que se pueden tratar con diferentes métodos computacionales. Por ejemplo, para una enzima, la parte del sistema en la que ocurre la catálisis y que incluye al sitio activo, los sustratos y los aminoácidos de participación directa, se trata con un nivel adecuado de mecánica cuántica. Para el resto de la enzima, que no tiene participación directa en la reacción e involucra un número mucho mayor de átomos, se utiliza la mecánica molecular, generalmente un campo de fuerza.<sup>5</sup>

Es importante mencionar que ninguno de estos métodos es completamente el mejor. Se debe de escoger aquel que se ajuste mejor a las necesidades de la tarea a realizar. La elección depende de ciertos factores, incluyendo: la naturaleza y tamaño de la molécula, el tipo de información que se busca, la disponibilidad de parámetros experimentales determinados y los recursos computacionales con los que se cuente.

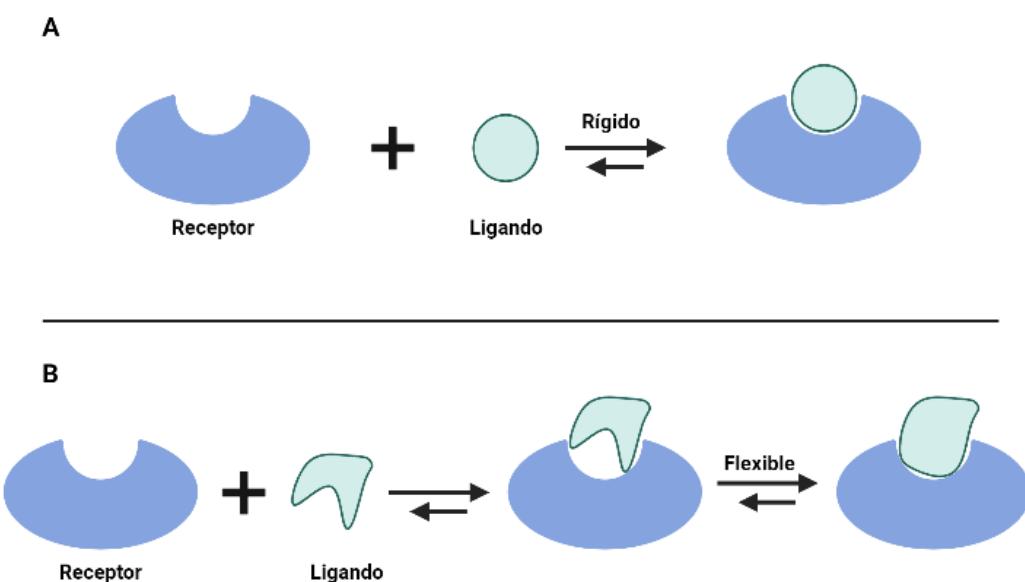
## **2 ¿Qué es el acoplamiento molecular (*docking*)?**

El acoplamiento molecular automatizado (o *docking* en inglés) es una simulación computacional cuyo objetivo se compone de dos pasos: predecir la conformación (comúnmente denominada “pose”) de un ligando en un complejo receptor-ligando y evaluar su afinidad de unión. También se puede definir como un proceso de simulación en donde se estima la posición de un ligando en un sitio diana de unión (o dentro de un blanco molecular).<sup>6-10</sup> En este proceso, el receptor es usualmente una proteína o ácido nucleico (ADN o ARN) y el ligando es una molécula pequeña, aunque también se puede buscar la posición y conformación más favorable de dos macromoléculas.<sup>8,9</sup> Estas predicciones se realizan con base en las interacciones electrostáticas, geométricas, de van der Waals, interacciones de Coulomb y la formación de puentes de hidrógeno. La suma de todas estas

interacciones se evalúa (o califica) por medio de un *score* de *docking*, mismo que representa el potencial de unión.<sup>11</sup>

### 3 Antecedentes del *docking*

La teoría más antigua para el mecanismo de unión ligando – receptor es la teoría de llave – cerradura propuesta por Fischer (**Fig. 1A**), en la que el ligando entra en el receptor como una llave en una cerradura. Los primeros métodos de *docking* reportados se basaban en esta teoría y tanto el ligando como el receptor se trataban como cuerpos rígidos. En este método el valor de la afinidad era directamente proporcional a la complementariedad geométrica entre sus formas.<sup>6,11</sup> Posteriormente el modelo de ajuste inducido, propuesto por Koshland (**Fig. 1B**), planteaba que el sitio activo de una proteína cambia su conformación dependiendo de las interacciones con sus ligandos. Esto sugiere que tanto el ligando como el receptor deberían de tratarse como cuerpos flexibles durante el *docking*, para que de esta manera las interacciones de unión se describieran con mayor exactitud que de manera rígida.<sup>6,11</sup> Considerando las limitaciones de los recursos computacionales, el *docking* se ha realizado con un ligando flexible y un receptor rígido por un largo tiempo y sigue siendo el método más usado hasta ahora.<sup>6</sup>

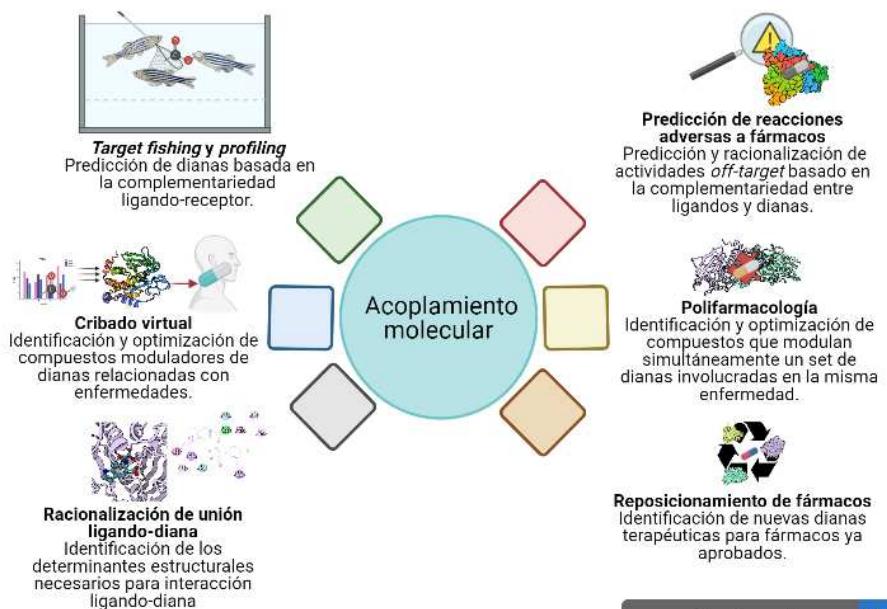


**Figura 1. Imagen de complejo enzima-sustrato.** (A) Teoría de llave – cerradura de Fischer (B) Modelo de ajuste inducido de Koshland. Tomado y modificado de <sup>12</sup>. Creado con biorender.com

### 4 Algunas aplicaciones

El *docking* se desarrolló en principio para investigar el reconocimiento molecular (o las interacciones atómicas) entre una molécula pequeña y el sitio de unión de una macromolécula.<sup>6,13</sup>

Actualmente esta herramienta tiene un sinfín de aplicaciones (**Fig. 2**), sobre todo en la Química Farmacéutica, entre las que se encuentran: identificación y optimización de compuestos *hits* y cabezas de serie (*leads*), reposicionamiento de fármacos y polifarmacología, entre otros.<sup>13</sup>



**Figura 2. Principales aplicaciones del docking en el descubrimiento de fármacos.** El *docking* se utiliza actualmente como auxiliar para comprender las interacciones que lleva a cabo un ligando hacia una diana de interés y para realizar campañas de cribado virtual basados en estructura, similar a cuando se desarrolló por primera vez. Además de estas aplicaciones, también se puede utilizar para identificar series de dianas para los que los ligandos presentan buena complementariedad (también conocidos como *target fishing y profiling*), algunos de los cuales podrían ser potencialmente responsables de reacciones adversas inesperadas (predicción de *off-targets*). Además, el *docking* también se emplea en la actualidad para la identificación de ligandos que se unan simultáneamente a un grupo de dianas de interés seleccionadas (polifarmacología) y para identificar nuevos usos para fármacos que ya tengan perfiles de seguridad optimizados (reposición de fármacos). Tomado y modificado de<sup>13</sup>. Creada con biorender.com

## 5 Flexibilidad en el docking

Existen tres tipos principales de *docking* (relacionado con la flexibilidad del ligando y la proteína), que son: *docking* rígido – rígido, rígido – flexible, y flexible – flexible. Cada uno de ellos tiene demandas computacionales (capacidad del procesador, memoria física y RAM) muy distintas, siendo el rígido – rígido el menos demandante de los tres y el flexible – flexible el de mayor demanda.<sup>6</sup>

### 5.1 Docking rígido

- *Docking* rígido – rígido: la proteína se trata como rígida y solo se exploran los grados de libertad traslacionales y rotacionales del ligando (al ligando se le considera como un cuerpo rígido sin grados de libertad internos o libertad conformacional).<sup>10</sup>

## 5.2 Docking flexible

Generalmente, los ligandos pequeños son muy flexibles y pueden generar un amplio rango de diferentes confórmeros (conocidos como poblaciones), los cuales adaptan su superficie para complementar la forma de la hendidura (o *pocket* en inglés) del receptor. Por esta razón, la mayoría de los programas de *docking* que se enfocan en ligandos pequeños son semiflexibles o flexibles.<sup>14</sup> El ligando se orienta en un espacio  $(6 + N)$ -dimensional que abarca variables translacionales, rotacionales y conformacionales en un ambiente anisotrópico del receptor.<sup>11</sup>

- **Docking rígido – flexible:** la proteína se trata como rígida, pero se exploran todos los grados de libertad del ligando (translacionales, rotacionales y conformacionales). La mayoría de los algoritmos utilizados actualmente consideran a la proteína como una entidad rígida y al ligando se le considera flexible, permitiendo que sus enlaces simples (ej. C – C de carbonos sp<sup>3</sup>) puedan rotar, pero se mantienen fijos la longitud y los ángulos de sus enlaces.<sup>10</sup>
- **Docking flexible – flexible:** la proteína se trata como total o parcialmente flexible y se exploran todos los grados de libertad del ligando.<sup>10</sup>

## 6 Algoritmos

La búsqueda para las soluciones de los candidatos durante el proceso de *docking* se aborda utilizando dos estrategias esenciales: 1) una búsqueda de la solución total del espacio conformacional y 2) una búsqueda guiada gradual y progresiva a través del espacio conformacional en estudio. La primera escanea la solución entera del espacio conformacional en una manera predefinida y sistemática. En contraste, la segunda estrategia escanea tanto una parte de la solución del espacio de forma tanto parcialmente aleatoria como guiada. La segunda estrategia incluye búsquedas por métodos Monte Carlo, alineamiento simulado, dinámica molecular (MD por sus siglas en inglés), algoritmos evolutivos tales como el algoritmo genético (GA por sus siglas en inglés) y Tabú.<sup>14</sup>

El proceso de *docking* comienza con la aplicación de un algoritmo, que posiciona al ligando en un sitio determinado sobre la macromolécula. Este paso aparentemente simple es realmente un reto pues incluso las moléculas pequeñas pueden tener muchos grados de libertad. Estos deben probarse con suficiente precisión para identificar la conformación que tenga mejor complementariedad con la estructura del receptor, pero también lo suficientemente rápido para poder evaluar miles de compuestos en una misma “corrida”.<sup>7</sup>

Los algoritmos de búsqueda se pueden dividir en tres categorías o grupos básicos dependiendo de la metodología que se emplee para explorar la flexibilidad del ligando: métodos sistemáticos, (construcción incremental, búsqueda conformacional, bases de datos), métodos aleatorios o estocásticos (Monte Carlo, algoritmos genéticos, búsqueda Tabú) y métodos de simulación o determinísticos (dinámica molecular, minimización de energía).<sup>7,10</sup>

### **6.1 Algoritmos sistemáticos**

Los algoritmos sistemáticos exploran todos los grados de libertad del ligando durante su búsqueda. Los métodos que emplean este enfoque se pueden clasificar como exhaustivo, construcción incremental y ensamblado conformacional. Las búsquedas exhaustivas exploran de manera sistemática el valor de cada uno de los grados de libertad de manera combinatoria, rotando todos los ángulos diedros del ligando de acuerdo con un rango de valores preestablecido y un conjunto de restricciones iniciales, como restricciones geométricas y químicas. Una desventaja es que los ligandos con más flexibilidad tendrán un mayor número de enlaces rotables, incrementando el problema de optimización significativamente.<sup>10</sup>

- Construcción incremental: este método acomoda al ligando en un sitio activo de manera fragmentada e incremental. Se rompen los enlaces rotables del ligando para dividirlo en varios fragmentos. Posteriormente uno de ellos se selecciona para ser *dockeado* en el sitio activo. Este fragmento, que funciona como ancla, es usualmente el fragmento más grande o aquel que tiene una función o interacción importante con la proteína. Luego, se unen covalentemente los otros fragmentos a la base. Se generan diferentes orientaciones para ajustar dentro del sitio activo, la cual divisa la flexibilidad del ligando.<sup>6,10</sup> Esta estrategia se utiliza ampliamente en el diseño *de novo* de ligandos, que tiene como objetivo descubrir compuestos novedosos uniendo los mejores fragmentos *dockeados* dentro del sitio de unión del receptor.<sup>10</sup> Este método ha sido utilizado en programas como: DOCK 4.0, FlexX, Hammerhead, SLIDE y eHiTS.<sup>6</sup>

### **6.2 Algoritmos aleatorios o estocásticos**

Los métodos estocásticos analizan el espacio conformacional al modificar de manera aleatoria la conformación del ligando (o la población de ligandos con diferentes conformaciones).<sup>6</sup> Estos métodos cambian de manera aleatoria todos los grados de libertad (traslacionales, rotacionales y conformacionales) del ligando, generando una gran diversidad de soluciones. Las poses generadas se evalúan de acuerdo con un criterio de probabilidad para decidir si serán o no rechazadas. La desventaja de estos métodos es que no existe garantía de convergencia para la solución óptima. Además, se requieren de corridas múltiples e independientes para maximizar la probabilidad de encontrar el mínimo global de energía.<sup>10</sup> Algunos algoritmos de este grupo son: Monte Carlo, algoritmos evolutivos, algoritmos genéticos, búsqueda Tabú y optimización de Swarm.<sup>6,10</sup>

- **Monte Carlo:** este método genera poses del ligando rotando sus enlaces o con movimientos de rotación o traslación de cuerpo rígido. La conformación que se obtiene pasa por una minimización de energía y posteriormente se selecciona o descarta de acuerdo con un criterio de energía, el factor de Boltzmann, que considera la energía de las soluciones antes y después de los cambios.<sup>10</sup> Si pasa este criterio, la pose se guarda y se sigue modificando para generar una nueva conformación. Esto seguirá hasta probar la cantidad de conformaciones previamente establecida. Su principal ventaja es que el cambio puede ser significativo permitiendo al ligando

cruzar barreras energéticas sobre el potencial energético de superficie, algo que no se logra fácilmente por métodos como la dinámica molecular.

Una variación importante de este método es el Alineamiento Simulado, el cual, aplica variaciones en la temperatura para incrementar la probabilidad de encontrar un mínimo conformacional global, utilizando el principio de que, al incrementar la temperatura se pueden superar algunas barreras energéticas. Algunos softwares que han aplicado este algoritmo son: una versión anterior de AutoDock, ICM, QXP y Affinity.<sup>6,10</sup>

- **Algoritmos evolutivos:** Son métodos de optimización basados en la teoría de la evolución de poblaciones biológicas de acuerdo con la selección natural para obtener soluciones óptimas a un problema particular. Estos métodos se dividen en tres grupos principales: algoritmos genéticos, programación evolutiva y estrategias de evolución. La ventaja que tienen es la facilidad con la que evitan mínimos locales y la posibilidad de identificar una gran diversidad de soluciones de baja energía.<sup>10</sup>
  - **Algoritmos genéticos (GA):** La idea detrás de este algoritmo surge de la teoría de la evolución de Darwin. Los grados de libertad del ligando se codifican como cadenas binarias denominadas “genes”. Estos genes conforman el “cromosoma” que representa la pose del ligando. Dentro de este algoritmo existen dos clases de operadores genéticos: mutaciones y *crossover*. Las mutaciones realizan cambios aleatorios a los genes; el *crossover* intercambia genes entre dos cromosomas. Cuando los operadores genéticos afectan a los genes, esto resulta en una nueva conformación del ligando. Estas nuevas conformaciones serán evaluadas con una función de calificación (*score*) y las que sobrevivan (sobrepasen una escala) se podrán usar para la siguiente generación. Estos algoritmos han sido usados en programas como Autodock, GOLD, DIVALI y DARWIN.<sup>6</sup>
- **Búsqueda Tabú:** Además de modificar de manera aleatoria los grados de libertad del ligando manteniendo el nivel más bajo de energía posible, este algoritmo analiza la similitud de una pose generada recientemente que no sea la de menor energía utilizando una “lista Tabú”, que contiene las soluciones de menor energía encontradas anteriormente. Una pose solo se mantiene si es distinta de todas las soluciones previamente encontradas. Este enfoque previene que el algoritmo visite regiones que ya fueron exploradas.<sup>10</sup>
- **Optimización de Swarm:** Este enfoque, también llamado “inteligencia Swarm”, se inspira en el comportamiento colectivo de agentes como aves y hormigas. De acuerdo con este algoritmo, los cambios realizados en una solución se dirigen para seguir la mejor pose de la población. Algunas variaciones de este enfoque son la optimización Swarm de partículas y la optimización de Colonia de Hormigas.<sup>10</sup>

### ***6.3 Métodos de simulación o determinísticos***

En estos métodos el estado actual del sistema determina las modificaciones que se harán, desembocando en su siguiente estado. El resultado final depende en gran medida de la estructura inicial que se haya dado pues, dada exactamente la misma configuración inicial del sistema y un conjunto particular de parámetros, el estado final siempre será el mismo. Algunos ejemplos de este algoritmo son la minimización de energía y la dinámica molecular.<sup>10</sup>

- **Minimización de energía:** Explora el panorama energético utilizando la dirección asociada con el gradiente de potencial de energía, llevando al sistema al mínimo local más cercano. El principal problema asociado con esta estrategia es su poca habilidad para sobrepasar barreras energéticas y lograr otros mínimos locales, explorando el panorama energético de manera más eficiente. La minimización de energía generalmente se emplea en conjunto con otras estrategias de *docking*, como la minimización de soluciones de *docking* obtenidas utilizando un método basado en fragmentos.<sup>10</sup>
- **Dinámica molecular:** Es un método en donde se simula el movimiento de un sistema a través del tiempo, considerando algunas variables como temperatura y presión. Este método representa la flexibilidad de manera más efectiva que la de otros algoritmos ya que durante la simulación se consideran todos los grados de libertad tanto del ligando como de la proteína. También son muy eficientes para la optimización local. La desventaja es que hacen progresos en pasos muy pequeños por lo que tienen dificultad para pasar sobre barreras conformacionales de alta energía, lo que puede desembocar en un muestreo inadecuado, además de su alta demanda computacional. Por lo tanto, una estrategia actual es utilizar un algoritmo aleatorio para identificar la conformación del ligando y posteriormente realizar la dinámica molecular.<sup>6,10</sup>

## **7 Funciones de evaluación o score**

Las funciones de *score* son métodos de aproximaciones matemáticas o fisicoquímicas que estiman la afinidad de unión entre una diana y un ligando. Estas funciones, además, pueden ser utilizadas como un método para localizar sitios de unión no identificados de las proteínas. Este *score* estima la constante de afinidad entre un ligando y receptor a partir de la ecuación de Gibbs-Helmholtz<sup>9</sup>:

$$\Delta G = \Delta H - T\Delta S$$

Donde  $\Delta G$  es la energía libre de unión,  $\Delta H$  es la entalpía,  $T$  es la temperatura en Kelvin y  $\Delta S$  es la entropía.

La constante de afinidad,  $K_i$ , se puede relacionar con la ecuación anterior de la siguiente manera:

$$\Delta G = -RT \ln K_i$$

Los algoritmos se complementan con funciones de score que están diseñadas para predecir la actividad biológica a través de la evaluación de las interacciones entre los compuestos y dianas potenciales.<sup>7</sup>

Las funciones de *score* se pueden clasificar en tres tipos:

- basados en campos de fuerza: estiman la energía de unión calculando la suma de las interacciones de no enlace (interacciones electrostáticas y de van der Waals).<sup>6,9</sup>
- basados en conocimiento: se basan en observaciones estadísticas de contactos intermoleculares identificados en bases de datos de estructuras. Se asume que mientras más favorable sea una interacción, esta ocurrirá con mayor frecuencia. El *score* se calcula favoreciendo los contactos preferidos y penalizando las interacciones repulsivas.<sup>6,9</sup>
- funciones de *score* empíricas: utilizan varios términos de interacciones intermoleculares, como puentes de hidrógeno, interacciones iónicas, efecto hidrofóbico y entropía de unión. Cada componente se multiplica por un coeficiente y luego se suman para dar un valor final. Los coeficientes se obtienen a través de un análisis de regresión lineal de complejos enzimáticos conocidos, donde los valores teóricos se acomodan para estar lo más cercanos posibles a los datos experimentales.<sup>6,9</sup>

## 8 Metodología del docking

Existen tres componentes clave en el *docking*:

1. Representación del sistema
2. Búsqueda en el espacio conformacional
3. Calificación de soluciones potenciales

El *docking* simula la interacción de un ligando con una superficie proteica. Por lo tanto, la primera pregunta es cómo definir la superficie de una proteína. Esta superficie se puede describir por modelos matemáticos, como descriptores de forma geométrica o una cuadrícula denominada *grid*. Además, esta superficie puede ser tratada de manera estática o dinámica, como por ejemplo rígido vs. flexible.<sup>14</sup>

El *docking* consiste en tres objetivos que son: predicción de pose, cribado virtual y estimación de la afinidad de unión. Una metodología exitosa debería de ser capaz de predecir correctamente, con cierto margen de error, la pose de unión de un ligando nativo dentro de su sitio diana además de sus interacciones moleculares. Adicionalmente, al probar bibliotecas de compuestos, el programa debería de ser capaz de distinguir entre moléculas que se unen y las que no se unen. También debería de poder calificar la afinidad de unión de los ligandos de la base de datos.<sup>10</sup>

## 9 Softwares

En la actualidad existen más de 60 *softwares* diferentes para la realización de cálculos de acoplamiento molecular. En la **Tabla 1** se agrupan algunos de los más usados y conocidos, junto con el algoritmo que usan y su método de evaluación. Algo que se debe de tener en cuenta siempre es que no todos los algoritmos de *docking* son adecuados para cualquier sistema. Generalmente, se recomienda utilizar más de un programa de *docking* para así poder establecer un consenso entre los resultados y predecir de manera más acertada las interacciones y las poses de unión del complejo ligando – receptor.<sup>11,15</sup>

**Tabla 1.** Algunos *softwares* para cálculos de acoplamiento molecular. Tomada y modificada de <sup>12</sup>.

Nombre	Algoritmo de búsqueda	Método de score	Velocidad de cálculo	Algunas características
<b>AutoDock 4.2</b>	GA (Algoritmo genético) LGA (algoritmo genético de Lamarckian)	Semiempírico	Media	<i>Docking</i> rígido–flexible. Es un <i>software</i> libre para uso académico
<b>Autodock VINA</b>	GA	Semiempírico	Rápida	<i>Docking</i> rígido–flexible. Emplea un optimizador global de búsqueda local iterada y es más rápido que AutoDock 4.2
<b>Dock</b>	Algoritmo de fragmentación	Campo de fuerza molecular	Rápida	<i>Docking</i> flexible
<b>FlexX</b>	Algoritmo de fragmentación	Semiempírico	Rápida	<i>Docking</i> rígido–flexible
<b>Glide</b>	Búsqueda sistemática exhaustiva	Semiempírico	Media	<i>Docking</i> flexible
<b>GOLD</b>	GA	Semiempírico	Rápida	<i>Docking</i> rígido–flexible. La precisión y fiabilidad de este programa han sido altamente evaluadas

<b>LeDOCK</b>	Alineamiento simulado GA	Campo de fuerza molecular	Rápida	<i>Docking</i> flexible. Es un <i>software</i> relativamente nuevo. Por su rapidez, se recomienda para cribado virtual
<b>RDOCK</b>	GA MC Minimización Simplex	Campo de fuerza molecular	Media	<i>Docking</i> rígido. Está especialmente diseñado para campañas de cribado virtual
<b>ZDOCK</b>	Complementariedad geométrica y dinámica molecular	Campo de fuerza molecular	Media	<i>Docking</i> rígido–flexible

## 10 Antes de comenzar...

A continuación, revisaremos todos los pasos necesarios para hacer los acoplamientos moleculares (*dockings*) en tres programas diferentes: Autodock4.2 (AD4.2), Autodock VINA (VINA) y GOLD. Por lo tanto, el primer paso es tener todos los programas instalados (para conocer cómo instalarlos, ver el **Anexo 1. Instalación de programas**).

En el siguiente esquema (Esquema 1), se presenta un comparativo sobre los archivos que requiere cada programa. Es importante tenerlos en cuenta para que no falte ninguno y no tengamos dificultades al realizar estos ejercicios.

AD4.2	VINA	GOLD
<ul style="list-style-type: none"><li>• ligando minimizado, en formato *.pdbqt</li><li>• proteína minimizada, en formato *.pdbqt</li><li>• archivo grid.gpf</li><li>• archivo dock.dlg</li><li>• ejecutable autogrid4.exe</li><li>• ejecutable autodock4.exe</li></ul>	<ul style="list-style-type: none"><li>• ligando minimizado, en formato *.pdbqt</li><li>• proteína minimizada, en formato *.pdbqt</li><li>• archivo conf.txt</li></ul>	<ul style="list-style-type: none"><li>• ligando minimizado, en formato *.mol2</li><li>• proteína minimizada, en formato *.pdb</li><li>• No se requiere un archivo, pero se deben tener apuntadas las coordenadas que se utilizarán</li></ul>

**Esquema 1.** Comparativo de archivos necesarios para cada programa

Como podemos ver, los tres programas tienen en común la necesidad de tener a la proteína y al ligando o ligandos en su forma minimizada. Por lo tanto, serán esos los primeros pasos que revisaremos a continuación.

Para este ejemplo, utilizaremos la estructura cristalizada de la Proteína Tirosina Fosfatasa 1B (PTP1B).

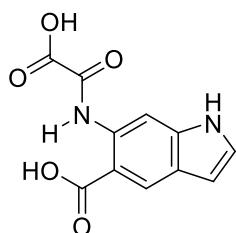
La Proteína Tirosina Fosfatasa 1B (PTP1B por sus siglas en inglés) es una enzima que participa en la desfosforilación del receptor de insulina y en la señalización de la leptina. Por tanto, su inhibición es de interés terapéutico para el tratamiento de la diabetes mellitus tipo 2 y la obesidad. En este aspecto, uno de los inhibidores que se une al sitio catalítico de la PTP1B, y que además se le puede encontrar co-cristalizado con esta enzima en la base de datos de *Protein Data Bank*, es el ácido 6-oxalilamino-(1*H*)-indol-5-carboxílico (OAI).<sup>16</sup>

## 11 Cómo preparar un ligando

Para preparar un ligando el primer paso es saber su nombre y estructura. Por tanto, lo primero que haremos es entrar al sitio web de *Protein Data Bank*, <https://www.rcsb.org/search> y buscar el PDB ID correspondiente a la proteína que nos interesa (en este ejemplo, PDB ID: **1C83**). En la página de 1C83 en RCSB, dice que el nombre del ligando es “6-(OXALYL-AMINO)-1H-INDOLE-5-CARBOXYLIC ACID”, además, nos indica la clave del mismo, OAI (que nos será de utilidad cuando preparemos la proteína).

Small Molecules				
Ligands 1 Unique				
ID	Chains	Name / Formula / InChI Key	2D Diagram	3D Interactions
OAI <a href="#">Query on OAI</a>	A	6-(OXALYL-AMINO)-1H-INDOLE-5-CARBOXYLIC ACID C <sub>11</sub> H <sub>8</sub> N <sub>2</sub> O <sub>5</sub> AHWMERGBWWROMM-UHFFFAOYSA-N		<a href="#">Ligand Interaction</a>
		<a href="#">Download Ideal Coordinates CCD File</a>		
		<a href="#">Download Instance Coordinates</a>		

También podemos ver que la estructura de este ejemplo es la siguiente:



ácido 6-(oxalilamino)-1*H*-indol-5-carboxílico

Para obtener nuestro ligando tenemos dos opciones:

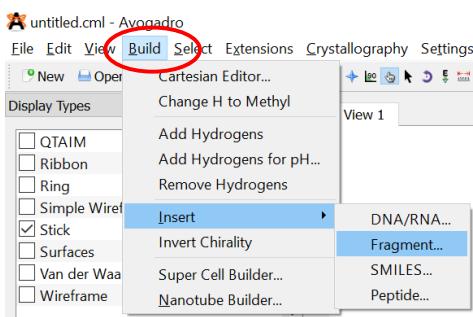
1. Construirlo desde cero en algún programa, como **Avogadro** (el cuál es de utilidad para moléculas novedosas)
2. Buscar el ligando (o el esqueleto del ligando) en alguna base de datos

A continuación, revisaremos ambas opciones.

### 11.1 *Construirlo desde cero*

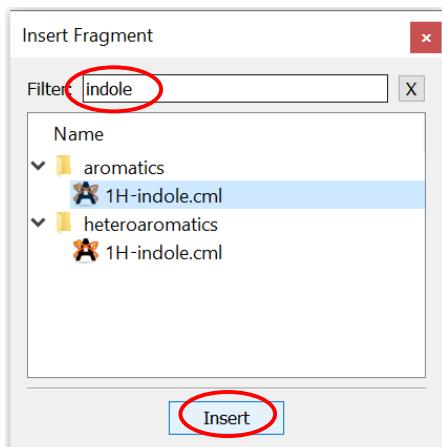
A continuación, veremos brevemente algunas herramientas de utilidad en Avogadro. (Si quieres saber más, puedes consultar algún manual o vídeo de Youtube,<sup>17,18</sup>)

Para comenzar, dibujaremos el núcleo base del ligando, el **indol**. Para ello, nos dirigiremos al menú “Build”, después a “Insert” y finalmente a “Fragment...”

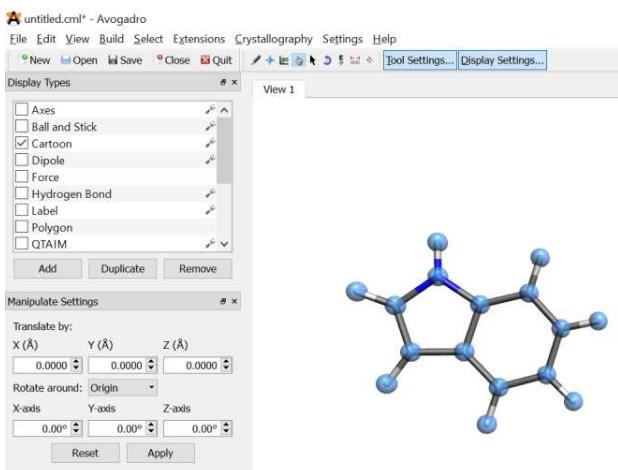


Una vez que hemos hecho esto, se desplegará un menú con diversos grupos funcionales que pueden ser de interés. También tenemos una barra de búsqueda, la usaremos para escribir “indole” y encontrar más rápidamente lo que buscamos.

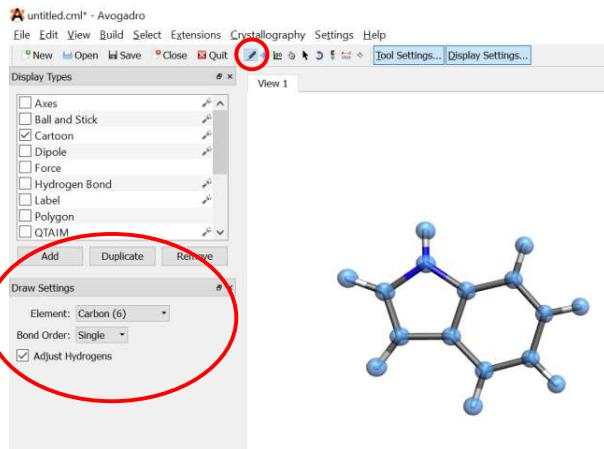
Cuando lo hayamos escrito, damos click sobre “Insert”



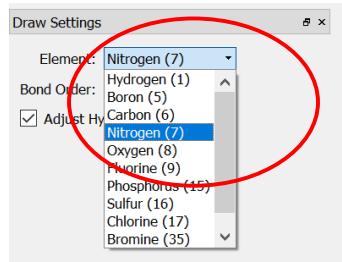
Veremos que aparece la estructura. El menú de fragmentos lo podemos cerrar.



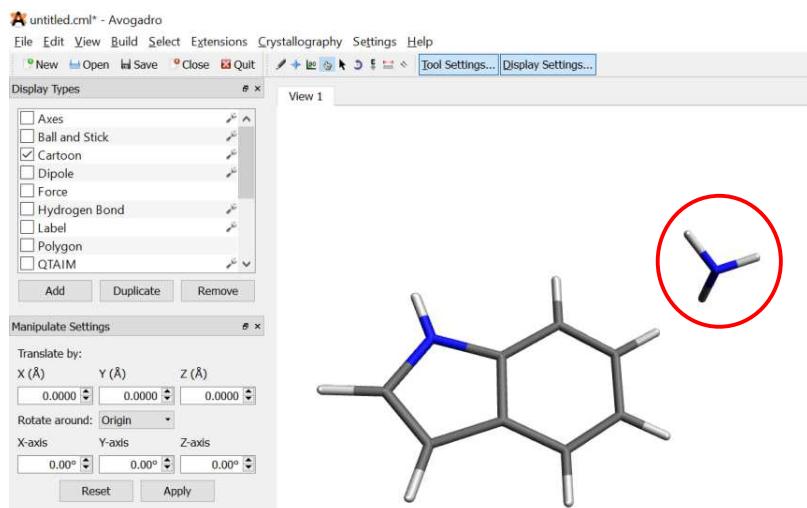
Para agregarle los sustituyentes correspondientes, utilizaremos el menú “Draw Settings” que tenemos del lado izquierdo. (Si no se visualiza este menú, podemos hacer click en el ícono del lápiz en la parte superior del programa para que se despliegue)



Una vez localizado el menú, en “Element” seleccionaremos el tipo de átomo que queremos, en este caso, nitrógeno.

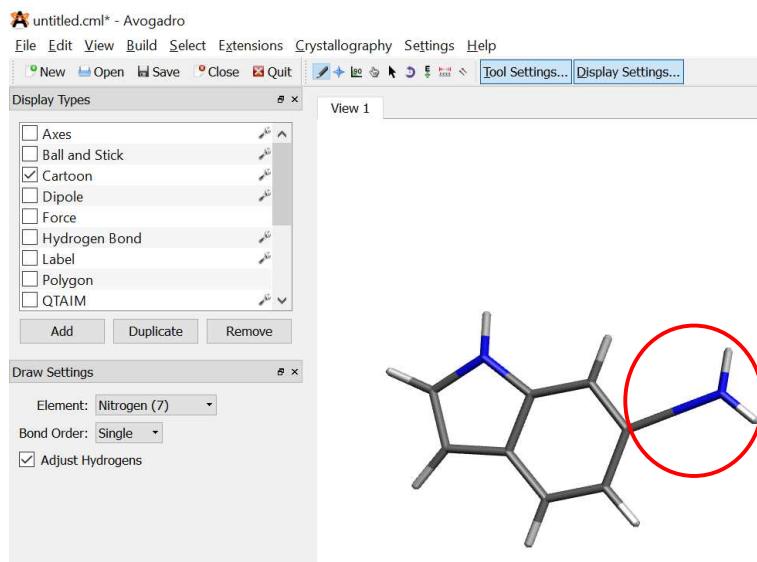


Posteriormente, hacemos click en la pantalla y veremos que aparece una molécula de NH<sub>3</sub>

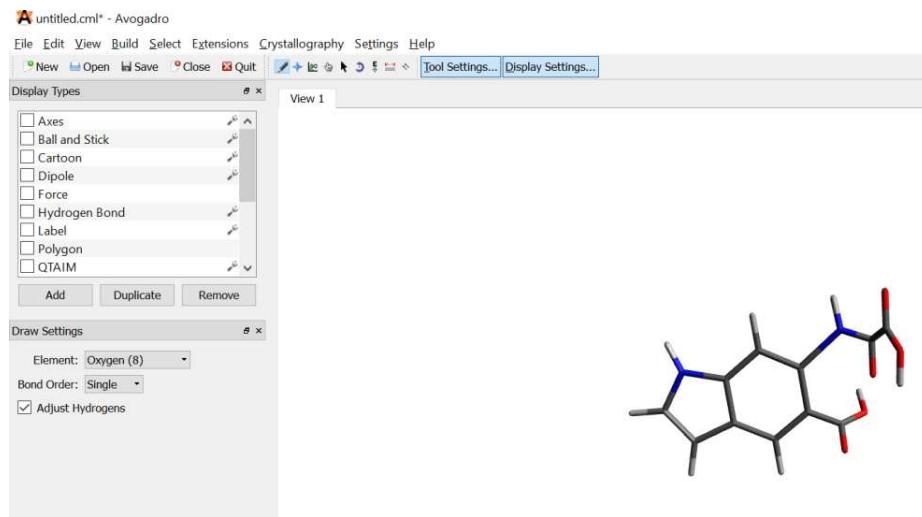


El siguiente paso es marcar los enlaces entre estas moléculas. Para ello, haremos click sobre el átomo de N del NH<sub>3</sub> y sin soltarlo, lo arrastraremos a donde queremos que se una, al C del anillo bencénico, y soltamos el botón. Veremos que se forma un enlace.

Otra manera de lograrlo es haciendo click directamente sobre el átomo de C en el que queremos generar el enlace.

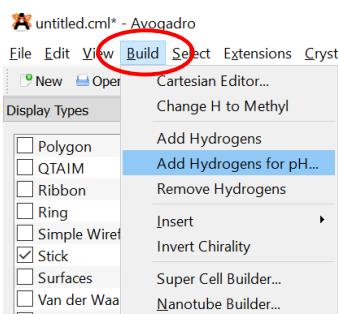


Repetimos los mismos pasos para agregar los átomos que nos faltan y obtendremos lo siguiente:

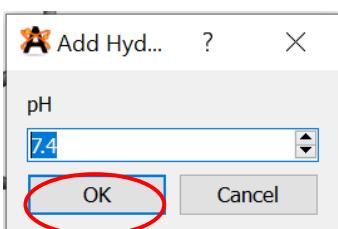


En este punto, ya tenemos la estructura que queremos. Sin embargo, todavía debemos de establecer dos cosas: la protonación o desprotonación de la molécula a pH fisiológico y su minimización.

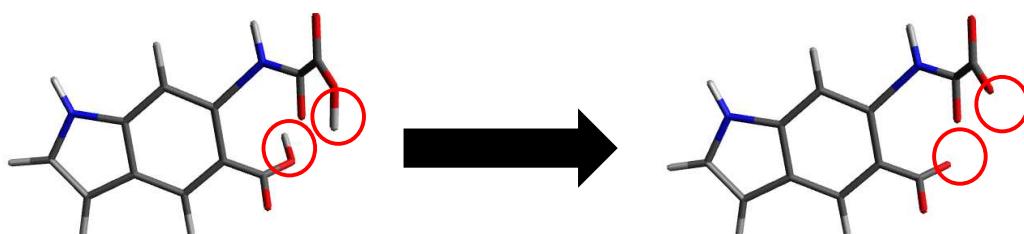
Recordemos que a pH fisiológico hay ciertos grupos funcionales que se protonan o desprotonan. Para que nuestros *dockings* sean más precisos, debemos de tomar esta situación en cuenta. Para ello, podemos ayudarnos de Avogadro. iremos al menú “Build”, seguido de “Add Hydrogens for pH...”



Aparecerá una ventana en donde podemos ajustar el valor del pH que queremos, en este caso, 7.4 y hacemos click en “OK”

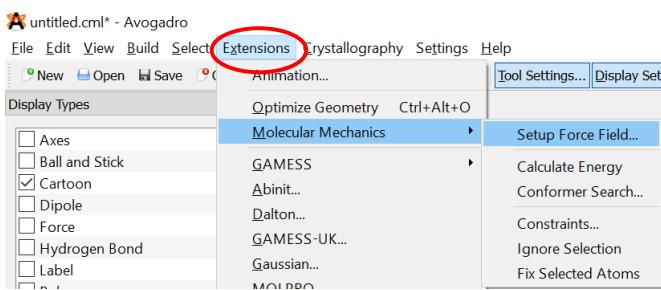


Veremos que se eliminan los hidrógenos correspondientes.

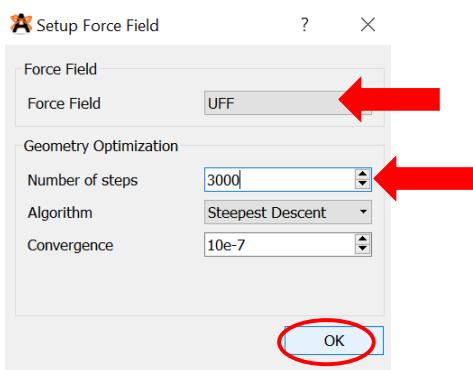


### 11.2 Cómo minimizar un ligando

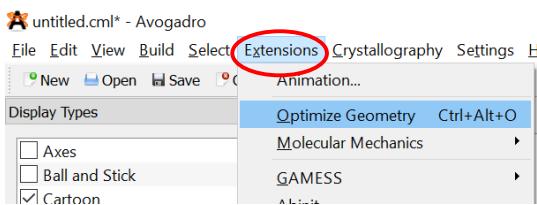
El siguiente paso es la minimización. Para ello, nos dirigiremos al menú “Extensions”, “Molecular Mechanics” y “set-up Force Field...”



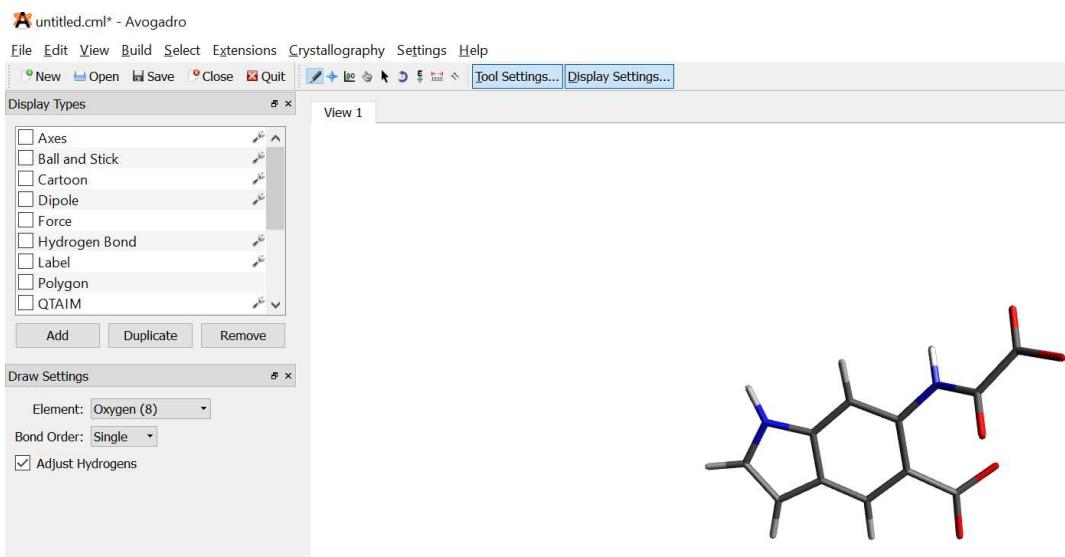
Aparecerá una pequeña ventana en donde podemos modificar los parámetros de la minimización. Para este ejemplo, escogeremos el campo de fuerza (Force Field) “UFF” y en número de pasos (Number of steps) escribiremos “3000”. Los demás parámetros no los modificaremos. Hacemos click sobre “OK”.



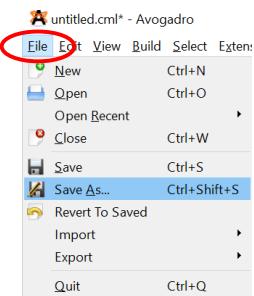
La ventana se cierra y no hay cambio aparente, pues lo que hicimos fue guardar los parámetros con los que queremos realizar la minimización. El siguiente paso es realizar la minimización. Para ello, vamos de nuevo al menú “Extensions” y seleccionamos la opción “Optimize Geometry”.



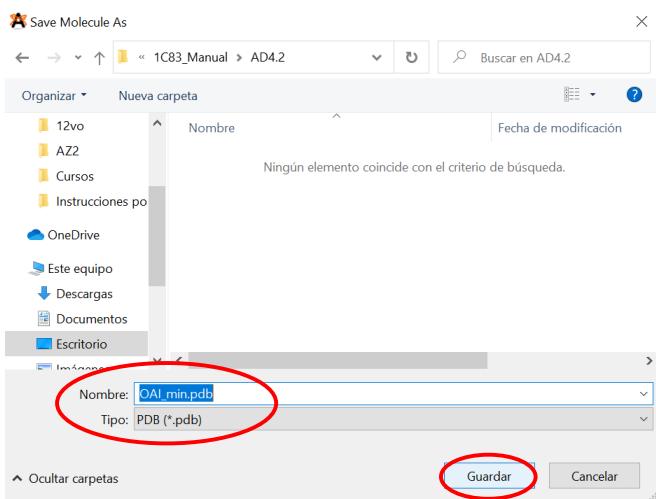
Veremos que la molécula se mueve para acomodarse en su forma más estable y termina así:



El siguiente paso es guardar el ligando ya minimizado. Para ello, vamos al menú “File”, “Save As...”



Escogemos el destino, nombramos nuestro archivo (se sugiere nombrarlo *nombre\_min*; min = minimizado), verificamos que se vaya a guardar en formato \*.pdb y hacemos click en “Guardar”.

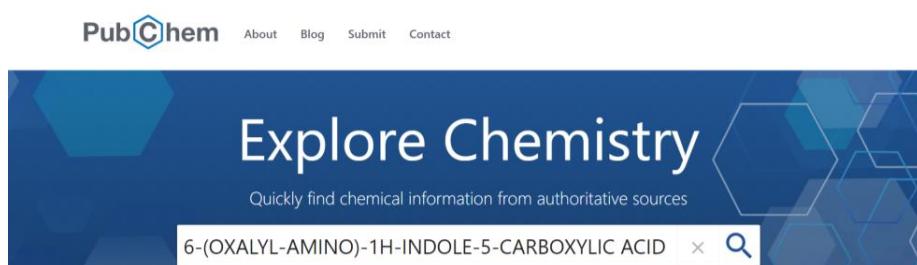


Es importante mencionar que el programa Avogadro no solo nos puede servir para generar archivos de moléculas conocidas (como el OAI), sino también podemos proponer, dibujar y cuantos ligandos estén en nuestras mentes para después llevar a cabo el *docking* con ellos.

### 11.3 Con una base de datos

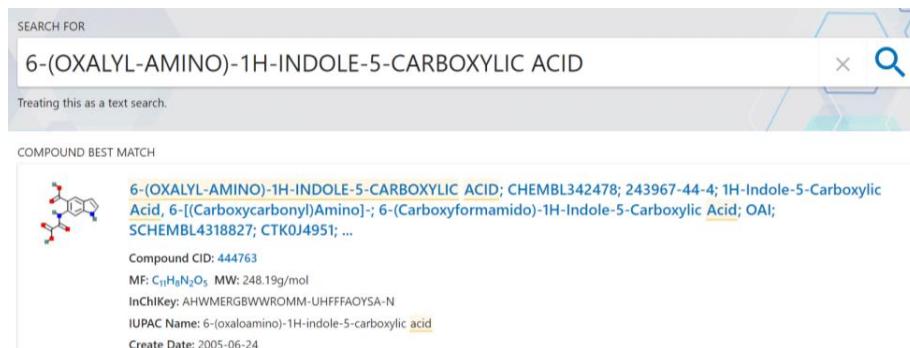
Existen diversas bases de datos de estructuras, como **PubChem** o **ZINC Database**. En este ejemplo, buscaremos nuestro ligando en **PubChem**.

Ya sabemos que el nombre del ligando es 6-(OXALYL-AMINO)-1H-INDOLE-5-CARBOXYLIC ACID. Podemos copiar el nombre y después pegarlo en el buscador de **PubChem**.



The screenshot shows the PubChem homepage with the title "Explore Chemistry" and the subtext "Quickly find chemical information from authoritative sources". Below the title is a search bar containing the text "6-(OXALYL-AMINO)-1H-INDOLE-5-CARBOXYLIC ACID".

Escogemos la opción que corresponda al ligando de interés, en este caso la primera opción.



The screenshot shows the PubChem compound details page for the first result. The search bar at the top contains the same query. The main content area is titled "COMPOUND BEST MATCH" and features a chemical structure of the compound. Below the structure, the full name is listed: "6-(OXALYL-AMINO)-1H-INDOLE-5-CARBOXYLIC ACID; CHEMBL342478; 243967-44-4; 1H-Indole-5-Carboxylic Acid, 6-[(Carboxycarbonyl)Amino]-; 6-(Carboxyformamido)-1H-Indole-5-Carboxylic Acid; OAI; SCHEMBL4318827; CTKOJ4951; ...". Other details include: Compound CID: 444763, MF: C10H8N2O5, MW: 248.19g/mol, InChIKey: AHWMERGBWWROMM-UHFFFAOYSA-N, IUPAC Name: 6-(oxaloamino)-1H-indole-5-carboxylic acid, and Create Date: 2005-06-24.

Aparecerá esta ventana y en el menú de la derecha, escogemos “Names and Identifiers” y hacemos click sobre “Canonical SMILES” para ir a esa sección.

## COMPOUND SUMMARY

# 6-(Oxaryl-amino)-1H-indole-5-carboxylic acid

Cite Download

**CONTENTS**

- Title and Summary**
- 1 Structures**
- 2 Names and Identifiers**   ←
- 2.1 Computed Descriptors
- 2.1.1 IUPAC Name
- 2.1.2 InChI
- 2.1.3 InChI Key
- 2.1.4 Canonical SMILES
- 2.2 Molecular Formula
- 2.3 Other Identifiers
- 2.3.1 CAS
- 2.3.2 DSSTox Substance ID

PubChem CID 444763

Structure

Find Similar Structures

Molecular Formula C<sub>11</sub>H<sub>8</sub>N<sub>2</sub>O<sub>5</sub>

6-(OXARYLAMINO)-1H-INDOLE-5-CARBOXYLIC ACID

Una vez en esa sección, seleccionamos el código SMILES y lo copiamos.

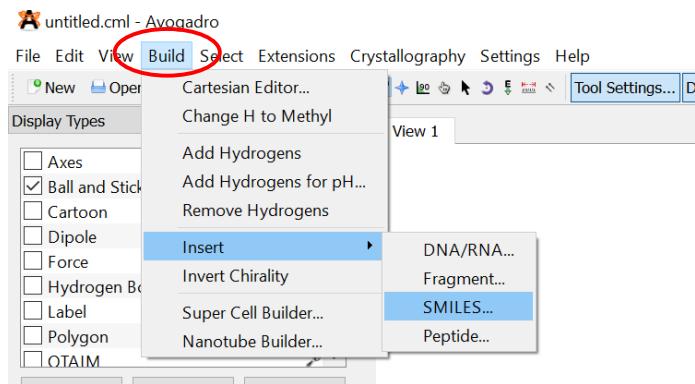
## 2.1.4 Canonical SMILES

**C1=CNC2=CC(=C(C=C21)C(=O)O)NC(=O)C(=O)O**

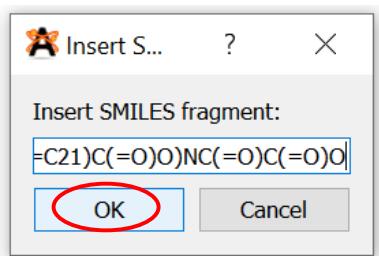
Computed by OEChem 2.1.5 (PubChem release 2019.06.18)

► PubChem

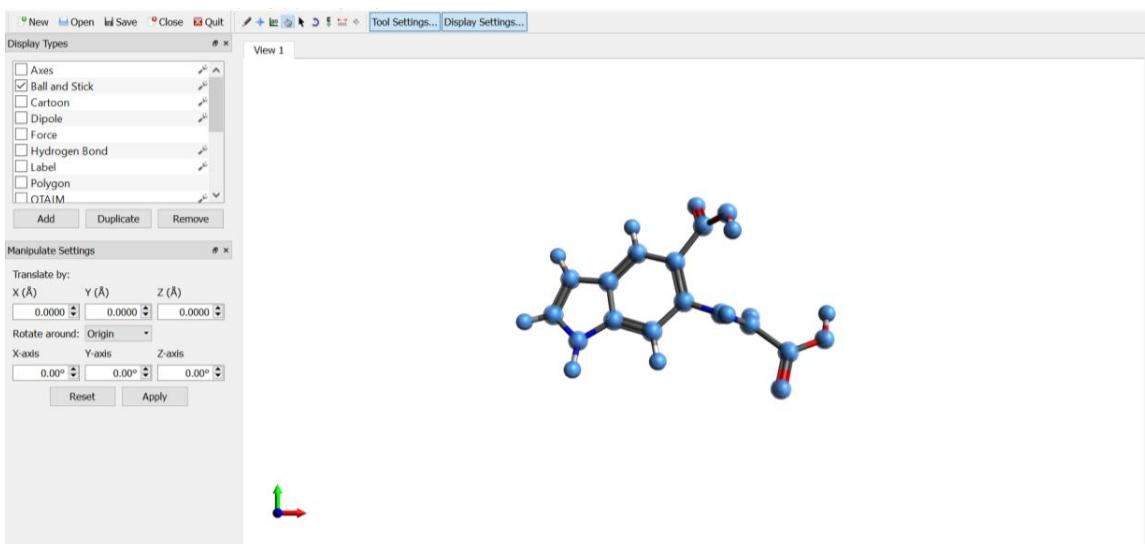
Posteriormente, en **Avogadro**, en el menú “Build”, hacemos click en “Insert” y luego en “SMILES”.



Aparecerá una ventana, en ella pegamos el código que copiamos previamente y damos click en “OK”.



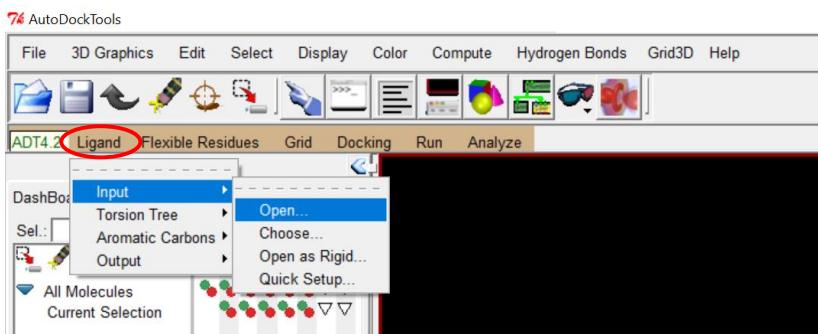
Después, aparecerá la molécula, el OAI.



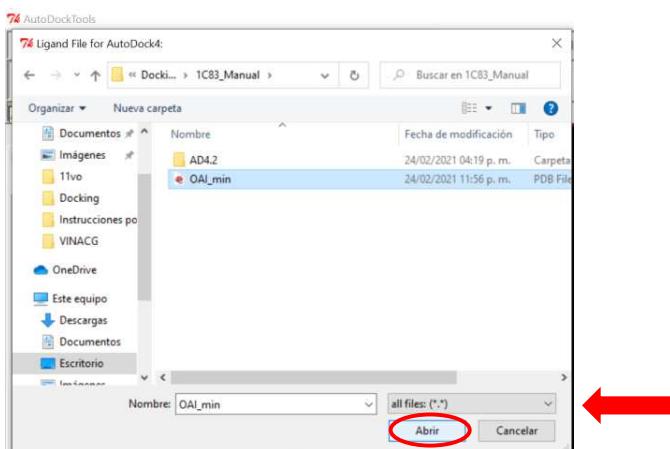
Para hacerle las modificaciones que faltan, primero debemos guardar la molécula. Podemos hacerlo siguiendo los pasos que mencionamos en el apartado anterior. Una vez que la hemos guardado, podemos proceder al ajuste de hidrógenos a pH fisiológico y a la minimización, siguiendo los mismos pasos descritos en la sección anterior. Una vez que lo hemos hecho, volvemos a guardar esta molécula y la nombramos como **OAI\_min.pdb**.

Si bien, este archivo \*.pdb puede ser utilizado para los programas GOLD, no es lo mismo para AD4.2 y VINA. Recordemos que estos dos últimos necesitan el ligando en archivo \*.pdbqt. Por lo tanto, el paso final será convertir el archivo \*.pdb en \*.pdbqt. Esto lo podemos realizar en **Autodock Tools (ADT)**.

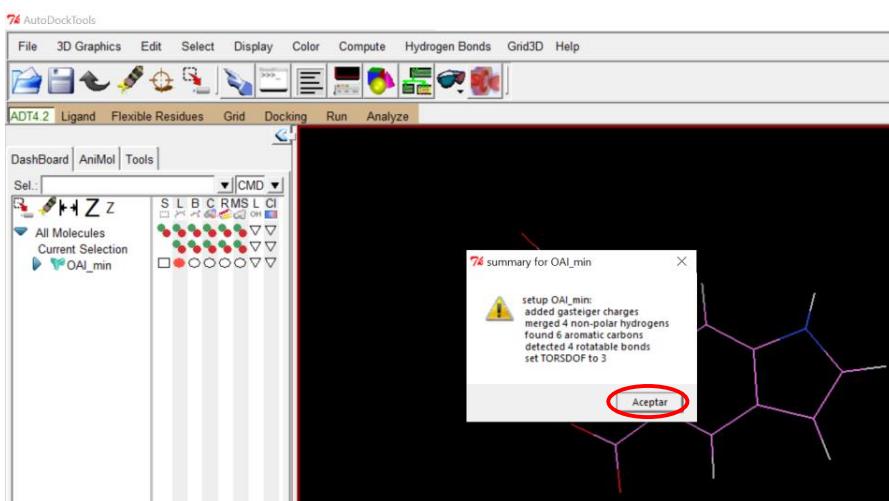
Para ello, abrimos el programa **Autodock Tools (ADT)**, hacemos click sobre la pestaña “Ligand”, después “Input” y “Open” y seleccionamos el archivo \*.pdb que generamos previamente en Avogadro.



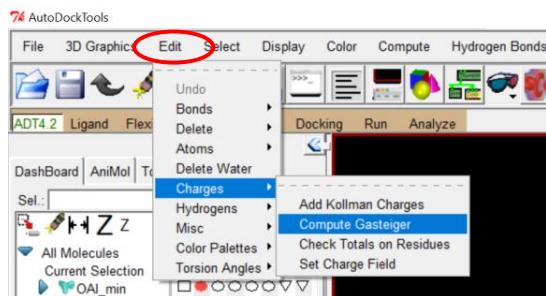
Algo importante de mencionar es que ADT por *default* solamente muestra los archivos \*.pdbqt. Por lo tanto, debemos cambiar la opción a “all files” para encontrar nuestro archivo. Lo seleccionamos y hacemos click sobre “Abrir”.



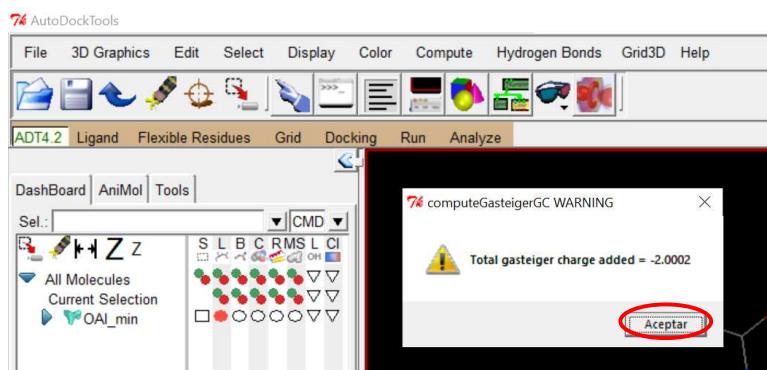
Una vez que abrimos el archivo, saldrá este mensaje. Hacemos click sobre “Aceptar”.



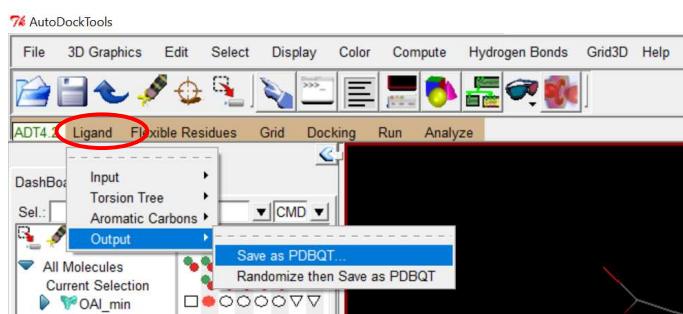
Después, nos dirigimos a la pestaña “Edit”, después “Charges” y “Compute Gasteiger”.



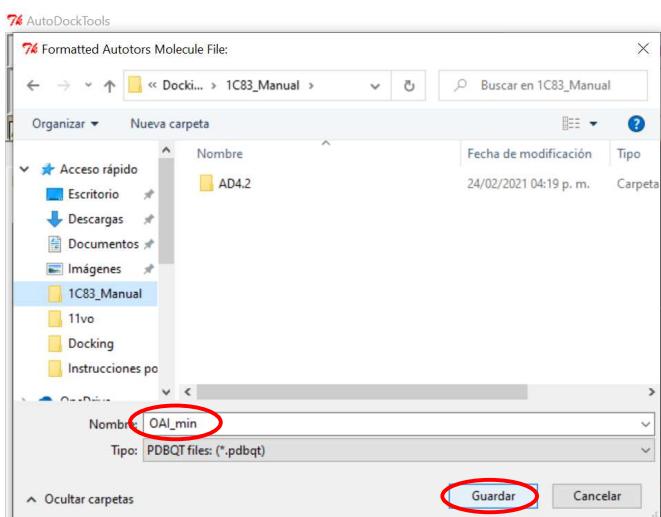
Una vez que las cargas se añadieron, saldrá este mensaje. Damos click en “Aceptar”.



El último paso es simplemente guardar el ligando (al que ya le adicionamos las cargas) en formato \*.pdbqt. Para hacerlo, vamos a la ventana “Ligand”, después “Output” y “Save as PDBQT”.



Después, nombramos a nuestro archivo, escogemos su destino y hacemos click en “Guardar”.



En este punto, ya tenemos a nuestro ligando listo. El siguiente paso es preparar la proteína.

## 12 Cómo preparar una proteína

Lo primero que necesitamos es el archivo \*.pdb de la proteína de interés para después abrirlo en **PyMOL**. Existen al menos **tres formas diferentes** de abrir este archivo en **PyMOL**, las cuales son:

1. Descargar el \*.pdb directamente de **RCSB Database** y después abrirlo en **PyMOL**
2. Buscar la proteína de interés utilizando el PDB Loader Service de PyMOL.
3. Escribir comandos en PyMOL.

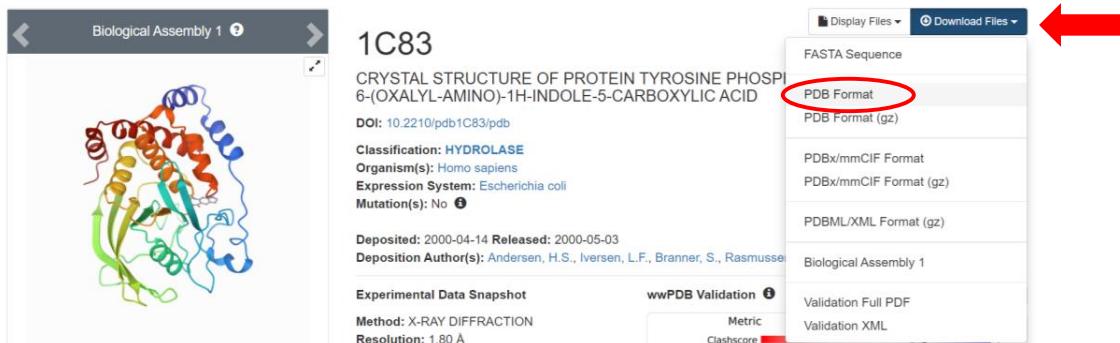
### 12.1 Con RCSB Database

Entramos al sitio web <https://www.rcsb.org/search> y buscamos el PDB ID correspondiente a la proteína que nos interesa (en este ejemplo, PDB ID: **1C83**).



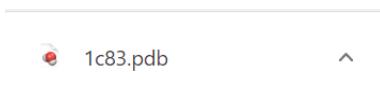
The screenshot shows the RCSB PDB website. At the top, there's a navigation bar with links like Deposit, Search, Visualize, Analyze, Download, Learn, More, and Documentation. Below the navigation is the RCSB PDB logo and a search bar containing the text '1c83'. A red circle highlights the search term '1c83'. To the right of the search bar is a magnifying glass icon. Below the search bar, there's a banner for 'Celebrating 50 YEARS OF Protein Data Bank'.

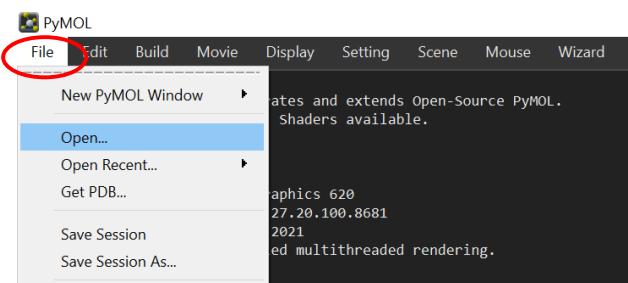
Después, nos aparecerá la página de 1C83. Hacemos click sobre “Download Files” y después escogemos “PDB Format”.



The screenshot shows the detailed view of the protein 1C83. On the left, there's a 3D ribbon diagram of the protein. In the center, the protein name '1C83' is displayed above its crystal structure. To the right, there's a sidebar with various details: DOI: 10.2210/pdb1C83/pdb; Classification: HYDROLASE; Organism(s): Homo sapiens; Expression System: Escherichia coli; Mutation(s): No; Deposited: 2000-04-14; Released: 2000-05-03; Deposition Author(s): Andersen, H.S., Iversen, L.F., Branner, S., Rasmussen. Below this, there's an 'Experimental Data Snapshot' section with Method: X-RAY DIFFRACTION and Resolution: 1.80 Å. To the right of the sidebar, there's a 'wwPDB Validation' section with a Clashscore of 1.80 Å. At the top right of the page, there's a 'Display Files' button and a 'Download Files' button. A red arrow points to the 'Download Files' button. A red circle highlights the 'PDB Format' option in the dropdown menu, which also includes 'PDB Format (gz)', 'PDBx/mmCIF Format', 'PDBx/mmCIF Format (gz)', 'PDBML/XML Format (gz)', and 'Biological Assembly 1'.

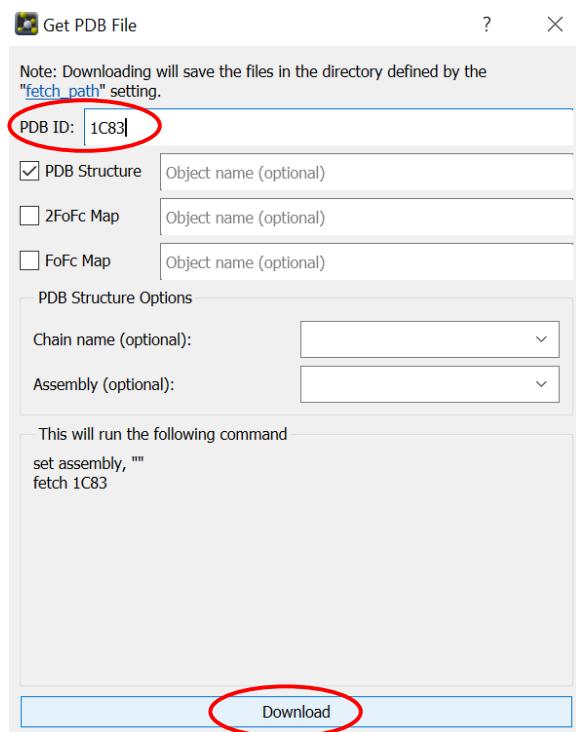
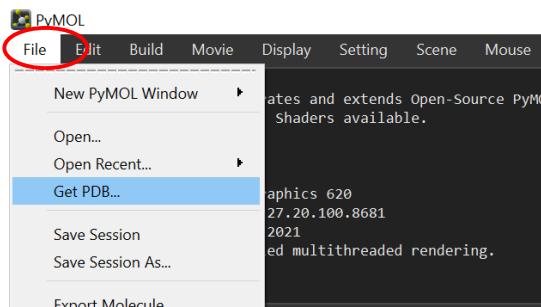
El archivo se descargará y se verá como en la imagen. En **PyMOL**, podemos abrir el archivo desde “File” después “Open”, seleccionamos el archivo de interés y hacemos click en “OK”.





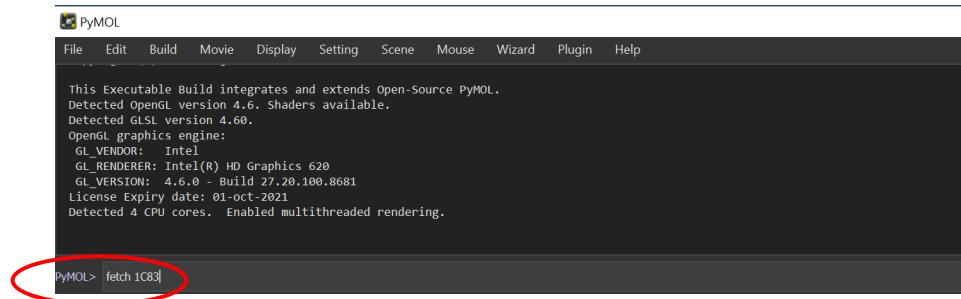
## 12.2 Con PDB Loader Service de PyMOL.

Para esta opción vamos a la pestaña “File”, después “Get PDB...”, escribimos el PDB ID y hacemos click sobre “Download”.

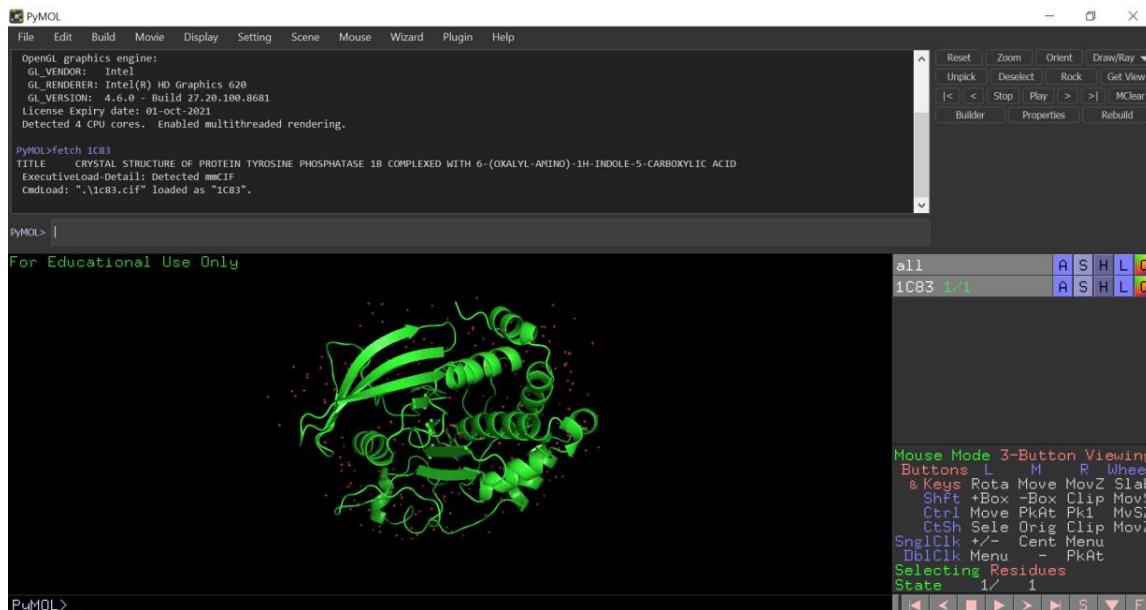


### 12.3 Con comandos en PyMOL.

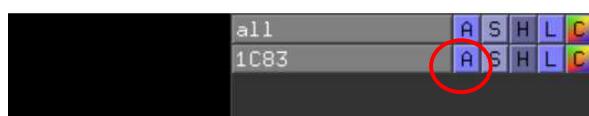
Otra manera de hacerlo (tal vez la más rápida y sencilla) es escribir directamente sobre la barra de comandos de PyMOL la palabra “fetch” seguido de la clave de la proteína y presionamos la tecla INTRO de nuestro teclado. Para este ejemplo escribiremos “**fetch 1C83**”.

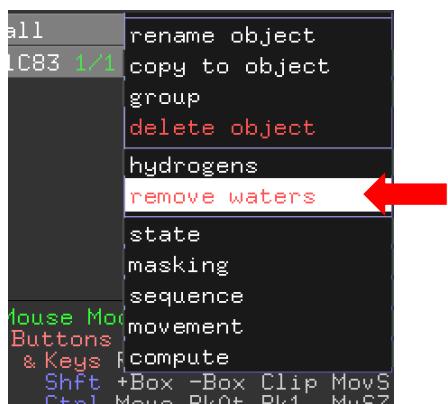


Independientemente de la manera en la que decidamos hacerlo, al final, deberíamos de ver la estructura de nuestra proteína en pantalla.

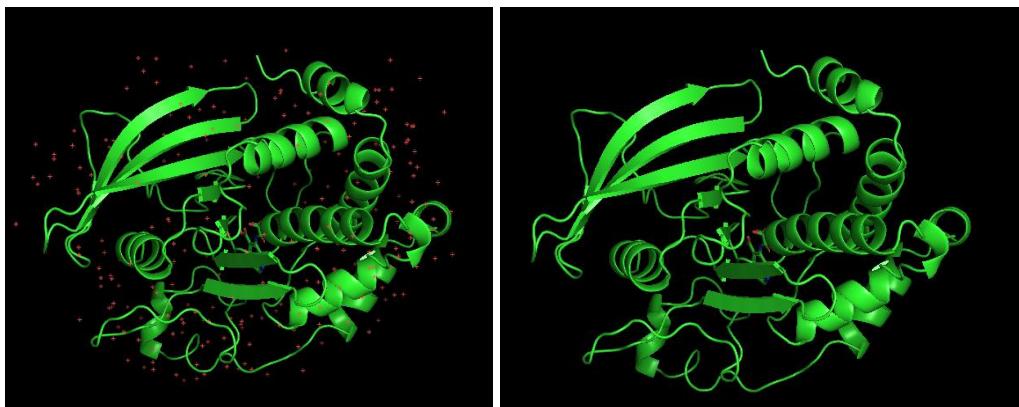


Generalmente, antes de minimizar la proteína a través de un servidor web, hay que removerle las moléculas de agua (representadas como cruces rojas). Esto lo realizaremos haciendo click sobre la “A” del panel derecho, en la barra de 1C83. Posteriormente, haremos click sobre “Remove Waters”.

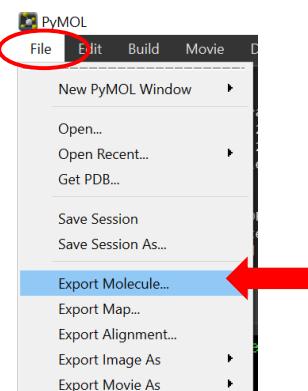




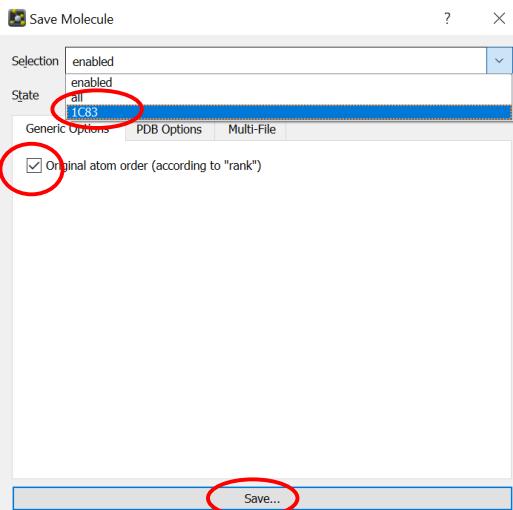
Como vemos, las moléculas de agua se eliminan de nuestra estructura.



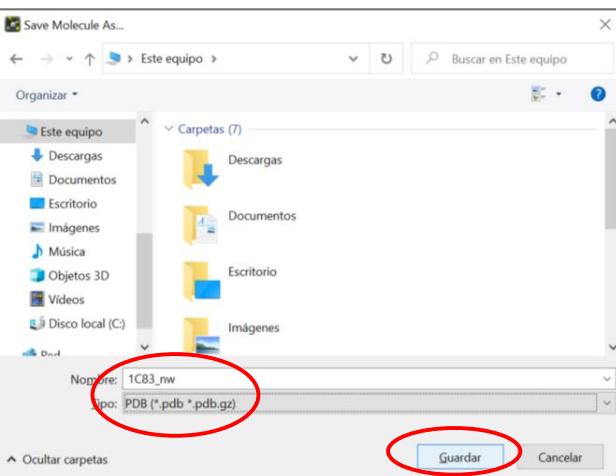
El siguiente paso es guardar la proteína sin moléculas de agua. Para ello, iremos al menú “File” y después a “Export Molecule...”



Posteriormente, saldrá una ventana en donde debemos asegurarnos de que se encuentre seleccionado lo que queremos guardar en la opción de “Selection” (en este caso, 1C83). Debemos marcar la casilla “Original atom order” y hacemos click sobre “Save”.



Después, escogemos el destino del archivo, lo nombramos [una sugerencia es nombrarlo como proteína\_nw.pdb (nw = no waters), por lo que, para este ejemplo, el nombre sería **1C83\_nw.pdb**], verificamos que se vaya a guardar en formato \*.pdb y hacemos click sobre “Guardar”.

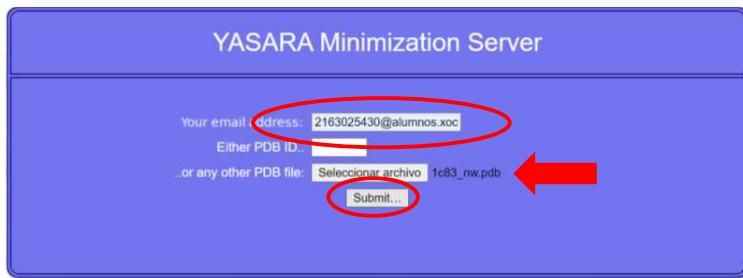


#### 12.4 Cómo minimizar una proteína

Ahora nuestra proteína está lista para ser minimizada. Esta minimización la realizaremos con el servidor web de **YASARA**. Por lo tanto, el primer paso es entrar a la página web <http://www.yasara.org/minimizationserver.htm>. Una vez ahí, veremos la siguiente pantalla:



Para utilizar este servidor, solamente necesitamos ingresar un correo electrónico (que es a donde nos enviarán los resultados) y subir el archivo de la proteína sin moléculas de agua. Una vez que hayamos agregado lo que se nos pide, simplemente hacemos click sobre “Submit...”.



YASARA envía dos correos, uno de confirmación y otro con los resultados. El correo de confirmación llega unos minutos después de haber subido el archivo de la proteína y se ve como en la imagen.

YASARA force field minimization server: your job has been queued Recibidos x

Eliza@yasara.org para mí ▾ 10:14 (hace 6 horas) ⭐ ↗ ⋮

Dear user of the YASARA force field minimizer,

thanks for using our server.

Your job has been put into the queue, the waiting time can range from minutes to days depending on the server load. Please do not submit a job multiple times. The results will be mailed to [2163025430@alumnos.xoc.uam.mx](mailto:2163025430@alumnos.xoc.uam.mx).

Best regards,  
the YASARA team.

El correo con los resultados tarda de horas a días en llegar e incluye una liga en la que podemos descargar el archivo de nuestra proteína minimizada (señalado con rojo).

YASARA force field minimization server: your job has been completed Recibidos x

**Eliza@yasara.org**  
para mí ▾

12:01 (hace 4 horas) ⚡ ↗ ⋮

Dear user of the YASARA force field minimizer,

many thanks for submitting your model to our server.

You can download the result from <http://download.yasara.org/server/minimization/00021988.sce>. Chrome has a bug which may require to right-click and 'Save as'. If your email client fails to handle the download, paste the link into your browser or use 'wget'.

\*) If you publish results obtained with this server, please cite 'Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K (2009) Proteins 77 Suppl 9, 114-122 PMID19768677'.

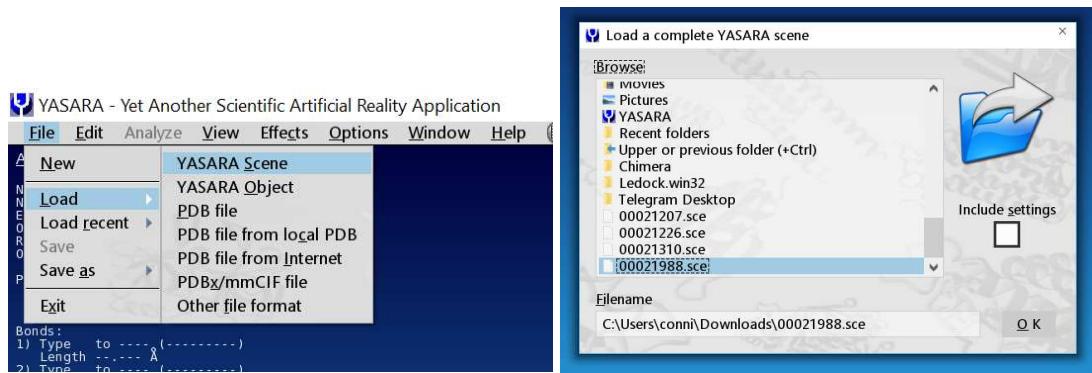
\*) The file above is an annotated YASARA scene, which you can open with the free YASARA View program (<http://www.yasara.org/viewdl>) by clicking 'File > Load > YASARA scene'.

\*) In case you want a different file format, click 'File > Save as' to export the structure. To get a PDB file, click for example 'File > Save as > PDB file', double-click on the 'End' object in the list on the left (to save the end conformation), then choose your filename and click OK. You can include the water shell by clicking on 'End' only once, holding down 'Ctrl' and selecting the 'Water' object too.

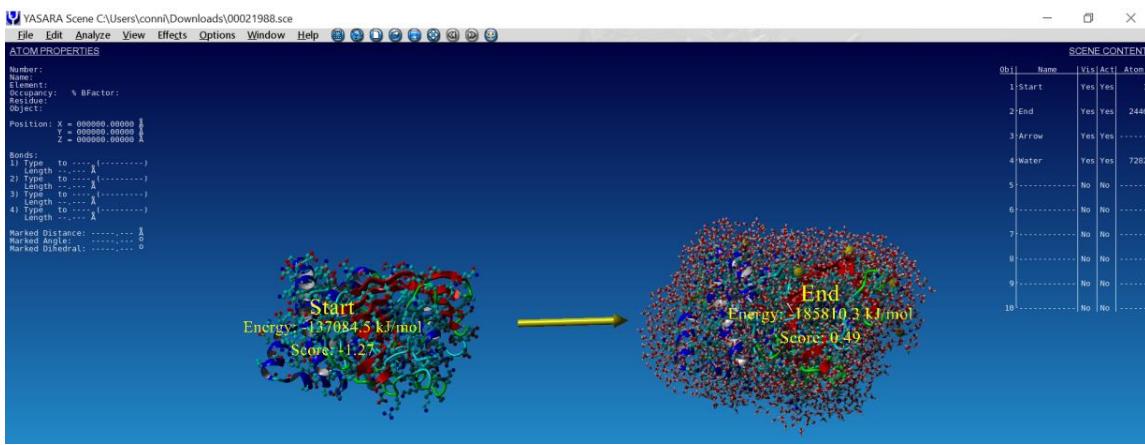
El archivo con los resultados es un archivo \*.sce



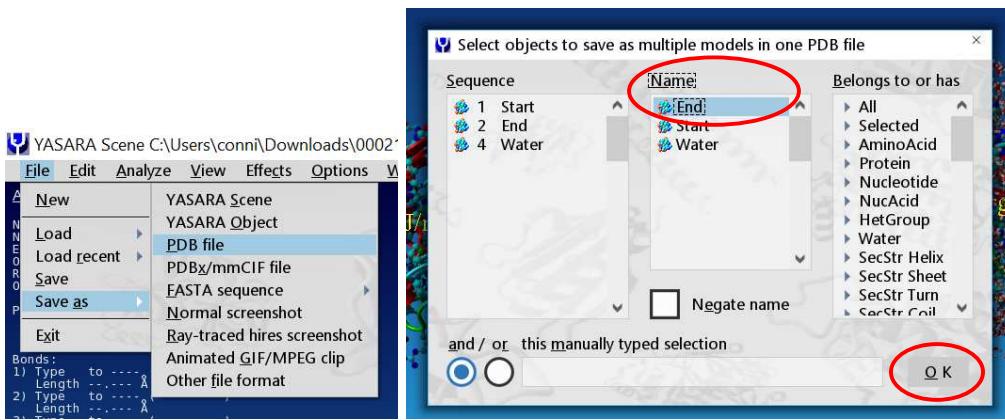
Posteriormente, abriremos este archivo en el programa de **YASARA View** con “File”, “Load”, “YASARA Scene”, seleccionamos nuestro archivo y hacemos click en “OK”



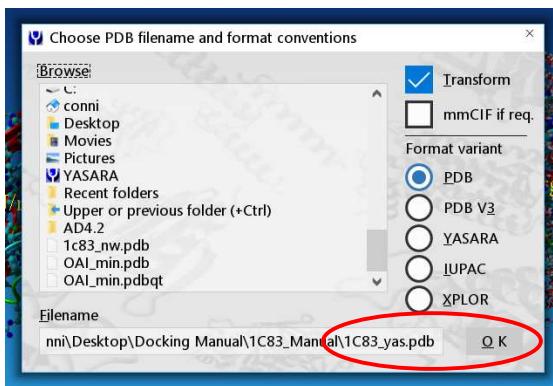
Una vez que abrimos el archivo, observaremos el estado pre y post minimización. Podemos ver que la minimización fue favorable porque la energía en la forma post minimización es menor que en la pre minimización. Esto quiere decir que los átomos de la proteína tienen menos interacciones desfavorables entre sí en la forma minimizada. Adicionalmente, podemos ver que el *score* mejoró, pasó de -1.27 a 0.49.



El siguiente paso es guardar la proteína minimizada en formato \*.pdb. Para esto, nos dirigimos a “File”, “Save as”, “PDB file”. Aparecerá una ventana como en la imagen de la derecha. Seleccionamos “END” en la ventana de “Name” y hacemos click sobre “OK”.

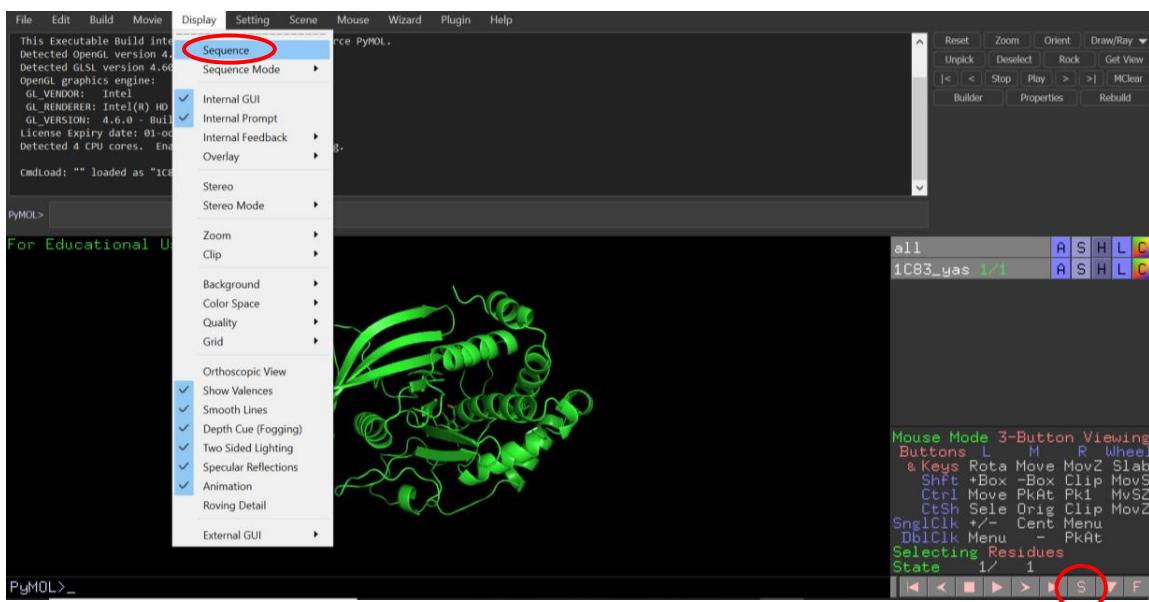


Finalmente, seleccionamos el destino del archivo y lo renombramos, por ejemplo, como proteína\_yas.pdb (yas = YASARA). En este caso, **1C83\_yas.pdb**.

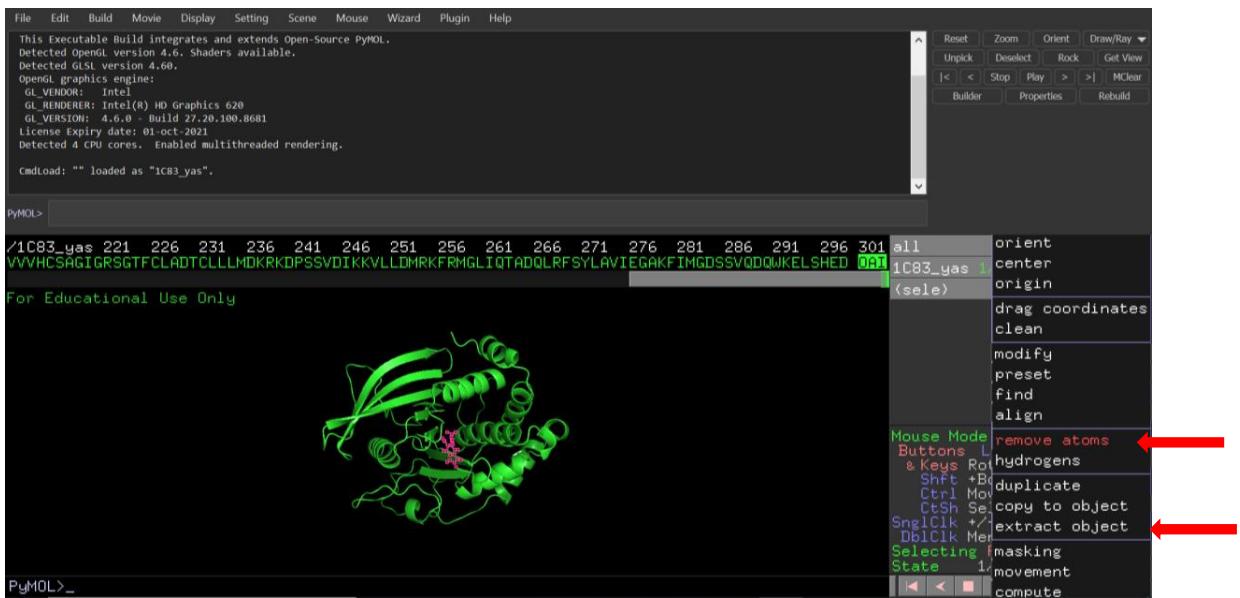


Recordemos que esta proteína todavía tiene a su ligando co-cristalizado. Para realizar los *dockings*, necesitamos solamente a la proteína, por lo tanto, debemos remover el ligando. Esto se puede lograr en **PyMOL**, de la siguiente manera.

Una vez que hemos abierto PyMOL, abrimos el archivo **1C83\_yas.pdb**. Posteriormente, debemos ver la secuencia de aminoácidos y ligandos que tiene nuestra proteína. Esta secuencia la podemos mostrar al hacer click en la pestaña “Display” con la opción “Sequence”, o bien sobre el botón “S” que se encuentra cerca de la esquina inferior derecha de la pantalla.



Con cualquiera de las dos opciones, nos aparecerá la secuencia de aminoácidos, ligandos y cofactores incluidos en la molécula en una barra en la parte superior de la pantalla, por la cual podemos desplazarnos. Generalmente, al final de la barra es en donde encontraremos a las moléculas que no sean aminoácidos, como los ligandos. Entonces, es ahí a donde tenemos que dirigirnos para remover al ligando. Como vimos desde la página de RCSB, el ligando tiene clave “**OAI**”, entonces una vez que lo encontramos, debemos seleccionarlo haciendo click sobre él. Posteriormente, del lado derecho saldrá una nueva barra “Sele”, que representa lo que hemos seleccionado (en este caso, el OAI). Hacemos click sobre “A” y después sobre “remove atoms” o “extract object”.

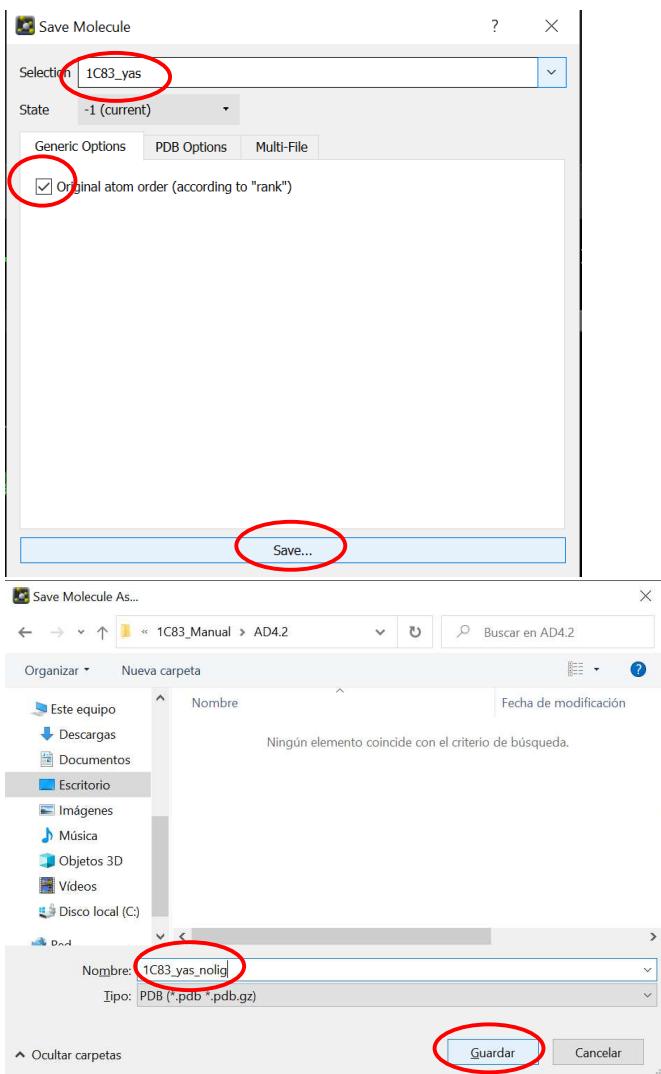


Si seleccionamos “remove atoms”, la selección (en este caso, el ligando) se borrará. Si seleccionamos la opción “extract object”, la selección se separará del resto de la molécula como si fueran dos “archivos” distintos. Guardar al ligando co-cristalizado y a la proteína en dos archivos separados puede ser de gran utilidad pues, podemos utilizar el archivo de la proteína para realizar cálculos de *docking*, mientras que el archivo del ligando co-cristalizado nos puede servir más adelante para validar el protocolo de docking que estamos planteando por medio de la determinación del RMSD entre el ligando co-cristalizado y el ligando *dockeado*. Por esta razón, en este ejercicio, seleccionaremos la opción “**extract object**”.

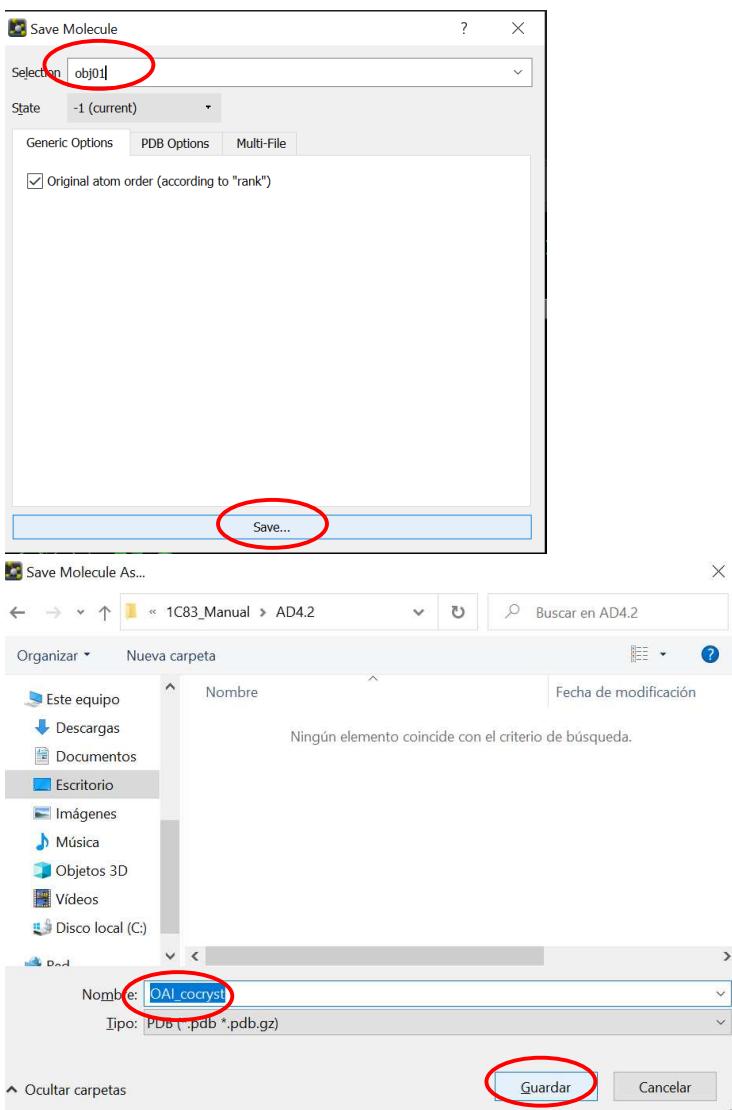
Una vez que hacemos click sobre esta opción, aparecerá una nueva barra en el panel del lado derecho con el nombre de “obj01”.



Por último, tenemos que guardar ambos archivos. Primero guardaremos a la proteína. Para ello, vamos a “File”, “Export molecule”, saldrá un menú como en la imagen, seleccionamos a la proteína y hacemos click sobre “Save”. Después escogemos el destino del archivo, lo nombramos de manera que distingamos entre la proteína con ligando y la que no tiene el ligando y damos click en “Guardar”. En este ejercicio, la nombramos como **1C83\_yas\_nolig.pdb** (donde *nolig* = no ligand).



Para guardar al ligando seguimos los mismos pasos que anteriormente. La única modificación será que debemos escoger “obj01”. De igual manera, escogemos destino, nombramos y guardamos. En este ejercicio, lo nombramos como “**OAI\_cocryst.pdb**” (donde *cocryst* = cocrystallized).



Hasta este punto, ya tenemos a nuestra proteína en formato \*.pdb. Este formato es compatible con GOLD. Sin embargo, para AD4.2 y VINA, es necesario transformarlo a formato \*.pdbqt. Ese paso lo veremos en el siguiente tema denominado “Docking en Autodock4.2”, en la sección “Generar el archivo \*.gpf” y “Generar el archivo \*.dpf”.

## 13 Docking en Autodock4.2

### 13.1 ¿Qué necesitamos?

Para correr un docking con AD4.2 son necesarios 4 archivos:

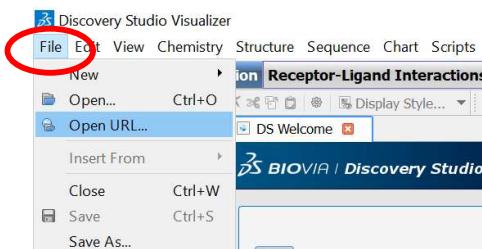
- PDBQT del ligando
- PDBQT de la proteína
- Archivo grid.gpf
- Archivo dock.dpf

### 13.2 Visualización de aminoácidos del sitio de unión

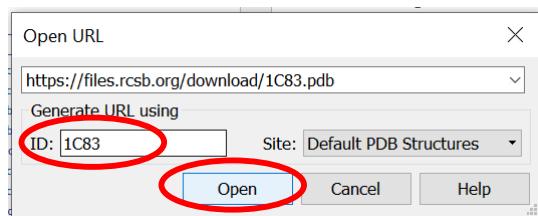
Como vimos en la parte de teoría, para realizar un docking es necesario determinar el espacio de la proteína en el que se hará el cálculo. Este espacio se conoce como **grid**.

El primer paso para generar un grid es saber qué residuos de aminoácidos de la proteína interactúan con el ligando. Esto lo podemos observar fácilmente con ayuda del programa **Discovery Studio**.

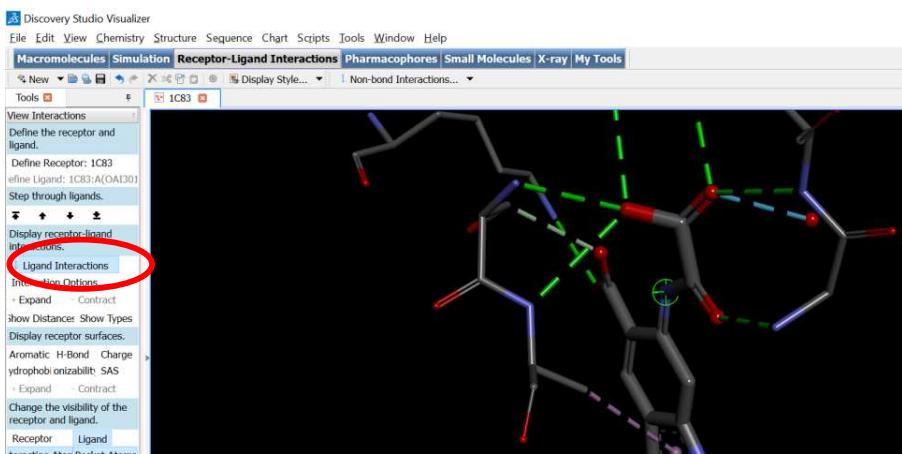
Abrimos el programa y posteriormente cargamos la proteína de interés (1C83) directamente desde PDB, con “File” y “Open URL...”.



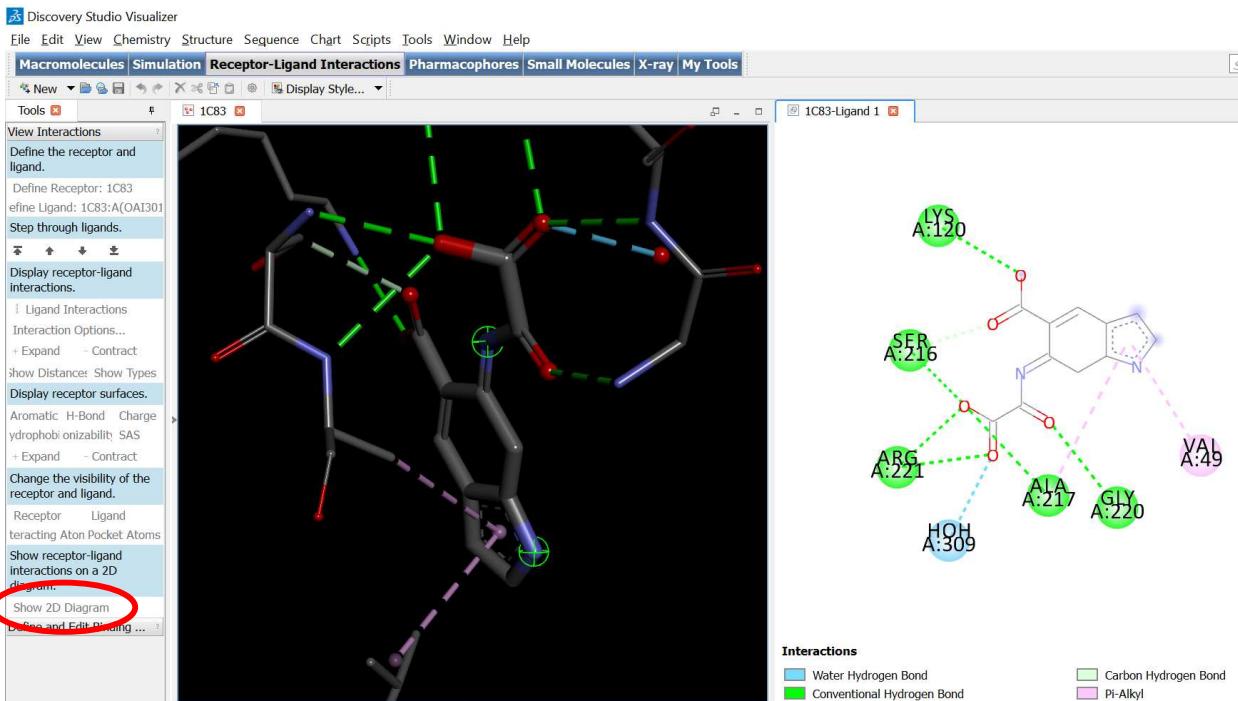
Se abrirá una pequeña ventana en la que debemos escribir el PDB ID de nuestra proteína en la caja de “ID” (para este ejercicio, 1C83) y hacemos click en “Open”.



Posteriormente, hacemos click en el panel del lado izquierdo, sobre “Ligand interactions”. Esta acción centrará a la proteína en el sitio de interacción con el ligando.



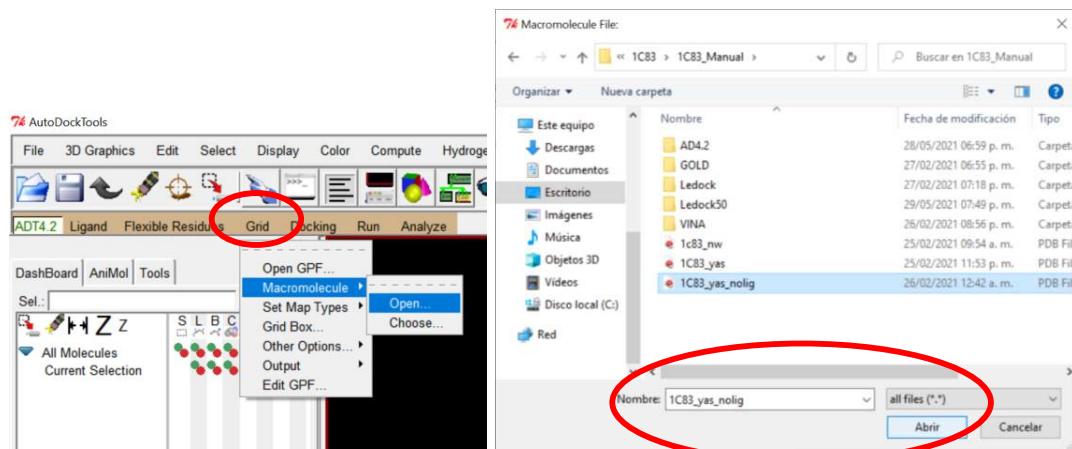
Después, para visualizarlo mejor, hacemos click del lado izquierdo hasta abajo sobre “Show 2D diagram”. Al hacerlo, nos aparecerá del lado derecho de la pantalla una representación más sencilla de las interacciones ligando-proteína. Es necesario que apuntemos los residuos de aminoácidos que se indican, porque nos servirán más adelante.



### 13.3 Preparación en ADT

Los siguientes pasos los realizaremos en el programa **ADT**. Una vez que esté abierto, lo primero es convertir nuestra proteína en formato \*.pdb a formato \*.pdbqt. Para ello, debemos ir a “Grid”, “Macromolecule”, “Open”. Por default aparecen solamente los archivos \*.pdbqt

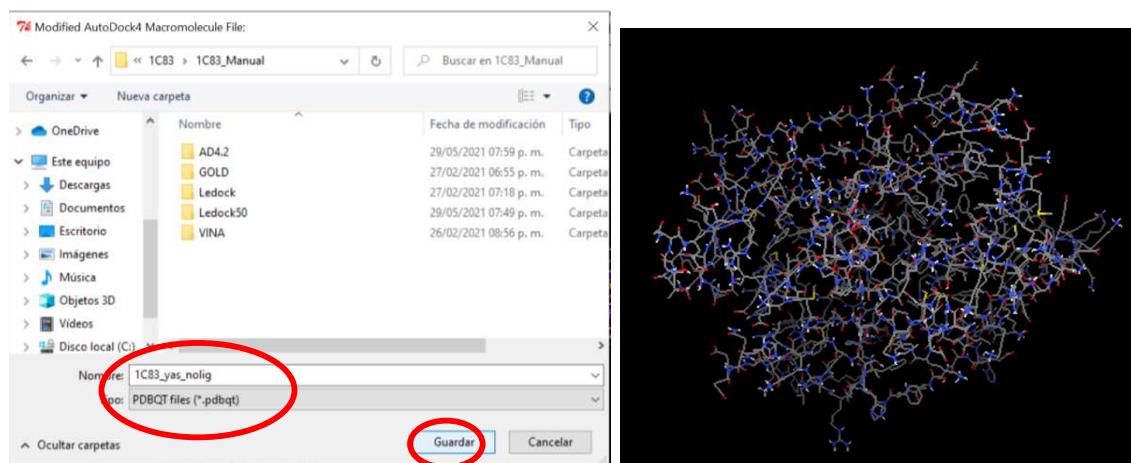
por lo cual debemos cambiar a “all files” para encontrar el archivo que necesitamos. Lo seleccionamos y hacemos click en “Abrir”.



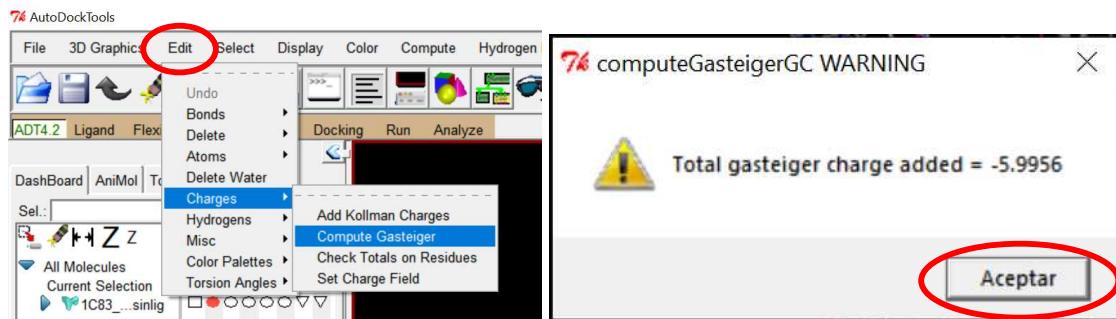
Cuando se abre el archivo, nos saldrá un mensaje como el de la imagen. Damos click en “Aceptar”



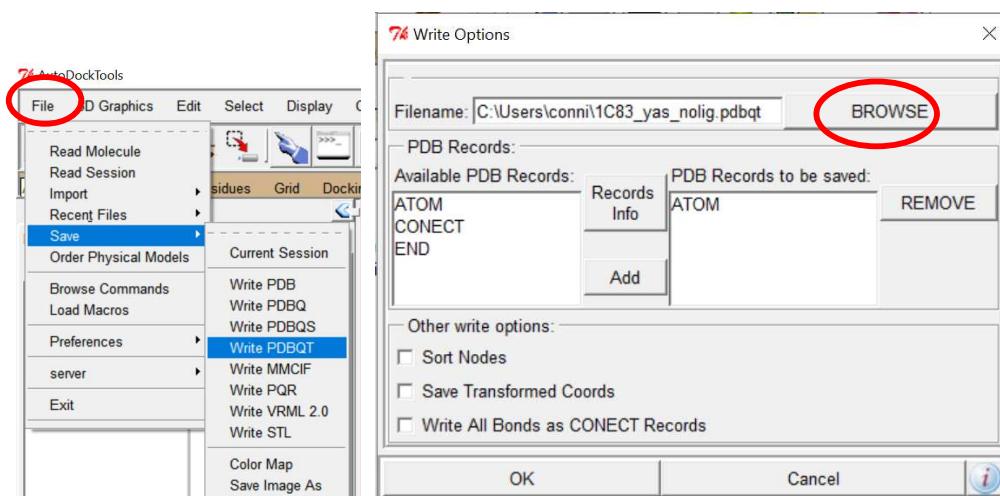
Inmediatamente después, saldrá una ventana para guardar nuestro archivo ahora en formato \*.pdbqt. Debemos simplemente seleccionar el destino del archivo y hacer click sobre “Guardar”. Una vez que lo guardamos, vemos que la proteína pasó de color rosa a gris.

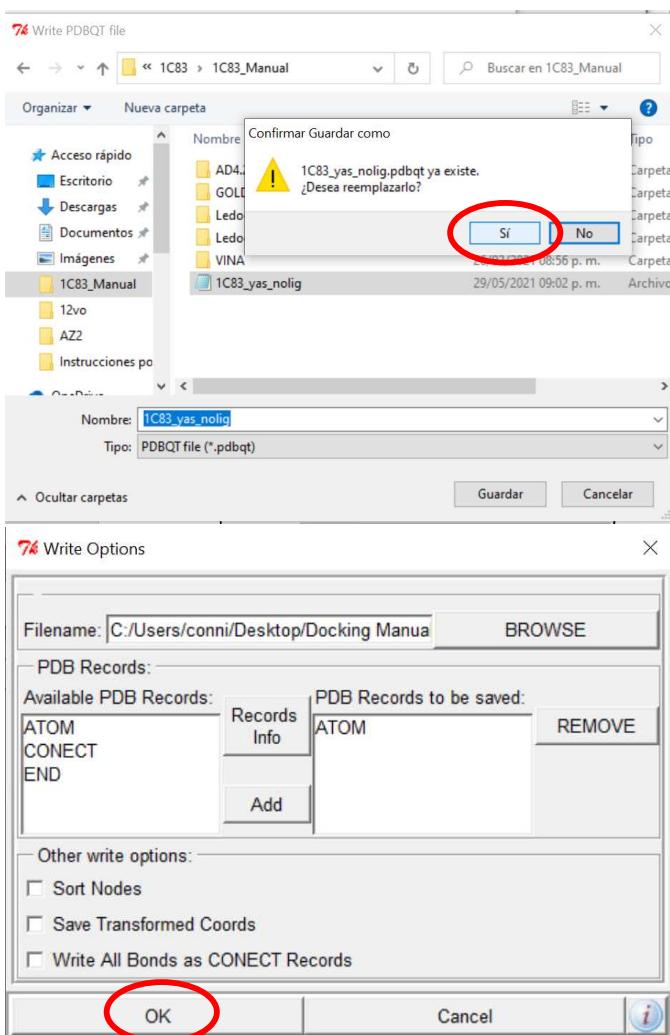


El siguiente paso es añadir cargas de Gasteiger. Para hacerlo, iremos a la pestaña “Edit”, “Charges” y “Compute Gasteiger”. Saldrá un mensaje que nos indicará cuántas cargas se agregaron y hacemos click en “Aceptar”.



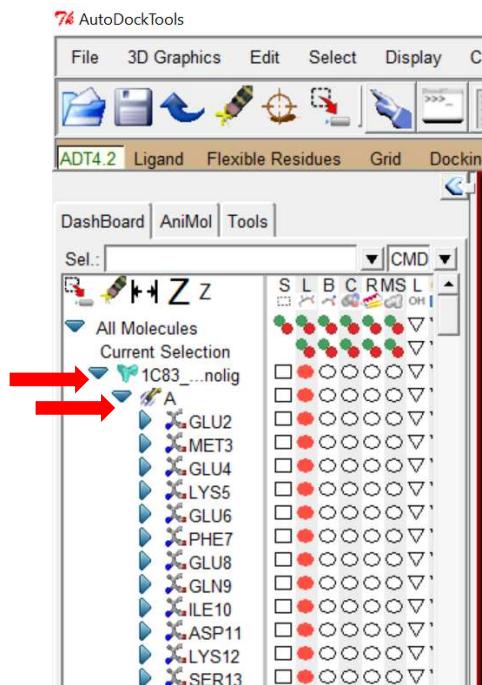
Como el primer \*.pdbqt que generamos de la proteína no tenía las cargas de Gasteiger, debemos de sobreescibirlo con el cambio que acabamos de realizar. Para esto, iremos a “File”, “Save”, “Write PDBQT”, buscamos el primer archivo \*.pdbqt que generamos, lo sobreescrivimos y hacemos click sobre “OK”.



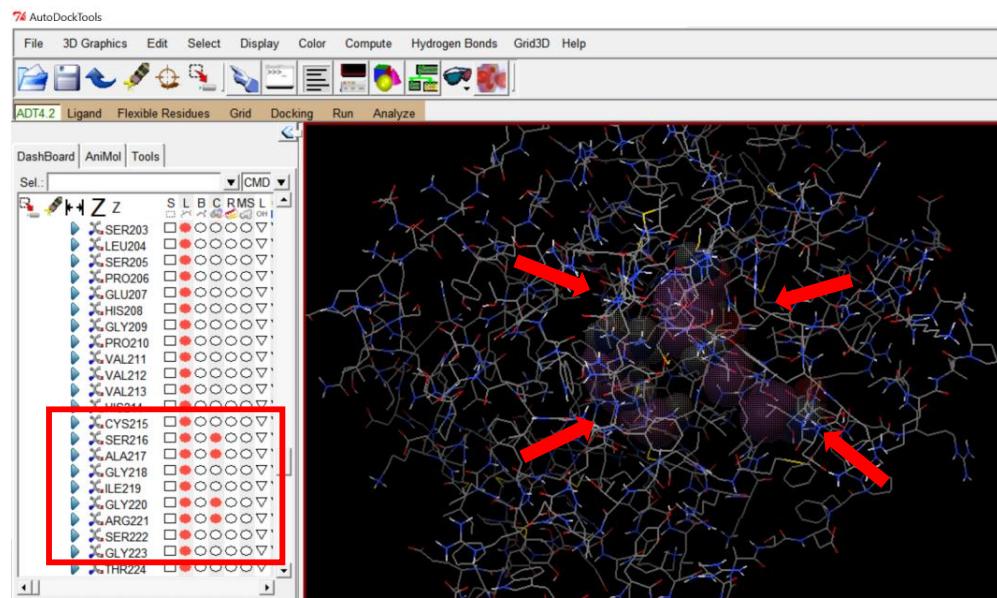


### 13.4 Generar el archivo \*.gpf

Lo siguiente es generar el **archivo \*.gpf**. Este archivo se genera al determinar en qué zona y de qué tamaño será nuestro *grid*. Es en este paso en donde necesitaremos saber los residuos de aminoácidos que apuntamos anteriormente. Para ver la secuencia de aminoácidos de la proteína, hacemos click sobre “>” y posteriormente desplegamos la lista.

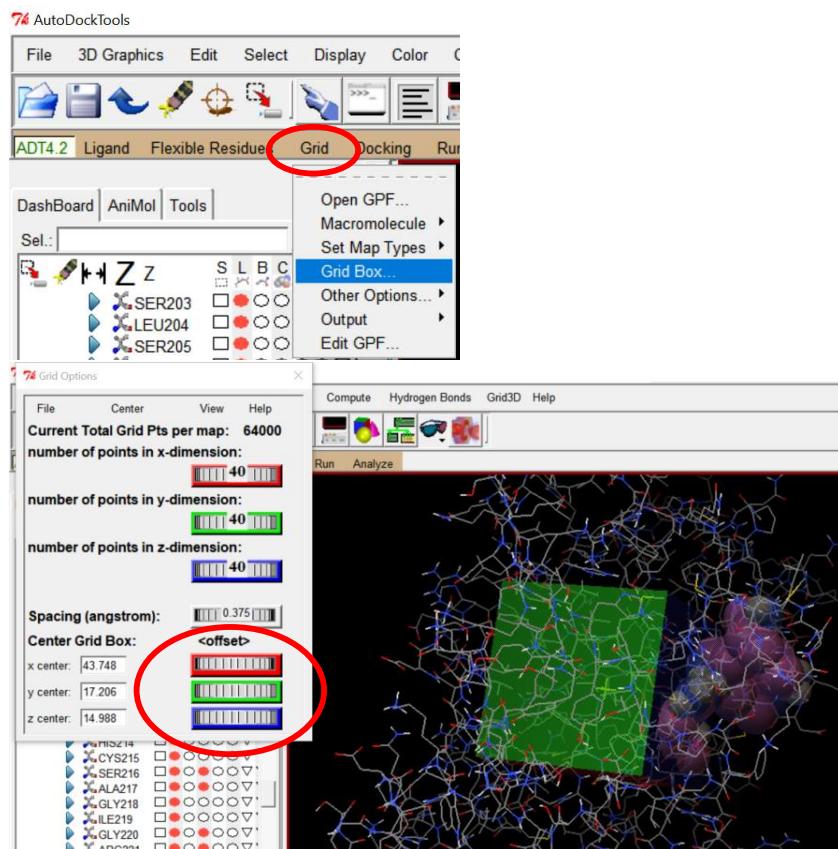


Después, buscamos en la lista los aminoácidos con los que el ligando interactúa y damos click en el tercer círculo a la derecha. Veremos que, al seleccionarlos en el panel de la izquierda, aparecerán también del lado derecho, representados como esferas de color rosa.

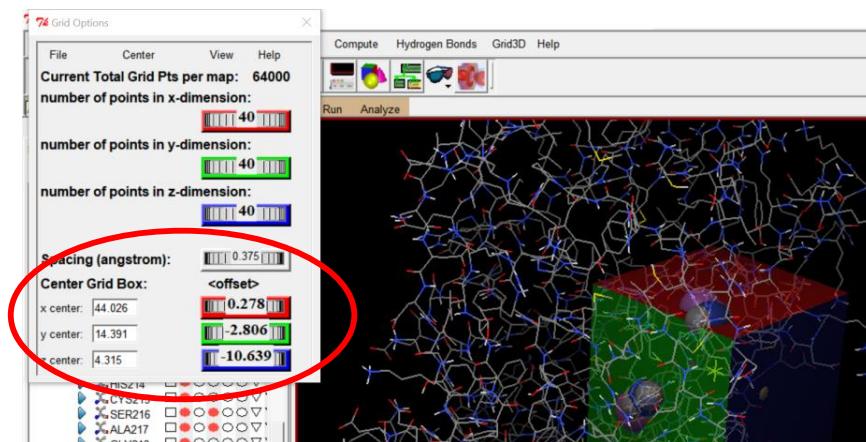


Esta selección que hicimos nos ayudará para poder centrar el *grid* en esa zona. Ahora iremos a “Grid”, “Grid box” y aparecerá una ventana como la que se ve a continuación. Con las

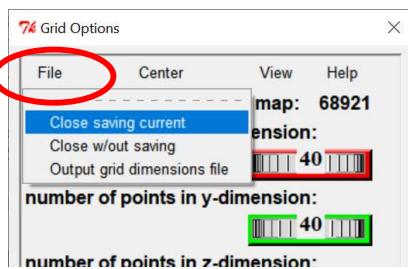
perillas de abajo (señaladas con un círculo rojo), podemos mover manualmente el *grid* (representado como un cubo de colores) e intentar centralizarlo en el área que corresponde.



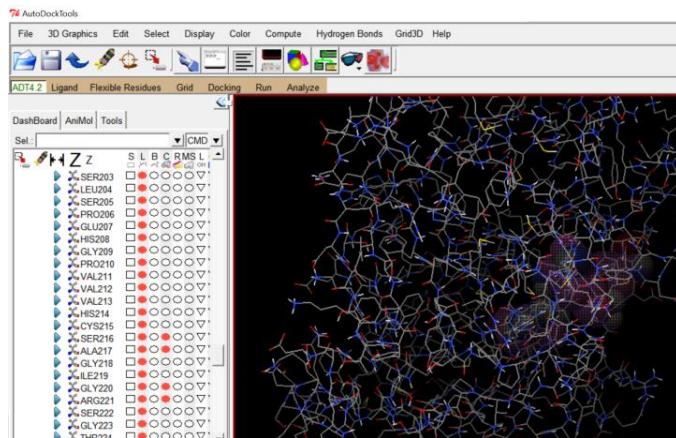
Una vez que centremos el *grid*, este debería de cubrir la mayor parte posible del área que señalamos. Es recomendable apuntar estas coordenadas x, y, z (en este ejemplo son x = 44.026, y = 17.206, z = 14.988) porque nos van a servir para trabajar en los programas VINA y GOLD.



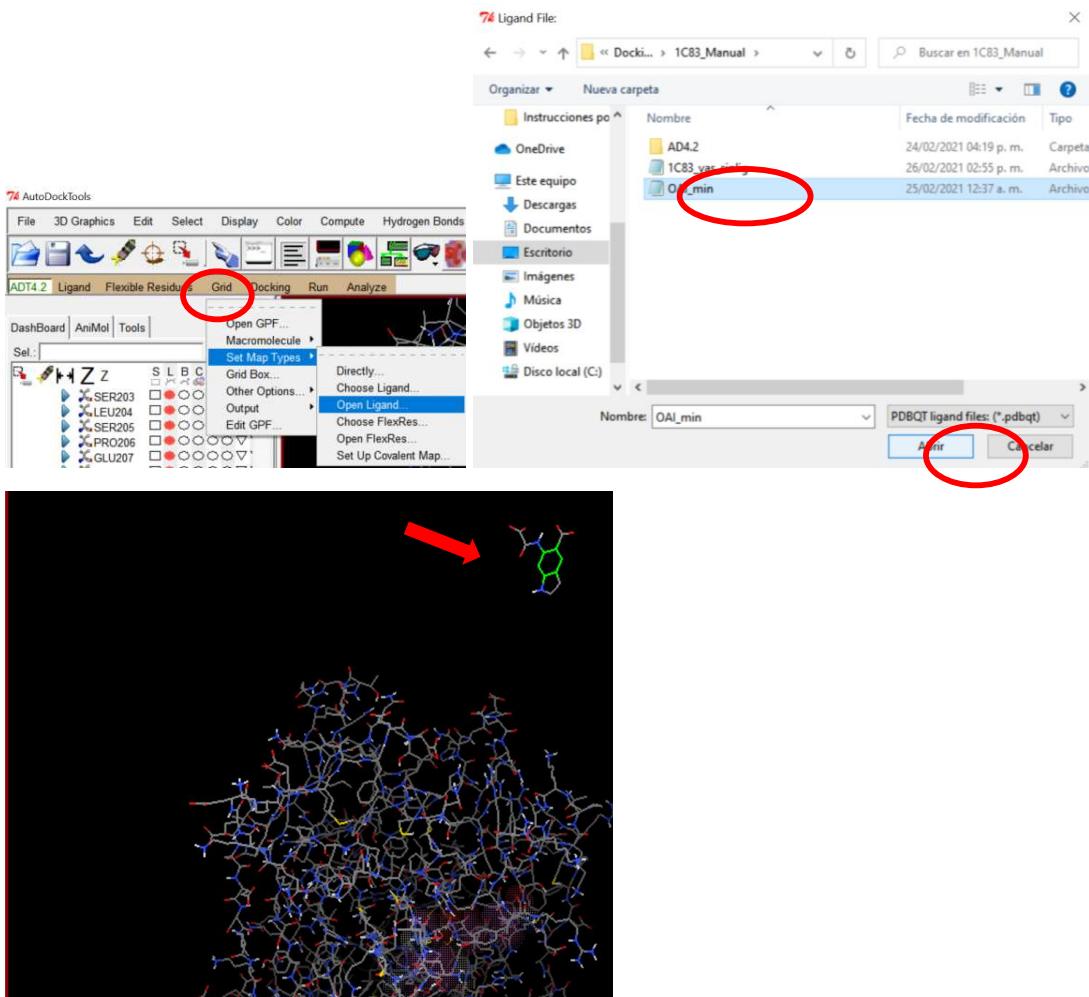
Para guardar el *grid*, en esa misma ventana, debemos ir a “File”, “Close saving current”.



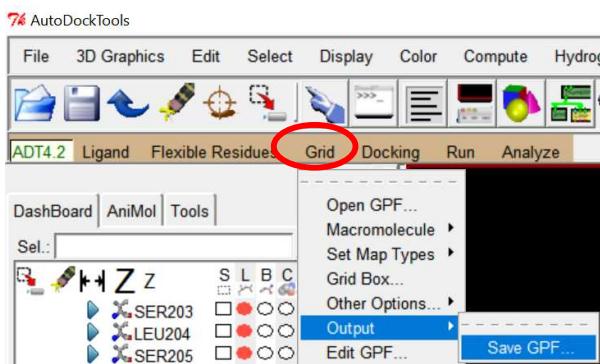
La ventana y el *grid* (el cubo de colores) desaparecerán, pero las coordenadas se habrán guardado.

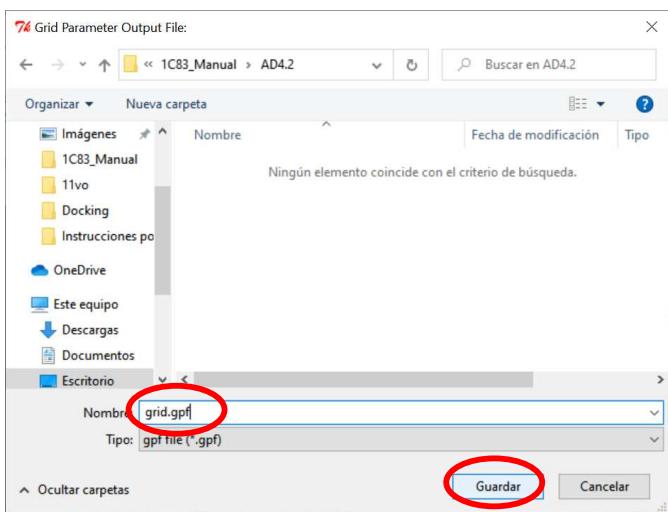


Ya que hemos determinado las coordenadas y el tamaño del *grid*, debemos ahora mencionar los tipos de átomos que tiene el ligando con el que realizaremos el *docking*, en este caso, el OAI. Para ello, iremos a “Grid”, “Set Map Types” y “Open ligand”. Seleccionamos el archivo \*.pdbqt de nuestro ligando minimizado y damos click en “Abrir”. Entonces, veremos la estructura del ligando en un color diferente al de la proteína (en este ejercicio, de color verde).



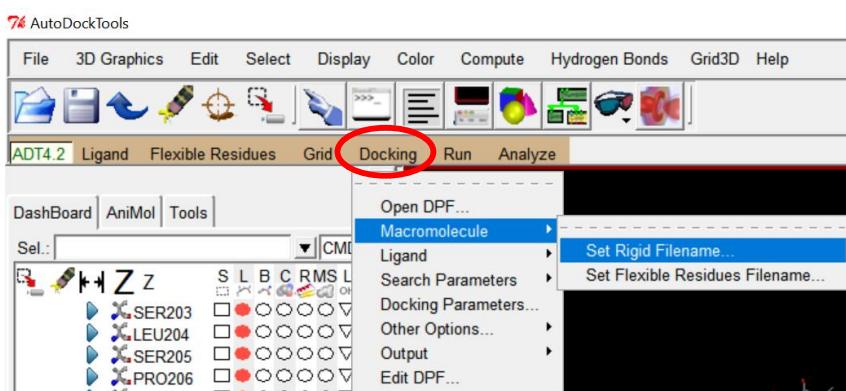
Con ello estaremos estableciendo las coordenadas en donde se realizará el docking, como el tipo de átomos que tiene nuestro ligando. Ahora para guardar el archivo \*.gpf (el *grid*) que hicimos, debemos ir a “Grid”, “Output”, “Save GPF”. Escogemos el destino, nombramos el archivo (es recomendable guardararlo como **grid.gpf**) y hacemos click sobre “Guardar”.

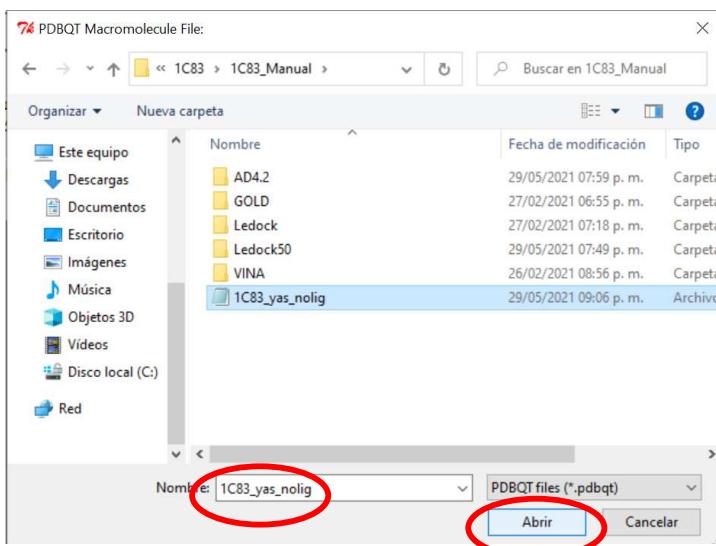




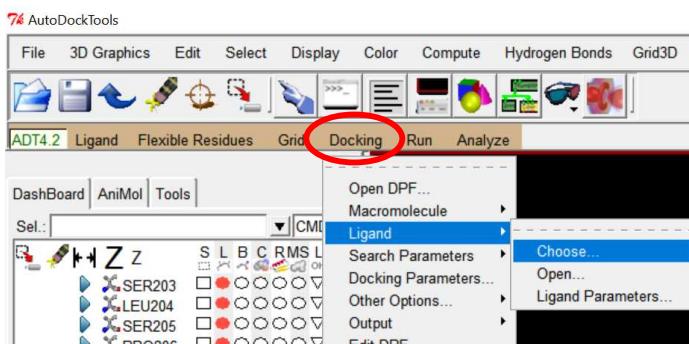
### 13.5 Generar el archivo \*.dpf

Para generar el archivo \*.dpf debemos ir a la pestaña “Docking”, “Macromolecule”, “Set Rigid Filename”. Escogemos el archivo \*.pdbqt de nuestra proteína minimizada y hacemos click en “Abrir”.





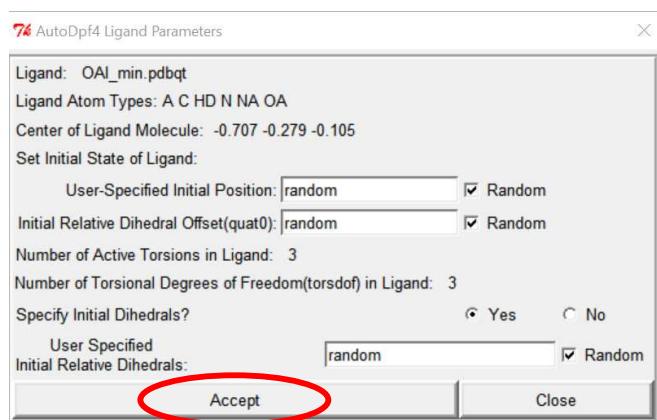
Parece que no hubiera pasado nada, pero sí se hizo la selección. El siguiente paso es escoger el ligando. Como ya tenemos abierto ese archivo porque lo usamos para el \*.gpf, podemos ir a “Docking”, “Ligand”, “Choose”. (Si, por el contrario, no tuviéramos abierto ese archivo, simplemente escogeríamos “Open” en lugar de “Choose”.)



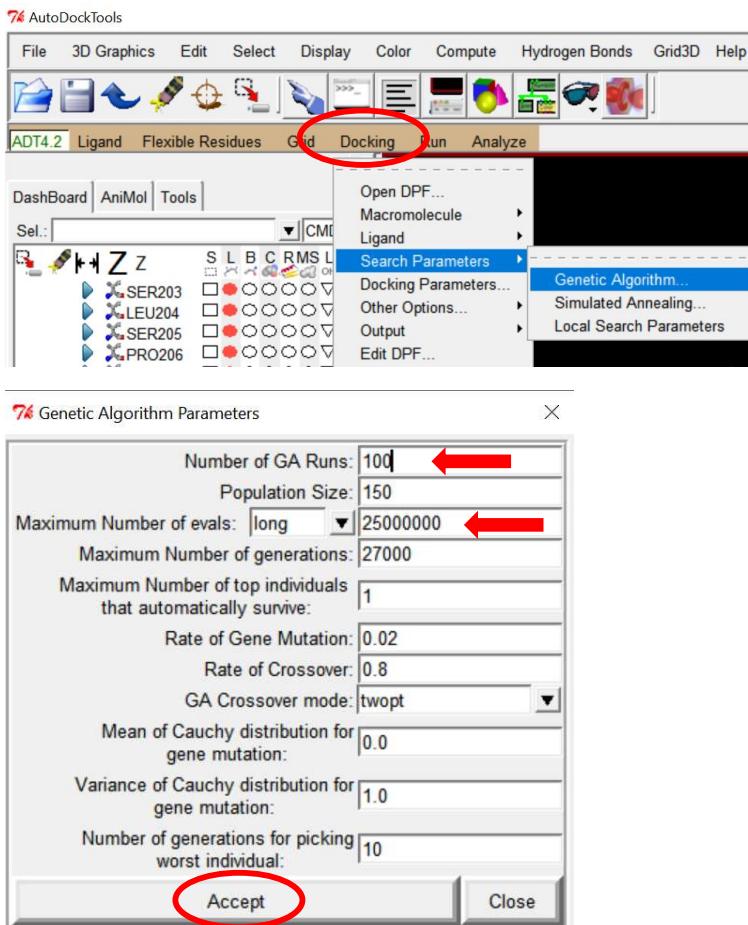
Se abrirá una ventana en donde tenemos que seleccionar a nuestro ligando y posteriormente hacer click sobre “Select ligand”.



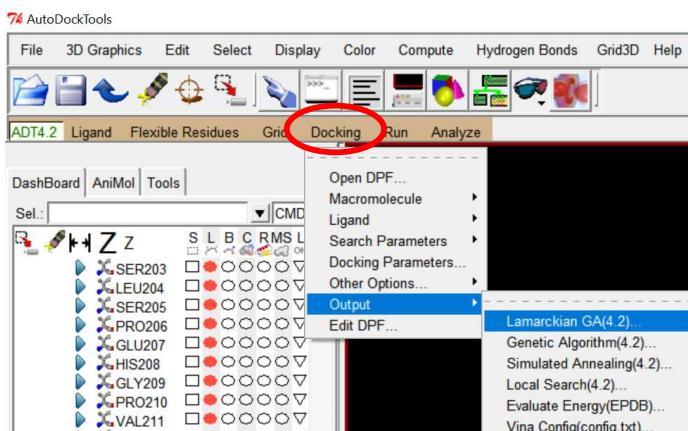
Posteriormente saldrá una ventana con algunos datos. En esta, solamente hacemos click sobre “Accept”.

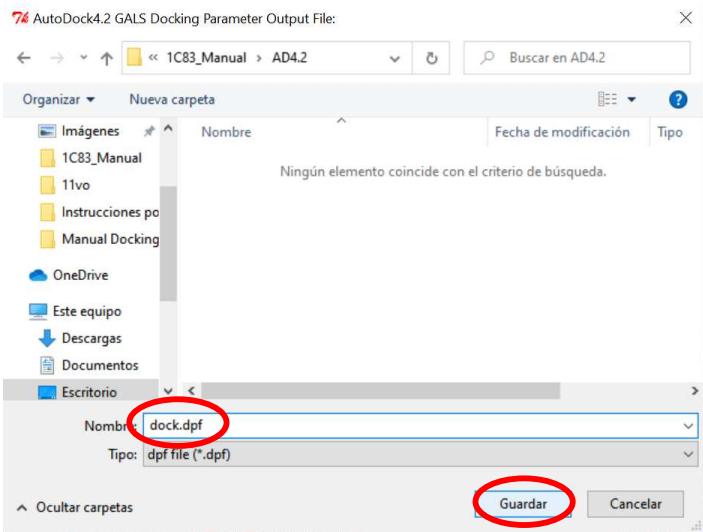


Después, nuevamente vamos a la pestaña de “Docking”, “Search Parameters” y “Genetic Algorithm”. Se desplegará una ventana con distintos parámetros. En “Number of GA Runs” es recomendable establecer el valor mínimo en “50”, pero preferentemente en “100”. También debemos corroborar que en “Maximum Number of evals” esté seleccionada la opción “long”. Todos los demás parámetros se dejan igual y se da click en “Accept”.



Finalmente, debemos guardar el archivo \*.dpf que generamos seleccionando a la pestaña “Docking”, “Output”, “Lamarckian GA (4.2)”. Escogemos el destino, nombramos al archivo (es recomendable nombrarlo **dock.dpf**) y hacemos click en “Guardar”.





En este punto ya tenemos todos los archivos que necesitamos. El siguiente paso es simplemente ejecutar el cálculo. Aprenderemos cómo en la siguiente sección “Para ejecutar el cálculo”.

### 13.6 *Para ejecutar el cálculo*

Es necesario tener bien organizados los archivos que generamos previamente, entonces es aconsejable crear una carpeta para cada ligando que vayamos a *dockear*. (En este caso solo es uno, pero si fuéramos a hacer *docking* de más de un ligando, se recomienda que sea una carpeta por ligando). También es aconsejable ponerle el nombre del ligando (sin usar espacios) a la carpeta para agilizar y simplificar su identificación.

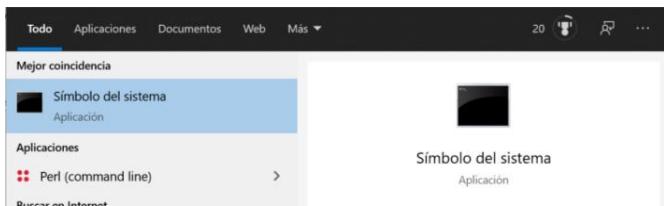
En esta carpeta, deben estar los cuatro archivos mencionados anteriormente y los dos ejecutables de AD4.2. Estos ejecutables se pueden descargar de la referencia <sup>19</sup>.

Por lo tanto, la carpeta debería verse así:

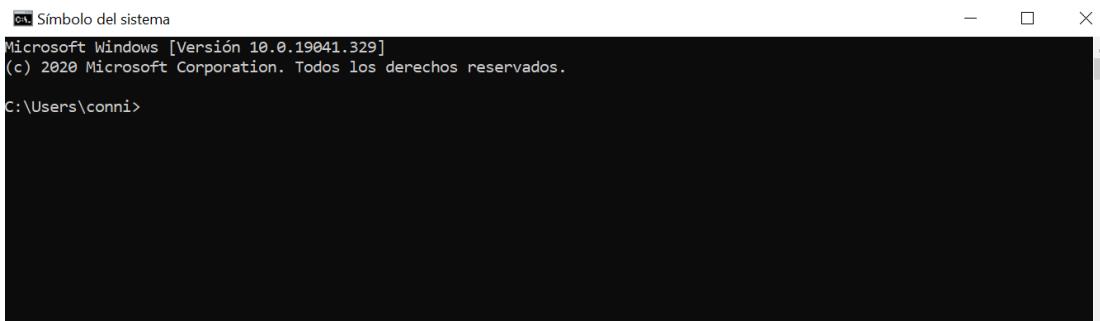
Nombre	Fecha de modificación	Tipo	Tamaño
1C83_yas_nolig	29/05/2021 09:06 p. m.	Archivo PDBQT	236 KB
autodock4	09/02/2020 03:00 a. m.	Aplicación	605 KB
autogrid4	09/02/2020 03:00 a. m.	Aplicación	1,883 KB
dock	26/02/2021 05:30 p. m.	Archivo DPF	3 KB
grid	26/02/2021 04:57 p. m.	Archivo GPF	2 KB
OAI_min	25/02/2021 12:37 a. m.	Archivo PDBQT	2 KB

### 13.7 Generar el archivo \*.glg

El primer paso es abrir el “Símbolo del sistema” (Se puede buscar en inicio como CMD).



Una vez abierto, se ve como en la imagen.



El siguiente paso es generar el archivo \*.glg. Para ello, nos serán de utilidad distintos comandos, como (→ representa un “INTRO”):

cd ^ nombre de la carpeta → para entrar a una carpeta (los ^ representan un espacio).

cd .. → para salir de una carpeta.

La tecla “Tab” → para completar el nombre de una carpeta o archivo.

dir → para desplegar el contenido de una carpeta.

Entonces, primero hay que ubicarnos dentro de la carpeta que corresponda al ligando que queremos dockear. En este ejemplo, dicho ligando se encuentra en la carpeta “AD4.2”, por lo que tenemos que llegar a ella ejecutando el comando *cd* y el nombre de las carpetas de interés, hasta llegar a “AD4.2”.

```
cmd Símbolo del sistema
Microsoft Windows [Versión 10.0.19041.804]
(c) 2020 Microsoft Corporation. Todos los derechos reservados.

C:\Users\conni>cd Desktop
C:\Users\conni\Desktop>cd "Docking Manual"
C:\Users\conni\Desktop\Docking Manual>cd 1C83_Manual
C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd AD4.2
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\AD4.2
```

Una vez dentro de la carpeta, escribiremos el siguiente comando para obtener el archivo \*.glg  
**autogrid4.exe^ -p^grid.gpf^ -l^grid.glg** (los ^ representan un espacio) seguido de la tecla *intro*. Este cálculo tarda unos cuantos minutos en realizarse.

```
cmd Símbolo del sistema - autogrid4.exe -p grid.gpf -l grid.glg
Microsoft Windows [Versión 10.0.19041.804]
(c) 2020 Microsoft Corporation. Todos los derechos reservados.

C:\Users\conni>cd Desktop
C:\Users\conni\Desktop>cd "Docking Manual"
C:\Users\conni\Desktop\Docking Manual>cd 1C83_Manual
C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd AD4.2
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\AD4.2>autogrid4.exe -p grid.gpf -l grid.glg
```

Una vez que el cálculo termina, aparece de nuevo la línea de comandos (señalada en rojo), que nos indica que ya podemos escribir un nuevo comando.

```
cmd Símbolo del sistema
Microsoft Windows [Versión 10.0.19041.804]
(c) 2020 Microsoft Corporation. Todos los derechos reservados.

C:\Users\conni>cd Desktop
C:\Users\conni\Desktop>cd "Docking Manual"
C:\Users\conni\Desktop\Docking Manual>cd 1C83_Manual
C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd AD4.2
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\AD4.2>autogrid4.exe -p grid.gpf -l grid.glg
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\AD4.2>
```

### 13.8 Generar el archivo \*.dlg

Lo siguiente ahora es correr el *docking*, generando el archivo \*.dlg para lo cual ejecutaremos el siguiente comando:

**autodock4.exe^ -p^dock.dpf^ -l^dock.dlg** (los ^ representan un espacio) seguido de la tecla *intro*. Este cálculo tarda varias horas en realizarse.

```

Símbolo del sistema - autodock4.exe -p dock.dpf -l dock.dlg
Microsoft Windows [Versión 10.0.19041.804]
(c) 2020 Microsoft Corporation. Todos los derechos reservados.

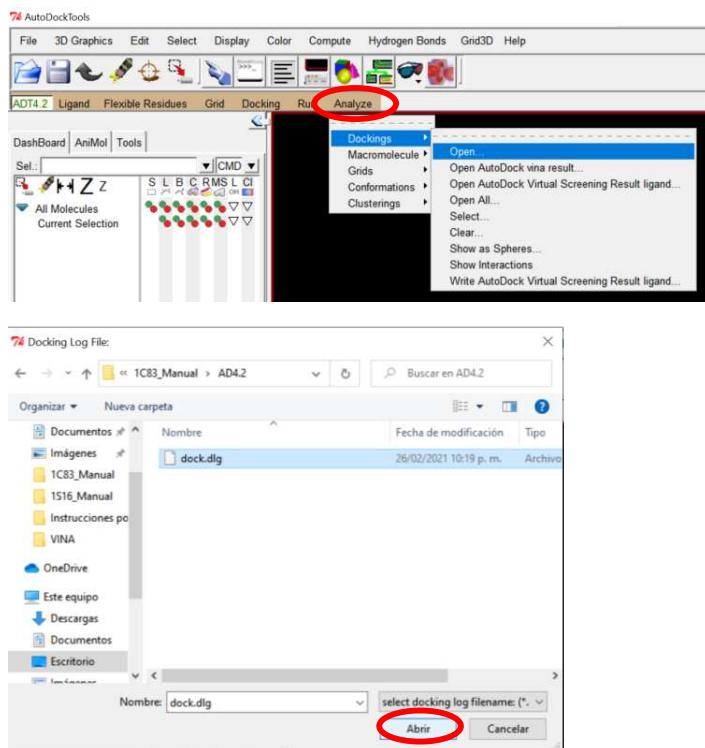
C:\Users\conni>cd Desktop
C:\Users\conni\Desktop>cd "Docking Manual"
C:\Users\conni\Desktop\Docking Manual>cd 1C83_Manual
C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd AD4.2
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\AD4.2>autogrid4.exe -p grid.gpf -l grid.glg
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\AD4.2>autodock4.exe -p dock.dpf -l dock.dlg

```

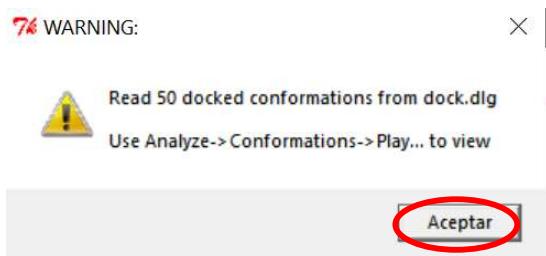
**NOTA:** es muy importante que cuando trabajemos con el “Símbolo del Sistema”, no cerremos las ventanas ni cerremos la computadora (en caso de usar una laptop), porque si lo hacemos, el cálculo se detiene (también aplica si la computadora se suspende o se hiberna).

### 13.9 Monitorear avance del docking

Si queremos saber qué tanto ha avanzado el *docking*, en ADT podemos ir a “Analyze”, “Dockings”, “Open”. Después, escogemos el archivo \*.dlg y hacemos click en “Abrir”.



Una vez que lo abrimos, nos aparecerá un mensaje donde se indica cuántas corridas se han realizado. En esta imagen vemos que se han realizado 50 corridas, por lo tanto, sabemos que nuestro cálculo va a la mitad. Después, podemos dar click en “Aceptar” y cerrar el programa.



Una vez que el cálculo ha terminado, nuevamente veremos que la línea de comandos ya nos permite escribir un nuevo comando. En este punto, ya podemos cerrar la ventana del “Símbolo del sistema”.

```

Símbolo del sistema
Microsoft Windows [Versión 10.0.19041.804]
(c) 2020 Microsoft Corporation. Todos los derechos reservados.

C:\Users\conni>cd Desktop
C:\Users\conni\Desktop>cd "Docking Manual"
C:\Users\conni\Desktop\Docking Manual>cd 1C83_Manual
C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd AD4.2
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\AD4.2>autogrid4.exe -p grid.gpf -l grid.glg
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\AD4.2>autodock4.exe -p dock.dpf -l dock.dlg
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\AD4.2>

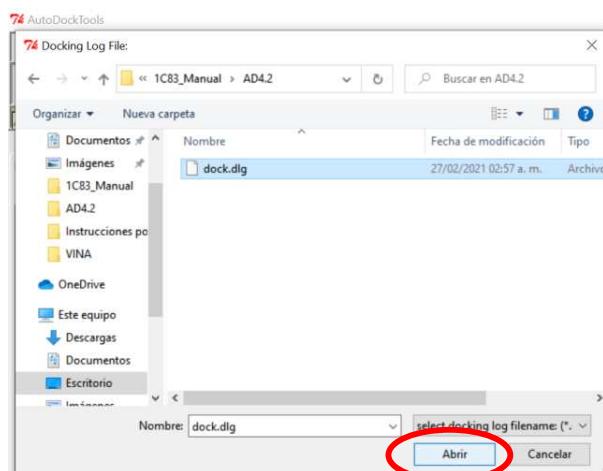
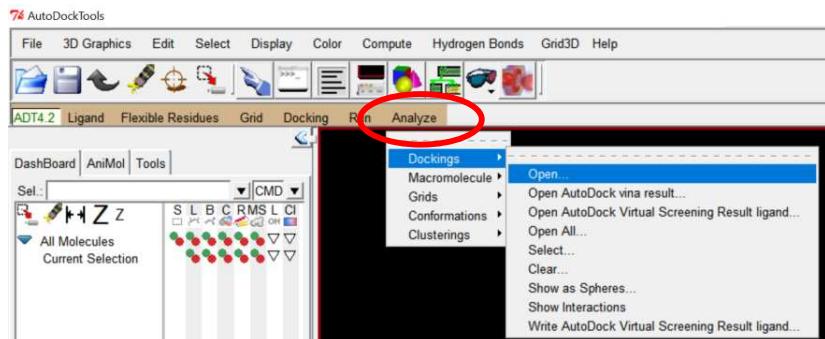
```

Podemos ver también, que en la carpeta en donde realizamos el *docking* (AD4.2), han aparecido nuevos archivos, incluyendo el archivo \*.dlg

Nombre	Fecha de modificación	Tipo	Tamaño
1C83_yas_nolig.A.map	26/02/2021 05:52 p. m.	Archivo MAP	612 KB
1C83_yas_nolig.C.map	26/02/2021 05:52 p. m.	Archivo MAP	613 KB
1C83_yas_nolig.d.map	26/02/2021 05:52 p. m.	Archivo MAP	472 KB
1C83_yas_nolig.e.map	26/02/2021 05:52 p. m.	Archivo MAP	489 KB
1C83_yas_nolig.HD.map	26/02/2021 05:52 p. m.	Archivo MAP	545 KB
1C83_yas_nolig.maps.fld	26/02/2021 05:52 p. m.	Archivo FLD	2 KB
1C83_yas_nolig.maps	26/02/2021 05:52 p. m.	XYZ File	1 KB
1C83_yas_nolig.N.map	26/02/2021 05:52 p. m.	Archivo MAP	598 KB
1C83_yas_nolig.NA.map	26/02/2021 05:52 p. m.	Archivo MAP	599 KB
1C83_yas_nolig.OA.map	26/02/2021 05:52 p. m.	Archivo MAP	593 KB
1C83_yas_nolig	26/02/2021 02:55 p. m.	Archivo PDBQT	236 KB
autodock4	09/02/2020 03:00 a. m.	Aplicación	605 KB
autogrid4	09/02/2020 03:00 a. m.	Aplicación	1,883 KB
<b>dock.dlg</b>	27/02/2021 02:57 a. m.	Archivo DLG	458 KB
dock	26/02/2021 05:30 p. m.	Archivo DPF	3 KB
grid.glg	26/02/2021 05:52 p. m.	Archivo GLG	391 KB
grid	26/02/2021 04:57 p. m.	Archivo GPF	2 KB
OAI_be	27/02/2021 11:07 a. m.	Archivo	238 KB
OAI_min	25/02/2021 12:37 a. m.	Archivo PDBQT	2 KB

### 13.10 Visualización de resultados

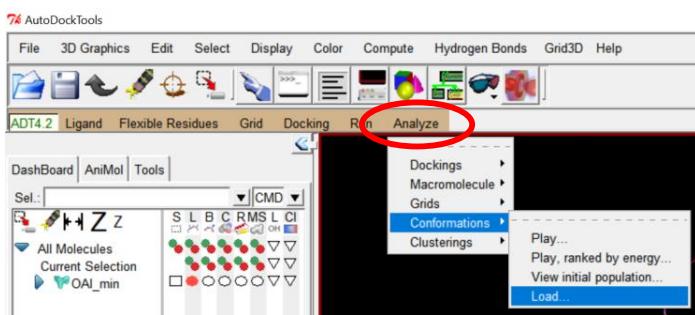
Para observar nuestros resultados abriremos el archivo **dock.dlg**. Para ello iremos a ADT, en “Analyze”, “Dockings”, “Open”. Seleccionamos el archivo y damos click en “Abrir”.



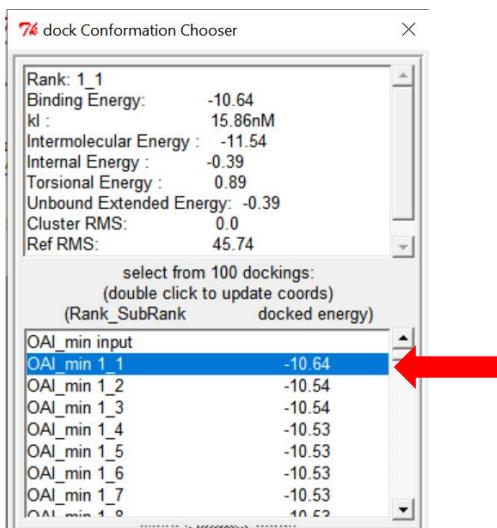
Al abrirlo, veremos el mensaje que nos dice que se realizaron 100 corridas. Damos click en “Aceptar”.



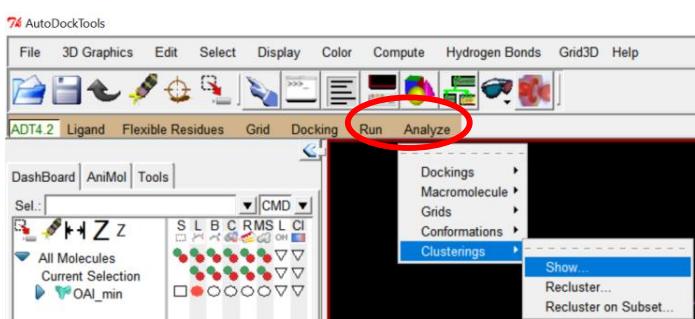
Posteriormente vamos a “Analyze”, “Conformations”, “Load”.

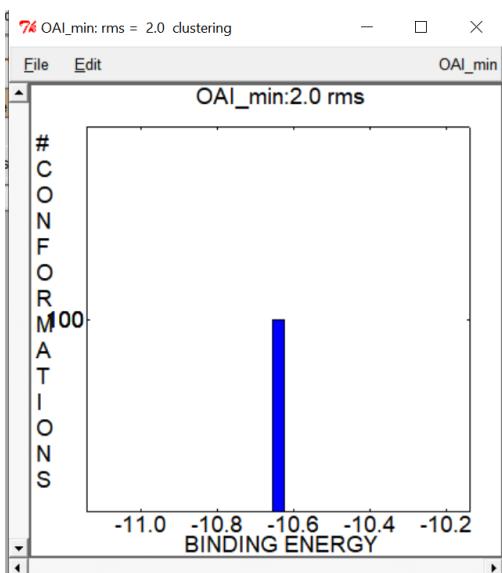


Veremos después, que aparece una nueva ventana con las energías de afinidad calculadas



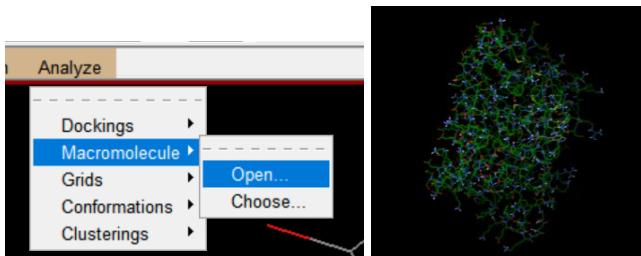
Estas energías de afinidad también están agrupadas en *clusters*, mismos que podemos observar al ir a “Analyze”, “Clusterings”, “Show”. Se desplegará una ventana con una gráfica donde podremos observar todos los *clusters*. En este ejemplo, solamente hubo un clúster que agrupa a 100 conformaciones, en dónde, si todas ellas se empalmaran entre sí, sus átomos tendrían una desviación de la media cuadrática (RMSD por sus siglas en inglés) menor o igual a 2 Å.



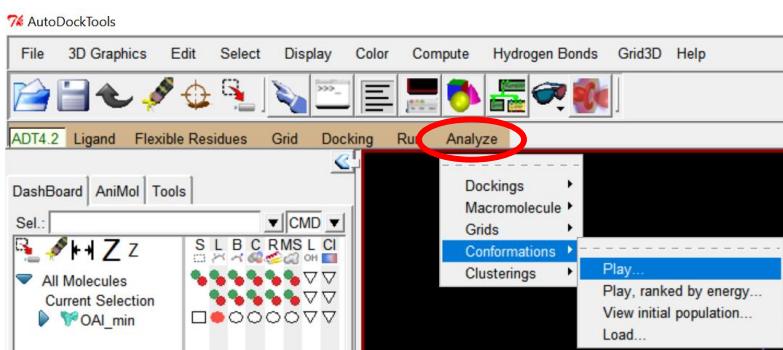


### 13.11 Guardar mejor conformación

Para guardar la mejor conformación que obtuvimos primero debemos abrir el archivo de la proteína. Para ello iremos a “Analyze”, “Macromolecule”, “Open”.



Una vez abierto, iremos a “Analyze”, “Conformations”, “Play”

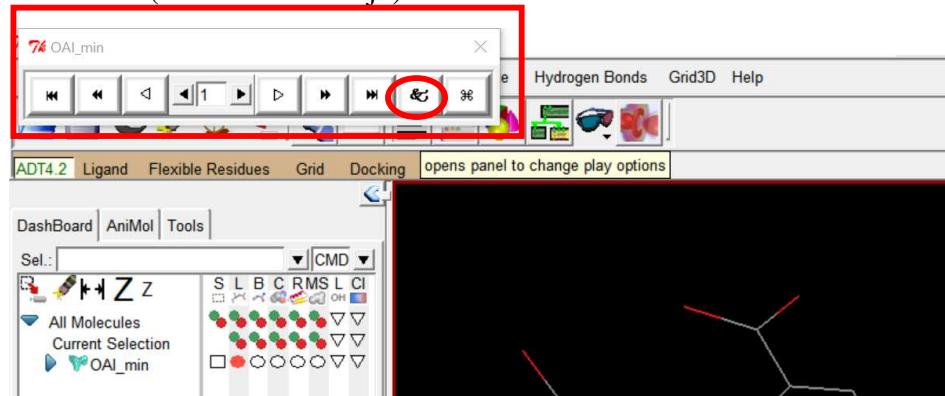


Veremos que aparece una barra como la de la imagen, con estos botones podemos ir “visualizando” cada una de las conformaciones que obtuvimos. Generalmente, la que más

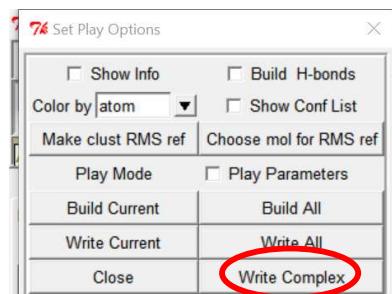
nos interesa es la primera porque tiene la mejor energía de unión, por lo tanto, nos situaremos en esta para guardar el complejo.

NOTA: a veces la pantalla se mueve y puede que no veamos al ligando, debemos movernos y encontrarlo.

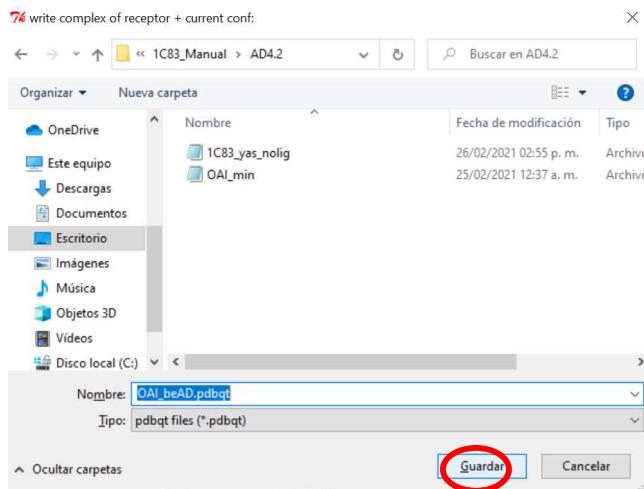
Una vez situados en la conformación que nos interesa, haremos click en el penúltimo botón de la barra (encerrado con rojo).



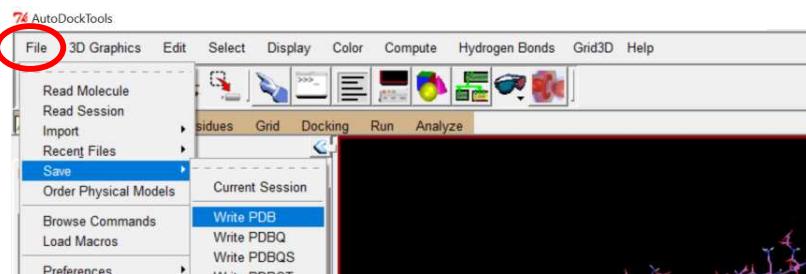
Posteriormente, aparecerá una ventana con diferentes opciones. Debemos hacer click sobre “Write Complex”.



Aparecerá una ventana para guardar el complejo en formato \*.pdbqt. Es recomendable guardarla como “ligando\_beAD.pdbqt” (be = best energy; AD = Autodock)



El siguiente paso es guardarlo en formato \*.pdb. Para ello, debemos cerrar todo lo que teníamos abierto en ADT (dando click derecho sobre el ligando y/o la proteína de la barra de la izquierda y después seleccionar “dismiss”) y posteriormente abrir el complejo que guardamos en el paso anterior, OAI\_beAD.pdbqt. Una vez que esté abierto, iremos a “File”, “Save”, “Write PDB”. Escogemos el destino del archivo y hacemos click sobre guardar.



El complejo que creamos nos servirá más adelante.

## 14 Docking en Autodock VINA

### 14.1 ¿Qué necesitamos?

Para realizar los cálculos de acoplamiento molecular con VINA, al igual que con AD4.2, se recomienda trabajar con distintas carpetas por cada ligando para una mejor organización. Antes de realizar el cálculo, en cada carpeta debe de haber tres archivos que son:

1. \*.pdbqt del ligando
2. \*.pdbqt de la proteína
3. conf.txt

A continuación, encontraremos cómo generar cada archivo.

### 14.2 \*.pdbqt del ligando

Seguimos los mismos pasos que para generar el \*.pdbqt en AD4.2 (ver sección 11.3), por lo tanto para este ejercicio simplemente podemos copiar el archivo \*.pdbqt del ligando y pegarlo en la carpeta correspondiente.

### 14.3 \*.pdbqt de la proteína

Seguimos los mismos pasos que para generar el \*.pdbqt en AD4.2 (ver sección 13.3), por lo tanto para este ejercicio simplemente podemos copiar el archivo \*.pdbqt de la proteína y pegarlo en la carpeta correspondiente.

### 14.4 Conf.txt

Para crear este archivo, es necesario saber el nombre del archivo \*.pdbqt de la proteína y del ligando. En este caso los nombres son **1C83\_yas\_nolig.pdbqt** (proteína) y **OAI\_min.pdbqt** (ligando), respectivamente.

También debemos tener a la mano las coordenadas de algún aminoácido que sea parte del sitio catalítico o del sitio en el que queremos realizar el *docking*. En este ejercicio, sabemos del ejemplo realizado en AD4.2 que las coordenadas son **x = 44.026; y = 14.391; z = 4.315**

De igual manera es necesario que determinemos el tamaño del *grid*. Para esto, podemos tomar como referencia el tamaño que utilizamos para el *docking* con AD4.2, que fue de 40 X 40 X 40 Å, pero con una ligera modificación: este valor debe multiplicarse por 0.375 (*spacing*) para arrojar el valor del *grid* en VINA, entonces:  $40 \times 0.375 = 15$ . Por lo tanto, el tamaño del *grid* en VINA será de **15 X 15 X 15 Å**.

El último dato que necesitamos conocer es cuántos ciclos (num\_modes) queremos hacer. Se recomienda realizar **500**.

Todos estos datos se deben escribir en un archivo de bloc de notas como se ve en la siguiente página. Nota: los \* representan pulsar la tecla “Intro” ↵

```
receptor =  
*  
*  
ligand =  
*  
*  
*  
*  
center_x =  
*  
*  
center_y =  
*  
*  
center_z =  
*  
*  
*  
size_x =  
*  
*  
size_y =  
*  
*  
size_z =  
*  
*  
*  
num_modes =
```

Una vez que hayamos escrito los datos correspondientes, el archivo conf.txt debería verse así:

```
*conf: Bloc de notas
Archivo Edición Formato Ver Ayuda
receptor = 1C83_yas_nolig.pdbqt

ligand = OAI_min.pdbqt

center_x = 44.026
center_y = 14.391
center_z = 4.315

size_x = 15
size_y = 15
size_z = 15

num_modes = 500
```

Posteriormente, guardamos el archivo con el nombre de **conf.txt**

#### 14.5 Para ejecutar el cálculo en VINA

Antes de ejecutar el cálculo, es prudente revisar que en la carpeta tengamos todos los archivos correspondientes. De ser así, esto es lo que veríamos:

Nombre	Fecha de modificación	Tipo	Tamaño
OAI_min	25/02/2021 12:37 a. m.	Archivo PDBQT	2 KB
1C83_yas_nolig	26/02/2021 02:55 p. m.	Archivo PDBQT	236 KB
conf	26/02/2021 07:11 p. m.	Documento de texto	1 KB

Una vez hecho esto, el siguiente paso es ejecutar el cálculo. Para ello, abriremos una ventana de “Símbolo del sistema”. Después, debemos de ingresar a la carpeta correspondiente. En

en este ejemplo la carpeta de trabajo se llama “VINA”, la cual está guardada dentro de una carpeta llamada “1C83\_manual”, esta última se encuentra dentro de la carpeta “Docking Manual”, ésta a su vez está en el escritorio o “Desktop”. Entonces para acceder a la carpeta “VINA”, primero tenemos que acceder a “Desktop” luego “Docking Manual” después “1C83\_Manual” y finalmente “VINA” utilizando el comando *cd* como se muestra en la figura de abajo.

```
□ Símbolo del sistema
Microsoft Windows [Versión 10.0.19041.804]
(c) 2020 Microsoft Corporation. Todos los derechos reservados.

C:\Users\conni>cd Desktop

C:\Users\conni\Desktop>cd "Docking Manual"

C:\Users\conni\Desktop\Docking Manual>cd 1C83_Manual

C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd VINA

C:\Users\conni\Desktop\Docking Manual\1C83_Manual\VINA>
```

Posteriormente, debemos escribir y ejecutar el siguiente comando:  
**“C:\Vina\vina.exe”^--config^conf.txt^--log^log.txt^--out^nombre\_dock.pdbqt** (los ^ representan un espacio).

En este ejemplo, reemplazaremos el “nombre\_dock.pdbqt” por “OAI\_dock.pdbqt”. Este último archivo contendrá todas las poses de mejor energía que realizó el ligando en el sitio de unión.

```
□ Símbolo del sistema
Microsoft Windows [Versión 10.0.19041.804]
(c) 2020 Microsoft Corporation. Todos los derechos reservados.

C:\Users\conni>cd Desktop

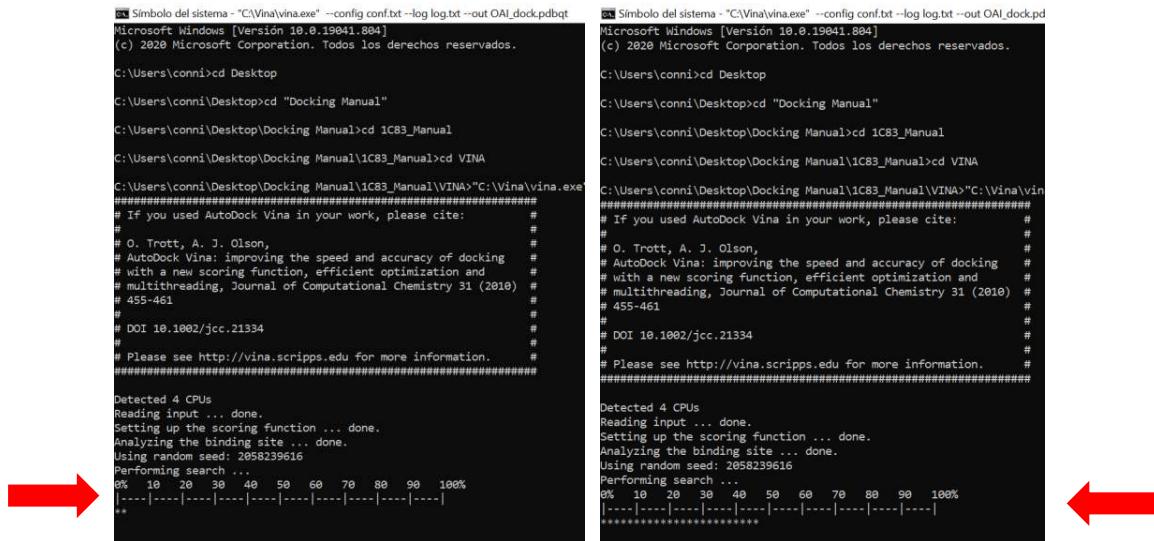
C:\Users\conni\Desktop>cd "Docking Manual"

C:\Users\conni\Desktop\Docking Manual>cd 1C83_Manual

C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd VINA

C:\Users\conni\Desktop\Docking Manual\1C83_Manual\VINA>"C:\Vina\vina.exe" --config conf.txt --log log.txt --out OAI_dock.pdbqt
```

Y damos *INTRO*. Al momento, aparecerán los datos como se muestra en la imagen y podremos monitorear el avance del *docking* en la barra de progreso.



```

Símbolo del sistema - "C:\Vina\vina.exe" --config conf.txt --log log.txt --out OAI_dock.pdbqt
Microsoft Windows [Versión 10.0.19041.804]
(c) 2020 Microsoft Corporation. Todos los derechos reservados.

C:\Users\conni>cd Desktop
C:\Users\conni\Desktop>cd "Docking Manual"
C:\Users\conni\Desktop\Docking Manual>cd 1C83_Manual
C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd VINA
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\VINA>"C:\Vina\vina.exe"
#####
# If you used AutoDock Vina in your work, please cite: #
#
# O. Trott, A. J. Olson,
# AutoDock Vina: improving the speed and accuracy of docking
# with a new scoring function, efficient optimization and
# multithreading, Journal of Computational Chemistry 31 (2010)
# 455-461
#
# DOI 10.1002/jcc.21334
#
# Please see http://vina.scripps.edu for more information.
#####

Detected 4 CPUs
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: 2058239616
Performing search ...
% 10 20 30 40 50 60 70 80 90 100%
|----|----|----|----|----|----|----|----|
**
```

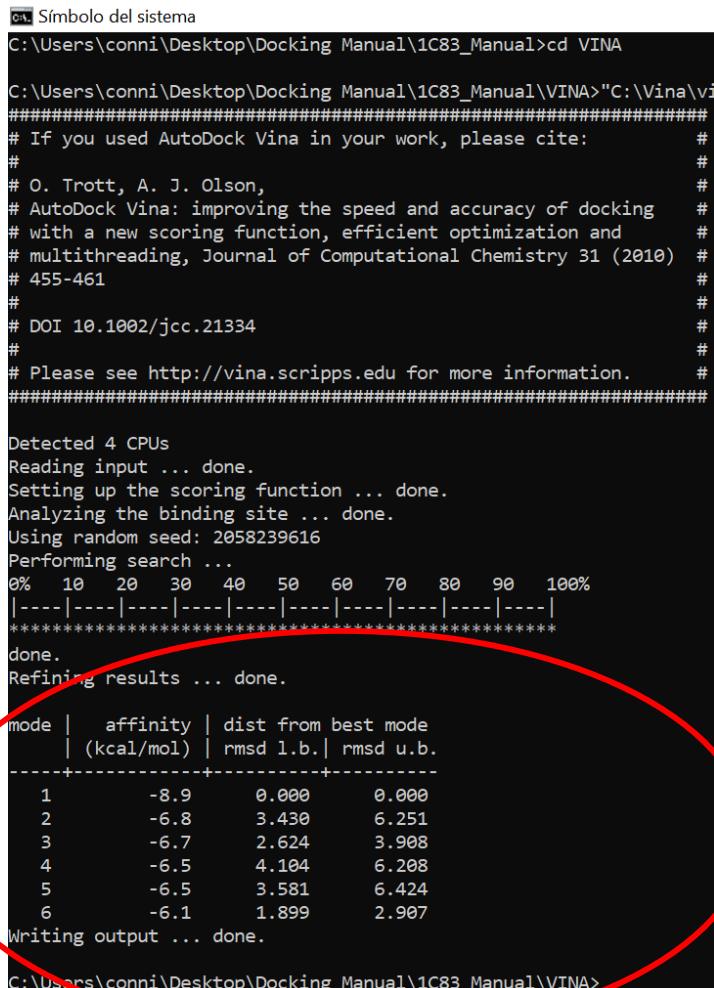
```

Símbolo del sistema - "C:\Vina\vina.exe" --config conf.txt --log log.txt --out OAI_dock.pdbqt
Microsoft Windows [Versión 10.0.19041.804]
(c) 2020 Microsoft Corporation. Todos los derechos reservados.

C:\Users\conni>cd Desktop
C:\Users\conni\Desktop>cd "Docking Manual"
C:\Users\conni\Desktop\Docking Manual>cd 1C83_Manual
C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd VINA
C:\Users\conni\Desktop\Docking Manual\1C83_Manual\VINA>"C:\Vina\vina.exe"
#####
# If you used AutoDock Vina in your work, please cite: #
#
# O. Trott, A. J. Olson,
# AutoDock Vina: improving the speed and accuracy of docking
# with a new scoring function, efficient optimization and
# multithreading, Journal of Computational Chemistry 31 (2010)
# 455-461
#
# DOI 10.1002/jcc.21334
#
# Please see http://vina.scripps.edu for more information.
#####

Detected 4 CPUs
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: 2058239616
Performing search ...
% 10 20 30 40 50 60 70 80 90 100%
|----|----|----|----|----|----|----|----|
*****
```

Cuando el cálculo termina, observamos una tabla de datos, además de la leyenda “Writing output...done” y la barra que nos indica que podemos escribir un nuevo comando.



```

Símbolo del sistema
C:\Users\conni\Desktop\Docking Manual\1C83_Manual>cd VINA
#####
# If you used AutoDock Vina in your work, please cite: #
#
# O. Trott, A. J. Olson,
# AutoDock Vina: improving the speed and accuracy of docking
# with a new scoring function, efficient optimization and
# multithreading, Journal of Computational Chemistry 31 (2010)
# 455-461
#
# DOI 10.1002/jcc.21334
#
# Please see http://vina.scripps.edu for more information.
#####

Detected 4 CPUs
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: 2058239616
Performing search ...
% 10 20 30 40 50 60 70 80 90 100%
|----|----|----|----|----|----|----|----|
*****
```

done.  
Refining results ... done.

mode	affinity	dist from best mode (kcal/mol)	rmsd 1.b.	rmsd u.b.
1	-8.9	0.000	0.000	
2	-6.8	3.430	6.251	
3	-6.7	2.624	3.908	
4	-6.5	4.104	6.208	
5	-6.5	3.581	6.424	
6	-6.1	1.899	2.907	

Writing output ... done.

C:\Users\conni\Desktop\Docking Manual\1C83\_Manual\VINA>

Al ir a la carpeta donde trabajamos, podemos observar que aparecieron dos nuevos archivos: el **log.txt** y el **OAI\_dock.pdbqt**

Nombre	Fecha de modificación	Tipo	Tamaño
OAI_min	25/02/2021 12:37 a. m.	Archivo PDBQT	2 KB
1C83_yas_nolig	26/02/2021 02:55 p. m.	Archivo PDBQT	236 KB
conf	26/02/2021 07:11 p. m.	Documento de texto	1 KB
log	26/02/2021 07:18 p. m.	Documento de texto	2 KB
OAI_dock	26/02/2021 07:18 p. m.	Archivo PDBQT	13 KB

Al abrir el archivo log.txt, encontraremos la misma tabla que se veía cuando el cálculo finalizó en la ventana de “Símbolo del Sistema”. En esta tabla podemos observar la energía de unión calculada para el ligando.

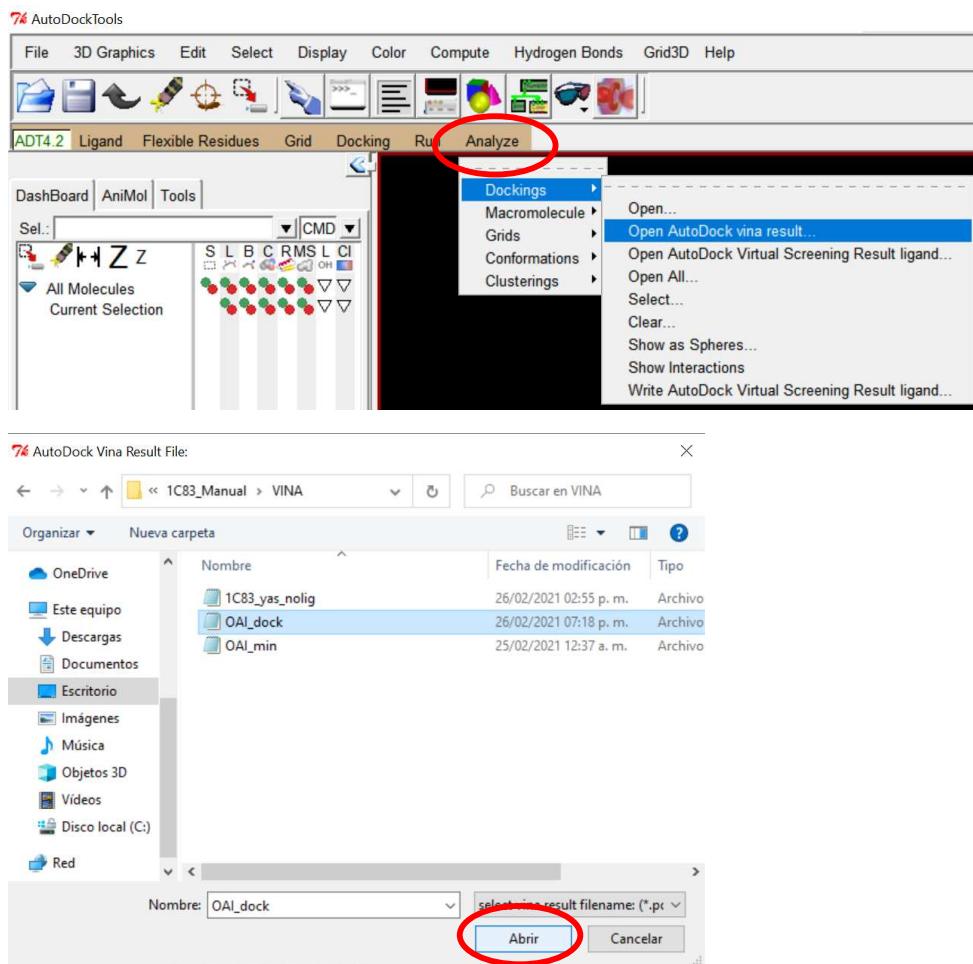
```
log: Bloc de notas
Archivo Edición Formato Ver Ayuda
#####
# If you used AutoDock Vina in your work, please cite: #
# O. Trott, A. J. Olson, #
# AutoDock Vina: improving the speed and accuracy of docking #
# with a new scoring function, efficient optimization and #
# multithreading, Journal of Computational Chemistry 31 (2010) #
# 455-461 #
# DOI 10.1002/jcc.21334 #
# Please see http://vina.scripps.edu for more information. #
#####

Detected 4 CPUs
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: 2058239616
Performing search ... done.
Refining results ... done.

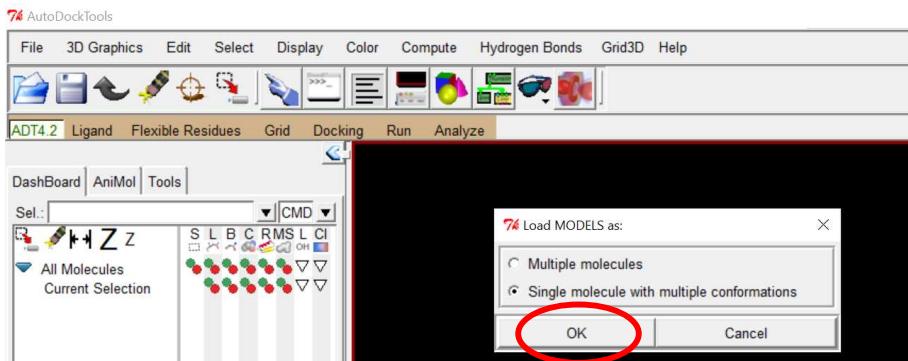
mode | affinity | dist from best mode
     | (kcal/mol) | rmsd l.b.| rmsd u.b.
-----+-----+-----+
 1    -8.9    0.000    0.000
 2    -6.8    3.430    6.251
 3    -6.7    2.624    3.908
 4    -6.5    4.104    6.208
 5    -6.5    3.581    6.424
 6    -6.1    1.899    2.907
Writing output ... done.
```

#### 14.6 Visualización de resultados

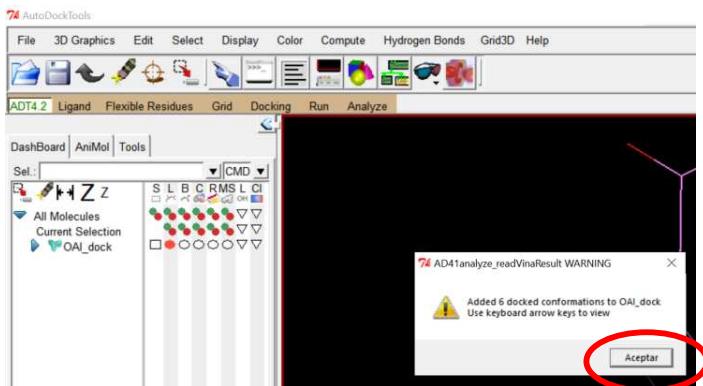
Por su parte, el archivo **OAI\_dock.pdbqt** se puede abrir desde ADT y podemos observar las conformaciones del ligando. Para ello, debemos ir a la pestaña “Analyze”, “Dockings”, “Open AutoDock vina result...”. Posteriormente, seleccionamos el archivo de interés y hacemos click en “Abrir”.



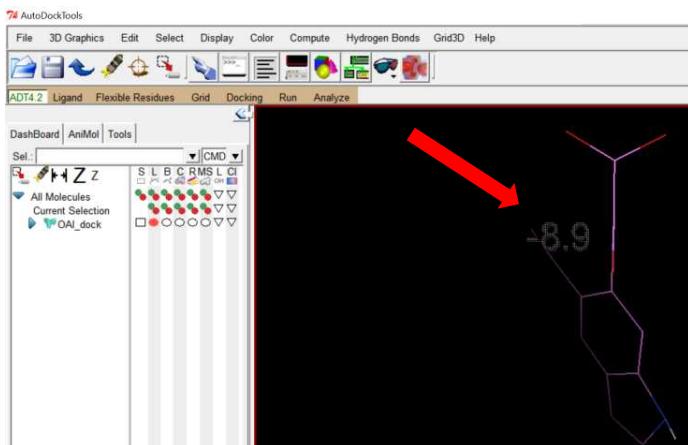
Aparecerá una ventana con dos opciones. Para ver la conformación con la mejor energía de unión, debemos seleccionar la opción que dice “single molecule with multiple conformations” y dar click en “OK”.



Aparecerá un mensaje que indica el número de conformaciones que se crearon. Hacemos click sobre “Aceptar”.

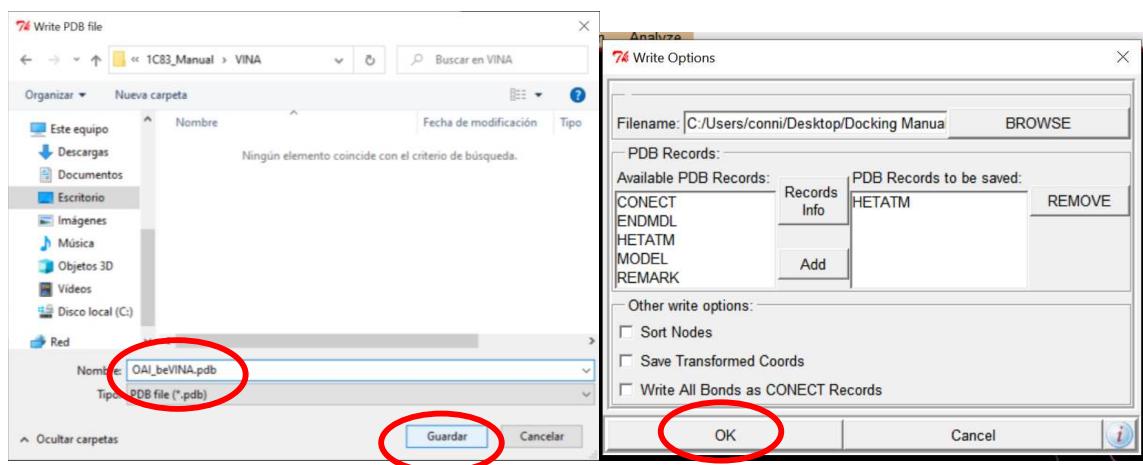
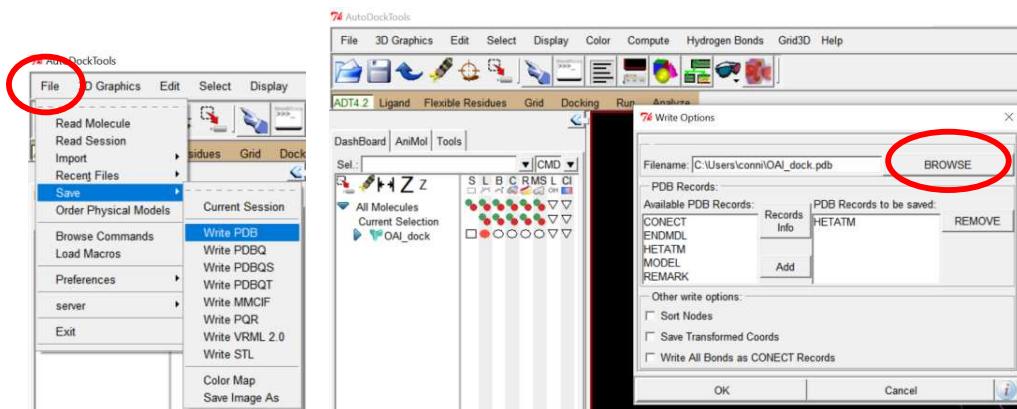


Veremos que aparecen las conformaciones en la pantalla acompañadas de su energía de unión



#### 14.7 Guardar mejor conformación

Para guardar la mejor conformación debemos ir a “File”, “Save”, “Write PDB”. Hacemos click en “Browse”, escogemos el destino del archivo, lo nombramos como **OAI\_beVINA.pdb** (be; Best Energy), hacemos click en “Guardar” y después en “OK”.



El complejo que creamos nos servirá más adelante.

## 15 Docking en GOLD

### 15.1 Scores de GOLD

Antes de comenzar con las instrucciones para este programa, es necesario hablar de una característica particular de GOLD: el *score* que utiliza. Recordemos que en *docking* el *score* es aquella función con la que un *software* califica la interacción ligando – proteína. Como vimos en las secciones anteriores, tanto AD como VINA utilizan la energía de unión como su *score*, donde, mientras más negativa sea esta energía, mejor calificado está el ligando. Sin embargo, en GOLD es al contrario pues mientras mayor sea el *score*, querrá decir que el ligando tiene una mejor calificación. Las funciones de *score* más utilizados de GOLD son: ChemPLP y GoldScore.

ChemPLP (Piecewise Linear Potential) es la función de *score* de GOLD por default. Es una función empírica optimizada de *score* para la predicción de poses. Modela la complementariedad estérica entre proteína y ligando. Esta función se puede utilizar para *docking* covalente, toma en cuenta la flexibilidad de la cadena lateral y moléculas de agua explícitas. La función de GoldScore es de tipo mecánica molecular optimizada para la predicción de sitios de unión del ligando y toma en cuenta factores como la energía de enlaces de hidrógeno, fuerzas de van der Waals, interacciones con metales y deformaciones por torsión.<sup>20</sup>

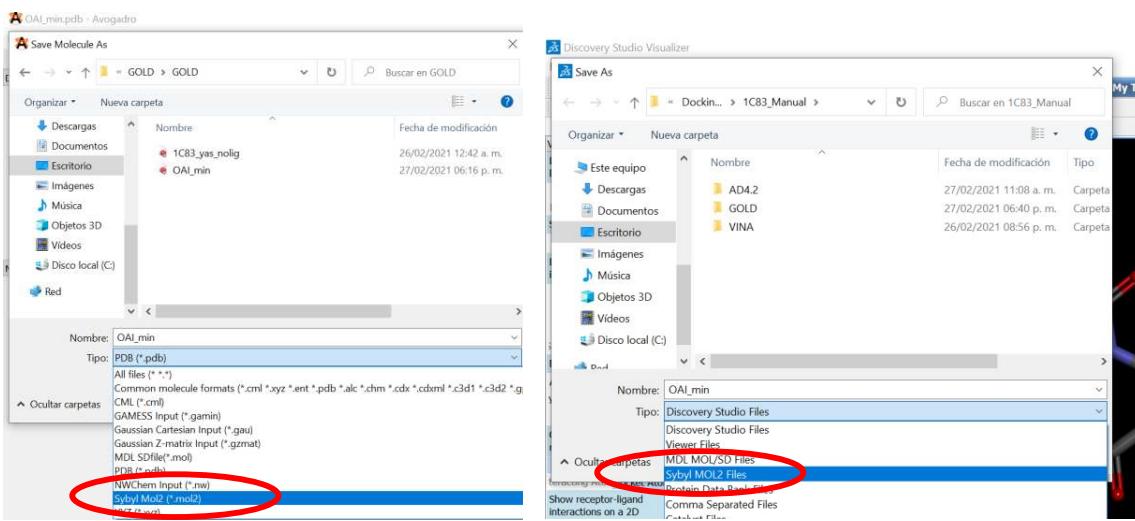
### 15.2 ¿Qué necesitamos?

Para hacer un *docking* en GOLD necesitamos el archivo \*.pdb de la proteína, además del ligando en formato \*.mol2.

Recordemos que tenemos ambos en \*.pdb. El \*.pdb de la proteína lo obtuvimos de YASARA y después le retiramos el ligando, mientras que el \*.pdb del ligando es el que generamos cuando construimos y minimizamos la estructura de OAI en Avogadro.

Nombre	Fecha de modificación	Tipo	Tamaño
1C83_yas_nolig	26/02/2021 12:42 a. m.	PDB File	386 KB
OAI_min	27/02/2021 06:16 p. m.	PDB File	3 KB

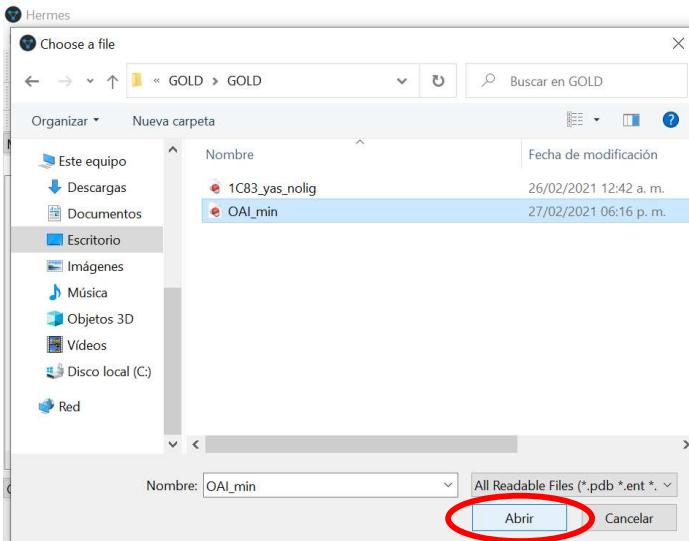
Podemos convertir fácilmente el archivo \*.pdb a \*.mol2 con **Avogadro** o con **Discovery Studio**. En ambos casos, el primer paso es abrir el programa y posteriormente abrir el \*.pdb del ligando (**OAI\_min.pdb**). Una vez abierto, solo debemos ir a “File”, “Save as” y escoger “Sybyl MOL2 Files”.



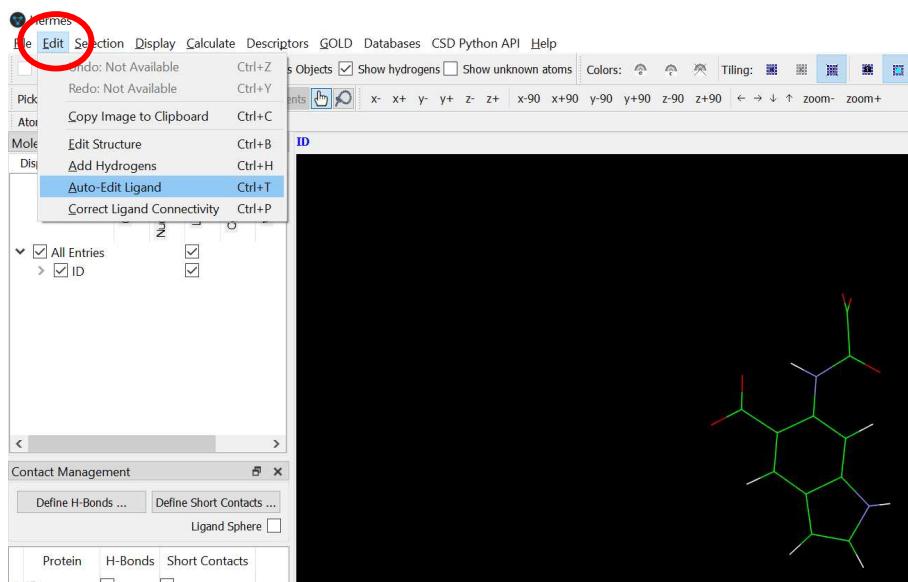
### 15.3 Verificar ligandos

Debido a que Avogadro es diferente de GOLD, a veces los archivos generados en el primero no se leen correctamente en el segundo, lo que podría causar problemas en los resultados del docking. Por lo tanto, debemos corroborar que GOLD tiene correctamente parametrizados átomos como los tipos de enlace del ligando.

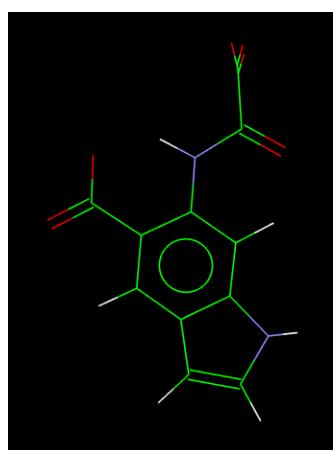
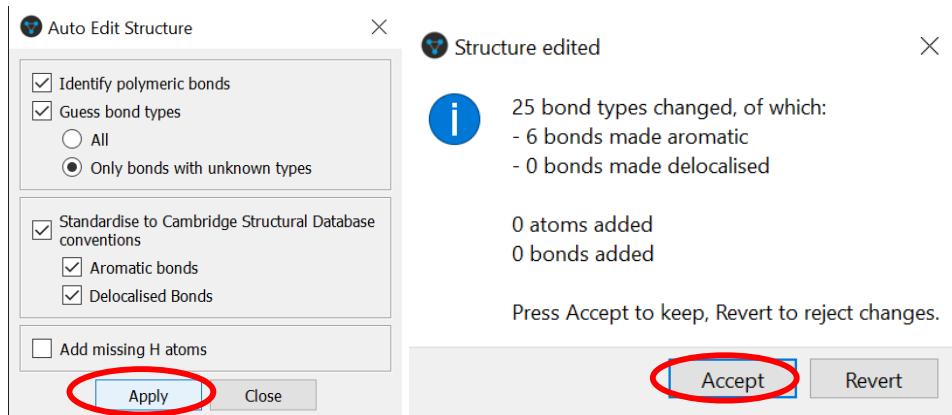
Para ello, el primer paso es abrir la interfaz **Hermes** de GOLD. Una vez abierto, iremos a “File”, “Open”. Despues, seleccionaremos el archivo \*.mol2 que recién generamos (OAI\_min.mol2) y daremos click en “Abrir”.



Una vez abierto, veremos nuestra estructura. También podremos observar, que no están bien representados algunos enlaces, incluyendo la aromaticidad. Para corregirlo, debemos ir a “Edit”, “Auto Edit ligand”



Se desplegará un pequeño menú, simplemente damos click en “Apply”. Aparecerá ahora un mensaje con las modificaciones que se le realizaron al ligando. Hacemos click en “Accept” y podremos ver a nuestra estructura, ahora representada de manera correcta.



Posteriormente, debemos guardar esta estructura seleccionando “File”, “Save as”, escogemos el destino de la estructura y podemos sobreescribirla o guardarla con un nombre diferente para no confundir estructuras, hacemos click en “Guardar”.

#### 15.4 Nombrar ligandos

Adicionalmente, otro problema que se presenta frecuentemente al utilizar GOLD con ligandos que se generaron en un programa diferente es que GOLD no los nombra de manera adecuada. Para revisar esto, debemos dirigirnos a la carpeta en donde guardamos el ligando corregido. Veremos este archivo \*.pdb, que abriremos a continuación con algún procesador de texto, como *Wordpad*. Para abrirlo, hacemos click derecho sobre el archivo, escogemos “Abrir con” y seleccionamos el procesador de texto.

En este caso en particular, como vemos donde señala la flecha roja, este archivo sí se nombró correctamente. Sin embargo, si surgiera un momento en donde al abrir el archivo vemos que el nombre aparece como \*\*\*\*\*, lo único que hay que hacer es borrarlo y escribir el nombre que deseamos (sin espacios), después guardamos los cambios en el documento y cerramos.

---

```
#      Created by:Discovery Studio
#      Creation time: 2021-02-27T18:27:19Z

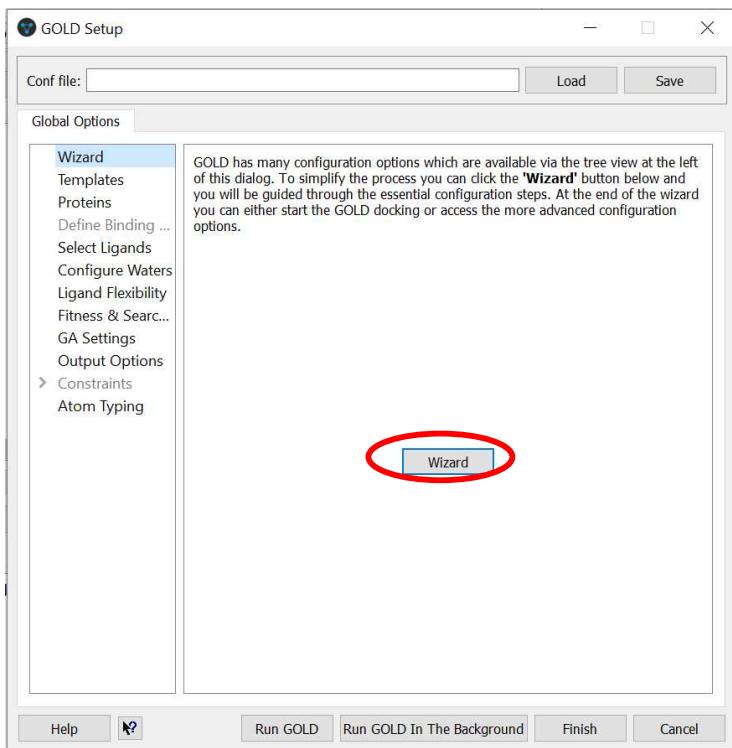
@<TRIPOS>MOLECULE
OAI_min    ←
  24      25      1 0 0
PROTEIN
USER_CHARGES

@<TRIPOS>ATOM
  1 C          2.973000   -1.973000    0.449000 C.2      1
UNK1        -0.0368
  2 C          3.693000   -0.808000    0.448000 C.2      1
UNK1        0.0017
  3 N          2.828000    0.292000    0.115000 N.p13    1
UNK1        -0.3611
  4 C          1.496000   -0.241000    0.087000 C.ar     1
UNK1        0.0475
  5 C          0.285000    0.406000   -0.105000 C.ar     1
UNK1        -0.0135
```

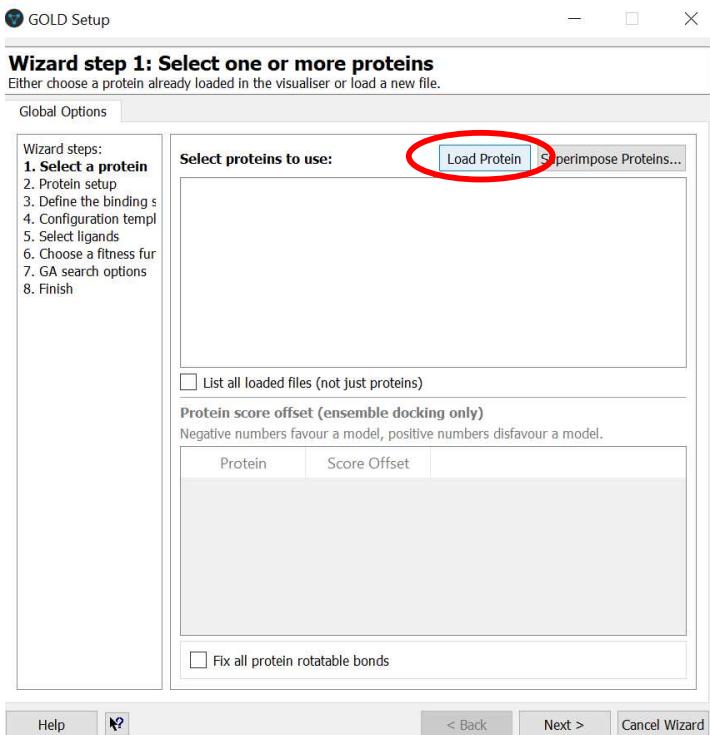
#### 15.5 Para ejecutar el cálculo

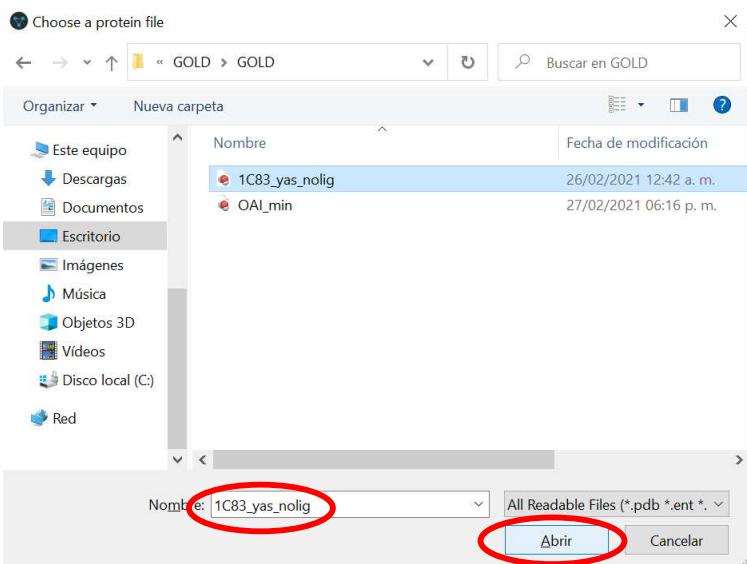
Para ejecutar el *docking* debemos abrir el programa GOLD. Aparecerá “Gold Setup”, que nos servirá para establecer todos los parámetros requeridos

Primero hacemos click sobre “Wizard”

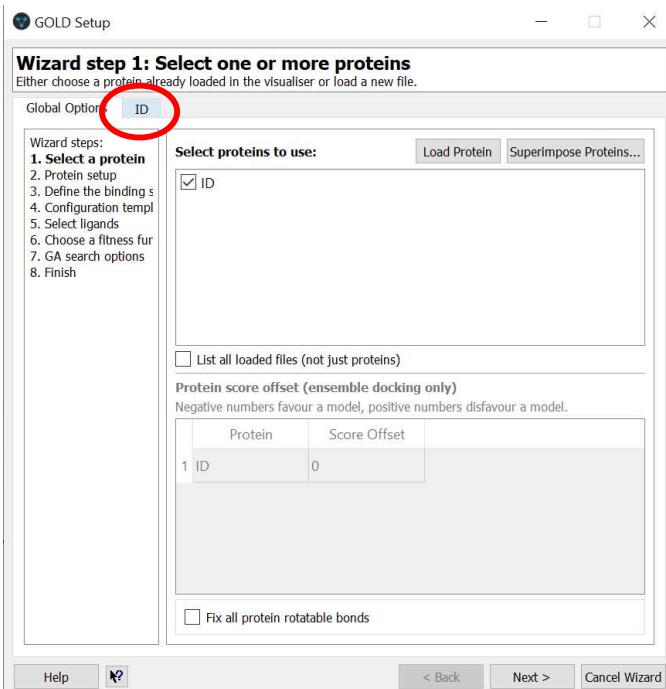


Posteriormente, hacemos click sobre “Load Protein”, escogemos el archivo \*.pdb de nuestra proteína (en este ejemplo, 1C83\_yas\_nolig.pdb) y hacemos click en “Abrir”.

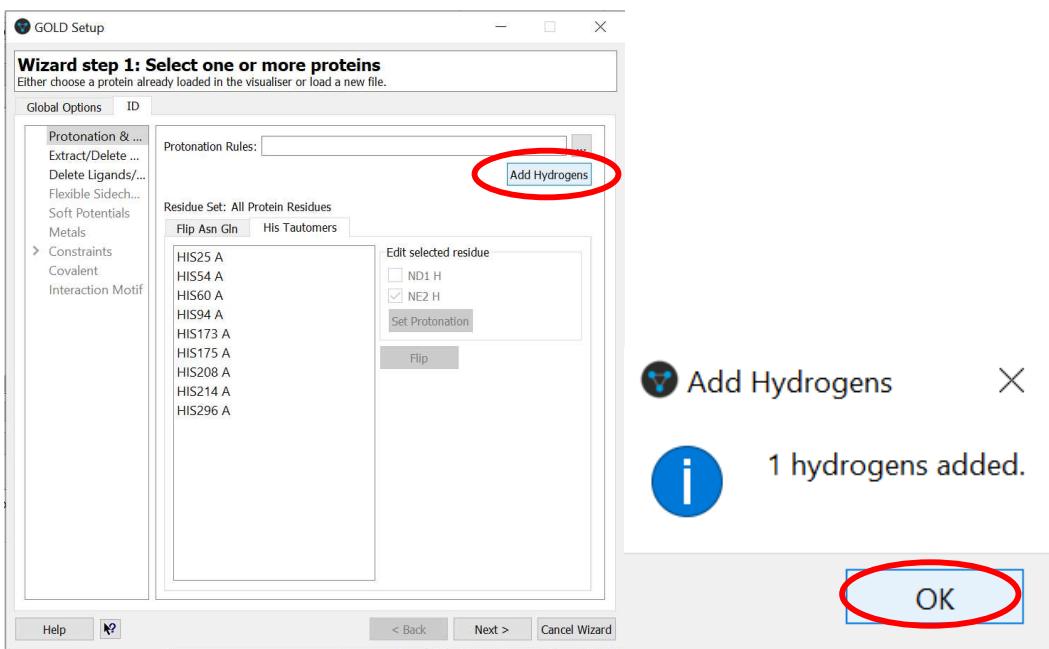




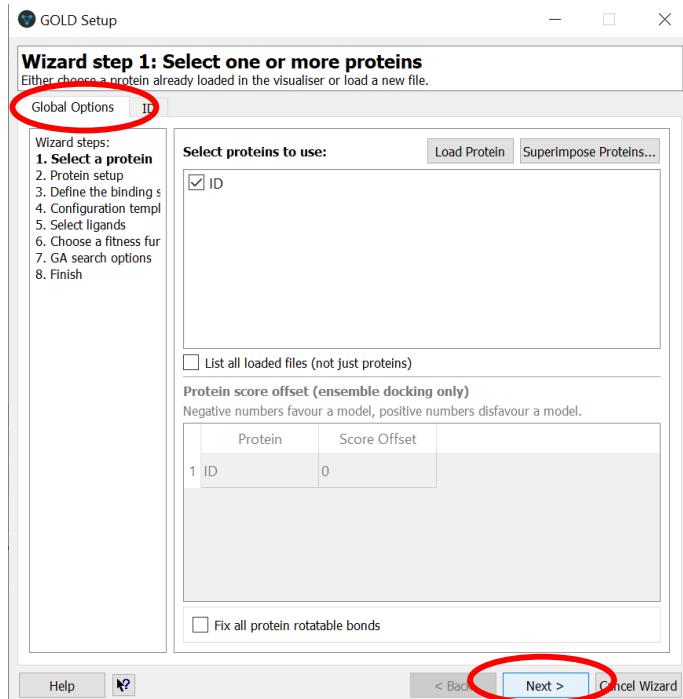
Posteriormente, hacemos click sobre la pestaña “ID”



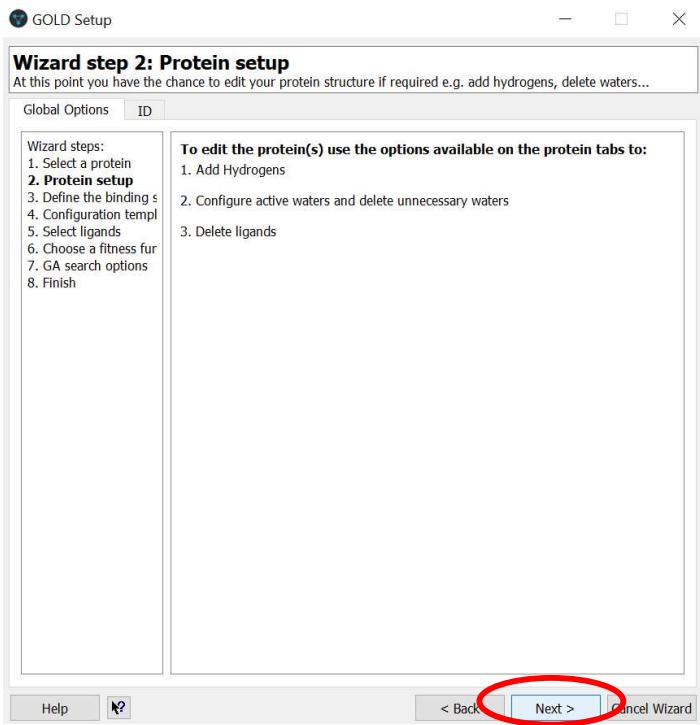
Una vez en la pestaña, haremos click sobre “Add Hydrogens”. Posteriormente, saldrá un mensaje con el número de hidrógenos que se agregaron y hacemos click sobre “OK”.



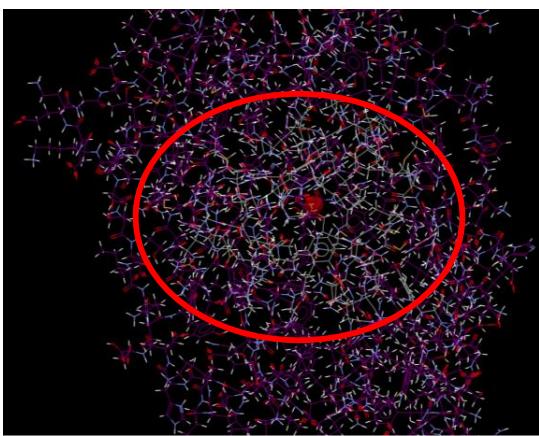
Después, regresamos a la pestaña de “Global Options” y hacemos click sobre “Next”.



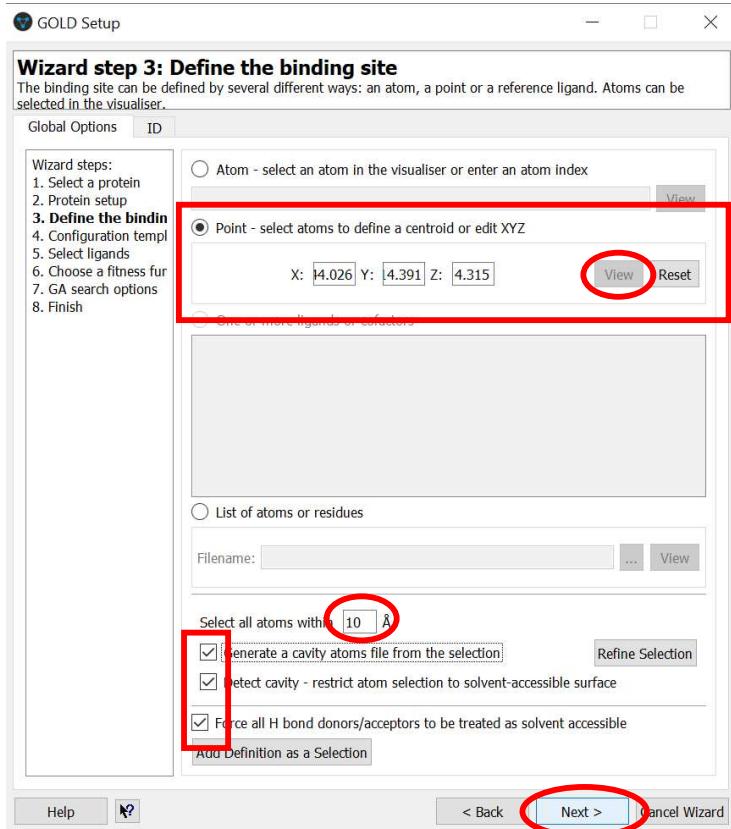
Nuevamente hacemos click sobre “Next”.



El siguiente paso es definir el “grid” en GOLD. Para ello, existen varias opciones como: seleccionar directamente los aminoácidos involucrados, escribir las coordenadas o centrar sobre algún cofactor. En este caso, como ya conocemos las coordenadas por los ejercicios anteriores en AD4.2 y VINA, podemos escoger esta opción. Por lo tanto, el siguiente paso será escribir las coordenadas en la caja correspondiente y hacer click sobre “View”. Debido al gran número de átomos que tenemos en pantalla, es un poco difícil distinguir cada uno de ellos. Sin embargo, es posible ver el “área” que se ha seleccionado para llevar a cabo el docking porque está en color blanco, mientras que todo lo demás, se queda de color morado.



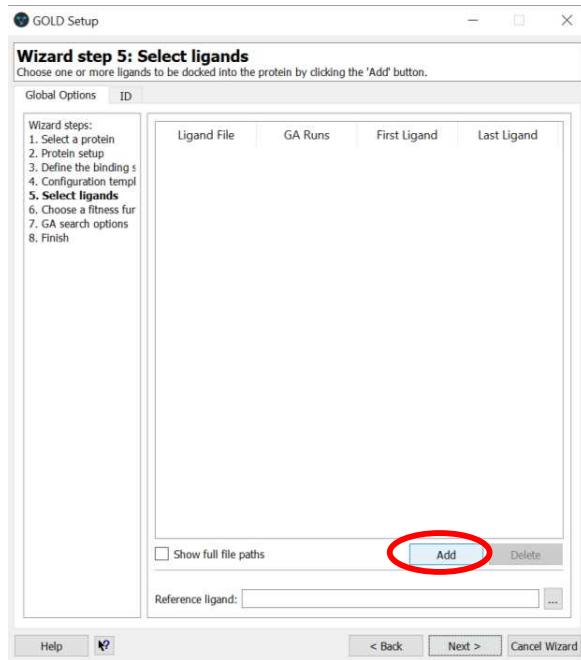
El siguiente paso, en esta misma ventana es establecer el tamaño del área en donde se realizará el docking (en este caso 10 Å) y también habilitar la opción de “Generate a cavity atom from the selection”. Daremos click en “Next”.



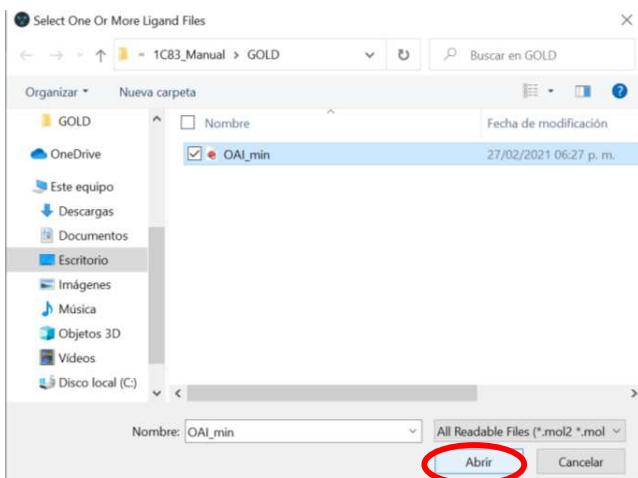
En el paso 4 del “Wizard”, nos da la opción de cargar un “template” para cierto tipo de proteínas. En este caso no utilizaremos ninguno porque no es necesario, sin embargo, son útiles para proteínas de la familia P450 y algunas metaloproteínas. Si quisieramos cargar un template, seleccionaríamos al que queremos y haríamos click sobre “Load Template” (señalado en azul). En este ejercicio, no seleccionaremos ninguno y simplemente haremos click sobre “Next”.



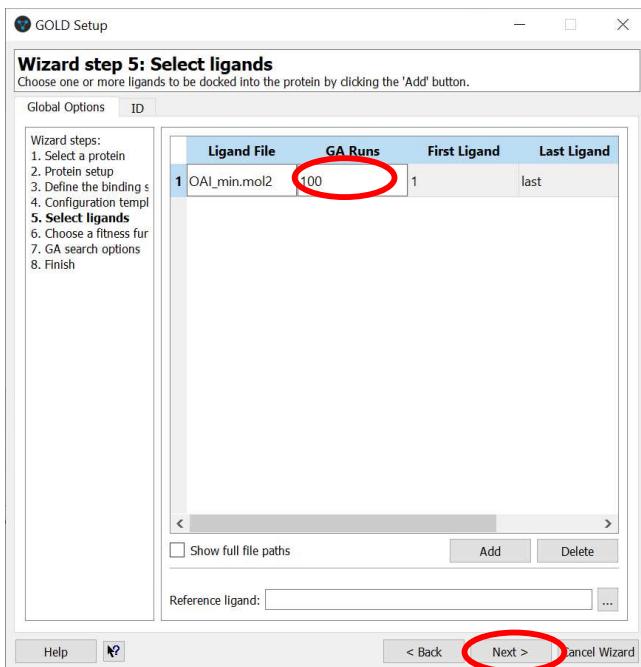
Lo siguiente es escoger los ligandos con los que realizaremos el cálculo. Para ello, debemos hacer click sobre “Add”.



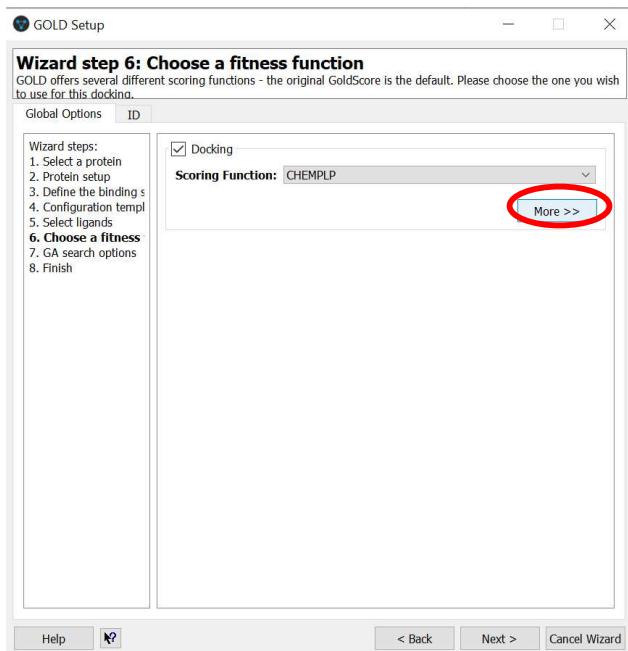
Posteriormente, seleccionamos a nuestro ligando en formato \*.mol2 y hacemos click en “Abrir”.



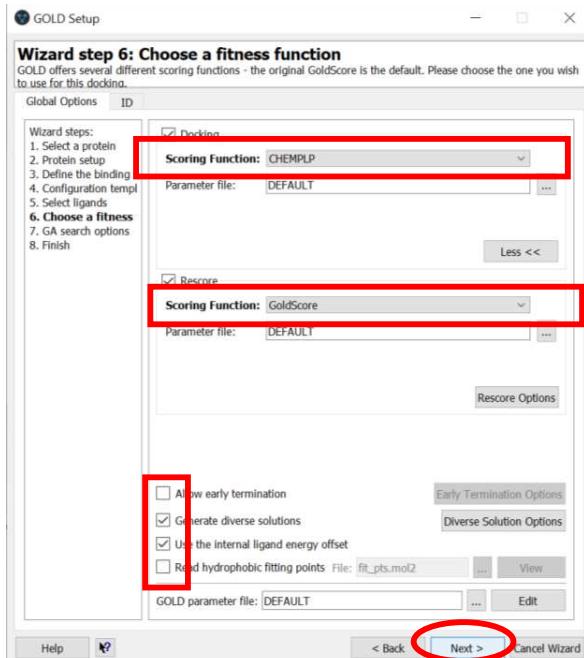
Lo siguiente es establecer el número de corridas que se hará. Se recomiendan 100. Para modificarlo, hacemos doble click sobre el rectángulo señalado con rojo y manualmente modificamos el valor numérico. Posteriormente hacemos click en “Next”.



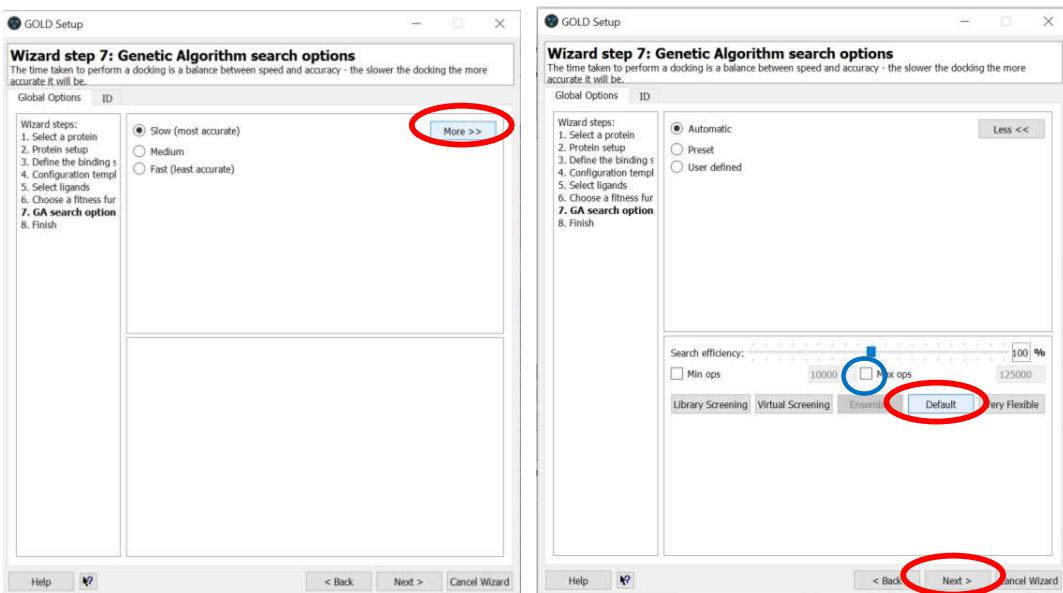
El siguiente paso es escoger las funciones de *score* que utilizaremos. Para ello haremos click sobre “More”



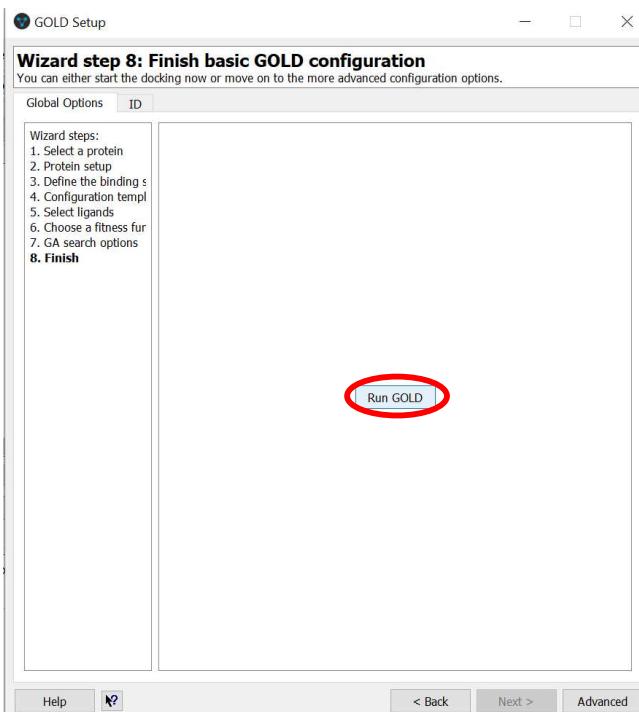
Posteriormente, escogeremos a CHEMPLP como nuestra función principal y a GoldScore para el *rescore*. Debemos deshabilitar la casilla “Allow early termination” y hacemos click en “Next”.



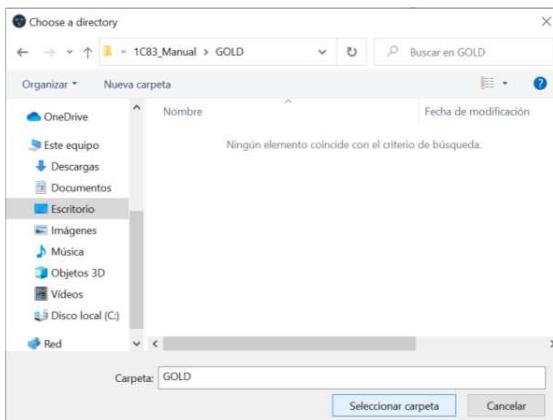
El siguiente paso es establecer algunas opciones referentes al GA. Si hacemos click sobre “More”, veremos que se encuentran más opciones. En este caso, nosotrosaremos click sobre “Default”. Sin embargo, es útil saber que estos parámetros pueden cambiarse conforme lo requiera el cálculo que realizaremos. Por ejemplo, en muchas metodologías se selecciona la casilla de “Max ops”, señalada en azul. Posteriormente, hacemos click en “Next”.



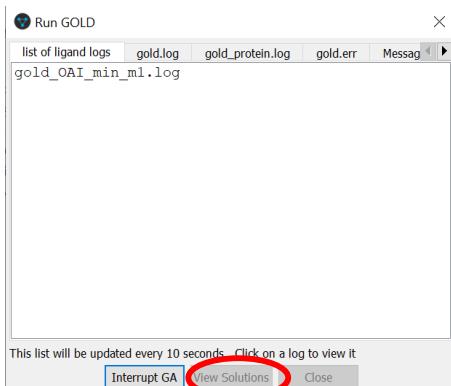
El último paso en el Wizard, es hacer click sobre “Run GOLD”.



Seleccionamos el destino del archivo y hacemos click en “Seleccionar carpeta”

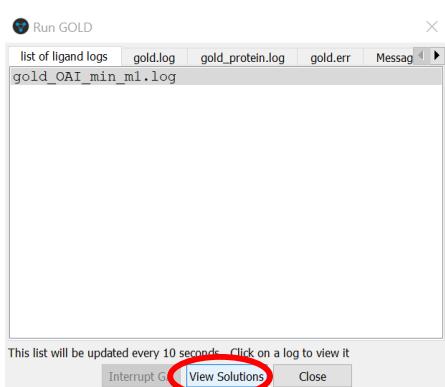


En cuanto empiece el cálculo, veremos una pequeña ventana llamada “Run GOLD”, donde podremos ir observando el progreso del mismo. El cálculo termina cuando la opción “View Solutions” esté activada.

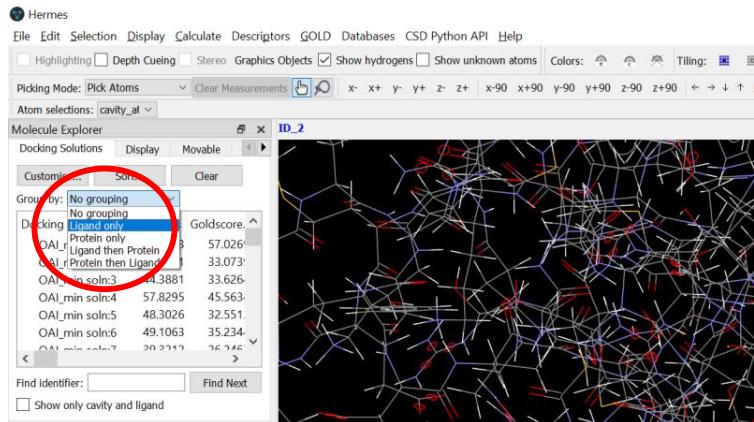


## 15.6 Visualización de resultados

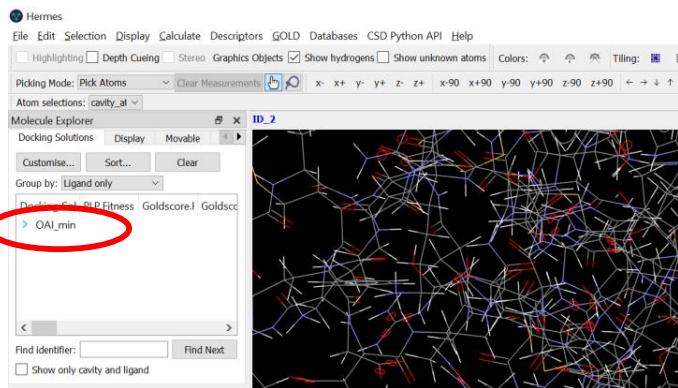
Una vez que el cálculo terminó y la opción “View Solutions” se ha activado, hacemos click sobre ella.



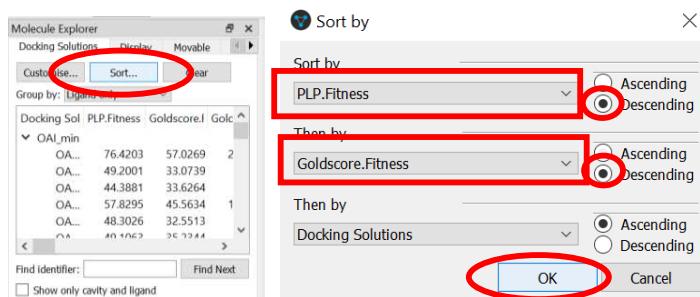
Ahora ya podemos cerrar tanto la ventana “Run GOLD”, como el “Wizard”. Ya que abrimos los resultados, debemos agruparlos por ligando, para su fácil interpretación. Para ello haremos click sobre las opciones desplegables de “Group by:” y escogeremos “Ligand only”.



Al escoger esta agrupación, veremos que todas las soluciones corresponden a nuestro único ligando, OAI\_min.



Como veremos, estos valores no están ordenados. Para ordenarlos, debemos ir a “Sort”, escogemos primero a “ChemPLP” y después a “GoldScore”, ambos en forma descendente y hacemos click sobre “OK”.

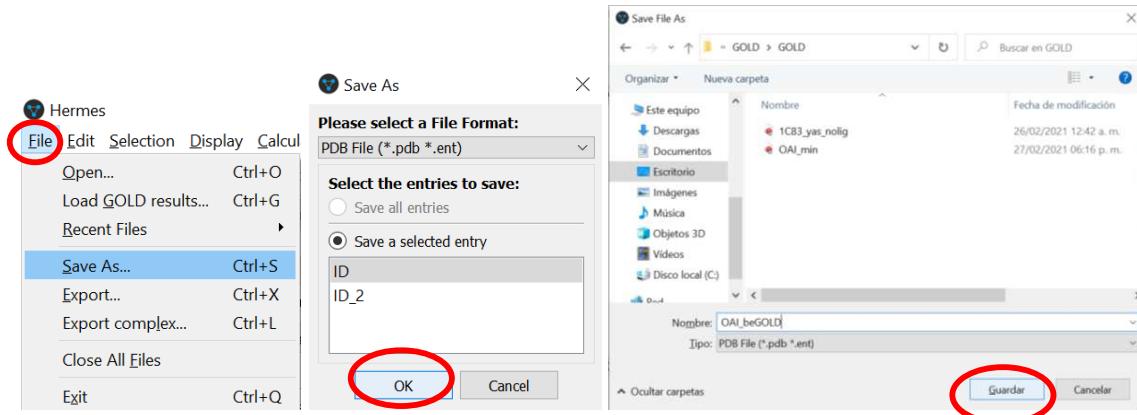


Como vemos, los cambios se han aplicado y ahora nuestros resultados están ordenados de forma descendente.

Molecule Explorer			
Docking Solutions	Display	Movable	Customise...
<input type="button" value="Sort..."/>	<input type="button" value="Clear"/>		
Group by: Ligand only			
Docking Sol	PLP.Fitness	Goldscore.I	Golc
▼ OAI_min			▲
OA...	76.4203	57.0269	2
OA...	57.8295	45.5634	1
OA...	54.8432	57.8518	1
OA...	52.7207	29.9973	1
OA...	49.2001	33.0739	
OA...	40.1062	35.2281	

### 15.7 Guardar mejor conformación

Si queremos guardar el complejo entre la mejor pose de unión y la proteína, debemos seleccionar la conformación que nos interesa y después ir a “File”, “Save As...”. Posteriormente, seleccionamos la entrada que nos interesa y hacemos click sobre “OK”. Despues, escogemos el destino del archivo, lo nombramos, revisamos que se guarde en el formato que nos interesa, \*.pdb y hacemos click en “Guardar”.



El complejo que creamos nos servirá más adelante.

## 16 Cómo calcular RMSD

### 16.1 Teoría

Los estudios de acoplamiento molecular, como cualquier otra metodología, deben de ser validados. Para validar *dockings* existen varios métodos, uno de los más usados es la “selección de pose”. En este método los programas de *docking* se utilizan para *redockear* un compuesto con una conformación y orientación conocida dentro del sitio activo de la proteína, generalmente de una estructura co-cristalizada. La raíz de la desviación cuadrática media (RMSD, por sus siglas en inglés *Root Mean Squared Deviation*) es una característica que se utiliza con regularidad para comparar diferentes conformaciones de un sistema molecular dado. En este sentido, la RMSD es muy útil para el análisis y la validación de los resultados al medir la calidad de la reproducción de una pose de unión conocida de una estructura cristalográfica y compararla con una pose de unión obtenida. En general, se considera que los programas que logran obtener poses con una RMSD  $\leq 1.5 - 2 \text{ \AA}$  (dependiendo del tamaño del ligando) se han desempeñado correctamente. Con base en este valor, se puede hacer el análisis y la comparación entre distintos *softwares* y determinar cuál de ellos arrojó los mejores resultados.<sup>21,22</sup>

Existen diversas maneras de obtener la RMSD, una de ellas es con el uso de servidores web, por ejemplo, **LigRMSD**. Este es un servidor web gratuito que elabora el cálculo de RMSD entre compuestos químicos idénticos o similares. Para realizarlo, LigRMSD requiere al menos un par de moléculas en formato \*.mol. De las moléculas que se escogen, una de ellas debe ser seleccionada como la “referencia”, para que así, todas las demás moléculas se comparen contra esta.

### 16.2 Ejemplo

A continuación, veremos los pasos para determinar RMSD de los tres programas con los que trabajamos (AD4.2, VINA y GOLD).

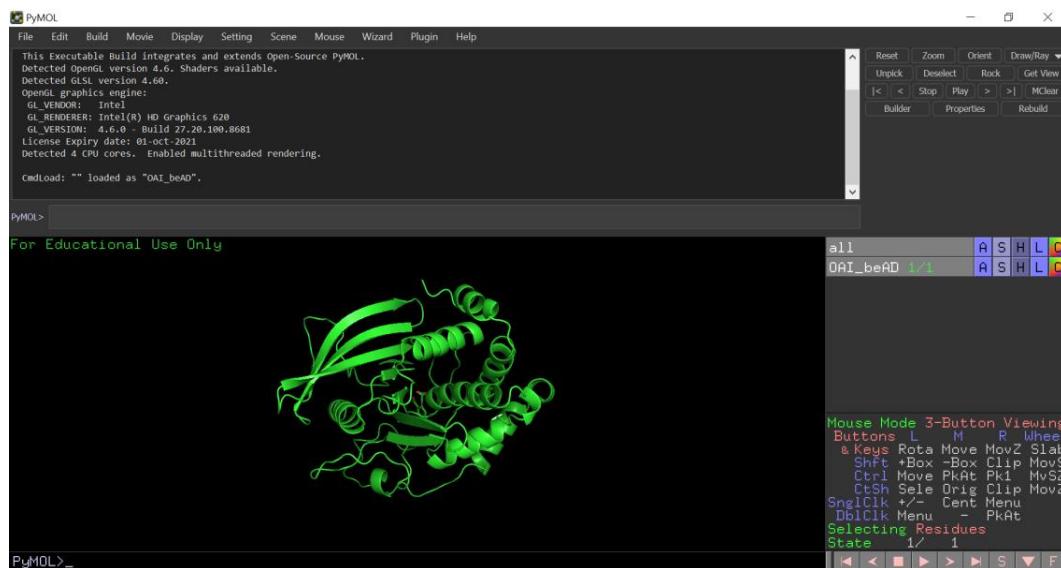
Como se mencionó anteriormente, necesitamos definir qué moléculas compararemos. En este ejemplo, realizamos los *dockings* del compuesto OAI, que es el mismo que se encuentra cocristalizado con la estructura, por lo tanto, podemos usar la mejor pose de unión que obtuvimos en cada uno de los programas para determinar la RMSD.

En este ejemplo, nuestro set estará compuesto por cuatro moléculas:

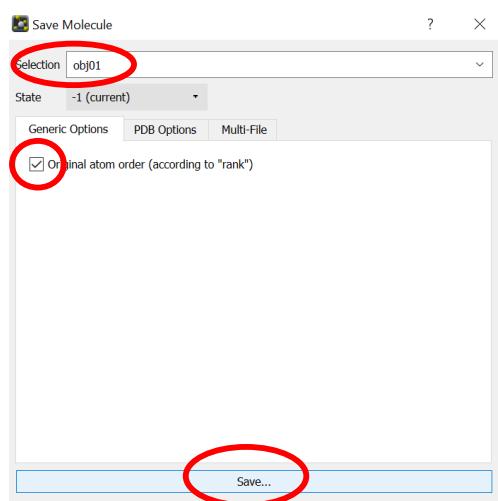
1. Ligando cocristalizado (OAI\_cocryst.pdb) Ver sección “cómo preparar una proteína”
2. Mejor pose de unión de AD4.2 (OAI\_beAD.pdb) Ver sección “*Docking* en Autodock4.2”
3. Mejor pose de unión de VINA (OAI\_beVINA.pdb) Ver sección “*Docking* en Autodock VINA”
4. Mejor pose de unión de GOLD (OAI\_beGOLD.pdb) Ver sección “*Docking* en GOLD”

Es muy importante recordar que las mejores poses de unión de los programas están complejadas con la proteína. Para el cálculo de la RMSD necesitamos solamente al ligando. Además, estos archivos los tenemos en formato \*.pdb y LigRMSD solamente acepta archivos en formato \*.mol. El primer inconveniente lo podemos solucionar con ayuda de PyMOL, mientras que la conversión a formato \*.mol se puede hacer en Avogadro.

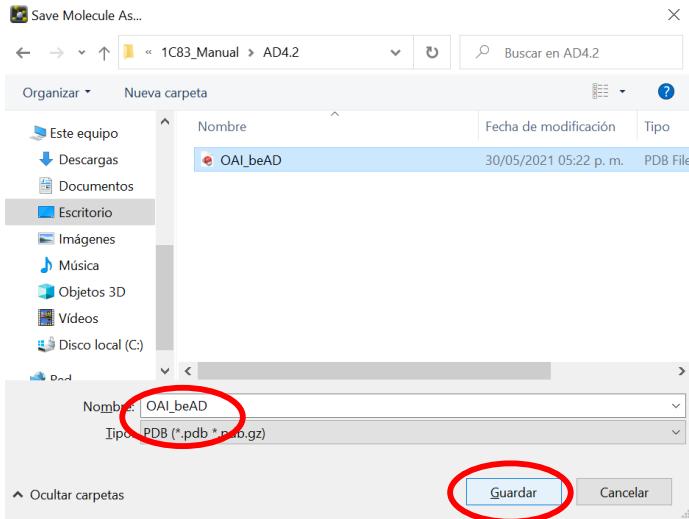
El primer paso será entonces abrir PyMOL. Posteriormente, abriremos el archivo **OAI\_beAD.pdb**.



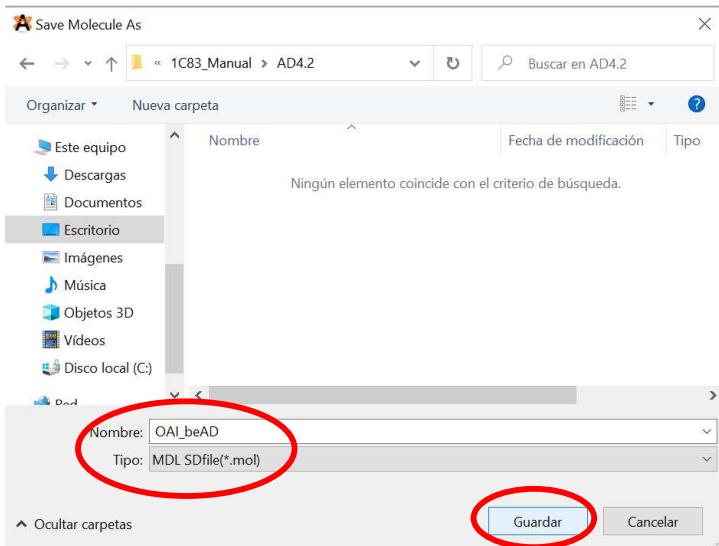
Centramos en el ligando y lo extraemos, como aprendimos en la sección “Cómo preparar una proteína”. Una vez que se ha extraído extraído, lo guardamos en formato \*.pdb, seleccionando “File”, “Export Molecule” y hacemos click en “Save”.



Posteriormente, aparecerá otra ventana, seleccionamos el destino del archivo y lo nombramos o sobreescribimos. El punto importante es no confundirnos con los archivos. En este ejemplo, lo sobreescribiremos.



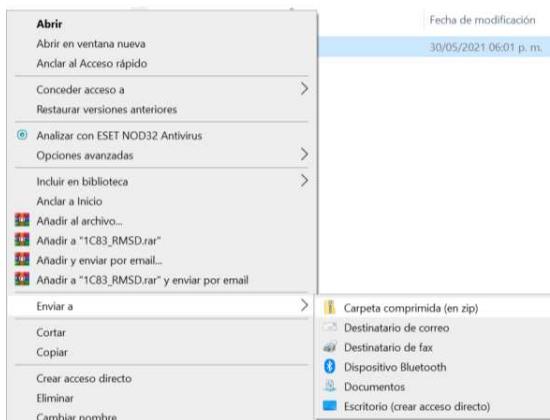
Después, nos dirigimos al programa Avogadro y abrimos el archivo que acabamos de guardar. Una vez abierto, simplemente vamos a “File”, “Save as...”, escogemos el destino y verificamos que esté seleccionado el formato que nos interesa, \*.mol y guardamos.



Ahora debemos repetir estos pasos para los demás complejos. Cuando tengamos los archivos \*.mol de todas las moléculas, debemos organizarlas en una misma carpeta, que debería de verse así:

Nombre	Fecha de modificación	Tipo	Tamaño
OAI_beAD	30/05/2021 05:53 p. m.	Mol File	2 KB
OAI_beGOLD	30/05/2021 05:56 p. m.	Mol File	3 KB
OAI_beVINA	30/05/2021 05:57 p. m.	Mol File	2 KB
OAI_cocryst	30/05/2021 05:58 p. m.	Mol File	3 KB

Posteriormente, debemos de comprimir esta carpeta a formato \*.zip. Esto lo podemos lograr, haciendo click derecho sobre la carpeta, después “enviar a” y seleccionamos “carpeta comprimida”.



Como veremos, aparece ahora una carpeta comprimida. Será esta la que subiremos a LigRMSD.

Nombre	Fecha de modificación	Tipo	Tamaño
1C83_RMSD	30/05/2021 06:01 p. m.	Carpeta de archivos	
1C83_RMSD	30/05/2021 06:01 p. m.	Archivo WinRAR ZIP	3 KB

En la siguiente imagen, veremos la portada del servidor web. Esta página se encuentra en <https://ligrmsd.appsbio.ualca.cl/>. Para comenzar, debemos hacer click sobre “Submit a Job”

Root Mean Squared Deviation (RMSD) is a commonly used feature for comparing different conformations of given molecular systems. In Molecular Dynamics (MD) and Protein-ligand Docking (PLD) approaches, for example, this measure is essential to the analysis and validation of the results. In the case of PLD, RMSD is routinely utilized to measure the quality of reproduction of a known binding pose by comparing the obtained ligand pose with experimental crystallographic source. For this, components might show a full atomic equivalence. Most methods available for RMSD calculations assume an equivalence between the atoms of the molecules, requiring uniquely the atom indices for the measures. Unfortunately, the vast majority of the 3D molecule files are generated using strategies of random atom labeling and rarely consider the possible symmetry of certain types of atoms or fragments (Coutsias and Wester, 2019). Thus, to avoid inaccurate RMSD values,

En la pantalla aparecerá ahora una opción en donde podemos seleccionar nuestro archivo. Hacemos click sobre ella.

The screenshot shows the LigRMSD v1.0 web interface. At the top, there are links for "Submit a Job", "Cases of use", "Step by step", and "Screencast". The main title is "LigRMSD: A web server for automatic structure matching and RMSD calculations among identical and similar compounds in protein-ligand docking". Below the title, there is a "Submit a Job" button. Under the "File" section, there is a text input field with the placeholder "Seleccionar archivo" and a tooltip "Ningún archi... seleccionado". Below the input field are two buttons: "Upload" (blue) and "Allowed format file" (red). The "Upload" button is circled in red. At the bottom of the page, there is a footer with the text "School of Bioinformatics Engineering | Department of Bioinformatics | Universidad de Talca | Talca, Chile".

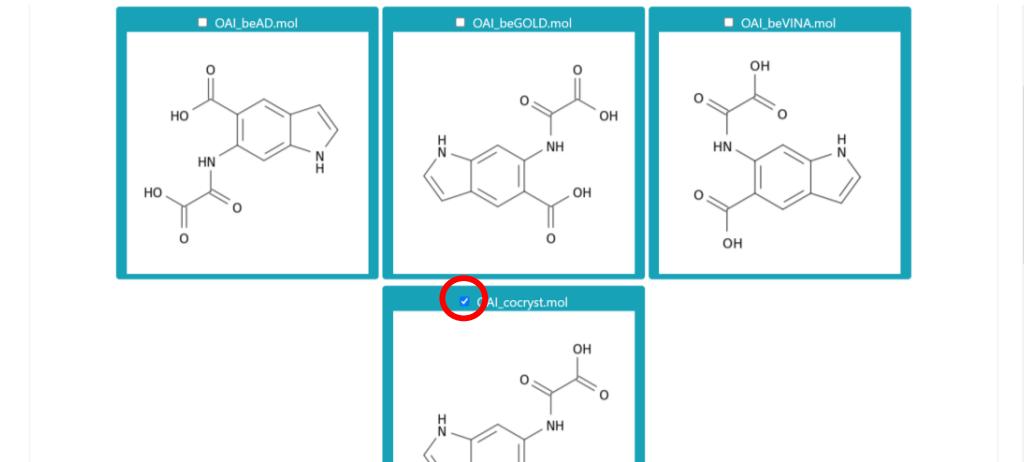
Buscamos nuestra carpeta, la seleccionamos y hacemos click sobre “Abrir”

The screenshot shows a Windows file explorer window titled "Abrir" (Open). The path is "Docking Manual > 1C83 > 1C83\_Manual > 1C83 >". The file "1C83\_RMSD.zip" is selected and highlighted with a red circle. In the bottom right corner of the file explorer, the "Abrir" (Open) button is also circled in red. The background shows the LigRMSD v1.0 web interface with the "File" section and the "Upload" button.

Ya que el archivo está cargado, podemos hacer click sobre “Upload”.

The screenshot shows the LigRMSD v1.0 web interface again. The "File" section now shows the uploaded file "1C83\_RMSD.zip" in the "Seleccionar archivo" input field. The "Upload" button, which was previously circled in red, is now highlighted with a blue border. The "Allowed format file" button is still visible next to it. The rest of the interface remains the same, including the "Submit a Job" button and the footer.

Cuando se cargan nuestros archivos, aparecerán como en la siguiente imagen. Es aquí donde tenemos que señalarle a LigRMSD quién es la referencia. En este caso, la referencia es el archivo “OAI\_cocryst.mol”, por lo tanto, marcamos su casilla.



Si nos desplazamos más abajo en la página, veremos que se nos piden algunos datos extra, como nombre del trabajo, límite y un correo electrónico. El nombre es obligatorio, mientras que el correo no lo es, el parámetro de límite lo dejaremos como default. Una vez que hemos llenado esos datos, hacemos click sobre “Execute”.

Job name  
1C83\_RMSD

Name of job.

Limit  
30

Represents the percent of match with reference by default is 30.0 percent but you could replace it with other values of your choice.

Email  
Email

An email will be sent including the link for your job. The access to your job results is with the URL link.  
Your email only will be used to send the link with the results.

Mientras se está realizando el cálculo, la página se verá como en la siguiente imagen.

LigRMSD v1.0 [Submit a Job](#)

Cases of use Step by step Screen cast

LigRMSD: A web server for automatic structure matching and RMSD calculations among identical and similar compounds in protein-ligand docking

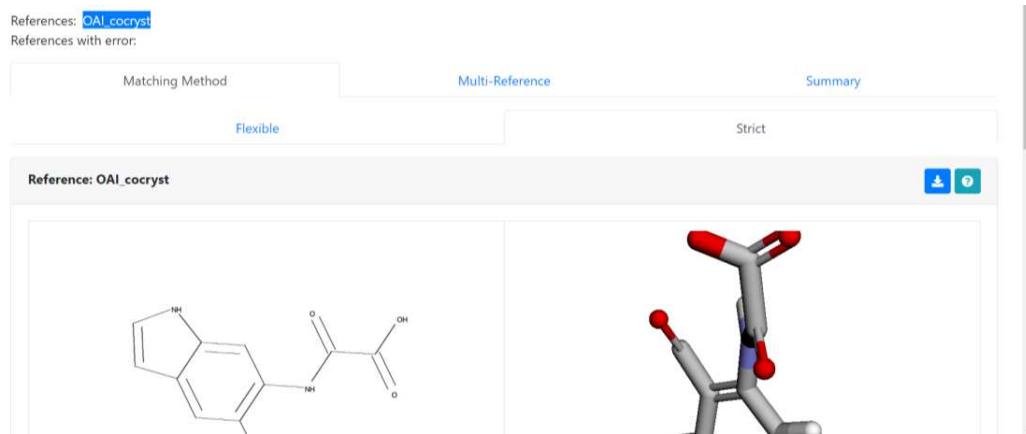
Your job is now running...

This page will refresh in 8 seconds.

The results will be displayed on this page.

Please bookmark this page or record the link below to return to your results:

Una vez que ha terminado, observaremos lo siguiente:



Si nos desplazamos más hacia abajo, encontraremos los valores exactos de RMSD para cada archivo. En este ejemplo, podemos observar que obtuvimos valores de  $\text{RMSD} \leq 1.5 \text{ \AA}$ . Por lo tanto, podemos decir que nuestros resultados están validados para los 3 programas. Si comparamos entre ellos, de los tres, el que obtuvo un mejor resultado fue VINA, ya que tiene el RMSD más bajo, seguido de AD y finalmente GOLD.

Molecule	Image	RMSD	Percent of ref match	Percent of molecule match	Details	Download
OAI_beAD		0.41	100.00	100.00	<a href="#">View</a>	<a href="#">Download</a>
OAI_beGOLD		0.55	100.00	100.00	<a href="#">View</a>	<a href="#">Download</a>
OAI_beVINA		0.34	100.00	100.00	<a href="#">View</a>	<a href="#">Download</a>

Es conveniente apuntar estos valores, ya que nos podrían servir después para discutir y analizar resultados.

## 17 Ejercicios

Ahora que hemos revisado los fundamentos teóricos y hemos realizado un ejemplo de cómo hacer un *docking*, es hora de poner en práctica lo aprendido y hacer algunos ejercicios.

Recordemos que la metodología del *docking* es cambiante y moldeable. Por lo tanto, no debería de sorprendernos encontrar diferentes especificaciones para algunos problemas en concreto.

Si nos surgieran dudas, podemos regresar y visitar las secciones anteriores, o bien, ir a las referencias que se nos proporcionan.

### ***17.1 Ejercicio 1***

La enfermedad de Chagas o tripanosomiasis de América es una enfermedad crónica causada por el parásito *Trypanosoma cruzi*. Actualmente, para su tratamiento solo se cuenta con dos medicamentos: benznidazol (N-benzyl-2-(2-nitro-1H-imidazole-1-yl)acetamide, Bnz) y nifurtimox (3-methyl-N-[5-nitrofuran-2-yl)methylidene]thiomorpholin-4-amine-1,1-dioxide, Nfx). En la última década, se ha propuesto a la trans-sialidasa como una diana terapéutica para fármacos anti-Chagas. Con base en esto, se han propuesto algunos derivados del ácido benzoico como posibles inhibidores de la trans-sialidasa, así como agentes anti-tripanosomales.<sup>23</sup>

Realice el acoplamiento molecular de los compuestos **11, 13, 14, 15, 19, 22, Nfx y Bzn**, con **los tres programas**, para la proteína **1MS8** con base en las siguientes especificaciones:

#### 1MS8

Nota: utilice el PDB ID: 1MS8, remueva moléculas de H<sub>2</sub>O, ligandos y la subunidad B. Envíe a minimizar en YASARA. Una vez minimizada, convierta a los formatos correspondientes de acuerdo con las necesidades de cada programa.

- Autodock4.2

Grid box: 60 x 60 x 60 Å

Spacing: 0.375

Coordenadas: x = 40.832      y = 63.762      z = -37.418

GA runs: 100

- Autodock VINA

Grid box: 22.5 x 22.5 x 22.5 Å

Coordenadas: x = 40.832      y = 63.762      z = -37.418

Num\_modes = 500

- GOLD

Tamaño del sitio: 10 Å

Coordenadas: x = 40.832      y = 63.762      z = -37.418

GA runs: 100

Número de operaciones = 125000

Llene la siguiente tabla con sus resultados y vea si existe correlación con la actividad inhibitoria sobre esta enzima.

## 17.2 Ejercicio 2

La familia de citocromos P450 (CYPs) es un grupo de hemoproteínas relacionadas con el metabolismo de una gran variedad de compuestos endógenos y xenobióticos. La enzima CYP1B1 es parte de esta familia y es una diana terapéutica interesante por varias razones, una de ellas es su asociación con la farmacorresistencia, puesto que metaboliza e inactiva algunos agentes anticancerosos como el docetaxel, paclitaxel y el cisplatino.<sup>24</sup>

Realice el acoplamiento molecular de los compuestos **2b**, **8b**, **13b**, **26** y **ANF** en **GOLD** para la proteína **3PM0** con base en las siguientes especificaciones:

### 3PM0

Nota: utilice el PDB ID: 3PM0, cárguelo directamente en GOLD. Con ayuda del Wizard, remueva moléculas de agua y extraiga el ligando (BHF) pero no el grupo hemo. Centre el “grid” en la molécula de naftoflavona (con clave BHF) que se extrajo.

Tamaño del sitio: 10 Å

Coordinadas: centrar en la molécula de BHF

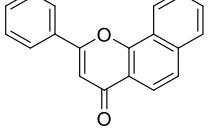
Template: goldscore\_p450\_csd

GA runs: 100

Número de operaciones = 125000

Llene la siguiente tabla con sus resultados y vea si existe correlación con la actividad inhibitoria sobre esta enzima.

Compuesto	Estructura	GOLD SCORE	CHEMPLP-Fitness score	CI <sub>50</sub> (μM)
2b				
8b				
13b				
26				

ANF				
-----	---	--	--	--

### 17.3 Ejercicio 3

La familia de citocromos P450 (CYPs) es un grupo grande de hemoproteínas que están involucradas en diversas reacciones de oxidación y reducción tanto de compuestos endógenos como xenobióticos de distintos tamaños. CYP1B1 es uno de los tres miembros que conforman esta familia. Este se expresa generalmente en tejidos como ovarios, testículos y glándulas adrenales, además de glándulas mamarias, útero y próstata. La inhibición de esta diana es una estrategia terapéutica prometedora, ya que tal inhibidor podría modular la bioactivación de procarcinógenos y al mismo tiempo reducir la fármaco-resistencia.<sup>25</sup>

Realice el acoplamiento molecular de los compuestos **3a, 3b, 4a, 4b, 6a, 6b, 7a, 7b, 11a, 11b, 12a, 12b, E1, E2** y **ANF** en **GOLD** con base en las siguientes especificaciones:

3PM0

Nota: utilice el PDB ID: 3PM0, cárguelo directamente en GOLD. Con ayuda del Wizard, remueva moléculas de agua y extraiga el ligando (BHF) pero no el grupo hemo. Centre el “grid” en la molécula de naftoflavona (con clave BHF) que se extrajo.

Tamaño del sitio: 10 Å

Coordenadas: centrar en la molécula de BHF

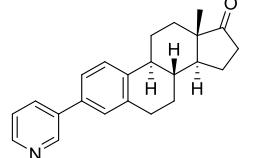
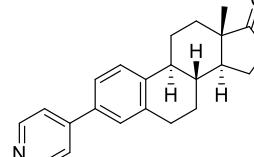
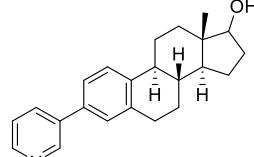
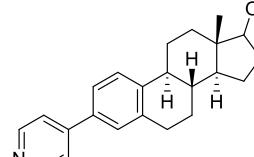
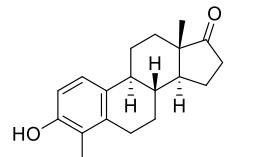
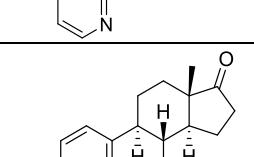
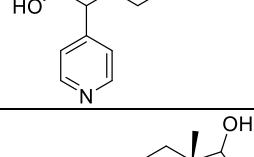
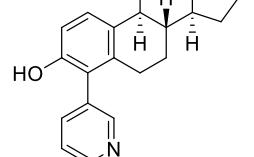
Template: goldscore\_p450\_csd

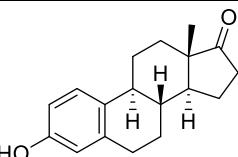
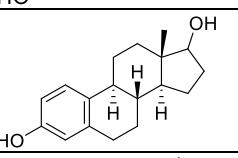
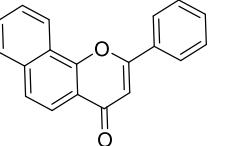
GA runs: 100

Número de operaciones = 125000

Llene la siguiente tabla con sus resultados y vea si existe correlación con la actividad inhibitoria sobre esta enzima.

Compuesto	Estructura	GOLD SCORE	CHEMPLP-Fitness score	IC <sub>50</sub> (μM)
3a				
3b				
4a				
4b				

6a				
6b				
7a				
7b				
11a				
11b				
12a				
12b				

E1				
E2				
ANF				

#### ***17.4 Ejercicio 4***

La diabetes mellitus tipo 2 (DM2) es una enfermedad crónica caracterizada por altos niveles de glucosa en plasma, que trae consigo falla de órganos, especialmente de los nervios, ojos, vasos sanguíneos, corazón y riñones. En la actualidad existen diferentes propuestas para su tratamiento. Una de ellas es la inducción de la sensibilización a la insulina a través del agonismo de los Receptores Activos Proliferadores de Peroxisomas (PPARs por sus siglas en inglés). Existen tres clases de PPARs que son: PPAR $\alpha$ , PPAR $\beta/\delta$  y PPAR $\gamma$ . La *Cordia morelosana*, comúnmente conocida como anacahuite o palo prieto, es un árbol que se utiliza en la medicina tradicional para el tratamiento de ciertos padecimientos, incluyendo la diabetes.<sup>26</sup>

Realice el acoplamiento molecular de los compuestos extraídos de la *Cordia morelosana*, **MR** y **RA**, con los tres programas, para dos proteínas (**1I7I** y **5HYK**) con base en las siguientes especificaciones:

##### **1I7I**

Nota: utilice el PDB ID: 1I7I, utilice la proteína sin minimizar, simplemente remueva moléculas de H<sub>2</sub>O y ligandos. Revise que la proteína tenga átomos de H. Convierta a los formatos correspondientes de acuerdo con las necesidades de cada programa.

- Autodock4.2

Grid box: 50 x 50 x 50 Å

Spacing: 0.375

Coordenadas: x = 16.847      y = 16.639      z = 12.538

GA runs: 100

- Autodock VINA

Grid box: 18.75 x 18.75 x 18.75 Å

Coordenadas: x = 16.847      y = 16.639      z = 12.538

Num\_modes = 500

- GOLD

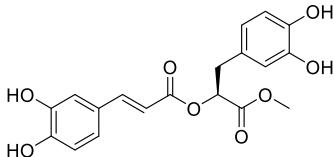
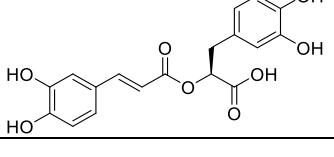
Tamaño del sitio: 11 Å

Coordenadas: x = 16.847      y = 16.639      z = 12.538

GA runs: 100

Número de operaciones = 125000

Llene la siguiente tabla con sus resultados y vea si existe correlación con la actividad inhibitoria sobre esta enzima.

Compuesto	Estructura	AD4.2	VINA	GOLD SCORE	CHEMPLP-Fitness score	Expresión relativa de niveles de mRNA del receptor PPAR $\gamma$
MR						
RA						

### 5HYK

Nota: utilice el PDB ID: 5HYK, remueva moléculas de H<sub>2</sub>O y mande a minimizar en YASARA con el ligando. Una vez minimizada, extraiga el ligando. Convierta a los formatos correspondientes de acuerdo con las necesidades de cada programa.

- Autodock4.2

Grid box: 60 x 60 x 60 Å

Spacing: 0.375

Coordenadas: x = 9.087                    y = 34.705            z = 17.038

GA runs: 100

- Autodock VINA

Grid box: 22.5 x 22.5 x 22.5 Å

Coordenadas: x = 9.087                    y = 34.705            z = 17.038

Num\_modes = 500

- GOLD

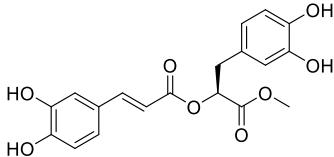
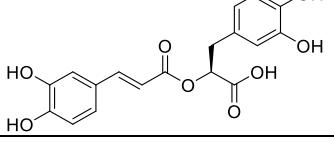
Tamaño del sitio: 11 Å

Coordenadas: x = 9.087      y = 34.705      z = 17.038

GA runs: 100

Número de operaciones = 125000

Llene la siguiente tabla con sus resultados y vea si existe correlación con la actividad inhibitoria sobre esta enzima.

Compuesto	Estructura	AD4.2	VINA	GOLD SCORE	CHEMPLP-Fitness score	Expresión relativa de niveles de mRNA del receptor PPAR $\alpha$
MR						
RA						

### **17.5 Ejercicio 5**

La *Actinidia arguta*, comúnmente conocida como kiwiberry o kiwiño, es una planta que se cultiva en Asia. Recientemente se extrajeron de ella nuevos triterpenos con constituyentes del tipo fenilpropanoides. La  $\alpha$ -glucosidasa es una enzima catalizadora de la digestión de carbohidratos. Su inhibición retardaría la digestión y absorción de carbohidratos, lo que resultaría en un efecto hipoglucémico. Los inhibidores de la  $\alpha$ -glucosidasa son de interés ya que podrían ser beneficiosos para el tratamiento y la prevención de la diabetes y obesidad al controlar los niveles de glucosa en plasma. En este sentido, los triterpenos con constituyentes fenilpropanoides podrían ser una opción.<sup>27</sup>

Realice el acoplamiento molecular de los compuestos **1, 2, 3, 4, 5, 6 y 14, con los tres programas**, para dos proteínas (**2QMJ** y **3TOP**) con base en las siguientes especificaciones:

#### 2QMJ

Nota: utilice el PDB ID: 2QMJ, remueva moléculas de H<sub>2</sub>O y mande a minimizar en YASARA con el ligando. Una vez minimizada, extraiga el ligando. Convierta a los formatos correspondientes de acuerdo con las necesidades de cada programa.

- Autodock4.2

Grid box: 50 x 50 x 50 Å

Spacing: 0.375

Coordenadas: x = -24.027    y = -10.733    z = -9.808

GA runs: 100

- Autodock VINA

Grid box: 18.75 x 18.75 x 18.75 Å

Coordenadas: x = -24.027    y = -10.733    z = -9.808

Num\_modes = 500

- GOLD

Tamaño del sitio: 10 Å

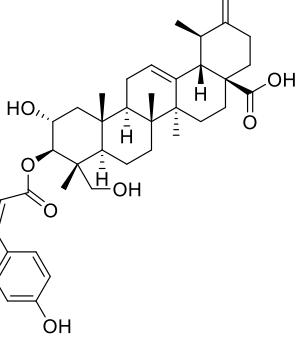
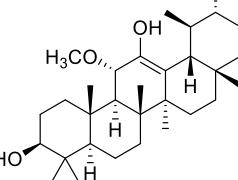
Coordenadas: x = -24.027    y = -10.733    z = -9.808

GA runs: 100

Número de operaciones = 125000

Llene la siguiente tabla con sus resultados y vea si existe correlación con la actividad inhibitoria sobre esta enzima.

Compuesto	Estructura	AD4.2	VINA	GOLD SCORE	CHEMPLP -Fitness score	% inhibición (100 µM)
1						
2						
3						
4						
5						

6						
14						

### 3TOP

Nota: utilice el PDB ID: 3TOP, remueva moléculas de H<sub>2</sub>O y mande a minimizar en YASARA con el ligando. Una vez minimizada, extraiga el ligando y borre la subunidad B. Convierta a los formatos correspondientes de acuerdo con las necesidades de cada programa.

- Autodock4.2

Grid box: 60 x 60 x 60 Å

Spacing: 0.375

Coordenadas: x = -34.13      y = 32.452      z = 30.044

GA runs: 100

- Autodock VINA

Grid box: 22.5 x 22.5 x 22.5 Å

Coordenadas: x = -34.13      y = 32.452      z = 30.044

Num\_modes = 500

- GOLD

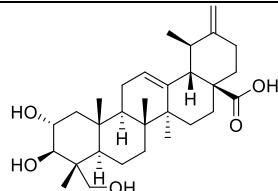
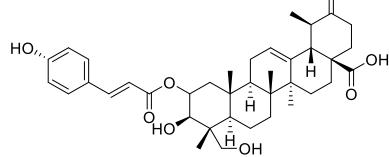
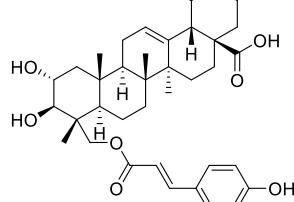
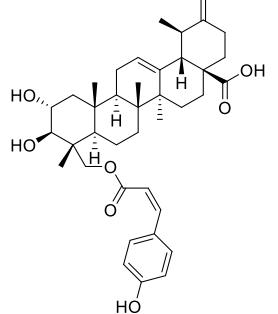
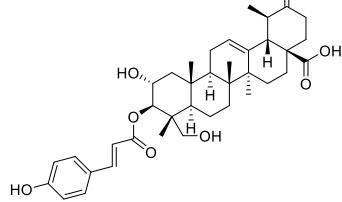
Tamaño del sitio: 10 Å

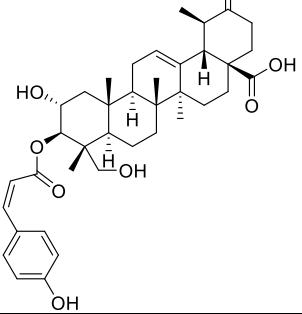
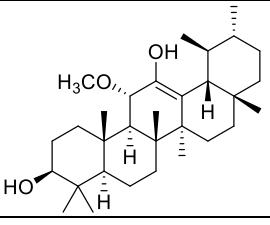
Coordenadas: x = -34.13      y = 32.452      z = 30.044

GA runs: 100

Número de operaciones = 125000

Llene la siguiente tabla con sus resultados y vea si existe correlación con la actividad inhibitoria sobre esta enzima.

Compuesto	Estructura	AD4.2	VINA	GOLD SCORE	CHEMPLP-Fitness score	% inhibición (100 µM)
1						
2						
3						
4						
5						

6					
14					

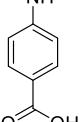
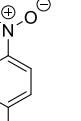
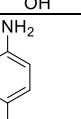
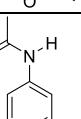
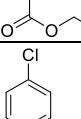
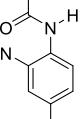
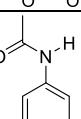
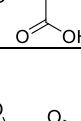
## 18 Referencias

1. Saldívar-González, F. I., Chávez-Ponce de León, D. E., López-López, E., Hernández-Luis, F., Lira-Rocha, A., Medina-Franco, J. L. (2018) *Manual de Químicoinformática* (2da ed.)
2. Vernin-Barbosa, N. S., Rocha de Almeida-Lima, E., Wanderley-Tavares, F. (2017) *Molecular Modeling in Chemical Engineering*. Elsevier
3. Jaitlet, L., Artemova, S., & Redon, S. (2017). IM-UFF: Extending the universal force field for interactive molecular modeling. *Journal of Molecular Graphics and Modelling*, 77, 350-362.
4. Smith, W. R., & Qi, W. (2018). Molecular simulation of chemical reaction equilibrium by computationally efficient free energy minimization. *ACS central science*, 4(9), 1185-1193.
5. Sousa, S. F., Ribeiro, A. J., Neves, R. P., Brás, N. F., Cerqueira, N. M., Fernandes, P. A., & Ramos, M. J. (2017). Application of quantum mechanics/molecular mechanics methods in the study of enzymatic reaction mechanisms. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 7(2), e1281.
6. Meng, X. Y., Zhang, H. X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design*, 7(2), 146–157.
7. Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews Drug discovery*, 3(11), 935-949.
8. Medina-Franco, J. L., Fernández-de Gortari, E., & Naveja, J. J. (2015). Avances en el diseño de fármacos asistido por computadora. *Educación química*, 26(3), 180-186.
9. Dias, R., de Azevedo, J., & Walter, F. (2008). Molecular docking algorithms. *Current drug targets*, 9(12), 1040-1047.
10. Guedes, I. A., de Magalhães, C. S., & Dardenne, L. E. (2014). Receptor–ligand molecular docking. *Biophysical reviews*, 6(1), 75-87.
11. Pagadala, N. S., Syed, K., & Tuszynski, J. (2017). Software for molecular docking: a review. *Biophysical reviews*, 9(2), 91-102.
12. Fan, J., Fu, A., & Zhang, L. (2019). Progress in molecular docking. *Quantitative Biology*, 1-7.
13. Pinzi, L., & Rastelli, G. (2019). Molecular docking: Shifting paradigms in drug discovery. *International journal of molecular sciences*, 20(18), 4331.
14. Halperin, I., Ma, B., Wolfson, H., & Nussinov, R. (2002). Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics*, 47(4), 409-443.
15. Prieto-Martínez, F. D., Arciniega, M., & Medina-Franco, J. L. (2018). Molecular docking: current advances and challenges. *TIP. Revista especializada en ciencias químico-biológicas*, 21.
16. Andersen, H. S., Iversen, L. F., Jeppesen, C. B., Branner, S., Norris, K., Rasmussen, H. B., ... & Møller, N. P. H. (2000). 2-(oxalamino)-benzoic acid is a general, competitive inhibitor of protein-tyrosine phosphatases. *Journal of Biological Chemistry*, 275(10), 7101-7108.
17. Avogadro: Molecular Editor and Visualization <https://avogadro.cc/docs/>
18. Drawing a molecule in Avogadro <https://www.youtube.com/watch?v=UWR4hPbtrHE>

19. AutoDock 4.2.6 Download Page <http://autodock.scripps.edu/downloads/autodock-registration/autodock-4-2-download-page/>
20. GOLD User Guide. A Component of the CSD-Discovery Suite. (2019). Retrieved 1 June 2021, from [https://www.ccdc.cam.ac.uk/support-and-resources/ccdcresources/GOLD\\_User\\_Guide.pdf](https://www.ccdc.cam.ac.uk/support-and-resources/ccdcresources/GOLD_User_Guide.pdf).
21. Hevener, K. E., Zhao, W., Ball, D. M., Babaoglu, K., Qi, J., White, S. W., & Lee, R. E. (2009). Validation of molecular docking programs for virtual screening against dihydropteroate synthase. *Journal of chemical information and modeling*, 49(2), 444-460.
22. Velázquez-Libera, J. L., Durán-Verdugo, F., Valdés-Jiménez, A., Núñez-Vivanco, G., & Caballero, J. (2020). LigRMSD: a web server for automatic structure matching and RMSD calculations among identical and similar compounds in protein-ligand docking. *Bioinformatics*, 36(9), 2912-2914.
23. Kashif, M., Moreno-Herrera, A., Villalobos-Rocha, J. C., Nogueda-Torres, B., Pérez-Villanueva, J., Rodríguez-Villar, K., ... & Rivera, G. (2017). Benzoic acid derivatives with trypanocidal activity: Enzymatic analysis and molecular docking studies toward Trans-Sialidase. *Molecules*, 22(11), 1863.
24. Dutour, R., Roy, J., Cortés-Benítez, F., Maltais, R., & Poirier, D. (2018). Targeting Cytochrome P450 (CYP) 1B1 Enzyme with four series of a-ring substituted estrane derivatives: Design, synthesis, inhibitory activity, and selectivity. *Journal of medicinal chemistry*, 61(20), 9229-9245.
25. Dutour, R., Cortés-Benítez, F., Roy, J., & Poirier, D. (2017). Structure-based design and synthesis of new estrane-pyridine derivatives as cytochrome P450 (CYP) 1B1 Inhibitors. *ACS medicinal chemistry letters*, 8(11), 1159-1164.
26. Giles-Rivas, D., Estrada-Soto, S., Aguilar-Guadarrama, A. B., Almanza-Pérez, J., García-Jiménez, S., Colín-Lozano, B., ... & Villalobos-Molina, R. (2020). Antidiabetic effect of Cordia morelosana, chemical and pharmacological studies. *Journal of ethnopharmacology*, 251, 112543.
27. Ahn, J. H., Park, Y., Yeon, S. W., Jo, Y. H., Han, Y. K., Turk, A., ... & Lee, M. K. (2020). Phenylpropanoid-Conjugated Triterpenoids from the Leaves of Actinidia arguta and Their Inhibitory Activity on  $\alpha$ -Glucosidase. *Journal of natural products*, 83(5), 1416-1423.

19 Respuestas

## 19.1 Ejercicio 1

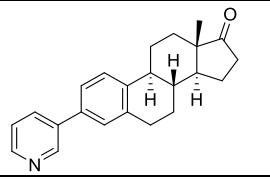
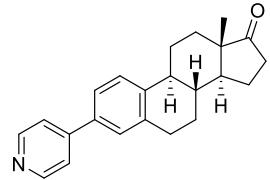
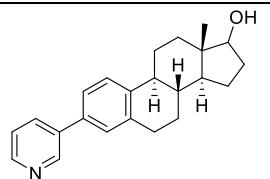
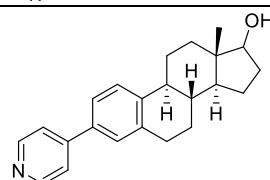
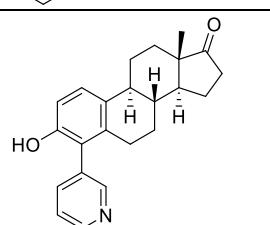
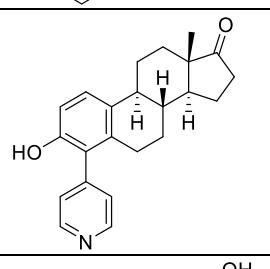
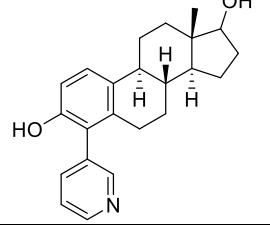
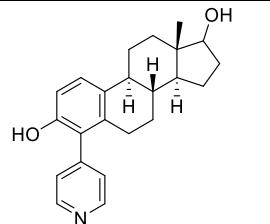
Compuesto	Estructura	AD4.2	VINA	GOLD SCORE	CHEMPLP-Fitness score	CL <sub>50</sub> (μM) NINOA	CL <sub>50</sub> (μM) INC-5
11		-4.42 [53]	-5.7	29.22	43.19	0.66 ± 0.39	0.58 ± 0.4
13		-4.45 [54]	-5.9	28.21	46.48	0.47 ± 0.16	0.46 ± 0.38
14		-4.28 [44]	-5.6	33.67	41.89	0.10 ± 0.041	0.10 ± 0.047
15		-5.31 [49]	-6.3	35.19	48.69	0.34 ± 0.18	0.21 ± 0.1
19		-6.31 [88]	-8.1	38.55	59.99	0.14 ± 0.08	0.0008 ± 0.0001
22		-5.22 [51]	-6.2	35.53	51.04	1.28 ± 0.26	1.28 ± 0.31
Nfx		-5.75 [54]	-6.2	41.79	48.90	0.213 ± 0.08	0.68 ± 0.17
Bzn		-6.07 [29]	-7.2	33.20	58.25	0.292 ± 0.12	0.62 ± 0.28

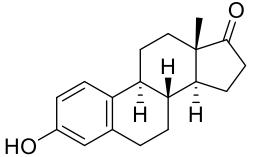
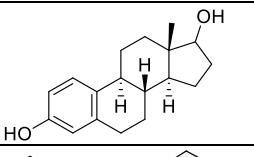
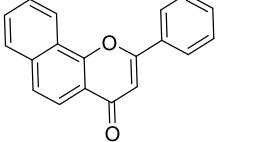
### 19.2 Ejercicio 2

Compuesto	Estructura	GOLD SCORE	CHEMPLP-Fitness score	$\text{IC}_{50}$ ( $\mu\text{M}$ )
2b		58.63	88.06	$0.24 \pm 0.05$
8b		71.57	94.75	$0.54 \pm 0.08$
13b		53.06	91.56	$0.24 \pm 0.04$
26		30.91	81.58	$0.37 \pm 0.09$
ANF		70.31	95.38	$0.24 \pm 0.08$

### 19.3 Ejercicio 3

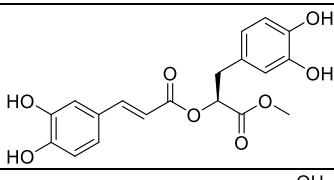
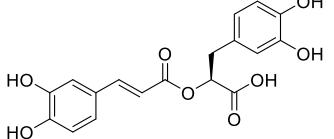
Compuesto	Estructura	GOLD SCORE	CHEMPLP-Fitness score	$\text{IC}_{50}$ ( $\mu\text{M}$ )
3a		50.82	88.13	0.063
3b		51.15	87.78	0.120
4a		60.50	91.30	0.011
4b		31.60	90.50	0.032

6a		19.91	89.98	ND
6b		13.53	78.25	ND
7a		46.62	91.16	ND
7b		22.87	87.03	ND
11a		41.28	83.14	ND
11b		37.55	82.83	ND
12a		35.35	80.44	ND
12b		31.00	80.51	ND

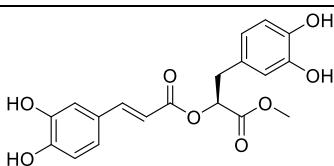
E1		42.09	85.39	ND
E2		40.67	85.13	ND
ANF		67.91	95.42	0.083

#### 19.4 Ejercicio 4

1I7I

Compuesto	Estructura	AD4.2	VINA	GOLD SCORE	CHEMPLP-Fitness score	Expresión relativa de niveles de mRNA del receptor PPAR $\gamma$
MR		-8.32 [25]	-8.3	51.8	81.63	3.9
RA		-7.88 [24]	-8.0	47.8	72.74	1.2

#### 5HYK

Compuesto	Estructura	AD4.2	VINA	GOLD SCORE	CHEMPLP-Fitness score	Expresión relativa de niveles de mRNA del receptor PPAR $\alpha$
MR		-8.5 [35]	-8.6	51.58	73.73	4.2

RA		-8.35 [15]	-9.1	62.39	79.09	1.2
----	--	---------------	------	-------	-------	-----

19.5 Ejercicio 5

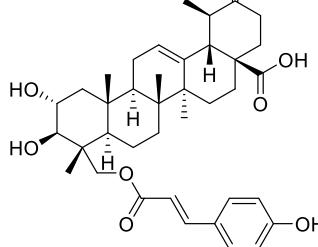
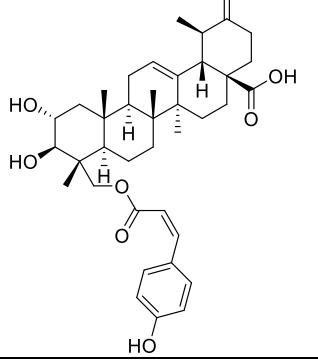
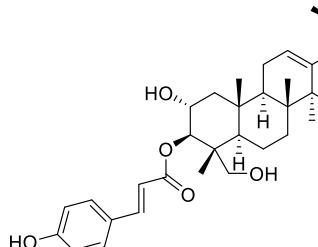
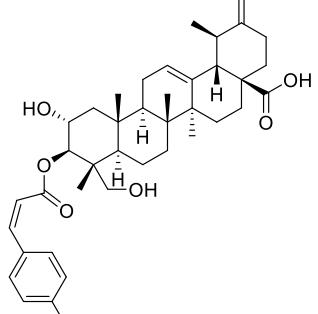
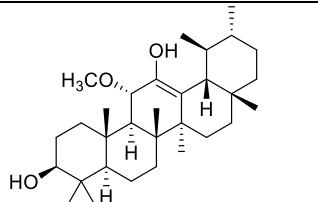
2QMJ

Compuesto	Estructura	AD4.2	VINA	GOLD SCORE	CHEMPLP-Fitness score	% inhibición (100 µM)
1		-7.88 [100]	-6.8	37.38	-24.46	13.0 ± 6.4
2		-8.62 [91]	-8.5	55.26	29.04	70.1 ± 3.5
3		-7.41 [86]	-7.5	65.35	47.98	59.9 ± 4.3
4		-7.6 [32]	-8.2	69.84	27.44	38.8 ± 3.3

5		-8.0 [41]	-7.0	58.1 9	17.54	65.8 ± 5.0
6		-8.15 [48]	-6.8	55.6 9	29.79	41.4 ± 3.3
14		-6.08 [60]	-6.9	39.2 4	20.07	23.9 ± 3.6

### 3TOP

Compuesto	Estructura	AD4.2	VINA	GOLD SCORE	CHEMPLP -Fitness score	% inhibición (100 µM)
1		-8.19 [91]	-7.1	50.9 9	11.65	13.0 ± 6.4
2		-9.85 [52]	-9.4	78.2 2	44.79	70.1 ± 3.5

3		- 8.33 [43]	-9.7 4	74.9	35.16	59.9 ± 4.3
4		- 9.39 [36]	- 10. 0	69.4	15.30	38.8 ± 3.3
5		- 8.28 [39]	-7.3 2	72.6	9.70	65.8 ± 5.0
6		-7.6 [46]	-7.6 7	66.5	17.07	41.4 ± 3.3
14		- 7.92 [44]	-6.9 2	51.5	34.75	23.9 ± 3.6

## 20 Anexo 1. Instalación de programas

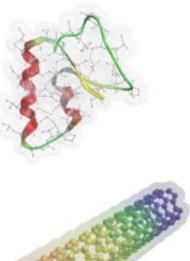
### 20.1 Avogadro

Ir a la página web <https://avogadro.cc/>

Hacer click sobre “Download”



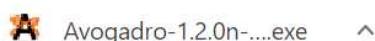
Avogadro is an advanced molecule editor and visualizer designed for cross-platform use in computational chemistry, molecular modeling, bioinformatics, materials science, and related areas. It offers flexible high quality rendering and a powerful plugin architecture.



En seguida, nos redireccionarán a otra página, en donde nuestra descarga **comenzará automáticamente**.

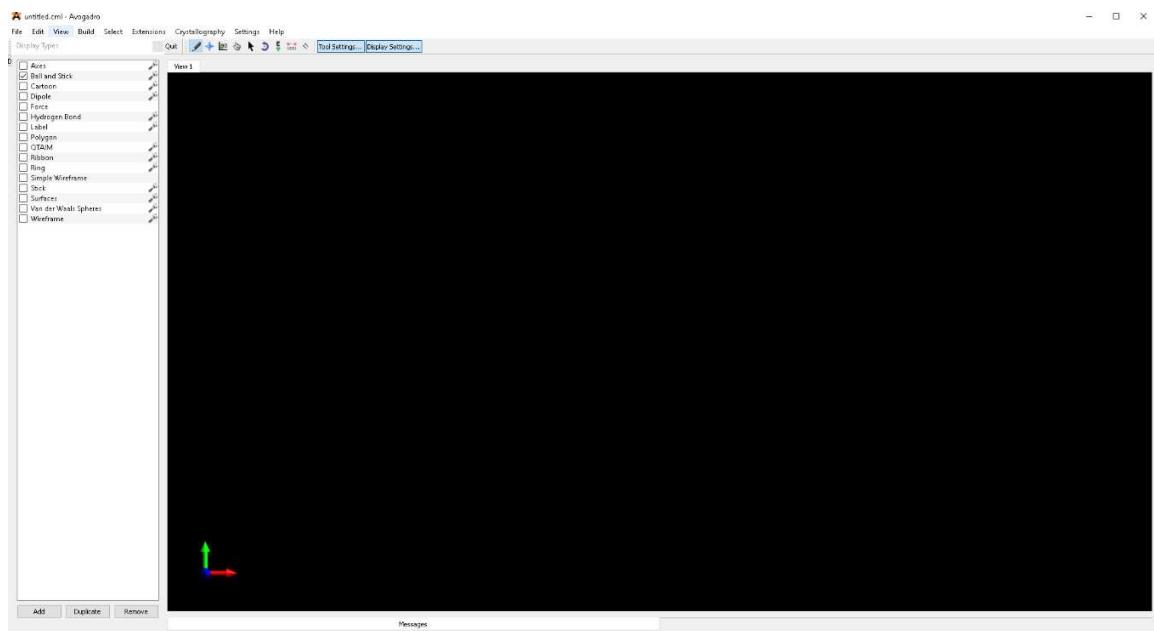
A screenshot of the SourceForge project page for Avogadro. The header includes links for Help, Create, Join, and Login. Below the header, there is an advertisement for Google. The main content area shows the Avogadro logo and a brief description: "An intuitive molecular editor and visualization tool. Brought to you by: cryosuk, ghutchis, timvdm". There is a link to "Learn more: check out screenshots, reviews, and more. We'll take you there in a few moments.". At the bottom, there are buttons for "Get Updates", "Share This", and "Problems Downloading?". A note at the bottom says "Avogadro-1.2.0n-win32.exe | Scanned for malware ✓". To the right, there is a sidebar with a "Mirror Provided by" section and a "Get latest updates about Open Source Projects, Conferences and News." section with a "Sign Up" button.

Una vez descargado, hacemos click sobre el ejecutable.



Una vez que se eligió la ruta de instalación y se instaló el programa, este debe iniciarse

haciendo doble click sobre el ícono de Avogadro (que puede estar en el escritorio o en el menú de inicio). La ventana debe verse así:



## 20.2 Autodock Tools (MGL Tools)

Ir a la página web <https://ccsb.scripps.edu/mgltools/downloads/>

Descargar la versión 1.5.6 para el sistema operativo Windows

---

Version 1.5.6

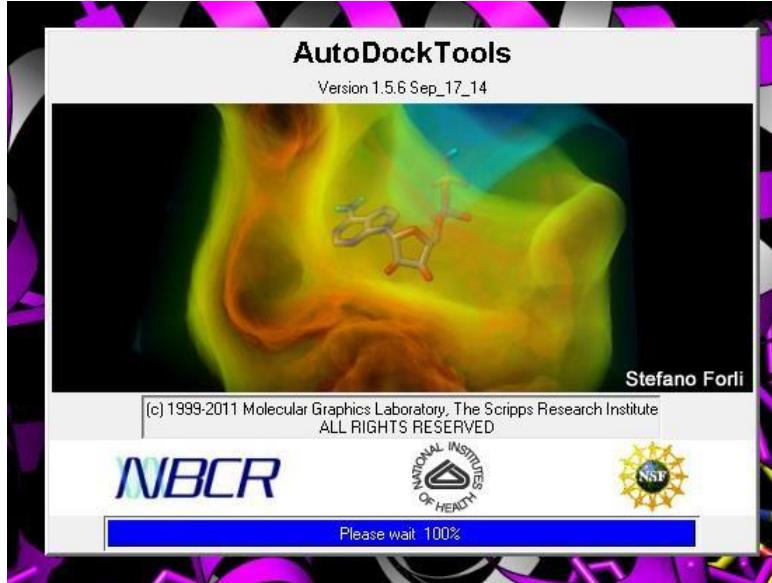
platform	installer
	mgltools_win32_1.5.6_Setup.exe (71Mb)
	mgltools_i86Darwin9_1.5.6.dmg (GUI installer 65Mb)
	mgltools_i86Darwin9_1.5.6.tar.gz (tarball installer 63Mb)
	mgltools_Linux-x86_64_1.5.6_Install (Linux 64 GUI installer 70Mb)
	mgltools_x86_64Linux2_1.5.6.tar.gz (Linux 64 tarball installer 68Mb)

---

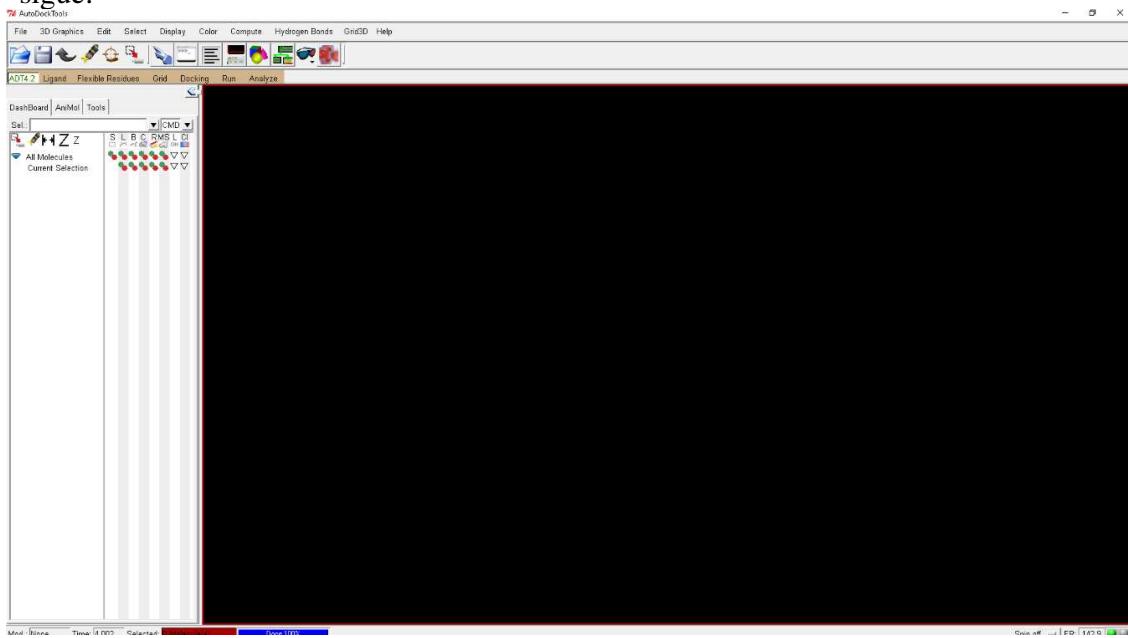
Una vez descargado, hacemos click sobre el ejecutable



Después de aceptar los términos y elegir la ruta de instalación, debemos iniciar el programa (AutoDockTools-1.5.6). Primero se observará una pequeña ventana que tiene una barra que se va llenando hasta alcanzar el 100%. Muchas veces cuando el programa no es compatible con la computadora, la barra suele quedarse pasmada en un porcentaje menor al 70%.



Si no pasó eso, entonces después de completar el 100%, el programa deberá verse como sigue:



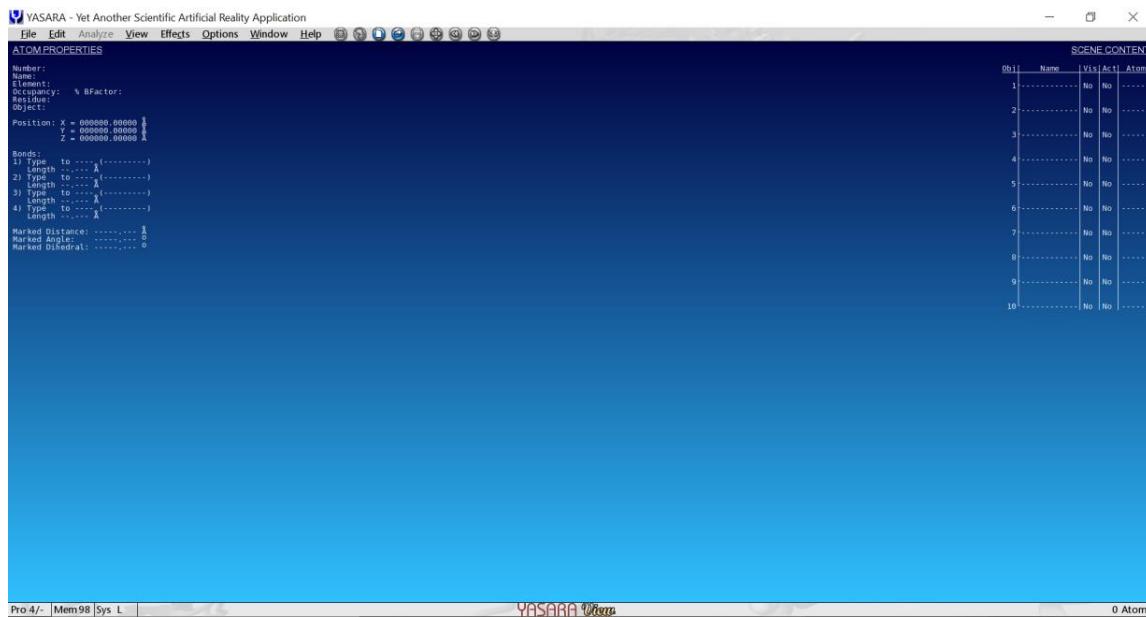
## 20.3 YASARA View

Ir a la página web <http://www.yasara.org/viewdl.htm>

Una vez en la página, llenar los datos que se nos piden y enviar



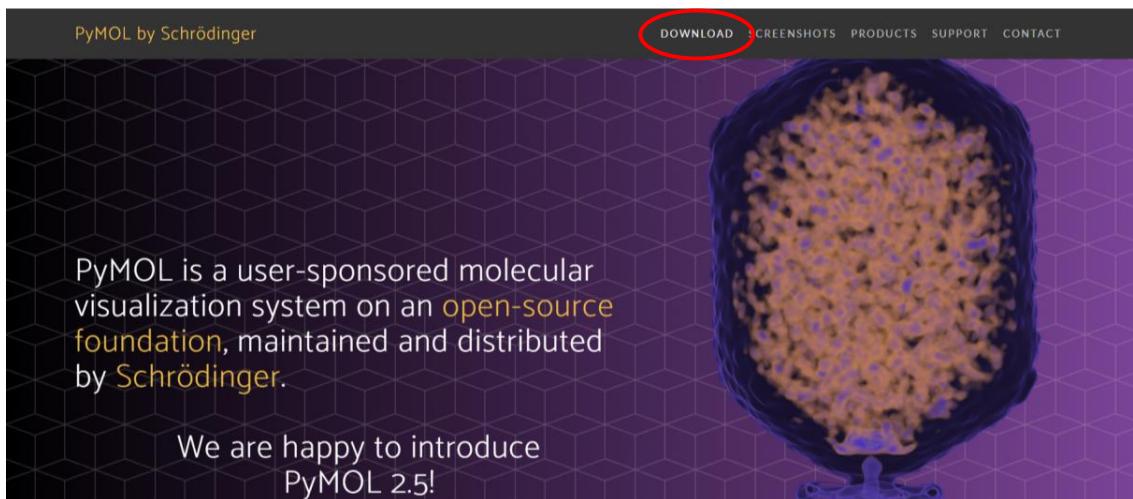
Al correo que hayamos puesto, nos llegará una liga de descarga para el programa. Una vez que lo descarguemos, hacemos click sobre el ejecutable y seguimos los pasos de instalación. Una vez completados esos pasos, el programa debería de verse así:



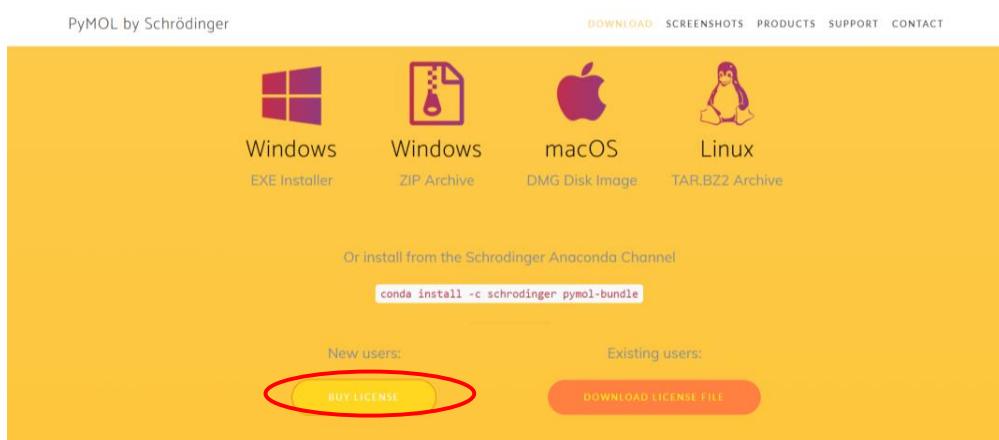
## 20.4 PyMOL

Ir a la página web <https://pymol.org/2/>

Una vez en la página, hacer click sobre “Download”



Posteriormente, hacer click sobre “Buy License”



Seleccionar la opción de “Student/Teacher”

## Order PyMOL and AxPyMOL Online

Please select your intended use from the list below to receive access within one business day:

- **Commercial**

for corporate organizations

- **Government**

for government institutions

- **Academic/Non-profit**

for academic and non-profit researchers/educators

- **Educational**

Educational subscriptions to PyMOL are available at no cost to full-time students and educators teaching full-time students. However, there is not an educational subscription for AxPyMOL at this time. Under this agreement, Schrödinger has no obligation to furnish updates, fix bugs, provide documentation, and meet other support needs. Please recognize that long-term reliance upon PyMOL for educational purposes should merit a subscription purchase in order to secure essential services and baseline

Llenamos los datos que se nos piden

### Registration For Educational-Use-Only PyMOL Builds

Schrödinger offers **Educational-use-only** PyMOL builds available at no cost to **teachers and high school and college students** (including online courses, homeschooling, etc.) for classroom instruction, homework assignments, and to provide a means for creating high quality figures. Please note that it is not provided for the purposes of academic research or publication.

-> [FAQ \(Frequently Asked Questions\)](#)

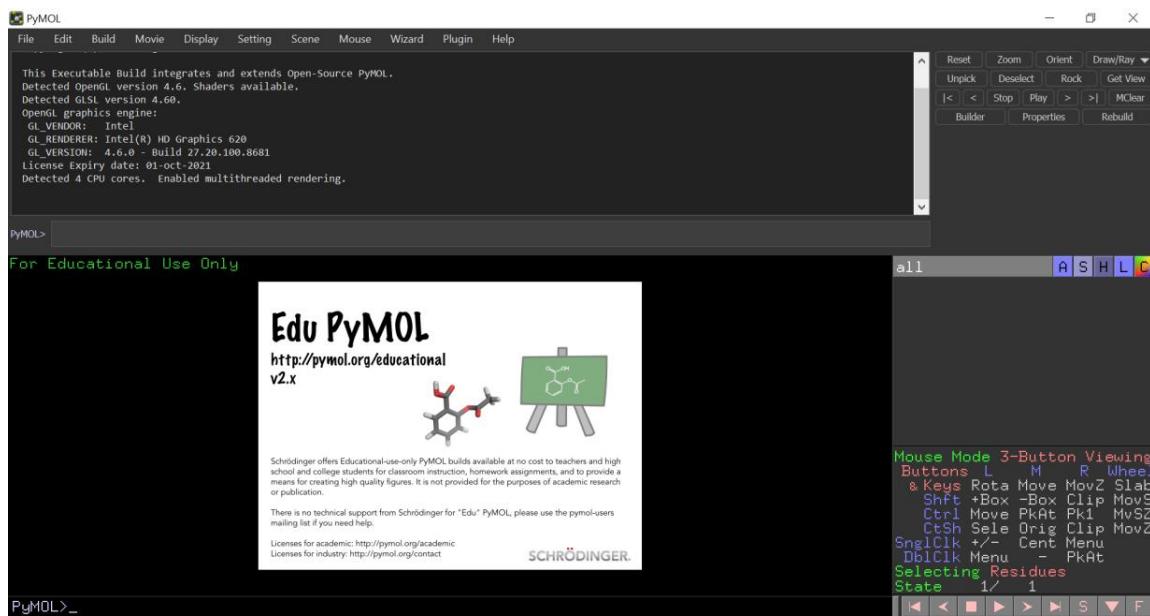
The Educational-use-only PyMOL builds are provided "AS IS" with no obligation to grant download access, fix bugs, furnish updates, provide documentation, or meet any other need related to the educational-use PyMOL builds.

If you intend to use PyMOL products for academic research or publication, please purchase an Academic PyMOL subscription, which includes access to technical support, screencasts, and additional resources. See <http://pymol.org/academic>.

I am a:	<input type="button" value="▼"/>
Your First Name:	<input type="text"/>
Your Last Name:	<input type="text"/>
Your Email Address:	<input type="text"/>
Your Telephone Number:	<input type="text"/>
Institution:	<input type="text"/>
Comments (optional):	<input type="text"/>

[Continue](#)

Al correo que hayamos puesto, nos llegará una liga de descarga para el programa. Una vez que lo descarguemos, hacemos click sobre el ejecutable y seguimos los pasos de instalación. Una vez completados esos pasos, el programa debería de verse así:



## 20.5 BIOVIA Discovery studio visualizer

Ir a la página web <https://discover.3ds.com/discovery-studio-visualizer-download>

Una vez en la página, escribir un correo electrónico en donde se indica

Llenar los datos que se nos solicitan



Una vez que llenemos nuestros datos y los enviamos, nos aparecerá la siguiente pantalla, donde podremos descargar el programa

**Thank you for your interest!**

You can download the most recent version by clicking the link below.

If you experience any issues downloading the files please try switching to another web browser program.

Please note that DS Visualizer is a free resource that is not supported by the BIOVIA Support Team. Instead, DS Visualizer is supported solely by the Discovery Studio user forum on the [BIOVIA Community](#).

[Link to End User License Agreement \(EULA\)](#)

Should you have any questions about purchasing Discovery Studio, please contact a BIOVIA representative at [BIOVIA.Marketing@3ds.com](mailto:BIOVIA.Marketing@3ds.com).

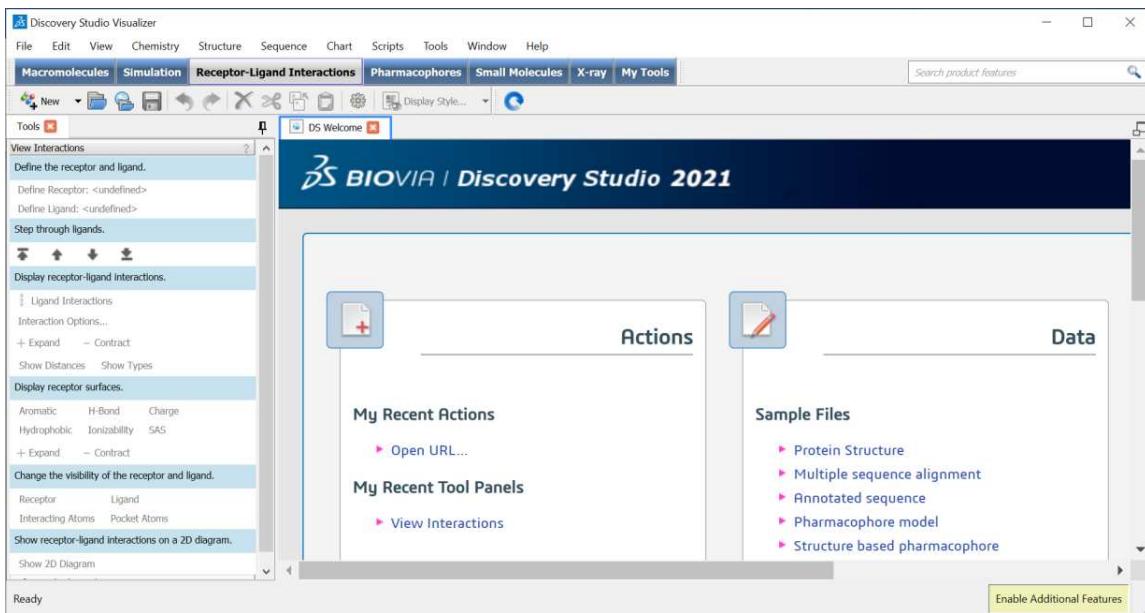
DS Visualizer Client (Windows 64 bit)      DS Visualizer Client (Linux 64 bit)

[Download](#)      [Download](#)

El siguiente paso es hacer click sobre el ejecutable



Una vez que se eligió la ruta de instalación y se instaló el programa, este debe iniciarse haciendo. El programa debe verse así:



## 20.6 Ejecutables de Autodock 4.2

Para descargar los ejecutables de AD4.2 debemos ir a la página web <http://autodock.scripps.edu/downloads/autodock-registration/autodock-4-2-download-page/>

Una vez en la página hacemos click para descargar desde el sistema operativo que tenemos

Select the platform and/or source code.	
Release 4.2.6 Notes.	
AutoDock4.2.6 features improved input checking and an output format suitable for automated analysis. Multiple search methods can be used in a single AutoDock4.2.6 job. AutoDock 4.2.6 is available for more platforms. The process of compiling new atomic parameter tables into AutoDock and AutoGrid is documented in the README file.	
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	<ul style="list-style-type: none"> <li>■ AutoDock 4.2.6 User Guide (PDF, 2.2 MB)</li> <li>■ Examples (11MB) md5sum 44148341b0b797c894a19b45bbb4210</li>   <li>■ Linux: Intel (32-bit) (667K) md5sum e3b18a7f399525c6edbea4b05f26e850           <ul style="list-style-type: none"> <li>■ Linux: Intel (64-bit) based on command 'uname -r' output:               <ul style="list-style-type: none"> <li>■ 2 - Linux: Intel (64-bit) (743K) md5sum 8c175d4f7b9b1529fdf8d3abf9c90772</li> <li>■ 3 - Linux: Intel (64-bit) (764K) md5sum 0f500576d03abd97cbe543af6e99dd2</li> </ul> </li> </ul> </li>   <li>■ Mac OS X including 10.5 (Leopard), 10.6 (Snow Leopard), 10.7 (Lion), 10.8 (Mountain Lion), 10.9 (Mavericks) that works on both 32-bit and 64-bit Power PC (PPC) and Intel (2.3MB) md5sum ce1333e17b53d1c7bd9734adfb1d7ca0</li>   <li><b>■ Windows (577K) md5sum adcd08691ec206060ba7e84d32a1cf8</b></li>   <li>■ Sun Solaris (Sparc) (808K) md5sum b4662f5023ee68e4da4e6afabe9af7fe</li>   <li>■ Source code (35MB) md5sum f4942c8e8c47aca7f3a2ae8794259067</li> </ul>

Una vez que se descargue, hacemos click sobre el archivo e instalamos. Debemos seleccionar el sitio en donde se instalará. Cuando termine, iremos al sitio que indicamos anteriormente. En esta ubicación podremos encontrar una carpeta que contiene los archivos ejecutables correspondientes.

: equipo > Disco local (C:) > Archivos de programa (x86) > The Scripps Research Institute > Autodock > 4.2.6

Nombre	Fecha de modificación	Tipo	Tamaño
autodock4	18/07/2014 05:53 p. m.	Aplicación	605 KB
autogrid4	18/07/2014 05:53 p. m.	Aplicación	1,883 KB

## 20.7 VINA

Ir a la página <http://vina.scripps.edu/>

Una vez en la página, hacer click sobre “Download”

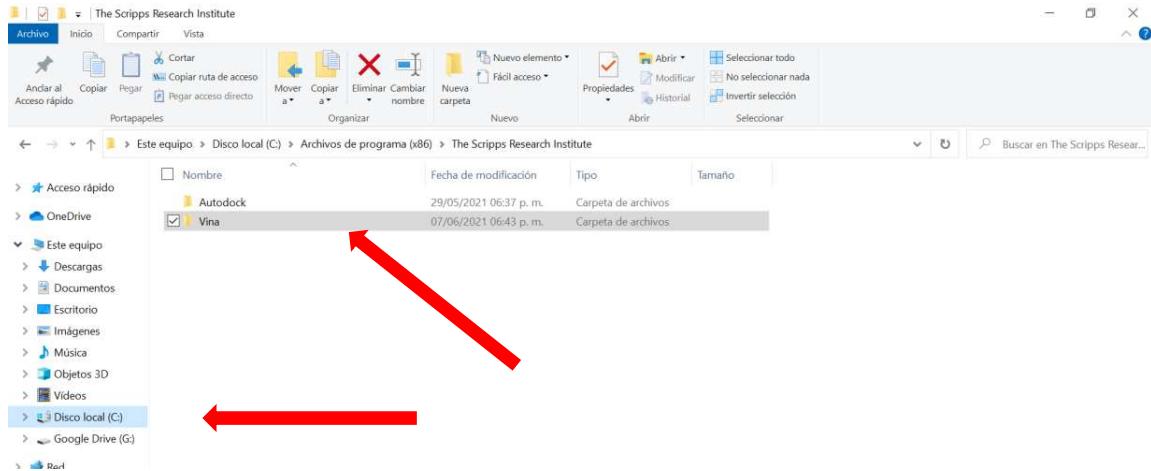
The screenshot shows the AutoDock Vina website. At the top, there is a navigation bar with links: Home, Features, Download (which is circled in red), Tutorial, FAQ, Manual, and Questions?. Below the navigation bar, there are four molecular docking results labeled 'a', 'b', 'c', and 'd'. To the right of these images, the text 'AutoDock Vina' is displayed, followed by a brief description of the program. At the bottom left, there is a link to 'vina.scripps.edu/download.html'.

Posteriormente, seleccionamos la opción para nuestro sistema operativo, Windows

The screenshot shows the 'Download' section of the AutoDock Vina website for the Windows operating system. The 'Windows' option is highlighted with a red circle. Below it, there are links for Linux, Mac OSX (64 bit, beta version), Mac OSX (32 bit), and Source. Each link includes the file name, size, and a compatibility note. At the bottom of the section, there is a note about seeing also 'GUIs, web interfaces, etc.' and links for 'Old versions' and 'History of changes'. There is also a 'Questions, problems or suggestions?' section and a link to 'See this site.'

Hacemos click sobre el archivo que descargamos y seguimos los pasos del asistente de instalación.

Una vez instalado, veremos que en el sitio donde se instaló se encuentra ahora una carpeta llamada Vina, como se ve en la imagen. Esta carpeta se debe cortar y pegar directamente en la unidad C de nuestra computadora.



## 20.8 GOLD

Para instalar este programa, debemos ir a la sección de Laboratorios Virtuales de BIDIUAM, en Cambridge Structural Database Enterprise

Cambridge Structural Database Enterprise

Anatomía 3D

Virtual Plant

Jove

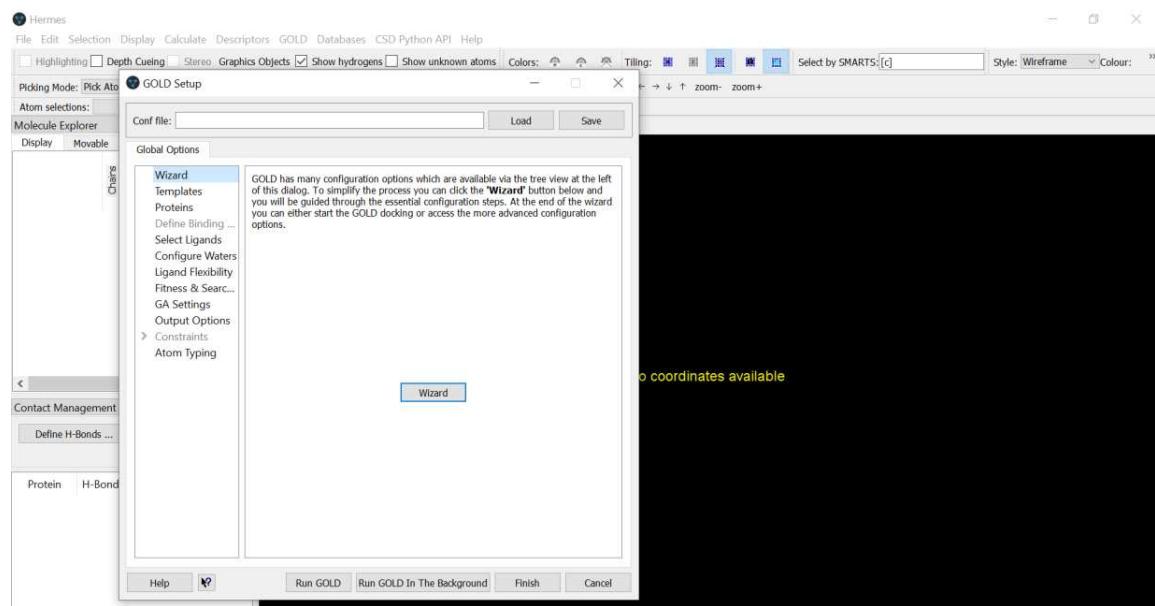
Cambridge Structural Database Enterprise

CSD-Enterprise es una paquetería de software especializado comprende los programas DASH, MERCURY, GOLD, SUPERSTAR y CSD- CROSSMINER, los cuales están enfocados en entender cómo se encuentran distribuidos los átomos en una molécula sencilla, entender de manera gráfica las interacciones en los sistemas biológicos entre macromoléculas, así

Una vez que lo hicimos, descargamos e instalamos todos los programas, podemos abrir GOLD haciendo doble click sobre su ícono



Una vez abierto, el programa debería de verse así



## **21 Anexo 2. Comandos y sitios web de interés**

### ***21.1 Símbolo del sistema***

cd INTRO → entrar a una carpeta

cd .. INTRO → salir de una carpeta

DIR INTRO → ver qué hay dentro de una carpeta

Tab → completar con texto escrito anteriormente

### ***21.2 AD4.2***

Comando para autogrid

`autogrid4.exe^-p^grid.gpf^-l^grid.glg`

Comando para autodock

`autodock4.exe^-p^dock.dpf^-l^dock.dlg`

### ***21.3 VINA***

`"C:\vina\vina.exe"^-config^conf.txt^-log^log.txt^-out^name_dock.pdbqt`

### ***21.4 Páginas web***

- RCSB PDB <https://www.rcsb.org/search>
- PubChem <https://pubchem.ncbi.nlm.nih.gov/>
- ZINC Database <https://zinc.docking.org/substances/home/>
- YASARA Minimization Server <http://www.yasara.org/minimizationserver.htm>
- LigRMSD <https://ligrmsd.appsbio.utalca.cl/>