



FACULTAD DE FARMACIA
UNIVERSIDAD COMPLUTENSE

TRABAJO FIN DE GRADO
DESCUBRIMIENTO DE NUEVOS FÁRMACOS
USANDO EL CRIBADO VIRTUAL DE
QUIMIOTECAS

Autor: Ismael Lastra Martínez

Fecha: Mayo 2020

Tutor: Almudena Perona Requena

Índice

1	Resumen.....	2
2	Introducción.....	3
3	Materiales y métodos	4
3.1	Docking.....	6
3.1.1	La función de puntuación o <i>scoring</i>	7
3.1.2	El componente estructural y las quimiotecas	8
3.1.3	El proceso de docking	10
3.2	Cribado Virtual	12
3.3	Superordenadores	14
4	Resultados y discusión	14
4.1	Reposicionamiento de fármacos	16
4.2	El éxito en los procesos basados en cribado virtual.....	17
5	Conclusión	18
6	Bibliografía.....	19

1 Resumen

Descubrir y desarrollar un fármaco es algo muy costoso en lo cual debemos invertir mucho tiempo y dinero, donde las moléculas que estudiamos no siempre resultan aptas o eficaces terapéuticamente. Una de las herramientas para solventar estos problemas son las técnicas computacionales basadas en cribados o filtrados virtuales, haciendo del proceso de descubrimiento algo más rápido y económico. Estas técnicas se basan en programas que recrean acoplamientos moleculares según las moléculas que se encuentren en las bases de datos independientes a estos. Así, el estudio *in silico* nos dará una aproximación sobre si dos moléculas pueden formar un complejo interesante para el estudio o no, puntuando dicha unión según una función de puntuación o *scoring*. Con el paso de los años estas técnicas teóricas han ido mejorando de forma muy notoria, aumentando así su utilidad, valor y fiabilidad.

Sin embargo debemos mirar hacia el futuro y seguir desarrollando y mejorando estas técnicas, ya que todavía quedan retos que afrontar, estimulando la innovación de los métodos que se emplean hoy en día en el descubrimiento y desarrollo de fármacos.

Palabras clave: Innovación, desarrollo, técnicas computacionales, *docking*, cribado virtual, quimioteca.

Abstract

Discovering and developing a drug is a very expensive thing that we should invest in a lot of time and money, where the molecules we study, would not always be suitable or therapeutically effective. One of the tools to solve that kind of problems are developing computational techniques based on virtual screening, the making of the discovery process to be faster and cheaper. These techniques are based on programs that recreate molecular docking, according to the molecules in which are found in independent databases. Thus, the in silico study will give us an approximation on whether or not the two molecules can form into an interesting complex for the studies, scoring that union through a scoring function. Over the years, these techniques have been improving in a clear notorious way, increasing his utility, his value and his reliability.

However we should look into the future and keep developing and improving these techniques, because we there are some challenges to face, stimulating the innovation of methods which nowadays are being used in developing and discovering of drugs.

Key words: Innovation, developing, computational techniques, docking, virtual screening, molecular database.

2 Introducción

Hoy en día, con el avance de las tecnologías, estamos desarrollando nuevos métodos para lograr ganar eficacia terapéutica y abaratar el descubrimiento del fármaco (1). El proceso de desarrollo de fármacos es largo y costoso, y la tasa de deserción de la nueva entidad química (NCE) suele ser muy alta, generalmente en las etapas posteriores del proceso de desarrollo del fármaco, es decir, durante varias fases de ensayos clínicos (2). Desde que se comienza a investigar una molécula hasta que ese medicamento llega al mercado pasan entre 10 y 15 años actualmente (24).

De ese modo, para ahorrar tiempo y dinero en descubrimiento y desarrollo de estos, se están empleando cada vez con más frecuencia técnicas de acoplamiento inducido (*docking*) basadas en cribados virtuales, con los cuales podremos predecir de forma teórica y usando técnicas computacionales, como dos moléculas se unirán y si llegarán a formar un complejo más estable. Gracias a las técnicas computacionales podemos probar millones de moléculas procedentes de bases de datos con las dianas terapéuticas de interés en muy poco tiempo. Gracias a estos estudios de acoplamiento, miles de moléculas pueden ser evaluadas como eficaces y seguras a un coste mucho más bajo que con métodos de descubrimiento tradicionales, en intervalos de tiempo mucho menores (2).

Sin embargo estas técnicas computacionales aún están por debajo de las expectativas iniciales creadas en torno a ellas, ya que el esfuerzo y dinero invertido aún no concuerdan con el escaso número de fármacos que llegan al mercado. Pero estos resultados también han repercutido de forma positiva en otros desarrollos de técnicas computacionales, tratando de eliminar fallos, para intentar racionalizar y, en base al conocimiento, acelerar y potenciar las etapas iniciales en el diseño de fármacos (1).

Otra preocupación sobre estas nuevas técnicas surge al extrapolar una molécula *lead* a humanos, ya que previamente se suele testar en animales. Según la *Administración de Drogas y Alimentos de los Estados Unidos* (FDA), más del 90% de las nuevas entidades químicas prometedoras fallan al ser probadas en humanos. Con el fin de evitar fallos tardíos en el desarrollo de una NCE para su uso válido como medicamento, los estudios de *docking* se llevan a cabo en procesos muy tempranos del desarrollo de fármacos (2). La mayoría de los compuestos que demuestran actividades *in vitro* con las dianas moleculares fallan en las

siguientes pruebas, debido de forma común a pobres actividades relacionadas con farmacocinética y toxicidad. Se estima que de cada 9000 moléculas biológicamente activas, solo una tendrá un uso clínico (5).

Otra técnica computacional que gana importancia con el paso de los años es el modelado molecular, el cual debemos llevar a cabo para obtener la estructura tridimensional de una proteína. Debemos diferenciar el proceso de modelado molecular del de *docking* (representa la unión entre moléculas) o del de dinámica molecular (evalúa el comportamiento de estas cuando están unidas) (26).

Por último podemos añadir que el descubrimiento y desarrollo de fármacos se ha producido durante muchos años con una única base de métodos experimentales, y se espera que este proceso se acelere con el uso de métodos *in silico* (hecho por computadora o vía computacional). Además, esto hace que se entiendan mecanismos de acción o se mejoren las propiedades de este. Un claro ejemplo es la optimización de los fármacos de origen natural, ya que hoy en día se están mejorando propiedades de estos medicamentos como reducir efectos secundarios o ampliar su potencia (3).

Algunos ejemplos de casos exitosos de fármacos descubiertos a través del cribado virtual son algunos inhibidores de la anhidrasa carbónica en el tratamiento del glaucoma, inhibidores de proteínas del SARS o de la proteasa del VIH o por ejemplo, fármacos para el tratamiento de la malaria entre otros (6).

3 Materiales y métodos

La investigación farmacéutica cada vez cuenta con más métodos computacionales, donde destacan una gran variedad de programas informáticos y técnicas teóricas, para estudiar sistemas biológicos y químicos complejos.

El desarrollo de fármacos ha progresado mucho gracias a los avances en métodos electroscópicos biomoleculares como la cristalografía de rayos X o la resonancia magnética nuclear (RMN), que han creado diversos avances en el conocimiento de la biología molecular y estructural. Más concretamente, estas técnicas han permitido la resolución de más de 100.000 estructuras de proteínas tridimensionales (3).

Los esfuerzos por almacenar y organizar la información molecular han generado una demanda muy alta de herramientas informáticas, por lo que los métodos experimentales o *in silico* se han asentado fuertemente en las últimas décadas gracias al uso de quimiotecas, para seleccionar así y lanzar al mercado aquellos candidatos que tienen una mayor probabilidad de reaccionar eficazmente con la diana y provocar el efecto terapéutico deseado (3).

Gracias a todos estos avances, los métodos de diseño de fármacos basados en la estructura de la proteína (SBDD) se han convertido en un componente muy importante para la química moderna. El acoplamiento molecular o *docking*, el cribado virtual basado en estructuras o la dinámica molecular, que más adelante desarrollaremos, se encuentran entre las estrategias SBDD más utilizadas al aportarnos información clave sobre estructuras moleculares y sus posibles interacciones (3).

Diseño de fármacos basado en estructuras (SBDD)

El SBDD hace referencia al modelado de fármacos sirviéndonos de una estructura proteica ya conocida, que obtenemos experimentalmente o mediante modelos de homología computacional. El objetivo así es de obtener fármacos con atributos electrostáticos y estereoquímicos específicos para lograr una alta afinidad al unirse al receptor, y que suponga dicha unión un efecto terapéutico deseado.

Por otro lado existe el LBDD (*Ligand base drug design*) el cual se basa en la estructura del ligando, por eso, comprender como estos reconocen estructuras y se unen a ellas es de gran importancia en investigación y desarrollo farmacéutico (2).

Una vez determinado el complejo ligando-diana, el proceso SBDD se reinicia, comenzando de nuevo para incorporar modificaciones moleculares y que estas aumenten la afinidad del ligando por el receptor (3).

Gracias al SBDD nos encontraremos ante procesos los cuales requerirán un tiempo y coste menor, al reducirnos el número de posibles ligandos eficaces (que debemos probar experimentalmente) para una diana concreta (3).

3.1 Docking

El acoplamiento molecular o *docking* es uno de los métodos más utilizados en SBDD debido a su capacidad para predecir como encajarían dos moléculas entre sí, es decir, consiste en aplicar modelos basados en algoritmos matemáticos fundamentados en las leyes de la física, donde sometemos a la proteína y al ligando a un campo de fuerzas y así se valora mediante una función de puntuación como de eficaz será esa unión. El acoplamiento molecular se ha convertido en una herramienta esencial en el descubrimiento de fármacos, desde que se produjera el desarrollo de los primeros algoritmos allá por la década de 1980 (2).

En primer lugar, debemos dejar claro que se asume un conocimiento previo de la estructura 3D, tanto de la diana como del ligando. A las moléculas que muestren algún tipo de actividad biológica las denominaremos *hits*. A posteriori, debemos encontrar compuestos que tengan propiedades farmacéuticas atractivas, con baja toxicidad, y otras propiedades farmacocinéticas de interés. A estas moléculas las llamaremos *leads*. Para llegar hasta estos, deberemos desarrollar los *hits* a través de modificaciones químicas. Este análisis sería muy complejo sin las técnicas computacionales, ya que se estima que el número de moléculas sintéticamente factible a desarrollar se encuentra entre 10^{20} y 10^{24} (5).

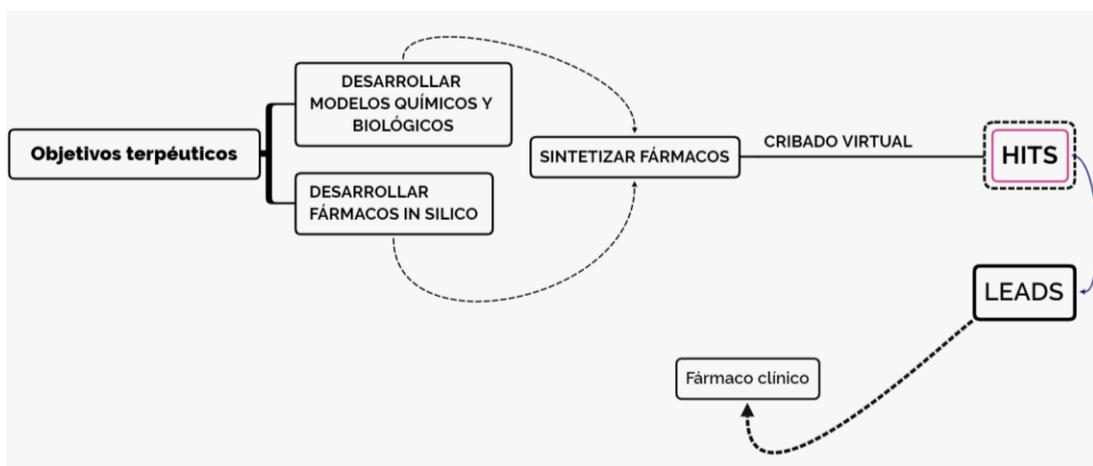


Figura 1. Representación esquemática del procedimiento que se da en una molécula hasta convertirse un fármaco aprobado clínicamente.

Una complicación del *docking* recae en el conocimiento de las estructuras. El ligando puede adoptar muchas posiciones dentro del sitio de unión de la diana (1), siendo la flexibilidad del

sitio de unión un aspecto clave en el proceso de desarrollar un fármaco eficaz. Las enzimas y los receptores pueden sufrir distintos tipos de cambios conformacionales durante el proceso de reconocimiento molecular (3). Aunque actualmente estas complicaciones ya están superadas, es cierto que pueden aumentar el tiempo en obtener un resultado válido, dificultando así el estudio.

3.1.1 La función de puntuación o *scoring*

Consiste en una ecuación matemática que nos da un valor único con el que cuantificaremos la fortaleza o eficacia con la que se una a el receptor un ligando cualquiera, donde se evaluará la unión sumando el número de interacciones favorables, tales como puentes de hidrógeno, interacciones del tipo carga-carga o las interacciones hidrofóbicas. Así que una vez que tengamos todas las posibles configuraciones del complejo ligando-diana (*poses*), esta función nos dirá basándose en dicho componente energético, como de buena es la unión en cada pose. Gracias a esta función conseguimos optimizar el proceso de descubrimiento de fármacos, ya que nos ayuda claramente a identificar ligandos de alta afinidad (1).

Esto es algo difícil de calcular de forma rápida y con precisión, por lo que aunque es muy útil para proponer modelos de unión, aún no es muy certero para calcular con precisión la energía con la que interaccionan ambos compuestos, debido a todas las aproximaciones que se hacen para calcular la energía de forma rápida, ya que por ejemplo no suelen considerarse detalles concretos como la entropía. Esto se solucionaría exponiendo ese modelo a cálculos más refinados con niveles más elevados de teoría, lo que supondría un tiempo de estudio mayor (5).

Conocemos 3 tipos de funciones de *scoring* diferenciadas entre sí por los datos usados en su derivación: empíricos (obtenidos a partir de análisis de regresión multilínea entre medidas experimentales de actividad y diversas propiedades), funciones de *scoring* basadas en el conocimiento (haciendo uso de bases de datos de estructuras 3D que más adelante comentaremos, donde buscamos cuales son las interacciones más frecuentes entre ligandos y dianas) y funciones de *scoring* basadas en campos de fuerza (donde se usan parámetros de mecánica molecular para descomponer las distintas energías de unión) (1,3).

Cada función de puntuación tiene sus virtudes y limitaciones, por lo que cada vez se da más el uso simultáneo de diferentes metodologías de puntuación. Esto es muy útil ya que combina

las ventajas y también atenúa las deficiencias de cada método. Algunos ejemplos destacables de funciones de puntuación de consenso son *MultiScore*, *X-Cscore* ó *SeleX-CS* (3).

3.1.2 El componente estructural y las quimiotecas

Es imprescindible conocer como es la estructura 3D para poder usarla en el proceso de *docking*. La estructura de la diana podemos obtenerla de dos formas (2):

- Métodos experimentales: De gran utilidad puesto que sirven para proteínas y ligandos. La primera que destaca es la cristalografía de rayos- X, la cual es una técnica de microscopía de muy alta resolución con la que podemos ver estructuras a niveles atómicos. Esto se da gracias a los rayos X, con los que podemos visualizar objetos que a simple vista no vemos, recurriendo a la utilización de fuentes de radiación electromagnética de una longitud de onda característica. A parte se necesitan cristales para la obtención de una estructura más detallada (7). Otra técnica que destaca es la espectroscopía de Resonancia Magnética Nuclear o *RMN*, donde se utiliza un campo magnético potente y ondas de radio de alta frecuencia para producir imágenes muy detalladas (8)
- Métodos teóricos: Se basan en modelados por homología, ya que aunque una proteína sea desconocida, si comparte cierto porcentaje de identidad con una proteína conocida, podremos obtener un modelo tridimensional razonable. Así podremos ver si dos secuencias son homólogas o no dependiendo del grado de similitud y la significancia estadística del alineamiento (9).

Bases de datos de estructuras de ligandos

Por otro lado, la estructura experimental de muchos ligandos se encuentra a través de la base de datos de *Cambridge Structural Database* (CSD) (o puede construirse también con programas de modelado molecular que más adelante desarrollaremos).

Esta se estableció en 1965, y es uno de los mayores almacenes virtuales en el mundo de pequeñas moléculas orgánicas y de estructuras cristalinas metalo-orgánicas. Contiene más de un millón de estructuras analizadas a partir de Rayos-X y análisis de difracción de neutrones y hoy en día es una base de datos de estructuras 3D muy precisa, siendo así una herramienta esencial para muchos científicos por todo el mundo.

El CSD es continuamente actualizado con nuevas estructuras (> 50,000 estructuras cada año) y gracias a actualizaciones regulares, cualquiera puede acceder a las nuevas estructuras publicadas (10).

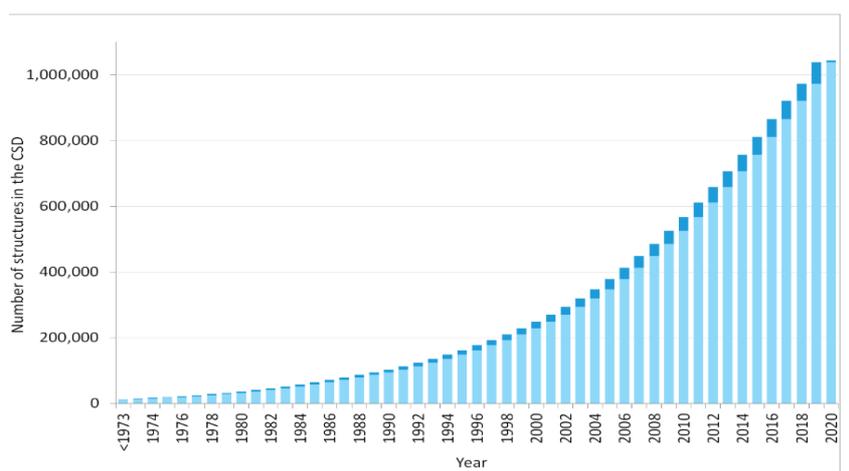


Figura 2. Número de estructuras registradas en el CSD por año desde 1973, en azul oscuro las añadidas durante ese año.

Finalmente podemos confirmar que la quimio-informática tiene aplicaciones especialmente valiosas en análisis y manejo de bases de datos moleculares. Algunos otros ejemplos de colecciones de ligandos usadas con frecuencia en el descubrimiento de fármacos son ZINC, DrugBank, ChEMBL, Binding Database y PubChem. Estas bases de datos contienen información para realizar cribado virtual, quimiogenómica, análisis estructura actividad y reposicionamientos de fármacos, entre otros (5).

Bases de datos de estructuras de proteínas

En el campo de las proteínas destaca una gran base de datos llamada *Protein Data Bank* (PDB) en la cual existen multitud de complejos ligando-receptor ya registrados (así quedan almacenadas las estructuras tridimensionales) y donde podemos encontrar de forma fácil la estructura de algunas dianas (1).

En la fase de descubrimiento de fármacos, esta base de datos es muy usada por los investigadores. La estructura de la proteína se puede seleccionar en función de la resolución de rayos X y las condiciones bajo las cuales se obtuvo la estructura cristalina de la proteína (11).

El PDB contiene en torno a 100,000 registros, la mayoría proteínas, pero también se registran ácidos nucleicos y complejos de ambos. Este archivo es de uso público sin ningún coste para los usuarios, y es actualizado semanalmente. Fue establecido en 1971 y contenía originalmente 7 estructuras. A esta base de datos virtual accede más de un millón de usuarios

al año y según un análisis económico establecido en 2016, se le otorgó un valor aproximado de 12 billones de dólares americanos (11).

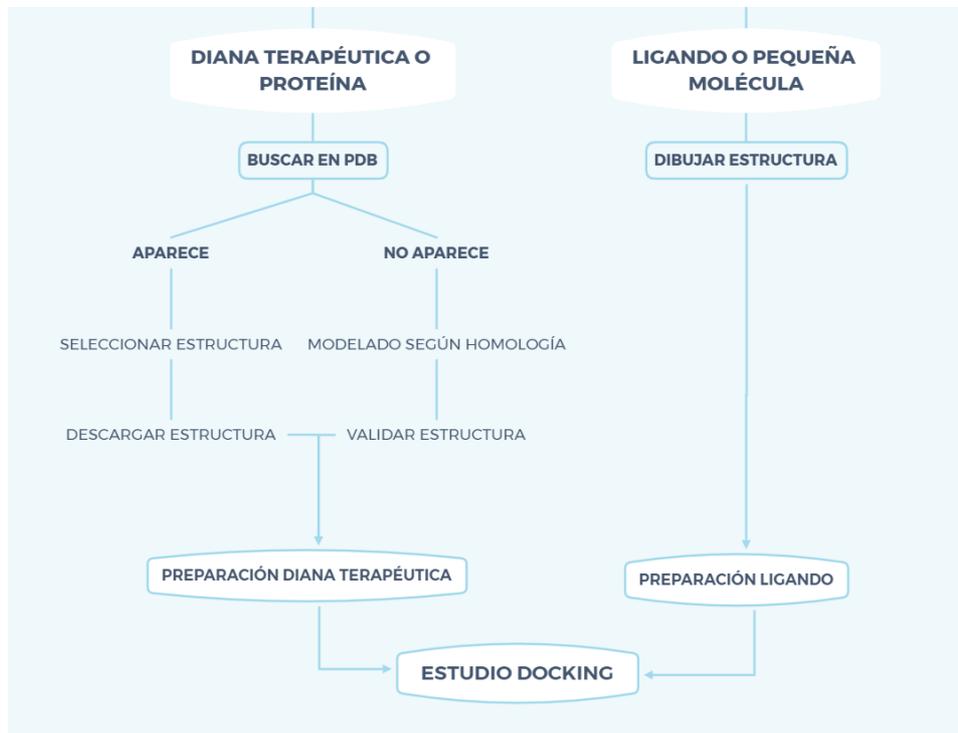


Figura 3. Representación estructural de la búsqueda de dos moléculas aptas para desarrollar un estudio docking del complejo que formen.

3.1.3 El proceso de docking

En primer lugar debemos tener en cuenta la flexibilidad de las estructuras ya que estas no son estáticas, aun así, si conocemos la zona de unión en la superficie de ambas moléculas, tendremos mayores probabilidades de éxito.

El protocolo a seguir es muy simple: Primero debemos someter a la proteína y al ligando a un campo de fuerzas, después medimos con la función de scoring la interacción entre ambas y por último las ordenamos dependiendo del valor de la función de scoring (1,2).

Existen diferentes softwares o programas para llevar a cabo estos procesos, dependiendo de las estructuras frente a las que nos encontremos. En el docking rígido (ambas moléculas no tienen grados de libertad internos) no hay demasiada complejidad y se usa para desarrollar programas de *docking* nuevos, a pesar de pecar de simplicidad (1).

Después, podemos hablar de *docking* donde el receptor es rígido pero el ligando adopta distintas formas (asumimos que la aproximación del modelo llave-cerradura es aceptable), donde aquí el principal objetivo será buscar el ligando (la llave) para la diana rígida. Para lograrlo se han ido desarrollando métodos o algoritmos implementados en procesos de acoplamiento molecular. Algunos, como FlexX, FRED o Surflex, se basan en dividir el ligando en fragmentos e ir analizando estos conformacionalmente, para predecir así interacciones proteína-ligando (1,2). Este método basado en ir cogiendo fragmentos conocidos del ligando y acoplarlos en una región complementaria al sitio de unión secuencialmente, hace que el algoritmo solo realice la búsqueda conformacional para los fragmentos que se van agregando, disminuyendo así el grado de libertad que podría tener la molécula respecto a otra que no fuera dividida previamente. El programa FRED examina todas las posibles posturas dentro del sitio activo, y las clasifica por complementariedad de forma y caracteres farmacóforos (3). Durante la búsqueda, son las 100 mejores poses (es decir, las que mayor puntuación obtengan) las que se someterán a una optimización sistemática. La mejor pose será la que se use para clasificar al ligando frente a otros ligandos en la base de datos que se emplee (22).

También existen otros programas como CRDOCK y GLIDE que generan las conformaciones del ligando antes de la operación del *docking*, para guardarla en sus bases de datos y usarla posteriormente (1,4). También se pueden generar conformaciones "in situ", como mediante el programa Dock (1) donde se evalúa geoméricamente la concordancia entre ligando y diana o los programas Gold y Autodock (1,2,3) donde mediante algoritmos genéticos (Basados en conceptos sobre evolución o teorías como la de evolución de Darwin o la de herencia de Lamarck) podemos analizar parámetros estructurales sobre la estructura inicial de un cromosoma y a partir de ese, conseguir generar poblaciones de cromosomas con un espacio conformacional menor a explotar (3).

Existen otros programas, como por ejemplo LigandFit (16) basados en el algoritmo de Montecarlo, el cual genera de forma completamente aleatoria grupos de rotación, translación y orientación de ligandos para luego evaluarlos mediante *scoring* (1). Este se combina con LigScore, una función de puntuación que produce tasas de aciertos amplias para un grupo de ligandos conocidos.

Por último hablaremos de *docking* flexible, donde se usan unos métodos u otros dependiendo del grado de flexibilidad de la diana. Aquí la efectividad de los softwares es menor y por eso tiene más valor, ya que las dos moléculas que se unirán pueden cambiar su conformación, dotando así de complejidad al proceso. La flexibilidad se puede abordar desde varios niveles: permitiendo cierto grado de interpenetración entre proteínas para que haya cambios conformacionales implícitamente (*soft docking*) o realizando múltiples recorridos de acoplamiento desde varias conformaciones (*cross docking*) (17,18).

También en el *docking* flexible se usa el *algoritmo de Montecarlo* ya mencionado anteriormente. Aquí lo usan programas como Gold, Glide, Autodock, FlexX o ICM, permitiéndonos realizar el proceso de *docking* usando diferentes conformaciones para la diana. Esto es muy útil para probar cambios en cadenas laterales del receptor.

El agua cristalográfica es un desafío importante en estos procesos de *docking*, ya que esta se encuentra unida fuertemente al receptor. Simulaciones basadas en el *algoritmo de Montecarlo* consiguen estimar la energía libre de unión para la molécula de agua dada, lo que nos permite poder discriminar posteriormente entre agua estructural desplazable y agua fuertemente unida (2).

Otro algoritmo importante en el proceso de descubrimiento de nuevos fármacos es el algoritmo ANCHOR, el cual se encarga de buscar cadenas laterales de aminoácidos en uniones del tipo proteína-proteína. Esto sirve para identificar posibles bolsillos de unión, ya que a veces las dianas cuentan con superficies cóncavas/convexas explícitas que pueden facilitar de forma muy clara la unión del ligando (2).

3.2 Cribado Virtual

El cribado virtual o *virtual screening* (VS) es la aplicación *in silico* de un filtrado computacional de posibles moléculas para seleccionar candidatos a *hits*, y así encontrar a aquel que encaje mejor con algunos criterios establecidos (3,5).

Esto es muy interesante desde el punto de vista del investigador y desarrollador de nuevos fármacos, ya que de esta forma, se reduce significativamente el número de ensayos biológicos que de forma tradicional serían necesarios para seleccionar una molécula apta (5).

Las técnicas para llevar a cabo el cribado virtual dependen de la complejidad de la base de datos que utilizemos y se pueden llevar a cabo diversos filtros, pudiendo usar un cribado virtual basado en la estructura (*docking*) si se conoce la estructura 3D, o buscando según similitud molecular basada en el ligando cuando no conocemos la diana (5).

El protocolo del *virtual screening* se basa en 3 pasos: El primer paso es de preparación de las estructuras, el siguiente se basa en el filtrado molecular y el último es la prueba *in silico*. El paso de la preparación es el más determinante, consistiendo en precisar de manera correcta las moléculas a emplear, asignándolas la geometría, carga y quiralidad adecuada. El filtrado consiste en seleccionar los complejos que mejor cumplen los parámetros o criterios de interés previamente establecidos. Posteriormente mediante las pruebas con *docking* obtendremos los *hits* candidatos a ser los precursores de un futuro fármaco clínicamente aprobado (20).

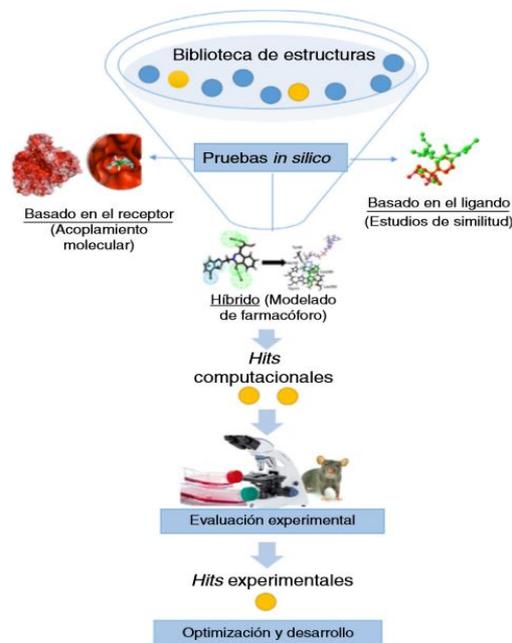


Figura 4. Representación esquemática del proceso de cribado virtual

Normalmente las simulaciones se suelen dar en una parte muy concreta de la proteína, como pasa en métodos como Glide, Dock y Autodock entre otros. A veces suponer que distintos ligandos van a interactuar con la proteína en la misma región lleva al error, descartando así otras áreas de dicha diana. Para solucionar esto se proponen ciertas alternativas como dividir toda la proteína en regiones definidas, donde se lleven a cabo simulaciones en todos los sitios especificados de la proteína simultáneamente. Esto nos podría descubrir nuevos lugares de unión de utilidad (6).

Existen paquetes de software disponibles libremente que estiman con precisión los descriptores moleculares para el filtrado de las bases de datos. Estos son muy útiles para predecir propiedades relacionadas con la actividad (como solubilidad o volumen molecular). Algunos son *Molinspiration*, *OSIRIS* ó *MoKa* entre otros (3).

3.3 Superordenadores

La frecuente necesidad de realizar cálculos cada vez más complejos nos ha llevado al uso de los superordenadores, los cuales tienen gran importancia en investigación científica e industria, ya que son máquinas de computación formadas por miles de procesadores que completan en minutos los cálculos que un ordenador convencional tardaría días. Pueden analizar grandes cantidades de información y así simular a escala molecular de forma más rápida las moléculas candidatas a *hits* para conseguir desarrollar nuevas terapias más efectivas o con menores efectos adversos (13).

Actualmente en España destacan superordenadores como el *MareNostrum 4* o el *IBM Power 9*, los cuales usan su capacidad de computación para simular reacciones y comportamientos de ciertos virus ante miles de compuestos. Estos superordenadores son capaces de buscar debilidades estructurales en el virus y poder plantear posibles flancos de ataque al mismo (14, 15).

4 Resultados y discusión

Cabe destacar que las técnicas *docking* han avanzado mucho en las últimas décadas, pero aún así, los científicos se encuentran ante un desafío debido a la frecuente complejidad del panorama energético de las interacciones entre moléculas. Esta complejidad se debe al hecho de que la función de energía se descompone en distintos términos (como el potencial de Lennard-Jones, solvatación entre moléculas, enlaces de hidrógeno, fuerzas electrostáticas...) (19).

El uso del cribado virtual ha conseguido abaratar en tiempo y dinero los procesos de descubrimiento de fármacos de manera muy notoria. 12 años de media y unos costes en torno a 1000 millones de euros es lo que supone llevar a cabo una investigación, hasta que se pone

el medicamento a disposición del paciente como molécula clínicamente apta. Esta industria es de alto riesgo ya que en torno al 98% de los fármacos que se comienzan a desarrollar no llegan al mercado. Esto es así ya que las empresas necesitan obtener beneficios para sustentar los costes de desarrollar un nuevo fármaco (25).

Todas estas técnicas ya mencionadas se están usando cada vez con más frecuencia para resolver problemas de salud pública. Las técnicas computacionales utilizadas con más frecuencia en epigenética son los procesos de *docking*, el modelado por homología y el filtrado computacional (20,21).

Un claro ejemplo es lo que ocurre al abordar el tratamiento terapéutico en algunos tipos de cánceres. La metilación del ADN es un proceso que se produce en el residuo de Citosina, más concretamente en la posición C5. Como los mecanismos de metilación no se conocen con certeza, han sido los aminoácidos que participan en el proceso los que han guiado a los investigadores en sus estudios de acoplamiento molecular. Gracias a estos estudios de *docking* se pudo observar que el diseño de inhibidores de la DNA Metil Transferasa (iDNMT) sería una estrategia perfecta para reducir la hipermetilación asociada a diversos tipos de cánceres (21).

En la actualidad se encuentran en uso clínico algunos de estos inhibidores como son la azacitidina y la decitabina (Ver imagen 5), ambos aprobados por la *Food and Drug Administration* (FDA) para el tratamiento de ciertos tipos de cáncer en los que las células sanguíneas de la médula ósea no maduran (síndromes mielodisplásicos). El problema recae en la citotoxicidad de estos compuestos, debido a su componente nucleosídico, por lo que se están usando asiduamente técnicas de acoplamiento molecular, para obtener fármacos por reposicionamiento (explicado a continuación) y así obtener *hits* a partir de procesos de cribado virtual (21).

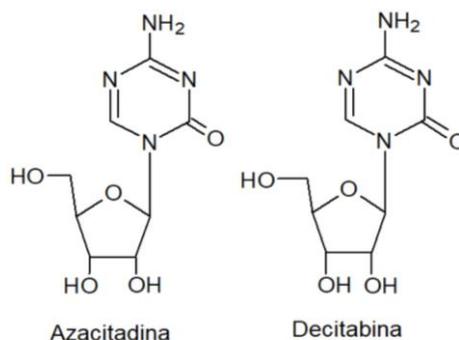


Imagen 5. Azacitidina y Decitabina

4.1 Reposicionamiento de fármacos

El reposicionamiento de fármacos ha contribuido a ahorrar hasta el 50% de las inversiones en tiempo y dinero, lo que ha supuesto un beneficio considerable tanto para la industria como para aquel que padece de dicha enfermedad cuyo fármaco se ha conseguido desarrollar (21). Podemos afirmar que moléculas similares tienen propiedades similares. Aunque esto no siempre sea así, gana cierta importancia ya que podemos reposicionar fármacos, encontrando nuevas aplicaciones terapéuticas con respecto a la que fueron diseñados. Así conseguimos reducir la investigación en tiempo y gastos. Un ejemplo muy característico es el reposicionamiento de la olsalazina, fármaco antiinflamatorio que fue identificado como agente hipometilante gracias a una búsqueda de similitud asistida computacionalmente en DrugBank, reposicionando así dicho fármaco como agente anticancerígeno (5).

Durante la última década se han usado mucho estas aproximaciones computacionales para cribar gran número de compuestos candidatos a fármacos, debido a su gran utilidad, y se ha ido desarrollando con rapidez este método en terapias como el tratamiento del VIH entre otras, ya que mediante aproximaciones basadas en el receptor o ligandos conocidos, tratamos de encontrar antagonistas de CXCR4 y CCR5. Gracias al VS, se ha conseguido mejorar los modelos de receptores a ambas dianas, llevando a cabo ensayos con ligandos conocidos de mayor afinidad. Para llegar a este punto, se han usado bases de datos de inhibidores de CXCR5/CCR5 y supuestos inactivos, muy parecidos en propiedades a los activos (12).

Esta optimización de fármacos ha sido algo clave para el desarrollo de muchos. Un caso concreto es el del Sildenafil, el cual se desarrolló como antihipertensivo pero no demostró efectividad. Por casualidad tras un proceso de reposicionamiento se le vio como altamente efectivo en el tratamiento contra la disfunción eréctil. Otros claros ejemplos de reposicionamiento son fármacos clásicos como la carbamazepina, la pregabalina o la gabapentina en el tratamiento contra la neuropatía diabética (21).

También destacan otros fármacos como los antihipertensivos Captopril o Aliskiren (*Ver imagen 5*), los cuales fueron descubiertos a partir del estudio de su diana (ECA y renina respectivamente). (6)

Nombre	Farmacéutica que lo desarrolló	Uso
Captopril	Bristol Myers-Squibb	Inhibidor de la ECA (Antihipertensivo)
Oseltamivir	Gilead Sciences	Inhibidor de la neuroaminidasa (Antiviral)
Dorzolamida	Merck	Inhibidor de la AC- α (Glaucoma)
Aliskiren	Novartis	Inhibidor de renina (Antihipertensivo)
Boceprevir	Schering-Plough	Hep-C (Antiviral)

Imagen 5. Tabla con algunos ejemplos de fármacos desarrollados por técnicas de cribado virtual

4.2 El éxito en los procesos basados en cribado virtual

A la hora de sacar conclusiones en la práctica, nos guiaremos por los valores *RMSD* (*Root Mean Square Deviation*), que se refieren a distancias y nos dan una idea de la diferencia estructural de dos topologías. Por esto, usaremos los valores de *RMSD* para saber cómo de alejadas están las soluciones de docking de una estructura de referencia o de otras soluciones. Gracias a estos resultados podemos comparar la fiabilidad de un programa de docking frente a una diana terapéutica de la que tenemos un complejo proteína-ligando o para comparar directamente distintas soluciones de docking entre sí. Así podemos afirmar que en un estudio de varias estructuras, la que menor valor de *RMSD* tenga, encajará mejor respecto a la estructura de referencia.

En la práctica vemos que las tasas de éxito en la unión entre dos moléculas en 0.5, 1.0, 1.5, 2.0 Å *RMSD* son de 19%, 47%, 60% y 64% respectivamente. Apenas hay ligandos acoplados entre 1.5 y 2.0 Å (4% promedio) y una pequeña fracción entre 1.5 y 2.0 Å (13% promedio). Esto nos indica que cuando un programa de acoplamiento molecular reconoce el modo correcto con el que se enlazan las moléculas, generalmente lo reproduce con alta precisión.

Para evaluar el rendimiento relativo, se calcula la probabilidad de que un método determinado sea el mejor para una sola ejecución. Los dos programas más diferenciables son Glide XP y FlexX, ya que este último tiene un 17% más de probabilidades de que funcione mejor en la reproducción de estructuras que Glide XP (*Ver tabla 6*) (22, 23).

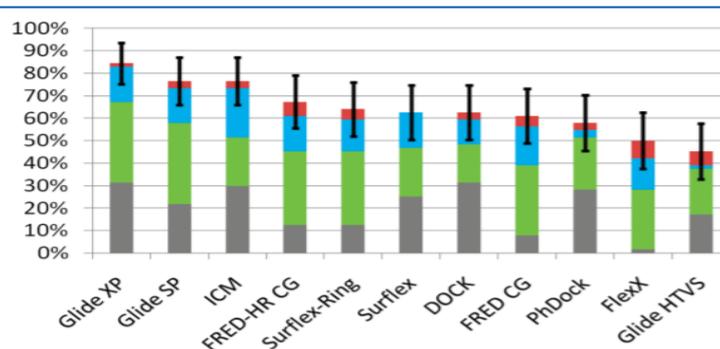


Tabla 6. Gráfica de barras que nos muestra valores RMSD que consiguen los distintos programas docking, donde la franja gris (0.5 Å), la verde (1.0 Å), la azul (1.5 Å) y la roja (2.0 Å) nos resumen la calidad de acoplamiento que genera dicho programa.

En definitiva, usar programas de *docking* para descubrir *hits* y poder optimizar a *leads* es un acto que realmente merece la pena, ya que se estima que se tienen 99% más de probabilidades de obtener un resultado satisfactorio que usando métodos aleatorios. Los mejores programas tienen aproximadamente un 75% de probabilidad de dar a una molécula activa una mejor puntuación de *scoring* que el de una inactiva (22).

5 Conclusión

En definitiva, está claro que las nuevas técnicas de acoplamiento molecular están abaratando los costes de las investigaciones y reduciendo el tiempo en el que estas se llevan a cabo y aunque no sean capaces de sustituir a las pruebas biológicas, ganan mucha importancia a la hora de optimizar tiempo y recursos. El avance de la tecnología se está aprovechando con rédito para así hacer llegar al mercado más fármacos y a menor coste, para el beneficio de la población.

Sin embargo debemos tener expectativas realistas respecto a lo que las técnicas computacionales nos ofrecen, ya que aún es bastante bajo el número de moléculas que llega a comercializarse comparado con las que se proponen. Pero esto no supone un problema ya que la mejoría con el paso de los años es evidente, simplemente hay que ir corrigiendo los problemas con los que las técnicas de cribado virtual se topan e ir puliendo estas para obtener moléculas *leads* de forma más precisa. Invertir en el desarrollo de fármacos debe ser algo crucial para conseguir resultados terapéuticos más precisos y beneficiar así a los ciudadanos.

6 Bibliografía

1. Morrale, A., Perona, A., Klett, J., Cortés-Cabrera, A., & Dos Santos, H. G. (2012). Diseño de fármacos asistido por ordenador. *Dinámica estructural y diseño de fármacos*, 401-424.
2. Gupta, M., Sharma, R., & Kumar, A. (2018b). Docking techniques in pharmacology: How much promising? *Computational Biology and Chemistry*, 76, 210-217. <https://doi.org/10.1016/j.compbiolchem.2018.06.005>
3. Ferreira, L, dos Santos, R., Oliva, G., & Andricopulo, A. (2015). Molecular Docking and Structure-Based Drug Design Strategies. *Molecules*, 20(7), 13384-13421. <https://doi.org/10.3390/molecules200713384>
4. Rachman, M., Barril, X., & Hubbard, R. (2018). Predicting how drug molecules bind to their protein targets. *Current Opinion in Pharmacology*, 42, 34-39. <https://doi.org/10.1016/j.coph.2018.07.001>
5. Saldívar-González, F., Prieto-Martínez, F. D., & Medina-Franco, J. L. (2017). Descubrimiento y desarrollo de fármacos: un enfoque computacional. *Educación Química*, 28(1), 51-58. <https://doi.org/10.1016/j.eq.2016.06.002>
6. Pérez-Sánchez, H., Cano, G., García-Rodríguez, J., & Cecilia, J. M. (2015). Descubrimiento de fármacos basado en cribado virtual refinado con enfoques neuronales paralelos. *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería*, 31(4), 207-211. <https://doi.org/10.1016/j.rimni.2014.06.004>
7. Bravo, J. (2012). Cristalografía de rayos X de macromoléculas. *Instituto de Biomedicina*, 1. https://doi.org/10.18567/sebbmdiv_rpc.2012.02.1
8. Manuales MSD. (2019). Resonancia magnética nuclear (RMN). Recuperado de <https://www.msmanuals.com/es-es/hogar/temas-especiales/pruebas-de-diagn%C3%B3stico-por-la-imagen-habituales/resonancia-magn%C3%A9tica-nuclear-rmn>
9. Díaz Caballero, A., Martínez Serrano, E., Vivas Reyes, R., Puerta Llerena, L., Menéndez Cuadro, D., Cabrales Salgado, R., & Padilla Rodríguez, A. (2012). Modelling by homology of Luxs protein in *Porphyromonas gingivalis* strain W83. *Rev. Clin. Periodoncia Implantología*, 105-113. Recuperado de <https://scielo.conicyt.cl/pdf/piro/v5n3/art01.pdf>
10. The Cambridge Structural Database. (2020). Recuperado de <https://www.ccdc.cam.ac.uk/solutions/csd-system/components/csd/>
11. Bank, R. P. D. (2017). RCSB PDB. Recuperado de https://www.rcsb.org/pdb/static.do?p=general_information/about_pdb/index.html
12. Pérez Nueno, V. (2009, mayo 26). Herramientas de cribado virtual aplicadas a inhibidores de entrada del VIH. Diseño de nuevos compuestos anti-VIH. Recuperado de <https://www.tdx.cat/handle/10803/9311?show=full>
13. Velasco, E. (2019, junio 8). ¿Para qué sirve un superordenador? Recuperado de <https://www.lavanguardia.com/ciencia/20190608/462719053611/superordenador-para-que-sirve-barcelona-marenostrum.html>

14. Krama. (2020, abril 7). La función de los superordenadores en la búsqueda de fármacos contra el Covid-19. Recuperado de <https://www.krama.es/blog-20-04-07-superordenadores-farmacos-covid-19.html>
15. López, J. (2019, marzo 8). Superordenadores en el diseño de nuevos fármacos | Bio (Ciencia+Tecnología). Recuperado de <https://www.madrimasd.org/blogs/biocienciatecnologia/2019/03/08/134220>
16. Venkatachalam, C. M., Jiang, X., Oldfield, T., & Waldman, M. (2003). LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics and Modelling*, 21(4), 289-307. [https://doi.org/10.1016/s1093-3263\(02\)00164-x](https://doi.org/10.1016/s1093-3263(02)00164-x)
17. Bonvin, A. M. (2006). Flexible protein–protein docking. *Current Opinion in Structural Biology*, 16(2), 194-200. <https://doi.org/10.1016/j.sbi.2006.02.002>
18. Jiang, F., & Kim, S. (1991). “Soft docking”: Matching of molecular surface cubes. *Journal of Molecular Biology*, 219(1), 79-102. [https://doi.org/10.1016/0022-2836\(91\)90859-5](https://doi.org/10.1016/0022-2836(91)90859-5)
19. Zarbafian, S., Moghadasi, M., Roshandelpoor, A., Nan, F., Li, K., Vakli, P., Paschalidis, I. (2018). Protein docking refinement by convex underestimation in the low-dimensional subspace of encounter complexes. *Scientific Reports*, 8(1), <https://doi.org/10.1038/s41598-018-23982-3>
20. Prieto-Martínez, F. D., & Medina-Franco, J. L. (2017). Diseño de fármacos asistido por computadora: Cuando la informática, la química y el arte se encuentran. *TIP Revista especializada en Ciencias Químico-biológicas*, 21(2), 124-135. <https://doi.org/10.22201/fesz.23958723e.2018.2.6>
21. 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. <https://doi.org/10.1016/j.eq.2015.05.002>
22. McGann, M. (2011). FRED Pose Prediction and Virtual Screening Accuracy. *Journal of Chemical Information and Modeling*, 51(3), 578-596. <https://doi.org/10.1021/ci100436p>
23. McGann, M., Nicholls, A., & Enyedy, I. (2015). The statistics of virtual screening and lead optimization. *Journal of Computer-Aided Molecular Design*, 29(10), 923-936. <https://doi.org/10.1007/s10822-015-9861-4>
24. MSD Salud. (2012). Recuperado de <https://www.msdsalud.es/informacion-practica/proceso-investigacion-farmaco.html>
25. Como se fabrica un fármaco. Fase 1: Investigación previa. (2015). Recuperado de <https://www.eupati.eu/es/descubrimiento-de-farmacos/como-se-fabrica-un-farmaco-fase-1-investigacion-previa/>
26. Font Maté, C. (2017, junio). MODELADO MOLECULAR COMO HERRAMIENTA PARA EL DESCUBRIMIENTO DE NUEVOS FÁRMACOS QUE INTERACCIONAN CON PROTEÍNAS. Recuperado de <http://147.96.70.122/Web/TFG/TFG/Memoria/CRISTINA%20FONT%20MATE.pdf>