

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 795 840**

51 Int. Cl.:

**G16C 20/50** (2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.05.2012 PCT/US2012/039001**

87 Fecha y número de publicación internacional: **29.11.2012 WO12162320**

96 Fecha de presentación y número de la solicitud europea: **22.05.2012 E 12789317 (0)**

97 Fecha y número de publicación de la concesión europea: **11.03.2020 EP 2715602**

54 Título: **Puntuación de afinidad de unión con penalización para la rotura de la conjugación entre grupos de ligandos aromáticos**

30 Prioridad:

**23.05.2011 US 201113113506**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

**24.11.2020**

73 Titular/es:

**SCHRODINGER, LLC (100.0%)  
120 West Forty-Fifth Street 17th Floor Tower 45  
New York, New York 10036-4041, US**

72 Inventor/es:

**FRIESNER, RICHARD A. y  
MURPHY, ROBERT**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 795 840 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Puntuación de afinidad de unión con penalización para la rotura de la conjugación entre grupos de ligandos aromáticos

## CAMPO TÉCNICO

5 La invención está en el campo general de los métodos basados en ordenador para la estimación de la afinidad de unión entre un ligando y una molécula de receptor.

## ANTECEDENTES

10 Muchos fármacos operan uniéndose químicamente a receptores moleculares específicos. Los receptores moleculares son normalmente proteínas (un término que incluye glucoproteínas y lipoproteínas) en un animal tal como un ser humano, y se pueden facilitar el diseño y la selección de fármacos estimando con exactitud la afinidad de unión de un fármaco a una proteína, o, más en general, estimando la afinidad de unión de un ligando por un receptor, siendo el término receptor usado para significar cualquier resto que se une específicamente al ligando.

15 Una forma de determinar la afinidad de unión receptor-ligando usa la estructura molecular que resulta cuando el ligando se une al receptor ("el complejo ligando-receptor"). Dichas estructuras pueden ser estudiadas por cristalografía de rayos X. El públicamente accesible Protein Data Bank (PDB) contiene ahora más de 70.000 estructuras cristalinas de rayos X, y las empresas farmacéuticas y de biotecnología tienen un orden de magnitud de más estructuras patentadas. Muchas de estas estructuras han sido cristalizadas conjuntamente con moléculas pequeñas unidas a ellas. El examen de dichas estructuras, y el empleo del conocimiento así obtenido para diseñar nuevos inhibidores, más potentes, y más específicos, se denomina el diseño de fármacos basado en estructura.

20 El modelado computacional facilita el diseño de fármacos basados en estructura. Un aspecto del modelado detallado a continuación implica puntuar funciones que usan técnicas de simulación, tales como dinámica molecular, Monte Carlo, o cálculos de electrostática continua. Las funciones de puntuación pueden ser problemáticas cuando se requiere calcular una diferencia muy pequeña (la afinidad de unión) entre dos números muy grandes (las energías libres del complejo y de la proteína y ligando separados). Un enfoque alternativo es desarrollar una función de puntuación empírica, basada en la geometría del complejo, que evalúa directamente la cantidad deseada. Dicho enfoque tiene la ventaja de ser extremadamente rápido, así como ser susceptible al ajuste de grandes cantidades de las que están ahora públicamente disponibles para los datos experimentales. El documento de patente US 2007/0061118 A1 del mismo solicitante "Predictive scoring function for estimating binding affinity" desvela dichas funciones de puntuación.

25 Es conveniente aumentar la exactitud y robustez de las funciones de puntuación haciendo mejoras materiales en la forma funcional que mejor refleja la realidad física. Diversas solicitudes de patente que son del mismo solicitante que la presente solicitud describen formas para mejorar las funciones de puntuación: documento de patente WO/2008/141260, titulado "Binding Affinity Scoring Function Including Factor For Environs Of Ring or Bulky Rigid Group"; N° de serie US 13/079.725, presentado el 4 de abril de 2011, titulado "Binding Affinity Scoring Function Penalizing Compounds which Make Unfavorable Hydrophobic Contacts With Quasi-localized Water Molecules in the Receptor Active Site"; y N° de serie US 13/079.489, presentado el 4 de abril de 2011, titulado "Scoring Function Penalizing Compounds which Desolvate Charged Protein Side Chains Structure". Un segundo problema importante con las funciones de puntuación es que pueden asignar mejores puntuaciones de afinidad de unión (más negativas) a los compuestos inactivos (es decir, compuestos que no se determinaría que se unirían al receptor en un protocolo de cribado experimental típico) que a los compuestos activos. Si un gran número de compuestos inactivos se encuentran a la cabeza de los compuestos activos, sería difícil llevar a cabo eficazmente una función principal de acoplamiento molecular —para descubrir nuevos compuestos activos contra un receptor especificado de una biblioteca de compuestos muy grande (frecuentemente millones de compuestos). Por tanto, reducir sustancialmente el número de "positivos falsos" (clasificación de compuestos inactivos como activos) mejoraría enormemente el rendimiento de la función de puntuación.

30 El documento de patente US 2008/0312894 A1 describe que las funciones de puntuación se pueden mejorar teniendo en cuenta el entorno del estado de los anillos del ligando (o de hecho otras estructuras de ligando rígidas voluminosas) en el ligando cuando el ligando se compleja con el receptor.

Gilson MK et al. describen en Annual Review of Biophysics and Biomolecular Structure, vol. 36, no. 1, 1 de junio de 2007 en las páginas 21-42, el cálculo de las afinidades de unión proteína-ligando.

## SUMARIO

35 La presente invención se define por la reivindicación independiente 1. Las reivindicaciones dependientes representan otras realizaciones de la invención.

40 Los presentes inventores han descubierto un método implementado por ordenador mejorado de puntuación de la afinidad de unión de una molécula de ligando propuesta por una molécula de receptor, que es una proteína. El método usa análisis informático y bases de datos informáticas para explicar el aumento en la energía requerida donde el acoplamiento molecular altera o altera parcialmente el carácter conjugado  $\pi$  del ligando cuando se une al receptor. El

método usa datos que representan una o más moléculas de ligando propuestas a puntuar y datos que representan la molécula de receptor. El análisis informático de los datos de la molécula de ligando propuesta determina si el ligando incluye al menos un resto conjugado  $\pi$  que tiene múltiples geometrías posibles, siendo una de las geometrías caracterizadas por menos deslocalización de electrones a través del resto conjugado  $\pi$  que la deslocalización de electrones que caracterizan otra de las geometrías. El análisis informático de la estructura ligando-receptor predicha determina si el ligando en la estructura ligando-receptor adopta la geometría caracterizada por menos deslocalización. Si es así, se impone explícitamente una penalización para la deslocalización reducida de electrones a través de los restos conjugados  $\pi$ .

En realizaciones preferidas, el análisis informático del ligando determina si el ligando tiene una o más trayectorias de enlace giratorio que conectan dos restos químicos que son capaces de una significativa deslocalización de electrones  $\pi$  entre los grupos en el enlace giratorio. El método determina entonces la configuración del ligando en un complejo ligando-receptor, y, para al menos una trayectoria de enlace giratorio, determina si la geometría de la configuración de acoplamiento molecular reduce sustancialmente la deslocalización de electrones de una geometría de ligando energéticamente accesible en disolución o en fase gaseosa. Por ejemplo, el método determina si el resto conjugado  $\pi$  incluye al menos un enlace giratorio cuya rotación cambia el grado de deslocalización de electrones a través del resto conjugado  $\pi$ . El resto conjugado  $\pi$  puede incluir al menos un anillo aromático. La magnitud de la penalización se puede establecer a una cantidad prevista para reflejar, en general, la reducción de la deslocalización de electrones entre la configuración del ligando propuesto cuando se compleja con el receptor y la deslocalización de una configuración sin complejar, más deslocalizada, del ligando propuesto. Frecuentemente, el resto conjugado  $\pi$  es más plano en la configuración caracterizada por mayor deslocalización que lo que es en la configuración caracterizada por menos deslocalización. En realizaciones específicas, las energías de las configuraciones del resto conjugado  $\pi$  se pueden medir por uno o más de: a) un campo de fuerzas de mecánica molecular, b) métodos de química cuántica, o c) desviaciones de la planaridad. Cuando se usa la última medición, los expertos en la técnica entenderán cómo medir la planaridad. Por ejemplo, se puede medir por los valores de los ángulos de torsión en la trayectoria conjugada. Eficazmente cero grados o 180 grados es planaridad completa. Las desviaciones de estos ángulos implican desviaciones de la planaridad. Un ángulo de 90 grados implica la máxima desviación de la planaridad.

Por tanto, en realizaciones preferidas, se pueden ignorar choques entre átomos que forman enlaces de hidrógeno internos. El ligando puede incluir múltiples trayectorias de enlace giratorio conjugados y se evalúa una penalización para cada trayectoria de enlace donde el ligando en la estructura ligando-receptor adopta la geometría caracterizada por menos deslocalización. Se puede usar un método de búsqueda conformacional para investigar la energía de al menos una conformación del ligando en disolución o en fase gaseosa y para identificar una confirmación que presenta conjugación entre los restos conectados por la trayectoria de enlace giratorio. El grado de conjugación a lo largo de la trayectoria de enlace giratorio se puede determinar por la desviación de la planaridad de los restos conectados por la trayectoria de enlace giratorio. Se puede usar un campo de fuerzas de mecánica molecular para calcular las energías de las diversas conformaciones en la búsqueda conformacional. Se pueden usar métodos de química cuántica para calcular las energías de las diversas conformaciones en la búsqueda conformacional. Se puede usar un modelo de interacción simplificada -- tal como uno que imponga distancias mínimas permitidas entre átomos de diversos tipos y basándose en la energía de conjugación tras la desviación de la planaridad -- para evaluar la energética de las diversas conformaciones en la búsqueda conformacional. La magnitud de la penalización se puede determinar por el grado de distorsión de la conformación de ligando en el complejo acoplado de una conformación de baja energía en disolución o en fase gaseosa. Se puede asignar una penalización de cero a una trayectoria de enlace giratorio conjugado si la conformación de energía más baja en disolución o en fase gaseosa no permite la sustancial conjugación entre los dos restos conectados por la trayectoria de enlace giratorio conjugado, debido a la gran desviación de la planaridad.

Una forma de enumerar las trayectorias de enlace aromático es identificar átomos aromáticos y enumerar las trayectorias de enlace de la conectividad de los átomos aromáticos identificados. Entonces se identifican los enlaces giratorios a lo largo de las trayectorias de enlace aromático y se calculan los ángulos de torsión de los enlaces giratorios. La penalización se puede estimar a partir de los valores de los ángulos de torsión. En una realización, se identifican los tipos de átomos aromáticos que componen un resto conjugado  $\pi$ , se enumeran las trayectorias de enlace aromático dentro de un resto conjugado  $\pi$ , se identifican enlaces giratorios aromáticos que se encuentran a lo largo de estas trayectorias de enlace aromático, se calculan los ángulos de torsión de estos enlaces giratorios aromáticos y se calcula una penalización por energía libre en función de los ángulos de torsión.

La divulgación también caracteriza medios legibles por ordenador que almacenan instrucciones que provocan que un ordenador realice el método de la reivindicación 1.

Otras características y ventajas de la invención son evidentes de la siguiente descripción, y de las reivindicaciones.

## BREVE DESCRIPCIÓN DE LOS DIBUJOS

Figura 1 La molécula señuelo 681094 representada aquí sufre una penalización por los ángulos diédricos adoptados por la molécula en un resto conjugado. Los diedros específicos que conducen a la penalización se representan en rojo.

## 5 DESCRIPCIÓN DETALLADA

La práctica de la invención empieza con una estructura de receptor (o una proteína "diana") que tiene resolución suficiente para permitir el uso de software computacional para "acoplar" un ligando de molécula pequeña en la correcta posición y orientarla en la cavidad del sitio activo del receptor y para calcular una afinidad de unión del ligando dada esta estructura. Los programas de software informático que realizan esta tarea se denominan los programas de "acoplamiento molecular".

Un programa de acoplamiento molecular normalmente lleva a cabo dos tareas distintas para modelar la unión receptor-ligando. Primero, se predice una estructura de un complejo receptor-ligando. Cuando falla esta suposición, se usa una estructura diferente del receptor como punto de partida. Es muy importante el problema de construir estructuras de receptor alternativas que se modifican para aceptar ligandos que requieren un cambio sustancial en la conformación de receptor ("ajuste inducido").

Se conocen bien diversos programas de predicción de estructuras de receptor-ligando adecuados y se pueden usar para el punto de partida, y los expertos en la técnica entenderán que la invención descrita en el presente documento se puede usar, en general, con programas de predicción receptor-ligando. Los ejemplos de dichas funciones que están fácilmente disponibles incluyen: Glide, GOLD, FRED, FlexX o AutoDock, entre muchas otras.

Una vez en posesión de una estructura de receptor en acuerdo razonable con datos experimentales, la segunda tarea del programa de acoplamiento molecular es calcular una afinidad de unión receptor-ligando, dada como una entrada de la estructura acoplada. Una función matemática empleada para calcular la afinidad de unión (o una contribución a la misma) se denomina una "función de puntuación". Las mejoras a dichas funciones de puntuación para calcular la afinidad de unión receptor-ligando son el sujeto de la presente invención.

Las estimaciones de las afinidades de unión receptor-ligando descritas a continuación son aplicables cuando una estructura del complejo receptor-ligando se representa por un modelo estructural adecuado. Existen varias formas de caracterizar la calidad de modelos estructurales de complejos receptor-ligando, mientras que el modelo esté adecuadamente de acuerdo con los datos experimentales. Se podrían usar medidas de concordancia estructural, tales como la puntuación RMSD, DME o SIFt, entre otras. Normalmente se obtienen puntuaciones, pero no exclusivamente, usando: i) una pequeña desviación cuadrática media (RMSD) de la estructura experimental (normalmente inferior a 2 Angstroms, aunque el valor requerido variará dependiendo de los detalles del complejo), ii) reconocimiento de la formación de sustancialmente todos los enlaces de hidrógeno observados en el complejo experimental, iii) ubicación apropiada de sustancialmente todos los grupos hidrófobos en las cavidades de receptor correctas, y iv) la ausencia de choques estructurales o electrostáticos incorrectos que podrían conducir a la asignación de términos de penalización sustancialmente incorrectos. Puesto que está en estudio la afinidad de unión de ligandos por un receptor dado, una compensación constante, como es en muchos casos ocasionada por la energía de reorganización del sitio activo del receptor para acomodar el ligando, no tiene efecto sobre las aplicaciones prácticas.

En muchos casos (aunque no todos), el receptor puede adoptar más de una conformación fundamentalmente diferente en respuesta a una clase de ligandos (por ejemplo, ligantes DFG-dentro y DFG-fuera para map cinasa p38); para comparar las energías libres de unión en tales casos, se pueden requerir diferentes parámetros de reorganización del núcleo para las diferentes conformaciones de receptor. Donde el cálculo de estos parámetros no sea práctico a partir de los primeros principios, se tratan como parámetros específicos de receptor ajustables. Otros parámetros están, sin embargo, contenidos en un modelo global que no es específico de receptor o incluso específico para una clase particular de receptores.

La presente invención mejora la capacidad de la función de puntuación para discriminar entre compuestos inactivos y activos contra un receptor dado. Mientras que los compuestos activos para un amplio intervalo de receptores se encuentran fácilmente en Protein Data Bank (PDB), los datos de los presentes inventores de más adelante que validan la invención también evaluaron compuestos inactivos conocidos como control negativo. El acceso a compuestos inactivos conocidos desde datos públicamente disponible es exigente. Por tanto, los presentes inventores han ideado un protocolo alternativo que se basa en el uso de una biblioteca aleatoria de 1000 compuestos similares a fármacos. Estos compuestos están acoplados molecularmente en una conformación del receptor, y se obtienen afinidades de unión predichas para cada compuesto usando diversas mejoras propuestas a la función de puntuación.

La función de puntuación de los presentes inventores se calibra de manera que los compuestos activos alcancen puntuaciones que normalmente son próximas a las afinidades de unión experimentales, con una desviación estándar de aproximadamente 1 unidad logarítmica de la afinidad de unión (~1,5 kcal/mol). Se considera, en general, que una "interacción positiva" experimental en un cribado aleatorio en la industria farmacéutica tiene una afinidad de unión de ~7,0 kcal/mol o más (~10  $\mu$ m de concentración). Dadas las fluctuaciones intrínsecas en la función de puntuación de 1,5 - 2,0 kcal/mol, los presentes inventores establecen el umbral computacional para estimar interacciones positivas

como una puntuación de -9,0 kcal/mol; así se predice que cualquier compuesto en la biblioteca de 1000 compuestos aleatorios que puntúa -9,0 o menos es una interacción positiva. Se predice que otros compuestos en el intervalo de -7,0 a -9,0 también son posibles interacciones positivas, pero dichas puntuaciones también pueden ser debidas al ruido en la función de puntuación que los presentes inventores son actualmente incapaces de reducir más (en parte debido a las limitaciones a los datos experimentales públicamente disponibles). Por tanto, los presentes inventores se centran en desarrollar las mejoras de los presentes inventores para asegurar que el número de compuestos aleatorios que puntúan más negativamente que -9,0 es compatible con las tasas de interacciones positivas experimentales para bibliotecas de compuestos aleatorios de tipo fármaco.

La tasa de interacciones positivas para los cribados experimentales variará dependiendo del receptor, pero una tasa de interacciones positivas "promedio" ilustrativa es 0,5 %, o 5 compuestos de los 1000. Así, si sustancialmente más de 5 compuestos alcanzan una puntuación de -9,0 kcal/mol, se hace la suposición de que los compuestos inactivos están recibiendo puntuaciones que son demasiado favorables.

En dicha situación, la función de puntuación se puede mejorar añadiendo términos de "penalización" positiva que reducen la magnitud de la afinidad de unión predicha. Dichos términos representan procesos físicos que hacen la unión menos favorable. Un ejemplo de un proceso de este tipo sería una desolvatación en la que un grupo polar de la proteína o ligando se bloquea por grupos no polares del ligando y pierde el acceso al agua. Esto da como resultado una gran pérdida en energía libre, inactivando el compuesto. Si la pérdida de energía libre es suficientemente grande, entonces dichas penalizaciones solo se observarán raramente (si se observan) en complejos con compuestos activos. Esto significa que los nuevos términos se deben obtener examinando las estructuras producidas cuando la biblioteca aleatoria se acopla molecularmente en el receptor.

La invención descrita en el presente documento se diseña para identificar un tipo particular de geometría para la que se puede asignar una gran energía de deformación del ligando, sustancialmente mayor que la presente en la gran mayoría de los compuestos activos en la base de datos de proteínas. La geometría implica la rotación alrededor de un enlace dentro de un resto conjugado  $\pi$  de un ligando, normalmente compuesto por 1 o más (sub)restos aromáticos. En esta situación, en la geometría apropiada, se deslocalizan los electrones en el resto conjugado  $\pi$  en una descripción de mecánica cuántica rigurosa, que reduce la energía de esta geometría en comparación con geometrías alternativas que alteran la deslocalización de electrones en el resto conjugado  $\pi$ . Cuando la geometría del sitio activo puede forzar al ligando en una geometría no óptica con deslocalización de electrones reducida en el resto conjugado  $\pi$ , se introduce una gran cantidad de energía de deformación, y se aplica una penalización sustancial, por ejemplo, 4 kcal/mol.

Aplicar el criterio anterior sería relativamente simple, una vez se definen la aromaticidad y las trayectorias de enlace que preservan la conjugación, si el ligando solo consistió en dos grupos conjugados, y estos grupos no interaccionaron, excepto mediante la conjugación. En este caso, la geometría ideal para reducir la energía mediante la conjugación es plana, y las desviaciones de la planaridad serían penalizadas según la magnitud de la desviación. Sin embargo, para muchos ligandos, la situación es mucho más complicada debido a otras interacciones, tales como repulsiones estéricas, enlaces de hidrógeno, que perturban la geometría ideal. La repulsión estérica en la geometría plana puede desestabilizar esta geometría y desplazar la energía mínima a una geometría no plana. El enlace de hidrógeno puede estabilizar una geometría que permite el enlace de hidrógeno. La invención descrita a continuación usa algoritmos matemáticos para identificar tipos de átomos y luego explora las diversas posibles geometrías alrededor del posible enlace de deformación, usando algoritmos empíricos para estimar la energía libre en función del ángulo de torsión en cuestión (que en muchos casos implica también buscar otros grados de libertad acoplados). Si se determina que la geometría del ligando en el receptor (obtenida, por ejemplo, a partir del acoplamiento molecular, de otros tipos de simulaciones, o de experimentos tales como cristalografía de rayos X) pierde energía de conjugación significativa en comparación con lo que es posible en disolución (esto última está afectado por los términos de repulsión estérica y enlace de hidrógeno), entonces se aplica una penalización de conjugación aromática.

La penalización energética anterior se podría asignar en principio llevando a cabo cálculos de mecánica cuántica, o mediante un campo de fuerzas de mecánica molecular altamente preciso. Sin embargo, los cálculos de química cuántica relevantes pueden ser muy caros computacionalmente (particularmente puesto que muchas geometrías pueden tener que ser examinadas para un ligando dado) y, actualmente, los campos de fuerzas de mecánica molecular no describen de forma fiable las energías de conjugación para una amplia variedad de química de ligandos. La invención proporciona una solución muy rápida a este problema crítico (aproximadamente >15 % por ciento de los ligandos en una base de datos aleatoria de 1 millón de moléculas de tipo fármaco tienen al menos un enlace que proporciona posiblemente una trayectoria de conjugación entre dos grupos aromáticos) y, como se muestra a continuación, es muy satisfactorio en centrar compuestos activos aleatorios a diferencia de conocidos para una amplia variedad de receptores.

Los detalles del algoritmo preferido que representa la invención son del siguiente modo, pero se pueden usar otros algoritmos que efectúan las etapas de la invención descritas aquí.

- 1) Asignar los siguientes átomos como aromáticos:

## ES 2 795 840 T3

a) Se marcan los átomos del anillo aromático por tipificación de átomos, tipificación de grupos u otros métodos; se realiza una comprobación para determinar que los carbonos y nitrógenos así marcados no tienen más de tres conexiones.

b) átomos en enlaces amida y conectores NH.

5 2) Asignar los siguientes enlaces como enlaces giratorios aromáticos:

a) cualquier enlace giratorio que consiste en dos átomos aromáticos como se marca en 1).

b) enlace(s) giratorio(s) que conectan dos anillos aromáticos.

10 3) Encontrar trayectorias aromáticas de enlaces giratorios a partir de un anillo aromático. Esto se puede hacer usando la lista de enlaces aromáticos anterior y el código de anclaje Glide XP™ (disponible de Schrodinger, LLC, Nueva York, NY) para buscar la trayectoria de enlaces que emanan de un anillo aromático. Se conocen bien por la técnica otros métodos de búsqueda de trayectorias, tales como aquellos descritos en el campo general de aplicaciones de métodos teóricos gráficos y métodos matemáticos combinatorios para problemas químicos. La trayectoria termina cuando se alcanza una conexión no aromática o se alcanza un anillo aromático separado. Obsérvese que una trayectoria aromática puede contener una unidad C=C o unidad -

15 N=X- aromática no giratoria. A efectos de los contactos átomo-átomo marcados, guardar la lista de átomos en cada trayectoria que incluya anillo(s).

4) Para cada trayectoria, construir una lista de torsiones a lo largo de la trayectoria usando la lista de enlaces giratorios de la trayectoria, así como átomos conectados a estos enlaces giratorios. Así, cada trayectoria tiene una lista asociada de torsiones.

20 5) Para cada torsión enumerada en (4), calcular el ángulo de torsión en la geometría de entrada, es decir, el valor (o  $180 +$  este valor) necesario para girar la torsión hasta planaridad.

25 6) Para cada trayectoria aromática, determinar si habrá choques entre átomos / anillos de la trayectoria cuando todas las torsiones de la trayectoria giren hasta la planaridad. La lista de átomos usada para comprobar los choques incluye átomos en el (los) anillo(s), átomos de los enlaces giratorios y átomos conectados al conjunto anillo/enlace giratorio. Puesto que cada torsión cumple la planaridad a un ángulo  $\Theta$  y un ángulo  $= 180^\circ + \Theta$ , se pueden muestrear  $2^n$  posibles combinaciones de ángulos de torsión a modo de dicho análisis de búsquedas conformacionales para probar si una no implica choques átomo-átomo. Los cortes usados para definir choques (entre pares de átomos pesados) son los siguientes.

a) por defecto de  $3,1 \text{ \AA}$  entre átomos pesados

30 b)  $3,25 \text{ \AA}$  si un átomo es un miembro de un anillo con más 6 átomos en él.

c) si uno de los átomos es un oxígeno de  $\text{SO}_2$ , se usa un corte de  $3,4 \text{ \AA}$ . La técnica entenderá a partir de esta memoria descriptiva que se pueden necesitar otros casos especiales.

35 En ciertos casos, se ignoran los choques. Por ejemplo, los átomos que chocan deben estar separados al menos 1-5 átomos en la tabla de conexión. No se consideran choques entre átomos que forman enlaces H internos (por ejemplo, CH con N).

Si se produce una geometría para la que no ocurren choques, entonces la trayectoria de torsión se mantiene en la lista de torsiones a medir en la formación de una penalización por no planaridad. En algunos casos, solo se puede rotar un subconjunto de la trayectoria aromática completa hasta la planaridad sin incurrir en choques y en estos casos considerar solo el subconjunto de torsiones de a continuación (en la función de penalización).

40 Las Etapas 1-6 son el procesamiento previo dada la geometría del ligando de entrada. Mientras se evalúa cualquier planteamiento dado durante el acoplamiento molecular, se evalúa la siguiente función que detecta torsiones no planas en las trayectorias.

7) Para cada torsión (i) guardada en (6), calcular el ángulo de torsión ' $\Theta_i$ '. Una simple penalización  $w_{\text{tor}_i}$  para esta torsión viene dada por;

45 
$$w_{\text{tor}_i} = 0,0 \text{ si } \Theta_i < \Theta_o$$

$$w_{\text{tor}_i} = A * \text{abs}(\text{sen}(\Theta_i - \Theta_o))$$

50  $\Theta_o$  es 20 grados para trayectorias de tipo no bifenilo y 40 grados para trayectorias de tipo bifenilo donde una trayectoria de tipo bifenilo es un enlace sencillo que conecta dos anillos aromáticos uno de los cuales es de tamaño 6. El valor de A dependerá de la escala de otros términos, pero en este ejemplo será entre 5 y 20 (por ejemplo, 10).

La penalización por torsión total a una geometría dada es simplemente la suma de  $w_{tor\_i}$ .

La eficacia de la invención anterior se demuestra del siguiente modo. Primero, los presentes inventores han ensamblado un conjunto de pruebas de los 622 complejos de proteína-ligando de compuestos activos. Como control para evaluar el método, los ejemplos de más adelante implican estructuras cristalinas conocidas disponibles en Protein Data Bank ("PDB") públicamente accesible de Research Collaborative for Structural Bioinformatics. Al llevar a cabo la optimización, los presentes inventores usan planteamientos acoplados con Glide XP™, una función de puntuación, descrita, en general, por ejemplo, en el documento de patente US 2007/0061118 A1, que filtra los (muy pocos) casos para los que el auto-acoplamiento molecular da estructuras inadecuadas. Usando estructuras acopladas, en vez de las estructuras cristalinas en sí, en el proceso de optimización de los presentes inventores, aumentan el realismo del modelo y también permiten su uso con Glide XP™. La invención se puede usar con cualquiera de un número sustancial de funciones de puntuación comerciales, que incluyen, sin limitación, FRED, FlexX o AutoDock

Las estructuras de PDB se pueden visualizar como un conjunto de entrenamiento grande y diverso para la función de puntuación. La invención no requiere el uso de ligandos para los que se conocen estructuras cristalinas, ni requiere el uso de Glide XP. La prueba de la función de puntuación en condiciones similares se puede realizar por empresas farmacéuticas y de biotecnología, usando conjuntos de datos patentados donde están disponibles estructuras cristalinas. Al llevar a cabo estas pruebas, no existe necesidad de liberar las estructuras o incluso de divulgar el nombre del receptor; se pueden realizar simplemente los cálculos, e informar de la capacidad de clasificar el orden de los compuestos como un coeficiente de correlación.

Estos complejos dan estructuras razonablemente precisas cuando el ligando se vuelve a acoplar molecularmente en su receptor nativo (por ejemplo, una RMSD máxima de 3,5 Å, con la gran mayoría de RMSD por debajo de 2 Å) y sus puntuaciones, usando la función de puntuación Glide XP™, están en promedio dentro de ~ 1 kcal/mol de la afinidad de unión experimental. Así, la función de puntuación en ausencia del término que constituye la invención funciona bien para complejos de compuestos activos tomados de PDB.

Los presentes inventores han añadido un nuevo término a la función de puntuación Glide XP™, que representa la penalización anteriormente descrita, y los presentes inventores han recalculado la afinidad de unión de todos los complejos activos.

La Tabla 1 informa una lista de complejos PDB que están afectados por este término, la afinidad de unión experimental de cada complejo (dG) y la afinidad de unión calculada con (Score\_np) y sin (Score\_p) el nuevo término. Se puede observar que existen muy pocos casos en los que los complejos activos cumplen las condiciones anteriores, pero cuando lo hacen, la penalización aplicada mejora frecuentemente significativamente la concordancia entre la teoría y el experimento. Estos resultados demuestran que los tipos de deformación penalizados por la invención como se describen no son, en general, características de complejos de receptores con moléculas de ligando activo.

Tabla 1. Ligandos de PDB nativos afectados significativamente (> 1 kcal/mol) por el término de penalización de torsión aromática. Score\_np es la puntuación sin penalizar y score\_p la puntuación penalizada, dG es la energía de unión libre experimental.

Sistema_Pdb	Score_np	Score_p	dG
Aff-47	-9,6	-8,4	-8,5
Cdk2_1h00	-6,3	-4,7	-6,0
Cdk2_1 g9x	-9,9	-8,9	-7,0
Jnk_2b1p	-12,6	-10,8	-11,6
Alr2_2hv5	-12,1	-11,1	-11,4
Lck_2ofv	-14,7	-12,1	-13,2
Pde5_3bjc	-9,2	-8,0	-9,5

El segundo criterio para eficacia de la invención es penalizar ligandos de bases de datos aleatorias a los que se asignan puntuaciones altamente favorables por la actual función de puntuación. En una base de datos aleatoria de 1000 compuestos, es muy poco probable que se encuentre experimentalmente un compuesto con una afinidad de unión que era más ajustada que 500 nm, o -9 kcal/mol. Por tanto, un término de penalización está mejorando la potencia discriminatoria cuando elimina compuestos con afinidades de unión tan buenas como o mejores que -9 kcal/mol sin penalizar inapropiadamente los compuestos activos. Debido a las fluctuaciones intrínsecas en la función de puntuación de 1,5-2 kcal/mol, observada anteriormente, los presentes inventores esperan, sin embargo, ver algunos compuestos que puntúan en el nivel de -9,0 kcal/mol (o un poco mejor); éstos representan compuestos activos que experimentalmente tendrían afinidades de unión en el intervalo de -7,0 a -9,0 kcal/mol, pero que alcanzarían una mejor puntuación debido a las fluctuaciones de la función de puntuación. Sin embargo, si la tasa de interacciones positivas experimentales para un cribado de 10 micromolar está en el orden de 0,5 % (típico para dianas farmacéuticamente

interesantes, aunque puede haber desviaciones significativas en cualquier dirección de este valor), entonces cabría esperar que no hubiera más de 5 compuestos de la biblioteca aleatoria que puntuaran por debajo de -9,0 kcal/mol. Por tanto, el reducir el número de dichos valores para el conjunto de receptores probados por debajo de 5, y en general reducir el número en la medida de lo posible, es una buena medida de la eficacia del término de penalización.

- 5 La Tabla 2 muestra el número de ligandos para la biblioteca aleatoria de 1000 compuestos estándar de los presentes inventores de moléculas de tipo fármaco cuyas puntuaciones son inferiores a -9 kcal/mol para varias funciones de puntuación diferentes y para 24 receptores de prueba. Esta comparación específica es significativa cuando las puntuaciones para compuestos activos son próximas a las puntuaciones experimentales para estos compuestos, por lo que en la primera columna de la Tabla 2 los presentes inventores presentan resultados obtenidos con Glide XP™ optimizado para reproducir las puntuaciones de complejos de PDB con un error promedio de ~1 kcal/mol. Los términos de penalización, tales como la presente invención, se deben entonces añadir a esta función de puntuación para mejorar el rendimiento en la discriminación de compuestos activos de inactivos. Un ejemplo de dicha molécula de señuelo discriminada se representa en la Figura 1.

- 15 **Tabla 2.** Número de ligandos señuelo con puntuaciones inferiores a -9 kcal/m para una versión de XP sin términos de penalización recientemente desarrollada (XP), XP con los términos de penalización más nuevos (newXP) y XP con la adición del término de torsión aromática de esta patente (XP\_tor). La columna cuatro tiene newXP sin el término de penalización por torsión aromática (newXP\_notor)

Sistema_Pdb	XP_	newXP	XP_tor	newXP_notor
Abl	24	15	18	20
Alr2	18	4	10	8
Jnk	7	2	4	3
Aur	2	2	2	2
Cdk2	10	3	8	5
Chk1	7	4	5	5
Dpp4	7	1	6	2
Er	4	0	2	0
Erk2	0	0	0	0
Err	3	2	2	3
FviiA	2	0	0	0
Fxa	0	0	0	0
Hivrt	11	2	8	7
Hsp90	2	0	0	1
Lck	10	3	4	7
Oppa	1	1	1	1
Pim1	46	3	28	6
Pka	0	0	0	0
Ppar	1	0	0	0
Ptp1b	1	1	1	1
Rho	5	2	2	5
Throm	2	1	1	2
Upa	0	0	0	0
P38	0	0	0	0

- 5 En la segunda columna (newXP), los presentes inventores presentan los resultados obtenidos cuando a la función de puntuación se aplican los términos de penalización descritos en el presente documento usados en combinación con otros términos de penalización. Esto da como resultado una reducción significativa en el número de señuelos que puntúan -9 y menos. Finalmente, para aislar el rendimiento específico de la invención descrito en el presente documento, la columna tres (XP\_tor) presenta los resultados de añadir el término inventado a la función de puntuación de la columna 1. El término no afecta fuertemente cada receptor, pero existe un número significativo de receptores que demuestran una mejora medible, altamente relevante. Finalmente, en la columna 4 (newXP\_tor), los presentes inventores presentan los resultados de borrar la invención de la función de puntuación de las mejores prácticas globales presentadas en la columna 3.
- 10 Tomados conjuntamente, los datos anteriores demuestran que la invención descrita en el presente documento hace una contribución sustancial a discriminar compuestos activos de inactivos en la función de puntuación empírica Glide XP. Se lograría mejora similar, en general, añadiendo el término de penalización a otras funciones de puntuación. La función de puntuación, que incluye la invención descrita aquí, da 23/24 receptores con 4 o menos ligandos de la base de datos aleatoria que puntúan por debajo de -9,0. El número total de ligandos de la base de datos que puntúan por
- 15 debajo de -9,0 kcal/mol que se eliminan por la invención (columna 1-columna 3), sumados con respecto a todos los receptores, es 57, una fracción no trivial del número total de ligandos que alcanzan esta puntuación en la función XP original (columna 1).

Otras realizaciones están dentro del alcance de las siguientes reivindicaciones.

## REIVINDICACIONES

1. Un método implemento por ordenador de la puntuación de afinidad de unión de una molécula de ligando propuesto por un receptor conocido, que es una proteína, método que usa un ordenador y datos almacenados en el ordenador, comprendiendo el método,
- 5 a. recibir, por uno o más ordenadores, datos que representan una o más moléculas de ligando propuesto a puntuar,
- b. recibir, por uno o más ordenadores, datos que representan una estructura predicha de un complejo entre un receptor y la molécula de ligando propuesto,
- c. puntuar la afinidad de unión del ligando propuesto por un método que comprende usar los datos de a. y b. en
- 10 i. análisis informático de los datos de la molécula de ligando propuesto para determinar si el ligando incluye al menos un resto conjugado  $\pi$  que tiene múltiples geometrías posibles, estando una de las geometrías caracterizada por menos deslocalización de electrones a través del resto conjugado  $\pi$  que la deslocalización de electrones que caracteriza otra de las geometrías,
- 15 ii. análisis informático de la estructura ligando-receptor predicha que incluye determinar si el ligando en la estructura de ligando-receptor adopta la geometría caracterizada por menos deslocalización,
- iii. si es así, incluir en la puntuación una penalización para la reducida deslocalización de electrones en los restos conjugados  $\pi$ , en donde la magnitud de la penalización es una cantidad que refleja la reducción de la deslocalización de electrones entre la configuración del ligando propuesto cuando se compleja con el receptor y la deslocalización de una configuración sin complejar, más deslocalizada, del ligando propuesto,
- 20 en cuya etapa c. i. incluye determinar si el ligando incluye una o más trayectorias de enlace giratorio que conectan dos restos químicos que son capaces de la significativa deslocalización de electrones  $\pi$  entre los grupos a través del enlace giratorio.
2. El método de la reivindicación 1, en que la etapa c. ii. incluye determinar la configuración del ligando en el complejo ligando-receptor y, para al menos una trayectoria de enlace giratorio identificada en la etapa c. i., determinar si la
- 25 geometría de la configuración de acoplamiento molecular reduce sustancialmente la deslocalización de electrones de una geometría de ligando energéticamente accesible en disolución o en fase gaseosa.
3. El método de la reivindicación 1, en que el resto conjugado  $\pi$  incluye al menos un enlace giratorio, en donde la rotación del enlace cambia el grado de deslocalización de electrones en el resto conjugado  $\pi$ .
4. El método de la reivindicación 1, en que el resto conjugado  $\pi$  incluye al menos un anillo aromático.
- 30 5. El método de la reivindicación 1, en que el resto conjugado  $\pi$  es más plano en la configuración caracterizada por mayor deslocalización que lo es en la configuración caracterizada por menos deslocalización.
6. El método de la reivindicación 1, en que las energías de las configuraciones del resto conjugado  $\pi$  se miden por uno o más de: un campo de fuerzas de mecánica molecular, métodos de química cuántica o desviaciones de la planaridad.
- 35 7. El método de la reivindicación 1, en que se ignoran los choques entre átomos que forman enlaces de hidrógeno internos.
8. El método de la reivindicación 1, en que el ligando incluye múltiples trayectorias de enlace giratorio conjugado y el método incluye evaluar una penalización para cada trayectoria de enlace cuando el ligando en la estructura de ligando-receptor adopta la geometría caracterizada por menos deslocalización.
- 40 9. El método de la reivindicación 1, en que se usa un método de búsqueda conformacional para investigar la energía de al menos una conformación del ligando en disolución o en fase gaseosa y para identificar una confirmación que presenta conjugación entre los restos conectados por la trayectoria de enlace giratorio.
10. El método de la reivindicación 9, en que el grado de conjugación a lo largo de la trayectoria de enlace giratorio se determina por la desviación de la planaridad de los restos conectados por la trayectoria de enlace giratorio.
- 45 11. El método de la reivindicación 9, en que se usa un campo de fuerzas de mecánica molecular para calcular las energías de las diversas conformaciones en la búsqueda conformacional.
12. El método de la reivindicación 9, en que se usan métodos de química cuántica para calcular las energías de las diversas conformaciones en la búsqueda conformacional.

13. El método de la reivindicación 9, en que se usa un modelo de interacción simplificado, tal como uno que impone distancias mínimas permitidas entre átomos de diversos tipos, y basándose en la energía de conjugación tras la desviación de la planaridad, para evaluar la energética de las diversas conformaciones en la búsqueda conformacional.
- 5 14. El método de la reivindicación 9, en que la magnitud de la penalización se determina por el grado de distorsión de la conformación del ligando en el complejo acoplado molecularmente a partir de la conformación de baja energía en disolución o en fase gaseosa.
- 10 15. El método de la reivindicación 9, en que se asigna una penalización de cero a una trayectoria de enlace giratorio conjugado si la conformación de energía más baja en disolución o en fase gaseosa no permite la sustancial conjugación entre los dos restos conectados por la trayectoria de enlace giratorio conjugado, debido a una gran desviación de la planaridad.

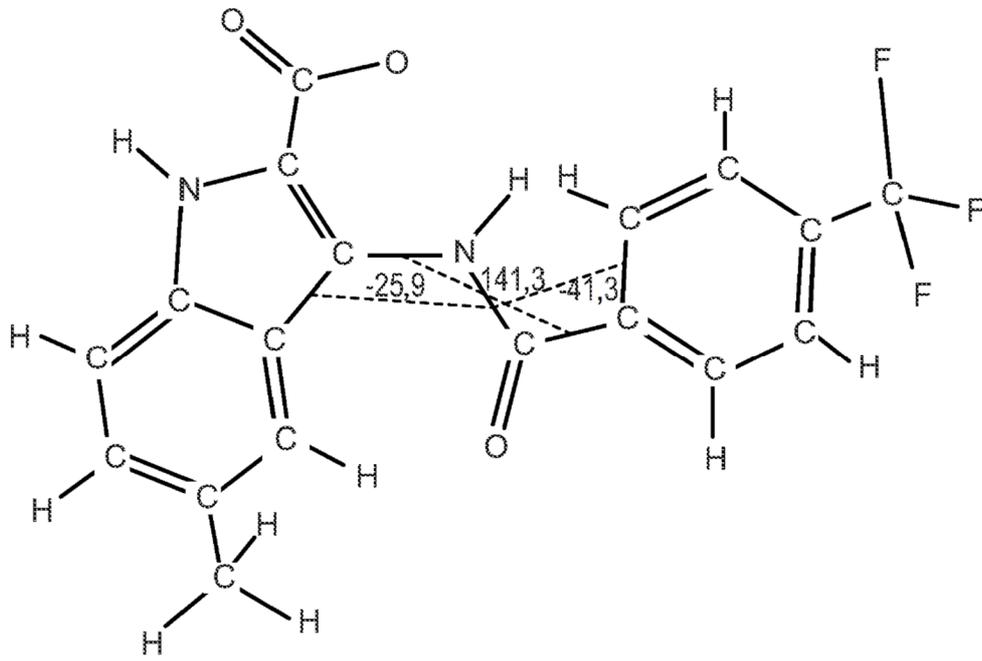


FIG. 1