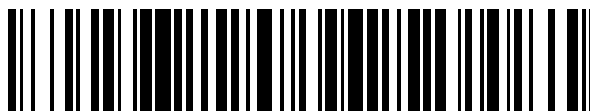


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 788 504**

51 Int. Cl.:

G16C 20/50 (2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **03.04.2012 PCT/US2012/031976**

87 Fecha y número de publicación internacional: **11.10.2012 WO12138640**

96 Fecha de presentación y número de la solicitud europea: **03.04.2012 E 12768363 (9)**

97 Fecha y número de publicación de la concesión europea: **26.02.2020 EP 2695103**

54 Título: **Función de puntuación de afinidad de enlace que incluye un término de penalización**

30 Prioridad:

04.04.2011 US 201113079725

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

21.10.2020

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 2 788 504 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

Función de puntuación de afinidad de enlace que incluye un término de penalización

5 Campo técnico

La invención se encuentra en el campo general de los procedimientos informáticos para estimar la afinidad de enlace entre un ligando y una molécula receptora, que es una proteína.

10 Antecedentes

Muchos fármacos funcionan enlazándose químicamente a receptores moleculares específicos. Los receptores moleculares son típicamente proteínas específicas (un término que incluye glicoproteínas y lipoproteínas) en un animal tal como un ser humano, y el diseño y selección de fármacos se puede facilitar estimando de manera precisa la afinidad de enlace de un fármaco a una proteína o, más generalmente, estimando la afinidad de enlace de un ligando a un receptor, donde el término receptor se utiliza para referirse a cualquier porción que se enlace específicamente al ligando.

20 Una forma de determinar la afinidad de enlace receptor-ligando utiliza la estructura molecular que resulta cuando el ligando se enlaza al receptor ("el complejo ligando-receptor"). Dichas estructuras se pueden estudiar mediante cristalografía de rayos X. El Banco de Datos de Proteínas (PDB) accesible al público contiene actualmente más de 70.000 estructuras cristalinas de rayos X, y las empresas farmacéuticas y biotecnológicas tienen un orden de magnitud de más estructuras propias. Muchas de estas estructuras han sido cristalizadas junto con moléculas pequeñas enlazadas a las mismas. El examen de dichas estructuras y el aprovechamiento de los conocimientos así
25 obtenidos para diseñar nuevos inhibidores más potentes y más específicos, se denomina diseño de fármacos basado en estructura.

El modelado computacional facilita el diseño de fármacos basado en estructura. Un aspecto del modelado detallado a continuación implica funciones de puntuación que utilizan técnicas de simulación, tales como dinámica molecular, Monte Carlo o cálculos electrostáticos continuos. Las funciones de puntuación pueden ser problemáticas cuando se requiere calcular una diferencia muy pequeña (la afinidad de enlace) 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 evalúe directamente la cantidad deseada. Este enfoque tiene la ventaja de ser extremadamente rápido, así como apto para adaptarse a los datos experimentales, donde una gran cantidad de los cuales está ahora a disposición del público. El documento US 2007/0061118 A1, de propiedad común, describe dichas funciones de puntuación.

30 Es deseable aumentar la precisión y robustez de las funciones de puntuación mediante mejoras materiales en la forma funcional que reflejen mejor la realidad física. La solicitud PCT WO/2008/141260, de propiedad común, describe mejoras destinadas a mejorar el orden de clasificación de las afinidades de enlace de una serie de ligandos enlazados a un receptor específico.

Un segundo problema importante con las funciones de puntuación es que pueden asignar mejores puntuaciones (más negativas) de afinidad de enlace a compuestos inactivos (es decir, compuestos que no se determinaría que se enlazan al receptor en un protocolo de cribado experimental típico) que a compuestos activos. Si un gran número de compuestos inactivos se clasifican por delante de compuestos activos, una función principal de acoplamiento, que es descubrir nuevos compuestos activos contra un receptor específico de una librería de compuestos muy grande (a menudo de millones de compuestos), se vuelve difícil de llevar a cabo de manera eficaz. Por lo tanto, reducir sustancialmente el número de "falsos positivos" (clasificación de compuestos inactivos como activos) mejoraría considerablemente el rendimiento de la función de puntuación.

35 Krammer A et al. describen en *Journal of Molecular Graphics and Modelling*, vol. 23, en las páginas 395-407, funciones empíricas de puntuación que consisten en tres términos distintos que describen la interacción de van der Waals, la atracción polar entre el ligando y la proteína, y la penalización de desolvatación atribuida al enlace de los átomos del ligando polar a la proteína, y viceversa.

Young T et al. describen en PNAS, vol. 104, n.º 3, 16 de enero de 2017, en las páginas 808-813, motivos para el reconocimiento molecular aprovechando la envoltura hidrofóbica en el enlace proteína-ligando.

60 Schnecke, V. et al. (2000), *Perspectives in Drug Discovery and Design*, vol. 20, páginas 171-190, describe un procedimiento para puntuar la afinidad de enlace de una molécula de ligando propuesta para una molécula receptora usando un ordenador y bases de datos informáticas.

Resumen

65 La presente invención se define mediante la reivindicación 1 independiente. Las reivindicaciones dependientes

ilustran otras formas de realización de la invención.

Los inventores han descubierto un procedimiento para puntuar la afinidad de enlace de una molécula de ligando propuesta para una molécula receptora utilizando un ordenador y bases de datos informáticas, que tiene en cuenta el aumento de energía necesaria cuando el acoplamiento separa moléculas de agua que están localizadas en los sitios de hidratación de un ligando. El procedimiento utiliza datos almacenados por ordenador que representan una estructura de ligando-receptor predicha (preferentemente una que está validada). El análisis de puntuación informatizado incluye determinar si el receptor incluye uno o más sitios de hidratación ocupados por agua localizada y, de ser así, evaluar una penalización si el acoplamiento del ligando en el receptor da como resultado una interacción desfavorable del ligando con una molécula de agua localizada que permanece en el sitio de hidratación del receptor (es decir, después del acoplamiento). La evaluación de la penalización para un caso dado se realiza a través de un análisis empírico, donde las penalizaciones específicas asignadas a un conjunto particular de interacciones del ligando con moléculas de agua localizadas que permanecen en los sitios de hidratación del receptor se computan previamente y vuelven a aplicarse de manera eficaz. Además, dicha evaluación de penalización solo requiere una única configuración del complejo receptor-ligando, y es significativamente más rápida (< 1/2 hora) que los enfoques alternativos basados en física pura, que intentan describir completamente toda la gama de la física manifestada por el ensamblaje del complejo ligando-receptor.

El término "agua localizada" incluye aguas que están localizadas o casi localizadas, es decir, cualquier agua que ocupa un sitio de hidratación donde la densidad de ocupación de agua local en el sitio de hidratación es sustancialmente mayor que la densidad numérica de fluido a granel. La presencia de dicha agua localizada en los sitios de hidratación se puede determinar realizando y analizando una trayectoria de dinámica molecular, o a partir de datos experimentales tales como densidades de agua obtenidas a partir de experimentos cristalográficos de rayos X. En contraste con estas penalizaciones basadas en agua localizada, las interacciones con moléculas de agua que residen en regiones de densidad aparente no reciben una penalización significativa. Por ejemplo, en formas de realización preferidas, se aplica una penalización a la puntuación de afinidad de enlace del ligando solo si se determina que la energía libre de desplazamiento de una molécula de agua desde el sitio del receptor hasta aguas a granel es más favorable que un valor de corte específico (por ejemplo, 0 kcal/mol) de la energía obtenida al desplazar el agua desde el sitio del receptor hasta el agua a granel. Debe observarse que se podrían utilizar otros estados de referencia para determinar la energía libre (y, por lo tanto, también el valor particular del corte) del agua en los sitios de hidratación para clasificar el agua en los sitios de hidratación. Por ejemplo, se podría utilizar la transferencia del agua desde el sitio de hidratación hasta la fase gaseosa, donde el corte específico empleado se variaría para incluir implícitamente la transferencia del agua desde la fase gaseosa hasta la fase líquida. En ese caso, el valor de corte se ajustaría para reflejar esta variación. Otros de dichos posibles estados de referencia podrían incluir el uso del desplazamiento del agua desde el sitio de hidratación hasta la fase cristalina, o el desplazamiento del agua desde el sitio de hidratación hasta cualquier otro entorno bien definido termodinámicamente. De forma alternativa, la favorabilidad energética de la molécula de agua en el sitio localizado se estima a partir de características del sitio, tales como la hidrofobicidad del sitio, interacciones específicas con grupos de receptor, la utilización de diversos tipos de cálculos, que incluyen, pero sin limitarse a, simulaciones de dinámica molecular, simulaciones Monte Carlo, enumeración exhaustiva de geometrías de la molécula de agua dentro del sitio y optimización (mediante, por ejemplo, minimización de energía) de la posición de la molécula de agua dentro del sitio. Los sitios de agua localizados se pueden identificar mediante análisis de solvatación de sitios activos, tal como el análisis WaterMap, para asignar energías libres de desplazamiento que se describen a continuación.

En formas de realización preferidas, los sitios de hidratación de receptor se determinan mediante muestreo informatizado de sitios para hidratación, seguido de asignación informatizada de ocupaciones de agua para los diversos sitios muestreados. Esto incluye el caso especial en el que las ocupaciones se asignaron discretamente como 0 o 1. Se evalúa una penalización para compuestos cuando el agua localizada dentro de un sitio de hidratación interactúa con un átomo de ligando hidrofóbico o cuando el agua localizada dentro de un sitio de hidratación interactúa con un átomo de ligando hidrofóbico, por ejemplo, a una distancia de 5 Å o menos. No se evalúa ninguna penalización para la interacción con un átomo de ligando hidrofóbico que tiene un átomo de ligando positivo a 4 átomos o más cerca del mismo (en función de la conectividad de enlace de los átomos en el ligando). No se evalúa ninguna penalización: a) cuando el átomo de ligando hidrofóbico es Cl o Br; b) cuando el átomo de ligando hidrofóbico no es parte de un grupo -CH₃ de ligando o no es un átomo de ligando a menos de 3 átomos de un grupo -CH₃, y si el agua localizada dentro de un sitio de hidratación tiene acceso al agua circundante debido a que bordea el borde de un bolsillo hidrofóbico; c) cuando el agua localizada dentro de un sitio de hidratación está a menos de 5 Å de un átomo de ligando con carga positiva independientemente de cualquier otra interacción; d) cuando el agua localizada dentro de un sitio de hidratación tiene enlaces de hidrógeno con el ligando, independientemente de cualquier otra interacción. Preferentemente, se determina que los sitios de receptor con agua localizada son aguas que se predice que tienen una ocupación en el marco de referencia de la proteína significativamente (por ejemplo, 1,5X) por encima de la ocupación de agua a granel. No se evalúa ninguna penalización si el agua localizada dentro de un sitio de hidratación tiene enlaces de hidrógeno al ligando, independientemente de cualquier otra interacción. En este caso se considera que el enlace de hidrógeno incluye interacciones C-H...O cuando el C-H es un grupo C-H aromático, con geometrías adecuadas, como se analiza en esta solicitud posteriormente, así como enlaces de hidrógeno más tradicionales. La puntuación puede incluir estimar la favorabilidad energética de la molécula de agua

en el sitio de hidratación a partir de la hidrofobicidad del sitio o a partir de interacciones específicas con grupos de receptor. La estimación de favorabilidad energética puede incluir cálculos tales como simulaciones de dinámica molecular, simulaciones Monte Carlo, enumeración exhaustiva de geometrías de la molécula de agua dentro del sitio y optimización (mediante, por ejemplo, minimización de energía) de la posición de la molécula de agua dentro del sitio. Las energías libres de desplazamiento de una molécula de agua desde el sitio de receptor hasta el agua a granel se pueden determinar mediante simulación de dinámica molecular seguida de la aplicación de solvatación no homogénea modificada para asignar energías libres de desplazamiento.

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

Descripción de los dibujos

El archivo de esta patente contiene al menos un dibujo a color. La Oficina de Patentes y Marcas proporcionará copias de esta patente con dibujos a color previa solicitud y pago de la tarifa necesaria.

Figura 1. Aquí se ilustra un ligando activo 2iq0 que forma contactos desfavorables con moléculas de agua 16 y 18. Este ligando sufre una penalización calculada de 1,0 kcal/mol, con lo que su afinidad de enlace estimada concuerda más estrechamente con el valor medido experimentalmente.

Figura 2. Aquí se ilustra un ligando activo 1r06 que forma contactos desfavorables con una molécula de agua 32. Este ligando sufre una penalización calculada de 2,0 kcal/mol, con lo que su afinidad de enlace estimada concuerda más estrechamente con el valor medido experimentalmente.

Figura 3. Aquí se ilustra un ligando señuelo 790331 que forma contactos desfavorables con una molécula de agua 53. Este ligando sufre una penalización calculada de 2,0 kcal/mol, con lo que su afinidad de enlace estimada es de -6,5 kcal/mol, que es más coherente con las afinidades de enlace estadísticamente esperadas de dichas moléculas señuelo en un cribado aleatorio.

Figura 4. La molécula activa 2o63 que se ilustra aquí evita la penalización sufrida por el señuelo 790331 al orientar un grupo hidroxilo para que interactúe con una molécula de agua 53.

Descripción detallada

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

Un programa de acoplamiento típicamente lleva a cabo dos tareas distintas para modelar el enlace receptor-ligando. En primer lugar, se predice una estructura de un complejo receptor-ligando acoplando el ligando en la estructura de receptor. Cuando este protocolo no produce una estructura precisa del complejo proteína-ligando, se requiere el uso de una estructura diferente del receptor como punto de partida. El problema de construir estructuras receptoras alternativas que se modifican para aceptar ligandos que requieren un cambio sustancial en la conformación del receptor ("ajuste inducido") es muy importante.

Diversos programas de predicción de estructuras adecuadas receptor-ligando son bien conocidos y se pueden utilizar como punto de partida, y los expertos en la técnica entenderán que ninguna metodología particular es crítica. 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 razonable coherencia con los datos experimentales, la segunda tarea del programa de acoplamiento es calcular una afinidad de enlace receptor-ligando, dada como entrada la estructura acoplada. Una función matemática empleada para calcular la afinidad de enlace (o una contribución a la misma) se denomina "función de puntuación". Las mejoras de dichas funciones de puntuación para calcular la afinidad de enlace receptor-ligado son el objeto de esta invención.

Las estimaciones de afinidades de enlace receptor-ligando descritas a continuación son aplicables cuando la estructura del complejo receptor-ligando está representada por un modelo estructural adecuado. Existen varias formas de caracterizar la calidad de los modelos estructurales de los complejos receptor-ligando, siempre y cuando el modelo concuerde adecuadamente con los datos experimentales. Podrían utilizarse medidas de concordancia estructural tales como RMSD, DME o puntuación SIFt, entre otras. Las puntuaciones exactas se obtienen típicamente, pero no exclusivamente, usando: i) una pequeña desviación cuadrática media raíz (RMSD) con respecto a la estructura experimental (típicamente menos de 2 Angstroms, aunque el valor requerido variará dependiendo de los detalles del complejo), ii) el reconocimiento de la formación de sustancialmente todos los enlaces de hidrógeno vistos en el complejo experimental, iii) la colocación adecuada de sustancialmente todos los grupos hidrofóbicos en las cavidades receptoras correctas, y iv) la ausencia de choques estructurales o

electrostáticos incorrectos que podrían dar lugar a la asignación de términos de penalización sustancialmente incorrectos. Dado que se está considerando la afinidad de enlace relativa de ligandos a un receptor dado, un desfase constante, generado en muchos casos por la energía de reorganización del sitio activo del receptor para acomodar el ligando, no tiene ningún efecto en las aplicaciones prácticas.

5 En muchos casos (aunque no en todos), el receptor puede adoptar más de una conformación fundamentalmente diferente en respuesta a una clase de ligandos (por ejemplo, ligantes *DFG-in* y *DFG-out* a p38 MAP quinasa); para comparar las energías libres de enlace en tales casos, se pueden requerir diferentes parámetros de reorganización de núcleo para las diferentes conformaciones de receptor. Cuando el cálculo de estos parámetros no resulta práctico de acuerdo con principios fundamentales, se tratan como parámetros ajustables específicos de receptor. Sin embargo, otros parámetros están contenidos en un modelo global que no es específico del receptor o incluso específico para una clase particular de receptores.

15 La presente invención se centra en mejorar la capacidad de la función de puntuación para discriminar compuestos inactivos y activos con respecto a un receptor dado. Si bien los compuestos activos para una amplia gama de receptores se encuentran fácilmente en el Banco de Datos de Proteínas (PDB), los siguientes datos, obtenidos por los inventores, que validan la invención también requerían compuestos inactivos conocidos evaluados o un modelo aproximado, pero suficientemente preciso, para una librería de dichos compuestos. El acceso a compuestos inactivos conocidos a partir de datos disponibles públicamente es un reto. Por lo tanto, los inventores han ideado un protocolo alternativo para representar un modelo aproximado, pero suficientemente preciso, de una librería de compuestos inactivos que se basa en el uso de una librería aleatoria de 1000 compuestos similares a fármacos. Estos compuestos se acoplan en una conformación del receptor y se obtienen afinidades de enlace predichas para cada compuesto usando varias mejoras propuestas en la función de puntuación.

25 La función de puntuación de los inventores se calibra para que los compuestos activos consigan puntuaciones que típicamente se acerquen a las afinidades de enlace experimentales, con una desviación típica de aproximadamente 1 unidad logarítmica de afinidad de enlace ($\sim 1,5$ kcal/mol). Generalmente se considera que un "acierto" experimental en un cribado aleatorio en la industria farmacéutica tiene una afinidad de enlace de $\sim 7,0$ kcal/mol o más (concentración de ~ 10 μ m). Dadas las fluctuaciones intrínsecas en la función de puntuación de 1,5 - 2,0 kcal/mol, se establece el umbral computacional para estimar aciertos en una puntuación de -9,0 kcal/mol; por lo tanto, se predice que cualquier compuesto en la librería aleatoria de 1000 compuestos que puntúe -9,0 o menos será un acierto. Otros compuestos en el intervalo de -7,0 a - 9,0 también se predicen como posibles aciertos, pero tales puntuaciones también pueden deberse al ruido en la función de puntuación que no pueden reducirse adicionalmente en la actualidad (en parte debido a las limitaciones en los datos experimentales disponibles al público). Por lo tanto, los esfuerzos se dirigen en desarrollar mejoras para garantizar que el número de compuestos aleatorios con una puntuación más negativa que -9,0 sea compatible con las tasas de aciertos experimentales para librerías aleatorias de compuestos similares a fármacos.

40 La tasa de aciertos para cribados experimentales variará dependiendo del receptor, pero una tasa de aciertos "promedio" ilustrativa es del 0,5%, o 5 compuestos de 1000. Por lo tanto, si sustancialmente más de 5 compuestos consiguen una puntuación de -9,0 kcal/mol, se asume que los compuestos inactivos están recibiendo puntuaciones que son demasiado favorables. En general, la reducción del número de compuestos de librería aleatoria con una puntuación más negativa que -9,0 se toma como una mejora de la función de puntuación, ya que en ausencia de fluctuaciones, cabría esperar incluso menos (0-2) compuestos en este intervalo.

45 En tal situación, la función de puntuación puede mejorarse añadiendo términos de "penalización" positivos que reducen la magnitud de la afinidad de enlace predicha. Dichos términos representan procesos físicos que hacen que el enlace sea menos favorable. Un ejemplo de un proceso de este tipo sería una desolvatación en la que un grupo polar de la proteína o ligando está bloqueado por grupos no polares del ligando y pierde acceso al agua. Esto da como resultado una gran pérdida de energía libre, haciendo que el compuesto esté inactivo. Si la pérdida de energía libre es suficientemente grande, entonces tales penalizaciones sólo se observarán rara vez (si es que se observan) en complejos con compuestos activos. Esto significa que los nuevos términos deben obtenerse examinando las estructuras producidas cuando la librería aleatoria está acoplada en el receptor.

55 La invención descrita en la esta solicitud utiliza una descripción de estructura de agua localizada (como se describió anteriormente) en el sitio activo de receptor. Una forma, pero no la única, de hacerlo es utilizar la tecnología denominada "WaterMap" (documento US 7.756.674). WaterMap utiliza simulaciones de dinámica molecular para introducir agua en varias ubicaciones en el sitio activo y asigna entalpías y entropías estimadas que se obtendrían si el agua se desplazara hacia un solvente a granel por el ligando. Las entalpías y entropías se estiman a través de una versión modificada de la teoría de la solvatación no homogénea. Se podrían utilizar otras técnicas para introducir aguas localizadas y estimar sus entropías y entalpías. La invención se puede aplicar siempre que se especifiquen las ubicaciones de las aguas localizadas. Las energías libres de desplazamiento pueden dar lugar a una mejor precisión, pero se puede obtener un beneficio significativo de la invención siempre que se puedan enumerar posiciones de agua localizada. Dichas posiciones se pueden obtener no solo a partir de procedimientos computacionales, sino también a partir de datos experimentales, tales como cristalografía de rayos X. La cristalografía determina algunas posiciones de agua localizada, pero no otras (típicamente dependiendo del grado

de localización). Se pretende que esta invención se aplique independientemente de cómo se determine la distribución de agua localizada. Otras formas de establecer la existencia de agua localizada incluyen: análisis apropiado de simulaciones Monte Carlo explícitamente solvatadas, aplicaciones de teorías tipo RISM o experimentos de NMR, entre otras técnicas.

5 El componente novedoso de la invención es el desarrollo de un algoritmo matemático específico que detecta en qué condiciones se debe aplicar un término de penalización. Los inventores han desarrollado este algoritmo basándose en una combinación de los principios básicos de química física, expuestos anteriormente, y en una optimización empírica en la que el algoritmo se ha diseñado para evitar penalizar inapropiadamente (es decir, aplicar una penalización que haría que la puntuación del compuesto coincidiera menos con los datos experimentales de afinidad de enlace que si no se aplicara la penalización) complejos conocidos de compuestos activos, tomados del Banco de Datos de Proteínas (PDB), mientras que al mismo tiempo se maximiza el número de compuestos de base de datos aleatoria de puntuación favorable que se penalizan. Los detalles del protocolo de optimización se describen a continuación, después de una descripción del algoritmo en sí.

15 La siguiente descripción detallada de la invención se utiliza para evaluar la asignación de un término de penalización a cada agua inestable generado por un cálculo de WaterMap en el sitio activo de receptor en ausencia de cualquier ligando. Un agua "inestable" de WaterMap se define en las mejores prácticas actuales de los inventores como un agua localizada dentro de un sitio de hidratación (o un "conglomerado de agua") con una energía libre estimada mayor (más positiva, por lo tanto más inestable en el sitio activo de proteína de lo que sería en el fluido a granel) que 1,0 kcal/mol. Podrían utilizarse otros límites y definiciones de lo que constituye un agua inestable.

25 La idea física básica de la invención es que un conglomerado de agua inestable de WaterMap se volverá aún más inestable si ese agua hace contacto hidrofóbico con uno o más átomos de ligando. Las definiciones estándar de lo que constituye un átomo hidrofóbico se utilizan en las mejores prácticas actuales de los inventores (véase el anexo I); también podrían utilizarse otras definiciones. Una molécula de agua no puede formar enlaces de hidrógeno con un átomo de ligando hidrofóbico. Agua en una solución a granel típicamente será capaz de alterar su distribución de geometrías para evadir el átomo o átomos hidrofóbicos del ligando y crear enlaces de hidrógeno con otras aguas a granel, a expensas de un cierto grado de entropía. Sin embargo, un agua considerada inestable por un cálculo de WaterMap está, por definición, localizada hasta cierto punto en un sitio activo confinado, y ya está siendo forzada a evadir las características de la proteína (que es lo que está causando que sea inestable). Por lo tanto, los contactos de esta agua con átomos de ligando hidrofóbicos causarán un mayor aumento desfavorable de energía libre que un contacto similar en solución a granel, y este diferencial es lo que se está modelando aproximadamente mediante la adición de un término de penalización a la afinidad de enlace.

35 La cantidad de penalización para cualquier molécula de agua dada dependerá de una serie de factores; la distancia específica del átomo de ligando hidrofóbico (más allá de cierta distancia no hay contacto efectivo y no se aplica ninguna penalización), la cantidad de contactos, si el ligando crea un enlace de hidrógeno con el agua así como un contacto hidrofóbico (en cuyo caso el agua ha compensado eficazmente el contacto hidrofóbico y no es necesario aplicar ninguna penalización) y otros factores detallados a continuación. Otra pregunta importante es qué cuenta como un enlace de hidrógeno. Un estudio empírico de la estructura ligando-agua en complejos activos indica concluyentemente que un enlace de hidrógeno C-H---O no debe recibir penalizaciones; sin embargo, tal interacción no es suficiente para quitar la penalización si hay otros contactos hidrofóbicos desfavorables. Por último, además de un posible término de penalización para cada agua inestable de WaterMap, se define un término de penalización general para el ligando en función de la evaluación de los contactos agua-ligando para todo el ligando. La idea subyacente a este término es que a medida que aumenta el número total de contactos, la perturbación de la estructura general del agua en el sitio activo también aumenta, y se alcanza un umbral cuando hay consecuencias desfavorables que van más allá de aguas individuales. Los detalles específicos de este término se describen a continuación y, nuevamente, se han desarrollado probando el término contra todo el conjunto de complejos del PDB con ligandos activos.

Los detalles específicos de la invención son los siguientes:

55 1) para cada agua de WaterMap calcular la cantidad de contactos de esta agua con átomos de ligando hidrofóbicos a menos de 5 angstroms del átomo de agua.

Omitir átomos de ligando que

- 60 a) estén a menos de una distancia 1-4 de un átomo de ligando positivo
 b) sean Cl o Br
 c) sean grupos CH₃ de ligando y átomos de ligando a menos de una distancia 1-3 de un grupo CH₃ si el agua de WaterMap está en el borde del bolsillo cerca de la solución. La solvatación del agua se determina usando el algoritmo de envasado de agua de Glide (no de WaterMap).
 65 d) Si el agua tiene 1 contacto de ligando usando las reglas anteriores y esta agua está a menos de 5 Å de un átomo de ligando con carga positiva, establecer el contacto de ligando a cero (negar el 1 contacto).

2) para cada agua de WaterMap calcular la cantidad de contactos de esta agua con átomos de proteína fóbica a menos de 5 angstroms del agua.

5 a) calcular la cantidad de contactos que el agua hace con átomos de proteína polares (O de C=O o N de NH) que no están unidos por hidrógeno a la proteína y están orientados hacia el agua. La orientación debe ser tal que el ángulo realizado por la terna C=O - agua o NH-agua sea mayor de 108 grados.

10 3) para cada agua de WaterMap calcular un número fraccionado que representa la medida en que el agua se desplaza mediante contactos con átomos pesados de ligando (que no son de hidrógeno). La fracción de desplazamiento se calcula de la siguiente manera:

15 a) para cada átomo de ligando i calcular la distancia d al agua de WaterMap.
 b) calcular la suma r del radio Vdw de ligando r_i y un radio fijo Vdw de agua de 2,2Å, $r=R_i + 2,2$.
 c) Si $d < r$ entonces la contribución vac_i que el átomo de ligando i hace a la evacuación de agua i es $vac_i = 1,0 - d/r$ $r \leq d$, $vac_i = 0,0$ $r > d$.
 d) la fracción total en que el agua i se elimina es la suma sobre tal v_i con la suma limitada a un máximo de 1,0.

20 4) calcular la cantidad de enlaces de hidrógeno que cada agua de WaterMap crea con el ligando. Los átomos de ligando unidos mediante H a aguas son átomos donantes y aceptores, además de los átomos F, Cl, Br, I de ligando. La distancia (d) a átomos de ligando de agua debe satisfacer:

25 a) átomos distintos de Cl $d \leq 3,8$ Å.
 b) átomos Cl $d \leq 5.2A$.
 Para átomos donantes de ligando o aceptores con una conexión (X-H), el ángulo de X-H con respecto al agua debe ser mayor que 130,0 grados para átomos no halógenos y 100 grados para átomos halógenos.
 Para átomos aceptores de ligando con dos conexiones se suman los dos vectores formados por los dos átomos conectados al aceptor, y el ángulo entre este vector de suma y el vector que conecta el aceptor y la molécula de agua debe ser mayor de 65 grados.

5) Asignar una penalización a cada agua de WaterMap que cumpla con los siguientes criterios:

35 a) tiene 1 o más contactos de ligando,
 b) es inestable en 1 kcal/mol o más,
 c) no tiene enlaces de hidrógeno al ligando.
 d) tiene una fracción de desplazamiento (véase 3) inferior al 60%.
 e) NO cumple ninguna de las siguientes condiciones

40 e1) si el agua tiene 3 o menos contactos de ligando y dicha agua tiene 2 o más contactos (a menos de 4,5 Å) con aguas vecinas que satisfacen lo siguiente:
 e1a) el vecino debe ser inestable en 0,5 kcal o más (comprobar esto, podría ser estable)
 e1b) el vecino debe ser evacuado en menos del 10% por contactos de ligando
 e1c) el vecino tiene 2 o menos contactos de ligando y menos de 5 contactos de proteína
 e1d) el vecino tiene menos de 4 enlaces de hidrógeno al ligando.

45 e2) si el agua tiene 3 o menos contactos de ligando y dicha agua está a menos de 5 Å de un átomo de anillo de ligando que contiene una carga positiva en el anillo.
 e3) si el agua tiene 1 o más contactos inferiores a 1,5 Å con átomos de proteína pesados.
 e4) tiene 10 o más contactos inferiores a 4 Å con átomos de proteína pesados (véase 2) y tiene 1 o más enlaces de hidrógeno con el ligando y 4 o más contactos con átomos de proteína polares (véase 2a).

50 e5) tiene 2 o más contactos con la proteína a menos de 3 Å y un contacto de ligando.
 e6) tiene entre 3 o más contactos a menos de 4 Å de átomos de proteína pesados y satisface estas condiciones:
 e6a) tiene menos de 10 contactos de proteína fóbica (2) y menos de 2 contactos de proteína polares (2a).
 e6b) si hay 1 contacto de proteína polar: debe tener menos de 8 contactos de proteína fóbica.
 e6c) tiene menos de 3 contactos de átomos pesado a menos de 3,5 Å.
 e7) tiene un enlace de hidrógeno aromático con el ligando definido por: un átomo de anillo aromático de ligando (C o N) a menos de 4,7 Å del agua y con el ángulo entre el vector de átomo de ligando aromático de agua y el vector de "par solitario en plano" del átomo de ligando es menor que 70 grados.

60 e8) el agua está a menos de 3,7 Å de un átomo de proteína con carga.
 e9) el agua está a menos de 5 Å de un átomo de proteína con carga y a menos de 5,5 Å de un átomo de ligando polar (O N).
 si el agua cumple con las condiciones (a-e9), asignar una penalización wpen al agua utilizando el

65

siguiente programa.

Si hay 1 contacto de ligando fóbico y 3 o más contactos de proteína fóbica y una fracción de eliminación inferior a 0,1, $w_{pen}=1,0$ kcal/m.

Si hay 2 contactos de ligando fóbicos, $w_{pen} = 2,0$ kcal/m.

Si hay 3 contactos de ligando fóbicos, $w_{pen} = 3,0$ kcal/m.

Si hay 4 contactos de ligando fóbicos, $w_{pen} = 4,0$ kcal/m.

6) Asignar una penalización w_{pen2} en función del número total de contactos de agua $waterwmap_ligando_fóbico$ $ilpc = \sum_i ilp(i)$, donde $ilp(i)$ es el número de contactos de ligando fóbicos para el agua i . Si el agua i cumple cualquiera de los siguientes criterios 6a-6d, entonces su contribución al total de $ilpc$ en esta suma se ignora (i se elimina de \sum_i).

a) el agua i tiene un contacto de ligando fóbico y 2 o más contactos de átomos de proteína pesados a menos de 3,5 Å.

b) el agua i no entra en contacto con ninguna otra agua a menos de 3, 5 Å que tenga una fracción de eliminación (3) superior al 70%.

c) el agua i tiene una penalización distinta de cero desde la etapa 5.

d) el agua i está a menos de 4 Å de un átomo de proteína con carga.

La función para w_{pen2} con respecto al total de contactos de ligando fóbicos $ilpc$ es:

si $ilpc = 8$ $w_{pen2} = 0,5$ kcal/m

si $ilpc = 9$ $w_{pen2} = 1,0$ kcal/m

si $ilpc = 10$ $w_{pen2} = 1,5$ kcal/m

si $ilpc = 11$ $w_{pen2} = 2,5$ kcal/m

si $ilpc = 12$ $w_{pen2} = 3,5$ kcal/m

si $ilpc \geq 13$ $w_{pen2} = 4,0$ kcal/m.

Finalmente, la penalización total para contactos hidrofóbicos es w_{pen} (etapa 5) + w_{pen2} (etapa 6).

La eficacia de la invención anterior se demuestra de la siguiente manera. En primer lugar, se ha ensamblado un conjunto de pruebas de 622 complejos proteína-ligando de compuestos activos. Como un control para evaluar el procedimiento, los siguientes ejemplos implican estructuras cristalinas conocidas disponibles en el Banco de Datos de Proteínas ("el PDB") accesible al público de *Research Collaborative for Structural Bioinformatics* (investigación colaborativa de bioinformática estructural). Para llevar a cabo la optimización se utilizan posiciones acopladas con Glide XP, una función de puntuación descrita, por ejemplo, en la patente n.º US 2007/0061118 A1, filtrando el (muy pequeño) número de casos para los que el acoplamiento automático produce estructuras inadecuadas. Mediante el uso de estructuras acopladas, en lugar de las propias estructuras cristalinas, en el proceso de optimización de los inventores, se aumenta el realismo del modelo, y también se permite que requiera el uso de Glide XP.

Las estructuras del PDB se pueden ver como un conjunto de entrenamiento grande y diverso para la función de puntuación. Las pruebas de la función de puntuación en condiciones similares se pueden realizar mediante acoplamiento para corregir pequeños errores geométricos en las estructuras cristalinas (por ejemplo, en distancias de enlace de hidrógeno) que pueden ser cruciales para asignar adecuadamente puntuaciones a estos términos. La invención no requiere el uso de ligandos para los cuales se conocen estructuras cristalinas, ni requiere el uso de Glide XP.

Las estructuras del PDB se pueden ver como un conjunto de entrenamiento grande y diverso para la función de puntuación. Las empresas farmacéuticas y biotecnológicas pueden realizar pruebas de la función de puntuación en condiciones similares, utilizando conjuntos de datos propios en los que se disponga de estructuras cristalinas. Al llevar a cabo estas pruebas, no es necesario publicar las estructuras o incluso divulgar el nombre del receptor; simplemente se pueden realizar los cálculos y notificar la capacidad de ordenar los compuestos como un coeficiente de correlación.

Estos complejos proporcionan estructuras razonablemente precisas cuando el ligando se vuelve a acoplar en su receptor nativo (la RMSD máxima es de 3,5 Å) y sus puntuaciones, utilizando la versión más reciente de la función de puntuación Glide XP, están en promedio dentro de ~ 1 kcal/mol de la afinidad de enlace experimental. Por lo tanto, la función de puntuación en ausencia del término que constituye la invención funciona bien para complejos de compuestos activos tomados del PDB.

Se ha añadido el nuevo término a la función de puntuación Glide XP™ y se han vuelto a ejecutar los cálculos activos con él en su lugar. En la Tabla 1 hay una lista de complejos de PDB que se ven afectados por este término, la afinidad de enlace experimental de cada complejo y la afinidad de enlace calculada con y sin el nuevo término. Se puede observar que hay muy pocos casos del total de 662 complejos en los que los complejos activos satisfagan las condiciones anteriores. En muchos de estos casos, el término de penalización mejora la concordancia de la afinidad de enlace predicha del complejo activo con el experimento. En un pequeño número de casos, la concordancia empeora pero, en general, el efecto del término de penalización promediado con respecto a los casos de prueba

presentados a continuación es favorable (hay 9 casos favorables y sólo 3 casos desfavorables). Estos resultados demuestran que la invención tal como se describe no es generalmente una característica de complejos de receptores con complejos activos, y cuando lo es, mejora la concordancia de la afinidad de enlace predicha con el experimento muchas más veces de las que la empeora. Ejemplos de la aplicación del término de penalización también se ilustran en las figuras 1 a 4.

El segundo criterio para la eficacia de la invención es penalizar ligandos de base de datos aleatorios a los que se asignan puntuaciones altamente favorables mediante la función de puntuación actual. En una base de datos aleatoria de 1000 compuestos, es muy poco probable que experimentalmente se encuentre un compuesto con una afinidad de enlace que sea más estrecha que 500 nm, o -9 kcal/mol. Por lo tanto, un término de penalización mejora el poder discriminatorio cuando elimina compuestos con afinidades de enlace tan buenas o mejores que -9 kcal/mol sin penalizar inadecuadamente ningún compuesto activo. A pesar de las fluctuaciones intrínsecas en la función de puntuación de 1,5-2 kcal/mol, señaladas anteriormente, cabe esperar, sin embargo, ver algunos compuestos puntuar en el nivel de -9,0 kcal/mol (o un poco mejor); estos representan compuestos activos que experimentalmente tendrían afinidades de enlace en el intervalo de -7,0 a -9,0, pero que logran una mejor puntuación debido a las fluctuaciones en la función de puntuación. Sin embargo, si la tasa de aciertos experimental para un cribado micromolar de 10 está en el orden del 0,5% (típico para dianas interesantes desde el punto de vista farmacéutico, aunque puede haber desviaciones significativas en cualquier dirección de este valor), entonces se esperaría que no hubiese más de 5 compuestos de la librería aleatoria que puntúen por debajo de -9,0 kcal/mol. Por lo tanto, el éxito en reducir el número de dichos valores para el conjunto de receptores probados por debajo de 5 y, en general, reducir el número tanto como sea posible, es una buena medida de la eficacia del término de penalización.

La Tabla 2 muestra la cantidad de ligandos de la librería aleatoria estándar de 1000 compuestos, de los inventores, de moléculas similares a fármacos 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 solo si las puntuaciones para los compuestos activos son cercanas a las puntuaciones experimentales para estos compuestos, por lo que en la primera columna de la Tabla 1 se muestran los resultados obtenidos con una versión de Glide XP que se ha optimizado para reproducir las puntuaciones de los complejos de PDB con un error promedio de ~ 1 kcal/mol. Los términos de penalización, tales como los utilizados en la presente invención, deben añadirse a esta función de puntuación para mejorar el rendimiento en la distinción de compuestos activos de inactivos.

En la segunda columna, se muestran los resultados obtenidos cuando todos los términos de penalización novedosos desarrollados recientemente por los inventores, incluidos los utilizados en la presente invención, se incluyen en la función de puntuación. Esto da como resultado una reducción significativa en el número de señuelos que puntúan -9 o menos. Finalmente, para aislar el rendimiento específico de la invención descrita en el presente documento, la columna tres muestra resultados de la adición del término inventado a la función de puntuación de la columna 1. El término no afecta considerablemente a todos los receptores, pero hay un número significativo de receptores que demuestran una mejora medible y altamente relevante. Por último, en la columna 4 se muestran los resultados de no aplicar la invención en la función de puntuación general de mejores prácticas presentada en la columna 3. Estos datos demuestran que la invención es un componente importante del procedimiento general de mejores prácticas; sin ella, el rendimiento se degrada en algunos receptores.

En conjunto, los datos anteriores demuestran inequívocamente que la invención descrita en el presente documento hace una contribución sustancial a la distinción de compuestos activos e inactivos en la función de puntuación empírica Glide XP™. Se observaría una mejora similar en cualquier función de puntuación que a su vez careciera de un término de penalización comparable; que se sepa, las funciones de puntuación publicadas existentes no contienen tal término en la actualidad. La versión completa de la nueva función de puntuación, que incluye la invención descrita en el presente documento, produce receptores 23/24 receptores con 4 o menos ligandos de base de datos aleatoria con una puntuación inferior a -9,0. La cantidad total de ligandos de base de datos con una puntuación inferior a -9,0 que se eliminan mediante la invención (columna 1 - columna 3), sumada a todos los receptores, es 34, una fracción no trivial de la cantidad total de ligandos que consiguen esta puntuación en la función XP original (columna 1).

Tabla 1. Dos ligandos pdb nativos afectados por los términos de contacto hidrofóbicos. score_np es la puntuación no penalizada y score_p la puntuación penalizada, dG es la energía libre experimental de enlace.

Pdb_system	Score_np	Score_p	dG
Alr2_2dv0	-12,3	-11,3	-11,4
Alr2_2hv5	-12,1	-11,1	-11,4
Alr2_2iq0	-7,9	-5,9	-5,7
Alr2_2isf	-5,2	-4,7	-4,7
Chk1_2br1	-7,6	-5,6	-7

Pdb_system	Score_np	Score_p	dG
Hivrt_1eet	-12,7	-11,7	-11,3
Jnk_1pmu	-11,5	-10,5	-8,5
Pde4_1y2k	-9,1	-7,1	-10,4
Pde4_1ro6	-11,3	-9,3	-8,7
Throm_1g32	-10,8	-9,8	-8,3
Throm_1gj4	-8,2	-7,2	-5,2
Upa_1gjc	-8,7	-7,2	-11

Tabla 2. Número de ligandos señuelo con puntuaciones inferiores a -9 kcal/m para una versión de XP sin términos de penalización recientemente desarrollados (XP), XP con los términos de penalización más recientes (new XP) y XP con la adición del término de contacto con agua de esta patente (XP_wcon). La columna 4 tiene el término de contacto con agua eliminado de newXP (newXP_nowcon).

Pdb_system	XP_	newXP	XP_wcon	newXP_nowcon
Abl	24	15	20	18
Alr2	18	4	8	4
Jnk	7	2	3	4
Aur	2	2	2	2
Cdk2	10	3	8	4
Chk1	7	4	5	5
Dpp4	7	1	7	1
Er	4	0	2	0
Erk2	0	0	0	0
Err	3	2	3	2
Fvii	2	0	0	0
Fxa	0	0	0	0
Hivrt	11	2	7	8
Hsp90	2	0	1	0
Lck	10	3	7	4
Oppa	1	1	1	1
Pim1	46	3	46	3
PKa	0	0	0	0
Ppar	1	0	0	0
Ptp1b	1	1	1	1
Rho	5	2	5	2
Throm	2	1	2	1
Upa	0	0	0	0
P38	0	0	0	0

Otras formas de realización se encuentran dentro del alcance de las siguientes reivindicaciones. Debe observarse que los valores de penalización anteriores solo son representativos. Típicamente en ligandos físicamente razonables, la penalización total evaluada en virtud de la presente invención será inferior a 5 kcal/mol.

Anexo I

5 Los tipos de átomos hidrofóbicos están codificados como patrones SMARTS estándar de la industria. Todas las porciones de ligando hidrofóbico son aquellas seleccionadas con ptype=1. El código de asignación ptype depende del orden, es decir, un grupo podría coincidir con múltiples patrones, pero obtiene el ptype para el patrón en la lista que se captura por primera vez.

	Matt
# El primer índice es el "ptype"	1 : pesado hidrofóbico
#	2 : hidrógeno donante
#	3 : heteroátomo aceptor
#	4 : polar general (enlace sin hidrógeno)
#	5 : catión metálico
#	0 : ninguna de las anteriores
# El segundo índice es la hibridación 0 : átomo libre	
#	1 : sp
#	2 : sp2
#	3 : sp3
# El tercer índice es el patrón SMARTS para correspondencias	
# El cuarto índice es una etiqueta descriptiva	
#	
# El orden de los patrones es IMPORTANTE - patrones superiores en la lista para un	
# elemento particular se emparejarán preferentemente en relación con un patrón inferior.	
#	
# Hidrógenos	
2 1 [#1][#7]	Hidrógeno en nitrógeno
2 1 [#1][S;X2]	Hidrógeno en tioles
2 1 [#1][O-]	Hidrógeno en ion hidróxido
2 1 [#1][O;X2]	Hidrógeno en alcoholes
0 1 [#1][c,n,o]	Hidrógeno aromático
0 1 [#1][C]	Hidrógeno en carbono
0 1 [#1]	Hidrógeno predeterminado
# Carbonos	
1 1 C#C	Carbono de alquino
4 1 C#N	Carbono de nitrilo
1 1 [CH2](C#[N-])	Nitrilo de metilo
4 2 [c](~[n])~[n]	Amidina aromática
1 2 [c]	Carbono aromático
1 2 C=C	Carbono de alqueno
4 3 [CX4](~[O,N,n])~[O,N, n]	Carbono tetraedro con dos acoplamientos N/O
1 2 [CX3](=[N])([NX3])[NX3]	Carbono de guanidina
4 2 C(=O)[O]	Carbono de carbonilo - ácido
4 2 C(=[O,S])[N;X3]	Carbono de carbonilo -amina/tioamina
4 2 C=[O,N,S]	Carbono de carbonilo predeterminado
1 1 [C;X2]	Carbono sp predeterminado
1 2 [C;X3]	Carbono sp2 predeterminado
1 3 [C;X4]	Carbono sp3 predeterminado
1 3 [#6]	Carbono predeterminado

# Nitrógenos	
3 1 [N]#C	Nitrógeno en nitrilo
4 2 [n;X3]	Nitrógeno aromático con tres vecinos
3 2 [n]	Nitrógeno aromático
3 2 N(=N=N)	Nitrógeno de azida
3 2 N(=N)=N	Átomo central de nitrógeno de azida
3 2 [N;X2]=C[N;X3]	Nitrógeno de amidina neutro
4 2 [N;X3][*]=[*]	Nitrógeno conjugado
4 2 [N;X3][c,n]~[c,n,0]	Nitrógeno conjugado aromáticamente
4 2 [NX3;+]	Nitrógeno
3 2 [NX2]	Nitrógeno de imina
4 3 [NH3]	Nitrógeno de amoníaco
4 3 [NX4;+]	Nitrógeno protonado
3 1 [N;X1]	Nitrógeno sp predeterminado
3 2 [N;X2]	Nitrógeno sp2 predeterminado
3 3 [N;X3]	Nitrógeno sp3 predeterminado
3 3 [#7]	Nitrógeno predeterminado
# Oxígeno	
3 2 [O;X1]~[N;X3]~[O;X1]	Átomos de oxígeno nitroso
3 2 [O;X1;-]C=[O;X1]	Ácido carboxílico oxígeno
3 2 [O-][P,S]	Fosfato/oxígeno sulfónico
3 3 [O-]	Anión de oxígeno
3 3 [o]	Oxígeno aromático
3 2 [O;X1]	Oxígeno sp2 predeterminado
3 3 [O;X2]	Oxígeno sp3 por defecto
3 2 [#8]	Oxígeno predeterminado
# Segunda fila de elementos	
4 3 [#15]	Fósforo predeterminado
1 2 [s]	Azufre aromático
4 2 [SX1]=[*]	Azufre de tiocarbonilo
3 3 [S-;X1]	Azufre de anión de tiolato
4 3 [SX2;H1]	Azufre de tiol
1 2 [SX2]	Azufre disustituido
4 3 S(=O)(=O)([C,N])([C,N])	Azufre en sulfona/sulfonamida
4 3 S(=O)([#6])([#6])	Azufre en sulfóxido
4 3 [SX4]	Azufre predeterminado
4 3 [#16]	Azufre predeterminado
# Halógenos	
3 0 [F-]	Ion de flúor
3 0 [Cl-]	Ion de cloro
4 0 [Br-]	Ion de bromo
4 0 [I-]	Ion de yodo
4 1 [#9]	Flúor predeterminado
1 1 [#17]	Cloro predeterminado
1 1 [#35]	Bromo predeterminado

ES 2 788 504 T3

1 1 [#53]

Iones

5 0 [*;X0]

Yodo predeterminado

Ion metálico libre

REIVINDICACIONES

1. Un procedimiento para puntuar la afinidad de enlace de una molécula de ligando propuesta para una molécula receptora, que es una proteína, usando un ordenador y bases de datos informáticas, comprendiendo el procedimiento,
- 5
- a. obtener datos almacenados por ordenador que representan una estructura de ligando-receptor predicha,
b. puntuar la afinidad de enlace de ligando mediante operaciones informatizadas que incluyen:
- 10
- determinar si el receptor incluye uno o más sitios de hidratación ocupados por agua localizada y, de ser así,
 - determinar
- 15
- si después de acoplar un agua localizada interactúa con un átomo de ligando hidrofóbico, y si el acoplamiento del ligando en el receptor da como resultado una interacción desfavorable del ligando con una molécula de agua localizada que ocupa un sitio de hidratación de receptor después del acoplamiento, donde la interacción desfavorable incluye un mayor aumento de energía libre en comparación con la interacción del ligando con una molécula de agua en un fluido a granel,
- 20
- determinar si una o más de las condiciones de excepción i. - iv. está presente:
 - i. el átomo de ligando hidrofóbico tiene un átomo de ligando con una carga neta positiva a 4 átomos o más cerca del mismo en función de la conectividad de enlace de los átomos en el ligando;
 - 25 ii. el átomo de ligando hidrofóbico no es parte de un grupo -CH₃ de ligando o no es un átomo de ligando a menos de 3 átomos de un grupo -CH₃, donde el agua localizada bordea el borde de un bolsillo hidrofóbico que le da acceso al agua circundante;
 - 30 iii. el agua localizada dentro de un sitio de hidratación está a menos de 5 Å de un átomo de ligando con carga positiva;
 - iv. el agua localizada dentro de un sitio de hidratación tiene enlaces de hidrógeno al ligando;
 - aplicar una penalización a una función de puntuación de afinidad de enlace si 1) el agua localizada interactúa con un átomo de ligando hidrofóbico como se determina en la etapa b, y 2) ninguna de las
 - 35 condiciones de excepción i. - iv. está presente.
2. El procedimiento según la reivindicación 1, en el que los sitios de hidratación de receptor se determinan mediante muestreo informatizado de sitios para hidratación, seguido de asignación informatizada de ocupaciones de agua para los diversos sitios muestreados.
- 40
3. El procedimiento según la reivindicación 1, en el que se evalúa una penalización para compuestos cuando el agua localizada dentro de un sitio de hidratación interactúa con un átomo de ligando hidrofóbico.
4. El procedimiento según la reivindicación 3, en el que el agua localizada dentro de un sitio de hidratación interactúa con un átomo de ligando hidrofóbico a una distancia de 5 Å o menos.
- 45
5. El procedimiento según la reivindicación 4, en el que el átomo de ligando hidrofóbico no es Cl o Br.
6. El procedimiento según la reivindicación 1, en el que se determina que los sitios de receptor con un agua localizada son aguas que se predice que tienen una ocupación en el marco de referencia de la proteína de > 1,5X la ocupación de agua a granel.
- 50
7. El procedimiento según la reivindicación 2, en el que se utilizan simulaciones de dinámica molecular para determinar la ocupación de agua.
- 55
8. El procedimiento según la reivindicación 7, en el que se utiliza un análisis de solvatación de sitio activo para identificar sitios de agua localizada.
9. El procedimiento según la reivindicación 1, en el que se aplica una penalización a la puntuación de afinidad de enlace del ligando cuando se determina que la energía libre de desplazamiento de una molécula de agua desde el sitio de receptor hasta aguas a granel es más favorable que un valor de corte especificado.
- 60
10. El procedimiento según la reivindicación 1, en el que se aplica una penalización a la puntuación de afinidad de enlace del ligando mediante un procedimiento que comprende estimar la favorabilidad energética de la molécula de agua en el sitio de hidratación a partir de la hidrofobicidad del sitio o a partir de interacciones específicas con grupos de receptor.
- 65

11. El procedimiento según la reivindicación 10, en el que la estimación de la favorabilidad energética del sitio de hidratación incluye cálculos seleccionados del grupo que consiste en simulaciones de dinámica molecular, simulaciones Monte Carlo, enumeración exhaustiva de geometrías de la molécula de agua dentro del sitio y optimización (mediante, por ejemplo, minimización de energía) de la posición de la molécula de agua dentro del sitio.
- 5
12. El procedimiento según la reivindicación 9, en el que el valor de corte es de al menos 0 kcal/mol.
13. El procedimiento según la reivindicación 9, en el que las energías libres de desplazamiento de una molécula de agua desde el sitio de receptor hasta el agua a granel se determinan mediante simulación de dinámica molecular seguida de la aplicación de la teoría de solvatación no homogénea modificada para asignar energías libres de desplazamiento.
- 10

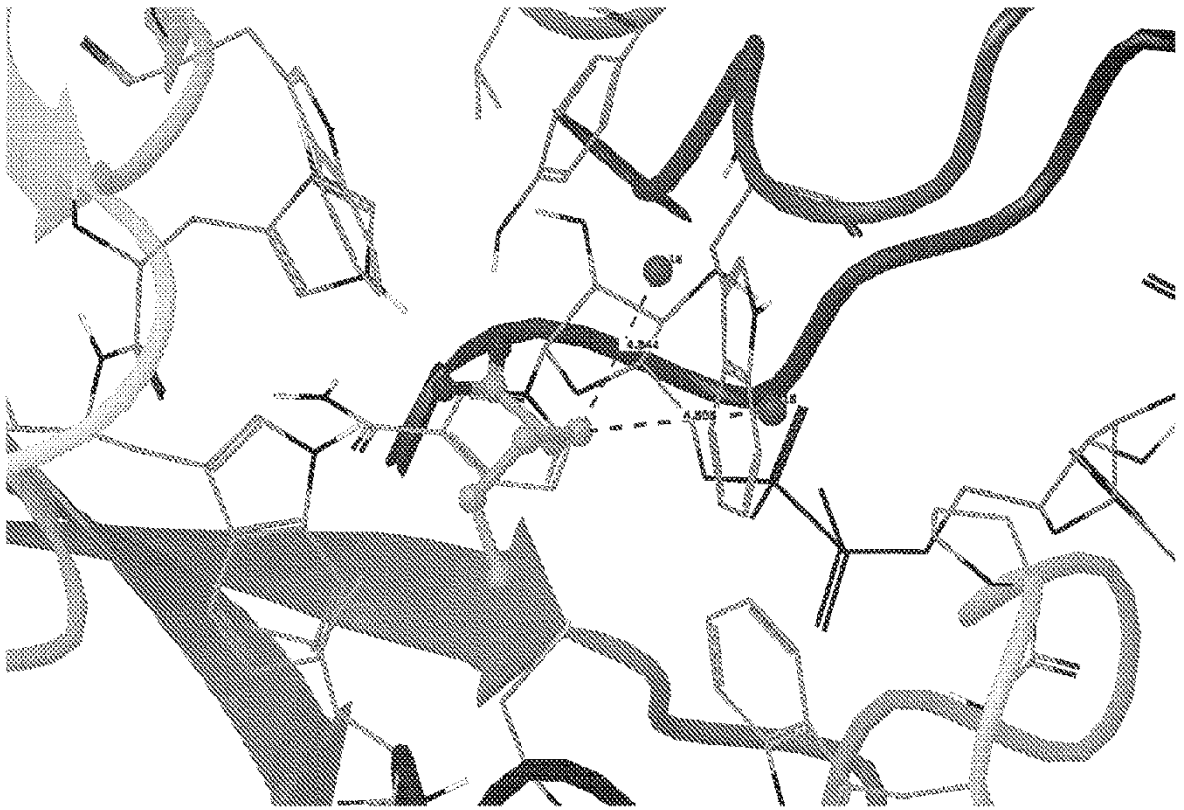


FIGURA 1

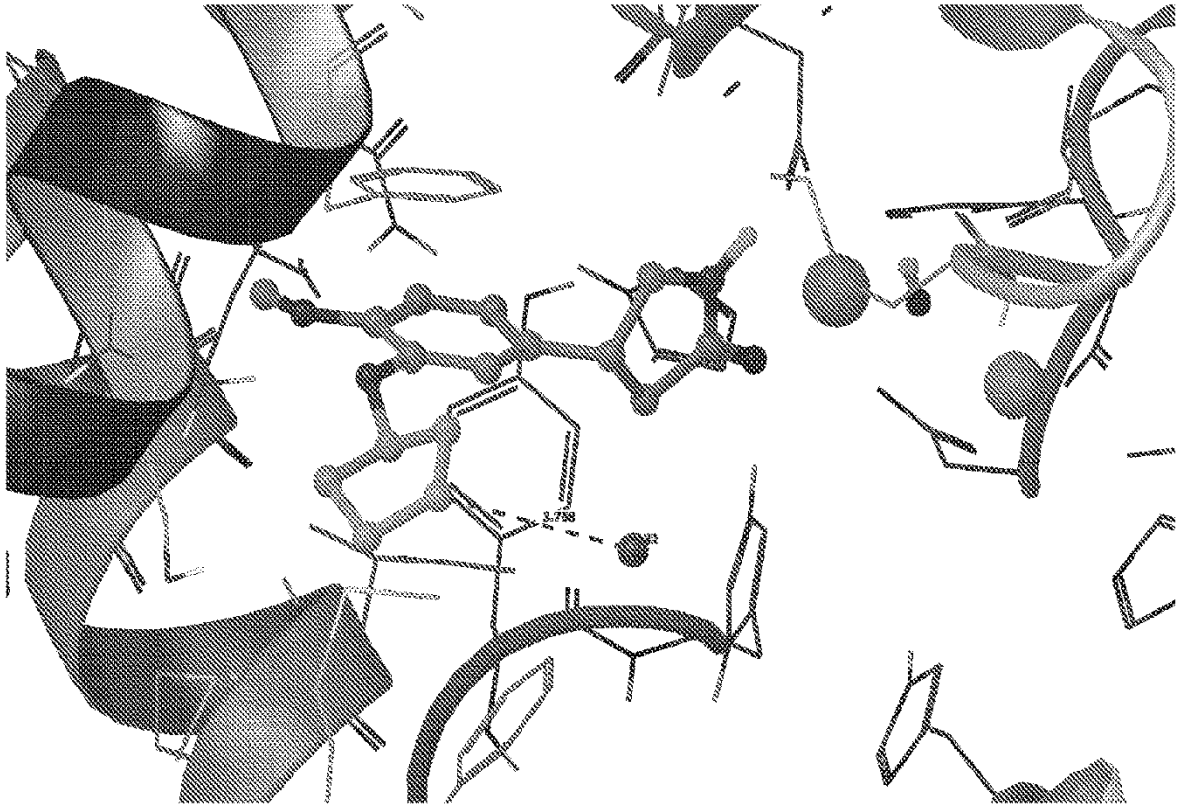


FIGURA 2

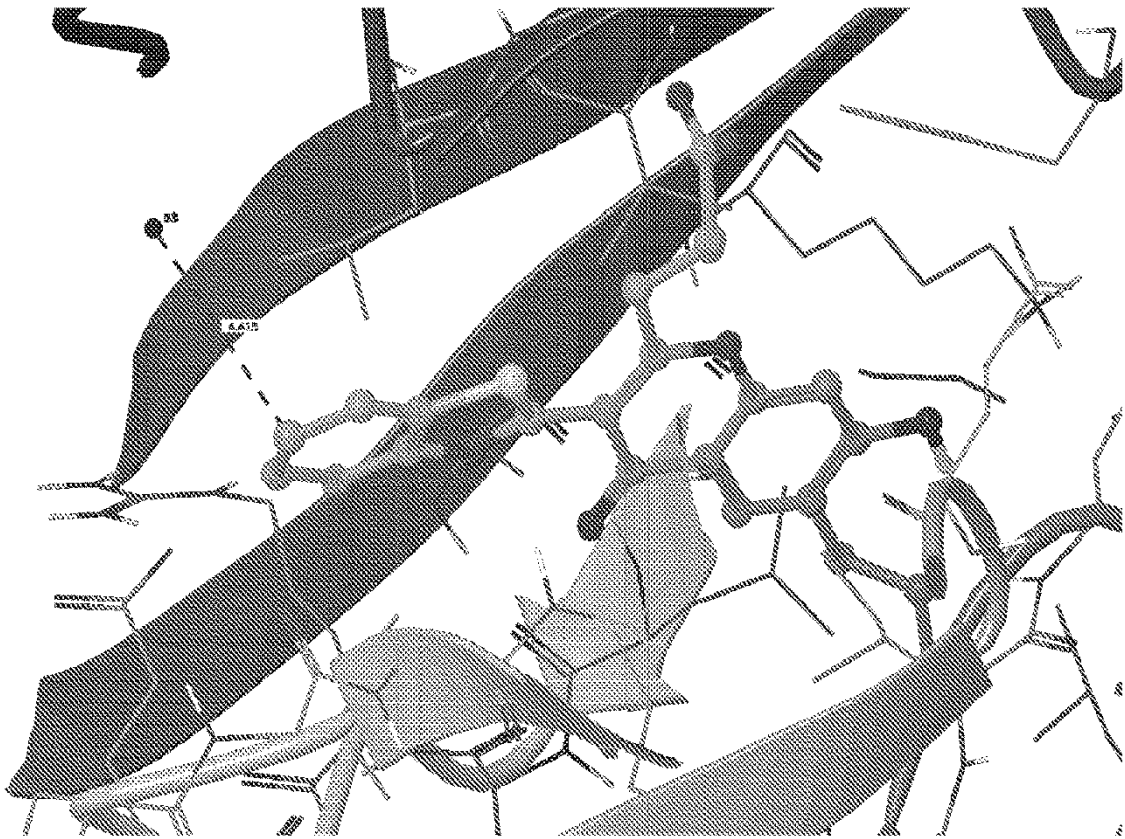


FIGURA 3

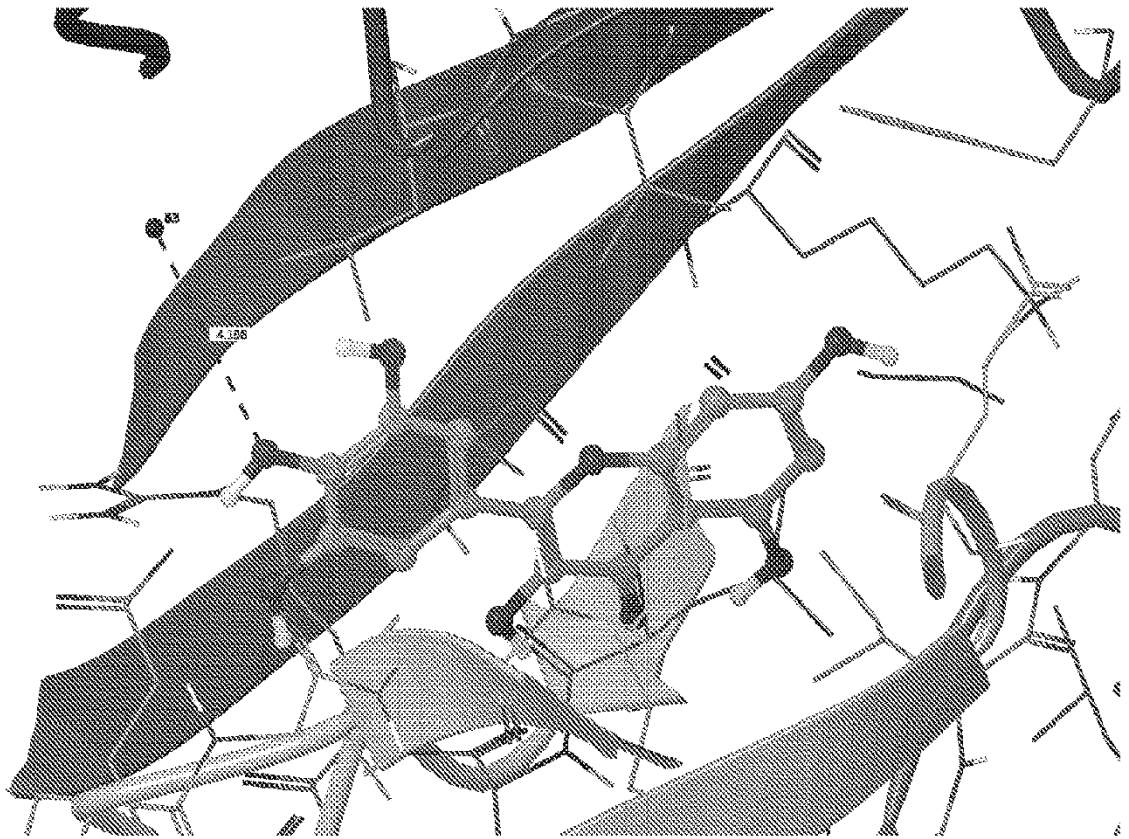


FIGURA 4