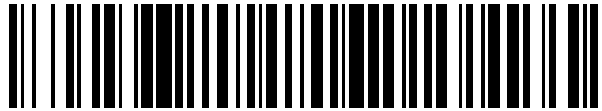


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 615 179**

51 Int. Cl.:

G06F 19/16 (2011.01)

G06F 19/00 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **12.05.2008 PCT/US2008/063364**

87 Fecha y número de publicación internacional: **20.11.2008 WO08141260**

96 Fecha de presentación y número de la solicitud europea: **12.05.2008 E 08755277 (4)**

97 Fecha y número de publicación de la concesión europea: **09.11.2016 EP 2162842**

54 Título: **Función de puntuación de la afinidad de unión que incluye un factor para los entornos del grupo del anillo**

30 Prioridad:

14.05.2007 US 930089 P

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

05.06.2017

73 Titular/es:

**SCHRÖDINGER, LLC (100.0%)
120 WEST FORTY-FIFTH STREET 29TH FLOOR,
TOWER 45
NEW YORK, NY 10036-4041, US**

72 Inventor/es:

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

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 615 179 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 la afinidad de unión que incluye un factor para los entornos del grupo del anillo

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 funcionan mediante la unión química 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 la selección de fármacos puede ser facilitada por la estimación precisa de la afinidad de unión de un fármaco a una proteína o, más generalmente, la estimación de la afinidad de unión de un ligando a un receptor, utilizándose el término receptor para dar a entender cualquier resto que se une específicamente al ligando.

- 15 Una forma de determinar la afinidad de unión receptor-ligando utiliza la estructura molecular del complejo receptor-ligando que resulta cuando el ligando se une al receptor. Tales estructuras pueden ser estudiadas por cristalografía de rayos X. El banco de datos de proteínas de acceso público (PDB) contiene ahora más de 10.000 estructuras cristalinas de rayos X, y las compañías farmacéuticas y de biotecnología tienen un orden de magnitud más estructuras de propiedad. Muchas de estas estructuras han sido co-cristalizadas con pequeñas moléculas unidas a ellas. Al examen de este tipo de estructuras y al despliegue del conocimiento ganado de este modo para diseñar nuevos inhibidores más potentes y más específicos, se les alude como diseño de fármacos basado en la estructura.

- 20 El modelado computacional facilita el diseño de fármacos basado en la estructura. Un aspecto de la modelización detallada a continuación incluye funciones de puntuación que utilizan técnicas de simulación tales como la dinámica molecular, Monte Carlo o cálculos electrostáticos de continuo. Estas funciones de puntuación pueden ser problemáticas, ya que 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 separada y el ligando). Un enfoque alternativo es desarrollar una función de puntuación empírica, en base a la geometría del complejo, que evalúa directamente la cantidad deseada. Un enfoque de este tipo tiene la ventaja de ser extremadamente rápido, además de ser susceptible de ajuste a los datos experimentales, grandes cantidades de las cuales ya están disponibles públicamente. La propia solicitud de la presente solicitante, solicitud de patente de EE.UU. número de serie de 25 11/373.684 (documentos US 2007/061118 y WO 2006/099178A2) titulada "Predictive Scoring Function for Estimating Binding Affinity", describe funciones de puntuación.

- 30 Es deseable aumentar la precisión y solidez de funciones de puntuación. Es particularmente deseable utilizar funciones de puntuación no sólo para distinguir ligandos activos de inactivos, sino también para clasificar ligandos en un grupo de ligandos, p. ej., de más activos a menos activos, para identificar candidatos más prometedores para ser estudiados delante de candidatos menos prometedores. Además de ello, la capacidad para jerarquizar un grupo de ligandos que tienen diferentes quimiotipos permitiría buscar una gama más amplia de candidatos a fármacos eficaces. Por último, la jerarquización de ligandos sin una gran cantidad de trabajo preliminar (p. ej., utilizando un conjunto de entrenamiento) aumentaría aún más la conveniencia y la rapidez del proceso.

SUMARIO

- 40 Los autores de la invención han descubierto que funciones de puntuación pueden ser notablemente mejoradas teniendo en cuenta los entornos de los anillos de ligando cuando el ligando forma un complejo con el receptor. En su forma más general, la invención presenta la cuantificación de un componente particular de la afinidad de unión entre un ligando y una molécula de receptor que tiene en cuenta la relación espacial entre la o las estructuras del anillo ligando y sus entornos ambiente, cuando el ligando se une a la molécula de receptor. Las etapas del método que 45 han desarrollado los autores de la invención se pueden utilizar cuando la cuantificación de un componente que refleja estas características de ligando particulares (p. ej., anillos de ligando) solo, o las etapas pueden ser parte de un método global de la cuantificación de la afinidad de unión que incluye numerosos otros factores que se relacionan con la afinidad de unión además de los anillos que reflejan el componente o estructuras rígidas voluminosos. Por

ejemplo, se puede calcular una afinidad de unión inicial y luego ajustar la afinidad de unión inicial por un factor obtenido, al menos en parte, sobre la base de la clasificación de la estructura del anillo.

5 En un aspecto, la invención presenta en general: a) la identificación de una estructura de anillo en el ligando; la clasificación de la estructura de anillo basada en la relación espacial de la estructura de anillo a sus entornos
ambientes cuando el ligando está unido a la molécula de receptor; la cuantificación de al menos el componente de la
afinidad de unión entre el ligando o parte del mismo y la molécula de receptor basado, al menos en parte, en la
clasificación de la estructura del anillo; y mostrar, a través del ordenador, información relacionada con la clasificación
de la estructura del anillo. En este contexto, los autores de la invención utilizan la expresión "estructura de anillo"
10 para dar a entender un anillo completo o un conjunto de anillos condensados. La estructura de anillo se clasifica de
acuerdo con la relación espacial entre la estructura de anillo y el receptor, en función de cómo se expone la
estructura de anillo al disolvente del entorno.

A la estructura de anillo se la asigna, por lo tanto, una de las siguientes clasificaciones: a) enterrada; b) expuesta al
disolvente; o c) expuesta al disolvente en una superficie, aunque también se pueden utilizar clasificaciones o sub-
clasificaciones adicionales. Preferiblemente, el método incluye una etapa que puntúa las interacciones hidrofóbicas
15 entre uno o más átomos de ligando y uno o más átomos de receptor mediante la concesión de un bono en cuanto a
la presencia de la envoltura hidrofóbica de uno o más átomos del ligando por la molécula de receptor. El bono es
indicativo de afinidad de unión mejorada entre el ligando y la molécula del receptor. Para clasificar la estructura de
anillo se puede obtener un valor representativo de la relación espacial entre la estructura del anillo y sus entornos
ambientes cuando se une a la molécula de receptor a) clasificando una estructura de anillo como enterrada si el
20 valor está por encima de una primera cantidad; b) clasificando una estructura de anillo como disolvente expuesto si
el valor está por debajo de una segunda cantidad, y c) clasificando una estructura de anillo que tiene una sola cara
expuesta al disolvente si el valor es una cantidad de entre la primera y segunda cantidad. En una realización: a) se
determina el número de contactos cercanos a diferentes distancias entre los átomos de los receptores y las dos
caras del anillo, b) una clasificación inicial del anillo se realiza en base al número de estos contactos y c) esta
25 clasificación inicial es seguida por el cálculo de una función de puntuación, depende de otros parámetros, lo cual
permite la clasificación final como enterrada, expuesta a disolvente o expuesta a disolvente en una superficie. El
número de estrechos contactos a diferentes distancias entre los átomos de los receptores y las dos caras del anillo
puede ser determinado con el fin de hacer una clasificación inicial del anillo basado en los número de estos
contactos, y esta clasificación inicial es entonces seguida por el cálculo de una función de puntuación, dependiendo
30 de otros parámetros que permiten la clasificación final como enterrada, expuesta a disolvente o expuesta a
disolvente en una superficie.

La función de puntuación que permite la clasificación de la estructura de anillo incluye calcular el grado de
envoltura de cada uno de los átomos de la estructura de anillo por átomos del receptor. También incluye el uso de
un parámetro que se correlaciona sustancialmente con el grado de envoltura de cada uno de los átomos de la
35 estructura de anillo por átomos del receptor. En los casos en los que los entornos ambientes de la estructura del
anillo cuando se une a la molécula del receptor incluyen moléculas de agua, la función de puntuación puede incluir el
uso de un parámetro correspondiente a un cierto número de moléculas de agua en una primera envoltura de la
estructura del anillo, o en una segunda envoltura de la estructura de anillo o en una combinación de la primera
envoltura y la segunda envoltura. La función de puntuación que permite la clasificación de la estructura de anillo
40 como enterrada, expuesta a disolvente o expuesta a disolvente en una superficie también puede incluir el uso de un
parámetro correlacionado sustancialmente con un cierto número de contactos estrechos en ambos lados de la
estructura de anillo o parte del mismo con la molécula de receptor. También puede incluir el uso de un parámetro
correspondiente a una interacción hidrofóbica de la estructura de anillo o parte de la misma con la molécula de
receptor. Por último, puede incluir el uso de un parámetro correlacionado sustancialmente con la puntuación del par
45 lipofílico-lipofílico entre la estructura de anillo o parte de la misma y la molécula de receptor.

En los casos en los que la estructura de anillo incluye anillos condensados, el método de cuantificación incluye al
menos una etapa de contabilizar de fusión del anillo. Por ejemplo, la contabilización de la fusión del anillo resulta en
una afinidad de unión que es menor que la afinidad obtenida sin dicha contabilización, por una cantidad relacionada
con el grado de fusión del anillo.

50 Etapas específicas en la cuantificación y en la puntuación de la clasificación de la estructura del anillo se encuentran
en el Anejo A. Estas etapas pueden ser utilizadas individualmente o en combinación. Alternativamente, se puede
obtener una afinidad de unión preliminar y ajustar esa afinidad de unión preliminar en base a una visualización de las
estructuras de anillo.

Cuando una estructura de anillo se designa expuesta al disolvente, se ignora sustancialmente en la cuantificación del componente de afinidad de unión, que no sea para reconocer los enlaces hidrógeno y otros parámetros que son independientes de la clasificación de la estructura de anillo. Más específicamente, la contribución hidrofóbica de estructuras de anillo expuestas a disolvente a la afinidad de unión se ignora sustancialmente en la cuantificación del componente de la afinidad de unión, pero se pueden reconocer otros aspectos de los anillos expuestos a disolvente (p. ej., enlaces hidrógeno).

El método anterior puede incluir métodos que identifican la energía de deformación inducida en el ligando cuando se une a la molécula de receptor y utilizar esa cantidad de energía de deformación cuando se cuantifica el componente de afinidad de unión. Por ejemplo, el método puede incluir: a) la identificación de una cantidad representativa de una energía de deformación inducida en el complejo ligando-receptor por el agregado de las estructuras de anillo identificadas como enterradas; b) la identificación de una cantidad representativa de una energía total de enlace de hidrógeno neutra-neutra; y c) la cuantificación del componente de la afinidad de unión entre el ligando y la molécula de receptor basándose, al menos en parte, en la cantidad representativa de la energía de deformación inducida en el receptor por el agregado de las estructuras de anillo enterradas y en la cantidad representativa de la energía total de enlace de hidrógeno neutra-neutra. Además, la afinidad de unión puede incluir la identificación de una energía que remata el enlace hidrógeno asociada con todo el ligando, en cuyo caso el componente afinidad de unión se cuantifica en base a una energía mayor de la energía que remata el enlace hidrógeno y la cantidad representativa de la energía de deformación inducida en el receptor por el agregado de las estructuras identificadas. El componente de afinidad de unión atribuible a la deformación puede ser cuantificado utilizando: dinámica molecular, mecánica de la molécula, búsqueda y/o minimización conformacional.

La información mostrada por el ordenador puede incluir una representación de la exposición a disolventes, en su caso, de la estructura del anillo. También puede incluir una representación del entierro, si existe, de estructura o estructuras de anillo. Otra información mostrada incluye una representación de al menos uno de los siguientes: a) el grado en el que la estructura de anillo está encerrada por átomos de la molécula de receptor; b) moléculas de agua que rodean la estructura de anillo en una primera envoltura o una segunda envoltura o tanto la primera como la segunda envoltura del ligando; c) un valor de una puntuación del par lipofílico-lipofílico de la estructura del anillo; y el número de los contactos estrechos de un lado de la estructura de anillo con la molécula de receptor.

El método puede incluir identificar un motivo de unión de la molécula de receptor con respecto al ligando, y cuantificar la afinidad de unión en base al motivo de unión. Una energía de reorganización del receptor puede ser identificada en base al motivo de unión. Más específicamente, el método puede incluir: a) identificar un motivo de unión de la molécula de receptor con respecto al ligando; b) identificar una energía de reorganización de la molécula de receptor en base al motivo de unión; y c) identificar una primera estructura de anillo como una contribución a la energía de reorganización. La cantidad representativa de energía de deformación se identifica con independencia de la clasificación de la primera estructura de anillo.

Después de haber realizado la cuantificación arriba descrita, se puede entonces realizar una prueba en una muestra física que incluye el ligando y la molécula de receptor, y se pueden seleccionar componentes de prueba en base, al menos en parte, a la afinidad de unión entre el ligando o parte del mismo y el receptor de molécula, o en el componente de dicha afinidad de unión. Por ejemplo, la prueba se puede utilizar para medir una eficacia terapéutica del receptor en el tratamiento de una enfermedad.

DESCRIPCIÓN DE LOS DIBUJOS

La FIG. 1 es una representación del complejo de proteína-ligando 1aq1 de la base de datos de proteínas (PDB). La Figura muestra la envoltura de los diversos sub-anillos en el amplio sistema de anillos condensados 1aq1 por la proteína, muchos de los cuales se definen como enterrados por el actual mejor esquema de clasificación en la práctica; también muestra los enlaces hidrógeno entre el ligando y la proteína.

La FIG. 2 es una representación del complejo de proteína-ligando 1d3d de la PDB. La Figura muestra la envoltura de los diversos anillos del ligando 1d3d por la proteína, que se definen como enterrados por el actual mejor esquema de clasificación en la práctica. Téngase en cuenta que hay muy pocos enlaces hidrógeno en la estructura.

La FIG. 3 es una representación del complejo de proteína-ligando 2d3u. El ligando contiene dos anillos (los del lado derecho de la Figura) que se definen como expuestos a disolvente por el actual mejor esquema de clasificación en la práctica. Un anillo enterrado se muestra en el lado izquierdo del ligando.

Las FIGs. 4-10 representan los resultados obtenidos con la actual mejor implementación en la práctica de la invención para los complejos de proteína-ligando en la PDB para 7 receptores diferentes, comparando las afinidades de unión calculados con los datos experimentales afinidad de unión. Los resultados obtenidos en las FIGs. 4-10 se discuten con más detalle en la Descripción Detallada que figura a continuación.

5 DESCRIPCIÓN DETALLADA

Teniendo en cuenta una estructura de receptor de alta resolución (siendo el receptor típicamente una proteína), se utiliza un software de cálculo para "anclar" un ligando de molécula pequeña en la posición correcta y orientarlo en la cavidad del sitio activo del receptor, y calcular una afinidad de unión del ligando dada esta estructura. A los programas informáticos que realizan esta tarea se les alude como programas de "anclaje".

10 Un programa de anclaje típicamente lleva a cabo dos tareas distintas para modelar la unión receptor-ligando. En primer lugar, se predice una estructura de un complejo receptor-ligando. Cuando este supuesto falla, se requiere el uso de una estructura diferente a la del receptor como punto de partida. El problema de construir estructuras de receptor alternativas que son modificadas para aceptar ligandos que requieren un cambio sustancial en la conformación del receptor ("ajuste inducido") es muy importante.

15 Los autores de la invención se centran aquí en los casos en donde el anclaje del ligando a un receptor produce una estructura en acuerdo razonable con los datos experimentales, y en particular se centran en una segunda tarea del programa de anclaje, que consiste en calcular una afinidad de unión receptor-ligando, dada como una entrada de la estructura anclada. Una función matemática empleada para calcular la afinidad de unión (o una contribución a la misma) que se conoce como una "función de puntuación".

20 Las estimaciones de las afinidades de unión receptor-ligando descritas a continuación son aplicables cuando una estructura del complejo receptor-ligando está representada por un modelo estructural adecuado. Existe un cierto número de formas de caracterizar la calidad de los modelos estructurales de los complejos ligando-receptor. Para obtener puntuaciones precisas se necesita generalmente una pequeña desviación de la raíz cuadrada media (RMSD) de la estructura experimental (típicamente menos de 2 Angstroms, aunque el valor requerido variará dependiendo de los detalles del complejo), la formación de sustancialmente todos los enlaces hidrógeno observados en el complejo experimental, la colocación apropiada de sustancialmente todos los grupos hidrofóbicos en las cavidades de los receptores correctos y la ausencia de choques estructurales o electrostáticos incorrectos que podrían conducir a la asignación de términos de penalización sustancialmente incorrectos. Sólo importa la afinidad de unión relativa de ligandos a un receptor dado; un desplazamiento constante, como es en muchos casos engendrado por la energía de reorganización del sitio activo del receptor para acomodar el ligando, no tiene efecto alguno en las aplicaciones prácticas. En muchos (aunque no todos) de los casos, el receptor puede adoptar más de una conformación fundamentalmente diferente en respuesta a una clase de ligandos (p. ej., aglutinantes DFG-in (activos) y DFG-out (inactivos) a p38 map quinasa); para comparar las energías libres de unión en tales casos, pueden ser necesarios diferentes parámetros básicos de reorganización del núcleo para las diferentes conformaciones del receptor. En los casos en los que el cálculo de estos parámetros no es práctico a partir de primeros principios, se tratan como parámetros ajustables, específicos para receptores. Sin embargo otros parámetros están contenidos en un modelo global que no es específico para el receptor o incluso específico para una clase particular de receptores.

40 En la presente solicitud, los autores de la invención se centran en el cálculo de componentes importantes de la afinidad de unión al ligando cuando está disponible una estructura adecuadamente precisa para el complejo ligando-receptor, tal como se ha definido anteriormente. El método es operativo para ligandos para los cuales no están disponibles estructuras cristalinas. Sin embargo, como un control para la evaluación del método, los siguientes ejemplos implican estructuras cristalinas conocidas disponibles en el Banco de Datos de Proteínas ("el PDB") públicamente accesible de Collaborative Research for Structural Bioinformatics. Al llevar a cabo la optimización, los autores de la invención utilizan poses anclados con Glide XP, una función de puntuación descrita p. ej., en la solicitud de patente de EE.UU., número de serie 11/373.684 (US 2007/061118) titulada "Predictive Scoring Function for Estimating Binding Affinity", filtrando el número (muy pequeño) de casos para los que el auto-anclaje proporciona estructuras inadecuadas. Al utilizar estructuras ancladas más que estructuras cristalinas en sí mismas, en el proceso de optimización, los autores de la invención aumentaron el realismo del modelo, y también permitieron el anclaje para corregir pequeños errores geométricos en las estructuras cristalinas (p. ej., en distancias de los enlaces hidrógeno) que puede ser crucial para asignar adecuadamente puntuaciones a estos términos. La invención no requiere el uso de ligandos para los que se conocen las estructuras cristalinas, ni requiere el uso de Glide XP.

Las estructuras del PDB pueden ser vistas como un amplio y diverso conjunto de entrenamiento para la función de puntuación. El ensayo de la función de puntuación en condiciones similares puede ser realizado por compañías farmacéuticas y de biotecnología, utilizando conjuntos de datos de propiedad, en los casos en los que estén disponibles estructuras cristalinas. En la realización de estos ensayos, no hay necesidad de liberar las estructuras o incluso de divulgar el nombre del receptor; simplemente se pueden realizar los cálculos, e informar de la capacidad para jerarquizar los compuestos como un coeficiente de correlación.

El comportamiento cualitativo de compuestos farmacéuticos que interactúan con receptores de proteínas es, en muchos casos, dominado por la interacción de los grupos voluminosos y rígidos en el ligando con el receptor. Un grupo es "voluminoso" si incluye cuatro o más átomos pesados (p. ej., no hidrógeno). Un grupo es "rígido" si sus enlaces constituyentes poseen menos de un umbral predeterminado de libertad de torsión. En algunas implementaciones, el umbral está entre 5 y 10 grados. Por ejemplo, las estructuras rígidas y voluminosas incluyen anillos y grupos terc.-butilo.

Por mucho, las estructuras rígidas y voluminosas más importantes son estructuras de anillo. Tipos de estructuras de anillo incluyen anillos aromáticos y saturados, anillos condensados y un subconjunto conectado de átomos en un solo anillo o anillo condensado. En lo que sigue, los autores de la invención se centrarán principalmente en un algoritmo para mejorar la caracterización de la contribución de la afinidad de unión hecha por estructuras de anillo con la afinidad total de unión al ligando. El algoritmo también puede aplicarse a otras estructuras rígidas voluminosas, sin embargo, aún no se ha desarrollado un modelo detallado de estructuras que no sean estructuras de anillo.

La afinidad de unión está típicamente influenciada por la interacción de estructuras de anillo del ligando con cavidades principalmente hidrofóbicas del receptor. Por ejemplo, la solicitud de patente de EE.UU. número de serie 11/373.684 (US 2007/061118) a la que se alude antes, identificó situaciones en las que particularmente se pueden obtener grandes incrementos de afinidad de unión, a saber, cuando un anillo está sujeto a una envoltura hidrofóbica por parte de la proteína, y cuando el anillo es complementario a cualesquiera grupos polares en la región encerrada (es decir, formando enlaces hidrógeno según sea necesario, pero que no presenten otros grupos polares o cargados a la cavidad hidrofóbicamente encerrada).

Los autores de la invención han investigado otros aspectos de cómo interactúa el anillo con el receptor. Este examen más general ha conducido al desarrollo de los métodos mejorados para el cálculo de los componentes de la afinidad de unión descritos en esta memoria. Estos componentes tienen una base física que se puede describir fácilmente y es objeto de investigación por medio de simulaciones atomizadas. Los componentes se pueden utilizar en y por sí mismos para entender mejor el efecto sobre la afinidad de unión o la separación de estructuras de anillo específicas para el ligando. También se pueden utilizar, en combinación con otros componentes de la función de puntuación, para producir una función de puntuación combinada que proporciona una mejor representación de la afinidad de unión que estaba disponible en trabajos anteriores. Específicamente, en combinación con la función de puntuación Glide XP descrita en la solicitud de patente de EE.UU. número de serie 11/373.684 (US 2007/061118), se puede desarrollar una función de puntuación que permite una jerarquización semicuantitativa de los ligandos con respecto a la afinidad de unión al receptor. Ejemplos de una jerarquización de este tipo se discuten más adelante.

En lo que sigue, los autores de la invención comienzan con una clasificación de estructuras de anillo del ligando en tres clases, dependiendo de la naturaleza de su interacción espacial con el receptor. A continuación, desarrollaron medidas cuantitativas para asignar cualquier estructura de anillo dada en una de las tres clases, utilizando una diversidad de descriptores. Por último, se presenta la implementación del esquema de clasificación en la función de puntuación, incluyendo fórmulas específicas para el cálculo de los componentes de energía libre de unión sobre una base de estructura por cada anillo, y el ensamblaje de estos términos en un componente de la afinidad de unión para todo el ligando.

Esquema de Clasificación del Anillo

Los autores de la invención clasifican estructuras de anillo en una de las tres categorías siguientes:

(1) *Estructuras de anillo enterradas*. Son estructuras que se insertan en una cavidad del receptor, con los átomos del receptor próximos a las dos caras del anillo, y predominantemente aisladas de disolvente acuoso. La clasificación de una estructura de anillo tal como se entierra se consigue mediante una función de puntuación dependiente de descriptores cuantitativos que miden la envoltura, el número de moléculas de disolvente próximas al ligando, el número de contactos estrechos en ambas caras del ligando y las interacciones hidrofóbicas con el receptor tal como

se describe con más detalle más adelante. Las FIGs. 1 y 2 muestran estructuras del PDB que contienen un número significativo de estructuras de anillo enterradas.

5 (2) *Estructuras de anillo de superficie.* Se trata de estructuras para las que el contacto principal con el receptor se realiza por una sola cara de la estructura de anillo, estando la otra cara predominantemente expuesta al disolvente. La cuestión de la frontera entre una estructura enterrada y una estructura de superficie - es decir, cuanto contacto con la proteína hace la segunda cara para ser colocada en orden en la categoría (1) anterior, en lugar de la categoría (2) - es una cuantitativa, y requiere la discriminación utilizando la función de puntuación numérica comentada anteriormente.

10 (3) *Estructuras expuestas a disolvente.* Son estructuras para la que ambas caras de la estructura del anillo están principalmente en contacto con el disolvente, a diferencia de la proteína. Una vez más, los criterios cuantitativos deben estar diseñados para separar estructuras de anillo expuestas a disolvente de estructuras de anillo de la superficie, en base al mismo esquema de puntuación numérica como se comentó anteriormente. Se podría preguntar por qué las estructuras de anillo están expuestas al disolvente, en lugar de hacer el mejor contacto disponible con el receptor con una cara de la estructura. A veces ese contacto es geoméricamente factible; en otros casos, la parte de la superficie que podría poner en contacto no es adecuada, poseyendo grupos polares o cargados que se desolvatarían en los casos en los que la estructura de anillo se aproximaría mucho al receptor. La FIG. 3 muestra un complejo del PDB que contiene dos anillos expuestos a disolvente y un anillo enterrado.

20 Habiendo definido las tres clases de estructuras de anillo anteriores, se comentan ahora los algoritmos numéricos que se utilizan para clasificar una estructura dada en una de las tres categorías anteriores. En algunas implementaciones de la clasificación de estructuras de anillo, incluyendo la utilizada en la implementación actual de mejor práctica de los autores de la invención, sólo se consideran los átomos pesados que comprende el núcleo de la estructura. Se pueden también considerar átomos directamente unidos a los átomos pesados. Para estructuras condensadas (es decir, estructuras que se componen de dos o más sub-estructuras solapantes), cada una de las sub-estructuras se evalúa por separado. Esto puede implicar la doble contabilización de algunos de los átomos compartidos entre las sub-estructuras. El doble recuento es tenido (aproximadamente) en cuenta en la función de puntuación final tal como se describe a continuación.

La primera etapa en la clasificación de la estructura es producir descriptores asociados con cada uno de los átomos en la estructura. Los autores de la invención emplean tres descriptores para cada uno de los átomos:

30 (1) *Una función de medición de la envoltura del átomo por el receptor.* Los autores de la invención utilizan para este fin la envoltura hidrofóbica de modelado función matemática descrita en el Anejo de la solicitud de patente de EE.UU. número de serie 11/373.684 (US 2007/061118), aunque con diferentes parámetros tal como se describe en el Anejo de la presente solicitud, principalmente para ponderar más fuertemente átomos más próximos. Una segunda e importante diferencia en la función que se utiliza aquí es que todos los átomos pesados de proteínas, en lugar de sólo átomos lipofílicos, se utilizan para computar la envoltura. Esta función evalúa directamente el grado de envoltura espacial del átomo en la estructura de anillo por el receptor.

40 (2) *El número de moléculas de agua en las primeras dos envueltas que rodean el átomo.* En la unión simulada, las aguas se añaden de forma explícita al complejo proteína-ligando y la asignación de aguas a envueltas se lleva a cabo mediante la definición de la primera envuelta como aguas en contacto directo (basado en el radio de van der Waals del átomo de ligando y una molécula de agua) con el átomo de ligando en cuestión, mientras que la segunda envuelta se define como moléculas de agua en el contacto de van der Waals con la primera envuelta. Este valor proporciona una medida directa del grado de exposición a disolventes del átomo.

(3) *La expresión par átomo-átomo lipofílico que representa interacciones del átomo del ligando con los átomos lipofílicos del receptor.* Esta función proporciona una medida de lo cerca que el átomo se encuentra a átomos lipofílicos en la superficie de la proteína (sin discriminar la disposición geométrica de los átomos circundantes).

45 La segunda etapa consiste en enumerar el número de estrechos contactos con el receptor hechos por las dos caras del anillo. Un contacto es "estrecho" a una cara si está dentro de 0,6 nm (6Å) de un átomo en el anillo, y satisface un criterio angular con respecto al plano del anillo definido en el Anejo. Para realizar la evaluación de cómo es más preciso dotar de forma más precisa a los contactos cercanos, los autores de la invención contaron el número de átomos pesados de proteína dentro de los diversos puntos de corte a distancia (p. ej., 0,4 nm (4Å), 0,45 nm (4,5Å),

0,5 nm (5Å) de cada una de las caras de la estructura. Los detalles de cómo los átomos se asignan a "recuento" para una cara dada de la estructura se dan en el Anejo.

Estos descriptores se combinan entonces para efectuar la clasificación de las estructuras. Los términos de envoltura, agua y par lipofílico se suman a lo largo de los átomos en la estructura para obtener un único valor de cada término. Cualitativamente, el uso de los descriptores en la función de clasificación es como uno esperaría. Con el fin de ser clasificado como expuesta a disolvente, una estructura que tiene un número relativamente grande de aguas circundantes, un valor relativamente pequeño del término envoltura y un número relativamente pequeño de contactos estrechos. Las estructuras de superficie tienen un valor relativamente pequeño del término envoltura, contactos predominantemente en un lado de la estructura (pero posiblemente muchos contactos en ese lado) y un número intermedio de las aguas circundantes. Los anillos encerrados tendrán un valor relativamente grande del término envoltura, un número relativamente pequeño de las aguas circundantes y los contactos estrechos en ambos lados del anillo. La adición de la expresión par lipofílico con el término envoltura produce un efecto de alisado que proporciona una medida algo mejor de cómo de cerca interactúa la estructura con la superficie de la proteína.

Un esquema de clasificación detallada se presenta en el Anejo 1. El esquema utiliza un enfoque de partición que incorpora algunos elementos no lineales; los términos contacto se agrupan para formar racimos, y luego dentro de cada uno de los racimos, se definen los valores de corte, que representan límites entre estructuras de anillo enterradas, de superficie y expuestas a disolvente, para la "puntuación del anillo" (término lipofílico más envoltura) y el número de las aguas circundantes. Los valores de corte han sido fijados para reproducir los datos experimentales para los receptores de proteínas que presentan más adelante en las Figs. 4 -10.

20 *Función de Puntuación de Afinidad de Unión*

Para convertir el esquema de clasificación de la estructura de anillo en una función de puntuación que proporciona un componente de la afinidad de unión predicha, los autores de la invención utilizan una serie de hipótesis físicas relativas a cómo cada uno de los tipos de estructura interactúa con el receptor.

Los autores de la invención enumeran las hipótesis físicas claves de la siguiente manera:

25 (1) Un término par átomo-átomo lipofílico típico tal como se utiliza en las funciones de puntuación Glide XP, Glide SP o GOLD (descritas en detalle en la solicitud de patente de EE.UU. número de serie 11/373.684 (US 2007/061118), que surge de estructuras expuestas a disolventes sobre-estima la ganancia de energía libre asociada a este tipo de estructuras. Puesto que ninguna cara de la estructura se encuentran protegidos frente al agua, ni es cualquiera de las superficies del receptor, a través de contactos lipofílicos, no hay ninguna razón para creer que la energía libre global de las moléculas de agua en el sistema se han hecho más negativa (ha disminuido). Dicha disminución de la energía libre es implícitamente el origen del término par átomo-átomo lipofílico. Surgen contribuciones favorables a la afinidad de unión de las estructuras expuestas a disolvente, porque el término típico par átomo-átomo lipofílico tiene una "cola" relativamente larga que se extiende 0,4 - 0,6 Nm (4 - 6 Angstroms) de los átomos del anillo, que pueden capturar una sustancial puntuación de afinidad de unión de los átomos de los receptores que no están en contacto de van der Waals con el ligando. Esta puntuación puede tener sentido para las estructuras de superficie o enterradas, pero los resultados para las estructuras expuestas a disolventes serán erróneos, en base a los datos de los autores de la invención. Una forma de superar esta sobre-estimación es simplemente para eliminar el término par átomo-átomo lipofílico (u otra contribución hidrofóbica favorable a la afinidad de unión) para cualquier estructura de anillo clasificada como expuesta a disolvente.

40 (2) Inserción de una estructura en un bolsillo cerrado del receptor puede inducir potencialmente energía de deformación en el receptor. En algunos casos, ya sea el receptor o el ligando, o ambos, tiene que ajustarse para optimizar cómo la estructura encaja en el bolsillo. Estructuras rígidas voluminosos tales como estructuras de anillo pueden crear tensión cuando se unen dentro de un bolsillo de este tipo, mientras que otros tipos de estructuras (p. ej., una cadena de hidrocarburo alifático lineal, que tiene una gran flexibilidad para ajustar los ángulos de torsión a un bajo coste energético) pueden responder mucho más fácilmente al confinamiento en el bolsillo cerrado.

(3) Empíricamente, los autores de la invención han encontrado que sólo los complejos de receptor-ligando con una importante contribución a la energía libre de unión de los enlaces hidrógeno neutro-neutro (enlaces hidrógeno entre un grupo de red-neutro en el ligando y un grupo de red-neutro en la proteína) parece manifestarse energía de deformación inducida de la hipótesis (2). Por ejemplo, los complejos con poca o ninguna energía de enlace

hidrógeno neutro-neutro rara vez, o nunca, parecen manifestarse como energía de deformación tal como se estima por la diferencia entre una función de puntuación que no incluya la tensión y la afinidad de unión experimental.

Esta observación, que une la tensión a la energética de unión de hidrógeno neutra-neutra tiene una base física intuitiva. A diferencia de los puentes de sales o puentes de hidrógeno de carga neutra, enlaces hidrógeno neutros-neutros requieren geometrías precisas para lograr interacciones electrostáticas decentes, debido a la naturaleza dipolo-dipolo de la electrostática, que cae rápidamente a medida que la distancia y/o configuración angular se vuelve no óptima. Otras funciones de puntuación intentan medir esta atenuación, pero no tienen en cuenta las situaciones en donde el receptor paga el costo de la energía mediante la reorganización para permitir que se construyan de manera óptima las geometrías de enlaces hidrógeno. La presencia de estructuras de anillo enterradas añade limitaciones geométricas adicionales. La satisfacción de estas limitaciones, al mismo tiempo que las requeridas para los enlaces hidrógeno neutros-neutros fuertes parece imponer dificultades en el complejo que sólo puede escapar de expender energía libre. La función de puntuación que se describe más adelante emplea un esquema de corrección empírico simple, en el que están vinculados entre sí tensión inducida y energía de enlace hidrógeno neutra-neutra.

Los autores de la invención muestran a continuación dos ejemplos de complejos del PDB que contiene un número sustancial de estructuras de anillo enterradas. El primero (Fig. 1) es el complejo 1aq1 del PDB que es un complejo entre un ligando (mostrado en la Figura) y la proteína CDK2. En este complejo hay un número significativo de enlaces hidrógeno neutro-neutro, incluyendo un "par especial" de enlaces hidrógeno que, en la función de la puntuación Glide XP, se asigna una cantidad sustancial de energía de enlace hidrógeno neutro-neutro adicional en comparación con enlaces hidrógeno ordinarios. La energía de enlace de hidrógeno neutro-neutro total es de más de 5 kcal/mol. El segundo (Fig. 2) es el complejo 1d3d del PDB que es un complejo entre un ligando (mostrado en la Figura) y la trombina de la proteína. Este complejo tiene una energía de enlace hidrógeno neutro-neutro muy baja, inferior a 0,5 kcal/mol.

Cuando es anclado y puntuado por Glide XP, el complejo 1aq1 está anclado con precisión, pero la puntuación está significativamente sobredimensionada en comparación con el experimento. La energía de deformación inducida por los anillos enterrados puede explicar este sobredimensionamiento (en ~ 6 kcal/mol), y la parametrización adecuada puede traer la afinidad de unión predicha de 1aq1 en buena concordancia con los experimentos. Para el complejo 1d3d, Glide XP hace de nuevo un excelente trabajo de anclaje. Sin embargo, la afinidad de unión predicha a partir de Glide XP está en buen acuerdo con el experimento (dentro de 1,0 kcal/mol) sin ningún término energía de deformación del anillo inducida. Si un anillo de energía de deformación inducida se introdujera en una escala comparable a la requerida para 1aq1, el acuerdo de la afinidad de unión predicha de 1d3d con el experimento podría empeorar significativamente, con errores de ~3-5 kcal/mol.

Una solución sencilla al dilema anterior es suponer que correcciones en la tensión del anillo sólo pueden imponerse en la medida en que existe una energía de enlace hidrógeno neutro-neutro que puede ser disminuida por los efectos de deformación hipotéticos. El mecanismo para imponer esta condición, una tapa simple en el anillo de energía de deformación calculada a partir de la energía total de enlace hidrógeno neutro-neutro se comenta más adelante.

(4) Los autores de la invención suponen que las estructuras de la superficie no implican a los fenómenos descritos anteriormente; es decir, su puntuación de par lipofílico es apropiada (que, después de todo, están blindando una cara del anillo, y una porción de tamaño similar de la proteína, frente al agua), y que no inducen la energía de deformación que interrumpe los enlaces hidrógeno (ya que la falta de la envoltura permite el ajuste de bajo costo del anillo para dar cabida a la superficie de la proteína, y viceversa). Por lo tanto, no se hacen correcciones para estructuras asignadas a esta categoría.

Para el montaje de la función de puntuación final, se añade un tipo adicional de término. El modelo para la energía de deformación descrito anteriormente se ocupa de la tensión genérica inducida mediante la inserción de una estructura voluminosa y rígida en un bolsillo cerrado. Sin embargo, también hay reorganizaciones de proteínas específicas necesarias para permitir a un tipo particular de ligando unirse a un receptor particular. Ejemplos incluyen la apertura del bolsillo alostérico en el VIH-RT para acomodar inhibidores de NNRTI, el movimiento del anillo de fenilo en CDK2 para permitir la unión de hidrógeno en la región de bisagra y el movimiento del bucle de activación en p38 map quinasa para permitir la unión de inhibidores DFG-out. En general, el coste de reorganización es claramente específico para el receptor; sin embargo, se puede esperar un penalización de la energía libre de "reorganización del núcleo" en otros casos también, que no sólo es específica para el receptor, sino que es específica para una clase de ligandos de unión al receptor. En principio se podría calcular los términos de reorganización de simulación atomística. Alternativamente, también se pueden tratar estos términos de reorganización como parámetros ajustables específicos para la clase de receptor y ligando (en los casos en los que

la clase de ligando se define por el tipo de conformación del ligando inducida en el receptor), para ser determinada empíricamente o por otro método. Los autores de la invención definen la energía de reorganización del núcleo para un ligando y receptor dados, pero computarizada como ECR.

- 5 Si todos los complejos de receptor-ligando para un receptor dado exhiben el mismo modo de unión al "núcleo", entonces la energía de reorganización del núcleo no tiene efecto alguno en las energías libres de unión relativas (ya que es idéntica para cada uno de los ligandos) y, por lo tanto, no tiene ningún efecto en la jerarquización. Por otro lado, si se quiere comparar ligandos con diferentes modos de unión al núcleo (p. ej., en inhibidores de DFG-in y DFG-out para p38 map quinasa), se asignan valores de reorganización del núcleo a las diferentes configuraciones de bucle. Tales asignaciones se pueden hacer mediante ajuste a unos pocos puntos de datos experimentales;
- 10 alternativamente, se podrían estimar los valores relativos de la reorganización del núcleo a partir de los receptores relacionados en un conjunto de datos disponibles, por ejemplo, el PDB. Por ejemplo, los autores de la invención utilizan el mismo modelo de reorganización del núcleo cuando un grupo con carga positiva se coloca en el bolsillo S1 tanto de la trombina como del factor Xa, mientras que los aglutinantes "inversos" en ambos sistemas tienen un parámetro diferente. El requisito de suministrar un modelo de reorganización del núcleo significa que la función de puntuación que los autores de la invención describen en esta memoria no es estrictamente un modelo "global"; se necesita una cierta cantidad mínima de "local", receptor y de entrada específica para la clase de ligando, al menos si el sistema en estudio tiene múltiples modos de unión al núcleo. Sin embargo, esto representa un pequeño número de parámetros, en los que se puede estimar al menos empíricamente (si es necesario) en base a las estructuras asociadas con los motivos de unión del núcleo.
- 15
- 20 Lo que sigue es un protocolo paso a paso para el cálculo de los componentes de la afinidad de unión ligando-receptor que puede derivarse de la metodología de clasificación del anillo descrita anteriormente:
- (1) Cada una de las estructuras de anillo se clasifica como se discutió anteriormente. Las matemáticas detalladas del sistema de clasificación se presentan en el Anejo.
- 25 (2) La contribución del par átomo-átomo lipofílico para estructuras de anillo expuestas al disolvente, designado ESOR, representa una reducción (por el valor ESOR) en la afinidad de unión ligando-receptor, suponiendo que un término favorable de este tipo haya sido calculado en un diferente componente de la función de puntuación. Alternativamente, se puede desde el principio eliminar cualquier contribución a la energía de interacción hidrofóbica de estructuras de anillo expuestas a disolvente.
- 30 (3) A una estructura (simple) de anillo no condensado se la asigna una "puntuación de estructura de anillo" de 1 si está enterrado, y se la asigna una puntuación de 0 en caso contrario. Para las estructuras condensadas, a la sub-estructura enterrada con la mayor puntuación de envoltura se la asigna una puntuación de estructura de anillo de 1; a otras sub-estructuras de anillo que se clasifican como enterradas se les asigna un valor de 0,5; a sub-estructuras expuestas a superficie o disolvente se les asigna un valor de cero. Las puntuaciones de todos los anillos del anillo se suman luego para producir una puntuación total de anillo R para la molécula.
- 35 (4) La puntuación de anillo R se ajusta si es necesario tener en cuenta la participación de la estructura de anillo en la energía de reorganización del núcleo ECR. En la actual mejor aplicación práctica de los autores de la invención, si una estructura de anillo está implicada específicamente en la reorganización del núcleo, la contribución de ese anillo R se resta de R. Por ejemplo, si un solo anillo enterrado está implicado en la reorganización del núcleo, el valor de R es revisado a R-1.
- 40 (5) La energía de enlace hidrógeno (EHB) neutro-neutro total se calcula incluyendo las contribuciones de los términos de enlace hidrógeno neutro-neutro "especiales" si la función de puntuación Glide XP se utiliza para asignar la energía de enlace hidrógeno neutro-neutro. Condiciones especiales de enlace hidrógeno neutro-neutro en la función de puntuación Glide XP se definen en la solicitud de patente de EE.UU. número de serie 11/373.684 (US 2007/061118),
- 45 (6) El ligando se clasifica como perteneciente a un motivo de unión particular y se le asigna una energía de reorganización del núcleo EC. Como se ha indicado anteriormente en (4), si una estructura de anillo en el ligando está involucrada de forma explícita en el motivo de unión del núcleo (como en CDK2), a continuación, esta estructura está en última instancia no incluida (p. ej., su contribución se resta de) la puntuación de estructura R para evitar un recuento doble.

(7) La energía de deformación ER total del anillo se calcula como $ER = R \cdot A$ kcal/mol. El parámetro A puede determinarse empíricamente o por ensayo y error. En algunas implementaciones, $0,5 \text{ kcal/mol} \leq A \leq 2,5 \text{ kcal/mol}$.

5 (8) La cantidad máxima de energía de enlace hidrógeno que puede verse afectada por la deformación del anillo del ligando, las energía de remate del enlace hidrógeno HBC se estima como $HBC = (EHB - B) \cdot C$. El parámetro B se puede determinar empíricamente o por ensayo y error. Este parámetro refleja el fenómeno de que la tensión se vuelve importante cuando múltiples enlaces hidrógeno se han de mantener en buenas posiciones geométricas. En algunas implementaciones, $0,5 \text{ kcal/mol} \leq B \leq 2 \text{ kcal/mol}$. El parámetro C es un factor de escala que mide la máxima fracción de la energía de enlace hidrógeno restante que se puede perder debido a la deformación. Este parámetro también se puede determinar empíricamente o por ensayo y error. En algunas implementaciones, $0,5 \leq C \leq 1,0$.

10 (9) Si ER es mayor que HBC, ER se sustituye por HBC. De lo contrario, ER permanece como computarizada.

(10) Al ligando se le asigna una energía de reorganización al núcleo ECR basada en el modo de unión del receptor y del ligando tal como se discutió anteriormente.

(11) La contribución total a la afinidad por la exposición a disolventes, la reorganización del núcleo y la energía de deformación de unión, como se define en el protocolo anterior, es entonces $ER + ECR - ESOR$.

15 (12) Los términos anteriores representan sólo una parte de la afinidad de unión al ligando total calculada. Para calcular la afinidad total de unión al ligando, los términos en (11) se añaden a una función de puntuación que representa estos otros términos tales como la función de puntuación Glide XP. Si el valor de la afinidad de unión definido por esta otra función de puntuación es E0, la afinidad de unión total después de la contribución de la exposición a disolventes, la reorganización del núcleo y la energía de deformación es entonces $E0 + ER + ECR - ESOR$.

20

Como una demostración inicial de la utilidad de la invención, los autores de la invención han aplicado su mejor práctica actual del protocolo (como se describe en 1-12 arriba y en el Anejo) a 7 receptores del PDB. Para los receptores del PDB realizaron un auto-anclaje para todos los complejos proteína-ligando relevantes derivados de ese receptor (excepto para el caso de ns5b, en donde utilizaron solamente el anclaje de ligandos en un sitio alostérico específico del receptor), llevaron a cabo un auto-anclaje con la metodología Glide XP, y eliminaron complejos en los casos en los que el ligando se ancló sustancialmente de forma errónea por Glide XP (que corresponde en general a menos del 5% de todos los complejos). Para los complejos restantes, los autores de la invención modificaron la puntuación Glide XP como se define en (12) anterior, utilizando los parámetros $A = 1,5$ kcal/mol, $B = 1,0$ kcal/mol y $C = 0,75$, que constituyen sus actuales mejores valores en la práctica, y parámetros de reorganización del núcleo para cada uno de los receptores según fuese apropiado (no se utilizaron más de dos de estos parámetros para cualquier receptor). Los resultados de estos cálculos de afinidad de unión del complejo se comparan con datos experimentales (para todos los complejos para los que estaba disponible) en las Figs. 4-10. Las Figs. 4-10 representan una afinidad de unión experimental frente a la afinidad de unión calculada, pero también presentan el coeficiente de correlación R^{**2} para la representación y la desviación absoluta media (MAD) de la representación. El intervalo de valores MAD (0,75 - 1,24 kcal/mol) y los valores R^{**2} (0,47-0,62) demuestran un alto grado de consistencia y representan una importante utilidad práctica en la discriminación de compuestos con diversos niveles de actividad. Considerando que el nuevo modelo es una función global de puntuación, con 1-2 parámetros específicos para receptores (sólo el segundo de los cuales afecta a la MAD y R^{**2}), los autores de la invención piensan que se ha logrado un avance significativo en comparación con las funciones de puntuación globales alternativas en la bibliografía.

25

30

35

40

Apéndice: Puntuación de Reorganización del Anillo

Este anejo describe los detalles del algoritmo de puntuación de reorganización del anillo.

A1) Puntuación de empaquetamiento del Anillo

45 La puntuación de empaquetamiento del anillo R_score de un anillo dado o componente de un anillo condensado está diseñado para describir tanto la envoltura tridimensional del anillo por parte de la proteína, como la proximidad del anillo a la proteína. Cuantitativamente esto significa que R_score es una función de la puntuación de la envoltura del anillo y la puntuación lipofílica del anillo.

La puntuación de la envoltura del anillo se deriva de la función de la envoltura XP-fóbica [] con la modificación de que a todos los átomos pesados de las proteínas se les permite contribuir en la envoltura en lugar de la clase fóbica de los átomos. Esto conduce a una puntuación de la envoltura r_{encl_i} para cada uno de los átomos i del anillo.

5 La puntuación lipofílica e_{lip_i} de un átomo i se deriva de la energía de par lipofílico e_{pair_i} (ref) y la energía del par vdw atómico e_{vdw_i}

$$e_{lip_i} = e_{pair_i} + e_{vdw_i}$$

La puntuación neta del anillo R_{score} que describe los efectos de la envoltura del anillo (r_{encl}) y la proximidad a la proteína (e_{lip}) es simplemente la suma,

$$10 \quad R_{score} = \text{Sum}_i e_{lip_i} + r_{encl_i}$$

A2) Puntuación Neta del Anillo

15 El objetivo del algoritmo descrito aquí es asignar un anillo dado a una puntuación 'condensada' o 'neta' ' R_{net} ' que clasifica su nivel de enterramiento utilizando R_{score} y descriptores adicionales que proporcionan un estrecho contacto entre el anillo y la proteína. Esta puntuación se utiliza en el cálculo de la energía de reorganización neta descrita en Sec.zz.

A2a) Esquema de Clasificación del Anillo

20 Un anillo dado se clasifica en 5 niveles de contactos con la proteína que oscilan entre contacto bajo (1) y alto (5). Estos niveles se determinan por el número de átomos de la proteína que contactan con el anillo en diferentes intervalos. Los contactos proteína - átomos que no son hidrógeno se han agrupado en intervalos de distancia r_c bin-1) $r_c < 4,0$, bin-2) $4,0 < r_c < 4,5$ y bin-3) $r_c > 4,5$. Además, los contactos se reparten en las dos caras del anillo con el que se hace contacto con "arriba" o 'abajo'. Marcar la combinación de las etiquetas de contacto bin/cara con las matrices $icup(1-3)$ e $icdown(1-3)$ por cada anillo. Definir adicionalmente los 'contactos estrechos' como $icup(1)$ + $icdown(1)$ y los 'contactos integrales' como $icup(1-3)$ + $icdown(1-3)$. Entonces los 5 grupos se pueden definir como:

Grupo 1: no hay contactos estrechos, número total de contactos menor que o igual a 3.

25 Grupo 2: 1 contacto estrecho, número total de contactos menor que o igual a 4, y no en el grupo 3.

Grupo 3: dos contactos estrechos, algo en cada lado, o un contacto estrecho, algo en cada lado, y al menos 4 contactos totales.

Grupo 4: Al menos 4 contactos en un lado, pero cero contactos en el otro lado.

Grupo 5: 5 o más contactos, algo en ambos lados o 3 contactos estrechos, algo en ambos lados.

30 A2b) Puntuación Neta del Anillo:

35 Un anillo dado se asigna a uno de los grupos de contacto (1-5) anteriores. Teniendo en cuenta esta asignación, una puntuación neta del anillo para el anillo i , ' R_{net_i} ' se define como 0,0, 0,5 ó 1,0. Esta asignación es una función del grupo, la puntuación del anillos R_{score} de arriba, y las 1ª + 2ª aguas totales de la envuelta en contacto con el anillo N_{water} . El R_{score} se asigna a la puntuación neta R_{net} utilizando dos puntos de corte específicos de grupos rg_1 y rg_2 , un par por cada uno de los grupos $g = 1-5$. El mapa es simplemente

$$\begin{aligned} \text{Si } R_{score} < rg_1 \quad R_{net_i} &= 0,0 \\ \text{Si } rg_1 < R_{score} < rg_2 \quad R_{net_i} &= 0,5 \\ \text{Si } R_{score} > rg_2 \quad R_{net_i} &= 1,0 \end{aligned}$$

40 A anillos de tamaño tres siempre se les asigna una puntuación de cero. Si el anillo es parte de un anillo condensado, entonces R_{net} es 1/2 de los valores anteriores (es decir, 0,0, 0,25 ó 0,5). Si N_{water} es menor que o igual a 10 y

$R_score > 2.1$, entonces al R_net_i completo se asigna al anillo (1,0 ó 0,5 para los anillos secundarios de un anillo condensado).

A3) Energía de Reorganización del Anillo

5 Como se resume en la Sec. XX la energía de reorganización del anillo, E_R , se compone de una energía de reorganización del núcleo E_rcore y una energía de deformación del anillo E_ring . En segundo lugar, E_ring tiene un enlace hidrógeno que cuelga de E_hbcap ,

$$E_R = E_rcore + \min(E_hbcap, E_ring)$$

La energía de deformación del anillo E_ring es una función de R_net , la suma de las puntuaciones de anillo R_net_i (eq.)

10
$$R_net = \text{Sum}_i R_net_i.$$

La energía de deformación del anillo E_ring es simplemente

$$E_ring = 1.5(R_net-1) \quad R_net > 1; \quad E_ring = 0 \quad R_net < 1$$

en donde la resta de 1 representa la reorganización del núcleo.

15 La energía de reorganización del núcleo es habitualmente una constante dependiente del sistema. La excepción es en casos tales como CDK2 o trombina en la que si un ligando se une sin un motivo de reconocimiento conocido (en especial enlaces H encerrados fóbicamente en CDK2 o puente de sal en trombina), entonces E_rcore se pone a cero.

El remate del enlace H es una cantidad que depende del sistema de la forma

$$E_hbcap = 0,75(HB_neut + a*HB_phob+ b*HB_phob_pr)$$

20 siendo HB_neut la puntuación del enlace hidrógeno neutro-neutro, excluyendo los grupos hidroxilo del ligando, HB_phob la puntuación de enlaces H simples encerrados fóbicamente (ref), y HB_phob_pr la puntuación de pares de enlace H encerrados fóbicamente (usar nomenclatura de papel XP), a y b son constantes, a es 1 ó 2 en los sistemas discutidos en este documento y b es cero, excepto para CDK2 en donde es 1. Para la trombina y Fxa E_hbcap tiene una forma diferente cuando no hay puente salino, entonces es sólo $E_hbcap = 0,5*(HB_neut - 1.0)$.

25 A4) 2) Corrección de la Exposición a Disolventes

Se observó en el trabajo XP [] que los términos par lipofílico chemscore sobre-estima contribuciones de unión de átomos del ligando expuestos a disolventes. Tales átomos del ligando expuestos a disolventes puede ser parte de un grupo de ligandos que tiene un lado frente a la proteína de la otra disolución o puede ser solo un átomo de disolvente expuesto de aproximadamente 4-5Å de unos pocos residuos en la superficie. En este trabajo los autores de la invención han incrementado la corrección XP para esta sobre-unión de la superficie. El término de corrección adicional para las regiones expuestas a disolventes, ESOR es simplemente la puntuación lipofílica y la puntuación vdw (e_lip_r Eq. XX), de un anillo sumada sobre la calificación de anillos expuestos a disolvente

$$ESOR = \text{Sum}_{r=\{rs\}} e_lip_rs$$

35 Téngase en cuenta que los átomos que tienen sus puntuaciones e_lip corregidas por el algoritmo XP () no se cuentan dos veces.

Las reglas para un anillo de calificar implican la etiqueta de grupo de contacto discutida en el esquema de clasificación de los anillos anterior, las aguas de la 1ª + 2ª envuelta totales alrededor del anillo, N_water , y la

ES 2 615 179 T3

puntuación de la envuelta del anillo r_{encl} . Los dos conjuntos de lógica que califican un anillo que necesitan la corrección de la exposición son;

1) el anillo está en los grupos de contacto 1 ó 2, $N_{water} > w1$ y $r_{encl} > 2.1$

5 2) el anillo no está en los grupos de contacto 1 ó 2, tiene cero contactos estrechos de proteínas en una o ambas caras, $N_{water} > w2$ y $r_{encl} > 1.4$

en que $w1$ y $w2$ son los puntos de corte de recuento de agua. Si un anillo satisface estas condiciones, entonces $elip_{rs}$ se añade al término ESOR como en Eq (). Las reglas para las configuraciones $w1$ y $w2$ dependen de la puntuación neta del anillo R_{net} y los contactos estrechos. Hay tres clasificaciones de un anillo que tiene cada clase específica de puntos de corte $w1/w2$. Consúltense la sección XX para la notación de contactos $icup/icdown$.

10 Clase 1: $R_{net} > 2,7$ e $icup$ (1) igual a 1, $icup$ (2) menor que 2, $icup$ (3) menor que 2 y no hay contactos estrechos en el lado 'abajo' / o invertir arriba/abajo

$R_{net} 2 \leq R_{net} \leq 2,7$ $w1 = 90$ $w2 = 90$
 $R_{net} 1,5 \leq R_{net} \leq 2,0$ $w1 = 86$ $w2 = 86$
 $R_{net} < 1,5$ $w1 = 81$ $w2 = 82$

15 Clase 2: $icup$ (1) o $icdown$ (1) no es cero y no en la clase 1

$w1 = 108$
 $w2 = 108$

Clase 3; no clases 1 ó 2; $w1 = 78$ $w2 = 83$

REIVINDICACIONES

1. Un método implementado por ordenador para cuantificar al menos un componente de afinidad de unión entre un ligando y una molécula de receptor, comprendiendo el método:
- 5 recibir datos que representan una estructura del ligando,
 recibir datos que representan una estructura de la molécula de receptor,
 identificar una estructura de anillo dentro del ligando, siendo la estructura de anillo un anillo entero o un anillo condensado;
 clasificar la estructura de anillo en base a la relación parcial de la estructura de anillo en sus entornos ambientes cuando el ligando está unido a la molécula de receptor, en donde dichos entornos ambientes de la estructura de anillo cuando se unen a la molécula de receptor incluyen moléculas de agua;
 10 en donde la estructura de anillo está asignada a una de las siguientes clasificaciones: a) enterrada; b) expuesta a disolvente; o c) expuesta a disolvente en una superficie;
 en donde determinar la clasificación se basa en:
- 15 i) un cálculo del grado de envoltura de cada uno de los átomos de no hidrógeno de la estructura de anillo por átomos no hidrógeno del receptor; y
 ii) el uso de un parámetro que corresponde a un cierto número de moléculas de agua en una primera envuelta de la estructura de anillo, o en una segunda envuelta de la estructura de anillo o en una combinación de la primera envuelta y la segunda envuelta; y
- 20 cuantificar al menos un componente de la afinidad de unión entre el ligando y la molécula de receptor en base a la clasificación de la estructura de anillo, en donde dicha cuantificación comprende calcular una afinidad de unión inicial y luego ajustar la afinidad de unión inicial por un factor obtenido en base a la clasificación de la estructura de anillo.
2. El método de la reivindicación 1, en el que la cuantificación de al menos el componente de la afinidad de unión incluye una etapa que puntúa interacciones hidrofóbicas entre uno o más átomos de ligando y uno o más átomos de receptor, otorgándole un bono en cuanto a la presencia de la envoltura hidrofóbica de uno o más átomos de dicho ligando por dicha molécula de receptor, siendo dicho bono indicativo de una afinidad de unión potenciada entre dicho ligando y dicha molécula de receptor.
- 25 3. El método de la reivindicación 1, en el que se determina el número de contactos estrechos a diferentes distancias entre los átomos de receptor y las dos caras del anillo, se realiza una clasificación inicial del anillo en base a los números de estos contactos y esta clasificación inicial es seguida a continuación por el cálculo de una función de puntuación, dependiendo de otros parámetros que permite la clasificación como enterrada, expuesta a disolvente o expuesta a disolvente en una superficie.
- 30 4. El método de la reivindicación 3, en el que la función de puntuación que permite la clasificación de la estructura de anillo como enterrada, expuesta a disolvente o expuesta a disolvente en una superficie incluye el uso de un parámetro que corresponde a una interacción hidrofóbica de la estructura de anillo o parte del mismo con la molécula de receptor.
- 35 5. El método de la reivindicación 3, en el que la función de puntuación que permite la clasificación de la estructura de anillo como enterrada, expuesta a disolvente o expuesta a disolvente en una superficie incluye utilizar un parámetro sustancialmente correlacionado con la contención del par lipofílico-lipofílico entre la estructura de anillo o parte de la misma y la molécula de receptor.
- 40 6. El método de la reivindicación 1, en el que las estructuras de anillo expuestas a disolvente en el ligando, si existen, se ignoran sustancialmente para cuantificar el componente de la afinidad de unión entre el ligando y las moléculas de receptor más que para reconocer enlaces hidrógeno y otros parámetros que son independientes de la clasificación de la estructura de anillo.
- 45 7. El método de la reivindicación 1, en el que la contribución hidrofóbica a la afinidad de unión a partir de estructuras de anillo clasificadas como expuestas a disolvente, si existen, se ignora sustancialmente para cuantificar el componente de la afinidad de unión.
8. El método de la reivindicación 1, en el que una estructura de anillo se clasifica como enterrada, y el método comprende, además:

identificar una cantidad representativa de una energía de deformación inducida en el complejo ligando-receptor por la estructura de anillo enterrada, en que la cuantificación del componente de afinidad de unión se basa adicionalmente, en parte, en la energía de deformación.

- 5 9. El método de la reivindicación 8, que comprende, además,
identificar una cantidad representativa de una energía de deformación inducida en el complejo ligando-receptor por el agregado de las estructuras de anillo identificadas como enterradas;
identificar una cantidad representativa de una energía de enlace hidrógeno neutra-neutra total; y
cuantificar el componente de la afinidad de unión entre el ligando y la molécula de receptor, basado, al menos en parte, en la cantidad representativa de energía de deformación inducida en el receptor por el agregado de las
10 estructuras de anillo enterradas y la cantidad representativa de la energía de enlace hidrógeno neutra-neutra total.

10. El método de la reivindicación 9, en el que
la cuantificación del componente de la afinidad de unión comprende, además,
identificar una energía de remate de enlace hidrógeno asociada con el ligando entero, y
15 el componente de la afinidad de unión se cuantifica en base a una mayor energía de remate de enlace hidrógeno y la cantidad representativa de la energía de deformación inducida en el receptor por el agregado de las estructuras identificadas.

11. El método de la reivindicación 8, en el que el componente de la afinidad de unión atribuible a la deformación se cuantifica utilizando al menos una de: dinámica molecular, mecánica molecular, búsqueda conformacional y minimización.

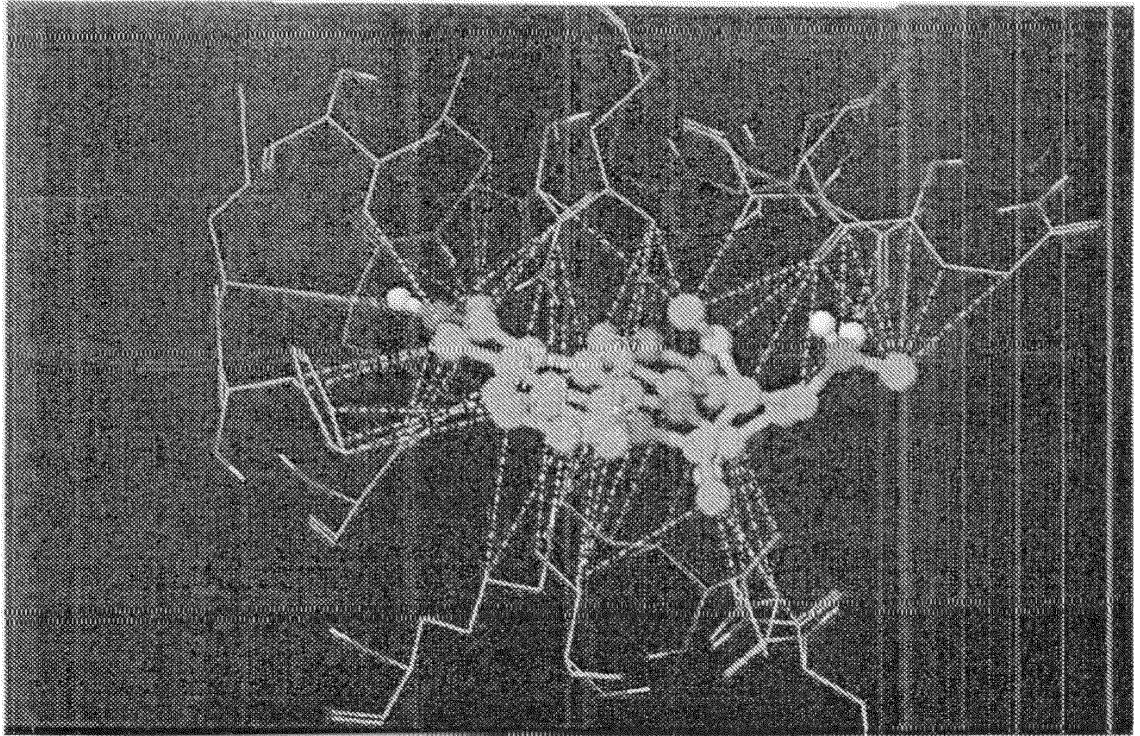


FIG. 1 (CDK2- 1aql)

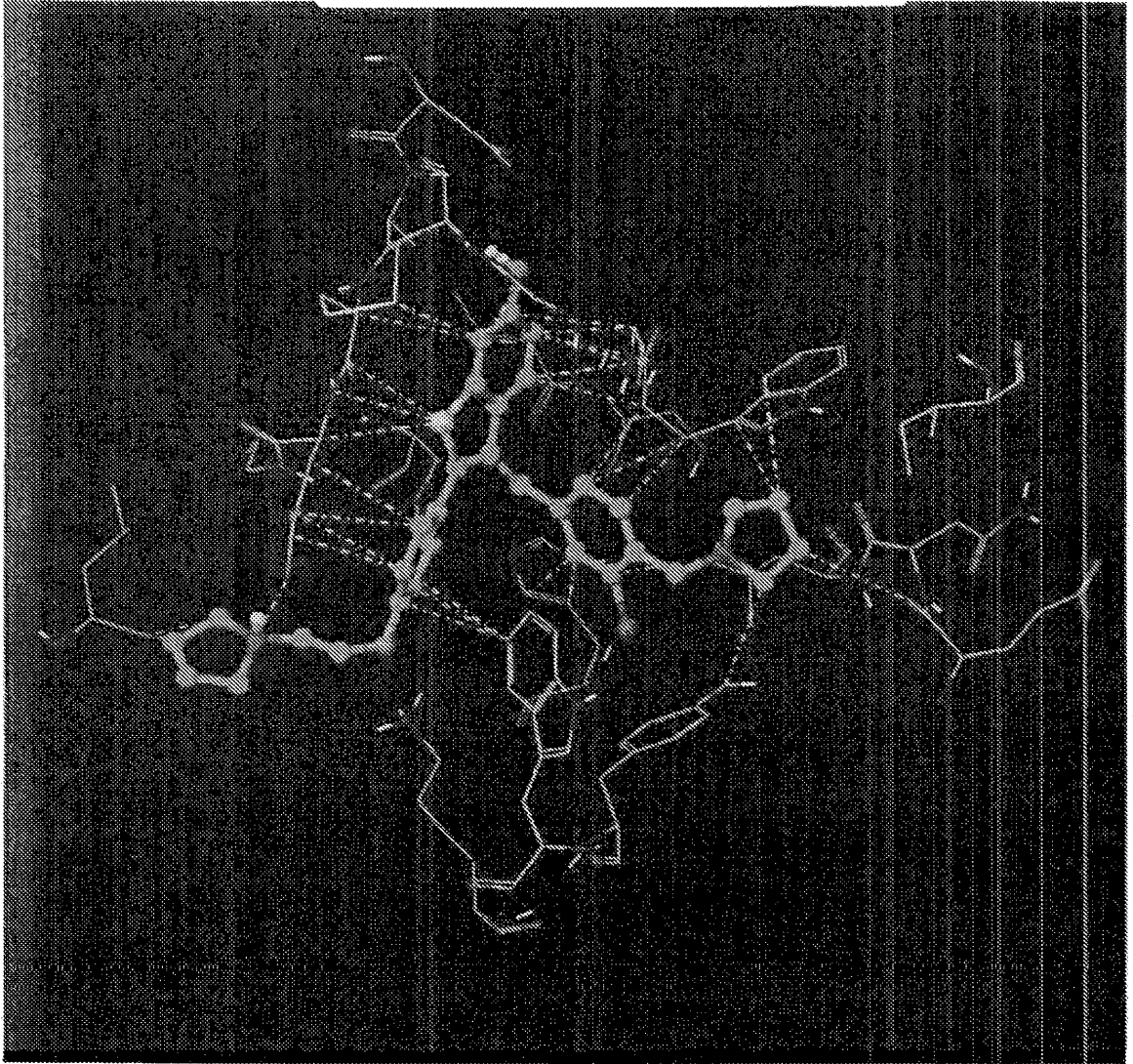


FIG. 2 (1d3d)

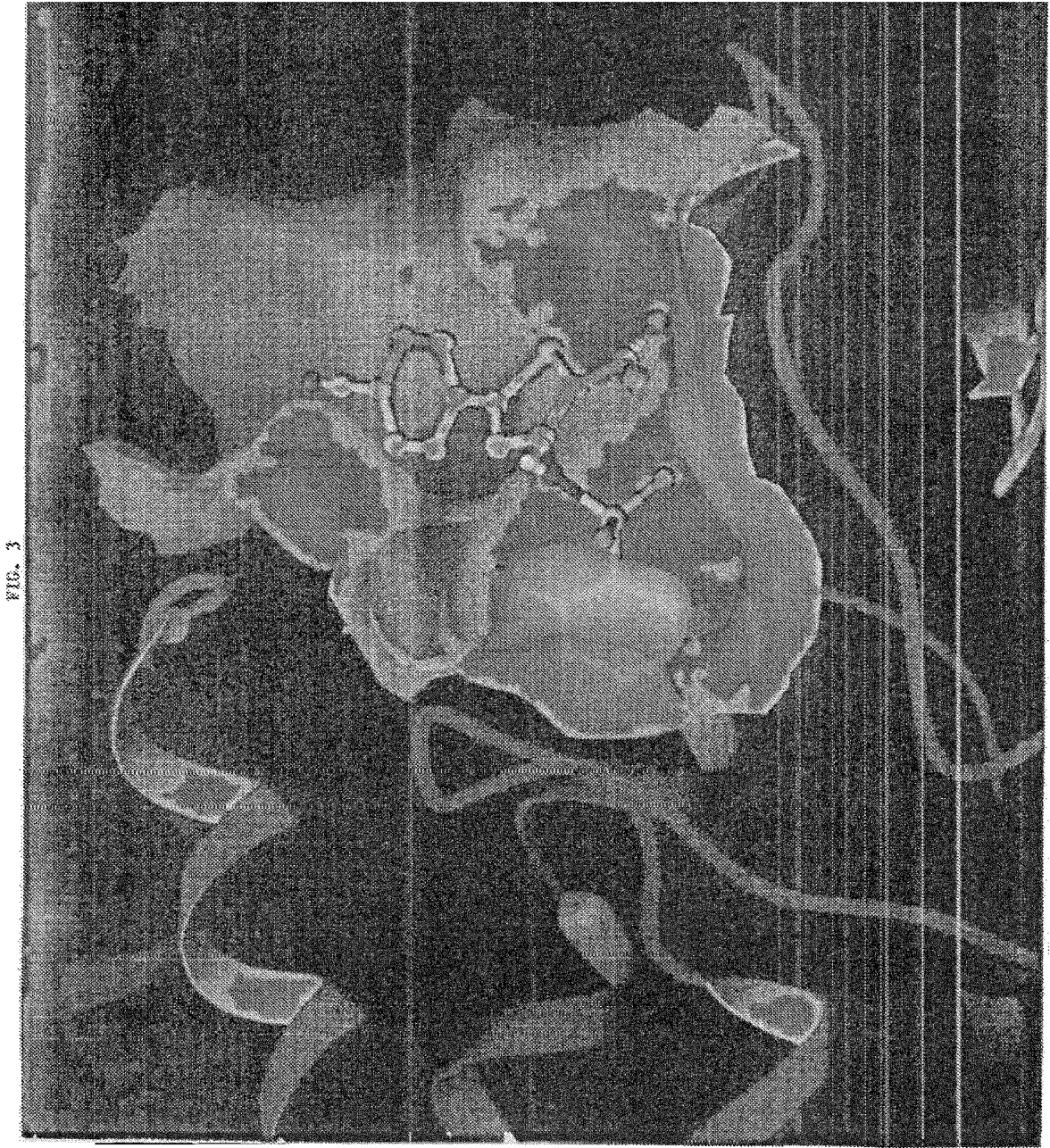
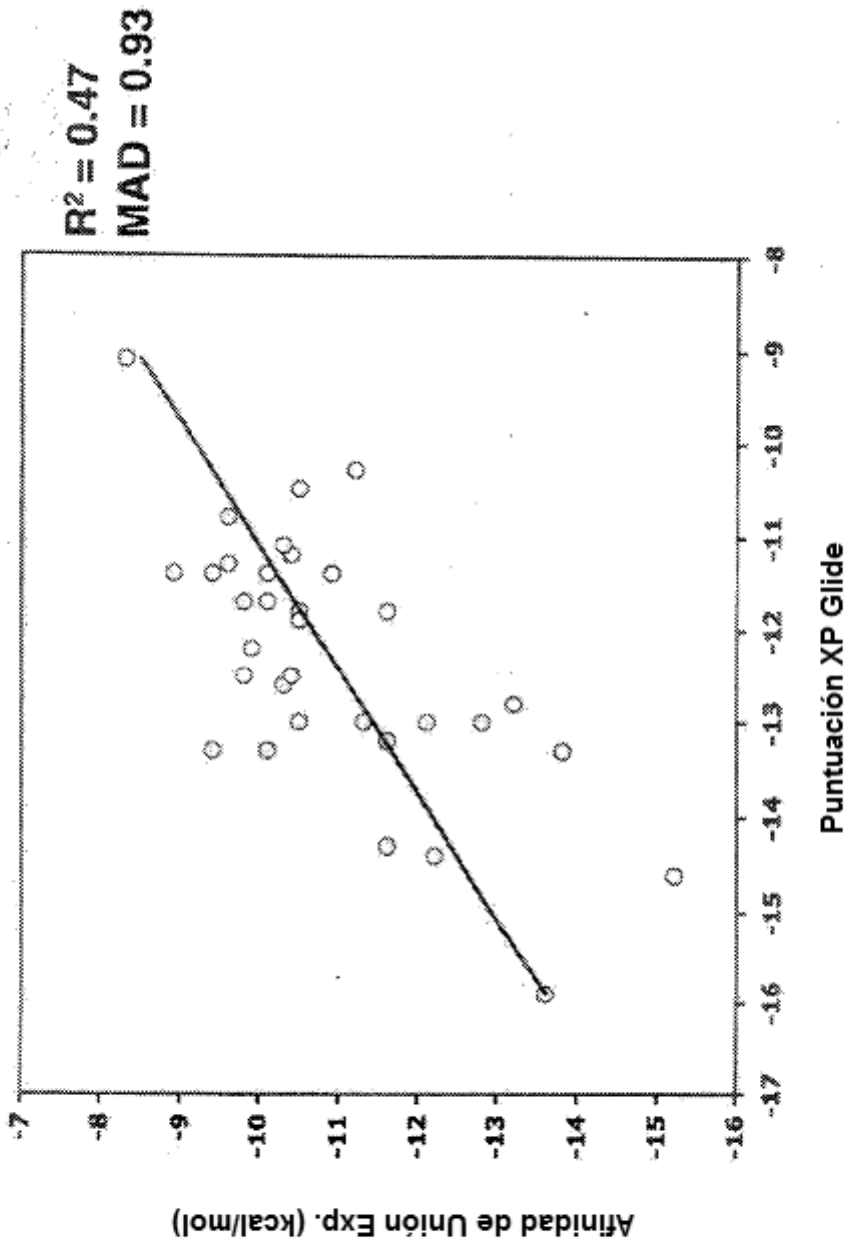


FIG. 3

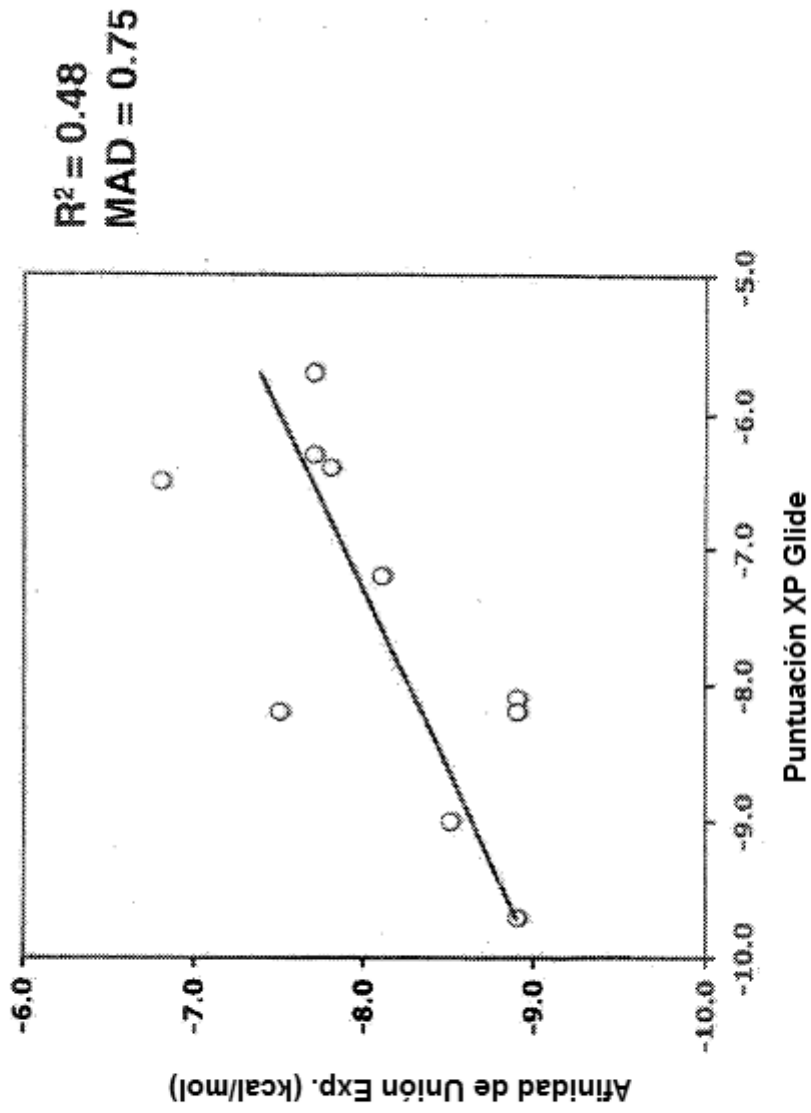
Factor Xa



SCHRÖDINGER

FIG. 4

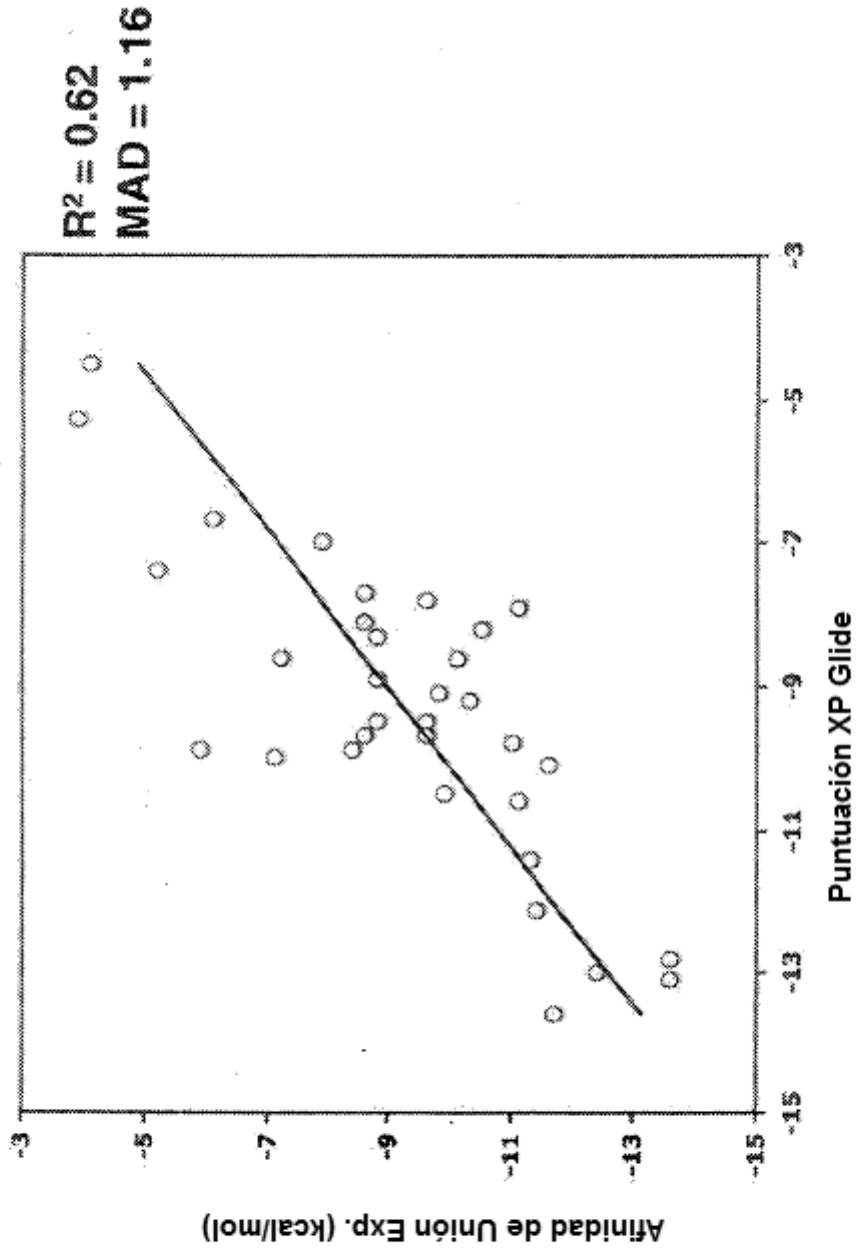
Virus NS5b de la Hepatitis C



SCHRÖDINGER

FIG. 5

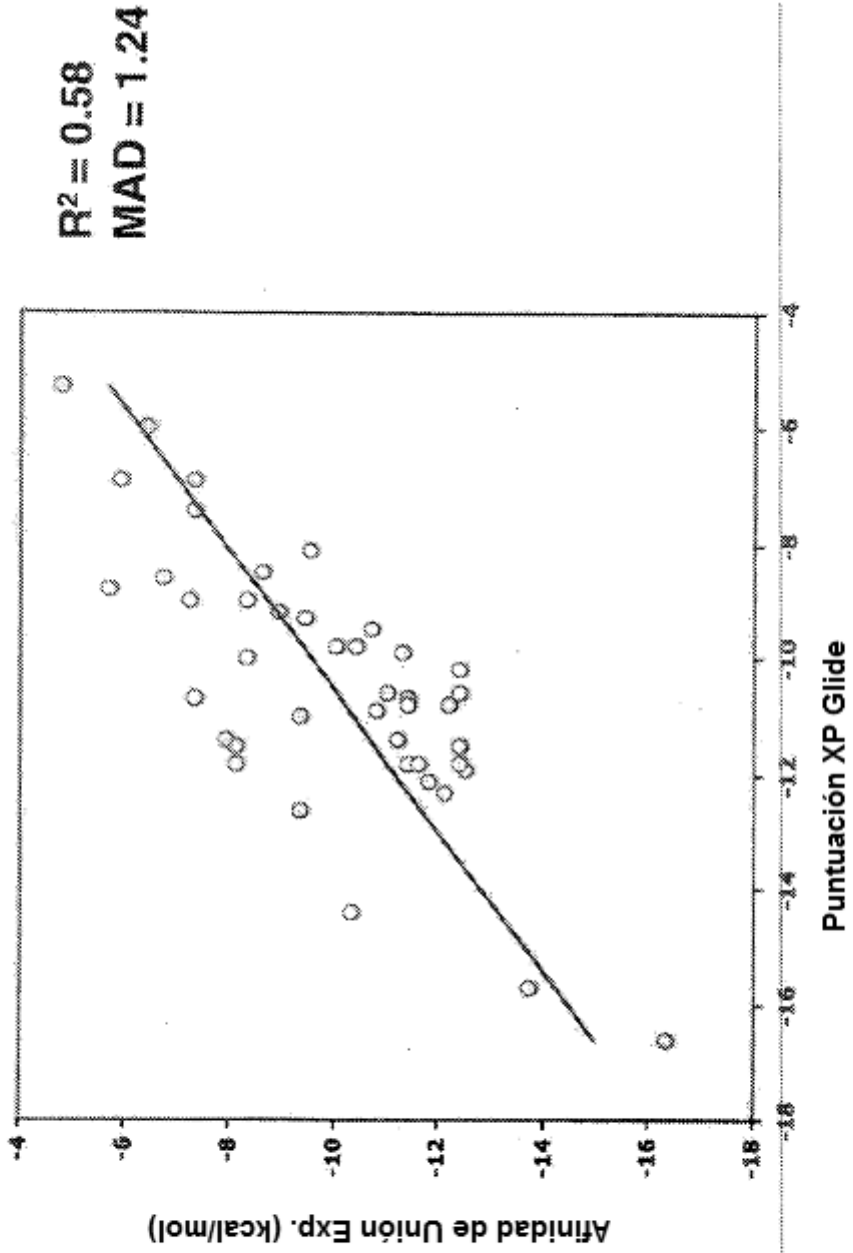
P38 MAP Quinasa



SCHRÖDINGER

FIG. 6

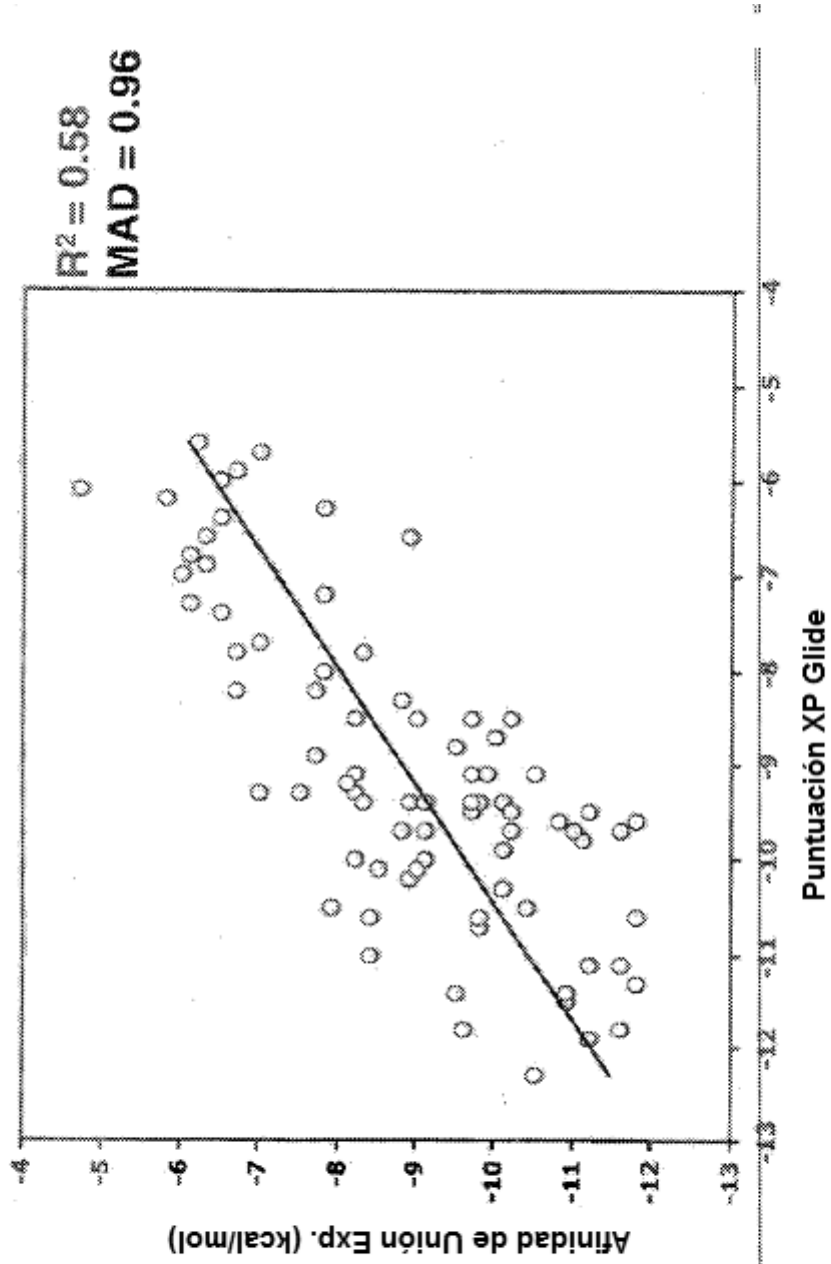
Trombina



SCHRÖDINGER

FIG. 7

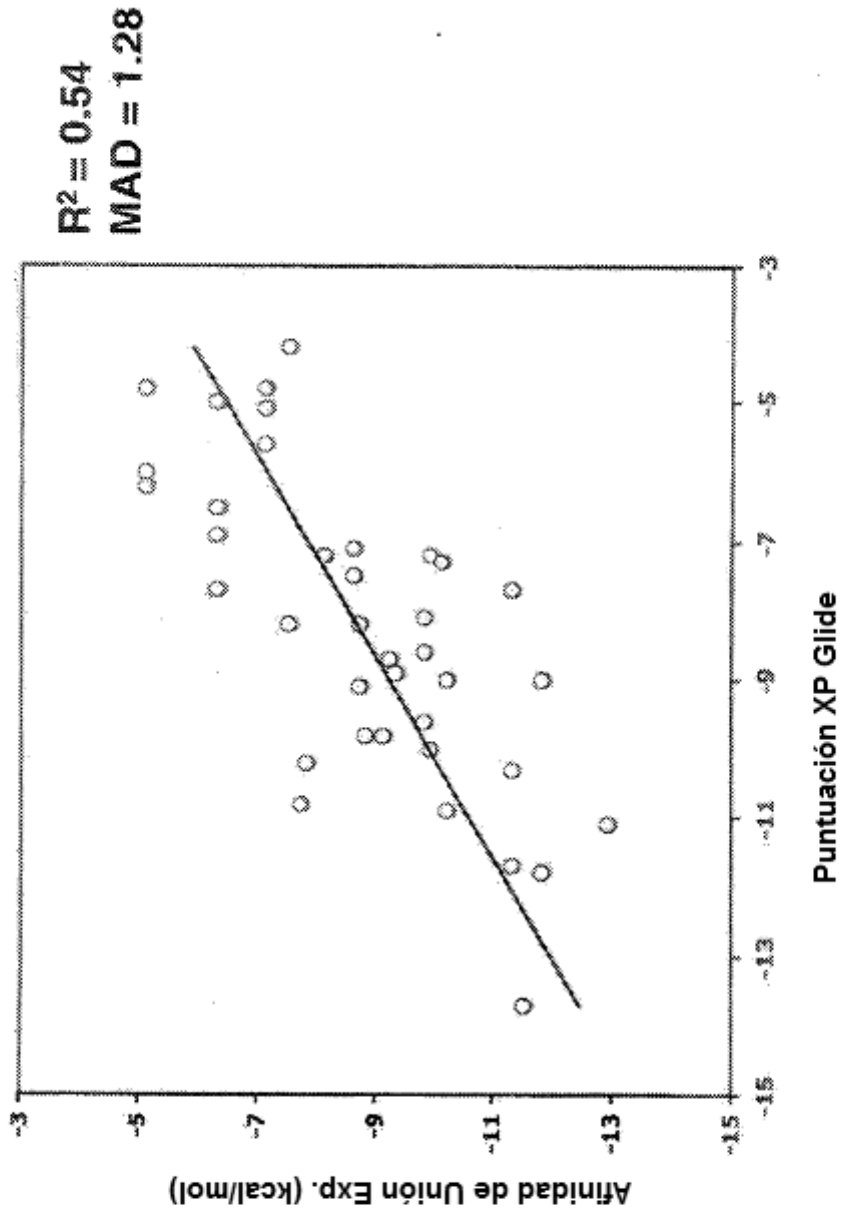
CDK2



SCHRÖDINGER

FIG. 8

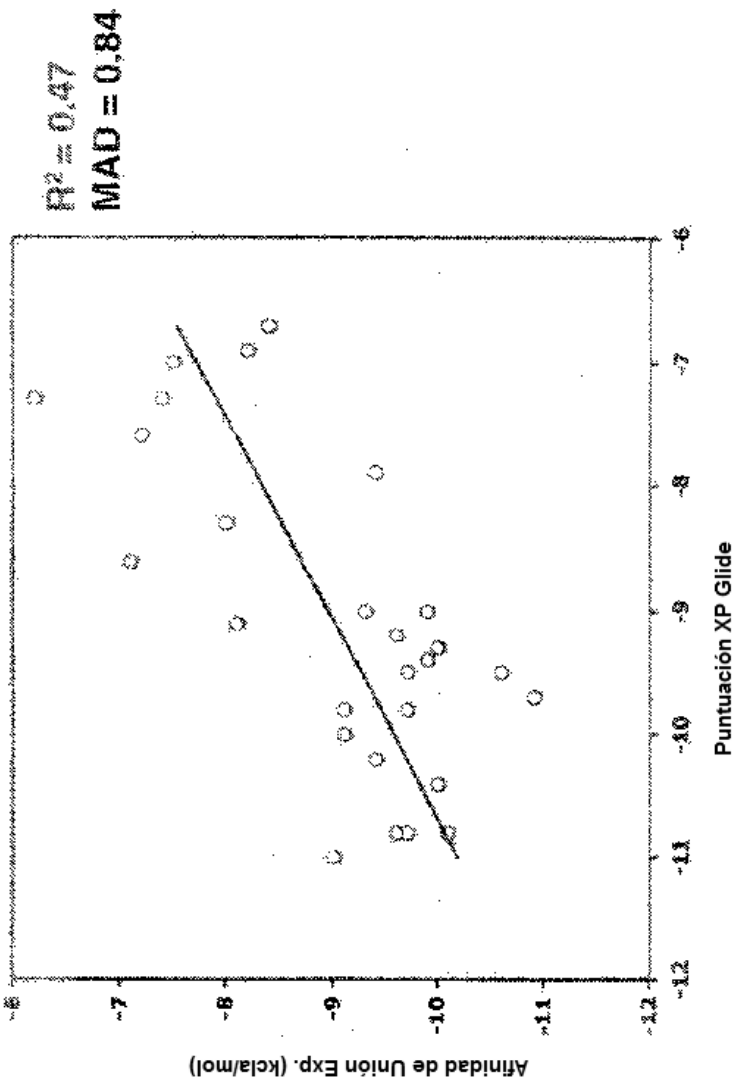
PKA



SCHRÖDINGER

FIG. 9

Oppa



SCHRÖDINGER

FIG. 10