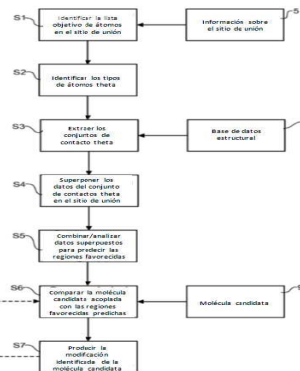


# DIRECCIÓN DE NUEVAS CREACIONES

## SOLICITUD FASE NACIONAL - PCT

Fig. 7



|                                              |                                                                                                                                                                                                                                                           |                                                    |                                                                     |
|----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------|---------------------------------------------------------------------|
| 1                                            | <b>Título de la Invención</b> (200 caracteres o espacios máximos)                                                                                                                                                                                         |                                                    |                                                                     |
| OBTENCIÓN DE UN LIGANDO TERAPÉUTICO MEJORADO |                                                                                                                                                                                                                                                           |                                                    |                                                                     |
| 2                                            | <b>Datos del Solicitante / Titular</b>                                                                                                                                                                                                                    |                                                    |                                                                     |
|                                              | <b>Nombre:</b>                                                                                                                                                                                                                                            | UCB BIOPHARMA SPRL                                 | <b>Dirección Electrónica:</b> clientes@cavelier.com                 |
|                                              | <b>Dirección:</b>                                                                                                                                                                                                                                         | 60 Allée de la Recherche, B-1070 Brussels, Bélgica | <b>Domicilio/País de constitución:</b> BELGICA - BRUSSEL - BRUSELAS |
|                                              | <b>Identificación:</b><br><input type="checkbox"/> CEDULA DE CIUDADANIA<br><input checked="" type="checkbox"/> EMPRESA EXTRANJERA<br><input type="checkbox"/> PASAPORTE<br><input type="checkbox"/> CEDULA DE EXTRANJERIA<br><input type="checkbox"/> NIT |                                                    |                                                                     |
|                                              | <b>Número:</b>                                                                                                                                                                                                                                            | 1021106-                                           |                                                                     |
| 3                                            | <b>Solicitantes</b>                                                                                                                                                                                                                                       |                                                    |                                                                     |

| Apellidos - Nombres o Razón Social                                                                                                                                                                                                                                                                                                         |                                 | Tipo                                                                                      | Identificación                                               |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------|
| 1.                                                                                                                                                                                                                                                                                                                                         | UCB BIOPHARMA SPRL              | EE                                                                                        | 1021106                                                      |
| 4                                                                                                                                                                                                                                                                                                                                          | Datos del Inventor              |                                                                                           |                                                              |
| Nombre:                                                                                                                                                                                                                                                                                                                                    |                                 | Jiye SHI                                                                                  | Dirección Electrónica: clientes@cavelier.com                 |
| Dirección:                                                                                                                                                                                                                                                                                                                                 |                                 | c/o UCB Biopharma SPRL, IP Department, 60 Allée de la Recherche, B-1070 Brussels, Bélgica | Domicilio/País de constitución: BELGICA - BRUSSEL - BRUSELAS |
| Identificación:                                                                                                                                                                                                                                                                                                                            |                                 |                                                                                           |                                                              |
| <div><div><input type="checkbox"/> CEDULA DE CIUDADANIA</div><div><input type="checkbox"/> EMPRESA EXTRANJERA</div><div><input type="checkbox"/> PASAPORTE</div></div> <div><div><input type="checkbox"/> CEDULA DE EXTRANJERIA</div><div><input type="checkbox"/> NIT</div><div><input checked="" type="checkbox"/> No Aplica</div></div> |                                 |                                                                                           |                                                              |
| Número:                                                                                                                                                                                                                                                                                                                                    |                                 | -                                                                                         |                                                              |
| 5                                                                                                                                                                                                                                                                                                                                          | Inventor(es)                    |                                                                                           |                                                              |
| Apellidos - Nombres                                                                                                                                                                                                                                                                                                                        |                                 | Domicilio                                                                                 |                                                              |
| 1.                                                                                                                                                                                                                                                                                                                                         | SHI Jiye                        | BELGICA                                                                                   |                                                              |
| 2.                                                                                                                                                                                                                                                                                                                                         | BAKER Terence Seward            | REINO UNIDO DE LA GRAN BRETAÑA                                                            |                                                              |
| 3.                                                                                                                                                                                                                                                                                                                                         | LAWSON Alastair David Griffiths | REINO UNIDO DE LA GRAN BRETAÑA                                                            |                                                              |
| 4.                                                                                                                                                                                                                                                                                                                                         | LIU Xiaofeng                    | REINO UNIDO DE LA GRAN BRETAÑA                                                            |                                                              |
| 6                                                                                                                                                                                                                                                                                                                                          | Datos Inventor(es)              |                                                                                           |                                                              |
| País de Residencia                                                                                                                                                                                                                                                                                                                         |                                 | Departamento/Estado                                                                       | Ciudad                                                       |
|                                                                                                                                                                                                                                                                                                                                            |                                 |                                                                                           | Dirección                                                    |

|    |                                |            |                   |                                                                                                |
|----|--------------------------------|------------|-------------------|------------------------------------------------------------------------------------------------|
| 1. | BELGICA                        | BRUSSEL    | BRUSELAS          | c/o UCB Biopharma SPRL, IP Department, 60 Allée de la Recherche, B-1070 Brussels, Bélgica      |
| 2. | REINO UNIDO DE LA GRAN BRETAÑA | INGLATERRA | SLOUGH, BERKSHIRE | c/o IPD, UCB Celltech, 208 Bath road, Slough Berkshire SL1 3WE, Reino Unido de la Gran Bretaña |
| 3. | REINO UNIDO DE LA GRAN BRETAÑA | INGLATERRA | SLOUGH, BERKSHIRE | c/o IPD, UCB Celltech, 208 Bath Road, Slough SL1 3WE, Reino Unido de la Gran Bretaña           |
| 4. | REINO UNIDO DE LA GRAN BRETAÑA | INGLATERRA | SLOUGH, BERKSHIRE | c/o IPD, UCB Celltech, 208 Bath Road, Slough Berkshire SL1 3WE, Reino Unido de la Gran Bretaña |

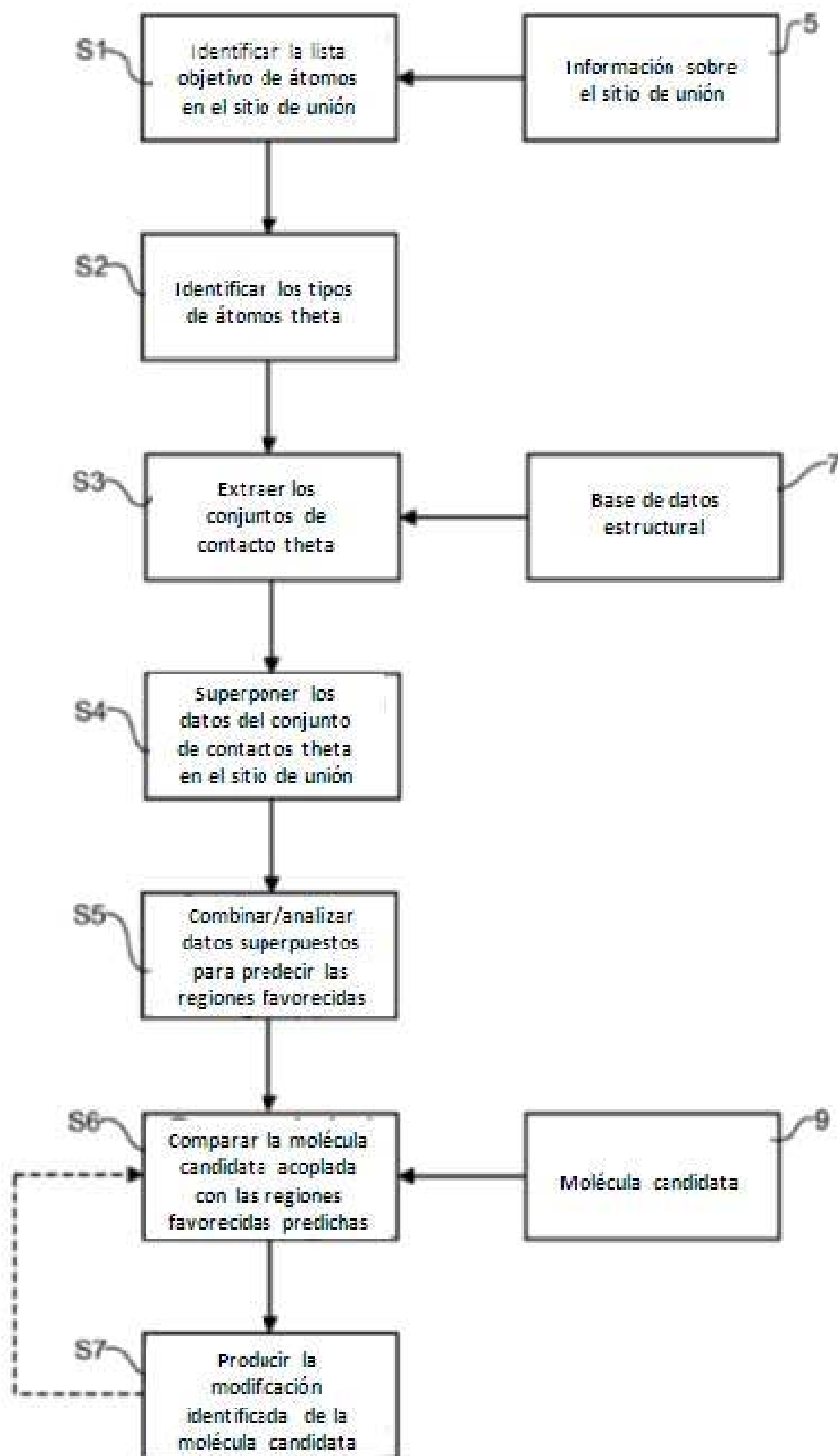
|                                                                                                                                                                                                                                                         |                                                  |                              |                                        |                                      |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|------------------------------|----------------------------------------|--------------------------------------|
| 7                                                                                                                                                                                                                                                       | <b>Datos del Representante Legal / Apoderado</b> |                              |                                        |                                      |
| <b>Nombre:</b>                                                                                                                                                                                                                                          |                                                  | JOSE ANDRES RINCON USCATEGUI | <b>Dirección Electrónica:</b>          | cavelier@cavelier.com                |
| <b>Dirección:</b>                                                                                                                                                                                                                                       |                                                  | CARRERA. 4 No. 72-35         | <b>Domicilio/País de constitución:</b> | COLOMBIA - BOGOTA D.C. - BOGOTA D.C. |
| <b>Identificación:</b><br><br><input checked="" type="checkbox"/> CEDULA DE CIUDADANIA <input type="checkbox"/> CEDULA DE EXTRANJERIA<br><input type="checkbox"/> EMPRESA EXTRANJERA <input type="checkbox"/> NIT<br><input type="checkbox"/> PASAPORTE |                                                  |                              |                                        |                                      |
| <b>Número:</b>                                                                                                                                                                                                                                          |                                                  | 79780910-                    |                                        |                                      |

|                              |  |  |  |
|------------------------------|--|--|--|
| <b>Presentación de Poder</b> |  |  |  |
| Año de Radicación            |  |  |  |
| Número de Radicación         |  |  |  |

|                          |                                  |                   |                         |            |
|--------------------------|----------------------------------|-------------------|-------------------------|------------|
| 8                        | <b>Datos Solicitud: PCT / WO</b> |                   |                         |            |
| <b>Número Solicitud:</b> |                                  | PCT/EP2014/062478 | <b>Fecha Solicitud:</b> | 13/06/2014 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                 |                              |                                                                    |                                 |            |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|------------------------------|--------------------------------------------------------------------|---------------------------------|------------|
| <b>Número Publicacion:</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                 | WO 2014/198951 A2            | <b>Fecha Publicacion:</b>                                          |                                 | 18/12/2014 |
| <b>9</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | <b>Declaraciones de prioridad</b>                               |                              | <input checked="" type="checkbox"/> SI <input type="checkbox"/> NO |                                 |            |
| 1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | <b>(33) País de origen</b><br>REINO UNIDO DE LA<br>GRAN BRETAÑA | <b>Codigo del país</b><br>GB | <b>(31) No. Solicitud</b><br>1310544.0                             | <b>(32) Fecha</b><br>13/06/2013 |            |
| <b>10</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | <b>Reivindicaciones</b>                                         |                              |                                                                    |                                 |            |
| <b>Número reivindicaciones:</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                 | 45                           | <b>Pago Reivindicaciones:</b>                                      |                                 | No         |
| <b>11</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | <b>Reducción de tasas.</b>                                      |                              |                                                                    |                                 |            |
| <p><i>Declaro que carezco de medios económicos para presentar la solicitud de patente.</i></p> <p><input type="checkbox"/> SI <input checked="" type="checkbox"/> NO</p> <p><b>Nota: En caso de ser persona natural y carecer de medios económicos, y por lo tanto, aplique la reducción de tasas a que se refiere la resolución vigente en tarifas, debe firmar la presente solicitud bajo la gravedad de juramento.</b></p>                                                                                                                                                                                                                                                                                                                                                                                |                                                                 |                              |                                                                    |                                 |            |
| <p><input type="checkbox"/> Micro, pequeñas y medianas empresas</p> <p><input type="checkbox"/> Universidades públicas o privadas</p> <p><input type="checkbox"/> Entidades sin ánimo de lucro</p> <p><b>Debe aportar los documentos que se indican en el numeral 17 de anexos</b></p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                 |                              |                                                                    |                                 |            |
| <b>12</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | <b>Documentos Anexos</b>                                        |                              |                                                                    |                                 |            |
| <p><input checked="" type="checkbox"/> Reivindicaciones</p> <p><input checked="" type="checkbox"/> Descripción</p> <p><input checked="" type="checkbox"/> Dibujos y/o Figuras</p> <p><input checked="" type="checkbox"/> Resumen</p> <p><input type="checkbox"/> Certificado Depósito Material Biológico</p> <p><input type="checkbox"/> Uso de Conocimiento tradicional</p> <p><input type="checkbox"/> Listado de secuencias</p> <p><input checked="" type="checkbox"/> Artes finales 12 x 12 cm</p> <p><input type="checkbox"/> Poderes, si fuere el caso</p> <p><input type="checkbox"/> Copia de la primera solicitud si se reivindica prioridad</p> <p><input type="checkbox"/> Traducción simple de la primera solicitud, si se reivindica prioridad</p> <p><input type="checkbox"/> Otros Anexos</p> |                                                                 |                              |                                                                    |                                 |            |

Fig. 7



## OBTENCIÓN DE UN LIGANDO TERAPÉUTICO MEJORADO

La presente invención hace referencia a la obtención de un ligando terapéutico mejorado, particularmente mediante la determinación de cómo un ligando existente o candidato puede ser modificado para mejorar la unión del ligando a un sitio de unión en una proteína diana o ayudando el diseño *de novo* de un posible candidato como precursor de un agente terapéutico.

Las moléculas terapéuticas (ligandos) se dividen en dos tipos de clases: entidades químicas (o nuevas entidades químicas, NCE) y productos biológicos. Las primeras son compuestos orgánicos de bajo peso molecular, usualmente de un peso molecular de 500 Daltones o menos, que han sido sintetizados químicamente o aislados a partir de productos naturales. Normalmente se derivan de productos químicos o 'hits' descubiertos mediante selección de librerías de productos químicos o naturales. Por lo general, estos hits tienen una afinidad de unión para el objetivo subóptima y se requiere la realización de pruebas considerables a base de eliminación de errores para obtener una mejor afinidad con el objetivo (usualmente de constante de afinidad ( $K_D$ ) baja, micromolar o menor). Es preferible que el hit tenga un peso molecular más bajo, como por ejemplo de 300 Daltones o menos, de tal manera que la modificación química posterior no exceda el límite de 500 Daltones. Usualmente estos hits se conocen como 'fragmentos'. La optimización del hit para obtener una molécula terapéutica o una molécula inicial aumenta mucho más por la información estructural; por ejemplo, mediante la obtención de una estructura cristalográfica de rayos x de la proteína en co-complejo con la molécula hit o fragmento. Estos datos dan una idea de la parte de la proteína diana a la que se une la molécula pequeña e indican, de manera importante, cómo las interacciones atómicas entre las dos dan cuenta de la unión. Adicionalmente, se revela la naturaleza topográfica de la superficie de la proteína que rodea inmediatamente al hit unido; y particularmente, si se trata de una hendidura o bolsillo, la estructura sugerirá cómo se puede elaborar el hit para llenar mejor el

espacio dentro del bolsillo y cómo hacer más interacciones con la proteína y de esta manera mejorar la afinidad de unión y la especificidad.

Existen algunos algoritmos a base de computador que el químico farmacéutico puede usar para tomar decisiones racionales para la elaboración química del hit. Estos son métodos basados en la física que tratan de calcular la energía libre de unión entre la molécula pequeña y la proteína a partir de principios básicos (p. Ej. el conjunto de software Schrodinger (RTM)) o posibles métodos estadísticos que dependen de una base de datos de interacciones atómicas tomadas de colecciones de estructuras de proteína-molécula pequeña (p. Ej. SuperStar).

Los productos biológicos son péptidos grandes o moléculas de proteínas (de peso molecular superior a 1000 Daltones). A menudo son anticuerpos o moléculas similares a un anticuerpo que reconocen y se unen a una molécula objetivo, usualmente con una mejor afinidad y especificidad comparado con las NCE ( $K_D$ ) baja, micromolar o menor). También pueden ser otros tipos de moléculas de proteínas, tales como hormonas, citoquinas, factores de crecimiento o receptores solubles.

La unión de una posible molécula terapéutica biológica a un sitio de unión puede ser modificada al mutar la molécula con potencial terapéutico. Esto puede resultar necesario para mejorar la afinidad de unión o para alterar la especificidad de unión. Sin embargo, llevar a cabo la mutación y probar la eficiencia de unión de la molécula mutada toma relativamente mucho tiempo. Puede que sea necesario realizar muchas mutaciones diferentes antes de obtener mejoras en la eficiencia de unión.

Es común usar computadores para predecir qué tipo de modificaciones pueden resultar más efectivas. Sin embargo, una molécula específica puede ser modificada de muchas maneras, y es difícil configurar un computador para lograr una predicción confiable en un periodo de tiempo práctico.

En "X-SITE: use of empirically derived atomic packing preferences to identify favourable interaction regions in the binding sites of proteins", Revista de Biología Molecular, 259, 175-201, Laskowski R A, Thomton J M, Humblet C & Singh J (1996) describen un método a base de computador para identificar las regiones de interacción favorables para distintos tipos de átomo en la superficie de una proteína, como por ejemplo en una interfaz de dímero o en un sitio de reconocimiento molecular o de unión. Las predicciones de Laskowski et al parten de una base de datos de información empírica sobre contactos intramoleculares no enlazantes observados en estructuras de proteínas de alta resolución.

En la estrategia de Laskowski et al, se rompen los 20 aminoácidos para obtener un total de 488 fragmentos posibles de 3 átomos. Teniendo en cuenta las similitudes químicas, esta cifra se reduce a un conjunto de 163 tipos de fragmentos que subdividen la base de datos. Cada fragmento contiene un primer átomo (conocido como "posición 3") con dos átomos adicionales que definen posiciones de triangulación (o normalización espacial). Se deriva una función de densidad al registrar las diferentes posiciones en las que un átomo (que se puede llamar "segundo átomo") se encuentra en contacto intramolecular no enlazante con el primer átomo de un fragmento de 3-átomos.

En Laskowski et al se obtiene una región de interacción favorable pronosticada para un tipo específico de átomo al trasplantar las funciones de densidad al sitio de unión. Cada función de densidad se trasplanta de tal manera que las coordenadas de los tres átomos del fragmento de 3 átomos que corresponden a la función de densidad se superponen a las coordenadas de un fragmento de 3 átomos correspondiente en el sitio de unión. Cuando se solapan las funciones de densidad de fragmentos de 3 átomos diferentes en el sitio de unión, se usa una "densidad" promedio para predecir la región de interacción favorable.

La estrategia de Laskowski et al es relativamente compleja y desecha datos potencialmente útiles. Las funciones de densidad de Laskowski et al se obtienen llenando una rejilla 3-D con las posiciones de los contactos de los segundos átomos con cada tipo de fragmento. Se usa una rejilla diferente para cada uno de los 163 tipos de fragmento. Luego, los datos para cada tipo de segundo átomo se transforman matemáticamente para dar la función de densidad. Al usar las definiciones de fragmento de Laskowski et al, el tipo de fragmento puede ser compartido por diferentes átomos en el mismo residuo y por átomos en diferentes residuos. Cuando la base de datos de Laskowski et al se ha construido sobre estos tipos de fragmentos, hay una sobre-abundancia de fragmentos de cadena principal que requiere un recorte en el paso de trasplante de las funciones de densidad al sitio de unión. Adicionalmente, un tipo de fragmento determinado puede incluir varios fragmentos reales con diferencias sutiles en longitudes y ángulos de enlace, así como diferencias concomitantes en las distribuciones del segundo átomo, las cuales se esconden cuando se combinan en estas divisiones.

La estructura secundaria de corto alcance en las proteínas usada para derivar los datos empíricos en Laskowski et al que puede llevar a sesgos y reduce la eficiencia con la que los datos empíricos indican las regiones de interacción favorables.

El objeto de la presente invención es abordar al menos uno de los problemas del estado de la técnica discutido anteriormente.

De acuerdo con un aspecto de la invención, se proporciona un método para diseñar un ligando *ab initio* que se enlazará a un sitio de unión de un destino macromolecular, o para identificar una modificación a un ligando con el fin de mejorar su afinidad con un sitio de unión de un destino macromolecular, el cual comprende:

a) identificar una lista de destino de los átomos que forman la superficie del sitio de unión de destino;

- b) identificar cada átomo, en adelante conocido como un átomo theta, en la lista de destino como un tipo particular de átomo theta;
- c) extraer, de una base de datos estructural de macromoléculas biológicas, información sobre contactos átomo a átomo intramolecular o intermolecular no enlazante, en donde el primer átomo en un par de átomos en contacto es un tipo de átomo theta particular y el segundo átomo opuesto del par, en adelante conocido como átomo iota, es un tipo de átomo iota particular, en donde dicha información comprende datos espaciales y/ o contextuales sobre el átomo iota respecto al átomo theta, y dichos datos recolectados para una pluralidad de contactos del tipo de átomo theta específico de dicha base de datos se conoce en adelante como un grupo de contacto theta;
- d) para cada átomo theta identificado en la lista de destino del paso b), superponer, en o alrededor del sitio de unión de destino, datos relacionados con tipo de átomo iota particular, o un grupo predeterminado de tipos de átomo iota relacionados, a partir del grupo de contacto theta correspondiente tomado en el paso c);
- e) combinar y/ o analizar los datos superpuestos de tal manera que se pueda predecir una o más regiones favorecidas del sitio de unión en donde el tipo de átomo iota particular o el grupo predeterminado de tipos de átomo iota relacionados tiene una alta propensión teórica; y
- f) con un posible ligando teóricamente acoplado al sitio de unión, comparar el tipo y la posición de uno o más de los átomos del posible ligando con las regiones favorecidas pronosticadas para los respectivos tipos de átomo iota, para identificar una modificación al posible ligando, en términos de átomos alternativos y/ o adicionales del posible ligando, que producirán una mayor intersección entre los átomos alternativos y/ o adicionales del posible ligando y las respectivas regiones favorecidas

del tipos de átomo iota, llevando a una mejora en la afinidad del posible ligando modificado con el sitio de unión, comparado con el posible ligando no modificado;

en donde cada contacto intramolecular o intermolecular no enlazante en la base de datos se define como un contacto entre los residuos opuestos de un pliegue de proteínas o entre unidades de monómero opuestas de un pliegue macromolecular o entre dos parejas macromoleculares que interactúan y es específicamente entre un átomo theta en un lado del pliegue o primera pareja de interacción y un átomo iota en el lado opuesto o segunda pareja de interacción; en un caso en el que se cumple la siguiente condición:

$s - R_w = t$ , en donde  $s$  es la separación entre los dos átomos del contacto,  $R_w$  es la suma de los radios de van de Waals de los dos átomos del contacto y  $t$  es una distancia umbral predeterminada; y

en donde el tipo de átomo theta se identifica de forma única en el paso b), de tal manera que no hay intersección entre los datos del grupo de contacto theta tomado en el paso c) para un tipo de átomo theta particular y los datos de cualquier otro grupo de contacto theta tomado en el paso c) para cualquier otro tipo de átomo theta, aparte de los datos relacionados con los contactos que involucran al átomo theta particular como el átomo iota.

Por consiguiente, cada tipo de átomo objetivo en el sitio de unión se clasifica de forma única y se asocia con información sobre un grupo de contactos tomados de la base de datos estructural que es única y que no se superpone con el conjunto de contactos asociados con cualquier otro tipo de átomo (aparte de los contactos que involucran al propio tipo de átomo objetivo como átomo iota). Esto significa que se puede combinar una distribución de ubicaciones teóricas para un tipo de átomo iota particular, o un grupo predeterminado de tipos de átomo iota relacionados, determinados con base en un átomo objetivo en el sitio de unión (p. Ej. mediante suma) de manera más eficiente (p. Ej. sin ponderación) con una distribución de ubicaciones teóricas para un tipo de

átomo iota, o un grupo predeterminado de tipos de átomo iota relacionados, determinado con base en otro átomo objetivo en el sitio de unión, por ejemplo para ofrecer una mejor predicción de una o más regiones favorecidas para el tipo de átomo iota o el grupo predeterminado de tipos de átomo iota relacionados. Además, debido a que cada átomo objetivo en el sitio de unión se clasifica de manera única, no hay que considerar variaciones en las longitudes o ángulos de enlace, y por lo tanto la ubicación teórica de un átomo iota particular es más precisa.

En una realización, se aplican normas simples para identificar de manera única los átomos vecinos para fines de triangulación. No es necesario hacer suposiciones sobre la naturaleza química de los átomos vecinos, cosa que sí es necesaria, por ejemplo, cuando los tipos de contacto se caracterizan en términos de los 163 tipos de fragmentos de 3 átomos de Laskowski et al.

En una realización, la información espacial extraída en el paso c) define la posición de cada átomo iota especificado en el grupo de contacto theta por referencia geométrica a la posición del átomo theta y a las posiciones del tercer y cuarto átomo, en donde el tercer átomo está unido covalentemente al átomo theta y el cuarto átomo está unido covalentemente al tercer átomo. En un ejemplo de dicha realización, para cada átomo iota especificado en el grupo de contacto theta, dicha información espacial tomada en el paso c) define aún más la posición de un quinto y sexto átomo por referencia geométrica a la posición del átomo theta y a las posiciones del tercer y cuarto átomo, en donde el quinto átomo está unido covalentemente al átomo iota y el sexto átomo o átomo iota.

En una realización, la superposición en o alrededor del sitio de destino del paso (d) comprende: analizar el grupo de contacto theta para extraer información espacial para los contactos que comprenden el tipo de átomo iota o uno o más del grupo predeterminado de tipos de átomo iota relacionados; y trazar esta información espacial para determinar ubicaciones teóricas que representan en dónde se ubicaría

cada tipo de átomo iota o cada uno del uno o más del grupo predeterminado de tipos de átomo iota relacionados si: i) el átomo theta del contacto estuviera ubicado en la posición del átomo theta correspondiente en el sitio de unión de destino; y ii) el tercer y cuarto átomo del contacto estuvieran ubicados en las posiciones del tercer y cuarto átomo del átomo theta correspondiente en el sitio de unión de destino. En una realización, la información espacial se analiza respecto a la información contextual antes del paso de trazado.

En una realización, una región en la que una densidad de las ubicaciones teóricas del tipo de átomo iota (o uno o más de los grupos predeterminados de tipos de átomo iota relacionados) está por encima de un umbral predeterminado se identifica como una de las regiones favorecidas. En un ejemplo de dicha realización, las ubicaciones teóricas para el tipo de átomo iota particular o para uno o más del grupo predeterminado de tipos de átomo iota relacionados se determinan por una pluralidad de los átomos theta en la lista de destino y una región en la que una densidad de los lugares teóricos acumulados está por encima de un umbral predeterminado se identifica como una de las regiones favorecidas.

Por consiguiente, las ubicaciones teóricas se combinan de forma acumulativa a partir de diferentes átomos en el sitio de unión antes de obtener la densidad de las ubicaciones teóricas con el fin de predecir las regiones favorecidas. Esto resulta en una representación estadística más precisa de la probabilidad de que un tipo de átomo iota particular, o en un grupo específico de tipos de átomo iota relacionados, estén ubicados en un lugar específico, pues toma en cuenta las contribuciones de todos los tipos de átomo relevantes en el sitio de unión de manera proporcionada e imparcial. En Laskowski et al., por lo contrario, las funciones de densidad se derivan para grupos de fragmentos de 3 átomos. Cada átomo puede ser asociado con diferentes grupos de fragmentos de 3 átomos, de tal forma que no se puede simplemente sumar las funciones de densidad de una manera compatible con las realizaciones de la presente invención. Por lo contrario, es necesario calcular el

promedio ponderado y/o promedio antes de combinar las funciones de densidad, lo que aumenta la complejidad y/o reduce la precisión.

En una realización, solo se usan los contactos entre átomos separados el uno del otro por residuos o más para identificar las regiones favorecidas. Esto reduce o evita significativamente el sesgo debido a la estructura secundaria de corta distancia. En una realización, los datos de contacto representan predominantemente los datos de proteínas entre pliegues de largo alcance.

De acuerdo con otro aspecto de la invención, se proporciona un método para generar una base de datos para uso en un método para diseñar un ligando *ab initio* que se enlazarán a un sitio de unión de un destino macromolecular, o para identificar una modificación a un ligando con el fin de mejorar su afinidad con un sitio de unión de un destino macromolecular, el cual comprende:

analizar las posiciones relativas de los átomos en cada uno de una pluralidad de proteínas u otras macromoléculas biológicas con el fin de identificar instancias de un contacto intramolecular no enlazante entre un primer átomo, conocido como átomo theta, y un segundo átomo, conocido como átomo iota, de la proteína o macromolécula; y

generar una base de datos que por cada contacto identificado especifica: el tipo del átomo theta, el tipo del átomo iota y la posición del átomo iota respecto al átomo theta;

en donde un contacto intramolecular no enlazante se define como una instancia en la que se cumplen las siguientes condiciones:

$s - R_w = t$ , en donde  $s$  es la separación entre los átomos theta y iota,  $R_w$  es la suma de los radios de van de Waals de los átomos theta y iota, y  $t$  es una distancia umbral predeterminada de 2,5 angstroms y preferiblemente de 0,8 angstroms; y

en donde, en el caso de las proteínas, los átomos theta y iota son residuos de aminoácidos separados el uno del otro por al menos cuatro residuos en un polipéptido lineal o que están en cadenas polipeptídicas separadas.

De acuerdo con otro aspecto de la invención, se proporciona un método para generar una base de datos para uso en un método para diseñar un ligando *ab initio* que se enlazaría a un sitio de unión de un destino macromolecular, o para identificar una modificación a un ligando con el fin de mejorar su afinidad con un sitio de unión de un destino macromolecular, el cual comprende:

analizar las posiciones relativas de los átomos en cada una de una pluralidad de proteínas o demás macromoléculas biológicas con el fin de identificar instancias de contacto intramolecular no enlazante entre un primer átomo conocido como átomo theta y un segundo átomo conocido como átomo iota de la proteína o macromolécula; y

generar una base de datos que por cada contacto identificado especifica: el tipo del átomo theta, el tipo del átomo iota y la posición del átomo iota respecto al átomo theta;

en donde un contacto intramolecular no enlazante se define como una instancia en la que se cumple la siguiente condición:

$s - R_w = t$ , en donde  $s$  es la separación entre los átomos theta y iota,  $R_w$  es la suma de los radios de van de Waals de los átomos theta y iota, y  $t$  es una distancia umbral predeterminada, usualmente de 2,5 angstroms y preferiblemente de 0,8 angstroms; y

en donde el método comprende la subdivisión de la base de datos para formar grupos de contactos identificados en donde el átomo theta es uno y solo uno de los 167 átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de las proteínas y el átomo iota es uno y solo uno de una pluralidad de grupos no superpuestos obtenidos al clasificar los 167 átomos distintos al hidrógeno presentes

en los 20 aminoácidos naturales de las proteínas en grupos de acuerdo con la similitud química.

A continuación se describen realizaciones de la invención, únicamente a modo de ejemplo, haciendo referencia a los dibujos adjuntos, en los que los símbolos de referencia correspondientes representan partes correspondientes, y en donde:

La Figura 1 es una ilustración esquemática de un ejemplo de nomenclatura para átomos en contacto intramolecular o intermolecular no enlazante y los átomos vecinos usados para normalización de coordenadas;

Las Figuras 2-6 ilustran un proceso de normalización de coordenadas para un átomo en contacto intramolecular o intermolecular no enlazante;

La Figura 7 es un diagrama de flujo que ilustra los pasos en un método para diseñar un ligando *ab initio* que se enlazará a un sitio de unión de un destino macromolecular, o para identificar una modificación a un ligando con el fin de mejorar su afinidad con un sitio de unión de un destino macromolecular;

La Figura 8 es una vista generada por computador que representa una mutación de treonina 30 a arginina 30 en una cadena ligera;

La Figura 9 es una vista generada por computador que representa una mutación de arginina 54 a serina 54 en una cadena ligera;

La Figura 10 es una vista generada por computador que representa una mutación de serina 56 a isoleucina 56 en una cadena ligera;

La Figura 11 es una vista generada por computador que representa una mutación de serina 60 a aspartato 60 en una cadena ligera;

La Figura 12 es una vista generada por computador que representa una mutación de treonina 72 a arginina 72 en una cadena ligera;

La Figura 13 es una vista generada por computador que representa la combinación de 5 mutaciones en la cadena ligera del anticuerpo 496i, resultando en una afinidad con IL17F mejorada 180 veces;

La Figura 14 es un diagrama de flujo que ilustra los pasos en un método para predecir los efectos de las mutaciones puntuales en la interfaz VH-VL de un fragmento Fab;

La Figura 15 es una vista generada por computador que representa una mutación de treonina 71 a arginina 71 en una cadena pesada;

La Figura 16 es una vista generada por computador que representa una mutación de serina 107 a ácido glutámico 107 en una cadena ligera;

La Figura 17 es una vista generada por computador que representa una mutación de treonina 109 a isoleucina 109 en una cadena ligera;

La Figura 18 es una vista generada por computador que representa la combinación de tres mutaciones en un fragmento Fab X, resultando en una T<sub>m</sub> de 81.2 °C.

El Banco Mundial de Datos de Proteínas ([www.PDB](http://www.PDB)) mantiene un archivos de datos estructurales macromoleculares que está libre y disponible públicamente para la comunidad global. En mayo del 2013 este conjunto de datos había alcanzado las 90000 estructuras. La mayoría de estas macromoléculas son proteínas que fueron determinadas principalmente por cristalografía de rayos X. De esta forma, los datos depositados contienen datos tridimensionales a nivel atómico en forma de coordenadas cartesianas de átomos individuales que componen la estructura de la proteína respectiva.

Los inventores plantearon la hipótesis de que es posible extraer información útil de este archivo que se puede aplicar para ayudar en el diseño de nuevas terapias. Las cadenas polipeptídicas de una proteína naciente se pliegan en estructuras terciarias y cuaternarias tridimensionales de una forma notablemente reproducible para producir

la proteína madura. Se conocen interacciones que afectan la formación de la estructura de proteínas secundaria, elementos como las hélices, láminas beta y giros. Sin embargo, las normas que predicen el plegamiento de proteínas de orden superior son poco conocidas. No obstante, los inventores se han dado cuenta de que debe haber normas precisas que regulen la interacción de átomos no enlazados pero sí "en contacto", ya sea dentro de la misma molécula, por ejemplo en caras opuestas de un pliegue de proteínas, o en diferentes moléculas.

En una realización, una base de datos estructural de macromoléculas biológicas (p. Ej. la wwPDB) es analizada para extraer dichas normas, y las normas se aplican para facilitar el descubrimiento de fármacos. A continuación se describe un ejemplo de dicho proceso.

En una realización, se identifican pares de átomos en contacto no enlazantes (conocidos respectivamente como átomos "theta" y "iota") para cada macromolécula (p. Ej. proteína), o un subconjunto de menos que todas las macromoléculas en la base de datos estructural de macromoléculas (p. Ej. la wwPDB). Dichos contactos pueden ocurrir, por ejemplo, entre residuos opuestos de un pliegue de proteínas o entre unidades de monómero opuestas de un pliegue macromolecular (entre cadenas separadas de una estructura macromolecular) o entre dos parejas macromoleculares que interactúan. Cada contacto se clasifica como estando entre un átomo theta a un lado del pliegue o primera pareja de interacción y un átomo iota en el lado opuesto o segunda pareja de interacción.

En una realización, los contactos intramoleculares o intermoleculares no enlazantes se definen como una instancia en la que se cumple la siguiente condición: 1)  $s - R_w = t$ , en donde  $s$  es la separación entre los dos átomos del contacto,  $R_w$  es la suma de los radios de van de Waals de los dos átomos del contacto, y  $t$  es una distancia umbral predeterminada. y, de manera opcional, también se cumple la siguiente condición: 2) los dos átomos del contacto están separados por al menos cuatro residuos a lo largo

de una cadena polipeptídica lineal o están en cadenas polipeptídicas separadas. En una realización, la distancia predeterminada es de 2,5 angstroms. En otra realización, la distancia predeterminada es de 1,5 angstroms. En otra realización, la distancia predeterminada es de 1,0 angstroms. En otra realización, la distancia predeterminada es de 0,8 angstroms.

En la siguiente descripción, toda referencia a un "contacto" se entiende como "contacto intramolecular o contacto intermolecular no enlazante" de acuerdo con la definición anterior.

Las bases de datos como la wwPDB pueden contener información sobre proteínas que son muy similares entre sí y/o que tienen estructuras relacionadas. En una realización, la base de datos es analizada con el fin de evitar/reducir el sesgo causado por dichas similitudes/ relaciones. En una realización, el análisis se lleva a cabo con base en la homología de secuencia primaria, por ejemplo de tal manera que solo se seleccione una estructura representativa de cada familia de proteínas similares/ relacionadas para el análisis. Adicional o alternativamente, se pueden usar uno o más criterios de selección; por ejemplo se pueden incorporar estructuras factoriales de alta resolución y de baja temperatura.

En una realización, se construye una base de datos secundaria a partir de la base de datos estructural de macromoléculas biológicas (primaria) (p. Ej. la wwPDB). La base de datos secundaria contiene información sobre los contactos intramoleculares o intermoleculares no enlazantes. En una realización, la base de datos secundaria contiene información sobre más de 1 millón de parejas de contacto, opcionalmente más de 5 millones de parejas de contacto, opcionalmente más de 11 millones de parejas de contacto. En una realización, la base de datos secundaria contiene información de más de 15 millones de parejas de átomos de contacto, extraídas de aproximadamente 20000 proteínas no homólogas.

En una realización, la base de datos secundaria contiene información sobre los tipos de átomos precisos del par de contacto. En una realización, la base de datos secundaria contiene información espacial que define la relación tridimensional del átomo theta respecto al átomo iota. En una realización, la base de datos secundaria también contiene información contextual relativa al entorno local del contacto. En una realización, la información contextual contiene información relativa al entorno local de cada par de contacto, incluyendo una o más de las siguientes en cualquier combinación: estructura secundaria, tipos de aminoácido u otros tipos de monómero que contienen el par de contacto, las unidades de monómero adyacentes y/o la geometría local de las mismas en una cadena de polímeros a cualquier lado del contacto, aminoácidos adyacentes en una cadena polipeptídica a cada lado del contacto, geometría local de dichas unidades de monómero adyacentes o aminoácidos, el factor de temperatura del átomo theta, el factor de temperatura del átomo iota, la superficie accesible del átomo theta, la superficie accesible del átomo iota, el número de diferentes contactos del átomo iota para el átomo theta particular, y el número de otros átomos theta en la misma unidad de monómero que el átomo theta.

En una realización, las coordenadas 3-D del par de contactos y los átomos adyacentes unidos covalentemente se normalizan, como grupo, a un marco de referencia de una base de datos común tal y como se describe a continuación. Esto simplifica el análisis posterior de posibles patrones o normas de contacto subyacentes y la aplicación de cualquiera de dichas normas al diseño del fármaco.

En una realización, el tipo de átomo theta se identifica como uno y solo uno de: los 167 tipos de átomos covalentes (excluyendo el hidrógeno) que conforman a los 20 bloques de construcción de aminoácidos naturales de las proteínas (en este caso la base de datos secundaria se puede dividir de conformidad y puede contener información sobre hasta 27889,  $167 \times 167$  tipos de contacto diferentes); y/o los 82 átomos distintos al hidrógeno presentes en los 4 nucleótidos del polímero de ácido desoxirribonucleico (ADN); y/o los 42 átomos distintos al hidrógeno presentes en los

nucleótidos de ADN metilado, citidina fosfato y fosfato de adenosina; y/o los 85 átomos distintos al hidrógeno presentes en los 4 fosfatos de nucleótidos del polímero de ácido ribonucleico (ARN); y/o los 89 átomos distintos al hidrógeno presentes en los fosfatos nucleótidos de ribosa 2-O' metilada del ARN; y/o los más de 400 átomos distintos al hidrógeno presentes en el ARN de base modificada más común después de la transcripción.

En una realización, el tipo de átomo iota se identifica como uno y solo uno de: los 167 tipos de átomos covalentes (excluyendo hidrógeno) que conforman los 20 bloques de construcción de aminoácidos naturales de las proteínas; y/o el átomo de oxígeno presente en moléculas de agua, estructuralmente relevantes, unidas a proteínas (esto puede ser útil porque las estructuras cristalinas en la base de datos primaria a menudo contienen moléculas de agua estructuralmente pertinentes, es decir, ciertos átomos de proteínas muestran interacciones definidas con las moléculas de agua enlazadas); y/o los 82 átomos distintos al hidrógeno presentes en los 4 nucleótidos del polímero de ácido desoxirribonucleico (ADN); y/o los 42 átomos distintos al hidrógeno presentes en los nucleótidos de ADN metilado, citidina fosfato y fosfato de adenosina; y/o los 85 átomos distintos al hidrógeno presentes en los 4 fosfatos de nucleótidos del polímero de ácido ribonucleico (ARN); y/o los 89 átomos distintos al hidrógeno presentes en los fosfatos nucleótidos de ribosa 2-O' metilada del ARN; y/o los más de 400 átomos distintos al hidrógeno presentes en el ARN de base modificada más común después de la transcripción.

En un par de contacto el átomo opuesto es visto y grabado desde cualquier lado en el contacto. A continuación se describe la nomenclatura en un ejemplo de realización y se ilustra de manera esquemática en la Figura 1.

El átomo en el lado de referencia del contacto se denomina el átomo theta 1, mientras que el átomo opuesto se denomina el átomo iota 2. En este ejemplo, los átomos

adicionales utilizados para la normalización de las coordenadas 3-D se definen de la siguiente manera. El siguiente átomo al que el átomo theta 1 está unido covalentemente en la dirección del átomo alfa C de ese aminoácido se conoce como el tercer átomo 3 y el siguiente átomo nuevamente como el cuarto átomo 4. El cuarto átomo 4 está unido covalentemente al átomo 3. El siguiente átomo al que el átomo iota 2 está unido covalentemente, en la dirección del átomo alfa C del respectivo aminoácido, se denomina el quinto átomo 5 y el siguiente átomo como sexto átomo 6. El sexto átomo está unido covalentemente al quinto átomo o al átomo iota 2.

En una realización, para evitar instancias de ambigüedad, el tercer y cuarto átomo se escogen únicamente para cada tipo de átomo theta especificado. En una realización, el quinto y el sexto átomo también se escogen de manera única. En una realización se aplica la siguiente convención. Si el átomo theta 1 es un átomo alfa C, entonces el tercer y cuarto átomo son respectivamente los átomos estructurales de carbono carbonílico y oxígeno. Si el átomo theta 1 es un carbono carbonílico estructural, el átomo 3 y el átomo 4 son respectivamente el carbono alfa C y el nitrógeno estructural. Si el átomo theta 1 es el nitrógeno estructural, el átomo 3 y el átomo 4 son respectivamente el carbono alfa C y el carbono carbonílico estructural. Si el átomo theta 1 es un átomo de carbono beta C, el átomo 3 y el átomo 4 son respectivamente el carbono alfa C y el carbono carbonílico estructural. En cadenas laterales de fenilalanina y tirosina en donde existe la posibilidad de escoger entre dos átomos de carbono épsilon para las posiciones del tercer y cuarto átomo, entonces se selecciona el átomo más cercano al átomo de nitrógeno estructural.

En una realización, la normalización de las coordenadas de cada contacto se lleva a cabo en los átomos, theta, iota, tercero y cuarto, y opcionalmente también en los átomos cinco y seis como grupo de tal manera que se mantenga su relación 3-D. Las coordenadas normalizadas resultantes se pueden llamar grupo de coordenadas normalizadas. En una realización, esto se logra llevando a cabo los siguientes pasos

en la secuencia, tal y como se ilustra en las Figuras 2-6.

La Figura 2 ilustra un átomo theta 1, un tercer átomo 3 y un cuarto átomo 4 de un contacto intramolecular o intermolecular no enlazante posicionado en relación a un marco de referencia, definido en relación a unos ejes X-, Y-, y Z-, de acuerdo con las coordenadas en una base de datos primaria (como por ejemplo la wwPDB). En una primera etapa de un ejemplo de proceso de normalización de coordenadas, las coordenadas del grupo de átomos se traducen de tal forma que el átomo theta 1 se encuentra en la coordenada cero (Figura 3). A continuación, se rota todo el grupo alrededor del eje Z hasta que el átomo 3 esté en  $y = 0$  (Figura 4). Luego, se rota el grupo alrededor del eje Y hasta que el tercer átomo esté en  $y=0$  y  $z=0$  (Figura 5). A continuación, se rota el grupo alrededor del eje x hasta que el cuarto átomo 4 esté en  $y=0$  y todo el grupo esté en el plano x-y (los tres átomos en  $y=0$ ; Figura 6). De esta forma, cada uno de los 167 primeros tipos de átomo se pueden superponer para ese tipo y se puede subdividir la base de datos secundaria como corresponde. A su vez, cada una de las primeras divisiones de átomos se puede subdividir en 167 tipos de átomo iota, facilitando el análisis de la distribución espacial de cada tipo de átomo iota respecto a cada tipo de átomo theta.

En una realización, los patrones de distribución de los átomos iota respecto a los átomos theta son analizados con el fin de identificar similitudes entre los patrones de distribución para tipos de átomo iota nominalmente diferentes. De esta forma, los tipos de átomo iota únicos (p. Ej. los 167 tipos de átomos covalentes mencionados anteriormente) se pueden combinar en diferentes grupos (conocidos en adelante como "grupos predeterminados de tipos de átomo iota relacionados") para simplificar el uso posterior de los datos. El agrupar los tipos de átomo de acuerdo con la similitud de los patrones de distribución reduce la carga computacional asociada con el método descrito a continuación, haciendo referencia a la Figura 7, por ejemplo, aumentando así la velocidad y/o reduciendo los gastos relacionados con hardware.

En una realización, este proceso se simplifica mediante el uso de coordenadas polares en vez de coordenadas cartesianas (en una realización, esto se logra mediante un proceso de conversión entre coordenadas cartesianas y coordenadas polares, por ejemplo cuando los datos en la base de datos primaria se presentan a través de coordenadas cartesianas). En una realización, se usan coordenadas polares 2-D, especificando las posiciones relativas de los átomos theta y iota en términos únicamente de los dos ángulos polares  $\theta$  (theta) y  $\phi$  (phi) (que corresponden a la latitud y longitud en un globo). las gráficas 2-D de latitud-longitud resultantes no muestran ninguna información respecto a variaciones en la distancia entre los átomos theta y iota. Sin embargo, se encuentra que esta distancia es relativamente constante, de tal forma que las gráficas theta-phi contienen la mayoría de la información relevante respecto al contacto. Reducir el análisis a un problema en dos dimensiones en vez de tres mejora considerablemente la eficiencia de los análisis posteriores. En una realización, se usan líneas de contorno para ilustrar las variaciones en la posición relativa del átomo iota. Las líneas de contorno pueden representar líneas de "densidad" constante o la probabilidad de un posicionamiento relativo de los átomos theta e iota.

El análisis de dichos gráficos de ángulo polar revelan que un factor particularmente importante que regula el patrón de frecuencias del átomo iota es la naturaleza elemental y el estado de hibridación de los átomo iota, a saber C sp<sup>3</sup>, C sp<sup>2</sup> (aromático), C sp<sup>2</sup> (no aromático), N sp<sup>3</sup>, N sp<sup>2</sup>, O sp<sup>3</sup>, O sp<sup>2</sup> o S. Como resultado, es posible mejorar la eficiencia del análisis agrupando los 167 tipos de átomos según estos ocho grupos identificados. En otras realizaciones se puede usar una agrupación diferente.

En general, los factores ambientales relacionados con el contacto, tales como la naturaleza de los aminoácidos adyacentes, hacen menos diferencia en el patrón de frecuencia iota, a excepción de la estructura secundaria. Como se podría esperar, los patrones de frecuencia de los átomos theta de nitrógeno amídico estructural versus

los átomos de oxígeno estructural y viceversa están sesgados por la estructura secundaria, particularmente en lo que respecta a si son o no son de lámina beta.

En una realización, la base de datos secundaria etiqueta los datos de contacto con el tipo de estructura secundaria local (hélice, lámina beta o espiral aleatoria). Esto proporciona la base para diferenciar cualquier posible influencia de la estructura secundaria sobre los patrones de contacto en una etapa posterior.

En una realización, se proporciona un método basado en lo anterior que ayuda a identificar modificaciones a un ligando que mejora la fuerza de unión, o afinidad, del ligando a un sitio de unión. En una realización, el método se usa con diseño de fármacos biológicos o de nuevas entidades químicas (NCE). Respecto al primero, el método puede ser útil para predecir 'puntos calientes' o posiciones de átomos farmacóforos en posibles puntos de unión de fármacos de proteínas diana. Esto puede facilitar el diseño de fármacos de novo. En situaciones donde haya una estructura disponible de materia química enlazada en un sitio de unión, el método puede sugerir tipos de átomo y posiciones para la elaboración de la química para obtener un ligando con mejores características de unión. En el caso de fármacos de proteínas, tales como anticuerpos, el método puede ser usado para predecir mutaciones en la proteína o sitio de unión del anticuerpo que llevarían a mejorar la afinidad o especificidad de unión. El método también puede ser usado para sugerir posiciones para la modificación dentro de una estructura macromolecular con el fin de mejorar las propiedades de la macromolécula. Por ejemplo, tal como se ilustra en los Ejemplos de la siguiente sección, el método puede ser usado para identificar mutaciones puntuales dentro de cadenas de anticuerpos VH y VL con el fin de mejorar la estabilidad térmica del anticuerpo. Las mutaciones están en cadenas separadas, pero aún están dentro de la macromolécula del anticuerpo.

La Figura 7 ilustra un ejemplo de un método para diseñar un ligando *ab initio* que se enlazará a un sitio de unión de un destino macromolecular, o para identificar una

modificación a un ligando con el fin de mejorar su afinidad con un sitio de unión de un destino macromolecular.

En el paso S1 se obtienen datos que representan el sitio de unión de destino de una proteína diana, por ejemplo de un dispositivo de memoria local o remota 5. Se identifica una lista de destino de átomos que forman la superficie del sitio de unión de destino.

En el paso S2 se identifica cada átomo en la lista de destino como un tipo de átomo theta particular.

En el paso S3, se extrae información de una base de datos estructural de macromoléculas biológicas (p. Ej. wwPDB), proporcionada por ejemplo por un dispositivo de memoria local o remota 7, sobre contactos intramoleculares o intermoleculares no enlazantes en donde el primer átomo en un par de átomos en contacto es un tipo de átomo theta particular y el segundo átomo del par es un tipo de átomo iota particular. La información extraída comprende información espacial y/o contextual sobre el átomo iota respecto al átomo theta. Se recolecta información para una pluralidad de contactos del tipo de átomo theta particular y el conjunto resultante de datos se conoce como un conjunto de contactos theta. En una realización, el conjunto de contactos theta comprende información recolectada para todos los contactos disponibles del tipo de átomo theta particular. La información extraída puede formar una base de datos que es un ejemplo de la "base de datos secundaria" discutida arriba. En una realización, la información extraída en el paso S3 se recoge en una base de datos secundaria que comprende uno y sólo un conjunto de contactos theta para cada uno de los tipos de átomo theta. En un ejemplo de dicha realización, los conjuntos de contactos theta de la base de datos secundaria se subdividen en una pluralidad de tipos de átomo iota que no se superponen o grupo de tipos de átomo iota relacionados que no se superponen. En un ejemplo de dicha realización, la base de datos se subdivide para formar grupos de contactos identificados, en donde el primer

átomo es uno y solo uno de los 167 átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de las proteínas y el segundo átomo es uno y solo uno de una pluralidad de grupos que no se solapan obtenidos al clasificar los átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de las proteínas en grupos basados en la similitud química.

En el paso S4, por cada átomo theta identificado en la lista de destino del paso en el paso S2, los datos relacionados con un tipo de átomo iota particular, o un grupo predeterminado de tipos de átomo iota relacionados, a partir del conjunto de contacto theta correspondiente extraído en el paso S3 se superpone en o alrededor del sitio de unión de destino. En una realización, la superposición comprende: analizar el grupo de contacto theta para extraer información espacial que comprende el tipo de átomo iota específico o uno o más del grupo predeterminado de tipos de átomo iota relacionados; y trazar estos datos espaciales para determinar ubicaciones teóricas que representan en dónde estaría cada tipo de átomo iota o cada uno de los uno o más grupos predeterminados de tipos de átomo iota relacionados si: i) el átomo theta del contacto estuviera ubicado en la posición del átomo theta correspondiente en el sitio de unión de destino; y ii) el tercer y cuarto átomo del contacto estuvieran ubicados en las posiciones del tercer y cuarto átomo del átomo theta correspondiente en el sitio de unión de destino. Cuando las ubicaciones teóricas determinadas entran en conflicto con los átomos del sitio de unión y/o están enterrados en la proteína diana, estas pueden ser retiradas de análisis posteriores. Por ejemplo, si se determina que la ubicación teórica de un átomo iota individual se cruza con la ubicación de un átomo de la macromolécula objetivo, más cerca que  $R_w - 0,2$  angstroms, entonces el átomo iota es excluido de análisis posteriores.

En el paso S5, los datos superpuestos se combinan y/o analizan a fin de predecir una o más regiones favorecidas del sitio de unión en donde el tipo iota particular, o el grupo predeterminado de tipos de átomo iota relacionados tienen una alta propensión teórica.

En el paso S6, un posible ligando está teóricamente acoplado al sitio de unión. Se pueden proporcionar datos que definen el posible ligando a partir de, por ejemplo, un dispositivo de memoria local o remota 9. Luego se hace una comparación del tipo y posición de uno o más de los átomos del posible ligando con las regiones favorecidas pronosticadas para los respectivos tipos de átomo iota. A partir de la comparación se identifican modificaciones al posible ligando, en términos de átomos adicionales o alternativos del posible ligando, que producirán una mayor intersección entre los átomos adicionales o alternativos del posible ligando y las regiones favorecidas del respectivo tipo de átomo iota, lo que lleva a una mejora en la afinidad del posible ligando modificado con el sitio de unión, comparado con el posible ligando no modificado.

En el paso S7 se saca el posible ligando modificado, ya sea como una propuesta de mejora a un ligando existente o como parte de un diseño *ab initio* de un nuevo ligando. De manera opcional, los pasos S7 y S6 pueden ser iterados para modificar aún más el ligando. El dispositivo de memoria local o remotas 5, 7 y 9 pueden ser implementados en una única pieza de hardware (p. Ej. un único dispositivo de almacenamiento) o en dos o más dispositivos diferentes separados.

En una modalidad, el posible ligando modificado se saca a un dispositivo de memoria de salida para su almacenamiento o transmisión y/o a una pantalla para visualización.

En una modalidad, el tipo de átomo theta particular es identificado única en el paso S2, de tal manera que no hay intersección entre el grupo de contactos para los que se extrae la información en el paso S3 para el átomo theta particular y el grupo de contactos para cualquier otro tipo de átomo theta (a excepción de los contactos que involucran el tipo de átomo theta particular como el átomo iota).

En una modalidad, el paso S5 comprende determinar una o más regiones favorecidas para cada uno de una pluralidad de diferentes tipos de átomo iota y/o grupos predeterminados de tipos de átomo iota relacionados. En dicha realización, el paso de

comparación S6 se puede repetir para cada uno de una pluralidad de tipos diferentes de átomo iota y/o grupos predeterminados de tipos de átomo iota relacionados, con el fin de identificar posibles modificaciones que involucren a los diferentes tipos de átomo iota o los grupos predeterminados de tipos de átomo iota relacionados.

En una modalidad, las etapas S2-S7 se llevan a cabo para una pluralidad de átomos diferentes en el sitio de unión. En una modalidad, tal y como se describe a continuación, las regiones favorecidas se pueden determinar de manera más precisa combinando de forma acumulativa (p. Ej. sumando) las distribuciones de las ubicaciones teóricas determinadas de los tipos de átomo iota, tal y como se deriva para una pluralidad de átomos diferentes en el sitio de unión.

En una modalidad, se extiende el análisis de tal manera que, para cada región favorecida, se derivan vectores que describen la posición del quinto átomo respecto a su correspondiente átomo iota. Se hace un análisis de los vectores para identificar un vector de enlace favorecido que represente una predicción de la unión covalente de consenso teórico de un átomo iota en la región. A continuación, el vector de enlace favorecido identificado puede ser usado para refinar el diseño del posible ligando y/o para refinar la modificación del posible ligando, según sea el caso. El vector de enlace favorecido identificado puede ser usado, por ejemplo, para indicar cómo se pueden unir los átomos iota en regiones favorecidas diferentes, ayudando así a identificar modificaciones que involucren múltiples adiciones o intercambios de átomos. En una realización, el análisis es un análisis de conglomerados.

La distribución de ubicaciones teóricas mide u ofrece una propensión de cómo un tipo particular de átomo iota (o grupo predeterminado de tipos de átomo iota relacionados) se verá favorecido en diferentes lugares en el sitio de unión. En una realización, una región en la que una densidad de las ubicaciones teóricas está por encima de un umbral predeterminado se identifica como una de las regiones favorecidas. La densidad de las ubicaciones teóricas es una medida del número de

ubicaciones teóricas que ocurren, por ejemplo, en un volumen espacial específico. En una realización, se determinan las ubicaciones teóricas del átomo *iota* para una pluralidad de átomos objetivo en el sitio de unión y una región en donde una densidad de las ubicaciones teóricas acumuladas para el átomo *iota* (o grupo predeterminado de tipos de átomo *iota* relacionados) para la pluralidad de átomos objetivo está por encima del umbral determinado es identificada como una de las regiones favorecidas. Las ubicaciones teóricas determinadas para diferentes átomos objetivo en el sitio de unión se pueden sumar, por ejemplo, para obtener las ubicaciones teóricas acumuladas. Este enfoque para tomar en cuenta los efectos de los átomos diferentes en el sitio de unión es computacionalmente eficiente y minimiza la pérdida de información sobre la interacción de entre el posible ligando y el sitio de unión. El enfoque es facilitado por la caracterización de los contactos en términos de pares de tipos simples de átomos o tipos simples de átomos combinados con átomos de grupos predeterminados de tipos simples de átomos. Dicho enfoque no es válido cuando los contactos se caracterizan en términos de fragmentos de 3 átomos, como es el caso en Laskowski et al., por ejemplo.

Las distribuciones de ubicaciones teóricas obtenidas se pueden transformar de distintas maneras para crear funciones de densidad de probabilidad, es decir, un potencial estadístico para la preferencia de un tipo de átomo *iota* específico en una posición específica en el sitio de unión. A su vez, las funciones de densidad de probabilidad pueden ser tratadas de una manera análoga a la densidad de electrones y convertidas a archivos ccp4 que son una forma estándar de visualizar estos mapas en software de gráficos moleculares, p. Ej. Pymol.

En una realización, en el paso S5, la una o más regiones favorecidas se expresa(n) en coordenadas polares, que opcionalmente comprenden únicamente los ángulos polares y azimutales, opcionalmente en donde el marco de referencia se normaliza por referencia al tercer y cuarto átomo 3,4.

En una realización, el paso S6 comprende: identificar una modificación del posible ligando que aumenta un grado de solapamiento entre un átomo del posible ligando (esté o no presente antes de la modificación) y una región favorecida pronosticada para un átomo del mismo tipo en el sitio de unión. En una realización, se inspeccionan las distribuciones generadas de ubicaciones teóricas y/o regiones favorecidas, por ejemplo a través de un software de computador o manualmente, como superposiciones sobre la estructura macromolecular objetivo en complejo con el respectivo posible ligando. Si por ejemplo el posible ligando se refiere a un anticuerpo, la interfaz entre el anticuerpo y la macromolécula de destino se puede examinar para determinar el grado de solapamiento entre los átomos de anticuerpos y las respectivas distribuciones de las ubicaciones teóricas de los átomos iota y/o las regiones favorecidas identificadas para ese átomo. En algunos casos el grado de solapamiento ya será alto. Sin embargo, en otras regiones de la interfaz de la superposición puede ser baja. Es en estas regiones en las que las mutaciones en el residuo de aminoácido adyacente del anticuerpo se puede identificar/ proponer de la manera más efectiva. En una realización, cada uno de los otros 19 aminoácidos naturales se considera a su vez en esta posición, en todas sus respectivas conformaciones de rotámeros y en cada caso se examinan el grado de solapamiento con las distribuciones de ubicación teórica de los átomos iota relevantes y/o las regiones favorecidas, con el fin de seleccionar los residuos con el máximo grado de solapamiento para las mutaciones propuestas. De esta manera, se proporciona un método racional para seleccionar mutaciones que puede mejorar la afinidad en el anticuerpo escogido. Se pueden generar mutaciones puntuales individuales en las diferentes regiones de la interfaz anticuerpo-proteína diana; las que llevan a mejoras en la afinidad se pueden ensayar en combinaciones de dos o más que pueden llevar a aumentos sinérgicos en la afinidad.

En una realización, el paso S6 comprende reemplazar cada uno de los uno o más de (opcionalmente todos) los residuos de aminoácidos del ligando que está(n) en

contacto directo con el sitio de unión de destino, o muy cerca al el sitio de unión de destino, en donde cada uno del uno o más de (opcionalmente todos) los residuos seleccionados de entre los otros 19 aminoácidos naturales. Cada uno de estos reemplazos se conoce en la presente como un “residuo de reemplazo” e incluye la modificación de un ligando mediante una única sustitución de un residuo por otro residuo. En una realización, para cada uno de estos residuos de reemplazo que no cree conflicto con átomos adyacentes del ligando u objetivo (p. Ej. solapamiento entre uno o más átomos del residuo de reemplazo y uno o más átomos diferentes del objetivo o ligando), el tipo y posición de cada átomo del residuo de reemplazo se compara con las respectivas regiones favorecidas del tipo de átomo para identificar si producirán o no una mayor intersección que los átomos del residuo original. En una realización, se saca a continuación una lista de los residuos de reemplazo identificados por producir una mayor intersección que los átomos del residuo original. En una realización, por cada uno de los residuos de reemplazo enumerados, el posible ligando se muta para producir un ligando mutado de único residuo modificado que incorpora el residuo de reemplazo. Luego, se puede ensayar la afinidad de cada ligando modificado con el sitio de unión de destino mediante un experimento para identificar las modificaciones que proporcionan la mejor afinidad para el posible ligando. Por ejemplo, se puede identificar un grupo de residuos de reemplazo que produce un residuo de reemplazo mayor que un umbral predeterminado. En una realización, el umbral predeterminado puede ser cero, de tal manera que el grupo seleccionado consiste únicamente de residuos de reemplazo que mejoran la afinidad hasta cierto punto. Luego, se pueden realizar modificaciones más avanzadas del posible ligando con base en esta información. Por ejemplo, en una realización se puede modificar el posible ligando para que incorpore una pluralidad de residuos de reemplazo, por ejemplo una pluralidad de esos residuos de reemplazo que se determinó que, individualmente, proporcionaban la mejor afinidad. De esta manera es posible diseñar un ligando con una afinidad mejorada aún más de lo posible solo al reemplazar un único residuo.

En algunas realizaciones, se pueden producir listas de residuos de reemplazo que cumplan con una mayor intersección que los átomos del residuo original (con un puntaje  $\Delta IOTA$  inferior a cero). Adicionalmente, se puede clasificar las listas con base en otros criterios. Por ejemplo, también se puede filtrar las listas con base en los puntajes  $\Delta\Delta G$  (ver abajo). Los residuos de reemplazos con los puntajes  $\Delta\Delta G$  de menos de cero implican interacciones más fuertes comparado con el residuo original. Por consiguiente, se puede producir una lista en donde los residuos cumplan con los dos criterios de un puntaje  $\Delta IOTA$  inferior a cero (puntaje  $\Delta IOTA$  negativo), y un  $\Delta\Delta G$  de menos de cero (un  $\Delta\Delta G$  negativo). Esto se ilustra en el Ejemplo 2 abajo.

En el caso de materia química, como por ejemplo una estructura de cristal de un fragmento químico de bajo peso molecular unido en un bolsillo de una proteína diana, las distribuciones y/o regiones favorecidas de ubicación teórica de átomo iota visualizadas en el sitio de unión pueden sugerir tipos de átomos y vectores de enlaces químicos para el crecimiento del fragmento que pueden producir un prototipo NCE de mayor potencia.

En una realización, se identifica una pluralidad de modificaciones al posible ligando. En este caso, el método también puede comprender la sección de un subgrupo de las modificaciones identificadas, como por ejemplo para identificar las modificaciones probablemente más efectivas en términos de mejorar la afinidad. La selección se puede llevar a cabo con base en la medida en que la intersección entre los átomos alternativos y/o del posible ligando y las respectivas regiones favorecidas del tipo de átomo iota es mayor en comparación con el posible candidato sin modificar. Por ejemplo, se pueden seleccionar las modificaciones que resultan en un aumento en la intersección que es superior a un umbral predeterminado y se pueden descartar las modificaciones que resultan en un aumento en la intersección que es inferior a un umbral predeterminado. Un ejemplo de dicho proceso de selección se discute a continuación en el contexto del "Ejemplo 2". El "Puntaje  $\Delta IOTA$ " es un ejemplo de una medida en que la intersección entre los átomos alternativos y/o adicionales del

posible ligando y las respectivas regiones favorecidas del tipo de átomo *iota* es mayor en comparación con el posible candidato sin modificar. Alternativa o adicionalmente, la selección se puede llevar a cabo con base en la medida en que uno o más factores que contribuyen a la energía total del complejo formado por la unión del posible ligando modificado al sitio de unión es/son reducido(s) comparado con los casos en los que el posible ligando no modificado está unido. Por ejemplo, pueden seleccionarse modificaciones que resulten en una disminución de los uno o más factores (por ejemplo, una disminución en una suma de los uno o más factores) por encima de un umbral predeterminado y se pueden descartar modificaciones que resulten en una disminución del uno o más factores (p. Ej. una disminución en una suma del uno o más factores) por debajo de un umbral predeterminado. Un ejemplo de dicho proceso de selección se discute a continuación en el contexto del “Ejemplo 2”. La “Puntaje Rosetta  $\Delta\Delta G$ ” es un ejemplo de una medición de la cuánto se reduce(n) uno o más factores que contribuyen a la energía total del complejo formado por la unión del posible ligando al sitio de unión. Ejemplos de factores que contribuyen a la energía total del complejo incluyen un término Lennard Jones, un término de solvatación implícita, un término de enlace de hidrógeno que depende de la orientación, potenciales de torsión de la cadena lateral y columna vertebral derivadas del PDB, un término electrostático de corto alcance basado en el conocimiento, y energías de referencia para cada uno de los 20 aminoácidos que modelan el estado desplegado, tal y como se discute a continuación.

En una realización, el método de identificación de una modificación a un posible ligando es un método implementado por computador. En una realización, cualquiera o más de los pasos S1-S7 se realizan por computador. En una realización, todos los pasos S1-S7 se realizan por computador. Adicionalmente, cualquier uno o más de los pasos S101-S109 de la Figura 14 (que ilustra un flujo de trabajo para predecir mutaciones en la interfaz VH-VL de un anticuerpo) pueden ser automatizados. Cualquier uno o más de los pasos S101-S109 se pueden realizar por computador. En

una realización, todos los pasos S101-S109 son automatizados. Todos los pasos S101-S109 se pueden realizar por computador.

Una amplia gama de configuraciones informáticas estándar, bien conocidas por la persona experta en la técnica, podría ser utilizada como plataformas para implementar el método. El método no está limitado a ninguna configuración de hardware, sistema operativo o medio de almacenamiento o transmisión de software particulares y necesarias para definir y/o implementar las etapas del método.

En una realización, se proporciona un medio o señal legible por computador que comprende instrucciones legibles por computador (p. Ej. un código en un lenguaje de programación) para hacer que un computador lleve a cabo el método.

En una realización, se proporciona un método de fabricación de un ligando terapéutico. En una realización, el método de fabricación comprende el diseño de un nuevo ligando o la modificación de un ligando existente de acuerdo con una o más de las realizaciones descritas arriba.

#### EJEMPLO 1: Maduración de afinidad de un fragmento Fab de un anticuerpo anti-IL7F

##### Introducción

Los métodos *in vitro* de maduración de afinidad de un anticuerpo son muy conocidos (ver US 8 303 953 B2 columna 13 líneas 19 a 33). En ejemplos recientes Fujino *et al* (Fujino *et al* (2012) “Robust *in vitro* affinity maturation strategy based on interface-focused high-throughput mutational scanning”, Biochem. Biophys. Res. Comm., 428, 395-400) reportan una estrategia de escaneo mutacional de alto rendimiento basada en una visualización ribosomal panorámica de librerías de fragmentos Fab mutantes de un sólo punto de una sola cadena en cada uno de los residuos antígenos de la interfaz del anticuerpo, seguido de una visualización ribosomal combinatoria de aglutinantes optimizados que resultaron en la identificación de un fragmento Fab con una mejora de 2000 veces en la afinidad. Dichos métodos requieren de una gran

inversión de recursos de laboratorio y por tanto varios grupos (reseñados por Kuroda *et al* (2012) en "Computer-aided antibody design", Prot. Eng. Design & Selection, 25, 507-521) han investigado métodos *in silico* que predicen mejoras en la afinidad del anticuerpo para así reducir o eliminar la necesidad de rastrear un gran número de variantes de anticuerpo mutadas para una afinidad mejorada. Estos protocolos de diseño de anticuerpos asistidos por computador son, por un lado, basados en el conocimiento; es decir, usan potenciales estadísticos derivados de datos de observación o, por otro, basados en física, por ejemplo usando funciones de energía derivadas de modelos de las interacciones físicas subyacentes. Lippow *et al* (2007) (Lippow *et al* (2007), "Computational design of antibody-affinity improvement beyond in vivo maturation", Nat Biotech., 25, 1171-1176) han logrado un éxito moderado con este último enfoque basado en las interacciones electrostáticas, pero nuestro entendimiento de parametrización de dichos métodos está aún muy lejos de ser completo. Los métodos basados en el conocimiento a la fecha tienden a identificar residuos individuales de anticuerpos para mutagénesis aleatoria (p. Ej. Barderas *et al* (2008) "...", PNAS, 105, 9029-9034), lo cual aún implica un trabajo considerable de laboratorio.

En este ejemplo, un enfoque basado en el conocimiento fue aplicado para madurar la afinidad de un fragmento Fab del anticuerpo anti-IL17F descrito en US 8 303 953 B2. El anticuerpo final con maduración de afinidad es descrito en WO 2012/ 095662 A1 como una molécula IgG1 de tamaño completo. Sin embargo, el método por el cual la afinidad del anticuerpo fue madurada no es revelado en esta última publicación.

## Métodos

### *Identificación de unos átomos (theta) objetivo que comprenden el epítipo IL17F*

Usando las coordenadas de la estructura compleja cocrislina IL17F / Fragmento Fab 496 descrita en WO 2009/ 130459 A2, todos los átomos IL17F dentro de 6 Å de cualquier átomo de fragmento Fab 496 fueron identificados como átomos epítopos y

están listados en la Tabla 1. De esta lista de 209 átomos theta, existen 86 tipos específicos de átomos theta.

*Tabla 1: Lista de átomos que comprenden el epítipo IL17F, donde la notación (1)-(2)-(3)-(4) designa (1) las respectivas cadenas F y 1 del homodímero IL17, (2) el residuo de aminoácido y (3) número de residuo y (4) el tipo de átomo.*

|              |              |               |              |
|--------------|--------------|---------------|--------------|
| F-ASN-53-CG  | F-ASN-89-C   | F-ILE-129-CA  | I-VAL-38-N   |
| F-ASN-53-OD1 | F-ASN-89-CA  | F-ILE-129-CB  | I-VAL-38-0   |
| F-ALA-70-C   | F-ASN-89-CB  | F-ILE-129-CD1 | I-SER-39-C   |
| F-ALA-70-0   | F-ASN-89-CG  | F-ILE-129-CG1 | I-SER-39-CA  |
| F-GLN-71-C   | F-ASN-89-N   | F-ILE-129-CG2 | I-SER-39-CB  |
| F-GLN-71-CA  | F-ASN-89-ND2 | F-ILE-129-N   | I-SER-39-N   |
| F-GLN-71-CB  | F-ASN-89-OD1 | F-ILE-129-0   | I-SER-39-OG  |
| F-GLN-71-CD  | F-SER-90-C   | F-HIS-130-C   | I-MET-40-C   |
| F-GLN-71-CG  | F-SER-90-CA  | F-HIS-130-N   | I-MET-40-N   |
| F-GLN-71-N   | F-SER-90-CB  | F-HIS-130-0   | I-MET-40-0   |
| F-GLN-71-NE2 | F-SER-90-N   | F-HIS-131-C   | I-SER-41-C   |
| F-GLN-71-0   | F-SER-90-0   | F-HIS-131-CA  | I-SER-41-CA  |
| F-GLN-71-OE1 | F-SER-90-OG  | F-HIS-131-CE1 | I-SER-41-CB  |
| F-CYS-72-C   | F-VAL-91-C   | F-HIS-131-N   | I-SER-41-N   |
| F-CYS-72-CA  | F-VAL-91-CA  | F-HIS-131-ND1 | I-SER-41-0   |
| F-CYS-72-CB  | F-VAL-91-CB  | F-HIS-131-0   | I-SER-41-OG  |
| F-CYS-72-N   | F-VAL-91-CG1 | F-VAL-132-C   | I-ARG-42-C   |
| F-CYS-72-0   | F-VAL-91-CG2 | F-VAL-132-CA  | I-ARG-42-CA  |
| F-CYS-72-SG  | F-VAL-91-N   | F-VAL-132-CB  | I-ARG-42-CB  |
| F-ARG-73-C   | F-VAL-91-0   | F-VAL-132-CG1 | I-ARG-42-CD  |
| F-ARG-73-CA  | F-PRO-92-C   | F-VAL-132-CG2 | I-ARG-42-CG  |
| F-ARG-73-CB  | F-PRO-92-CA  | F-VAL-132-N   | I-ARG-42-CZ  |
| F-ARG-73-CG  | F-PRO-92-CB  | F-VAL-132-0   | I-ARG-42-N   |
| F-ARG-73-N   | F-PRO-92-CD  | F-GLN-133-C   | I-ARG-42-NE  |
| F-ARG-73-0   | F-PRO-92-CG  | F-GLN-133-CA  | I-ARG-42-NH1 |
| F-ASN-74-C   | F-PRO-92-N   | F-GLN-133-CB  | I-ARG-42-NH2 |

|              |               |               |               |
|--------------|---------------|---------------|---------------|
| F-ASN-74-CA  | F-PRO-92-0    | F-GLN-133-CD  | I-ARG-42-0    |
| F-ASN-74-CB  | F-GLN-94-CD   | F-GLN-133-CG  | I-ASN-43-N    |
| F-ASN-74-CG  | F-GLN-94-NE2  | F-GLN-133-N   | I-ASN-43-OD1  |
| F-ASN-74-N   | F-GLN-94-OE1  | F-GLN-133-0   | 1-1 LE-44-CA  |
| F-ASN-74-ND2 | F-GLU-114-OE1 | F-GLN-133-OE1 | I-ILE-44-CB   |
| F-ASN-74-0   | F-LEU-117-CB  | F-GLN-133-OXT | I-ILE-44-CD1  |
| F-LEU-75-C   | F-LEU-117-CD1 | I-ILE-32-CG2  | I-ILE-44-CG1  |
| F-LEU-75-CA  | F-LEU-117-CD2 | I-ASN-33-CB   | 1-1 LE-44-CG2 |
| F-LEU-75-CB  | F-LEU-117-CG  | I-ASN-33-CG   | I-ARG-47-CD   |
| F-LEU-75-CD1 | F-THR-119-CB  | I-ASN-33-ND2  | I-ARG-47-CZ   |
| F-LEU-75-CD2 | F-THR-119-CG2 | I-ASN-33-OD1  | I-ARG-47-NE   |
| F-LEU-75-CG  | F-THR-119-OG1 | I-GLN-36-C    | I-ARG-47-NH1  |
| F-LEU-75-N   | F-VAL-125-CB  | I-GLN-36-CA   |               |
| F-LEU-75-0   | F-VAL-125-CG1 | I-GLN-36-CB   |               |
| F-GLU-84-OE1 | F-VAL-125-CG2 | I-GLN-36-CD   |               |
| F-ILE-86-C   | F-PRO-127-C   | I-GLN-36-NE2  |               |
| F-ILE-86-CA  | F-PRO-127-CA  | I-GLN-36-OE1  |               |
| F-ILE-86-CB  | F-PRO-127-CB  | I-ARG-37-C    |               |
| F-ILE-86-CD1 | F-PRO-127-CD  | I-ARG-37-CA   |               |
| F-ILE-86-CG1 | F-PRO-127-CG  | I-ARG-37-CB   |               |
| F-ILE-86-CG2 | F-PRO-127-N   | I-ARG-37-CD   |               |
| F-ILE-86-0   | F-PRO-127-0   | I-ARG-37-CG   |               |
| F-SER-87-C   | F-VAL-128-C   | I-ARG-37-CZ   |               |
| F-SER-87-CA  | F-VAL-128-CA  | I-ARG-37-N    |               |
| F-SER-87-N   | F-VAL-128-CB  | I-ARG-37-NE   |               |
| F-SER-87-0   | F-VAL-128-CG1 | I-ARG-37-NH1  |               |
| F-MET-88-C   | F-VAL-128-CG2 | I-ARG-37-NH2  |               |
| F-MET-88-CA  | F-VAL-128-N   | I-ARG-37-0    |               |
| F-MET-88-N   | F-VAL-128-0   | I-VAL-38-C    |               |
| F-MET-88-0   | F-ILE-129-C   | I-VAL-38-CA   |               |

### *Creación de base de datos de IOTA*

Una base de datos secundaria de más de 11 millones de datos de contacto atómico intramolecular fue extraída de más de 20000 estructuras proteínicas no homólogas, en donde la solución fue  $\approx 2 \text{ \AA}$ . Los contactos fueron definidos como dos átomos cualquiera en lados opuestos de un pliegue de proteínas separados por una distancia de  $1\text{\AA} +$  la suma de sus respectivos radios Van de Waals o menos, y estaba limitados a átomos sobre residuos al menos 4 residuos aparte en la secuencia peptídica lineal. El primer átomo del par de contacto fue designado como átomo theta y el segundo átomo, el átomo iota. La base de datos fue dividida en 167 grupos de contactos de acuerdo con el tipo del átomo theta, habiendo allí 167 tipos de átomos distintos de hidrógeno dentro de los 20 residuos naturales de aminoácidos que contienen proteínas. Dentro de un grupo de contactos, las coordenadas relativas de cada posición del átomo iota fueron grabadas luego de la normalización de las coordenadas pares theta -iota. Lo anterior se consiguió al ubicar al átomo theta en  $X,Y,Z = 0,0,0$ ; el siguiente átomo unido covalentemente (3<sup>er</sup> átomo) al átomo theta (en la dirección del esqueleto peptídico) en  $X,Y,Z = x',0,0$  y el siguiente nuevamente unido covalentemente (4<sup>to</sup> átomo) en  $X,Y,Z = x'',0,z'$ . Una convención consistente fue usada para definir el tercer y cuarto átomo.

Cada conjunto de contactos theta fue subdividido en 167 tipos de átomo iota, pero por conveniencia estos fueron concatenados en 26 subgrupos de acuerdo con su tipo químico basado en la definición de Engh y Huber (Engh y Huber (1991) "Accurate Bond and Angle Parameters for X-ray Protein Structure Refinement", Acta Cryst., A47, 392-400).

En este ejemplo, los siguientes subgrupos iota fueron utilizados:

| Descripción     | Átomos                                                                        |
|-----------------|-------------------------------------------------------------------------------|
| Carbonilo O     | todos esqueletos de O, Asn Od1, Gln Oel                                       |
| CH2 tetraédrico | todo C $\beta$ 9 excepto Ala, Ile, Thr, Val) Arg C d, Cy, Gln Cy, Glu Cy, Ile |

|                 |                                                                      |
|-----------------|----------------------------------------------------------------------|
|                 | Cyl, Lys Cy, C d, Ce, Met Cy                                         |
| CH3 tetraédrico | Ala Cß, lle Cdl , Cy2, Leu Cdl , Cd2, Met Ce, Thr Cy2, Val Cy 1, Cy2 |
| NH              | todos esqueletos N (excepto Pro), Arg Ne, His Ndl ,Ne2, Trp Nel      |
| Hidroxilo O     | Ser Oy, Thr Oyl , Tyr On                                             |
| Carboxilo O     | Asp Od1, Od2, Glu Oel , Oe2                                          |
| NH2             | Asn Nd2, Gln Ne2                                                     |

### *Superposición de los datos iota sobre la superficie del epítipo IL17F*

Para cada uno de los 209 átomos theta que comprenden el epítipo IL17F, el conjunto de contacto theta correspondiente fue seleccionado de la base de datos IOTA y, a partir de esta, se seleccionó un subgrupo iota apropiado, por ejemplo oxígeno de carbonilo. Las coordenadas relativas iota de este subgrupo fueron transpuestas respecto al marco de referencia del átomo theta del epítipo IL17F proporcionado (la Tabla 2 muestra datos de ejemplo). Un conjunto de datos iota para un subgrupo dado fue entonces acumulado a lo largo de todo el epítipo IL17F. En casos en los que la ubicación de un punto de datos iota dado intersectaba con un átomo de IL17F, más cerca que la suma de sus respectivos radios Van de Waals menos  $0.2\text{\AA}$ , entonces estos puntos de datos eran excluidos del conjunto de datos. El proceso se repitió para todos los subgrupos iota relevantes para producir una serie de conjuntos de datos iota para el epítipo IL17F.

Table 2:

| id       | out01:0000098  | out01:0000127  | out01:0000249 | out01:0000280    | out01:0000282    | out01:0000283    | out01:0000346   | out01:0000366   | out01:0000527 | out01:0000542 |
|----------|----------------|----------------|---------------|------------------|------------------|------------------|-----------------|-----------------|---------------|---------------|
| pdcode   | 16vpA          | 16vpA          | 1a0cA         | 1a0pA            | 1a0pA            | 1a0pA            | 1a0sP           | 1a0sP           | 1a12A         | 1a12A         |
| res      | 2.1            | 2.1            | 2.5           | 2.5              | 2.5              | 2.5              | 2.4             | 2.4             | 1.7           | 1.7           |
| rval     | 0.26           | 0.26           | 0.177         | 0.287            | 0.287            | 0.287            | 0.228           | 0.228           | 0.219         | 0.219         |
| org      | Herpes simplex | Herpes simplex | Thermoaerob   | Escherichia coli | Escherichia coli | Escherichia coli | Salmonella typh | Salmonella typh | Homo sapiens  | Homo sapiens  |
| Tchain   | A              | A              | A             | A                | A                | A                | P               | P               | A             | A             |
| Taalnd   | 96             | 235            | 36            | 126              | 137              | 137              | 85              | 138             | 257           | 12            |
| Taanum   | 142            | 281            | 36            | 145              | 156              | 156              | 155             | 208             | 277           | 32            |
| Taaname  | THR            | THR            | THR           | THR              | THR              | THR              | THR             | THR             | THR           | THR           |
| Taass    | C              | E              | E             | H                | E                | E                | H               | E               | C             | C             |
| Taaphi   | -135.7         | -129.6         | -71.7         | -87.1            | -105.2           | -105.2           | -118.5          | -123.9          | -114.7        | -127.1        |
| Taapsi   | 163.3          | 160            | 157.7         | -34.9            | 174              | 174              | 18.9            | 132.4           | 1             | -10.7         |
| Tomg     | -179.5         | 178.5          | 177.5         | 178.7            | 177.6            | 177.6            | -178.3          | 179.3           | -178.5        | -171.7        |
| Tc1      | 61.3           | 52.2           | 66.6          | 60.5             | 64.4             | 64.4             | 53.1            | -54.5           | 49.2          | 59.2          |
| Tc2      | 999.9          | 999.9          | 999.9         | 999.9            | 999.9            | 999.9            | 999.9           | 999.9           | 999.9         | 999.9         |
| Tc3      | 999.9          | 999.9          | 999.9         | 999.9            | 999.9            | 999.9            | 999.9           | 999.9           | 999.9         | 999.9         |
| Tc4      | 999.9          | 999.9          | 999.9         | 999.9            | 999.9            | 999.9            | 999.9           | 999.9           | 999.9         | 999.9         |
| TUpAA    | ALA            | LEU            | LYS           | ALA              | LEU              | LEU              | SER             | GLY             | GLY           | SER           |
| TDnAA    | ARG            | VAL            | MET           | GLY              | MET              | MET              | SER             | GLY             | GLU           | GLU           |
| Tnum     | 724            | 1867           | 288           | 1032             | 1111             | 1111             | 626             | 1048            | 1919          | 99            |
| Tname    | C              | C              | C             | C                | C                | C                | C               | C               | C             | C             |
| Tbval    | 23.52          | 24.07          | 23.92         | 31.8             | 38.02            | 38.02            | 16.02           | 15.07           | 20.11         | 12.34         |
| Tasa     | 0              | 0              | 0             | 0                | 0                | 0                | 1.84            | 0.19            | 0             | 0.21          |
| Tcdist   | 0              | 0              | 0             | 0                | 0                | 0                | 0               | 0               | 0             | 0             |
| Todist   | 1.234          | 1.233          | 1.24          | 1.215            | 1.239            | 1.239            | 1.226           | 1.222           | 1.247         | 1.249         |
| Tndist   | 2.463          | 2.461          | 2.432         | 2.482            | 2.455            | 2.455            | 2.536           | 2.412           | 2.492         | 2.502         |
| Tcadist  | 1.525          | 1.532          | 1.527         | 1.525            | 1.53             | 1.53             | 1.541           | 1.519           | 1.516         | 1.521         |
| Icdist   | 3.472          | 3.089          | 3.359         | 3.618            | 3.647            | 3.199            | 3.565           | 3.589           | 3.432         | 3.67          |
| Iodist   | 3.677          | 3.072          | 4.523         | 3.286            | 3.657            | 4.381            | 2.796           | 2.512           | 3.207         | 3.652         |
| Indist   | 4.507          | 4.556          | 4.306         | 4.708            | 4.178            | 4.537            | 5.735           | 3.963           | 4.235         | 4.639         |
| Itidist  | 3.472          | 3.089          | 3.359         | 3.618            | 3.647            | 3.199            | 3.565           | 3.589           | 3.432         | 3.67          |
| Inum     | 1644           | 1896           | 244           | 2089             | 1535             | 1519             | 550             | 1213            | 1845          | 372           |
| Iname    | CD1            | OD1            | O             | CD1              | CB               | O                | NE2             | N               | O             | CG1           |
| Ichain   | A              | A              | A             | A                | A                | A                | P               | P               | A             | A             |
| Iaalnd   | 206            | 213            | 30            | 258              | 190              | 188              | 75              | 159             | 247           | 50            |
| Iaanum   | 252            | 259            | 30            | 279              | 209              | 207              | 145             | 229             | 267           | 70            |
| Iaaname  | LEU            | ASP            | GLU           | TYR              | LEU              | ASP              | GLN             | GLY             | SER           | VAL           |
| Iaass    | H              | E              | C             | H                | C                | C                | E               | E               | C             | E             |
| Iaaphi   | -94.9          | -110.9         | -60.4         | -69.5            | -73.2            | -83.2            | -125.6          | -109.2          | -149.7        | -101.1        |
| Iaapsi   | -34.2          | 127.4          | 132.7         | -18.6            | -32.6            | -14.4            | 142.7           | 141.1           | 38.2          | -21           |
| Iomg     | -173.5         | -178.2         | -179.6        | -177.8           | -177.1           | -179.8           | 177.8           | 178.2           | -179.1        | 179.5         |
| Ic1      | -63.9          | -171.8         | 176.2         | -79.4            | 167.9            | 66               | -66.4           | 999.9           | -164          | -65.2         |
| Ic2      | 167.9          | 57.6           | 172.1         | -57.1            | 69.8             | -21.2            | -63.5           | 999.9           | 999.9         | 999.9         |
| Ic3      | 999.9          | 999.9          | 7             | 999.9            | 999.9            | 999.9            | -52.7           | 999.9           | 999.9         | 999.9         |
| Ic4      | 999.9          | 999.9          | 999.9         | 999.9            | 999.9            | 999.9            | 999.9           | 999.9           | 999.9         | 999.9         |
| IUpAA    | ASP            | CYS            | GLU           | ILE              | VAL              | ILE              | GLY             | TYR             | LEU           | VAL           |
| IDnAA    | PHE            | LEU            | VAL           | THR              | PHE              | VAL              | THR             | ARG             | ASN           | GLN           |
| Ibval    | 6.09           | 32.79          | 27.18         | 37.22            | 8.29             | 43.39            | 11.84           | 20.76           | 9.22          | 8.96          |
| Iasa     | 0.13           | 0              | 0             | 0                | 0                | 0                | 0.69            | 0.25            | 0             | 1.04          |
| K1cdist  | 4.988          | 4.114          | 4.557         | 4.099            | 4.209            | 4.342            | 4.034           | 4.42            | 3.962         | 4.187         |
| K1odist  | 5.113          | 3.819          | 5.728         | 3.526            | 4.6              | 5.553            | 3.425           | 3.283           | 3.395         | 4.575         |
| K1ndist  | 5.733          | 5.402          | 5.198         | 5.704            | 4.874            | 5.334            | 5.781           | 4.513           | 5.257         | 4.346         |
| K1tdist  | 4.988          | 4.114          | 4.557         | 4.099            | 4.209            | 4.342            | 4.034           | 4.42            | 3.962         | 4.187         |
| K1num    | 1642           | 1895           | 243           | 2088             | 1532             | 1518             | 548             | 1214            | 1844          | 371           |
| K1name   | CG             | CG             | C             | CG               | CA               | C                | CD              | CA              | C             | CB            |
| K1aanum  | 252            | 259            | 30            | 279              | 209              | 207              | 145             | 229             | 267           | 70            |
| K1aaname | LEU            | ASP            | GLU           | TYR              | LEU              | ASP              | GLN             | GLY             | SER           | VAL           |
| K2cdist  | NoValue        | NoValue        | NoValue       | 3.857            | NoValue          | NoValue          | NoValue         | NoValue         | NoValue       | NoValue       |
| K2odist  | NoValue        | NoValue        | NoValue       | 3.951            | NoValue          | NoValue          | NoValue         | NoValue         | NoValue       | NoValue       |
| K2ndist  | NoValue        | NoValue        | NoValue       | 4.353            | NoValue          | NoValue          | NoValue         | NoValue         | NoValue       | NoValue       |
| K2tdist  | NoValue        | NoValue        | NoValue       | 3.857            | NoValue          | NoValue          | NoValue         | NoValue         | NoValue       | NoValue       |
| K2num    | NoValue        | NoValue        | NoValue       | 2091             | NoValue          | NoValue          | NoValue         | NoValue         | NoValue       | NoValue       |
| K2name   | NoValue        | NoValue        | NoValue       | CE1              | NoValue          | NoValue          | NoValue         | NoValue         | NoValue       | NoValue       |
| K2aanum  | NoValue        | NoValue        | NoValue       | 279              | NoValue          | NoValue          | NoValue         | NoValue         | NoValue       | NoValue       |
| K2aaname | NoValue        | NoValue        | NoValue       | TYR              | NoValue          | NoValue          | NoValue         | NoValue         | NoValue       | NoValue       |
| Cx       | 0              | 0              | 0             | 0                | 0                | 0                | 0               | 0               | 0             | 0             |
| Cy       | 0              | 0              | 0             | 0                | 0                | 0                | 0               | 0               | 0             | 0             |
| Cz       | 0              | 0              | 0             | 0                | 0                | 0                | 0               | 0               | 0             | 0             |
| CsphR    | 0              | 0              | 0             | 0                | 0                | 0                | 0               | 0               | 0             | 0             |
| CsphT    | 0              | 0              | 0             | 0                | 0                | 0                | 0               | 0               | 0             | 0             |
| CsphP    | -3.142         | 0              | 0             | 3.142            | 0                | 0                | 0               | 0               | 0             | 0             |

Table 2 (continued):

| Id     | out01:0000098 | out01:0000127 | out01:0000249 | out01:0000280 | out01:0000282 | out01:0000283 | out01:0000346 | out01:0000366 | out01:0000527 | out01:0000542 |
|--------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| CcylR  | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             |
| CcylT  | -3.142        | 0             | 0             | 3.142         | 0             | 0             | 0             | 0             | 0             | 0             |
| CcylZ  | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             |
| Ox     | -0.628        | -0.629        | -0.648        | -0.659        | -0.666        | -0.666        | -0.609        | -0.619        | -0.609        | -0.582        |
| Oy     | 0.998         | 1.004         | 0.978         | -0.754        | 1.028         | 1.028         | -1.009        | 0.713         | -1.084        | -1.095        |
| Oz     | 0.364         | 0.341         | 0.401         | -0.689        | 0.182         | 0.182         | 0.341         | 0.777         | 0.085         | -0.146        |
| OsphR  | 1.234         | 1.233         | 1.24          | 1.216         | 1.238         | 1.238         | 1.227         | 1.223         | 1.246         | 1.249         |
| OsphT  | 1.271         | 1.291         | 1.241         | 2.173         | 1.423         | 1.423         | 1.289         | 0.882         | 1.503         | 1.688         |
| OsphP  | 2.332         | 2.13          | 2.156         | -2.289        | 2.146         | 2.146         | -2.114        | 2.286         | -2.083        | -2.059        |
| OcylR  | 1.234         | 1.233         | 1.24          | 1.216         | 1.238         | 1.238         | 1.227         | 1.223         | 1.246         | 1.249         |
| OcylT  | 2.332         | 2.13          | 2.156         | -2.289        | 2.146         | 2.146         | -2.114        | 2.286         | -2.083        | -2.059        |
| OcylZ  | 0.364         | 0.341         | 0.401         | -0.689        | 0.182         | 0.182         | 0.341         | 0.777         | 0.085         | -0.146        |
| Nx     | 2.055         | 2.054         | 2.006         | 2.096         | 2.039         | 2.039         | 2.158         | 1.982         | 2.11          | 2.128         |
| Ny     | 1.357         | 1.356         | 1.375         | 1.328         | 1.368         | 1.368         | 1.332         | 1.374         | 1.327         | 1.316         |
| Nz     | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             |
| NsphR  | 2.463         | 2.461         | 2.432         | 2.481         | 2.455         | 2.455         | 2.536         | 2.412         | 2.493         | 2.502         |
| NsphT  | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         |
| NsphP  | 0.584         | 0.583         | 0.601         | 0.565         | 0.591         | 0.591         | 0.553         | 0.606         | 0.561         | 0.554         |
| NcylR  | 2.463         | 2.461         | 2.432         | 2.481         | 2.455         | 2.455         | 2.536         | 2.412         | 2.493         | 2.502         |
| NcylT  | 0.584         | 0.583         | 0.601         | 0.565         | 0.591         | 0.591         | 0.553         | 0.606         | 0.561         | 0.554         |
| NcylZ  | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             |
| Cax    | 1.525         | 1.532         | 1.527         | 1.525         | 1.53          | 1.53          | 1.541         | 1.519         | 1.516         | 1.521         |
| Cay    | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             |
| Caz    | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             |
| CasphR | 1.525         | 1.532         | 1.527         | 1.525         | 1.53          | 1.53          | 1.541         | 1.519         | 1.516         | 1.521         |
| CasphT | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         | 1.571         |
| CasphP | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             |
| CacyR  | 1.525         | 1.532         | 1.527         | 1.525         | 1.53          | 1.53          | 1.541         | 1.519         | 1.516         | 1.521         |
| CacyT  | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             |
| CacyZ  | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             | 0             |
| Tx     | -12.209       | 3.16          | 75.678        | 17.53         | 11.229        | 11.229        | -34.605       | -58.479       | 39.484        | 5.281         |
| Ty     | -17.125       | -7.179        | 45.43         | -33.077       | -23.148       | -23.148       | -6.201        | -0.92         | -28.232       | -11.189       |
| Tz     | -3.618        | -14.31        | 44.154        | 38.22         | 26.483        | 26.483        | -24.486       | -14.58        | 0.224         | 9.116         |
| TsphR  |               |               |               |               |               |               |               |               |               |               |
| TsphT  |               |               |               |               |               |               |               |               |               |               |
| TsphP  |               |               |               |               |               |               |               |               |               |               |
| TcylR  |               |               |               |               |               |               |               |               |               |               |
| TcylT  |               |               |               |               |               |               |               |               |               |               |
| TcylZ  |               |               |               |               |               |               |               |               |               |               |
| Ix     | -0.948        | -0.804        | 1.174         | -1.669        | -0.297        | 0.852         | -2.689        | -1.363        | 1.533         | -0.362        |
| Iy     | 0.629         | -0.729        | -2.203        | 1.536         | 1.127         | -1.85         | -0.806        | 3.054         | -2.417        | -0.092        |
| Iz     | -3.281        | 2.871         | -2.247        | -2.818        | -3.455        | -1.176        | 2.198         | 1.302         | -1.893        | -3.651        |
| IsphR  | 3.473         | 3.069         | 3.359         | 3.617         | 3.646         | 3.199         | 3.565         | 3.589         | 3.432         | 3.67          |
| IsphT  | 2.808         | 0.361         | 2.304         | 2.464         | 2.816         | 1.947         | 0.907         | 1.2           | 2.155         | 3.04          |
| IsphP  | 2.556         | -2.405        | -1.081        | 2.398         | 1.828         | -1.28         | -2.85         | 1.991         | -1.006        | -2.893        |
| IcylR  | 3.473         | 3.069         | 3.359         | 3.617         | 3.646         | 3.199         | 3.565         | 3.589         | 3.432         | 3.67          |
| IcylT  | 2.556         | -2.405        | -1.081        | 2.398         | 1.828         | -1.28         | -2.85         | 1.991         | -1.006        | -2.893        |
| IcylZ  | -3.281        | 2.871         | -2.247        | -2.818        | -3.455        | -1.176        | 2.198         | 1.302         | -1.893        | -3.651        |
| K1x    | -1.194        | -1.35         | 1.899         | -2.979        | 0.085         | 1.65          | -2.134        | -1.06         | 0.644         | 0.446         |
| K1y    | 1.1           | -0.243        | -2.896        | 1.098         | -0.13         | -3.76         | -0.567        | 3.344         | -3.182        | 1.142         |
| K1z    | -4.717        | 3.878         | -2.962        | -2.593        | -4.206        | -1.413        | 3.376         | 2.69          | -2.271        | -4.003        |
| K1sphR | 4.989         | 4.113         | 4.557         | 4.099         | 4.209         | 4.342         | 4.034         | 4.421         | 3.962         | 4.187         |
| K1sphT | 2.81          | 0.34          | 2.278         | 2.256         | 3.105         | 1.902         | 0.579         | 0.917         | 2.181         | 2.844         |
| K1sphP | 2.397         | -2.963        | -0.99         | 2.788         | -0.992        | -1.157        | -2.882        | 1.878         | -1.371        | 1.198         |
| K1cylR | 4.989         | 4.113         | 4.557         | 4.099         | 4.209         | 4.342         | 4.034         | 4.421         | 3.962         | 4.187         |
| K1cylT | 2.397         | -2.963        | -0.99         | 2.788         | -0.992        | -1.157        | -2.882        | 1.878         | -1.371        | 1.198         |
| K1cylZ | -4.717        | 3.878         | -2.962        | -2.593        | -4.206        | -1.413        | 3.376         | 2.69          | -2.271        | -4.003        |
| K2x    | NoValue       | NoValue       | NoValue       | -1.258        | NoValue       | NoValue       | NoValue       | NoValue       | NoValue       | NoValue       |
| K2y    | NoValue       | NoValue       | NoValue       | 2.771         | NoValue       | NoValue       | NoValue       | NoValue       | NoValue       | NoValue       |
| K2z    | NoValue       | NoValue       | NoValue       | -2.37         | NoValue       | NoValue       | NoValue       | NoValue       | NoValue       | NoValue       |
| K2sphR |               |               |               | 3.857         |               |               |               |               |               |               |
| K2sphT |               |               |               | 2.232         |               |               |               |               |               |               |
| K2sphP |               |               |               | 1.997         |               |               |               |               |               |               |
| K2cylR |               |               |               | 3.857         |               |               |               |               |               |               |
| K2cylT |               |               |               | 1.997         |               |               |               |               |               |               |
| K2cylZ |               |               |               | -2.37         |               |               |               |               |               |               |
| Inter  | MS            | MS            | MM            | MS            | MS            | MM            | MS            | MM            | MM            | MS            |
| Icount | 1             | 6             | 3             | 4             | 4             | 4             | 4             | 2             | 4             | 4             |
| Tcount | 1             | 1             | 1             | 1             | 2             | 2             | 1             | 1             | 1             | 1             |

| <b>Tabla 2: Clave</b> |                                                                                                                                                                                |
|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Encabezado</b>     | <b>Descripción</b>                                                                                                                                                             |
| Id                    | Identificador incrementado de un contacto. Cada identificador único es una interacción theta-iota única. El id es ocho dígitos con ceros (0) iniciales (p. Ej. 1 es 00000001). |
| pdbcode               | El código PDB que se dio en el archivo de entrada. (p. Ej. lmu4B para la cadena B de lmu4, o lmu4 para la estructura).                                                         |
| res                   | La resolución de la estructura PDB.                                                                                                                                            |
| rval                  | El valor R de la estructura PDB.                                                                                                                                               |
| org                   | Fuente del organismo para la estructura PDB.                                                                                                                                   |
| Tchain                | El identificador de cadena para el átomo theta.                                                                                                                                |
| Taaind                | El índice de aminoácidos para el aminoácido theta (el índice empieza en 1 para el primer aminoácido en la estructura y aumenta con cada aminoácido)                            |
| Taanum                | El número de aminoácidos del átomo theta.                                                                                                                                      |
| Taaname               | El nombre del aminoácido del átomo theta.                                                                                                                                      |
| Taass                 | La estructura secundaria del aminoácido del átomo theta.                                                                                                                       |
| Taaphi                | El ángulo phi del aminoácido del átomo theta.                                                                                                                                  |
| Taapsi                | El ángulo psi del aminoácido del átomo theta.                                                                                                                                  |
| Tomg                  | El ángulo omega del aminoácido del átomo theta.                                                                                                                                |
| Tchi1                 | El ángulo chi 1 del aminoácido del átomo theta.                                                                                                                                |
| Tchi2                 | El ángulo chi 2 del aminoácido del átomo theta.                                                                                                                                |
| Tchi3                 | El ángulo chi 3 del aminoácido del átomo theta.                                                                                                                                |
| Tchi4                 | El ángulo chi 4 del aminoácido del átomo theta.                                                                                                                                |
| TUpAA                 | El aminoácido que está corriente arriba del aminoácido                                                                                                                         |
| TDnAA                 | El aminoácido que está corriente abajo del aminoácido                                                                                                                          |
| Tnum                  | El número de átomos para el átomo theta.                                                                                                                                       |
| Tname                 | El nombre del átomo para el átomo theta.                                                                                                                                       |
| Tbval                 | El valor B o el factor de temperatura para el átomo theta.                                                                                                                     |
| Tasa                  | El Área de Superficie Accesible del átomo theta.                                                                                                                               |
| Tcdist                | La distancia del átomo theta al Carbono estructural.                                                                                                                           |
| Todist                | La distancia del átomo theta al Oxígeno estructural.                                                                                                                           |
| Tndist                | La distancia del átomo theta al Nitrógeno estructural.                                                                                                                         |
| Tcadist               | La distancia del átomo theta al Carbono Alfa estructural.                                                                                                                      |
| Icdist                | La distancia del átomo iota al Carbono estructural del aminoácido theta.                                                                                                       |
| Iodist                | La distancia del átomo iota al Oxígeno estructural del aminoácido theta.                                                                                                       |
| Indist                | La distancia del átomo iota al nitrógeno estructural del aminoácido theta.                                                                                                     |
| Itdist                | La distancia del átomo iota al átomo theta.                                                                                                                                    |
| Inum                  | El número de átomos del átomo iota.                                                                                                                                            |

|           |                                                                                                                                                    |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------|
| Iname     | El nombre del átomo para el átomo iota.                                                                                                            |
| Ichain    | El identificador de cadena para el átomo iota.                                                                                                     |
| Iaaind    | El índice de aminoácidos para el aminoácido iota (el índice empieza en 1 para el primer aminoácido en la estructura y aumenta por cada aminoácido) |
| Ia anum   | El número de aminoácidos del átomo iota.                                                                                                           |
| Iaaname   | El nombre del aminoácido del átomo iota.                                                                                                           |
| Iaass     | La estructura secundaria del aminoácido del átomo iota.                                                                                            |
| Iaaphi    | El ángulo phi del aminoácido del átomo iota.                                                                                                       |
| Iaapsi    | El aminoácido ángulo psi átomo iota.                                                                                                               |
| Iomg      | El ángulo omega del aminoácido para el átomo iota.                                                                                                 |
| Ichi 1    | El ángulo chi 1 del aminoácido para el átomo iota.                                                                                                 |
| Ichi2     | El ángulo chi 2 del aminoácido para el átomo iota.                                                                                                 |
| Ichi3     | El ángulo chi 3 del aminoácido para el átomo iota.                                                                                                 |
| Ichi4     | El ángulo chi 4 del aminoácido para el átomo iota.                                                                                                 |
| IUpAA     | El aminoácido que está corriente arriba del aminoácido iota.                                                                                       |
| IDnAA     | El aminoácido que está corriente abajo del aminoácido iota.                                                                                        |
| Ibval     | El valor B o factor de temperatura del átomo iota.                                                                                                 |
| lasa      | El Área de Superficie Accesible para el átomo iota.                                                                                                |
| Kl cdist  | La distancia del primer átomo kappa al Carbono estructural del aminoácido theta.                                                                   |
| Klodist   | La distancia del primer átomo kappa al Oxígeno estructural del aminoácido theta.                                                                   |
| Kl ndist  | La distancia del primer átomo kappa al nitrógeno estructural del aminoácido theta.                                                                 |
| Kl tdist  | La distancia del primer átomo kappa al átomo theta.                                                                                                |
| Kl num    | El número de átomos del primer átomo kappa.                                                                                                        |
| Kl name   | El nombre del átomo del primer átomo kappa.                                                                                                        |
| Kl anum   | El número de aminoácidos del primer átomo kappa.                                                                                                   |
| Kl aaname | El nombre del aminoácido del primer átomo kappa.                                                                                                   |
| K2cdist   | La distancia del segundo átomo kappa al Carbono estructural del aminoácido theta.                                                                  |
| K2odist   | La distancia del segundo átomo kappa al Oxígeno estructural del aminoácido theta.                                                                  |
| K2ndist   | La distancia del segundo átomo kappa al Nitrógeno estructural del aminoácido theta.                                                                |
| K2tdist   | La distancia del segundo átomo kappa al átomo theta.                                                                                               |
| K2num     | El número de átomos del segundo átomo kappa.                                                                                                       |
| K2name    | El nombre del átomo del segundo átomo kappa.                                                                                                       |
| K2aanum   | El número de aminoácidos del segundo átomo kappa.                                                                                                  |
| K2aaname  | El nombre del aminoácido para el segundo átomo kappa.                                                                                              |

|       |                                                                                                 |
|-------|-------------------------------------------------------------------------------------------------|
| Cx    | La coordenada x theta superpuesta para el átomo de Carbono estructural de theta AA.             |
| Cy    | La coordenada y theta superpuesta para el átomo de Carbono estructural de theta AA.             |
| Cz    | La coordenada z theta superpuesta para el átomo de Carbono estructural de theta AA.             |
| CsphR | La distancia polar esférica theta superpuesta para el Carbono estructural de theta AA.          |
| CsphT | El ángulo de latitud polar esférica theta superpuesta para el Carbono estructural de theta AA.  |
| CsphP | El ángulo de longitud polar esférica theta superpuesta para el Carbono estructural de theta AA. |
| CcylR | La distancia polar cilíndrica theta superpuesta para el Carbono estructural de theta AA.        |
| CcylT | El ángulo polar cilíndrico theta superpuesta para el Carbono estructural de theta AA.           |
| CcylZ | La coordenada z polar cilíndrica theta superpuesta para el Carbono estructural de theta AA.     |
| Ox    | La coordenada x theta superpuesta para el átomo de oxígeno estructural de theta AA.             |
| Oy    | La coordenada y theta superpuesta para el átomo de oxígeno estructural de theta AA.             |
| Oz    | La coordenada z theta superpuesta para el átomo de oxígeno estructural de theta AA.             |
| OsphR | La distancia polar esférica theta superpuesta para el oxígeno estructural de theta AA.          |
| OsphT | El ángulo de latitud polar esférica theta superpuesta para el oxígeno estructural de theta AA.  |
| OsphP | El ángulo de longitud polar esférica theta superpuesta para el oxígeno estructural de theta AA. |
| OcylR | La distancia polar cilíndrica theta superpuesta para el oxígeno estructural de theta AA.        |
| OcylT | El ángulo polar cilíndrico theta superpuesto para el oxígeno estructural de theta AA.           |
| OcylZ | La coordenada z polar cilíndrica theta superpuesta para el oxígeno estructural de theta AA.     |
| Nx    | La coordenada x theta superpuesta para el átomo de nitrógeno estructural de theta AA.           |
| Ny    | La coordenada y theta superpuesta para el átomo de nitrógeno estructural de theta AA.           |
| Nz    | La coordenada z theta superpuesta para el átomo de nitrógeno estructural de theta AA.           |

|        |                                                                                                   |
|--------|---------------------------------------------------------------------------------------------------|
| NsphR  | La distancia polar esférica theta superpuesta para el nitrógeno estructural de theta AA.          |
| NsphT  | El ángulo de latitud polar esférica theta superpuesta para el nitrógeno estructural de theta AA.  |
| NsphP  | El ángulo de longitud polar esférica theta superpuesto para el nitrógeno estructural de theta AA. |
| NcylR  | La distancia polar cilíndrica theta superpuesta para el nitrógeno estructural de theta AA.        |
| NcylT  | El ángulo polar cilíndrico theta superpuesto para el nitrógeno estructural de theta AA.           |
| NcylZ  | La coordenada z polar cilíndrica theta superpuesta para el nitrógeno estructural de theta AA.     |
| Cax    | La coordenada x theta superpuesta para el átomo de Carbono Alfa de theta AA.                      |
| Cay    | La coordenada y theta superpuesta para el átomo de Carbono Alfa de theta AA.                      |
| Caz    | La coordenada z theta superpuesta para el átomo de Carbono Alfa de theta AA.                      |
| CasphR | La distancia polar esférica theta superpuesta para el Carbono Alfa de theta AA.                   |
| CasphT | El ángulo de latitud polar esférica theta superpuesto para el Carbono Alfa de theta AA.           |
| CasphP | El ángulo de longitud polar esférica theta superpuesto para el Carbono Alfa de theta AA.          |
| CacylR | La distancia polar cilíndrica theta superpuesta para el Carbono Alfa de theta AA.                 |
| CacylT | El ángulo polar cilíndrico theta superpuesto para el Carbono Alfa de theta AA.                    |
| CacylZ | La coordenada z polar cilíndrica theta superpuesta para el Carbono Alfa de theta AA.              |
| Tx     | La coordenada x theta superpuesta para el átomo theta.                                            |
| Ty     | La coordenada y theta superpuesta para el átomo theta.                                            |
| Tz     | La coordenada z theta superpuesta para el átomo theta.                                            |
| TsphR  | La distancia polar esférica theta superpuesta para el átomo                                       |
| TsphT  | El ángulo de latitud polar esférica theta superpuesto para el átomo theta.                        |
| TsphP  | El ángulo de longitud polar esférica theta superpuesto para el átomo theta.                       |
| TcylR  | La distancia polar cilíndrica theta superpuesta para el átomo theta.                              |
| TcylT  | El ángulo polar cilíndrico theta superpuesto para el átomo                                        |
| TcylZ  | La coordenada z polar cilíndrica theta superpuesta para el átomo theta.                           |

|          |                                                                                     |
|----------|-------------------------------------------------------------------------------------|
| Ix       | La coordenada x theta superpuesta para el átomo iota.                               |
| Iy       | La coordenada y theta superpuesta para el átomo iota.                               |
| Iz       | La coordenada z theta superpuesta para el átomo iota.                               |
| IsphR    | La distancia polar esférica theta superpuesta para el átomo                         |
| IsphT    | El ángulo de latitud polar esférica theta superpuesto para el átomo iota.           |
| IsphP    | El ángulo de longitud polar esférica theta superpuesto para el átomo iota.          |
| IcylR    | La distancia polar cilíndrica theta superpuesta para el átomo                       |
| IcylT    | El ángulo polar cilíndrico theta superpuesto para el átomo                          |
| IcylZ    | La coordenada z polar cilíndrica theta superpuesta para el átomo iota.              |
| K l x    | La coordenada x theta superpuesta para el primer átomo                              |
| K l y    | La coordenada y theta superpuesta para el primer átomo                              |
| K l z    | La coordenada z theta superpuesta para el primer átomo                              |
| K l sphR | La distancia polar esférica theta superpuesta para el primer átomo kappa.           |
| K l sphT | El ángulo de latitud polar esférica theta superpuesto para el primer átomo kappa.   |
| K l sphP | El ángulo de longitud polar esférica theta superpuesto para el primer átomo kappa.  |
| K l cylR | La distancia polar cilíndrica theta superpuesta para el primer átomo kappa.         |
| K l cylT | El ángulo polar cilíndrico theta superpuesto para el primer átomo kappa.            |
| K l cylZ | La coordenada z polar cilíndrica theta superpuesta para el primer átomo kappa.      |
| K2x      | La coordenada x theta superpuesta para el segundo átomo                             |
| K2y      | La coordenada y theta superpuesta para el segundo átomo                             |
| K2z      | La coordenada z theta superpuesta para el segundo átomo                             |
| K2sphR   | La distancia polar esférica theta superpuesta para el segundo átomo kappa.          |
| K2sphT   | El ángulo de latitud polar esférica theta superpuesto para el segundo átomo kappa.  |
| K2sphP   | El ángulo de longitud polar esférica theta superpuesto para el segundo átomo kappa. |
| K2cylR   | La distancia polar cilíndrica theta superpuesta para el segundo átomo kappa.        |
| K2cylT   | El ángulo polar cilíndrico theta superpuesto para el segundo átomo kappa.           |
| K2cylZ   | La coordenada z polar cilíndrica theta superpuesta para el segundo átomo kappa.     |

|        |                                                                                                                                                                                        |
|--------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| inter  | El tipo de interacción para el contacto theta & iota: MM - Cadena principal a Cadena principal, MS - Cadena principal a Cadena lateral, SM. ..., SS. .. en donde la primera letra hace |
| Icount | El recuento de átomos iota para el átomo theta particular.                                                                                                                             |
| Tcount | El recuento de átomos theta para el aminoácido particular.                                                                                                                             |

#### *Visualización de conjuntos de datos iota IL17F*

Cada conjunto de datos iota se visualizó con relación a la estructura IL17F / Fragmento Fab 496 usando un software de gráficas moleculares como Pymol. Esto se pudo hacer con trazado directo del conjunto de datos iota como puntos individuales o transformando primero matemáticamente el conjunto de datos a una función de densidad y a un formato compatible para visualizar la gráfica molecular por ejemplo ccp4, de tal forma que los mapas de contorno de mayor densidad pudieran ser visualizados sobre el epítipo IL 17.

#### *Inspección de mapas de densidad iota para intersección con átomos paratopo del Fragmento Fab 496*

La interfaz IL17F / Fragmento Fab 496 fue examinada para determinar el grado de intersección entre los átomos individuales de fragmento Fab 496 por residuo y los correspondientes mapas de densidad iota. Se identificaron residuos en donde había poca o no se encontraba intersección. En estos casos, los residuos alternativos fueron sustituidos por medio de un software de gráficas moleculares para determinar si la mejor intersección entre átomos de residuo y mapas de densidad iota relevantes podía ser alcanzada. Sustituciones de aminoácidos con una buena intersección de mapa de densidad iota fueron listadas para producción y pruebas in vitro como mutaciones puntuales simples como versiones intactas IgG del fragmento Fab 496.

#### *Manipulaciones de ADN y métodos generales*

*E. coli* INVaF (Invitrogen) fue usada para transformación y crecimiento de cultivo de rutina. Las enzimas de restricción y modificación de ADN se obtuvieron de Roche Diagnostics Ltd. y New England Biolabs. Las preparaciones de plásmidos se realizaron utilizando kits de purificación de Maxi Plasmid (Qiagen, catálogo No. 12165). Las reacciones de secuenciación de ADN se realizaron usando un kit de secuenciación del terminador ABI Prism Big Dye (catálogo No. 4304149) y se ejecutaron en un secuenciador automático ABI 3100 (Applied Biosystems). Los datos fueron analizados usando el programa Auto Assembler (Applied Biosystems). Los oligonucleótidos se obtuvieron de Invitrogen. La concentración de IgG fue determinada por el ensamblado IgG ELISA.

#### *Maduración de la afinidad del anticuerpo CA028\_00496*

CA028\_0496 es un anticuerpo humanizado neutralizador que une las isoformas IL17A y IL17F. Se compone de las regiones variables injertadas, denominadas gL7 y gH9, cuyas secuencias se presentaron en WO 2008/047134. El fragmento Fab salvaje de este anticuerpo (Fab 496) y las variantes mutantes fueron preparadas como se indica a continuación: Se diseñaron y construyeron secuencias cebadoras de oligonucleótidos con el fin de introducir mutaciones puntuales en la región variable de cadena ligera (gL7) como por residuos y posiciones determinadas en la lista corta presentada anteriormente. Cada cadena ligera mutada fue subclonada separadamente en un vector de expresión de cadena ligera humana UCB Celltech pKH10.1, la cual contenía ADN que codificaba la región constante humana C-kappa (Alotipo Km3). La secuencia de región variable de cadena pesada (gH9) fue subclonada en un vector de expresión UCB Celltech pVhg1 Fab6His, la cual contenía ADN que codificaba una región constante gama-1 de cadena humana pesada, CH1. Plásmidos de cadena ligera y pesada que codifican fueron co-transfectados en células HEK293 usando el procedimiento 293fectin™ de acuerdo con las instrucciones del fabricante (Invitrogen. Catálogo No. 12347-019). Los niveles del anticuerpo IgG secretados en el cultivo de sobrenadantes después de 10 a 12 días de cultivo fueron evaluados por ELISA y la

cinética de unión fue evaluada por resonancia de plasmones superficiales (ver abajo). Los mutantes que mostraban una unión mejorada o similar al IL17F fueron preparadas y probadas en combinación como mutaciones de cadena ligera dobles, triples, cuádruples y quíntuples como se mencionó anteriormente.

#### *Resonancia de Plasmones Superficiales (SPR, por sus siglas en inglés)*

Todos los experimentos SPR fueron realizados en un sistema Biacore 3000 (Biacore AB) a 25°C usando un buffer de corrida HBS-EP (10 mM HEPES pH 7,4, 150 mM NaCl 3mM EDTA 0.005% (v/v) surfactante P20, Biacore AB). El anticuerpo específico Goat F(ab')<sub>2</sub> anti-IgG Fab' (Jackson Labs. Código de producto 109-006-097) fue anexado covalentemente a la superficie de un chip sensor CM5 (GE Healthcare) por medio del método de acoplamiento de aminas, según la recomendación de los fabricantes. Brevemente, la superficie de carboximetil dextrano fue activada con una mezcla fresca de 50 mM N-hidroxisuccinimida y 200 mM l-etil-3- (3- dimetilaminopropil) carbodiimida por 5 minutos a una tasa de flujo de 10 µl / min. Un anticuerpo Anti-Fab a 50 µg / ml en 10 mM de regulador de acetato de sodio con pH 5.0 fue inyectado por 60 segundos con la misma tasa de flujo. Finalmente, la superficie fue desactivada con un pulso de 10 minutos de 1 M etanolamina.HCl pH8,5, dejando de 4000 a 5000 unidades de respuesta (RU, por sus siglas en inglés) de anticuerpos inmovilizados en el chip. Una célula de flujo de referencia se preparó en el mismo chip omitiendo la proteína del procedimiento anterior.

Los fragmentos Fab 496 de tipo salvaje y mutados fueron cosechados de sobrenadantes de cultivos en el rango 3 a 30 µg / ml y se diluyeron los sobrenadantes crudos en un buffer de corrida en el rango 0,5 a 2 µg / ml. Con el propósito de evaluar la cinética de unión a IL17F, cada anticuerpo fue capturado inicialmente en una superficie anti-Fab' por inyección a 10 µl / min durante 60 segundos para producir una señal RU adicional de 150 a 250. El IL17F humano recombinante fue valorado de 10 nM en buffer de corrida e inyectado a 30 µl / ml, para producir una fase de

asociación de más de 180 segundos, seguida de una fase de disociación de 300 segundos. Al final de cada ciclo, la superficie fue regenerada con un pulso de 60 segundos de 40 mM HCl seguido de un pulso de 30 segundos de 5 mM NaOH at 10  $\mu$ l / min. Para cada fragmento Fab, se realizó un ciclo de control en el que la inyección IL17F fue reemplazada por una inyección de buffer de corrida.

Los sensogramas fueron corregidos por sustracción de la señal del flujo de células de referencia, luego sustrayendo el sensograma de control de ciclo para el respectivo fragmento Fab'. Las constantes de velocidad de disociación ( $k_d$ ) y las constantes de velocidad de asociación ( $k_a$ ) fueron incluidas dentro de los datos usando el software Biaevaluation (Biacore AB). Las afinidades del fragmento Fab ( $K_D$ ) fueron calculadas como  $K_D = k_d / k_a$ .

## Resultados

### *Intersección de los mapas de densidad iota con átomos de fragmento Fab 496*

La inspección de esa parte de la superficie del fragmento Fab 496 que conforma la interfaz con IL17F reveló que en muchas áreas había una intersección completa entre un fragmento Fab 496 dado y el mapa de densidad iota correspondiente. Este fue el caso particular para los residuos aminoácidos que contienen la cadena pesada del fragmento Fab 496. Sin embargo, hubo un número de regiones que contienen la cadena ligera en donde había poca o ninguna intersección de átomos de fragmento Fab 496 y sus correspondientes mapas de densidad iota (Panel A, Figs. 8 a 13).

Al sustituir residuos alternativos en la estructura del fragmento Fab 496 en estas posiciones, fue posible lograr la intersección entre uno o más átomos de residuo sustituido y sus correspondientes mapas de densidad iota. (Panel B, Figs. 8 a 13). Por tanto, en la cadena ligera del fragmento Fab 496, el átomo de treonina 30 Cy2 no interseca con el mapa de densidad iota del carbono de metilo y el átomo Oy1 no se interseca con el mapa de densidad iota del oxígeno hidroxilo. Sin embargo, cuando la

treonina 30 es mutada a arginina, existe una intersección entre el átomo Ne y el mapa de densidad iota de la amida secundaria (Fig. 8).

La Figura 9 muestra la densidad iota tetraédrica de metileno y la densidad iota de la amida secundaria dispersas alrededor de la arginina de cadena ligera 54 sin una intersección con respectivos átomos de la cadena lateral de este residuo. Por el contrario, cuando se mutan en una serina en esta posición, existe una buena intersección de densidad iota de oxígeno hidroxilo y el átomo de oxígeno gama de la cadena lateral de serina. El átomo de oxígeno gama de serina de cadena ligera 56 no interseca con la densidad iota de oxígeno hidroxilo. Sin embargo, cuando este residuo es mutado a isoleucina existe intersección del átomo delta 1 metil de la cadena lateral de la isoleucina y la densidad iota metil (Fig. 10). En el caso de la serina 60, nuevamente el oxígeno gama está muy distante de la densidad iota de oxígeno hidroxilo para lograr una intersección, (Fig. 11), pero mutando este residuo a aspartato trae como resultado una intersección de la cadena lateral del átomo de oxígeno delta con una densidad iota de oxígeno carboxilo.

La treonina 72 de cadena ligera no es un residuo CDR sino parte del marco 3, su átomo de metilo de cadena lateral no interseca con la densidad iota del carbono de metilo ni su átomo de oxígeno de cadena lateral con la densidad iota de oxígeno hidroxilo. Sin embargo, la mutación de arginina 72 permite la intersección de la cadena lateral de ambos costados del átomo de carbono delta con la densidad iota del carbono de metileno y de los átomos de nitrógeno eta de cadena lateral con la densidad iota del nitrógeno de guanidinio.

#### *Efectos de las mutaciones iota diseñadas en el anticuerpo CA028\_00496 en cinética de unión*

Cinco mutaciones de cadena ligera de punto sencillo diseñadas en el fragmento Fab 496, T30R, R54S, S56I, S60D y T72R, mostraron pequeñas mejoras en su afinidad de unión que van desde 1,7 a 3,6 veces. La mejora observada con la mutación no CDR,

T72R, fue sorprendente ya que los residuos en esta posición no suelen contactar el antígeno. Para todas menos para una de estas mutaciones (S60D), la mejora fue impulsada por una reducción en la constante de velocidad de disociación (Tabla 3). La combinación de estas mutaciones en pares resultó en una mejora sinérgica en la unión, con 3,8 a 7,5 veces más reducción de la constante de velocidad de la disociación; con combinaciones triples o cuádruples que producen más reducciones de paso en la constante de velocidad de disociación (Tabla 3).

La combinación de todas las mutaciones de cadena ligera produjo la mejora más grande en afinidad de unión (Tabla 4) para dar un valor de afinidad de 11 pM a IL17F, algunos 180 veces mejor que el original fragmento Fab 496. Un hallazgo importante fue que no había un efecto perjudicial en la unión a la isoforma IL17FA, de hecho una mejora, con afinidad constante a 2 pM comparado a 14 pM para el fragmento Fab 496 original. Es interesante que las combinaciones de las mutaciones diseñadas producen una mejora sinérgica de la afinidad; una explicación es que están separadas razonablemente a lo largo de la superficie de paratopo de cadena ligera (Fig. 13) y por tanto evitan los efectos negativos de la interacción.

*Tabla 3: Efecto de mutaciones de cadena ligera de un residuo simple y de un residuo combinado del anticuerpo CA028\_000496 en la constante de velocidad de disociación IL17F*

| <b>Mutaciones de la cadena ligera<br/>CA028_00496</b> | <b><math>k^d(s^{-1})</math></b> | <b>Mejora en<br/>veces</b> |
|-------------------------------------------------------|---------------------------------|----------------------------|
| wt                                                    | 4.2E-03                         |                            |
| T30R                                                  | 3.2E-03                         | 1.3                        |
| R54S                                                  | 2.1 E-03                        | 2.0                        |
| S56I                                                  | 3.0E-03                         | 1.4                        |
| S60D                                                  | 4.1 E-03                        | 1.0                        |
| T72R                                                  | 2.5E-03                         | 1.7                        |
| T30R / R54S                                           | 7.0E-04                         | 5.8                        |
| T30R / S56I                                           | 1.1 E-03                        | 3.8                        |
| R54S / S56I                                           | 6.2E-04                         | 6.6                        |

|                                  |         |     |
|----------------------------------|---------|-----|
| R54S / T72R                      | 5.5E-04 | 7.5 |
| S56I / T72R                      | 8.0E-04 | 5.1 |
| S60D / T72R                      | 1.2E-03 | 3.6 |
| T30R / R54S / T72R               | 2.5E-04 | 19  |
| T30R / S56I / T72R               | 4.1E-04 | 10  |
| T30R / S56I / S60D / T72R        | 3.1E-04 | 14  |
| T30R / R54S / S56I / T72R        | 1.3E-04 | 35  |
| T30R / R54S / S60D / T72R        | 9.3E-05 | 44  |
| T30R / R54S / S56I / S60D / T72R | 5.4E-05 | 104 |

*Tabla 4: Constante de afinidad del anticuerpo CA028\_000496 versus la variante que comprende 5 mutaciones de cadena ligera*

|                                   | <b>K<sub>a</sub> (M<sup>-1</sup> s<sup>-1</sup>)</b> | <b>K<sub>d</sub> (s<sup>-1</sup>)</b> | <b>K<sub>D</sub> (M)</b> | <b>K<sub>D</sub> (pM)</b> |
|-----------------------------------|------------------------------------------------------|---------------------------------------|--------------------------|---------------------------|
| Tipo salvaje                      | 2,0E+06                                              | 4,1E-03                               | 2,0E-09                  | 2000                      |
| T30R / RSRS / S56I<br>S60D / T72R | 2,3E+06                                              | 2,6E-05                               | 1,1E-11                  | 11                        |

### Conclusión

El método de crear mapas de densidad iota sobre la superficie del epítipo de IL17F con el fin de predecir mutaciones favorables en el anticuerpo correspondiente, fragmento fab 496, ha demostrado ser exitoso en cuanto que la constante de afinidad ha sido mejorada 180 veces a un K<sub>D</sub> de 11 pM. Este método no asume que sólo mutaciones CDR puedan ser usadas sino que demuestra que las mutaciones de región marco también son importantes para la maduración de la afinidad.

### EJEMPLO 2: Utilización de un proceso automatizado de invención en la región entre las cadenas ligeras y pesadas Ab para mejorar la estabilidad

#### Métodos

*Identificación de átomos (theta) objetivos de cadena ligera y pesada de la interfaz Fv del fragmento Fab X*

Usando las coordenadas de la estructura de cristal de un fragmento Fab de anticuerpo "Fab X" complicado con el antígeno, todos los átomos de cadena pesada dentro de 6Å de cualquier átomo de cadena ligera fueron identificados como átomos epítomos (167 átomos theta listados en la Tabla 5). Similarmente, todos los átomos de cadena ligera dentro de 6Å de cualquier átomo de cadena pesada fueron identificados como átomos epítomos (159 átomos theta listados en la Tabla 5).

*Tabla 5: Lista de átomos que comprenden epítomos de cadena ligera y pesada, en los que la notación (1)-(2)-(3) designa (1) Las respectivas cadenas H (pesada) y L (ligera) del fragmento Fab X, (2) número residuo y (3) tipo de átomo.*

**Bolsillo de unión (THETA) = H Pareja de unión (IOTA) = L**

|           |           |            |            |
|-----------|-----------|------------|------------|
| H-47-CA   | H-57-0    | H-104-0H   | H-117-N    |
| H-47-CB   | H-60-CB   | H-109-CB   | H-117-0    |
| H-47-CG 1 | H-60-CD 1 | H-109-CG 1 | H-117-0H   |
| H-47-CG2  | H-60-CG 1 | H-109-CG2  | H-118-C    |
| H-49-CB   | H-60-CG2  | H-112-CD 1 | H-118-CA   |
| H-49-CD   | H-62-CE2  | H-112-CE 1 | H-118-CB   |
| H-49-CG   | H-62-CH2  | H-112-CE2  | H-118-CD 1 |
| H-49-NE2  | H-62-CZ2  | H-112-CZ   | H-118-CD2  |
| H-49-0E 1 | H-62-NE 1 | H-112-0H   | H-118-CE 1 |
| H-53-C    | H-68-C    | H-113-C    | H-118-CE2  |
| H-53-CA   | H-68-CA   | H-113-CA   | H-118-CG   |
| H-53-0    | H-68-CB   | H-113-CB   | H-118-CZ   |
| H-54-C    | H-68-CD 1 | H-113-0    | H-118-N    |
| H-54-CA   | H-68-CD2  | H-113-0G   | H-118-0    |
| H-54-N    | H-68-CE 1 | H-114-C    | H-119-C    |
| H-54-0    | H-68-CE2  | H-114-CA   | H-119-CA   |
| H-55-C    | H-68-CG   | H-114-CB   | H-119-CB   |
| H-55-CA   | H-68-CZ   | H-114-CG2  | H-119-CG   |
| H-55-CB   | H-68-N    | H-114-N    | H-119-N    |
| H-55-CD 1 | H-68-0    | H-114-0    | H-119-0    |
| H-55-CD2  | H-69-C    | H-114-0G 1 | H-121-C    |
| H-55-CG   | H-69-CA   | H-115-C    | H-121-CA   |
| H-55-N    | H-69-N    | H-115-CA   | H-121-CB   |
| H-55-0    | H-69-0    | H-115-CB   | H-121-CD 1 |
| H-56-C    | H-70-C    | H-115-N    | H-121-CD2  |
| H-56-CA   | H-70-CA   | H-115-0    | H-121-CE2  |

|          |           |           |           |
|----------|-----------|-----------|-----------|
| H-56-CB  | H-70-CB   | H-116-C   | H-121-CE3 |
| H-56-N   | H-70-N    | H-116-CA  | H-121-CG  |
| H-56-0   | H-70-0    | H-116-CB  | H-121-CH2 |
| H-57-C   | H-71-C    | H-116-CD  | H-121-CZ2 |
| H-57-CA  | H-71-CA   | H-116-CG  | H-121-CZ3 |
| H-57-CB  | H-71-CB   | H-116-N   | H-121-N   |
| H-57-CD1 | H-71-CG2  | H-116-0   | H-121-NE1 |
| H-57-CD2 | H-71-N    | H-117-C   | H-121-0   |
| H-57-CE2 | H-71-0G1  | H-117-CA  | H-122-C   |
| H-57-CE3 | H-72-N    | H-117-CB  | H-122-CA  |
| H-57-CG  | H-104-CD1 | H-117-CD1 | H-122-N   |
| H-57-CH2 | H-104-CD2 | H-117-CD2 | H-122-0   |
| H-57-CZ2 | H-104-CE1 | H-117-CE1 | H-123-C   |
| H-57-CZ3 | H-104-CE2 | H-117-CE2 | H-123-CA  |
| H-57-N   | H-104-CG  | H-117-CG  | H-123-0   |
| H-57-NE1 | H-104-CZ  | H-117-CZ  |           |

**Bolsillo de unión (THETA) = L Pareja de unión (IOTA) = H**

|          |         |          |           |
|----------|---------|----------|-----------|
| L-11-0D1 | L-52-C  | L-60-CZ  | L-106-CG2 |
| L-41-0   | L-52-0  | L-60-N   | L-106-N   |
| L-42-C   | L-53-C  | L-60-0   | L-106-0   |
| L-42-CA  | L-53-CA | L-61-C   | L-107-C   |
| L-42-CB  | L-53-N  | L-61-CA  | L-107-CA  |
| L-42-0   | L-53-0  | L-61-CB  | L-107-CB  |
| L-43-C   | L-54-C  | L-61-CD  | L-107-N   |
| L-43-CA  | L-54-CA | L-61-CG  | L-107-0   |
| L-43-CB  | L-54-CB | L-61-N   | L-107-0G  |
| L-43-CD1 | L-54-N  | L-61-0   | L-108-C   |
| L-43-CD2 | L-54-0  | L-61-0E1 | L-108-CA  |
| L-43-CE1 | L-55-C  | L-61-0E2 | L-108-CB  |
| L-43-CE2 | L-55-CA | L-62-N   | L-108-CG  |
| L-43-CG  | L-55-CB | L-64-NZ  | L-108-N   |
| L-43-CZ  | L-55-CD | L-66-CG2 | L-108-0D1 |
| L-43-N   | L-55-CG | L-98-CB  | L-108-0D2 |
| L-43-0   | L-55-N  | L-98-CD1 | L-109-C   |
| L-44-C   | L-55-0  | L-98-CD2 | L-109-CA  |
| L-44-CA  | L-56-C  | L-98-CE1 | L-109-CB  |
| L-44-N   | L-56-CA | L-98-CE2 | L-109-CG2 |
| L-44-0   | L-56-CB | L-98-CG  | L-109-N   |
| L-45-C   | L-56-N  | L-98-CZ  | L-109-0   |

|          |          |           |           |
|----------|----------|-----------|-----------|
| L-45-CA  | L-56-0   | L-98-0H   | L-109-0G1 |
| L-45-CB  | L-57-C   | L-100-C   | L-110-CA  |
| L-45-N   | L-57-CA  | L-100-CA  | L-110-N   |
| L-45-0   | L-57-CB  | L-100-N   | L-111-CA  |
| L-45-0G  | L-57-CD1 | L-100-0   | L-111-CB  |
| L-47-CB  | L-57-CD2 | L-101-C   | L-111-CD1 |
| L-47-CD1 | L-57-CG  | L-101-CA  | L-111-CD2 |
| L-47-CD2 | L-57-N   | L-101-N   | L-111-CE1 |
| L-47-CE1 | L-59-C   | L-101-0   | L-111-CE2 |
| L-47-CE2 | L-59-0   | L-102-C   | L-111-CG  |
| L-47-CG  | L-60-C   | L-102-CA  | L-111-CZ  |
| L-47-CZ  | L-60-CA  | L-102-N   | L-111-N   |
| L-47-0H  | L-60-CB  | L-105-0   | L-111-0   |
| L-49-CB  | L-60-CD1 | L-106-C   | L-112-0   |
| L-49-CD  | L-60-CD2 | L-106-CA  | L-113-C   |
| L-49-CG  | L-60-CE1 | L-106-CB  | L-113-CA  |
| L-49-NE2 | L-60-CE2 | L-106-CD1 | L-113-0   |
| L-49-0E1 | L-60-CG  | L-106-CG1 |           |

### *Superposición de datos iota sobre la superficie de interfaz de cadenas ligeras y pesadas*

Para cada uno de los 167 átomos theta que contienen el epítipo de cadena pesada, el conjunto de contactos theta correspondiente fue seleccionado de la base de datos IOTA y, de ahí, un subgrupo iota apropiado fue seleccionado, por ejemplo el oxígeno carbonilo. Las coordenadas relativas iota de este subgrupo fueron transpuestas respecto al marco de referencia del átomo theta dado del epítipo de cadena pesada. Un conjunto de datos iota de un subgrupo dado fue entonces acumulado a lo largo de todo el epítipo de cadena pesada. En los casos en los que la ubicación de un punto de datos iota dado intersectara con un átomo de la cadena pesada, más cerca que la suma de sus respectivos radios Van de Waals menos  $0,2\text{\AA}$ , entonces estos puntos de datos eran excluidos del conjunto de datos. El proceso se repitió para todos los subgrupos iota relevantes para producir una serie de conjunto de datos iota para el epítipo de cadena pesada.

Para cada uno de los 159 átomos theta que contienen un epítipo de cadena ligera, el conjunto de contactos theta correspondiente fue seleccionado de una base de datos IOTA y, de ahí, un subgrupo iota apropiado fue seleccionado, por ejemplo oxígeno carbonilo. Las coordenadas relativas iota de este subgrupo fueron transpuestas respecto al marco de referencia del átomo theta dado del epítipo de cadena pesada. Un conjunto de datos iota de un subgrupo dado fue entonces acumulado a lo largo de todo el epítipo de cadena ligera. En los casos en los que la ubicación de un punto de datos iota dado intersectaba con un átomo de cadena ligera, más cerca que la suma de sus respectivos radios Van de Waals menos  $0,2\text{\AA}$ , entonces esos puntos de datos eran excluidos del conjunto de datos. El proceso se repitió para todos los subgrupos iota relevantes para producir una serie de conjuntos de datos iota para el epítipo de cadena ligera.

#### *Inspección de los mapas de densidad iota para intersección con átomos de cadenas pesadas-ligeras*

Todo el proceso fue realizado automáticamente con un *Rosetta python library script* personalizado internamente hecho a la medida para la identificación de posiciones mutables, generación de puntos singulares mutantes, enumeración del estado rotámero de baja energía, computación de puntaje cuantitativa IOTA, estimación de la energía de unión de cadenas VH (muy pesadas) –VL (muy ligeras), y priorización de puntos mutantes.

Dos métodos de puntaje fueron usados para categorizar los mutantes:

##### 1. Puntaje $\Delta$ IOTA

Los mapas de densidad IOTA generados, mencionados anteriormente fueron usados para computar los valores de intersección especial entre cada átomo de residuo pesado en cada posición mutable y los puntos críticos de densidad en los tipos de mapas cercanos correspondientes. El puntaje IOTA es la suma de las superposiciones

volumétricas entre átomos pesados de un residuo con el máximo de densidades IOTA con los correspondientes tipos de definición, lo cual refleja el grado de intersección entre átomos de fragmento Fab X individuales por residuo y el mapa de densidad iota correspondiente. El puntaje IOTA es numéricamente negativa, en donde los valores más bajos implican una mayor intersección. El puntaje IOTA es el cambio de puntaje IOTA entre el residuo mutante y el de tipo salvaje; similarmente, entre más negativo es el valor de puntaje  $\Delta$ IOTA más grande es la implicación de que el mutante es más favorecido que el de tipo salvaje.

## 2. Puntaje Rosetta $\Delta\Delta G$

La función de energía Rosetta es una combinación lineal de períodos que modela fuerzas de interacción entre átomos, efectos de solvatación y energías de torsión. Más específicamente, Puntaje 12, la función por defecto de un átomo lleno de energía en Rosetta está compuesto por un período Lennard- Jones, un período de solvatación implícita, un período de unión de hidrógeno dependiente de la orientación, potenciales de torsión de la cadena lateral y de la espina dorsal derivados del PDB, un período electrostático de corto rango basado en el conocimiento, y energías de referencia para cada uno de los aminoácidos que modelan el estado desplegado. La fuerza de unión entre una pareja de unión, o  $\Delta G$ , puede computarse al sustraer el puntaje Rosetta de las partes individuales con la de la estructura compleja formada por ambas partes de la unión. Un bajo  $G$  implica una unión más fuerte.  $\Delta\Delta G$  es el cambio de  $\Delta G$  entre el complejo mutante y el de tipo salvaje; entre más negativo el valor de  $\Delta G$  más grande la implicación de que la afinidad del mutante es más alta que la de tipo salvaje.

La Figura 14 muestra el flujo de trabajo para predecir *in silico* el punto de mutación en la interfaz VH-VL de la estructura de fragmento Fab X.

En el paso S101, todos los residuos de la cadena pesada con al menos un átomo pesado dentro de 8Å de cualquier cadena ligera de átomos pesados fueron

identificados como posiciones mutables. Similarmente, todos los residuos en la cadena ligera con al menos un átomo pesado dentro de 8Å de cualquier átomo pesado de cadena pesada fueron identificados como posiciones mutables.

En el paso S102, para la estructura de cristal del fragmento Fab X de tipo salvaje, el puntaje IOTA de tipo residuo y la energía de unión  $\Delta G$  son computadas, respectivamente. En el paso S 102.1, el puntaje IOTA para el residuo de tipo salvaje en la posición mutable actual con los mapas de densidad IOTA cercanos correspondientes es computada, denominada como ( $IOTAScore_{wt}$ ,  $Position_j$ ); en el paso S 102.2, la energía de unión de las cadenas VH y VL del fragmento Fab X es computada con la función 12 de el puntaje Rosetta, denominada  $\Delta G_{wt}$ .

En el paso S 103, el residuo de tipo salvaje en la posición mutable actual identificado en el paso S 101 es reemplazado (mutado) por los otros tipos de aminoácidos. De los 20 tipos de aminoácidos naturales, la prolina y cisteína son excluidas de la mutación. Los otros 18 tipos (alanina, arginina, asparagina, ácido aspártico, ácido glutámico, glutamina, glicina, histidina, isoleucina, leucina, lisina, metionina, fenilalanina, serina, treonina, triptófano, tirosina, valina) excepto el de tipo salvaje como tal, son mutados en cada posición mutable uno por uno.

En el paso S 104, para cada tipo de residuo mutado a cada posición mutable, el top 100 de estados rotámeros de menor energía (en términos de la función de puntaje de Rosetta) son generados usando Rosetta. Los otros estados rotámeros de alta energía son descartados.

En el paso S105, para cada estado rotámero del residuo mutante generado en el S104, el puntaje IOTA es computada de la misma manera que en S102, denominado ( $IOTAScore_{mutant}$ ,  $Position_j$  Type<sub>k</sub>, Rotamer<sub>i</sub>).

En el paso S 106, el puntaje  $\Delta IOTA$  para la combinación actual del estado rotámero, el tipo de residuo mutante, y la posición mutable es computada sustrayendo

( $IOTAScore_{wt}$ ,  $Position_j$ ) y ( $IOTAScore_{mutant}$ ,  $Position_j$   $Type_k$ ,  $Rotamer_i$ ), lo cual es denominado ( $\Delta IOTAScore$ ,  $Position_j$   $Type_k$ ,  $Rotamer_i$ ). Los pasos S 105 y S 106 fueron repetidos para computar todos los puntajes  $\Delta IOTA$  para cada estado rotámero para el tipo actual de residuo mutante y su posición mutable.

En el paso S 107, el estado rotámero óptimo del actual tipo de residuo mutante y la posición mutable es determinado por el valor de puntaje  $\Delta IOTA$  más bajo, como se muestra en el paso S 107.1. La energía de unión del mutante con el estado rotámero óptimo es computada en el paso S 107.2 de la misma manera que en el paso S 102.2, denominado ( $\Delta G_{mutant}$ ,  $Position_j$   $Type_k$ ). En el paso S 107.3, el cambio de energías de unión  $\Delta \Delta G$  entre el mutante y el de tipo salvaje es calculado por la sustracción de  $\Delta G_{wt}$  y  $\Delta G_{mutant}$ . Después de que el estado rotámero óptimo es priorizado, los pasos S 103 a S 107 son repetidos para el siguiente tipo de aminoácido mutante en la posición mutable actual.

En el paso S 108, para la posición mutable actual, únicamente los mutantes candidatos que satisfacen los criterios de ambas puntuaciones  $\Delta IOTAScore < 0$  y  $\Delta \Delta G < 0$  son mantenidos para una posterior categorización. Los demás son descartados. Los pasos S 102 a S 108 son repetidos para ir por todas las posiciones mutables y generar todos los candidatos mutantes que satisfacen los mismos criterios.

En el paso S 109, todas las estructuras mutantes candidatas fueron emitidas para un análisis de visualización posterior. La lista final de los mutantes candidatos fue ordenada y clasificada por las menores puntuaciones  $\Delta IOTA$ .

El comando ejecutado y los parámetros usados fueron como se muestran a continuación:

Para predicción de mutaciones de cadena ligera, el comando era:

```
"python multiRotamersFabinterfaceIOTAScan.py --pdb FabX.pdb --only_chains L -  
region all --useIOTA --IOTAtype 167 --output_mutant"
```

Para predicción de mutaciones de cadena pesada, el comando era:

```
"python multiRotamersFabInterfaceIOTAScan.py --pdb FabX.pdb --only_chains H --region all --useIOTA --IOTAtype 167 --output_mutant"
```

Los parámetros relevantes extra Rosetta fueron inicializados añadiendo el siguiente código al

```
"multiRotamersFabInterfaceIOTAScan.py":
```

```
"init(extra_options = "-ex l -ex2 -score:weights score l 2 -no_his_his_pairE -constant_seed -edensity:mapreso 3.0 -correct -mute all"
```

### *Manipulaciones de ADN y métodos generales*

La cepa *E. coli* INVaF (Invitrogen) fue usada para transformación y crecimiento de cultivo de rutina. Las enzimas de restricción y modificación de ADN se obtuvieron de Roche Diagnostics Ltd. y New England Biolabs. Las preparaciones de plásmidos se realizaron utilizando kits de purificación de Maxi Plasmid (Qiagen, catálogo No. 12165). Las reacciones de secuenciación de ADN se realizaron usando un kit de secuenciación del terminador ABI Prism Big Dye (catálogo No. 4304149) y se ejecutaron en un secuenciador automático ABI 3100 (Applied Biosystems). Los datos fueron analizados usando el programa Auto Assembler (Applied Biosystems). Los oligonucleótidos se obtuvieron de Invitrogen. La concentración de IgG fue determinada por el ensamblado IgG ELISA.

### *Mejora de la termoestabilidad del fragmento Fab X a través de la maduración de la afinidad de la interfaz de cadena pesada-ligera*

El fragmento Fab de tipo salvaje del Fab X y las variantes mutantes fueron preparados como se explica a continuación: las secuencias cebadoras de oligonucleótidos fueron diseñadas y construidas con el fin de introducir mutaciones

de punto sencillo tanto en las regiones variables de cadena ligera como de cadena pesada como por residuos y posiciones determinados en la lista corta presentada anteriormente. Cada cadena ligera mutada fue subclonada de manera separada en un vector de expresión de cadena ligera humana UCB Celltech pKH10.1, el cual contenía el ADN que codifica la región constante humana C-kappa (alotipo Km3). Cada secuencia de región variable de cadena pesada mutada fue subclonada separadamente en un vector de expresión UCB Celltech pVhg l Fab6His, el cual contenía el ADN que codifica la región constante gamma-1 humana de cadena pesada, CH1. Los plásmidos codificadores de cadena ligera y pesada fueron co-transfectadas en células HEK293 usando el procedimiento 293fectin™ de acuerdo con las instrucciones del fabricante (Invitrogen. Catálogo No. 12347-019). Los niveles del anticuerpo IgG1 Fab secretados en el cultivo de los supernadantes después de 10 a 12 días de cultivo fueron evaluados por ELISA y la cinética de unión fue evaluada por resonancia de plasmones superficiales (ver abajo).

Luego, los mutantes que mostraron mejora en la termoestabilidad fueron preparados y probados en combinación como mutaciones dobles o triples como se mostró anteriormente.

#### *Resonancia de plasmones superficiales (SPR)*

Todos los experimentos SPR fueron realizados en un BIAcore T200 (GE Healthcare). El fragmento Affinipure F(ab')<sub>2</sub> goat anti-humano IgG, F(ab')<sub>2</sub> fragmento específico (Jackson ImmunoResearch) fue inmovilizado en un Chip Sensor CM5 por medio de una química de acoplamiento de amina a un nivel de captura de ~5000 unidades de respuesta (RUs). Se usó un buffer HBS-EP (10mM HEPES pH 7.4, 0.15 M NaCl, 3 mM EDTA, 0.05 % Surfactant P20, GE Healthcare) como el buffer de corrida con una velocidad de flujo de 10 µL/ min. Una inyección de 10 µL de fragmento Fab X a 0.75µg/mL fue usada para captura por el anti-humano inmovilizado IgG-F(ab')<sub>2</sub>. El antígeno fue valorado sobre el fragmento Fab

X capturado en varias concentraciones (50nM to 6.25nM) a una velocidad de flujo de 30  $\mu\text{L}/\text{min}$ . La superficie fue regenerada por una inyección de 2 x 10  $\mu\text{L}$  de 50 mM HCl, seguida de una inyección de 5  $\mu\text{L}$  de 5 mM NaOH a una velocidad de flujo de 10  $\mu\text{L}/\text{min}$ . Las curvas de unión para sustracción de fondo fueron analizadas usando el software de evaluación T200 (versión 1.0) siguiendo procedimientos estándar. Los parámetros de cinética fueron determinados del algoritmo de ajuste.

#### *Ensayo de termoestabilidad*

Un ensayo de termofluor fue realizado para evaluar las estabilidades térmicas de moléculas purificadas. Las proteínas purificadas (0,1 mg/ml) fueron mezcladas con SYPRO® Orange dye (Invitrogen), y la mezcla dispensada se cuadruplica en una placa óptica 384 PCR. Las muestras fueron analizadas en un sistema de 7900HT Fast Real-Time PCR (Agilent Technologies) en un rango de temperatura de 20°C a 99°C, con una tasa de rampa de 1,1°C/min. Los cambios de intensidad fluorescente por pozo se trazaron contra la temperatura y los puntos de inflexión de las laderas resultantes fueron usados para generar el  $T_m$ .

#### Resultados

##### *Intersección de los mapas de densidad iota con átomos de cadena ligera y pesada de la interfaz*

El método automatizado usando un scan Rosetta produce una tabla de mutaciones categorizada por puntuaciones  $\Delta\text{IOTA}$  (Tabla 6).

##### *Efectos de las mutaciones diseñadas de iota en el fragmento Fab X en termoestabilidad*

Seis mutaciones de punto sencillo diseñadas en fragmento Fab X, H-T71R, H-T71K, H-T71N, H-T71H, L-S107E y L-T109I, mostraron pequeñas mejoras en termoestabilidad desde 0,5°C a 2,9°C sobre el de tipo salvaje (Tabla 7). Las Figuras 15, 16 y 17 muestran visualizaciones generadas por computadora mostrando los efectos de estas

mutaciones de punto sencillo. En particular, estas Figuras muestran que mientras H-T71, L-S107 y L-T109 no tienen intersección de densidad, H-R71 interseca con la densidad de amida, L-E 107 interseca con la densidad carboxilato y L-I109 interseca con la densidad metilo. Las combinaciones de estas mutaciones en pares resultaron en una mejora sinérgica en la termoestabilidad; con combinaciones triples que producen más mejoras de paso en la termoestabilidad (Tabla 8).

La combinación de H-T7 1 R, L-S 107E y L-T 109I, las mutaciones produjeron la mayor mejora en termoestabilidad (Tabla 8) para dar un Tm de 8 1.2°C, unos 5.8°C mejor que el fragmento Fab Xoriginal. Esta combinación de las tres mutaciones se muestra en la Figura 18. Un hallazgo importante fue que no hubo una pérdida significativa en la unión a su antígeno.

*Tabla 6: Mutaciones propuestas generadas por el método de escaneo Rosetta y la categorización de acuerdo con el puntaje IOTA*

| Clasificación<br>(por Puntaje<br>IOTA) | Cadena<br>pesada | $\Delta\Delta G$ | Puntaje<br>$\Delta IOTA$ | Cadena<br>ligera | $\Delta\Delta G$ | Puntaje<br>$\Delta IOTA$ |
|----------------------------------------|------------------|------------------|--------------------------|------------------|------------------|--------------------------|
| 1                                      | T71R             | -0,37            | -56,27                   | T109H            | -1,49            | -92,37                   |
| 2                                      | V109R            | -0,04            | -50,58                   | T109K            | -0,01            | -67,33                   |
| 3                                      | V109K            | -0,51            | -40,67                   | T109I            | -1,69            | -45,17                   |
| 4                                      | T71H             | -1,06            | -38,94                   | I106L            | -2,11            | -42,11                   |
| 5                                      | T71K             | -0,08            | -36,09                   | T109L            | -1,58            | -37,44                   |
| 6                                      | T71W             | -0,67            | -25,95                   | I106H            | -3,5             | -31,77                   |
| 7                                      | T71Y             | -0,67            | -25,64                   | S107E            | -0,62            | -18,71                   |
| 8                                      | T71N             | -0,03            | -24,89                   | A53Y             | -0,93            | -17,68                   |
| 9                                      | D119W            | -0,89            | -23,58                   | A53F             | -1,13            | -15,64                   |
| 10                                     | T71Q             | -0,38            | -23,28                   | I106N            | -0,47            | -13,57                   |
| 11                                     | V109I            | -0,56            | -22,64                   |                  |                  |                          |
| 12                                     | V109H            | -0,04            | -21,73                   |                  |                  |                          |

*Tabla 7: Termoestabilidad de mutaciones comparadas con la termoestabilidad de tipo salvaje de 75. 7 °C*

| Clasificación | Cadena | $\Delta\Delta G$ | Puntaje | Tm °C | Cadena | $\Delta\Delta G$ | Puntaje | Tm °C |
|---------------|--------|------------------|---------|-------|--------|------------------|---------|-------|
|---------------|--------|------------------|---------|-------|--------|------------------|---------|-------|

|    | pesada |       | $\Delta IOTA$ |      | ligera |       | $\Delta IOTA$ |      |
|----|--------|-------|---------------|------|--------|-------|---------------|------|
| 1  | T71R   | -0,37 | -56,27        | 78,6 | T109H  | -1,49 | -92,37        | 75,2 |
| 2  | V109R  | -0,04 | -50,58        | 75,6 | T109K  | -0,01 | -67,33        | ND   |
| 3  | V109K  | -0,51 | -40,67        | 75,6 | T109I  | -1,69 | -45,17        | 79,2 |
| 4  | T71H   | -1,06 | -38,94        | 76,6 | I106L  | -2,11 | -42,11        | 74,5 |
| 5  | T71K   | -0,08 | -36,09        | 78,6 | T109L  | -1,58 | -37,44        | 77,6 |
| 6  | T71W   | -0,67 | -25,95        | 75,6 | I106H  | -3,5  | -31,77        | 72,2 |
| 7  | T71Y   | -0,67 | -25,64        | 75,3 | S107E  | -0,62 | -18,71        | 77,3 |
| 8  | T71N   | -0,03 | -24,89        | 76,2 | A53Y   | -0,93 | -17,68        | 68,8 |
| 9  | D119W  | -0,89 | -23,58        | 73,7 | A53F   | -1,13 | -15,64        | 67,2 |
| 10 | T71Q   | -0,38 | -23,28        | 75,8 | I106N  | -0,47 | -13,57        | 71,8 |
| 11 | V109I  | -0,56 | -22,64        | ND   |        |       |               |      |
| 12 | V109H  | -0,04 | -21,73        | 75,1 |        |       |               |      |

ND=No determinado

*Tabla 8: Termoestabilidad y afinidad de las combinaciones de mutaciones.*

| Combinación                | Rosetta ddG (VH/VL) | Rosetta dE (Fv) | Tm (°C) | Tm SD | KD (nM) |
|----------------------------|---------------------|-----------------|---------|-------|---------|
| H-T71R + L-S107E + L-T109I | -3,72               | -1,73           | 81,2    | 0,0   | ND      |
| H-T71R + L-T109I           | -2,3                | -1,17           | 80,1    | 0     | 0,88    |
| H-T71R + L-S107E           | -2,13               | -0,93           | 80,1    | 0     | 3,90    |
| H-T71K + L-S107E           | -0,3                | 0,76            | 79,4    | 0,6   | 3,50    |
| H-T71K + L-S107E + L-T109L | -4,11               | -1,41           | 79,3    | 0,3   | ND      |
| H-T71K + L-T109I           | -1,96               | -0,98           | 79,1    | 0,3   | 1,39    |
| H-T71R + L-T109L           | -2,77               | -0,95           | 78,0    | 0,6   | 0,96    |
| H-T71K + L-T109L           | -2,2                | -0,7            | 77,7    | 0,1   | 1,59    |
| H-T71R                     | -0,37               |                 | 78,0    | 0,3   | 0,80    |
| H-T71K                     | -0,08               |                 | 77,9    | 0,3   | 1,27    |
| L-S107E                    | -0,62               |                 | 77,3    | 0,5   | 5,52    |
| L-T109I                    | -1,69               |                 | 77,9    | 0,3   | 1,47    |
| L-T109L                    | -1,58               |                 | 76,9    | 0,5   | 1,73    |
| WT                         |                     |                 | 75,4    | 0,2   | 1,35    |

## REIVINDICACIONES

1. Un método para diseñar un ligando *ab initio* que se enlazará a un sitio de unión de un destino macromolecular, o para identificar una modificación a un ligando con el fin de mejorar su afinidad con un sitio de unión de un destino macromolecular, el cual comprende:

a) identificar una lista de destino de los átomos que forman la superficie del sitio de unión objetivo;

b) identificar cada átomo, en adelante conocido como un átomo theta, en la lista de destino como un tipo particular de átomo theta;

c) tomar, de una base de datos estructural de macromoléculas biológicas, información sobre contactos átomo a átomo intramoleculares o intermoleculares no enlazantes en donde el primer átomo en un par de átomos en contacto es un tipo de átomo theta particular y el segundo átomo opuesto del par, en adelante conocido como átomo iota, es un tipo de átomo iota particular, en donde dicha información comprende datos espaciales y/o contextuales sobre el átomo iota respecto al átomo theta, y dichos datos recolectados para una pluralidad de contactos del tipo de átomo theta específico de dicha base de datos se conoce en adelante como un grupo de contacto theta;

d) para cada átomo theta identificado en la lista de destino del paso b), superponer, en o alrededor del sitio de unión de destino, datos relacionados con tipo de átomo iota particular, o un grupo predeterminado de tipos de átomo iota relacionados, a partir del grupo de contacto theta correspondiente tomado en el paso c);

e) combinar y/o analizar los datos superpuestos de tal manera que se pueda predecir una o más regiones favorecidas del sitio de unión en donde el tipo de átomo iota particular o el grupo predeterminado de tipos de átomo iota relacionados tiene una alta propensión teórica; y

f) con un posible ligando teóricamente acoplado al sitio de unión, comparando el tipo y la posición de uno o más de los átomos del posible ligando con las regiones favorecidas pronosticadas para los respectivos tipos de átomo iota, para identificar una modificación al posible ligando, en términos de átomos alternativos y/o adicionales del posible ligando, que producirán una mayor intersección entre los átomos alternativos y/o adicionales del posible ligando y las respectivas regiones favorecidas del tipos de átomo iota, llevando a una mejora en la afinidad del posible ligando modificado con el sitio de unión, comparado con el posible ligando no modificado;

en donde cada contacto intramolecular o intermolecular no enlazante en la base de datos se define como un contacto entre los residuos opuestos de un pliegue de proteínas o entre unidades de monómero opuestas de un pliegue macromolecular o entre dos parejas macromoleculares que interactúan y es específicamente entre un átomo theta en un lado del pliegue o primera pareja de interacción y un átomo iota en el lado opuesto o segunda pareja de interacción; en un caso en el que se cumple la siguiente condición:

$s - R_w = t$ , en donde  $s$  es la separación entre los dos átomos del contacto,  $R_w$  es la suma de los radios Van de Waals de los dos átomos del contacto, y  $t$  es un umbral de distancia predeterminado; y

en donde el tipo de átomo theta se identifica de forma única en el paso b), de tal manera que no hay intersección entre los datos de un conjunto de contactos theta extraídos en el paso c) para un determinado tipo de átomo theta y los datos de cualquier otro conjunto de contactos theta extraídos en el paso c) para cualquier otro tipo de átomo theta, aparte de los datos sobre el átomo theta particular como el átomo iota.

2. Un método conforme a la reivindicación 1, en donde por cada contacto intramolecular no enlazante extraído de la base de datos estructural de miembros de proteína en el paso c) también se cumple la siguiente condición:

el átomo theta y el átomo iota del contacto están en residuos diferentes, separados por al menos cuatro residuos a lo largo del polipéptido lineal o están en cadenas polipeptídicas separadas.

3. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde el tipo de átomo theta se identifica como uno y solo uno de:

los 167 átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de las proteínas;

los 82 átomos distintos al hidrógeno presentes en los 4 nucleótidos del polímero de ácido desoxirribonucleico (ADN);

los 42 átomos distintos al hidrógeno presentes en los nucleótidos de ADN metilado, citidina fosfato y fosfato de adenosina;

los 85 átomos distintos al hidrógeno presentes en los 4 fosfatos de nucleótidos del polímero de ácido ribonucleico (ARN);

los 89 átomos distintos al hidrógeno presentes en los fosfatos nucleótidos de ribosa 2-O' metilada del ARN;

los más de 400 átomos distintos al hidrógeno presentes en el ARN de base modificada más común después de la transcripción.

4. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde la información extraída en el paso c) se recolecta en una base de datos secundaria que comprende uno y solo un conjunto de contactos theta para cada tipo de átomos theta.

5. Un método conforme a la reivindicación 4, en donde cada uno de dichos conjuntos de contactos thetas de la base de datos secundaria se subdivide en una pluralidad de tipos de átomo iota que no se solapan o grupos de tipos de átomo iota relacionados que no se solapan.

6. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde el tipo de átomo iota se identifica como uno y solo uno de:

Los 167 átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de las proteínas;

el átomo de oxígeno presente en moléculas de agua, estructuralmente relevantes, unidas a proteínas;

los 82 átomos distintos al hidrógeno presentes en los 4 nucleótidos del polímero de ácido desoxirribonucleico (ADN);

los 42 átomos distintos al hidrógeno presentes en los nucleótidos de ADN metilado, citidina fosfato y fosfato de adenosina;

los 85 átomos distintos al hidrógeno presentes en los 4 fosfatos de nucleótidos del polímero de ácido ribonucleico (ARN);

los 89 átomos distintos al hidrógeno presentes en los fosfatos nucleótidos de ribosa 2-O' metilada del ARN; y/o

los más de 400 átomos distintos al hidrógeno presentes en el ARN de base modificada más común después de la transcripción.

7. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde dicho grupo predeterminado de tipos de átomo iota relacionados es uno de una pluralidad de grupos que no se solapan obtenidos al clasificar los 167 átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de proteínas en

grupos de tipo químico similar.

8. Un método conforme a la reivindicación 7, en donde los tipos de átomo  $\iota$  se clasifican en la pluralidad de grupos que no se solapan de acuerdo con uno o más de los siguientes factores: naturaleza elemental del tipo de átomo, estado de hibridación del tipo de átomo.

9. Un método conforme a la reivindicación 7 o 8, en donde los tipos de átomo  $\iota$  se clasifican en una pluralidad de grupos que no se solapan que comprenden los siguientes:  $C sp^3$ ,  $C sp^2$ (aromáticos),  $C sp^2$ (no aromáticos),  $N sp^3$ ,  $N sp^2$ ,  $O sp^3$ ,  $O sp^2$ , S.

10. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde:

dichos datos espaciales extraídos en el paso c) definen la posición de cada átomo  $\iota$  especificado en el conjunto de contactos  $\theta$  por referencia geométrica a la posición del átomo  $\theta$  y a las posiciones del tercer y cuarto átomo;  
el tercer átomo está unido covalentemente al átomo  $\theta$ ; y  
el cuarto átomo está unido covalentemente al tercer átomo.

11. Un método conforme a la reivindicación 10, en donde:

por cada átomo  $\iota$  especificado en el conjunto de contactos  $\theta$ , dicha información espacial extraída en el paso

c) define la posición del quinto y sexto átomo por referencia geométrica a la posición del átomo  $\theta$  y a las posiciones del tercer y cuarto átomo;

el quinto átomo está unido covalentemente al átomo iota; y

el sexto átomo está unido covalentemente al quinto átomo o al átomo iota.

12. Un método conforme a la reivindicación **11**, en donde la superposición en o alrededor del sitio de destino del paso (d) comprende:

analizar el grupo de contacto theta para extraer información espacial que comprende el tipo de átomo iota específico o uno o más del grupo predeterminado de tipos de átomo iota relacionados; y

trazar esta información espacial para determinar ubicaciones teóricas que representan en dónde se ubicaría cada tipo de átomo iota o cada uno del uno o más del grupo predeterminado de tipos de átomo iota relacionados si: i) el átomo theta del contacto estuviera ubicado en la posición del átomo theta correspondiente en el sitio de unión de destino; y ii) el tercer y cuarto átomo del contacto estuvieran ubicados en las posiciones del tercer y cuarto átomo del átomo theta correspondiente en el sitio de unión de destino.

13. Un método conforme a la reivindicación 12, en donde:

los datos espaciales extraídos se analizan contra dichos datos contextuales antes de dicho paso de trazado.

14. Un método conforme a las reivindicaciones 12 and 13, en donde una región en la que una densidad de las ubicaciones teóricas del tipo de átomo iota particular, o para el uno o más de los grupos predeterminados de tipos de átomo iota relacionados, está por encima de un umbral predeterminado se identifica como una de las regiones favorecidas.

15. Un método conforme a cualquiera de las reivindicaciones 12-14, en donde las ubicaciones teóricas para el tipo de átomo iota particular o para uno o más del grupo predeterminado de tipos de átomo iota relacionados se determinan por una pluralidad de los átomos theta en la lista de destino y una región en la que una densidad de los lugares teóricos acumulados está por encima de un umbral predeterminado se identifica como una de las regiones favorecidas.

16. Un método conforme a cualquiera de las reivindicaciones 12-15, en donde:

si la ubicación teórica de un átomo iota individual se intersecta con la ubicación de un átomo de la macromolécula objetivo, más cerca que  $R_w - 0,2$  angstroms, entonces el átomo iota es excluido de análisis posteriores.

17. Un método conforme a cualquiera de las reivindicaciones 11-16, en donde:

El tercer y cuarto átomo son escogidos de manera única para cada tipo específico de átomo theta.

18. Un método conforme a cualquiera de las reivindicaciones 11-17, en donde:

El quinto y sexto átomo son escogidos únicamente para cada tipo específico de átomo iota.

19. Un método conforme a cualquiera de las reivindicaciones 11-18, en donde:

por cada región favorecida, se derivan vectores que describen la posición del quinto átomo respecto a su correspondiente átomo iota y se hace un análisis de dichos vectores para identificar un vector de enlace favorecido que represente una predicción de la unión covalente consenso teórica de un átomo iota en la región. El

vector de enlace favorecido identificado se usa para refinar el diseño del posible ligando o la modificación del posible ligando.

20. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde:

dichos datos contextuales extraídos en el paso (c), contienen información contextual sobre el ambiente local de cada par de contacto en el conjunto de contactos theta, incluyendo uno o más de los siguientes en cualquier combinación: estructura secundaria, tipos de aminoácido u otros tipos de monómero que contienen el par de contacto, las unidades de monómero adyacentes y/ o la geometría local de las mismas en una cadena de polímeros a cualquier lado del contacto, aminoácidos adyacentes en una cadena polipeptídica a cada lado del contacto, geometría local de dichas unidades de monómero adyacentes o aminoácidos, el factor de temperatura del átomo theta, el factor de temperatura del átomo iota, la superficie accesible del átomo theta, la superficie accesible del átomo iota, el número de diferentes contactos del átomo iota para el átomo theta particular, y el número de otros átomos theta en la misma unidad de monómero que el átomo theta.

21. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde el paso (f) comprende: identificar una modificación del posible ligando que aumenta un grado de solapamiento entre uno o más átomos del posible ligando y una región o regiones favorecida(s) pronosticada(s) para un tipo de átomo iota o un grupo predeterminado de tipos de átomo iota relacionados en el sitio de unión.

22. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde se identifica una pluralidad de modificaciones al posible ligando en el paso (f) y el método comprende, además, seleccionar un subconjunto de las modificaciones identificadas con base en uno o ambos de los siguientes factores: 1) la medida en que

la intersección entre los átomos alternativos y/ o adicionales del posible ligando y las respectivas regiones favorecidas del tipo de átomo iota es mayor en comparación con el posible candidato sin modificar; y 2) la medida en que uno o más factores que contribuyen a la energía total del complejo formado por la unión del posible ligando modificado al sitio de unión es/son reducido(s) comparado con los casos en los que el posible ligando no modificado está unido.

23. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde  $t = 2,5$  angstroms.

24. Un método conforme a cualquiera de las anteriores reivindicaciones, en donde  $t = 0,8$  angstroms.

25. Un método conforme a cualquiera de las anteriores reivindicaciones, que comprende además: salida de datos que representan la modificación identificada en el paso (f).

26. Un método conforme a cualquiera de las reivindicaciones 1-25, en donde el ligando es una proteína.

27. Un método conforme a la reivindicación 26, en donde el ligando es un anticuerpo.

28. Un método conforme a la reivindicación 26 o 27, en donde el paso (f) comprende reemplazar cada uno o más de los residuos de aminoácido del ligando que está(n) en contacto directo con el sitio de unión de destino o muy cerca al sitio de unión de destino, en donde cada uno o más de los residuos alternativos escogidos de entre los otros 19 aminoácidos naturales, en donde cada uno de estos reemplazos se conoce en la presente como un “residuo de reemplazo”,

en donde para cada uno de estos residuos de reemplazo que no cree conflicto con átomos adyacentes del ligando u objetivo, el tipo y posición de cada átomo del residuo de reemplazo se compara con las respectivas regiones favorecidas del tipo de átomo iota para identificar si producirán o no una mayor intersección que los átomos del residuo original.

29. Un método conforme a la reivindicación 28, que comprende además:  
la generación de una lista de los residuos de reemplazo identificados por producir una mayor intersección que los átomos del residuo original;  
por cada residuo de reemplazo enumerado, mutar el posible ligando para producir un ligando modificado que incorpora el residuo de reemplazo;  
ensayar la afinidad de cada ligando modificado con el sitio de unión de destino para identificar los residuos de reemplazo que proporcionan una mejora en la afinidad mayor a un umbral predeterminado.

30. Un método conforme a la reivindicación 29, que comprende además:  
modificar el posible ligando para que incorpore una pluralidad de residuos de reemplazo, los cuales se ha determinado que resultan en una mayor afinidad por encima del umbral predeterminado.

31. Un medio o señal legible por computador que comprende instrucciones legibles por computador para hacer que un computador lleve a cabo el método de cualquiera de las anteriores reivindicaciones.

32. Un medio o señal conforme a la reivindicación 31, en donde se hace que el computador realice al menos los pasos (c) - (e).

33. Un medio o señal conforme a la reivindicación 31 o 32, en donde se hace que el computador realice al menos el paso (f).

34. Un método de fabricación de un ligando terapéutico, que comprende:

diseñar un ligando terapéutico conforme al método de cualquiera de las reivindicaciones 1-30; y fabricar el ligando terapéutico diseñado de este modo.

35. Un ligando terapéutico fabricado conforme al método de la reivindicación 34.

36. Un método o ligando conforme a la reivindicación 34 o 35, en donde el ligando es una proteína.

37. Un método o ligando conforme a la reivindicación 36, en donde la proteína es un anticuerpo.

38. Un método para generar una base de datos para uso en un método para diseñar

un ligando *ab initio* que se enlazará a un sitio de unión de un destino macromolecular, o para identificar una modificación a un ligando con el fin de mejorar su afinidad con un sitio de unión de un destino macromolecular, el cual comprende:

analizar las posiciones relativas de los átomos en cada una de una pluralidad de proteínas o demás macromoléculas biológicas con el fin de identificar instancias de contacto intramolecular no enlazante entre un primer átomo conocido como átomo theta y un segundo átomo conocido como átomo iota de la proteína o macromolécula; y

generar una base de datos que por cada contacto identificado especifica: el tipo del átomo theta, el tipo del átomo iota y la posición del átomo iota respecto al átomo theta;

en donde un contacto intramolecular no enlazante se define como una instancia en la que se cumplen las siguientes condiciones:

$s - R_w = t$ , en donde  $s$  es la separación entre los átomos theta y iota,  $R_w$  es la suma de los radios de van de Waals de los dos átomos theta y iota, y  $t$  es una distancia umbral predeterminada, usualmente de 2,5 angstroms y preferiblemente de 0,8 angstroms; y en donde, en el caso de las proteínas, los átomos theta y iota son residuos de aminoácidos separados el uno del otro por al menos cuatro residuos en un polipéptido lineal o que están en cadenas polipeptídicas separadas.

39. Un método conforme a la reivindicación 38, en donde el método comprende la subdivisión de la base de datos para formar grupos de contactos identificados, en donde el átomo theta es uno y solo uno de los 167 átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de las proteínas y el átomo iota es uno y solo uno de una pluralidad de grupos que no se solapan, obtenidos al clasificar los 167 átomos distintos al hidrógeno de los 20 aminoácidos naturales de las proteínas en grupos de acuerdo con la similitud química.

40. Un método para generar una base de datos para uso en un método para diseñar un ligando *ab initio* que se enlazará a un sitio de unión de un destino macromolecular, o para identificar una modificación a un ligando con el fin de mejorar su afinidad con un sitio de unión de un destino macromolecular, el cual comprende:

analizar las posiciones relativas de los átomos en cada una de una pluralidad de proteínas o demás macromoléculas biológicas con el fin de identificar instancias de contacto intramolecular no enlazante entre un primer átomo conocido como átomo theta y un segundo átomo conocido como átomo iota de la proteína o macromolécula;  
y

generar una base de datos que por cada contacto identificado especifica: el tipo del átomo theta, el tipo del átomo iota y la posición del átomo iota respecto al átomo theta;

en donde un contacto intramolecular no enlazante se define como una instancia en la que se cumple la siguiente condición:

$s - R_w = t$ , en donde  $s$  es la separación entre los átomos theta y iota,  $R_w$  es la suma de los radios de van de Waals de los dos átomos theta y iota, y  $t$  es una distancia umbral predeterminada, usualmente de 2,5 angstroms y preferiblemente de 0,8 angstroms; y en donde el método comprende la subdivisión de la base de datos para formar grupos de contactos identificados, en donde el átomo theta es uno y solo uno de los 167 átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de las proteínas y el átomo iota es uno y solo uno de una pluralidad de grupos que no se solapan, obtenidos al clasificar los 167 átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de las proteínas en grupos de acuerdo con la similitud química.

41. Un método conforme a la reivindicación 39 o 40, en donde:

por cada contacto, la posición de cada átomo iota se define por referencia geométrica

a la posición del átomo theta y a las posiciones del tercer y cuarto átomo, en donde el tercer átomo está unido covalentemente al átomo theta y el cuarto átomo está unido covalentemente al tercer átomo. El método comprende, además:

normalización de las coordenadas del átomo iota, el átomo theta, el tercer átomo y el cuarto átomo de cada contacto como un grupo para generar un grupo de coordenadas normalizadas;

por cada uno o más de los tipos de átomo theta, usar los grupos de coordenadas normalizadas para una pluralidad de contactos que involucran el tipo de átomo theta y un tipo de átomo iota particular para generar un gráfico polar de dos dimensiones que representa una distribución de direcciones del átomo iota particular, en términos de latitud y longitud, respecto al átomo theta;

repetir el paso anterior para tipos de átomo iota diferentes;

comparar los gráficos polares de dos dimensiones resultantes para identificar grupos de tipo de átomo iota que producen distribuciones similares de las direcciones y usar estos grupos como los grupos basados en la similitud química para clasificar los 167 átomos distintos al hidrógeno presentes en los 20 aminoácidos naturales de las proteínas en una pluralidad de grupos que no se solapan.

42. Un método para generar una base de datos conforme a cualquiera de las reivindicaciones 38-41, que comprende: extraer información de contacto de al menos 2000 proteínas u otras macromoléculas biológicas, en donde la información extraída contiene información de al menos dos millones de pares de átomos de contacto.

43. Un método para generar una base de datos conforme a cualquiera de las reivindicaciones 38-42, que comprende:

extraer información de contacto de al menos 10000 proteínas u otras macromoléculas biológicas, en donde la información extraída contiene información de al menos diez millones de pares de átomos de contacto.

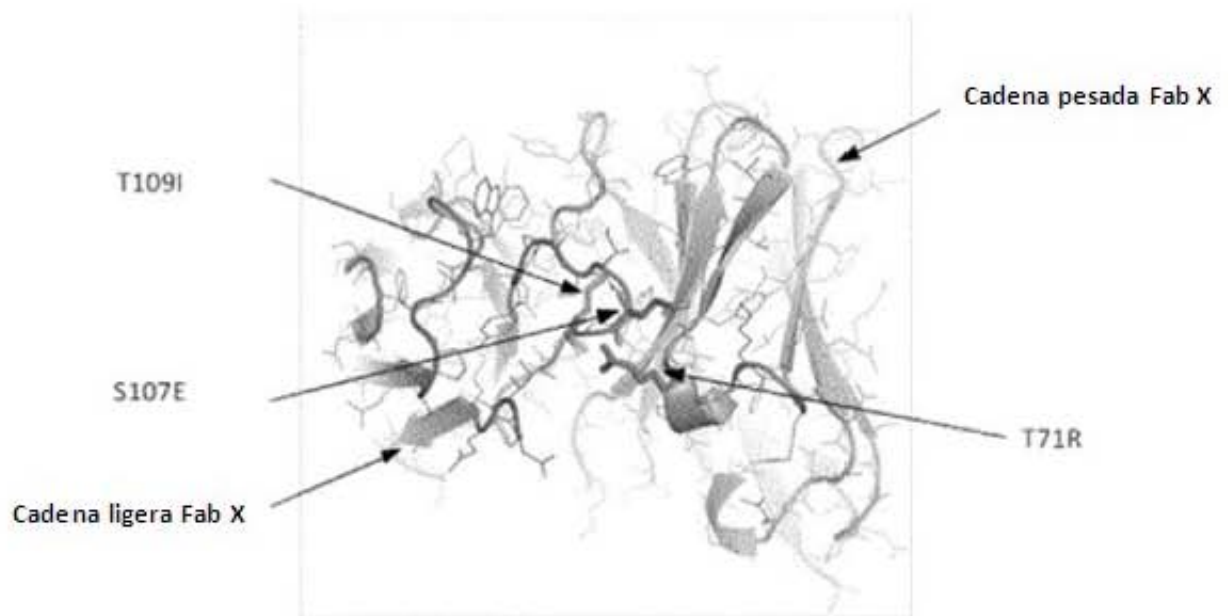
44. Un medio legible por computador que almacena una base de datos generada conforme a cualquiera de las anteriores reivindicaciones.

45. Un método conforme a cualquiera de las reivindicaciones 1 a 30 o que usa bases de datos compiladas conforme a cualquiera de las reivindicaciones 38 a 43, en donde: para un complejo de anticuerpo-antígeno particular, se predicen mutaciones específicas a los residuos de aminoácido en o alrededor del sitio de destino del anticuerpo para producir una mayor afinidad de unión del anticuerpo al antígeno.

## RESUMEN:

Métodos y aparatos asociados que involucran el diseño de un ligando *ab initio* que se unirá a un sitio de unión de un destino macromolecular, o de identificación de una modificación a un ligando para mejorar la afinidad del ligando a un sitio de unión de un destino macromolecular, que comprende el uso de información de contactos átomo a átomo intramoleculares o intermoleculares no enlazantes tomados de una base de datos de macromoléculas biológicas para identificar las regiones favorecidas adyacentes al sitio de unión para tipos particulares de átomos y modificar un posible candidato con el fin de aumentar la intersección entre los átomos del posible candidato y las regiones favorecidas. Uno o más pasos de los métodos pueden ser realizados por computador.

Fig. 18



**Fig. 17**

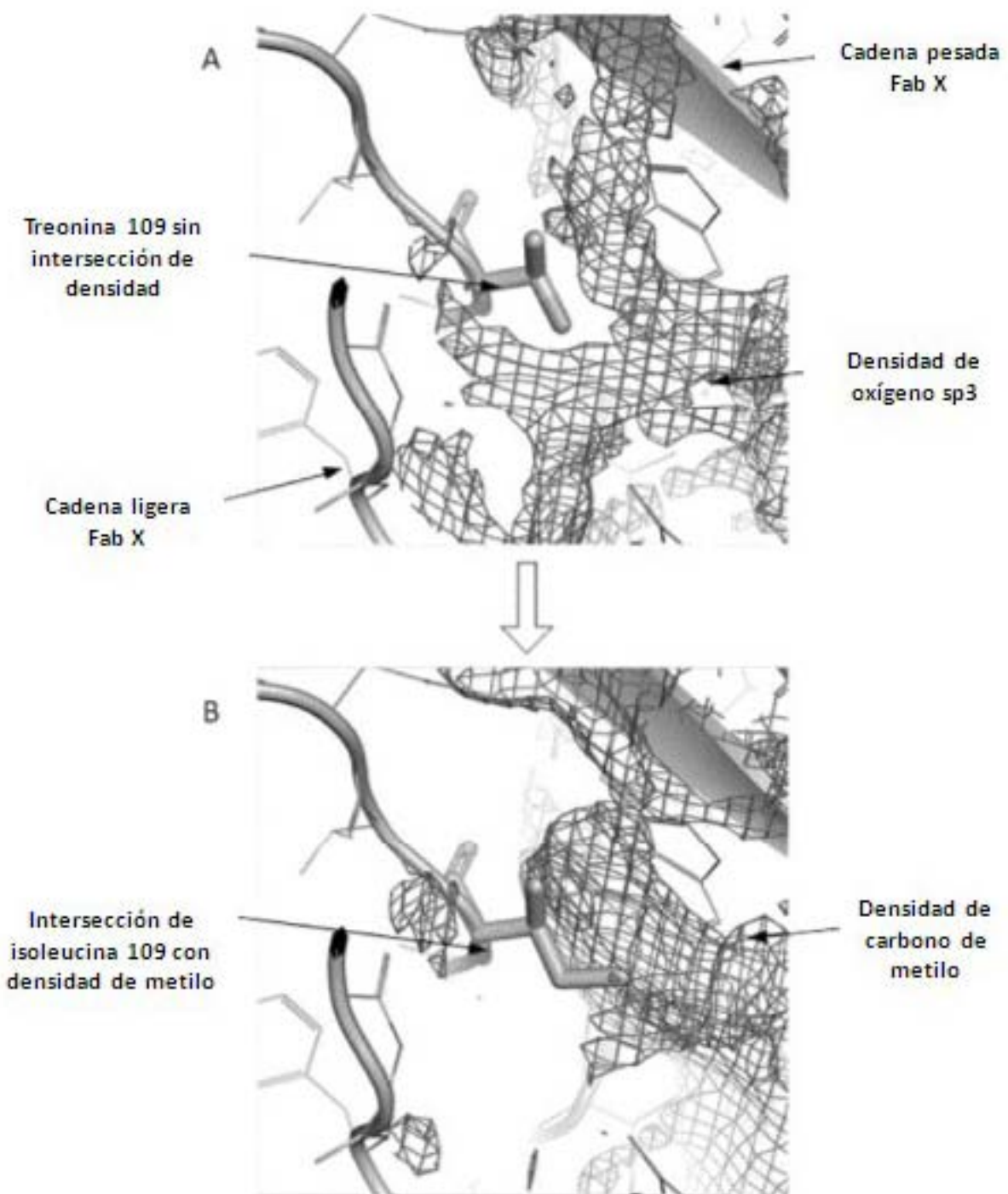


Fig. 16

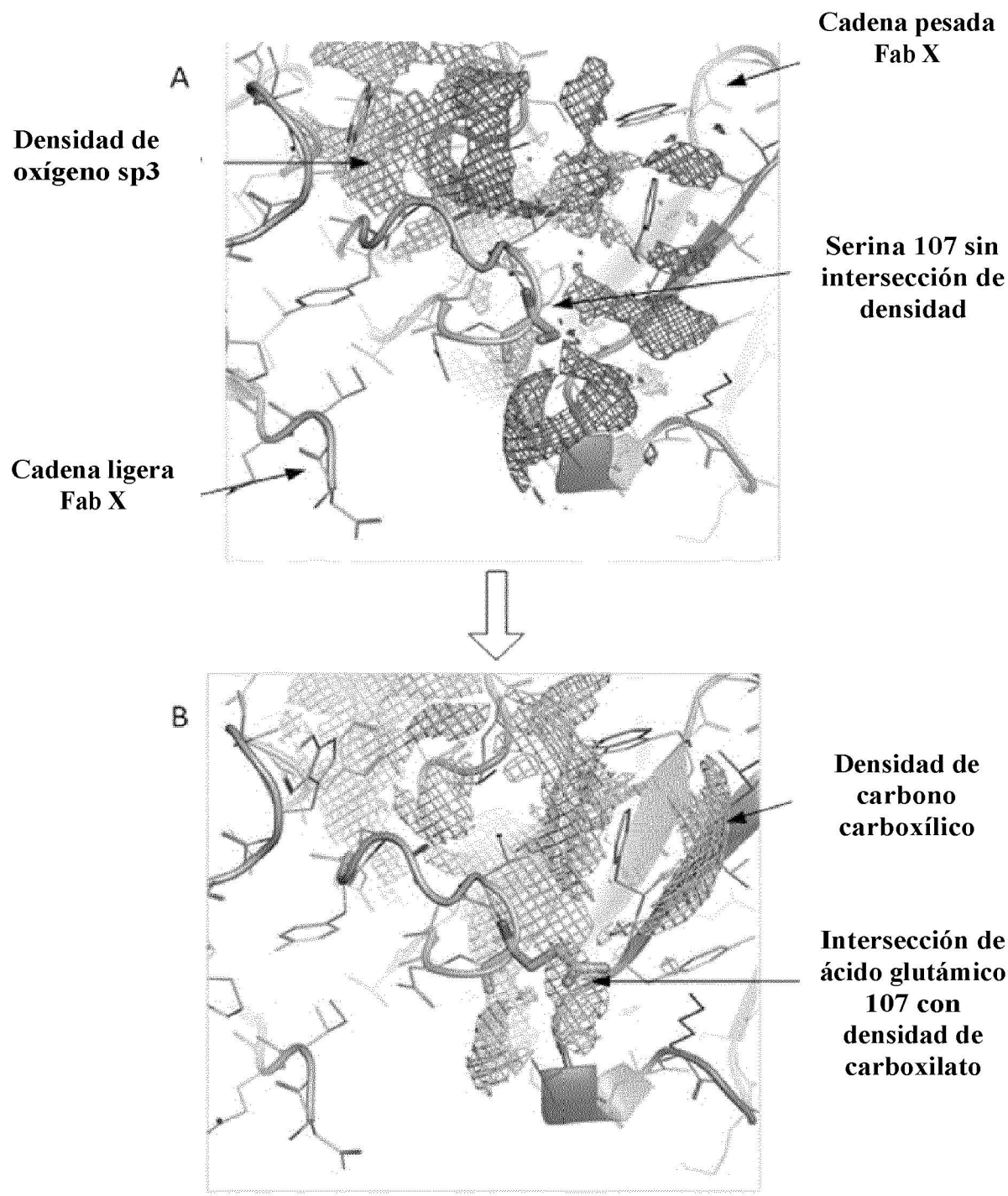


Fig. 15

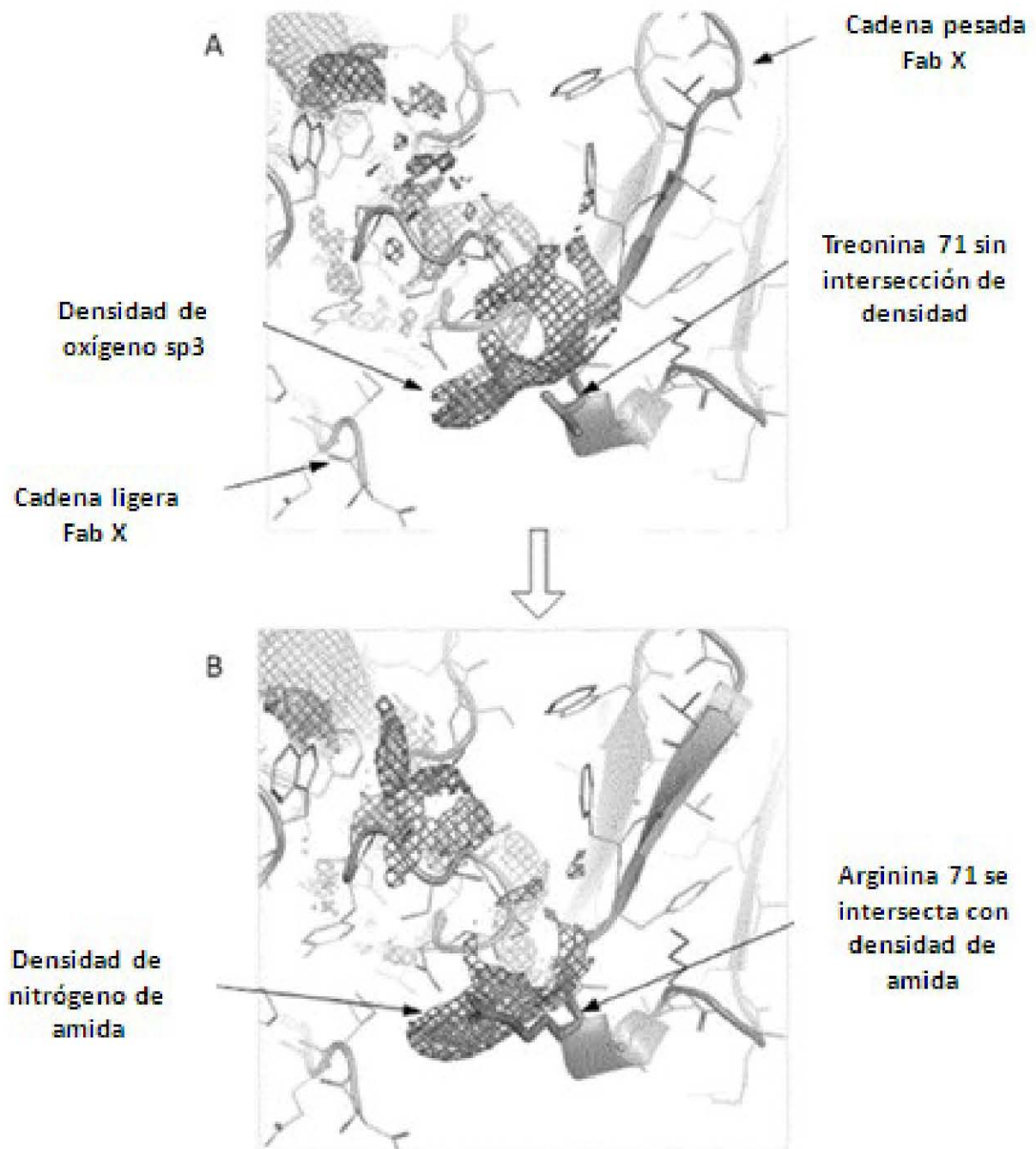


Fig. 14

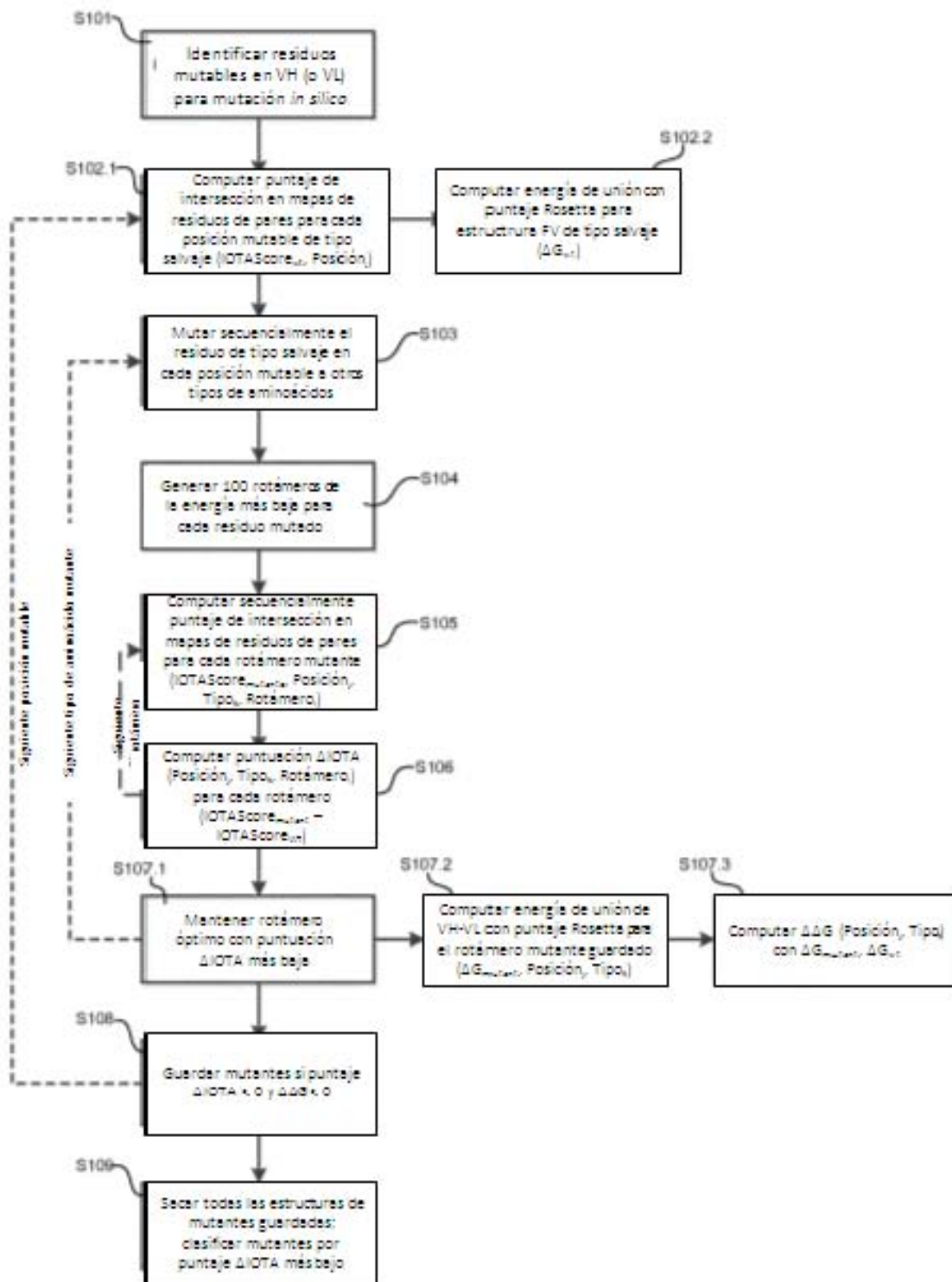
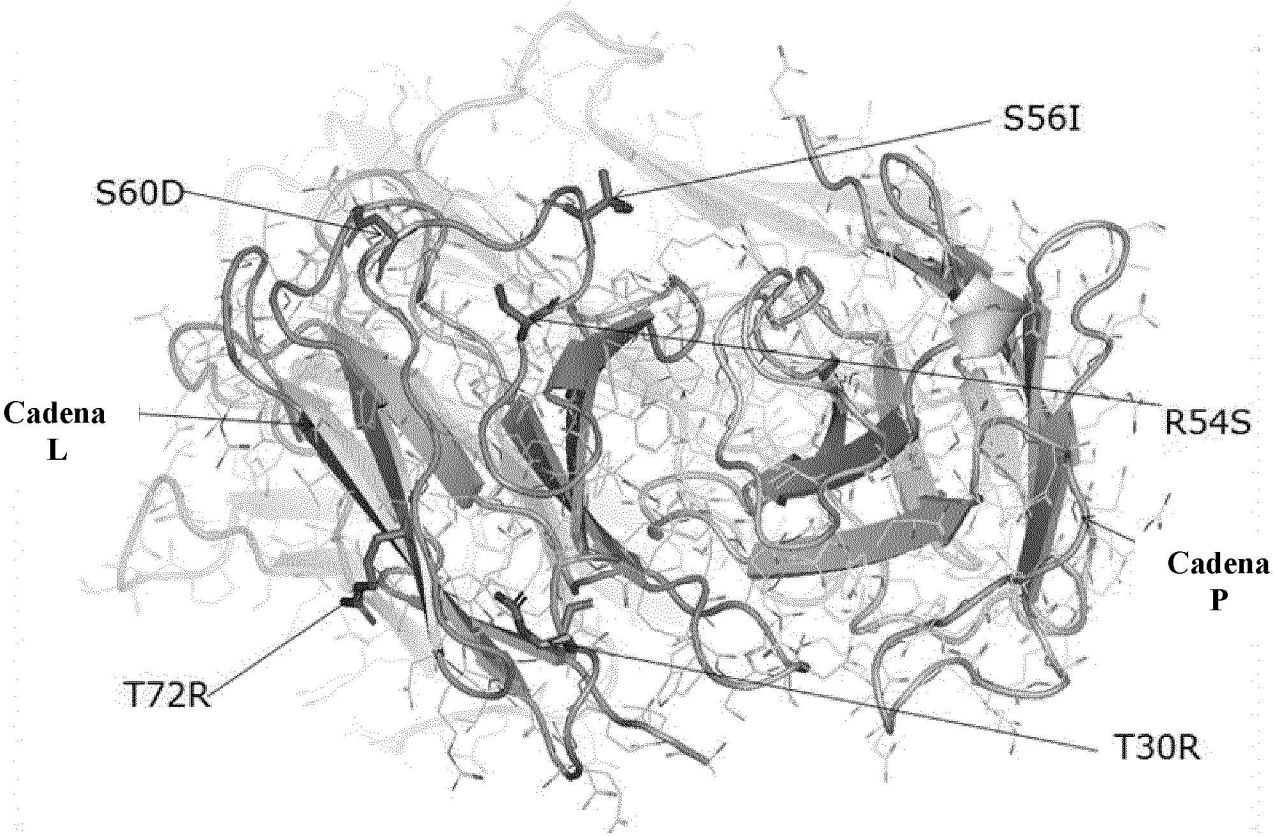


Fig. 15



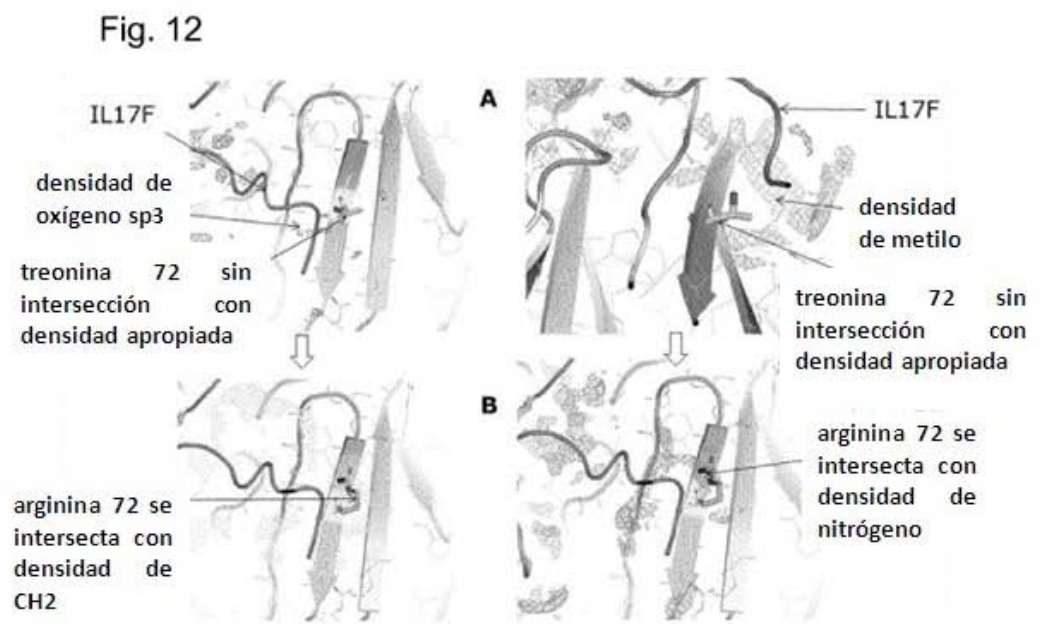
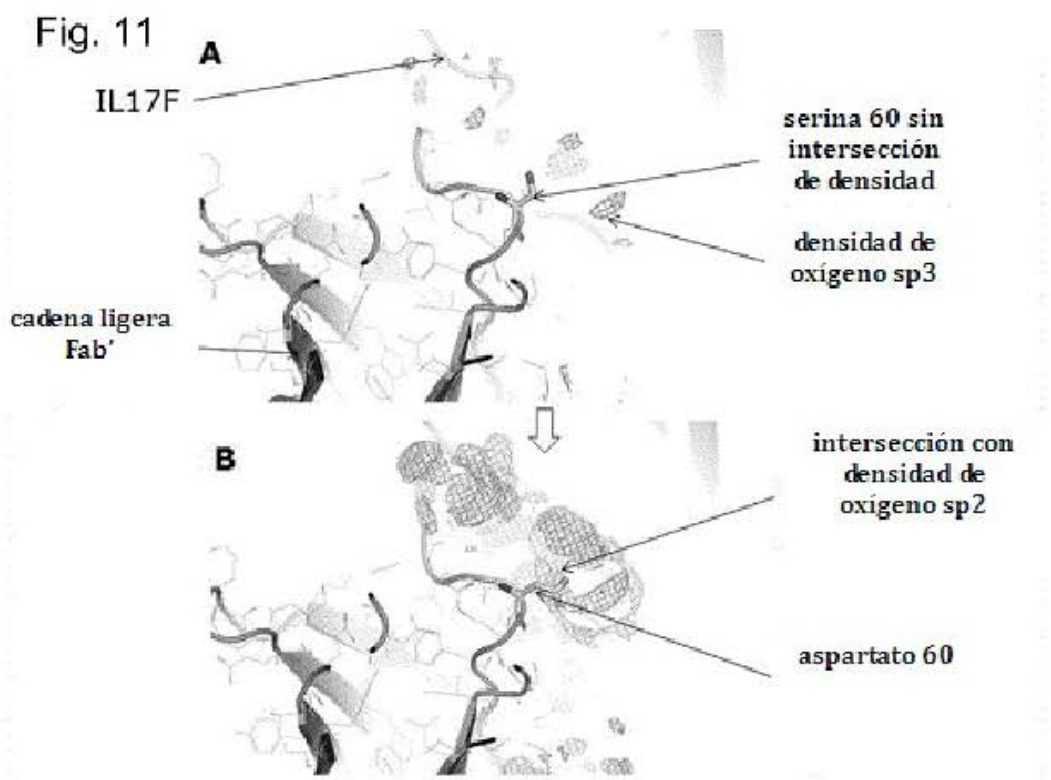


Fig. 9

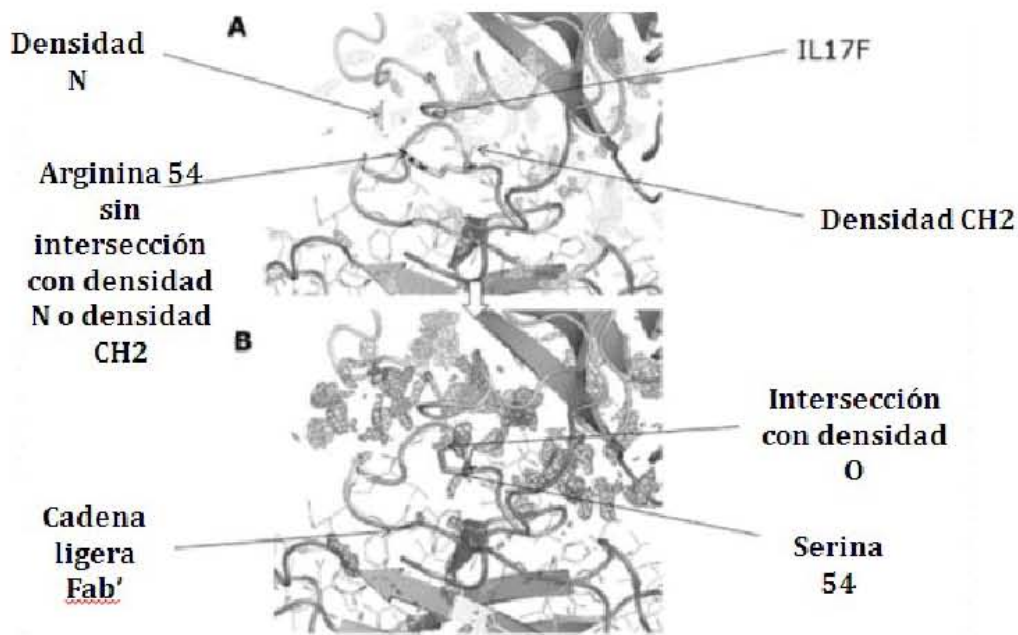


Fig. 10

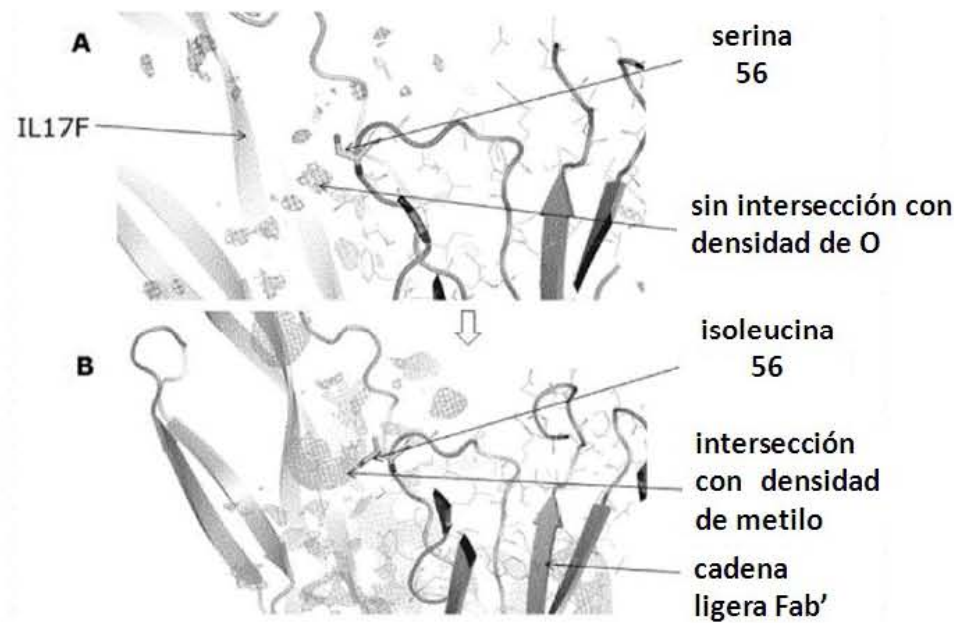


Fig. 8

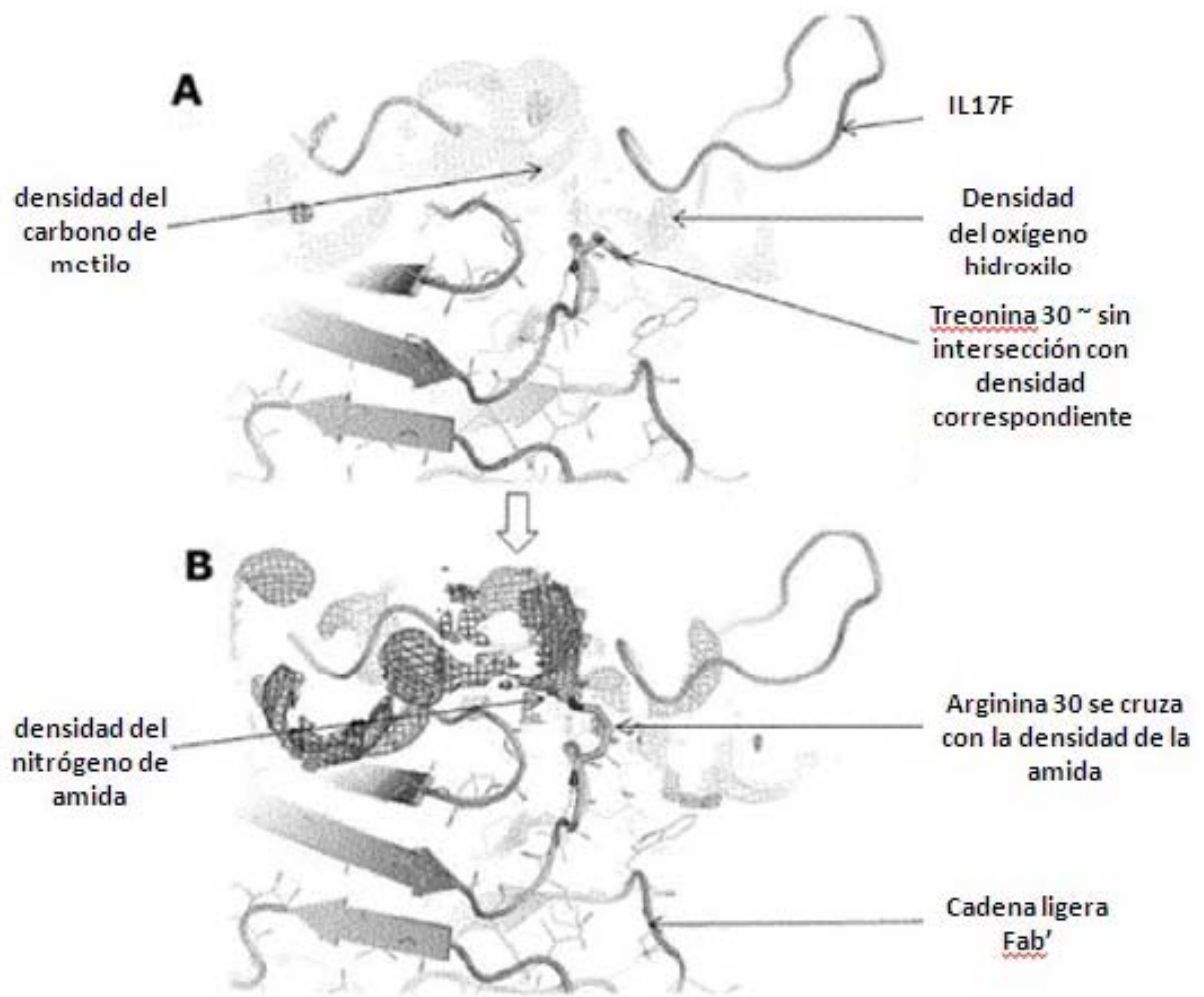


Fig. 7

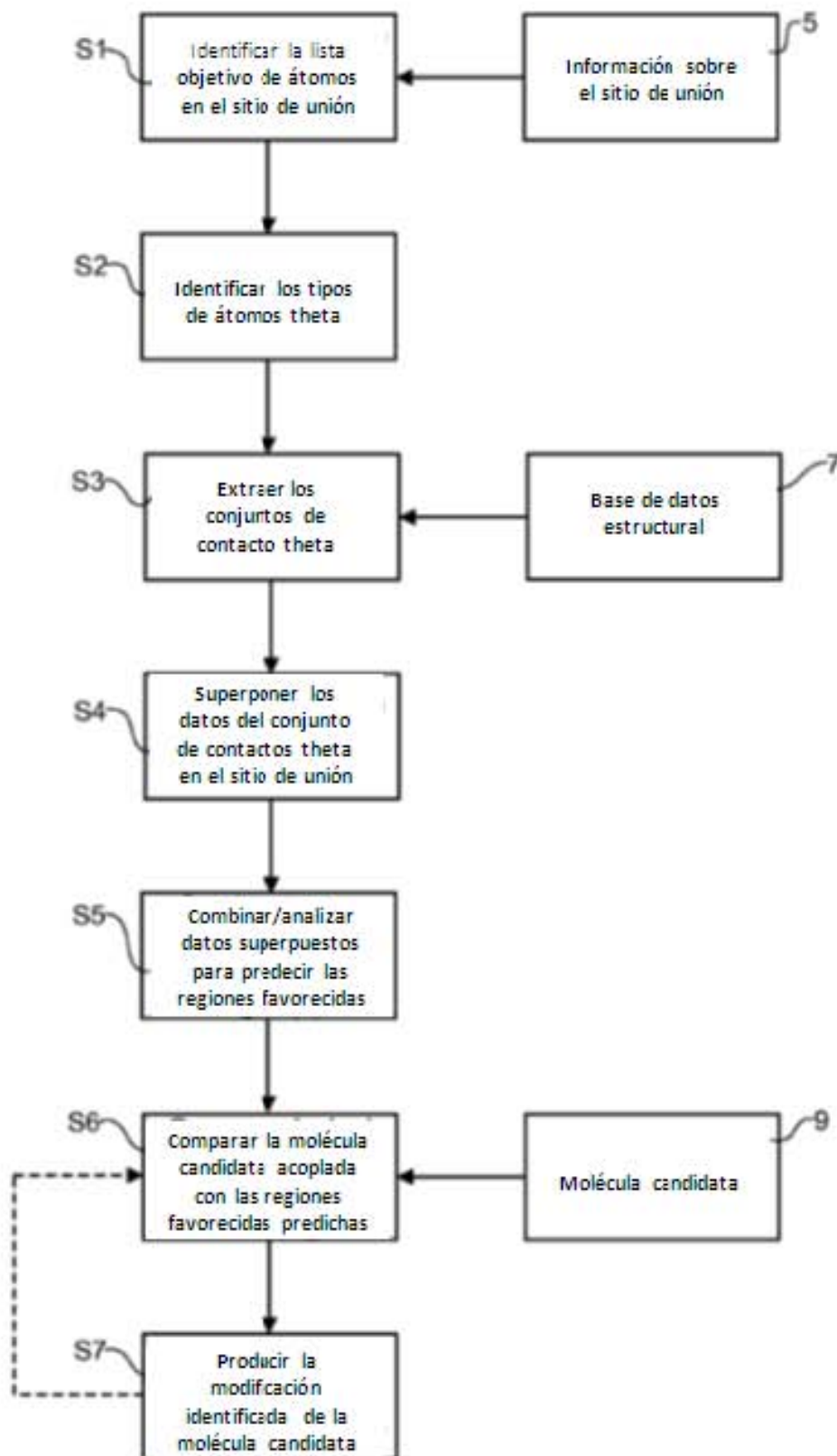


Fig. 6

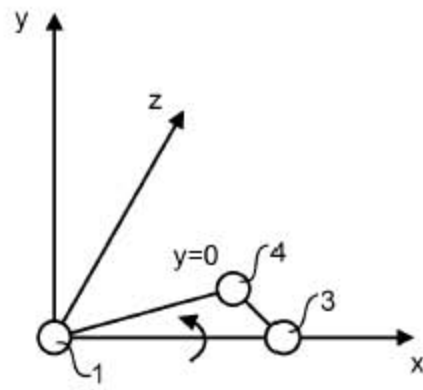


Fig. 1

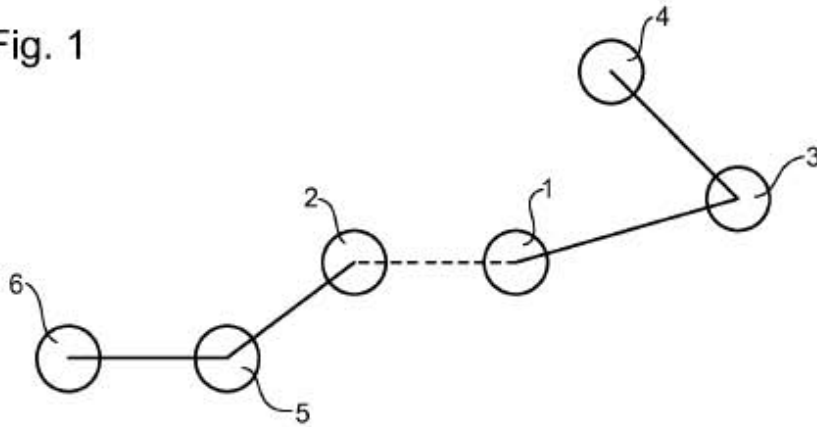


Fig. 2

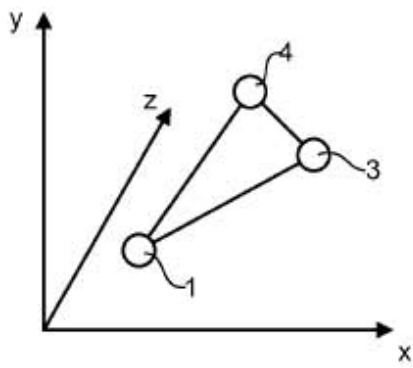


Fig. 3

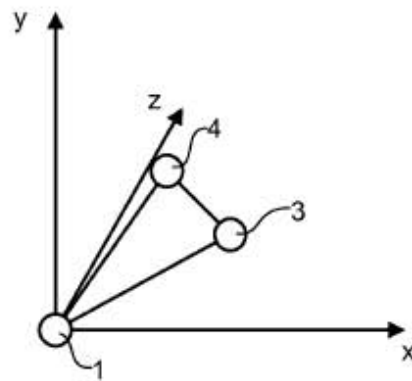


Fig. 4

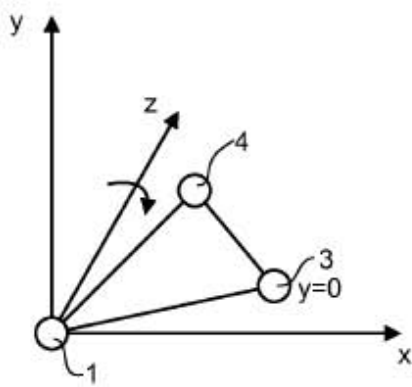


Fig. 5

