



①9



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

①1 Número de publicación: **2 356 620**

⑤1 Int. Cl.:
C07K 1/00 (2006.01)

⑫

TRADUCCIÓN DE PATENTE EUROPEA

T3

⑨6 Número de solicitud europea: **02757870 .7**

⑨6 Fecha de presentación : **30.03.2002**

⑨7 Número de publicación de la solicitud: **1373301**

⑨7 Fecha de publicación de la solicitud: **10.10.2002**

⑤4 Título: **Reducción de la inmunogenicidad de las proteínas de fusión.**

③0 Prioridad: **30.03.2001 US 280625 P**

④5 Fecha de publicación de la mención BOPI:
11.04.2011

④5 Fecha de la publicación del folleto de la patente:
11.04.2011

⑦3 Titular/es: **MERCK PATENT GmbH**
Frankfurter Strasse 250
64293 Darmstadt, DE

⑦2 Inventor/es: **Gillies, Stephen, D.**

⑦4 Agente: **Carvajal y Urquijo, Isabel**

ES 2 356 620 T3

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

DESCRIPCIÓN

Reducción de la inmunogenicidad de las proteínas de fusión.

Campo de la Invención

- 5 La presente invención hace referencia en general métodos y composiciones para la producción y el uso de proteínas de fusión modificada sin inmunogenicidad o con inmunogenicidad reducida como agentes terapéuticos. Más específicamente, la invención hace referencia a proteínas de fusión, producidas con una menor inmunogenicidad mediante la identificación de los epítopos candidatos de células T y la modificación de la secuencia de aminoácidos para eliminar tales epítopos.

Antecedentes de la Invención

- 10 Muchas proteínas terapéuticas son proteínas humanas normales. Por ejemplo, la interleucina-2, la eritropoyetina, y las hormonas del crecimiento son todas proteínas humanas que son suministradas a humanos que ya producen habitualmente niveles endógenos de estas proteínas. En general, las respuestas inmunes contra las proteínas humanas completamente normales son poco frecuentes cuando estas proteínas se utilizan como agentes terapéuticos.

- 15 Recientemente se ha vuelto evidente que muchas proteínas de fusión con actividades artificiales son útiles como proteínas terapéuticas. Por ejemplo, Enbrel es una fusión del dominio extracelular de un receptor TNF con una región Fc de la IgG1. Enbrel se utiliza para tratar artritis reumatoide, y se cree que funciona por titulación del TNF y la supresión de la actividad del TNF. Sin embargo, una incidencia significativa de anticuerpos anti-Enbrel se han descubierto en pacientes tratados con Enbrel.

- 20 Otro ejemplo de una clase terapéuticamente útil de proteínas de fusión es la inmunocitocinas. Estas proteínas incluyen una fracción de anticuerpo y una fracción de citocina, y son útiles para dirigir las citocinas hacia células enfermas diana, tales como las células cancerosas. Sin embargo, el uso terapéutico de muchas de estas proteínas de fusión es reducido debido a su inmunogenicidad en mamíferos, especialmente en humanos.

- 25 La patente WO 02/066514 revela las proteínas de fusión de inmunocitocina con una inmunogenicidad reducida, en donde también la región de unión entre la molécula Ig y la citocina puede ser menos inmunogénica mediante la eliminación de los epítopos de células T de esa región. Sin embargo, este documento no muestra modificaciones específicas dentro de dicha región de unión.

- 30 La patente WO 00/34317 muestra un concepto general de reducción de la inmunogenicidad de la proteína mediante la identificación y modificación de los epítopos de células T. Sin embargo, este documento no menciona las proteínas de fusión.

Por lo tanto, existe la necesidad de producir proteínas de fusión con una inmunogenicidad reducida a fin de utilizar estas proteínas terapéuticamente.

Resumen de la Invención

- 35 La presente invención ofrece métodos y composiciones útiles para la producción de proteínas de fusión con una inmunogenicidad reducida para su uso terapéutico. Por ejemplo, la invención ofrece inmunocitocinas, inmunofusinas, inmunoligandos, otros anticuerpos y proteínas de fusión Fc, proteínas de fusión citocina-citocina, y proteínas de fusión de albúmina con inmunogenicidad disminuida.

- 40 La presente invención hace referencia en parte a la percepción de que las proteínas de fusión contienen secuencias que son "no propias." Por ejemplo, incluso en una fusión entre dos proteínas humanas, la región que rodea la unión de fusión comprende una secuencia peptídica que no está habitualmente presente en el cuerpo humano. Por ejemplo, un fármaco proteico tal como Enbrel se deriva a partir de dos proteínas humanas normales: el receptor TNF y la IgG1. Sin embargo, la unión entre el receptor TNF y la IgG1 es una secuencia peptídica que no se encuentra habitualmente en el cuerpo humano.

- 45 Los métodos preferentes de la invención incluyen la reducción de la inmunogenicidad de una proteína de fusión mediante la reducción de la habilidad de un epítipo de unión (péptido de unión) de interactuar con un receptor de células T mediante la reducción de su habilidad de unirse (su afinidad de unión) a las moléculas MHC. De acuerdo con la invención, el epítipo o péptido de unión es preferentemente "no-propio." En general, las proteínas, incluyendo las proteínas terapéuticas, son inmunogénicas, en parte porque las células presentadoras de antígeno endocitan las proteínas y éstas son proteolizadas, y los péptidos resultantes se unen a las moléculas llamadas
- 50 Complejo Mayor de Histocompatibilidad (MHC) que presentan los péptidos a las células T. El complejo antigénico péptido-MHC sobre la superficie de una célula presentadora de antígeno (APC) activa a las células T para proliferar, diferenciar y liberar citocinas. Paralelamente, se induce la diferenciación de células B y la producción de anticuerpos lo que puede además limitar la efectividad de la proteína terapéutica debido a la remoción. De ese modo, el péptido antigénico, si se deriva a partir de una proteína terapéutica, es capaz de inducir una serie de respuestas inmunes

indeseables. La efectividad de la proteína terapéutica se encuentra limitada por los anticuerpos debido a su titulación, y la inducción de respuestas de las células T y las células B es habitualmente perjudicial debido a las reacciones inflamatorias y alérgicas en el paciente.

La presente invención proporciona (1) la identificación de nuevas secuencias de aminoácidos en la región de la unión inmunoglobulina-proteína diana con uno o más epítopos de células T candidatos; y (2) la modificación de estas secuencias de aminoácidos para reducir o eliminar la presencia de péptidos, derivados a partir de la secuencia de unión, que funcionan como epítopos de células T.

La invención proporciona dos clases de composiciones y métodos generales que se relacionan con la reducción de la inmunogenicidad. De acuerdo con una realización de la invención, los epítopos potenciales de células T no propias se identifican en las secuencias que abarcan la unión de fusión. Por ejemplo, los epítopos potenciales de células T no propias se identifican mediante métodos computacionales basados en el modelado de enlace de péptidos a moléculas MHC Clase II. Las sustituciones se realizan, entonces, de modo tal que la habilidad de los péptidos que se derivan a partir de la región de unión para el enlace con MHC Clase II se reduce o se elimina. Este proceso de identificación y modificación de péptidos que se unen a MHC Clase II se denomina "desinmunización" y las moléculas de proteínas modificadas resultantes se denominan "desinmunizadas".

De acuerdo con otra realización de la invención, uno o más sitios de glicosilación se introducen en la unión de fusión. Se utiliza preferentemente un sitio de glicosilación N-ligada, aunque también puede utilizarse un sitio de glicosilación O-ligada. De acuerdo con una realización preferente, los aminoácidos en una región de unión que rodea una unión de fusión de una secuencia de tipo salvaje se mutan de modo que el último aminoácido del compañero de fusión del extremo N-terminal se muta a una asparagina, y los dos primeros aminoácidos del segundo compañero de fusión se mutan a una glicina seguida por una serina o una treonina.

De acuerdo con la invención, se prefiere la eliminación del enlace MHC Clase II en situaciones en donde se producirá una proteína en bacterias o en un organismo que no genere un patrón de glicosilación como en los mamíferos, tales como la levadura o las células de insectos.

Puede preferirse la introducción de sitios de glicosilación cuando la proteína ha de ser producida en una línea celular de mamífero o en una línea celular que crea un patrón de glicosilación que es inocuo para los mamíferos.

En una realización preferente, un componente de la proteína de fusión es una citocina. El término "citocina" se utiliza en la presente para describir proteínas recombinantes o naturales, análogas de éstas, y fragmentos de las mismas que provocan una respuesta específica en una célula que tiene un receptor para esa citocina. Preferentemente, las citocinas son proteínas que pueden ser producidas y excretadas por una célula. Preferentemente, las citocinas incluyen a las Interleucinas tales como Interleucina-2 (IL-2), IL-3, IL-4, IL-5, IL-6, IL-7, IL-10, IL-12, IL-13, IL-14, IL-15, IL-16 y IL-18, factores hematopoyéticos tales como el factor de estimulación de colonias de granulocitos-macrófagos (GM-CSF), G-CSF y eritropoyetina, los factores de necrosis tumoral (TNF) tales como el TNF α , las linfoquinas tales como la linfoxina, los reguladores de los procesos metabólicos tales como la leptina, y los interferones tales como el interferón α , el interferón β , y el interferón γ y las quimiocinas. Preferentemente, la proteína de fusión anticuerpo-citocina de la presente invención muestra una actividad biológica específica de la citocina.

En otra realización, un componente de la proteína de fusión es una citocina antiobesidad. Por ejemplo, un componente es la leptina, el CNTF, o una porción de Acrp30.

En una realización alternativa, un componente de la proteína de fusión es una hormona. Por ejemplo, un componente puede ser insulina, hormona de crecimiento, o un péptido similar al glucagón tipo 1 (GLP-1).

En otra realización alternativa, un componente de la proteína de fusión es una proteína de unión al ligando con actividad biológica. En una realización, se utiliza un dominio extracelular del receptor TNF.

De acuerdo con una serie de realizaciones, una proteína de fusión de la invención comprende el extremo N-terminal de una fracción que no es anticuerpo fusionada a un extremo C-terminal de una fracción de anticuerpo. De acuerdo con otra serie de realizaciones, una proteína de fusión de la invención comprende el extremo C-terminal de una fracción que no es anticuerpo fusionada a un extremo N-terminal de una fracción de anticuerpo. De acuerdo con la invención, una fracción de anticuerpo puede ser una inmunoglobulina intacta o una porción de una inmunoglobulina intacta. Una porción de una inmunoglobulina puede incluir una región variable o una región constante o ambas. Las inmunoglobulinas preferentes incluyen las regiones Fc o porciones de las mismas. Una realización preferente de la invención incluye un isotipo de inmunoglobulina IgG1, o una porción de la misma, modificada para ser menos inmunogénica y/o para tener una vida media en suero más larga. Por ejemplo, se prefiere una IgG1 con modificación de los residuos de aminoácidos cercanos a la unión CH3-citocina. Para ciertas aplicaciones, se prefieren fracciones de anticuerpos a partir de los isotipos IgG2 o IgG4.

Las inmunocitocinas son sólo un ejemplo de una terapia de proteínas de fusión cuya diana es un tumor. También se pueden utilizar otras moléculas tóxicas cuya diana sean tumores mediante la fusión a anticuerpos

específicos del tumor. Además, las proteínas de fusión a anticuerpos pueden atacar otros tipos de células enfermas, tales como las células infectadas por virus. Se ha utilizado otra aproximación para la ingeniería de las proteínas de fusión dirigidas a dianas a partir de la tecnología Fc-X y X-Fc donde X es un polipéptido. Estas tecnologías utilizan el conocimiento de que la producción y la obtención de una proteína diana se mejora si el polipéptido de interés se enlaza a una porción Fc de una inmunoglobulina. Para las proteínas de fusión Fc-X, un péptido señal, seguido por el fragmento Fc de un gen inmunoglobulina es el compañero de fusión del extremo N-terminal a la proteína diana. En algunos casos es específicamente ventajoso desarrollar una proteína de fusión en la orientación X-Fc. Con estas construcciones la proteína diana es la proteína de fusión del extremo N-terminal y le sigue el fragmento Fc. Para algunas proteínas esta aproximación es útil, como se ha mostrado con la glicoproteína de la superficie celular linfocítica (LHR) (patente estadounidense US 5.428.130), y el péptido similar al glucagón tipo 1 (GLP-1).

En consecuencia, los métodos y composiciones de la invención proporcionan formas de las proteínas de fusión Fc-X y X-Fc con una inmunogenicidad reducida. De acuerdo con la invención, la inmunogenicidad de una proteína de fusión puede comprobarse de acuerdo con un método conocido en el arte o revelado en este documento.

Los métodos y las composiciones de la invención también proporcionan proteínas de fusión de albúmina con inmunogenicidad reducida. La albúmina de suero humana (HSA), debido a su notablemente larga vida media, su amplia distribución in vivo y su falta de funciones enzimáticas o inmunológicas, se ha utilizado como un portador para los péptidos o proteínas terapéuticos (Yeh et al, PNAS 89:1904-1908, 1992). Una fusión genética de un péptido bioactivo a la HSA resulta útil para la recuperación de un derivado de HSA terapéutica segregada. Sin embargo, de acuerdo con la invención, las proteínas de fusión de albúmina tales como la HSA-CD4 tienen una unión novedosa que generalmente contiene uno o más epítopos de células T capaces de ser presentados en las moléculas MHC clase II. La invención proporciona menos formas inmunogénicas de proteínas de fusión de albúmina, y métodos para la reducción de la inmunogenicidad de las proteínas de fusión de albúmina. De acuerdo con la invención, las proteínas de albúmina útiles incluyen a las variantes de especie, mutantes y alélicas de albúmina, incluyendo los fragmentos de las mismas. Las proteínas de albúmina preferentes retienen las propiedades funcionales y estructurales de una proteína de albúmina del tipo natural tales como la HSA.

En otro aspecto, la invención proporciona las proteínas de fusión de anticuerpos desimmunizados con isotipos normales, mutantes, o híbridos que comprenden mutaciones útiles. Estas mutaciones pueden estar cerca de la unión o en posiciones separadas de la región de la unión.

Por ejemplo, la invención proporciona una inmunocitocina desimmunizada, modificada en la unión, con una mutación de punto en la unión entre las fracciones IgG y no IgG. La fracción citocina incluye cualquier citocina pero preferentemente IL-2 o IL-12. En una realización, los cambios del aminoácido involucran el cambio de la lisina del extremo C-terminal de la fracción de anticuerpo a un aminoácido hidrofóbico tal como la alanina o la leucina. Una ventaja clave de la combinación de tales mutaciones con una modificación que desimmuniza de la invención es que las mutaciones actúan en conjunto para aumentar la vida media en suero y para disminuir la inmunogenicidad. Los métodos descritos en este documento para la combinación de la desimmunización de una unión de fusión con una mutación que altera la vida media en suero son útiles para mejorar significativamente la eficacia clínica de estas proteínas de fusión.

En otro aspecto, la invención proporciona inmunocitocinas que comprenden una fracción de anticuerpo híbrido que incluye dominios a partir de diferentes isotipos Ig, preferentemente a partir de isotipos tanto de IgG1 como de IgG2, y una modificación desimmunizante en la unión de fusión. Por ejemplo, la invención proporciona una inmunocitocina modificada en la unión desimmunizada utilizando un híbrido IgG2 y un híbrido IgG2h (IgG2 modificado en la región bisagra a la IgG1). En una realización preferente, la proteína de fusión híbrida consiste en una fracción de inmunoglobulina desimmunizada compuesta de una fracción IgG ($\gamma 1$:CH1-H)($\gamma 2$: CH2-CH3) y una fracción citocina.

En otro aspecto, la invención proporciona secuencias de ácido nucleico novedosas que codifican proteínas de fusión con inmunogenicidad reducida o facilitan la expresión, producción, y secreción de proteínas de fusión con inmunogenicidad reducida. Tales ácidos nucleicos se generan de acuerdo con técnicas de ADN recombinante estándares.

En una realización preferente, una molécula de ácido nucleico codifica una proteína de fusión de inmunocitocina. Una inmunocitocina preferente incluye una citocina, por ejemplo, la Interleucina 2, y un anticuerpo monoclonal específico de tumor tales como un anticuerpo a la molécula de adhesión celular epitelial humana KSA (EP-CAM)(huKS).

En otra realización preferente, las moléculas de ácido nucleico codifican proteínas de fusión Fc en diversas configuraciones. La molécula de ácido nucleico codifica de modo secuencial en una dirección de 5' a 3', (i) una secuencia señal, una región Fc de inmunoglobulina y una secuencia de proteína diana o (ii) una secuencia señal, una proteína diana, y una región Fc de inmunoglobulina. De ese modo la molécula de ácido nucleico resultante codifica una estructura Fc-X o X-Fc donde X e Y son una proteína diana. En una realización alternativa, un ácido nucleico codifica una proteína Fc-X o X-Fc sin una secuencia señal.

En otra realización preferente, un ácido nucleico de la invención codifica una proteína de fusión Ig con isotipos híbridos o mutantes. Específicamente, el ácido nucleico proporciona fracciones de anticuerpos con isotipos híbridos, o alternativamente con regiones bisagra alteradas. Por ejemplo, la proteína de fusión consiste de una IgG2, modificada para contener menos enlaces disulfuro en la región bisagra, o una región IgG2 CH2 y CH3 en la cual la región bisagra deriva a partir de otro anticuerpo, preferentemente una región bisagra IgG1 normal o mutante.

Un ácido nucleico de la invención se incorpora preferentemente en una asociación operativa dentro de un vector de expresión replicable que se introduce dentro de una célula hospedadora de mamífero competente para producir la proteína de fusión. La proteína de fusión resultante se produce eficientemente y se segrega a partir de la célula hospedadora de mamífero. La proteína de fusión segregada se recolecta con posterioridad a partir del medio de cultivo sin lisar la célula hospedadora de mamífero. El producto de proteína se analiza en busca de actividad y/o se purifica utilizando reactivos comunes según se desee, y/o se divide de compañero de fusión, todo ello mediante técnicas convencionales.

De este modo, la presente invención también proporciona métodos para la producción de proteínas de fusión con inmunogenicidad reducida.

Los métodos y composiciones de la invención son también útiles para proporcionar tratamiento terapéutico utilizando una proteína de fusión que se ha transformado en menos inmunogénica. Un objeto general de la invención es proporcionar procesos que son tanto eficientes como económicos así como proteínas que son menos inmunogénicas. Las composiciones terapéuticas preferentes de la invención incluyen una cantidad terapéuticamente efectiva de proteína de fusión desinmunizada. Preferentemente, la proteína de fusión desinmunizada se administra junto con un portador farmacéuticamente aceptable.

Los anteriores y otros aspectos, características y ventajas de la presente invención serán más evidentes a partir de la descripción detallada, los dibujos, y reivindicaciones a continuación.

Descripción detallada de la invención

Todas las proteínas, incluyendo los anticuerpos, que se administran a un paciente para uso terapéutico tienen el potencial de inducir una respuesta inmune en el receptor. Esta respuesta inmune se coordina mediante linfocitos T (células T) que después activan linfocitos B (células B) para producir anticuerpos. La producción de anticuerpos contra el agente terapéutico es perjudicial dado que conlleva una eliminación más rápida del agente terapéutico y puede inducir una respuesta alérgica.

La presente invención proporciona métodos para reducir la inmunogenicidad de proteínas de fusión. De acuerdo con un método de esta invención, los epítopos potenciales de células T se identifican en la región de unión de una unión de fusión en una proteína de fusión. Los epítopos de células T se identifican mediante una variedad de métodos informáticos y no informáticos, incluyendo la predicción basada en el modelado computacional estructural o mediante la síntesis de péptidos y pruebas de unión a moléculas MHC Clase II específicas o en un ensayo de inmunogenicidad.

Según la presente invención, una unión de fusión se define como la posición entre el último aminoácido (C-terminal) de una primera proteína o péptido y el primer aminoácido (N-terminal) de una segunda proteína o péptido en una proteína de fusión. En consecuencia, una unión de fusión incluye cualquiera de los aminoácidos entre el último aminoácido de una proteína y el primer aminoácido de una segunda proteína. En una realización, la unión de fusión incluye un ligador.

De acuerdo con la invención, una región de unión es la región de una proteína de fusión que rodea o abarca la unión de fusión entre dos proteínas. Una región de unión preferentemente incluye entre 1 y aproximadamente 100 aminoácidos, más preferentemente entre 1 y aproximadamente 50 aminoácidos, o entre 1 y aproximadamente 25 aminoácidos, e incluso más preferentemente entre 1 y aproximadamente 15 aminoácidos, o entre 1 y 9 aminoácidos. En una realización, una región de unión comprende un espaciador o ligador péptido insertado en el punto de unión entre dos proteínas. De acuerdo con la invención, una región de unión que incluye un espaciador o ligador péptido también puede desinmunizarse para minimizar la respuesta de un paciente a una proteína de fusión incluyendo el espaciador o ligador.

De acuerdo con la invención, el epítipo de células T de unión se define como una secuencia peptídica capaz de unir un MHC Clase II que contiene por lo menos un aminoácido derivado de cada una de por lo menos dos proteínas compañeras de fusión diferentes. Por ejemplo, Paul (Fundamental Immunology, Capítulo 8, Tabla 8, p. 276 [2000] 4th ed.) ilustra segmentos de 10 aminoácidos que pueden unirse a una molécula MHC Clase II. En un epítipo de células T de unión, estos 10 aminoácidos péptidos se derivan a partir de diferentes compañeros de fusión. De acuerdo con la invención un epítipo de células T candidato o potencial que abarca la unión de fusión (un epítipo de células T de unión candidato) preferentemente incluye de 1 a 8 aminoácidos a partir de cualquiera de los lados de la unión, y más preferentemente de 1 a 10 o de 1 a 11 aminoácidos a partir de cualquiera de los lados de la unión. Los epítopos candidatos tienen un largo de preferentemente 9, 11, ó 12 aminoácidos. En consecuencia, dado que un epítipo de células T de unión de la invención incluye por lo menos un aminoácido de cada lado de la unión, los epítopos de células T candidatos preferentes son epítopos de unión que incluyen 1-8 (ó 1-10, u 11) aminoácidos de

un lado de la unión y también incluyen un número complementario de aminoácidos del otro lado de la unión para ofrecer como resultado un epítipo que tiene 9-12 aminoácidos, y más preferentemente 9 aminoácidos.

De acuerdo con la presente invención, los residuos de anclaje dentro de un epítipo de células T de unión se mutan para prevenir la unión a una molécula MHC Clase II. En general, se tiene cuidado de no introducir epítopos potenciales de células T adicionales, y de preservar la función de cada compañero de fusión.

De acuerdo con la invención, una fusión de secuencias de tipo natural es una fusión en la cual las secuencias en los lados de los extremos N-terminal y C-terminal de la unión de fusión se derivan directamente a partir de secuencias naturales.

De acuerdo con la invención, una unión de fusión desinmunizada es una secuencia de unión en la cual se ha introducido una o más mutaciones de sustitución relativas a una unión de secuencias del tipo natural. En una realización más preferente, la desinmunización de una unión de fusión no involucra la introducción de un ligador, tal como un ligador Gli-Ser 'no inmunogénico', y la relación espacial entre los compañeros de fusión no se altera en una proteína de fusión desinmunizada. De acuerdo con la invención, se puede sustituir o cambiar uno o más aminoácidos en la región de unión en forma N-terminal a la unión de fusión o bien en forma C-terminal a la unión de fusión, o tanto en forma N-terminal como C-terminal a la unión de fusión.

De acuerdo con la invención, un epítipo potencial de células T es una secuencia que, cuando se considera como un péptido aislado, se predice para la unión a una molécula MHC Clase II o un equivalente en una especie no humana. Un epítipo potencial de células T se define sin considerar otros aspectos del proceso antigénico, tales como la eficiencia de la captación de proteínas dentro de las células que presentan antígenos, la eficiencia de la división en los sitios en una proteína intacta para producir un péptido que puede unirse a un MHC Clase II, etcétera. De este modo, el conjunto de epítopos de células T que se presentan realmente sobre un MHC Clase II tras la administración de una proteína a un animal es un subconjunto de los epítopos potenciales de células T.

De acuerdo con la invención, un epítipo de células T es un epítipo sobre una proteína que interactúa con una molécula MHC clase II. Sin intención de estar sujeto a una teoría, se entiende que un epítipo de células T es una secuencia de aminoácidos en una proteína o una proteína de fusión, que no se sometió al proceso de selección negativa de células T durante el desarrollo de células T y por lo tanto se espera que sea presentada mediante una molécula MHC Clase II y que se reconozca mediante un receptor de células T. En una realización preferente de la invención, los epítopos de células T no propias están presentes en la región de unión en la unión de fusión de dos proteínas que forman una proteína de fusión.

La invención proporciona métodos no informáticos para la reducción o eliminación del número de epítopos de células T en una unión de proteína de fusión sin requerir simulaciones computacionales elaboradas o estructuras de proteínas tridimensionales. En una realización, un método de la invención toma ventaja del hecho de que un segmento principal de nueve aminoácidos interactúa tanto con la molécula MHC clase II como con el receptor de células T durante la presentación del antígeno. El aminoácido más cercano al extremo N-terminal se denomina residuo de posición "de anclaje" que se une a un bolsillo profundo dentro de la molécula MHC clase II. Uno de los siguientes aminoácidos está típicamente presente en la posición de anclaje que es importante para la unión a una molécula MHC clase II: Leucina, Valina, Isoleucina, Metionina, Fenilalanina, Tirosina y Triptófano. De acuerdo con la invención, 2 a 3 aminoácidos adicionales adyacentes a los 9 aminoácidos principales también afectan la interacción con las moléculas MHC. Además, el aminoácido más cercano al extremo C-terminal en la primera proteína de la proteína de fusión generalmente puede sustituirse. Esto es útil especialmente cuando se sabe que el compañero de fusión del extremo N-terminal o la primera proteína está activa cuando se fusiona al compañero de fusión del extremo C-terminal o a la segunda proteína en el extremo C-terminal de la primera proteína.

Un método general de la invención incluye la mutación de cualquiera de las Leucinas, Valinas, Isoleucinas, Metioninas, Fenilalaninas, Tirosinas o Triptófanos que están presentes en los ocho aminoácidos más cercanos al extremo C-terminal de un compañero de fusión del extremo N-terminal en una proteína de fusión. En una realización, uno o más de estos aminoácidos en un epítipo de células T de unión candidato aminoácidos se muta de modo preferente a una Treonina, una Alanina o una Prolina. Esto retiene algo de la naturaleza hidrofóbica del aminoácido que se reemplaza. En realizaciones adicionales de la invención, uno o más de los aminoácidos antes mencionados se eliminan del epítipo de células T de unión potencial o candidato, o se reemplazan por un aminoácido apropiado análogo. De acuerdo con la invención, si un aminoácido se elimina para destruir un epítipo potencial de células T, se tiene cuidado de no producir un nuevo epítipo de células T que incluya aminoácidos cerca de la eliminación.

De acuerdo con la invención, es habitualmente útil fabricar una construcción de plásmido intermedio de expresión generalizada que comprenda la secuencia de codificación para un compañero de fusión del extremo N-terminal que contiene una mutación de uno o más residuos hidrofóbicos en los últimos ocho aminoácidos. Generalmente, tal plásmido tiene uno o más sitios de enzimas de restricción convenientes en o cerca del ADN que codifica el extremo C-terminal del compañero de fusión con extremo N-terminal.

El propósito de la construcción de un plásmido intermedio es fabricar plásmidos de expresión que codifiquen una proteína de fusión en la cual uno o más compañeros de fusión del extremo N-terminal tienen una o más sustituciones de una Leucina, Valina, Isoleucina, Metionina, Fenilalanina, Tirosina, o Triptófano a otro aminoácido en los ocho aminoácidos de C-terminal. La construcción de tales plásmidos de expresión final puede lograrse mediante una variedad de otros métodos que son bien conocidos en el arte, tales como la generación de fragmentos de PCR o síntesis de ácidos nucleicos, seguidos por la ligadura del fragmento dentro de un vector o un vínculo apropiado con otras secuencias a través de técnicas de PCR bien conocidas.

Las realizaciones preferentes específicas incluyen los plásmidos de fusión de Fc-X, los plásmidos de fusión de albúmina-X, los plásmidos de fusión de scFv-X, y los plásmidos de fusión de Fab-X. En el caso de Fc(gama)-X, es útil introducir mutaciones dentro de la secuencia de codificación para provocar sustituciones de aminoácidos del segmento Leucina-Serina-Leucina-Serina cerca del extremo C-terminal la región Fc de una molécula IgG1, IgG2, IgG3, o IgG4, como se diagrama aquí para la IgG1: Las secuencias de aminoácidos de las regiones Fc humanas derivadas a partir de IgG1, IgG2, IgG3 y IgG4 se ilustran en las SEQ ID Nos. (Números de Identificación de Secuencias): 1, 2, 3 y 4 respectivamente.

En un ejemplo, KSLSLSPGK (SEQ ID No.: 5) se cambia por KSATATPGK (SEQ ID No.: 6). Esta mutación se designa para eliminar los epítomos de células T de unión potenciales y también para eliminar un epítipo de células T en el cual la Fenilalanina o Tirosina ascendente sirve como un residuo de anclaje de posición 1.

Alternativamente, a veces es útil combinar mutaciones que eliminar los epítomos de células T de unión candidatos con una mutación que extiende la vida media en suero. Por ejemplo, mediante el cambio de KSLSLSPGK (SEQ ID No.: 5) por KSATATPGA (SEQ ID No.: 7).

Otras realizaciones incluyen sustituciones en el segmento LSLS por otros aminoácidos tales como Glicina o Prolina.

En el caso de los vectores de expresión utilizados para la realización de proteínas de fusión IgA, es útil eliminar algunos de los aminoácidos del extremo C-terminal, de modo que se elimina la cisteína cerca del extremo C-terminal que participa en la oligomerización de IgA. Por ejemplo, pueden eliminarse quince aminoácidos, tales como los extremos de la secuencia de cadena pesada de IgA con Prolina-Treonina-Histidina antes de ser fusionada a una segunda proteína. Además, es útil introducir los siguientes cambios cerca del extremo C-terminal del dominio CH3 de la región Fc de IgA:

QKTIDRLAGKPTH (SEQ ID No.: 8) se cambia por QKTADRTAGKPTH (SEQ ID No.: 9)

Las secuencias desinmunizadas adicionales en una proteína de fusión de IgA-X son:

QKTPTRTAGKPTH (SEQ ID No.: 10)

QKTPTRPAGKPTH (SEQ ID No.: 11)

QKTATRPAGKPTH (SEQ ID No.: 12).

En el caso de una fusión de albúmina-X, es útil introducir los siguientes cambios en un plásmido de expresión de albúmina-X de modo que el extremo C-terminal de albúmina se modifica de la siguiente forma:

KKLVAASQAALGL (SEQ ID No.: 13) se cambia por KKLVAASQAATTA (SEQ ID No.: 14).

De este modo, la presente invención proporciona secuencias de ácido nucleico y proteínas que son útiles en la construcción de proteínas de fusión menos inmunogénicas. Específicamente, la invención proporciona proteínas con mutaciones de cualquiera Leucinas, Valinas, Isoleucinas, Metioninas, Fenilalaninas, Tirosinas, o Triptófanos en los últimos ocho aminoácidos. Las proteínas son preferentemente proteínas humanas con secuencias que generalmente corresponden a secuencias encontradas en el cuerpo humano. La invención también proporciona secuencias de ácido nucleico que codifican tales proteínas. Las secuencias de ácido nucleico para este aspecto de la invención pueden existir como plásmidos, fragmentos generados por PCR, o ácidos nucleicos producidos por síntesis química.

La invención también proporciona plásmidos de expresión que codifican una proteína de fusión en la cual uno o más compañeros de fusión del extremo N-terminal tienen una o más mutaciones de una Leucina, Valina, Isoleucina, Metionina, Fenilalanina, Tirosina, o Triptófano a otro aminoácido en los ocho aminoácidos de C-terminal.

Por ejemplo, la invención proporciona los plásmidos que codifican una Fc-IL2 o la proteína de fusión del anticuerpo completo-IL2 en donde la región se muta como se describe con anterioridad. Además, la invención proporciona fusiones que comprenden una región Fc mutada como se describió con anterioridad a formas normales o mutadas de eritropoyetina, tales como las formas de eritropoyetina descritas en WO01/36489.

La invención también proporciona un método para la reducción de inmunogenicidad de una unión de proteína de fusión mediante la introducción de un sitio de glicosilación N-ligada u O-ligada cerca, o preferentemente,

en una unión de fusión. Por ejemplo, los aminoácidos Asparagina, Serina o Treonina, y un tercer residuo se introducen de la siguiente forma. Considere una secuencia en la cual las Xs representan los aminoácidos de un compañero de fusión del extremo N-terminal, y las Zs representan los aminoácidos de un compañero de fusión del extremo C-terminal.

5 $X_1X_2X_3X_4X_5X_6Z_1Z_2Z_3Z_4Z_5Z_6Z_7Z_8Z_9$

$X_1X_2X_3X_4X_5N\ G\ S\ Z_3Z_4Z_5Z_6Z_7Z_8Z_9$

De acuerdo con este método, no se bloquea necesariamente la unión de un péptido de unión mediante la introducción de un sitio de glicosilación. Sin embargo, cualquier péptido que se une en el surco del MHC Clase II y tiene la asparagina glicosilada del extremo C-terminal al residuo de anclaje más cercano al extremo N-terminal no funcionará como un epítipo de células T. La presencia de la fracción de la gran glicosilación impedirá de modo estérico el reconocimiento del complejo MHC Clase II/péptido. Un sitio de glicosilación preferente incluye la secuencia Asn-X-Ser o Asn-X-Tr en donde X es preferentemente Gli, pero puede ser cualquier aminoácido.

Además, la introducción de mutaciones que introducen residuos de Glicina y Serina no crea epítopos nuevos de células Ts. Ni la Glicina ni la Serina pueden actuar como un residuo de anclaje. Durante el procesamiento del antígeno, una proteína de fusión, en principio, se divide entre la Asparagina glicosilada y la Glicina o entre la Glicina y la Serina. En cualquiera de los dos casos, los péptidos resultantes tienen los residuos de Glicina y/o Serina mutante del extremo N-terminal a un residuo de anclaje, y de este modo los residuos de Glicina y/o Serina mutante no se reconocen mediante un receptor de células T, dado que los residuos del extremo N-terminal a un residuo de anclaje se encuentran fuera de la región reconocida por el TCR.

En una variación de este método, una región de unión de fusión ya contiene una Serina o Treonina precedida por residuos de aminoácidos tales como Glicina, Serina, Alanina, etc. El segundo método se utiliza preferentemente cuando una región de unión es flexible y se desplaza desde el núcleo hidrofóbico de cada compañero de fusión, de modo que la glicosilación N-ligada novedosa no interfiere con el plegamiento o la función de ninguno de los compañeros de fusión.

Es una cuestión evidente para los expertos en el arte de la ingeniería de proteínas determinar cuándo es factible la introducción de un sitio de glicosilación. Por ejemplo, se puede conocer la estructura tridimensional de cada compañero de fusión, u homólogos cercanos de los compañeros de fusión. Es usual el caso donde algunos aminoácidos en el extremo N-terminal o C-terminal de una proteína no se resuelven en una estructura de rayos X, o exhiben muchas conformaciones posibles en una estructura por RMN. En los casos donde tres o más aminoácidos se desordenan en cualquiera de los dos lados de un sitio de glicosilación, existe algo de confianza en que la proteína de fusión resultante se plegará correctamente y que ambos compañeros estarán activos. Algunas experimentaciones de rutina son necesarias para determinar si una construcción de proteína de fusión dada será funcional.

En las realizaciones preferentes de la invención, el compañero de la proteína de fusión tanto del extremo N-terminal como del C-terminal son proteínas humanas. Los epítopos potenciales de células T en tales proteínas de fusión se crean a partir de los 8 aminoácidos finales del compañero del extremo N-terminal (primera proteína) combinada con los primeros 8 aminoácidos del compañero del extremo C-terminal (segunda proteína). Esto proporciona una serie de 8 híbridos 9-mer creados a partir de las primeras y la segundas proteínas. Cualquier residuo alifático o aromático (Leucina, Valina, Isoleucina, Metionina, Fenilalanina, Triptófano o Tirosina) en los últimos 8 aminoácidos de la primera proteína presenta un alto riesgo de crear un péptido de unión del MHC con el aminoácido en la primera posición (posición de anclaje) que une el bolsillo de la molécula del MHC. Por lo tanto, la sustitución de cualquiera de los aminoácidos antes mencionados, con un aminoácido que no es uno de los aminoácidos antes mencionados, y preferentemente con Alanina, Prolina, o Treonina, eliminará un epítipo de células T candidato.

45 Por ejemplo, en el caso de una proteína de fusión Fc que contiene la secuencia:

HNHYTQKSLSLSPGKGGGGSGGGSGGGGS (SEQ ID No.: 15),

los residuos de leucina crean dos epítopos potenciales. Por lo tanto, la secuencia puede desinmunizarse como;

HNHYTQKSATATPGKGGGGSGGGSGGGGS (SEQ ID No.: 16),

50 al cambiar L por A y S por T. Estos cambios remueven los epítopos con la Leucina como el primer aminoácido en el bolsillo de unión del MHC y la Tirosina como el primer aminoácido en el bolsillo de unión del MHC, respectivamente.

Estas sustituciones para la desinmunización funcionan en humanos para todas las proteínas Fc, tanto con secuencias ligadoras como aquellas sin éstas, preferentemente cuando 1) ambas proteínas en la proteína de fusión

son proteínas humanas; 2) se ignoran los péptidos de unión del MHC en las secuencias naturales de ambas proteínas; y 3) se ignoran también las 9-mer idénticas a las secuencias originales.

Los métodos de la invención son generalmente aplicables en todos los organismos vertebrados, preferentemente en mamíferos y más preferentemente en humanos. La invención se ilustra además mediante los siguientes ejemplos no limitativos.

Ejemplos

Ejemplo 1: Deducción de los epítopos reactivos inmunogénicos de la inmunocitocina huKS-IL2.

La HuKS-IL2 consiste en las regiones humanizadas V_H y V_L combinadas con las regiones constantes de cadena H y L humana. La cadena H se fusionó en su terminal carboxilo a la secuencia madura de IL-2 humana como se describió con anterioridad. Esta cadena H es del isotipo γ 1 y tiene gran afinidad por los receptores Fc. Debido a esta gran afinidad la HuKS-IL2 se eliminó rápidamente de la circulación. Sin la intención de estar sujeto a una teoría, la eliminación de la HuKS-IL 2 tiene lugar probablemente mediante las células con FcR en el hígado (células de Kupffer) y el vaso (células presentadoras de antígenos).

Se estableció con anterioridad que ciertos pacientes habían tenido respuestas inmunes a alguna porción de la molécula de huKS-IL2; sin embargo, los epítopos reconocidos por estos anticuerpos no son conocidos. Para deducir los epítopos reactivos, se comparó las reactividades relativas de los sueros de paciente con huKS-IL2 con otras proteínas relacionadas:

(1) Hu14.18-IL2, una molécula con regiones V humanizadas completamente diferentes pero exactamente las mismas regiones C y unión de fusión con la IL-2;

(2) VH1, una forma desinmunizada de huKS-IL2 sin epítopos de células T en las regiones VH y VL, derivada a partir de regiones V de ratón con epítopos de células B de ratón expuestos en la superficie incrustados a residuos humanos.

(3) VH2, una forma desinmunizada de huKS-IL2 con un epítipo de células T restante en CDR3, derivado de regiones V de ratón con epítopos de células B de ratón expuestos en la superficie incrustados a residuos humanos, en el cual la VH contiene un epítipo de células T.

(4) 425-IL2 construida con regiones KOL o bien con regiones EU Cy1 (en lugar de KS) (para comparar la reactividad alotípica);

(5) huKS-mIL2 – una molécula con las regiones V huKS fusionadas a las regiones C de ratón e IL-2 de ratón;

(6) Fc-IL2 humana;

(7) sólo Fc humana;

(8) sólo IL-2 humana.

Las proteínas de fusión y los fragmentos de inmunoglobulina se purificaron mediante cromatografía en proteína A Sefarosa y se recubrieron sobre placas de 96 pozos en amortiguador de bicarbonato y después se bloqueó con suero de cabra al 1% con BSA al 1%. Las disoluciones de los sueros de paciente se incubaron y después se eliminó el material suelto mediante tres lavados con PBS-Tween. Los anticuerpos humanos unidos a partir de los sueros del paciente se detectaron con diferentes anticuerpos conjugados con HRP según la proteína unida. Generalmente, el conjugado HRP de cadena λ de cabra antihumano se usó porque la mayoría de las proteínas unidas en la placa consistieron en Fc humana y cadenas κ humanas.

Los sueros de ciertos pacientes mostraron una reactividad clara al huKS-IL2 que no fue detectable en los sueros pre-inyección a partir de los mismos pacientes. Se utilizaron antisueros preinmunes para establecer un control no inmunizado de partida. La reactividad vista en los sueros del paciente puede atribuirse a (1) la reactividad de la anti-IL2, (2) la reactividad de la anti Fc (alotípica), (3) la reactividad a la secuencia de unión novedosa o (4) a la reactividad con el idiotipo KS, o una combinación de reactividades.

Ninguno de los sueros del paciente reaccionaron significativamente con la IL-2 recombinante o a la región Fc (1 y 2 más arriba). Algunos pacientes mostraron reactividad anti-idiotípica a las regiones V del KS. Todos los sueros de paciente mostraron reactividad con la Fc-IL2. Tres de cuatro pacientes mostraron reactividad a la Fc-IL2. La presencia de reactividad contra la Fc-IL2 pero no contra la Fc o la IL2 sugiere que los anti-sueros de pacientes reconocieron la unión entre la Fc y la IL2.

Ejemplo 2: Modificación de los residuos de aminoácidos en la unión de una proteína de fusión de anticuerpo-citocina para reducir la inmunogenicidad mediante la eliminación de los motivos de unión del MHC Clase II

El análisis de reconocimiento de plegamiento del péptido identificó dos segmentos péptidos superpuestos con fuerte potencial de unión del MHC en la unión entre la porción Fc y IL2 de la inmunocitocina. El reconocimiento de plegamiento del péptido y la identificación de los epítomos potenciales de células T se realizó como se revela en Carr (WO00/34317). Los cambios de aminoácidos se introdujeron de modo tal que los epítomos de unión del MHC Clase II potenciales existentes se eliminaron, pero no se introdujeron nuevos epítomos del MHC Clase II potenciales.

La modificación de una secuencia de unión LSLSPGK-AP (SEQ ID No.: 17) por ATATPGA-AP (SEQ ID No.: 18) ("LSLS por ATAT"), donde el guión es la unión de la huKS-IL2 inmunocitocina, produjo secuencias de péptido derivado de unión incapaces de unirse a cualquier MHC Clase II humano con una afinidad suficientemente alta para dar como resultado inmunogenicidad.

Ejemplo 3: Modificación de residuos de aminoácidos en la unión de proteínas de fusión de inmunocitocina para reducir la inmunogenicidad

La modificación de una secuencia de unión LSLSPGK-AP (SEQ ID No.: 17) por LNLSPGA-AP (SEQ ID No.: 19) ("LSLS por LNLS"), donde el guión es la unión de la huKS-IL2 de inmunocitocina, da como resultado secuencias de péptido derivado de unión que son aún capaces de unirse a ciertas moléculas del MHC Clase II. Sin embargo, cuando la proteína KS-IL2 se expresa en las células de mamíferos y se segrega, la proteína es N-glicosilada cerca de la unión debido a la secuencia de NXS/T.

Los péptidos derivados de unión resultantes no son efectivos como epítomos de células T, porque cuando el MHC Clase II presenta los péptidos derivados de unión a las células T, la gran fracción N-glicosilación evita el acoplamiento (docking) específico entre un receptor de células T y el MHC Clase II.

Ejemplo 4: Caracterización de la reactividad inmune de las células presentadoras de antígeno a la huKS-IL2 inmunocitocina en comparación con una inmunocitocina huKS-IL2 desinmunizada.

La reducción de la inmunogenicidad debido a la modificación del epítipo reactivo mediante la mutación de LSLS a ATAT se prueba directamente como se muestra a continuación. Los péptidos sintéticos que imitan esta secuencia alteran la respuesta inmune de una célula presentadora de antígeno clásica tal como una célula dendrítica (DC). Los siguientes péptidos sintéticos

KSLSLSPGK-APTS (SEQ ID No.: 20) y

KSATATPGK-APTS (SEQ ID No.: 21),

donde el guión es la unión KS-IL2, se utilizan para estimular la presentación de antígeno mediada por DC a células T autólogas. La habilidad de aquellas células T para proliferar en respuesta a un desafío posterior con el antígeno péptido sirve como una medida de inmunogenicidad de ese péptido.

Específicamente, las células mononucleadas de sangre periférica (PBMC) se aíslan a partir de leucopaqetes mediante técnicas de gradiente de densidad estándares. Las células mononucleadas se resuspenden en un medio de cultivo Aim-V sin suero y se les permite adherirse. Después de 2 horas a 37°C se eliminan las células no adherentes. Las células adherentes se cultivan durante 7 días en un medio que contiene GM-CSF (50ng/ml) humano e IL-4 (20ng/ml) para derivar células dendríticas inmaduras (DC). Después de 7 días, las células se cosechan y se las caracteriza por el fenotípica mediante citometría de flujo con Abs marcadas con FITC apropiados para el MHC clase I, MHC clase II, CD80 y CD40 para confirmar el fenotipo DC inmaduro.

Las células no adherentes se cultivan con IL2 e IL 7 para obtener células efectoras autólogas (células T) para ser utilizadas en estudios funcionales posteriores. Para los estudios funcionales, se agregan células T a células dendríticas inmaduras (proporción 10:1) y se co-cultivan con huKS, huKS desinmunizada, secuencia 13 mer de unión peptídica (KSLSLSPGK-APTS) (SEQ ID No.: 20) y la secuencia peptídica 13 mer desinmunizada modificada (KSATATPGK-APTS) (SEQ ID No.: 21). La comparación del índice de proliferación, como se mide mediante la incorporación de timidina tritiada tras la exposición a cada una de las inmunocitocinas o los péptidos desinmunizados modificados e inmunogénicos, demuestra el grado de inmunogenicidad de cada molécula. A saber, un aumento en la incorporación radiactiva es aproximadamente proporcional a la habilidad de cada péptido de unirse a una molécula MHC clase II en DC y ser presentado a células T.

Ejemplo 5: Deducción de epítomos reactivos inmunogénicos encontrados en las proteínas de fusión de albúmina y la modificación de residuos aminoácidos en una unión de fusión para reducir la inmunogenicidad.

La albúmina de suero humana (HSA), debido a su notablemente larga vida media, su amplia distribución *in vivo* y su falta de funciones enzimáticas o inmunológicas, se ha utilizado como un portador para péptidos/proteínas terapéuticos. Se ha mostrado que un híbrido HSA-CD4 desarrollado genéticamente bloquea la entrada del virus de inmunodeficiencia humano dentro de las células CD4+ y al mismo tiempo muestra propiedades *in vitro* antivirales

similares a aquellas del CD4 soluble (Yeh et al, PNAS 89:1904-1908, 1992). De este modo, la fusión genética de los péptidos bioactivos a la HSA es útil para el diseño y la recuperación de derivados HSA terapéuticos segregados. Sin embargo, como con todas las proteínas de fusión, la HSA-CD4 tiene una unión novedosa que puede ser inmunogénica y contiene epítopos de células T capaces de ser presentados en moléculas MHC clase II. El análisis de una unión entre HSA y CD4 utilizando los métodos de los Ejemplos 1, 2, 3, y 4 identifica péptidos con potencial de unión del MHC. Las secuencias potencialmente inmunogénicas se modifican para disminuir o eliminar los epítopos de células T y B potenciales a fin de reducir la inmunogenicidad. De forma similar, un sitio de glicosilación novedoso puede introducirse dentro de la región de unión a fin de reducir la inmunogenicidad.

Secuencia de CD4 de secuencia de albúmina

10 TCFAEEGKKLVAASQAALGL - KKVVLGKKGDTVELTCTAS (SEQ ID No.: 22).

La invención contempla que la región de unión de la proteína de fusión de HSA-IFN α contiene tres epítopos de células T candidatos,

KKLVAASQAALGL (SEQ ID No.: 13);

KLVAASQAALGL (SEQ ID No.: 23); y

15 LGLCDLPQTHSLG (SEQ ID No.: 24).

Los epítopos de células T ilustrados en las SEQ ID Nos.:13 y 23 se superponen y pueden desinmunizarse mediante el cambio de LV (en negrita) por cualquiera excepto F, I, L, M, V, W e Y. De modo alternativo, la puntuación del reconocimiento de plegamiento del péptido puede reducirse significativamente al cambiar LG por TT. El epítipo de células T en la SEQ ID No.: 24 puede desinmunizarse mediante el cambio del segundo L (en negrita) por un A.

20 por un A.

Además, se contempla que en el caso de una fusión HSA-X, en donde X puede ser cualquier proteína, la desinmunización de la unión de fusión se logra mediante el cambio de la secuencia de aminoácido AALGL (SEQ ID No.: 25) por TATTA (SEQ ID No.: 26).

CFAEEGKKLVAASQTATTA (SEQ ID No.: 27).

25 Ejemplo 6: Proteínas de fusión X-Fc y la modificación de residuos de aminoácidos en la unión de fusión para reducir la inmunogenicidad.

En algunos casos es específicamente ventajoso desarrollar una proteína de fusión en la orientación X-Fc. Con estas construcciones, una proteína diana es una proteína de fusión del extremo N-terminal y sigue un fragmento Fc. Por ejemplo, el péptido similar al glucagón tipo 1 (GLP-1) requiere un extremo N-terminal libre para su actividad, por lo tanto una fusión GLP-1-Fc es útil.

30 por lo tanto una fusión GLP-1-Fc es útil.

Una proteína de fusión GLP-1-Fc se construye de acuerdo con técnicas estándar descritas en el arte. Esta proteína de fusión tiene el extremo C-terminal de la GLP-1 unida a la bisagra de la cadena pesada γ 1. Se utiliza la secuencia de bisagra γ 1 que contiene una *Cis* a una mutación Ser (residuo 5) que elimina el residuo de *Cis* que forma un enlace disulfuro con la cadena liviana en IgG1 (Lo et al., (1998) Protein Engineering 11:495-500). La secuencia Fc no mutante es

35 la secuencia Fc no mutante es

EPKSCDKTHTCPPCPAPELLG (SEQ ID No.: 28)

con la región bisagra subrayada, seguida por el comienzo de la secuencia del dominio CH2.

La unión de fusión entre GLP-1 (7-37) y la Fc mutante es:

HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG –

40 EPKSSDKTHTCPPCPAPELLG (SEQ ID No.: 29).

La unión de fusión entre GLP-1 (7-37) y la Fc normal es:

SYLEGQAAKEFIAWLVKGRG –

EPKSCDKTHTCPPCPAPELLG (SEQ ID No.: 30)

45 Se identifican tres epítopos potenciales mediante el reconocimiento de plegamiento del péptido en la unión de fusión de GLP-1-Fc.

KEFIAWLVKGRGE (SEQ ID No.: 31)

EFIAWLVKGRGEP (SEQ ID No.: 32)

AWLVKGRGEPKSS (SEQ ID No.: 33).

El análisis de las uniones de fusión entre la GLP-1(texto en negrita) y la Fc (texto simple), realizado como en los Ejemplos 1-3, identifica péptidos con potencial de unión del MHC. Después de la identificación de sitios potenciales mediante el análisis de reconocimiento de plegamiento de péptidos, las secuencias potencialmente inmunogénicas se modifican mediante la sustitución de aminoácidos para reducir o eliminar los epítomos de unión de células T y B potenciales y disminuir la inmunogenicidad.

Los epítomos potenciales de células T mencionados con anterioridad ilustrados en las SEQ ID Nos.: 31, 32 y 33 se desinmunizan mediante la realización de sustituciones de aminoácidos individuales. Por ejemplo, el péptido que se muestra en la SEQ ID No.: 31 se desinmuniza mediante el cambio de Lisina (que se muestra en negrita) por una Treonina y la Arginina (que se muestra en negrita) por una Treonina. El péptido que se muestra en la SEQ ID No.: 32 se desinmuniza mediante el reemplazo de la Isoleucina (que se muestra en negrita) con una Alanina o una Prolina y el péptido en la SEQ ID No.: 33 se desinmuniza mediante el reemplazo de la Leucina con una Alanina o una Prolina. La unión desinmunizada resultante es:

HAEGTFTSDVSSYLEGQAAKEFAAWAVTIGT - EPKSSDKTHTCPPCPAPELLG (SEQ ID No.: 34).

De acuerdo con un método ejemplar para la introducción de un sitio de glicosilación en una unión de fusión se introdujeron los siguientes cambios:

SYLEGQAAKEFIAWLVKGRN - **GSKSSDKTHTCPPCPAPELLG** (SEQ ID No.: 35).

Ejemplo 7: Deducción de epítomos reactivos inmunogénicos de Enbrel, una proteína de fusión de TNFR-Fc y la modificación de residuos aminoácidos en una unión de fusión para reducir la inmunogenicidad.

ENBREL o etanercept, una proteína de fusión de X-Fc aprobada por la FDA, es un inhibidor del factor de necrosis tumoral (TNF) utilizado para tratar la artritis reumatoide. ENBREL es una proteína de fusión dimérica que consiste en un dominio extracelular de unión al ligando del receptor TNF enlazado a una proteína Fc de IgG1 humana. La TNFR-Fc exhibe en forma competitiva la unión del TNF a su receptor e inactiva biológicamente la unión TNF, dando como resultado una reducción significativa en la actividad inflamatoria. Como se describió con anterioridad para la GLP-1-Fc, la TNFR-Fc tiene una unión novedosa que contiene epítomos potenciales de células T.

La unión entre una fusión directa de una porción del extremo C-terminal del TNF-R (texto en negrita) al extremo N-terminal de la bisagra g1 (texto simple con la región subrayada que representa la región bisagra) es **STSFLPMGSPSPAEGSTGD** - **EPKSCDKTHTCPPCPAPELLG** (SEQ ID No.: 36)

El análisis de una unión entre el TNF-R y la Fc, realizado como en los Ejemplos 1-4, identifica péptidos con potencial de unión del MHC. Tras la identificación de sitios potenciales mediante el análisis de reconocimiento del plegamiento de péptidos, las secuencias potencialmente inmunogénicas se modifican mediante la sustitución de aminoácidos para reducir o eliminar los epítomos de unión de células T y B potenciales y disminuir la inmunogenicidad.

De acuerdo con un método ejemplar para la introducción de un sitio de glicosilación en una unión de fusión se introdujeron los siguientes cambios:

STSFLPMGSPSPAEGSTGN - **GSKSCDKTHTCPPCPAPELLG** (SEQ ID No.: 37).

Ejemplo 8: Deducción de epítomos inmunogénicos reactivos para las proteínas de fusión Fc-X-Y tales como Fc-IL12-IL2 y la modificación de residuos de aminoácidos en la unión de fusión para reducir la inmunogenicidad.

Las proteínas de fusión de una orientación Fc-X-Y tal como la Fc-IL12-IL2 tienen múltiples uniones de fusión novedosas que son potencialmente inmunogénicas. Por ejemplo, la Fc-IL12 tiene una unión de fusión similar a otras proteínas de fusión de Fc-X o inmunocitocinas (Ejemplo 1) pero es novedosa debido a la utilización de la citocina TL12. La unión de fusión se analiza en busca de sitios de unión inmunogénica y se modifica en consecuencia. Segundo, existe una unión de fusión de X-Y comparable a esa descrita en el Ejemplo 5, con dos citocinas diferentes que constituyen una proteína de fusión. El análisis del reconocimiento de plegamiento de péptidos se utiliza para cada una de las uniones de fusión.

Análisis de las uniones:

(1) **MHEALHNHYTQKSLSLSPGK** - RNLVPATPDGPMFPCLLH SQ (SEQ ID No.: 38)

entre el extremo C-terminal de la Fc (texto en negrita) y el extremo N-terminal de la IL12p35 (texto simple), y

(2) **RAQDRYSSSWSEWASVPCS** - APTSSSTKKTQLQLEHLLLD (SEQ ID No.: 39)

entre el extremo C-terminal de IL12p40 (texto en negrita) y el extremo N-terminal de IL2 (texto simple) mediante reconocimiento de plegamiento del péptido se identifica péptidos con potencial de unión del MHC. Las secuencias potencialmente inmunogénicas se modifican para disminuir o eliminar los epítopos potenciales de células T.

5 Por ejemplo, en la secuencia (1) más arriba, se realizaron los siguientes cambios:

MHEALHNHYTQKSATATPGK - RNLPVATPDGPMFPC LHHSQ (SEQ ID No.: 40).

Estos cambios reducen o eliminan el potencial de unión del MHC Clase II de diversos epítopos de células T en una unión de la Fc y la subunidad p35 de la IL12.

10 En otro ejemplo, la secuencia (2) más arriba se modifica para introducir un sitio de glicosilación mediante la introducción de una Asparagina y una Glicina en las primeras dos porciones dentro de la IL-2. Esta estrategia utiliza la Treonina natural en la posición 3 de la IL-2 madura. Además, es importante no interrumpir la formación de un enlace disulfuro en la fracción p40, por lo tanto es útil separar el sitio de glicosilación mediante por lo menos uno o dos aminoácidos a partir de la Cisteína en p40.

RAQDRYSSSWSEWASVPCS - NGTSSSTKKTQLQLEHLLD (SEQ ID No.: 41).

15 En el caso de la fusión de IL12p40-IL2, la introducción de un sitio de glicosilación como se expuso con anterioridad crea los siguientes epítopos potenciales de células T.

SEWASVPCSNNGTS (SEQ ID No.: 42)

ASVPCSNNGTSSST (SEQ ID No.: 43)

20 Sin embargo, la glicosilación del epítipo de células T evita la unión del MHC Clase II lo cual da como resultado inmunogenicidad reducida.

Ejemplo 9: Modificación de los residuos de aminoácidos en la unión de fusión de inmunocitocinas e inmunofusinas preparadas con un isotipo híbrido para eliminar los epítopos de células T.

25 A menudo es útil producir un anticuerpo o una proteína de fusión en base a un anticuerpo con un isotipo híbrido, de modo que las características útiles de diferentes isotipos puedan combinarse dentro de una única molécula. Las proteínas de fusión con isotipos híbridos pueden modificarse de acuerdo con la invención para reducir la inmunogenicidad.

30 Una proteína de fusión de anticuerpo con los siguientes componentes se construye mediante técnicas de ADN recombinante estándares: una cadena liviana y una cadena pesada, las regiones V que reconocen un antígeno específico de tumor, donde la cadena liviana es una cadena liviana típica, y la cadena pesada comprende los dominios CH1, CH2, y CH3 a partir de IgG2 y una región bisagra de IgG1, con una citocina fusionada al extremo C-terminal de la cadena pesada que involucra una unión de fusión como se describió con anterioridad.

35 Esta proteína contiene uniones novedosas entre CH1g2 y g1 bisagra, y g1 bisagra y CH2g2. La identificación y modificación de los epítopos de células T en estas uniones se realiza de la siguiente forma. Para las inmunocitocinas y las proteínas de fusión de Fc-X preparadas con un isotipo IgGe o bien un isotipo IgG2h, estas modificaciones son idénticas a aquellas expuestas en los Ejemplos 1, 2, 3, y 8 más arriba. Para las inmunofusinas X-Fc IgG2h, la unión novedosa también es idéntica dado que el extremo N-terminal de la Fc se ubica dentro de la región bisagra de la proteína IgG2h que se ha modificado a un tipo IgG1. Sin embargo, existen dos uniones de fusión novedosas en que la bisagra IgG1 insertada dentro de una inmunoglobulina IgG2 crea dos uniones novedosas entre la IgG2 CH1 y la bisagra IgG1 y la bisagra IgG1 y la IgG2 CH2.

40 IgG2 CH1 – bisagra IgG1 IgG2 CH2-IgG2 CH3 – proteína diana.

De este modo, el análisis de las uniones

qytcnvdhkpnsntkvdktv - epkscgdkthtppcp (SEQ ID No.: 49)

entre el extremo C-terminal de la IgG2 CH1 (texto en negrita) y el extremo N-terminal de la bisagra IgG1 (texto simple), y

45 **epkscgdkthtppcp** – appvagpsvflfpkpkdtl (SEQ ID No.: 50)

entre el extremo C-terminal de la bisagra IgG1 (texto en negrita) y el extremo N-terminal de la IgG2 CH2 F (texto simple) mediante el reconocimiento de plegamiento del péptido se debería identificar péptidos con potencial de unión del MHC. Las secuencias potencialmente inmunogénicas se modifican para disminuir o eliminar los epítopos de células T y B potenciales a fin de reducir la inmunogenicidad.

50 Dos epítopos potenciales de células T en la unión de fusión de bisagra IgG2CH1-IgG1 son:

TKVDKTVEPKSCD (SEQ ID No.: 51) y KTVEPKSCDKTHT (SEQ ID No.: 52).

[0108] La unión de fusión de bisagra IgG2CH1-IgG1 se desinmuniza mediante el cambio de la V (en negrita) por una A, una T o una P. La secuencia de la unión de fusión modificada se ilustra en la SEQ ID No.: 53.

qytycnvdhkpsntkadt - epkscdkthtcppcp (SEQ ID No.: 53).

- 5 Como se indicó con anterioridad, la secuencia bisagra y1 en pdCs-huFcy1 puede contener una Cis a una mutación Scr (subrayada) que elimina el residuo de Cis que forma un enlace disulfuro con la cadena liviana en IgG1 (Lo et al., (1998) protein Engineering 11:495-500), creando de este modo dos uniones de fusión potencialmente inmunogénicas adicionales para el análisis y la modificación:

(3) qytycnvdbkpsntkvdv - epkscdkthtcppcp (SEQ ID No.: 54)

- 10 **(4)epkscdkthtcppcp - appvagsvflfppkpkdtl** (SEQ ID No.: 55).

Ejemplo 10: Generación de una proteína de fusión Fc-EPO utilizando componentes Fc de isotipo híbrido de IgG1 e IgG4.

- 15 Para producir una proteína de fusión de Fc-eritropoyetina, el siguiente plásmido de expresión se construyó utilizando técnicas de biología molecular estándar. Se utilizó un fragmento de ADN de XmaI-XhoI que contiene una forma de la secuencia de codificación de eritropoyetina humana con mutaciones que dan como resultado las sustituciones de aminoácidos His32Gli, Cis33Pro, Trp88Cis, y Pro90Ala, como se revela en WO01/36489. La secuencia de proteínas correspondientes se muestra en la SEQ ID No.: 56.

**APPRLICDSRVLLERYLLEAKEAENITTCGACGPSLNENITVPTKVNFYAWKRMEVGQQAQVEV
WQGLALLSEAVLRGQALLVNSSQPCGLQLHVDKAVSGRLSLTLLRALGAQKEAISPPDAAS
AAPLRITITADITFRKLFRVYSNFLRGRLKLYTGEACRTGDR**

- 20 Este fragmento de ADN de XmaI-XhoI se insertó dentro de un vector plásmido que codifica una región bisagra a partir de IgG1 y una región CH2 y CH3 a partir de la IgG2, excepto que existieran dos grupos de mutaciones que den como resultado sustituciones de aminoácidos en la región de la CH3 del extremo C-terminal, de modo que la secuencia en la unión de la CH3 del extremo C-terminal y la Epo del extremo N-terminal es de la siguiente manera:

...TQKSATATPGA-APPRLI(SEQ ID No.: 57)

- 25 El primer grupo de mutaciones, que cambia la secuencia KSLSLSPG (SEQ ID No.: 58) de la región IgG2 CH3 por KSATATPG (SEQ ID No.: 59), se revela en la Solicitud de Patente Estadounidense con el número de solicitud 60/280625. El efecto de la sustitución de Leu-Ser-Leu-Ser (posición 3 a posición 6 de la SEQ ID No.: 58) con Ala-Tr-Ala-Tr (posición 3 a posición 6 de la SEQ ID No.: 59) es eliminar los epítopos potenciales de células T no propias humanas que podrían aparecer porque la unión entre la Fc humana y la eritropoyetina humana contiene secuencias de péptidos no propias. El segundo grupo que consiste en la sustitución de aminoácidos individuales K al A en el aminoácido del extremo C-terminal de la región CH3 se revela en la Solicitud de Patente Estadounidense con el número de solicitud 09/780668.

- 35 El plásmido resultante se transfectó dentro de las células NS/0 y la proteína de fusión Fc-Epo se expresó y purificó de acuerdo con los procedimientos conocidos en el arte. Después de la purificación en base a la unión con la proteína A, la proteína huFcy2h-huEpo que contiene las sustituciones de IgG2 CH3 y eritropoyetina descritas con anterioridad se caracterizó mediante cromatografía por exclusión de tamaño y se descubrió que consiste en 97% monómero y 90% monómero en dos preparaciones independientes. Se descubrió que la proteína huFcy2hhuEpo que contiene las sustituciones de IgG2 CH3 y eritropoyetina descritas con anterioridad es aproximadamente tan activa, sobre una base molar, como la eritropoyetina humana en un ensayo en base a una célula que midió la habilidad de una proteína de eritropoyetina para estimular la división celular de TF-1. El ensayo se realizó como se describe en WO01/36489.

- 40 Además, se caracterizaron las fusiones de eritropoyetina humana no mutante al extremo C-terminal de una región Fc que consiste en IgG1(bisagra-CH2-CH3), IgG2(bisagra-CH2-CH3), o bien IgG1(bisagra)-IgG2(CH2-CH3). Los plásmidos de expresión que comprenden secuencias Fc humanas no mutantes y secuencias de eritropoyetina no mutante se construyeron de forma análoga a los plásmidos descritos con anterioridad. Las células NS/0 se transfectaron con los plásmidos de expresión Fcy1-Epo, Fcy2-Epo y Fcy2h-Epo, y los clones estables se aislaron después de la evaluación de un número aproximadamente igual de clones para cada plásmido. Los clones de mejor producción produjeron 50µg/ml para Fcy1-Epo, 20µg/ml para Fcy2-Epo, y 120µg/ml para Fcy2h-Epo.

- 50 El siguiente ejemplo describe en detalle un método preferente para la identificación de las regiones de secuencia inmunogénica (epítopos de células T) dentro de las secuencias de las proteínas de fusión como se divulgó

en esta invención. Sin embargo, se debería señalar que dichas moléculas pueden obtenerse mediante otros métodos conocidos.

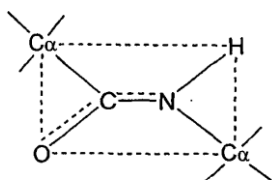
Ejemplo 11. Identificación de los epítomos de células T mediante métodos informáticos

De acuerdo con la invención, los epítomos en una región de unión de una proteína de fusión pueden modificarse utilizando métodos para la introducción de mutaciones dentro de las proteínas para modular su interacción con el sistema inmune. De acuerdo con la invención, los métodos conocidos en el arte que pueden adaptarse de acuerdo con la invención incluyen aquellos descritos en el arte previo (WO 92/10755 y WO 96/40792 (Novo Nordisk), EP 0519 596 (Merck & Co.), EP 0699 755 (Centro de Inmunología Molecular), WO 98/52976 y WO 98/59244 (Biovation Ltd.) o métodos relacionados.

Sin embargo, es posible obtener proteínas mutantes ventajosas si la identificación de dichos epítomos se realiza mediante el siguiente método nuevo que se describe junto en este documento en detalle y se aplica a la región de unión de las proteínas de fusión de acuerdo con la invención.

Existen varios factores que juegan un papel importante en la determinación de la estructura total de una proteína, un polipéptido o una inmunoglobulina. Primero, el enlace peptídico, es decir, el enlace que une los aminoácidos en la cadena, es un enlace covalente. Este enlace es planar en las estructuras, esencialmente una amida sustituida. Una "amida" es cualquiera de un grupo de compuestos orgánicos que contienen el grupo -CONH-.

El enlace peptídico planar que enlaza el C α de los aminoácidos adyacentes puede representarse como se ilustra a continuación:



Dado que los átomos O=C y los C-N yacen en un plano relativamente rígido, la rotación libre no tiene lugar alrededor de estos ejes. Por lo tanto, hace referencia a veces a un plano esquemáticamente ilustrado mediante la línea interrumpida como un plano "amida" o un "plano peptídico" en donde yacen los átomos de oxígeno (O), carbono (C), nitrógeno (N), e hidrógeno de la cadena principal del péptido. En los ángulos opuestos de este plano amida se ubican los átomos de C α . Dado que no existe sustancialmente rotación alrededor de los átomos O=C y C-N en el plano peptídico o plano amida, una cadena polipeptídica comprende en consecuencia una serie de enlaces peptídicos planares que unen los átomos de C α .

Un segundo factor que juega un papel importante en la definición de la estructura total o la conformación de un polipéptido o una proteína es el ángulo de rotación de cada plano amida alrededor del enlace C α común. Los términos "ángulo de rotación" y "ángulo de torsión" de aquí en más se consideran términos equivalentes. Suponiendo que los átomos O, C, N, y H permanecen en el plano amida (que es habitualmente una suposición válida, aunque puede haber algunas ligeras desviaciones a partir de la planaridad de estos átomos para algunas conformaciones), estos ángulos de rotación definen la conformación de la cadena principal de los polipéptidos N y R, es decir, la estructura como existe entre los residuos adyacentes. Estos dos ángulos se conocen como ϕ y ψ . Un grupo de los ángulos ϕ_1 , ψ_1 , donde el subíndice i representa un residuo particular de una cadena polipeptídica, de modo que efectivamente define la estructura secundaria del polipéptido. Las convenciones utilizadas para definir los ángulos ϕ , ψ , es decir, los puntos de referencia en los cuales los planos amida forman un ángulo de cero grado, y la definición de cual ángulo es ϕ , y cual ángulo es ψ , para un polipéptido dado, se definen en la literatura. Véase, por ejemplo "Ramachandran et al. Adv. Prot. Chem. 23:283-437 (1968), en las páginas 285-94.

El presente método puede aplicarse a cualquier proteína, y se basa en parte en el descubrimiento de que en los humanos la posición de anclaje 1 del Bolsillo primario del surco de unión de la molécula del MHC Clase II tiene una especificidad bien diseñada para cadenas laterales de aminoácidos particulares. La especificidad de este bolsillo se determina mediante la identidad del aminoácido en la posición 86 de la cadena beta de la molécula del MHC Clase II. Este sitio se ubica en el fondo del Bolsillo 1 y determina el tamaño de la cadena lateral que puede albergarse en este bolsillo. Marshall, K.W., J. Immunol., 152:4946-4956 (1994). Si este residuo es una glicina, todos los aminoácidos aromáticos y alifáticos hidrofóbicos (donde son alifáticos hidrofóbicos: valina, leucina, isoleucina, metionina y son aromáticos: fenilalanina, tirosina y triptófano) pueden albergarse en el bolsillo, siendo la preferencia para las cadenas laterales aromáticas. Si este residuo de bolsillo es una valina, la cadena lateral de este aminoácido se proyecta dentro del bolsillo y restringe el tamaño de las cadenas laterales peptídicas que pueden albergarse de modo que sólo pueden albergarse las cadenas laterales alifáticas hidrofóbicas. Por lo tanto, en una secuencia de residuo de aminoácido, cada vez que se descubre un aminoácido con una cadena lateral aromática o alifática hidrofóbica, existe el potencial para que un epítomo de células T del MHC Clase II restringido esté presente. Si la

cadena lateral es alifática hidrofóbica, sin embargo, es aproximadamente dos veces más probable que esté asociada con un epítipo de células T que una cadena lateral aromática (suponiendo una distribución aproximadamente pareja de tipos de Bolsillo 1 a lo largo de la población global).

Un método informático que incorpora la presente invención perfila la probabilidad de que las regiones peptídicas contengan epítipos de células T de la siguiente forma: (1) Se escanea la secuencia primaria de un segmento peptídico de un largo predeterminado, y se identifican todas las cadenas laterales aromáticas y alifáticas hidrofóbicas presentes. (2) Se asigna a las cadenas laterales alifáticas hidrofóbicas un valor mayor que el asignado a las cadenas laterales aromáticas; preferentemente aproximadamente el doble del valor asignado a las cadenas laterales aromáticas, por ejemplo, un valor de 2 para una cadena lateral alifática hidrofóbica y un valor de 1 para una cadena lateral aromática. (3) Se suman los valores determinados que han de estar presentes para cada segmento (ventana) de residuo de aminoácido superpuesto de largo uniforme predeterminado dentro del péptido, y el valor total para un segmento (ventana) particular se asigna al residuo de un aminoácido individual en una posición intermedia del segmento (ventana), preferentemente a un residuo en aproximadamente el punto medio del segmento (ventana) de muestra. Este procedimiento se repite para cada segmento (ventana) de residuo de aminoácido superpuesto de muestra. De este modo, se asigna a cada residuo de aminoácido del péptido un valor que se relaciona con la probabilidad de que un epítipo de células T esté presente en ese segmento (ventana) en particular. (4) Se pueden graficar los valores calculados y asignados como se describe en el Paso 3, con anterioridad, contra las coordenadas de aminoácidos de la secuencia de residuo de aminoácidos completa que se está analizando. (5) Se considera probable que todas las porciones de la secuencia que tienen una puntuación de un valor predeterminado, por ejemplo, un valor de 1, contienen un epítipo de células T y pueden modificarse, si se desea.

Este aspecto particular de la presente invención proporciona un método general mediante el cual pueden describirse las regiones de péptidos que tienen probabilidad de contener epítipos de células T. Las modificaciones a los péptidos en estas regiones tienen el potencial de modificar las características de unión del MHC Clase II.

De acuerdo con otro aspecto de la presente invención, los epítipos de células T se pueden predecir con mayor precisión mediante el uso de un método informático más sofisticado que toma en cuenta las interacciones de los péptidos con los modelos de alelos del MHC Clase II.

La predicción informática de los epítipos de células T presentes dentro de un péptido de acuerdo con este aspecto particular contempla la construcción de modelos de por lo menos 42 alelos del MHC Clase II en base a las estructuras de todas las moléculas del MHC Clase II y un método para el uso de estos modelos en la identificación informática de los epítipos de células T, la construcción de bibliotecas de cadenas principales de péptidos para cada modelo a fin de permitir la variabilidad conocida en las posiciones relativas de los carbonos alfa (C α) en la cadena principal del péptido, la construcción de bibliotecas de conformaciones de cadenas laterales de aminoácidos para cada acoplamiento de cadena principal con cada modelo para cada uno de los 20 aminoácidos alternativos en posiciones críticas para la interacción entre el péptido y la molécula del MHC Clase II, y el uso de estas bibliotecas de cadenas principales y conformaciones de cadenas laterales conjuntamente con una función de puntuación para seleccionar la cadena principal óptima y la conformación de cadena lateral para un péptido en particular acoplado con una molécula del MHC Clase II en particular y la derivación de una puntuación de unión a partir de esta interacción.

Los modelos de las moléculas del MHC Clase II pueden derivarse mediante el modelado por homología a partir de varias estructuras similares encontradas en el Banco de Datos de Proteínas Brookhaven ("PDB"). Éstos pueden realizarse mediante el uso de software de modelado por homología semi-automático (Modeller, Sali A. & Blundell T.L., 1993. J. Mol Biol 234:779-815) que incorpora una función de hibridación simulada, conjuntamente con el campo de fuerza CHARMM para la minimización de energía (disponible en Molecular Simulations Inc., San Diego, Ca.). También pueden utilizar métodos de modelado alternativos.

El presente método difiere significativamente de otros métodos informáticos que utilizan bibliotecas de datos de unión derivados experimentalmente de cada aminoácido alternativo en cada posición en el surco de unión para un pequeño grupo de moléculas del MHC Clase II (Marshall, K.W., et al., Biomed. Pept. Proteins Nucleic Acids, 1 (3):157-162) (1995) o aún otros métodos informáticos que utilizan datos de unión experimental similares a fin de definir las características de unión de tipos particulares de bolsillos de unión dentro del surco, de nuevo utilizando un subgrupo relativamente pequeño de moléculas del MHC Clase II, y después 'mezclando y haciendo coincidir' tipos de bolsillos a partir de esta biblioteca de bolsillos para crear artificialmente moléculas 'virtuales' adicionales del MHC Clase II (Sturniolo T., et al., Nat. Biotech, 17(6): 555-561 (1999)). Los dos métodos anteriores sufren la principal desventaja de que, debido a la complejidad de los ensayos y la necesidad de sintetizar grandes números de variantes peptídicas, sólo un número pequeño de moléculas del MHC Clase II se pueden escanear experimentalmente. Por lo tanto el primer método sólo puede realizar predicciones para un número pequeño de moléculas del MHC Clase II. El segundo método anterior también asume que un bolsillo revestido con aminoácidos similares en una molécula tendrá las mismas características de unión en el contexto de un alelo de Clase II diferente y sufre desventajas adicionales en que sólo aquellas moléculas del MHC Clase II se pueden crear 'virtualmente' las cuales contienen bolsillos contenidos dentro de la biblioteca de bolsillos. Utilizando la aproximación de modelado descrito en este documento, se puede deducir la estructura de cualquier número y tipo de moléculas del MHC Clase II, por lo tanto los alelos se pueden seleccionar específicamente para ser representativos de la población global.

Además, el número de moléculas del MHC Clase II escaneadas puede aumentarse mediante la realización de modelos adicionales más allá de tener que generar datos adicionales mediante experimentación compleja.

El uso de una biblioteca de cadenas principales permite la variación en las posiciones de los átomos de C α de los diferentes péptidos que se están escaneando cuando están acoplados con moléculas particulares del MHC Clase II. Esto contrasta nuevamente con los métodos informáticos alternativos previos descritos con anterioridad que confían en el uso de cadenas principales de péptidos simplificadas para el escaneo de la unión de aminoácidos en bolsillos particulares. No es probable que estas cadenas principales simplificadas sean representativas de conformaciones de cadenas principales que se han encontrado en péptidos 'reales' lo que lleva a imprecisiones en la predicción de uniones peptídicas. La presente biblioteca de cadenas principales se crea mediante la superposición de las cadenas principales de todos los péptidos unidos a moléculas del MHC Clase II encontrados dentro del Banco de Datos de Proteínas y observando la desviación media cuadrática (RMS) entre los átomos de C α de cada uno de los once aminoácidos ubicados dentro del surco de unión. Mientras que esta biblioteca puede derivarse a partir de un pequeño número de estructuras humanas y estructuras de ratón disponibles adecuadas (actualmente 13), a fin de permitir la posibilidad de incluso mayor variabilidad, la cifra de RMS para cada posición de C"- α se aumenta un 50%. Se determina la posición C α promedio de cada aminoácido y se dibuja una esfera alrededor de este punto cuyo radio es igual a la desviación de RMS en esa posición más un 50%. Esta esfera representa todas las posiciones C α permitidas.

Trabajando a partir del C α con la menor desviación de RMS (la del aminoácido en el Bolsillo 1 como se mencionó con anterioridad, equivalente a la Posición 2 de los residuos 11 en el surco de unión), se retícula tridimensionalmente la esfera, y cada vértice dentro de la retícula se utiliza como una ubicación posible para un C α de ese aminoácido. El plano amida subsiguiente, que corresponde al enlace peptídico con el subsiguiente aminoácido se injerta sobre cada uno de estos Cas y los ángulos ϕ y ψ se rotan gradualmente a intervalos determinados a fin de posicionar el subsiguiente C α . Si el subsiguiente C α cae dentro de la "esfera de posiciones permitidas" para este C α se acepta la orientación del dipéptido, mientras que si cae fuera de la esfera se rechaza el dipéptido. Este proceso se repite para cada una de las subsiguientes posiciones de C α , de modo que el péptido crece a partir de la "semilla" C α del Bolsillo 1, hasta que los subsiguientes nueve Cas se han posicionado a partir de todas las permutaciones posibles de los Cas precedentes. El proceso se repite una vez más para el bolsillo 1 anterior de C α único para crear una biblioteca de cadenas principales de posiciones C α ubicadas dentro del surco de unión.

El número de cadenas principales generadas depende de varios factores: El tamaño de las "esferas de posiciones permitidas"; la finura del reticulado de la "primera esfera" en la posición de Bolsillo 1; la finura de la rotación gradual de los ángulos ϕ y ψ utilizada para posicionar los posteriores Cas. Utilizando este proceso, se puede crear una extensa biblioteca de cadenas principales. Cuanto más extensa sea la biblioteca de cadenas principales, mayores serán las probabilidades de que se encuentre el encaje óptimo para un péptido determinado dentro del surco de unión de una molécula del MHC Clase II. Dado que no todas las cadenas principales serán adecuadas para acoplarse con todos los modelos de moléculas del MHC Clase II debido a los conflictos con aminoácidos de los dominios de unión, se crea para cada alelo un subgrupo de la biblioteca que comprende las cadenas principales que pueden ser albergadas por esos alelos. El uso de la biblioteca de cadenas principales, conjuntamente con los modelos de moléculas del MHC Clase II crea una base de datos exhaustiva que consiste en las conformaciones de cadenas laterales permitidas para cada aminoácido en cada posición del surco de unión para cada molécula del MHC Clase II acoplada con cada cadena principal permitida. Este grupo de datos se genera utilizando una función de solapamiento estérico simple donde una molécula del MHC Clase II se acopla con una cadena principal y una cadena lateral de aminoácido se injerta sobre la cadena principal en la posición que se desea. Cada uno de estos enlaces rotables de la cadena lateral se rotan gradualmente a intervalos determinados y las posiciones resultantes de los átomos dependen de ese enlace destacado. La interacción del átomo con los átomos de las cadenas laterales del surco de unión se destaca y las posiciones se aceptan o bien se rechazan de acuerdo con los siguientes criterios: La suma total del solapamiento de todos los átomos posicionados hasta el momento no debe exceder un valor predeterminado. De modo que el rigor de la búsqueda conformacional es una función del intervalo utilizado en la rotación gradual del enlace y el límite predeterminado para el solapamiento total. Este último valor puede ser pequeño si se sabe que un bolsillo en particular es rígido, sin embargo el rigor puede reducirse si se sabe que las posiciones de las cadenas laterales de bolsillo son relativamente flexibles. De modo que se pueden realizar concesiones para imitar las variaciones en flexibilidad dentro de los bolsillos del surco de unión. Esta búsqueda conformacional se repite para todos los aminoácidos en todas las posiciones de cada cadena principal cuando está acoplada con cada una de las moléculas del MHC Clase II para crear la base de datos exhaustiva de conformaciones de cadenas laterales.

Se utiliza una expresión matemática adecuada para estimar la energía de unión entre los modelos de moléculas del MHC Clase II conjuntamente con las conformaciones peptídicas ligandos que tienen que ser derivadas empíricamente mediante el escaneo de la extensa base de datos de las conformaciones de cadenas principales/cadenas laterales descritas con anterioridad. Por lo tanto, se escanea una proteína para encontrar epítomos potenciales de células T sometiendo a cada péptido posible con un largo que varía entre 9 y 20 aminoácidos (aunque el largo se mantiene constante para cada escaneo) a los siguientes cálculos: Una molécula del MHC Clase II se selecciona junto con una cadena principal del péptido permitida para esa molécula y se injertan sobre ella las cadenas laterales que corresponden a la secuencia peptídica deseada. Los datos de la identidad del

átomo y la distancia interatómica relacionados con una cadena lateral determinada en una posición particular sobre la cadena principal se recogen para cada conformación permitida de ese aminoácido (obtenida a partir de la base de datos descrita con anterioridad). Esto se repite para cada cadena lateral junto con las puntuaciones de la cadena principal y el péptido derivadas utilizando una función de puntuación. La mejor puntuación para esa cadena principal se retiene y el proceso se repite para cada cadena principal permitida para el modelo seleccionado. Las puntuaciones a partir de todas las cadenas principales permitidas se comparan y la puntuación más alta se considera la puntuación peptídica para el péptido deseado en ese modelo del MHC Clase II. Este proceso se repite para cada modelo con todos los péptidos posibles derivados a partir de la proteína que se está escaneando, y se muestran las calificaciones de los péptidos contra los modelos.

En el contexto de la presente invención, cada ligando presentado para el cálculo de afinidad de unión es un segmento de aminoácido seleccionado entre un péptido o proteína como se expuso con anterioridad. De modo que, el ligando es un tramo de aminoácidos con un largo de aproximadamente 9 a 20 aminoácidos derivados a partir de un péptido, un polipéptido o una proteína de secuencia conocida. Los términos "aminoácidos" y "residuos" de aquí en adelante se consideran términos equivalentes. El ligando, en la forma de los aminoácidos consecutivos del péptido que se va a examinar injertado sobre una cadena principal de la biblioteca de cadenas principales, se posiciona en la hendidura de unión de una molécula del MHC Clase II a partir de una biblioteca modelo de moléculas del MHC Clase II mediante las coordenadas de los átomos de C^α de la cadena principal peptídica y se selecciona una conformación permitida para cada cadena lateral a partir de la base de datos de conformaciones permitidas. Las distancias interatómicas y las identidades de los átomos relevantes se obtienen también a partir de esta base de datos y se utilizan para calcular la puntuación de unión peptídica. Los ligandos con una alta afinidad de unión para el bolsillo de unión del MHC Clase II se marcan como candidatos para la mutagénesis sitio-dirigida. Las sustituciones de aminoácidos se realizan en el ligando marcado (y por lo tanto en la proteína de interés) que después se vuelve a probar utilizando la función de puntuación a fin de determinar los cambios que reducen la afinidad de unión por debajo de un valor umbral predeterminado. Estos cambios pueden incorporarse dentro de la proteína de interés para eliminar los epítomos de células T.

La unión entre el ligando peptídico y el surco de unión de la molécula del MHC Clase II involucra interacciones no covalentes incluyendo, pero sin limitarse a: enlaces de hidrógeno, interacciones electrostáticas, interacciones hidrofóbicas (lipofílicas) e interacciones de Van der Waals. Éstas se incluyen en la función de puntuación peptídica como se describe en detalle más abajo. Se debe entender que un enlace de hidrógeno es un enlace no covalente que se puede formar entre grupos cargados o polares y consiste en un átomo de hidrógeno compartido por otros dos átomos. El hidrógeno del donador de hidrógeno tiene una carga positiva donde el aceptor del hidrógeno tiene una carga negativa parcial. A los efectos de las interacciones de proteínas/peptídicas, los donadores de enlaces de hidrógeno pueden ser nitrógenos con hidrógeno unido o bien hidrógenos unidos a oxígeno o nitrógeno. Los átomos aceptores del enlace de hidrógeno pueden ser oxígenos no unidos a hidrógeno, nitrógenos sin hidrógenos unidos y una o más conexiones, o azufres con sólo una conexión. Ciertos átomos, tales como los oxígenos unidos a hidrógenos o nitrógenos imina (por ejemplo C=NH) pueden ser tanto aceptores como donadores de hidrógeno. Las energías de enlace del hidrógeno oscilan entre 3 y 7Kcal./mol y son mucho más fuertes que los enlaces de Van der Waals, pero más débiles que los enlaces covalentes. Los enlaces de hidrógeno también son altamente direccionales y tienen su mayor fuerza cuando el átomo donador, el átomo de hidrógeno y el átomo aceptor son colineales. Los enlaces electrostáticos se forman entre los pares de iones de carga opuesta y la fuerza de la interacción es inversamente proporcional al cuadrado de la distancia entre los átomos de acuerdo con la ley de Coulomb. La distancia óptima entre pares de iones es de aproximadamente 2,8Å. En las interacciones de proteínas/péptidos, los enlaces electrostáticos pueden formarse entre arginina, histidina o lisina y aspartato o glutamato. La fuerza del enlace dependerá del pKa del grupo ionizante y la constante dieléctrica del medio aunque ellos sean aproximadamente similares en fuerza a los enlaces de hidrógeno.

Las interacciones lipofílicas son contactos hidrofóbico-hidrofóbico favorables que tienen lugar entre la proteína y el ligando peptídico. Habitualmente, éstas tendrán lugar entre las cadenas laterales de aminoácidos hidrofóbicas del péptido enterrado dentro de los bolsillos del surco de unión de modo que no están expuestas al solvente. La exposición de los residuos hidrofóbicos al solvente es altamente desfavorable dado que se fuerza a las moléculas de solvente circundantes al enlace de hidrógeno entre sí formando estructuras de clatrato tipo jaula. La disminución en entropía resultante es altamente desfavorable. Los átomos lipofílicos pueden ser azufres que no son ni aceptores de hidrógeno ni polares y átomos de carbono que no son polares.

Los enlaces de Van der Waals son fuerzas no específicas encontradas entre los átomos que tienen una separación de 3-4Å. Son más débiles y menos específicos que los enlaces de hidrógeno y los electrostáticos. La distribución de la carga electrónica alrededor de un átomo cambia con el tiempo y, en un instante, la distribución de la carga no es simétrica. Esta asimetría momentánea en la carga electrónica induce una asimetría similar en los átomos vecinos. Las fuerzas de atracción resultantes entre los átomos alcanzan un máximo en la distancia de contacto de Van der Waals pero disminuye muy rápidamente en aproximadamente 1Å a aproximadamente 2Å. Por otro lado, dado que los átomos se separan por menos de la distancia de contacto, las fuerzas de repulsión cada vez más fuertes se vuelven dominantes a medida que se superponen las nubes de electrones exteriores de los átomos. Aunque las fuerzas de atracción son relativamente débiles comparadas con los enlaces de hidrógeno y electrostáticos (aproximadamente 0,6Kcal./mol), las fuerzas de repulsión en particular pueden ser muy importantes para determinar si un ligando peptídico puede unirse exitosamente a una proteína.

En una realización, la función de puntuación de Böhm (aproximación SCORE1) se utiliza para estimar la constante de unión. (Böhm, H.J., J. Comput Aided Mol. Des., 8(3):243-256 (1994)). En otra realización, la función de puntuación (aproximación SCORE2) se utiliza para estimar las afinidades de unión como un indicador de un ligando que contiene un epítipo de células T (Böhm, H.J., J. Comput Aided Mol. Des., 12(4):309-323 (1998)). Sin embargo, las funciones de puntuación de Böhm como se describen en las referencias anteriores se utilizan para estimar la afinidad de unión de un ligando a una proteína donde ya se sabe que el ligando se une exitosamente a una proteína y se ha resuelto la estructura del complejo proteína/ligando, con la estructura resuelta presente en el Banco de Datos de Proteínas ("PDB", por sus siglas en inglés). Por lo tanto, la función de puntuación se ha desarrollado con el beneficio de los datos de unión positivos conocidos. A fin de permitir la discriminación entre ligandos positivos y negativos, se debe agregar un término de repulsión a la ecuación. Además, un estimado más satisfactorio de la energía de unión se logra mediante el cálculo de las interacciones lipofílicas por pares en lugar de utilizar el término de energía en base al área de las funciones de Böhm anteriores. Por lo tanto, en una realización preferente, la energía de unión se estima utilizando una función de puntuación de Böhm modificada. En la función de puntuación de Böhm modificada, la energía de unión entre proteína y ligando (ΔG_{bind}) se estima considerando los siguientes parámetros: La reducción de la energía de unión debido a la pérdida general de entropía rotacional y traslacional del ligando (ΔG_0); las contribuciones a partir de enlaces de hidrógeno ideales (ΔG_{hb}) donde por lo menos un compañero es neutral; las contribuciones a partir de interacciones iónicas no perturbadas (ΔG_{ionic}); las interacciones lipofílicas entre átomos ligandos lipofílicos y átomos aceptores lipofílicos (ΔG_{lipo}); la pérdida de energía de unión debido a la congelación de los grados internos de libertad en el ligando, es decir, se reduce la libertad de rotación alrededor de cada enlace C-C (ΔG_{rot}); la energía de la interacción entre la proteína y el ligando (E_{vdw}). La consideración de estos términos da la ecuación 1:

$$(\Delta G_{\text{bind}}) = (\Delta G_0) + (\Delta G_{\text{hb}} \times N_{\text{hb}}) + (\Delta G_{\text{ionic}} \times N_{\text{ionic}}) + (\Delta G_{\text{lipo}} \times N_{\text{lipo}}) + (\Delta G_{\text{rot}} \times N_{\text{rot}}) + (E_{\text{vdw}}).$$

Donde N es el número de interacciones calificadas para un término específico y, en una realización, ΔG_0 , ΔG_{hb} , ΔG_{ionic} , ΔG_{lipo} y ΔG_{rot} son constantes a las que se les dan los valores: 5,4, -4,7, -4,7, -0,17, y 1,4 respectivamente.

El término N_{hb} se calcula de acuerdo con la ecuación 2:

$$N_{\text{hb}} = \sum_{\text{h-bonds}} f(\Delta R, \Delta \alpha) \times f(N_{\text{neighb}}) \times f_{\text{pcs}}$$

$f(\Delta R, \Delta \alpha)$ es una función de penalización que representa grandes desviaciones de los enlaces de hidrógeno del ideal y se calcula de acuerdo con la ecuación 3:

$$f(\Delta R, \Delta \alpha) = f1(\Delta R) \times f2(\Delta \alpha)$$

Donde:

$$f1(\Delta R) : \text{si } \Delta R \leq \text{TOL}$$

o

$$= 1 - (\Delta R - \text{TOL}) / \text{si } \Delta R \leq 0.4 + \text{TOL}$$

o

$$\text{si } \Delta R > 0.4 + \text{TOL}$$

Y:

$$f2(\Delta \alpha) : \text{si } \Delta \alpha < 30^\circ$$

o

$$= 1 - (\Delta\alpha - 30), \text{ si } \Delta\alpha \leq 80^\circ$$

0

$$\text{si } \Delta\alpha > 80^\circ$$

TOL es la desviación tolerada en el largo de enlace de hidrógeno = 0,25Å

5 ΔR es la desviación del largo de enlace de hidrógeno H-O/N a partir del valor ideal = 1,9Å

$\Delta\alpha$ es la desviación del ángulo de enlace de hidrógeno $\angle N/O-H..O/N$ a partir de su valor idealizado de 180°

$f(N_{\text{neighb}})$ distingue entre las partes cóncavas y convexas de la superficie de una proteína y por lo tanto asigna mayor peso a las interacciones polares encontradas en los bolsillos que a aquellas encontradas en la superficie de la proteína. Esta función se calcula de acuerdo con la ecuación 4 a continuación:

$$10 \quad f(N_{\text{neighb}}) = (N_{\text{neighb}}/N_{\text{neighb},0})^\alpha \text{ donde } \alpha = 0.5$$

N_{neighb} es el número de átomos de proteínas no hidrógenos que están más próximos que 5Å a cualquier átomo de proteína dado.

$N_{\text{neighb},0}$ es una constante = 25

15 f_{pcs} es una función que permite el área de superficie de contacto polar por enlace de hidrógeno y por lo tanto distingue entre enlaces de hidrógeno fuertes y débiles y su valor se determina de acuerdo con los siguientes criterios:

$$f_{\text{pcs}} \text{ cuando } \text{len } A_{\text{polar}}/N_{\text{HB}} < 10 \text{ \AA}^2$$

0

$$f_{\text{pcs}} \text{ cuando } \text{len } A_{\text{polar}}/N_{\text{HB}} > 10 \text{ \AA}^2$$

20 A_{polar} es el tamaño de la superficie de contacto ligando-proteína polar

N_{HB} es el número de enlaces de hidrógeno

β es una constante cuyo valor = 1,2

25 Para la implementación de la función de puntuación de Böhm modificada, las contribuciones a partir de las interacciones iónicas, ΔG_{ionic} , se calculan en una forma similar a aquellas a partir de los enlaces de hidrógeno descritos con anterioridad dado que se asume la misma dependencia geométrica. El término N_{lipo} se calcula de acuerdo con la ecuación 5 a continuación:

$$N_{\text{lipo}} = \sum_{1L} f(r_{1L})$$

$f(r_{1L})$ se calcula para todos los átomos de ligando lipofílico, 1, y todos los átomos de proteína lipofílica, L, de acuerdo con los siguientes criterios:

$$f(r_{1L}) = 1 \text{ cuando } r_{1L} \leq R1 \quad f(r_{1L}) = (r_{1L} - R1)/(R2 - R1) \text{ cuando } R2 < r_{1L} < R1$$

30

$$R1$$

$$f(r_{1L}) \text{ cuando } r_{1L} \geq R2$$

Donde:

$$R1 = r_1^{vdw} + r_L^{vdw} + 0.5$$

y

$$R2 = R1 + 3.0$$

y r_1^{vdw} es el radio de Van der Waals del átomo 1

5 y r_L^{vdw} es el radio de Van der Waals del átomo L

El término N_{rot} es el número de enlaces rotables de la cadena lateral de aminoácidos y se toman como el número de enlaces acíclicos $sp^3 - sp^3$ y $sp^3 - sp^2$. Las rotaciones del terminal $-CH_3$ o $-NH_3$ no se toman en cuenta.

El término final, E_{vdw} , se calcula de acuerdo con la [ecuación 6](#) a continuación:

$$E_{vdw} = \epsilon_1 \epsilon_2 \left((r_1^{vdw} + r_2^{vdw})^{12} / r^{12} - (r_1^{vdw} + r_2^{vdw})^6 / r^6 \right),$$

10 donde:

ϵ_1 y ϵ_2 son constantes dependientes de la identidad del átomo

$r_1^{vdw} + r_2^{vdw}$ son los radios atómicos de Van der Waals

r es la distancia entre un par de átomos.

15 Por lo que se refiere a la ecuación 6, en una realización, se da a las constantes ϵ_1 y ϵ_2 los valores de átomo: C: 0,245, N: 0,283, O: 0,316, S: 0,316, respectivamente (es decir, para los átomos de Carbono, Nitrógeno, Oxígeno y Azufre, respectivamente). En cuanto a las ecuaciones 5 y 6, se da a los radios de Van der Waals los valores de átomo C: 1,85, N: 1,75, O: 1,60, S: 2,00Å.

20 Se debe entender que todos los valores y constantes predeterminados dados en la ecuaciones más arriba se determinan dentro de las restricciones de los actuales conocimientos de las interacciones de ligandos-proteínas particularmente con respecto al tipo de cálculo que se está aplica en este documento. Por lo tanto, es posible que estos valores y constantes cambien, a medida que esta función de puntuación se refine más, por consiguiente se podría utilizar cualquier valor numérico adecuado que ofrece los resultados deseados en materia de estimar la energía de unión de una proteína a un ligando, y por consiguiente podría estar incluido dentro del alcance de la presente invención.

25 Como se describió con anterioridad, la función de puntuación se aplica a los datos extraídos a partir de la base de datos de las conformaciones de cadenas laterales, las identidades de átomos, y las distancias interatómicas. A los fines de la presente invención, el número de moléculas del MHC Clase II incluido en esta base de datos es de 42 modelos más cuatro estructuras resueltas. Debería ser evidente a partir de las descripciones anteriores que la naturaleza modular de la construcción del método informático de la presente invención significa
30 que los modelos nuevos pueden agregarse y escanearse simplemente con la biblioteca de cadenas principales de péptidos y la función de búsqueda conformacional de cadenas laterales para crear grupos de datos adicionales que pueden procesarse mediante la función de puntuación peptídica como se describió con anterioridad. Esto permite aumentar fácilmente el repertorio de moléculas del MHC Clase II escaneadas, o reemplazar las estructuras y datos asociados si los datos están disponibles para crear modelos más precisos de los alelos existentes.

35 El presente método de predicción puede calibrarse contra un grupo de datos que comprenda un gran número de péptidos cuya afinidad a diferentes moléculas del MHC Clase II se ha determinado experimentalmente con anterioridad. Mediante la comparación de los datos calculados con los experimentales, se puede determinar un corte de valor por encima del cual se sabe que todos los epítomos de células T determinados experimentalmente se predicen correctamente.

40 Se debería entender que, aunque la función de puntuación anterior es relativamente simple comparada con algunas metodologías sofisticadas que están disponibles, los cálculos se realizan extremadamente rápido. Se debería entender también que el objetivo no es calcular la energía de unión real per se para cada péptido acoplado en el surco de unión de una proteína del MHC Clase II seleccionada. El objetivo subyacente es obtener datos de energía de unión comparativos como una ayuda para predecir la ubicación de los epítomos de células T en base a la
45 estructura primaria (es decir, secuencia de aminoácidos) de una proteína seleccionada. Una energía de unión relativamente alta o una energía de unión por encima de un valor umbral seleccionado sugeriría la presencia de un

epítopo de células T en el ligando. Después puede someterse al ligando a por lo menos una ronda de sustituciones de aminoácidos y recalcular la energía de unión. Debido a la naturaleza rápida de los cálculos, estas manipulaciones de la secuencia peptídica pueden realizarse interactivamente dentro de la interfaz de usuario del programa sobre un hardware informático disponible de modo eficiente y rentable. No se requiere por consiguiente una inversión importante en hardware informático.

Será evidente para una persona con conocimientos en el arte que es posible utilizar otro software disponible para los mismos propósitos. En particular, puede utilizarse software más sofisticado que es capaz de acoplar ligandos dentro de sitios de unión de proteínas conjuntamente con la minimización de energía. Los ejemplos de software de acoplamiento son: DOCK (Kuntz et al., J. Mol. Biol., 161:269-288 (1982)), LUDI (Böhm, H.J., J. Comput Aided Mol. Des., 8:623-632 (1994)) y FLEXX (Rarey M., et al., ISMB, 3:300-308 (1995)). Los ejemplos de software de modelado y manipulación incluyen: AMBER (Tripos) y CHARMM (Molecular Simulations Inc.). El uso de estos métodos de cálculo limitaría gravemente el rendimiento del método de esta invención debido a la extensión de tiempo de procesamiento requerida para realizar los cálculos necesarios. Sin embargo, es factible que tales métodos pudieran utilizarse como una "prueba secundaria" para obtener cálculos más precisos de la energía de unión para los péptidos que se descubre que son "ligandos positivos" mediante el método de la presente invención. La limitación del tiempo de procesamiento para cálculos dinámicos moleculares o mecánicos moleculares sofisticados es algo que se define tanto por el diseño del software que realiza estos cálculos como por las limitaciones tecnológicas actuales del hardware informático. Puede anticiparse que, en el futuro, con la escritura de un código más eficiente y los continuos aumentos en velocidad de los procesadores informáticos, puede ser factible realizar tales cálculos dentro de un plazo más aceptable. Se puede encontrar información adicional sobre las funciones de energía aplicadas a las macromoléculas y la consideración de las diferentes interacciones que tienen lugar dentro de una estructura de proteína plegada en: Brooks, B.R., et al., J. Comput. Chem., 4:187-217 (1983) y se puede encontrar información adicional sobre las interacciones proteína-ligando generales en: Dauber-Osguthorpe et al., Proteins 4(1):31-47(1988), que se incorporan a este documento a modo de referencia en su totalidad. También se puede encontrar información útil sobre otros antecedentes, por ejemplo, en Fasman, G.D., ed., Prediction of Protein Structure and the Principles of Protein Conformation, Plenum Press, New York, ISBN: 0-306 4313-9.

Listado de secuencias

<110> Lexigen Pharmaceuticals Corp.

<120> Reducción de la inmunogenicidad de las proteínas de fusión

<130> LEX-017PCT

<150> US 60/280,625

<151> 2001-03-30

<160> 59

<170> Patente en versión 3.0

<210> 1

<211> 330

<212> PRT

<213> Homo sapiens

<220>

<221> característica_misc

<223> región C cadena pesada Ig gamma humana

<400> 1

Ala	Ser	Thr	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys
1			5					10						15	
Ser	Thr	Ser	Gly	Gly	Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr
		20						25					30		
Phe	Pro	Glu	Pro	Val	Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser
		35					40					45			
Gly	Val	His	Thr	Phe	Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser
	50					55					60				
Leu	Ser	Ser	Val	Val	Thr	Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr
65				70					75						80
Tyr	Ile	Cys	Asn	Val	Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys
			85						90					95	
Lys	Val	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys
		100						105					110		
Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro
		115					120					125			
Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys
	130					135					140				
Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp
145					150					155					160
Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu
			165						170					175	
Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val	Leu	Thr	Val	Leu

				180					185					190			
His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys	Cys	Lys	Val	Ser	Asn		
		195					200					205					
Lys	Ala	Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr	Ile	Ser	Lys	Ala	Lys	Gly		
	210					215					220						
Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu	Pro	Pro	Ser	Arg	Asp	Glu		
225					230					235					240		
Leu	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys	Leu	Val	Lys	Gly	Phe	Tyr		
			245						250					255			
Pro	Ser	Asp	Ile	Ala	Val	Glu	Trp	Glu	Ser	Asn	Gly	Gln	Pro	Glu	Asn		
		260						265					270				
Asn	Tyr	Lys	Thr	Thr	Pro	Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	Phe	Phe		
		275					280					285					
Leu	Tyr	Ser	Lys	Leu	Thr	Val	Asp	Lys	Ser	Arg	Trp	Gln	Gln	Gly	Asn		
	290					295					300						
Val	Phe	Ser	Cys	Ser	Val	Met	His	Glu	Ala	Leu	His	Asn	His	Tyr	Thr		
305					310					315					320		
Gln	Lys	Ser	Leu	Ser	Leu	Ser	Pro	Gly	Lys								
				325					330								

$\langle 210 \rangle$ 2

<211> 326

<212> PRT

<213> Homo sapiens

$\langle 220 \rangle$

<221> característica_misc

<223> región C cadena Ig gamma-2 humana

$\langle 400 \rangle$ 2

Ala	Ser	Thr	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Cys	Ser	Arg
1				5					10					15	
Ser	Thr	Ser	Glu	Ser	Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr
			20					25					30		
Phe	Pro	Glu	Pro	Val	Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser
		35					40					45			
Gly	Val	His	Thr	Phe	Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser
	50					55					60				
Leu	Ser	Ser	Val	Val	Thr	Val	Pro	Ser	Ser	Asn	Phe	Gly	Thr	Gln	Thr
65					70					75					80
Tyr	Thr	Cys	Asn	Val	Asp	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys
				85					90					95	
Thr	Val	Glu	Arg	Lys	Cys	Cys	Val	Glu	Cys	Pro	Pro	Cys	Pro	Ala	Pro
			100					105					110		

Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
115 120 125

Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
130 135 140

Val Ser His Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly
145 150 155 160

Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn
165 170 175

Ser Thr Phe Arg Val Val Ser Val Leu Thr Val Val His Gln Asp Trp
180 185 190

Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro
195 200 205

Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly Gln Pro Arg Glu
210 215 220

Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn
225 230 235 240

Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
245 250 255

Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
260 265 270

Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys
275 280 285

Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
290 295 300

Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
305 310 315 320

Ser Leu Ser Pro Gly Lys
325

<210> 3

<211> 362

5 <212> PRT

<213> Homo sapiens

<220>

<221> característica_misc

<223> región constante Ig3 humana

10 <400> 3

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
 1 5 10 15
 Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
 20 25 30
 Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
 35 40 45
 Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
 50 55 60
 Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
 65 70 75 80
 Tyr Thr Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
 85 90 95
 Arg Val Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr His Thr Cys Pro
 100 105 110
 Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
 115 120 125
 Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys
 130 135 140
 Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro
 145 150 155 160
 Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys
 165 170 175
 Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Gln Phe Lys Trp
 180 185 190
 Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Trp Glu
 195 200 205
 Glu Gln Tyr Asn Ser Thr Phe Arg Val Val Ser Val Leu Thr Val Leu
 210 215 220
 His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn
 225 230 235 240
 Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly
 245 250 255
 Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu
 260 265 270
 Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr
 275 280 285
 Pro Ser Asp Ile Ala Met Glu Trp Glu Ser Ser Gly Gln Pro Glu Asn
 290 295 300
 Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe
 305 310 315 320
 Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
 325 330 335
 Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr
 340 345 350
 Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys

5

<210> 4

<211> 327

<212> PRT

<213> Homo sapiens

<220>

<221> característica_misc

<223> región C cadena Ig gamma-4 humana

<400> 4

```

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
1          5          10          15
Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
          20          25          30
Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
          35          40          45
Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
          50          55          60
Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Lys Thr
65          70          75          80
Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp Lys
          85          90          95
Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro Ala Pro
          100          105          110
Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
          115          120          125
Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val
130          135          140
Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp
145          150          155          160
Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe
          165          170          175
Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp
180          185          190
Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu
195          200          205
Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
210          215          220
Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys
225          230          235          240
Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
          245          250          255

```

Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
260 265 270
Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
275 280 285
Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser
290 295 300
Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
305 310 315 320
Leu Ser Leu Ser Leu Gly Lys
325

<210> 5

<211> 9

<212> PRT

5 <213> Secuencia artificial

<220>

<223> epítipo potencial de células T

<400> 5

Lys Ser Leu Ser Leu Ser Pro Gly Lys
1 5

10 Lis Ser Leu Ser Leu Ser Pro Gli Lis 1 5

<210> 6

<211> 9

<212> PRT

<213> Secuencia artificial

15 <220>

<223> epítipo potencial de células T mutado

<400> 6

Lys Ser Ala Thr Ala Thr Pro Gly Lys
1 5

Lis Ser Ala Tr Ala Tr Pro Gli Lis 1 5

20 <210> 7

<211> 9

<212> PRT

<213> Secuencia artificial

<220>

25 <223> epítipo potencial de células T mutado

<400> 7

Lys Ser Ala Thr Ala Thr Pro Gly Ala
1 5

<210> 8

<211> 13

5

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia cercana al extremo C-terminal del dominio CH3 de la región IgA Fc

<400> 8

Gln Lys Thr Ile Asp Arg Leu Ala Gly Lys Pro Thr His
1 5 10

10

<210> 9

<211> 13

<212> PRT

<213> Secuencia artificial

15

<220>

<223> secuencia mutada cercana al extremo C-terminal del dominio CH3 de la región IgA Fc

<400> 9

Gln Lys Thr Ala Asp Arg Thr Ala Gly Lys Pro Thr His
1 5 10

<210> 10

20

<211> 13

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia desinmunizada en una fusión IgA-X

25

<400> 10

Gln Lys Thr Pro Thr Arg Thr Ala Gly Lys Pro Thr His
1 5 10

<210> 11

<211> 13

<212> PRT

30

<213> Secuencia artificial

<220>

<223> secuencia desinmunizada en una fusión IgA-X

<400> 11

Gln Lys Thr Pro Thr Arg Pro Ala Gly Lys Pro Thr His
1 5 10

<210> 12

<211> 13

5

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia desinmunizada en una fusión IgA-X

<400> 12

Gln Lys Thr Ala Thr Arg Pro Ala Gly Lys Pro Thr His
1 5 10

10

<210> 13

<211> 13

<212> PRT

<213> Secuencia artificial

15

<220>

<223> epítipo potencial de células T en la unión HSA-IFNalfa

<400> 13

Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
1 5 10

<210> 14

20

<211> 13

<212> PRT

<213> Secuencia artificial

<220>

<223> extremo C-terminal de la albúmina modificado

25

<400> 14

Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Thr Thr Ala
1 5 10

<210> 15

<211> 30

<212> PRT

30

<213> Secuencia artificial

<220>

<223> secuencia en una proteína de fusión de Fc

<400> 15

```

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly
1          5          10          15
      /
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
      20          25          30
    
```

<210> 16

5 <211> 30

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia modificada en una proteína de fusión de Fc

10 <400> 16

```

His Asn His Tyr Thr Gln Lys Ser Ala Thr Ala Thr Pro Gly Lys Gly
1          5          10          15
Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
      20          25          30
    
```

<210> 17

<211> 9

<212> PRT

15 <213> Secuencia artificial

<220>

<223> secuencia de unión

<400> 17

```

Leu Ser Leu Ser Pro Gly Lys Ala Pro
1          5
    
```

20 <210> 18

<211> 9

<212> PRT

<213> Secuencia artificial

<220>

25 <223> secuencia de unión modificada

<400> 18

Ala Thr Ala Thr Pro Gly Ala Ala Pro
1 5

<210> 19

<211> 9

<212> PRT

5 <213> Secuencia artificial

<220>

<223> secuencia de unión modificada

<400> 19

Leu Asn Leu Ser Pro Gly Ala Ala Pro
1 5

10 <210> 20

<211> 13

<212> PRT

<213> Secuencia artificial

<220>

15 <223> péptido sintético que contiene un epítipo reactivo

<400> 20

Lys Ser Leu Ser Leu Ser Pro Gly Lys Ala Pro Thr Ser
1 5 10

<210> 21

<211> 13

20 <212> PRT

<213> Secuencia artificial

<220>

<223> péptido sintético modificado que contiene un epítipo reactivo

<400> 21

Lys Ser Ala Thr Ala Thr Pro Gly Lys Ala Pro Thr Ser
1 5 10

25 <210> 22

<211> 39

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia de unión de albúmina-CD4

<400> 22

```

    Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala
    1             5             10             15

    Ala Leu Gly Leu Lys Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val
              20             25             30

    Glu Leu Thr Cys Thr Ala Ser
              35

```

5 <210> 23

<211> 13

<212> PRT

<213> Secuencia artificial

<220>

10 <223> epítopo potencial de células T en fusión de HSA-IFNalfa

<400> 23

```

    Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu Cys
    1             5             10

```

<210> 24

<211> 13

15 <212> PRT

<213> Secuencia artificial

<220>

<223> epítopo potencial de células T en fusión de HSA-IFNalfa

<400> 24

```

    Leu Gly Leu Cys Asp Leu Pro Gln Thr His Ser Leu Gly
    1             5             10

```

20 <210> 25

<211> 5

<212> PRT

<213> Secuencia artificial

25 <220>

<223> secuencia de albúmina del extremo C-terminal

<400> 25

Ala Ala Leu Gly Leu
1 5

<210> 26

<211> 5

<212> PRT

5 <213> Secuencia artificial

<220>

<223> secuencia mutada de albúmina del extremo C-terminal

<400> 26

Thr Ala Thr Thr Ala
1 . 5

10 <210> 27

<211> 19

<212> PRT

<213> Secuencia artificial

<220>

15 <223> región de unión de albúmina modificada

<400> 27

Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Thr Ala
1 5 10 15

Thr Thr Ala

<210> 28

<211> 21

20 <212> PRT

<213> Secuencia artificial

<220>

<223> secuencia de Fc no mutada

<400> 28

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
1 5 10 15

Pro Glu Leu Leu Gly
20

25 <210> 29

<211> 52

<212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión de Fc GLP-1-mutante

<400> 29

```

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1          5          10          15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Glu
          20          25          30

Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro
          35          40          45

```

5

Glu Leu Leu Gly
50

<210> 30

<211> 41

<212> PRT

10

<213> Secuencia artificial

<220>

<223> unión de fusión de Fc GLP-1-normal

<400> 30

```

Ser Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val
1          5          10          15

Lys Gly Arg Gly Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
          20          25          30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
          35          40

```

15

<210> 31

<211> 13

<212> PRT

<213> Secuencia artificial

<220>

20

<223> epítipo potencial de células T en la fusión de GLP-1-Fc

<400> 31

```

Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Glu
1          5          10

```

<210> 32

<211> 13

<212> PRT
 <213> Secuencia artificial
 <220>
 <223> epítipo potencial de células T en la unión de fusión de GLP-1-Fc
 5 <400> 32
 Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Glu Pro
 1 5 10
 <210> 33
 <211> 13
 <212> PRT
 10 <213> Secuencia artificial
 <220>
 <223> epítipo potencial de células T en unión de fusión de GLP-1-Fc
 <400> 33
 Ala Trp Leu Val Lys Gly Arg Gly Glu Pro Lys Ser Ser
 1 5 10
 15 <210> 34
 <211> 52
 <212> PRT
 <213> Secuencia artificial
 <220>
 20 <223> unión de fusión de GLP-1Fc desinmunizada
 <400> 34
 His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1 5 10 15
 Gln Ala Ala Lys Glu Phe Ala Ala Trp Ala Val Thr Gly Thr Gly Glu
 20 25 30
 Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro
 35 40 45
 Glu Leu Leu Gly
 50
 <210> 35
 <211> 41
 25 <212> PRT
 <213> Secuencia artificial
 <220>

<223> unión de fusión de GLP-1-Fc con un sitio de glicosilación

<400> 35

```

Ser Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val
1           5           10           15

Lys Gly Arg Asn Gly Ser Lys Ser Ser Asp Lys Thr His Thr Cys Pro
          20           25           30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
          35           40

```

<210> 36

<211> 41

<212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión de TNF-R-gamma-1

<400> 36

```

Ser Thr Ser Phe Leu Leu Pro Met Gly Pro Ser Pro Pro Ala Glu Gly
1           5           10           15

Ser Thr Gly Asp Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
          20           25           30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
          35           40

```

<210> 37

<211> 41

<212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión de TNF-R-Fc

<400> 37

```

Ser Thr Ser Phe Leu Leu Pro Met Gly Pro Ser Pro Pro Ala Glu Gly
1           5           10           15

Ser Thr Gly Asn Gly Ser Lys Ser Cys Asp Lys Thr His Thr Cys Pro
          20           25           30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
          35           40

```

<210> 38

<211> 40

<212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión de Fc-IL12p35

<400> 38

```

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
1          5          10          15

Ser Pro Gly Lys Arg Asn Leu Pro Val Ala Thr Pro Asp Pro Gly Met
          20          25          30

Phe Pro Cys Leu His His Ser Gln
          35          40

```

5 <210> 39

<211> 40

<212> PRT

<213> Secuencia artificial

<220>

10 <223> unión de fusión de IL-12p40-IL2

<400> 39

```

Arg Ala Gln Asp Arg Tyr Tyr Ser Ser Ser Trp Ser Glu Trp Ala Ser
1          5          10          15

Val Pro Cys Ser Ala Pro Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu
          20          25          30

Gln Leu Glu His Leu Leu Leu Asp
          35          40

```

<210> 40

<211> 40

15 <212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión de Fc-IL12p35 modificada

<400> 40

```

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Ala Thr Ala
1          5          10          15

Thr Pro Gly Lys Arg Asn Leu Pro Val Ala Thr Pro Asp Pro Gly Met
          20          25          30

Phe Pro Cys Leu His His Ser Gln
          35          40

```

20 <210> 41

<211> 40

<212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión de IL12p40-IL2 modificada

<400> 41

```

Arg Ala Gln Asp Arg Tyr Tyr Ser Ser Ser Trp Ser Glu Trp Ala Ser
1          5          10          15

Val Pro Cys Ser Asn Gly Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu
          20          25          30

Gln Leu Glu His Leu Leu Leu Asp
          35          40

```

5

<210> 42

<211> 13

<212> PRT

<213> Secuencia artificial

<220>

<223> epítipo potencial de células T en fusión de IL12p40-IL

<400> 42

```

Ser Glu Trp Ala Ser Val Pro Cys Ser Asn Gly Thr Ser
1          5          10

```

<210> 43

<211> 13

<212> PRT

<213> Secuencia artificial

<220>

<223> epítipo potencial de células T en unión de fusión de IL12p40-IL2

<400> 43

```

Ala Ser Val Pro Cys Ser Asn Gly Thr Ser Ser Ser Thr
1          5          10

```

<210> 44

<211> 41

<212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión de IL4-Fc

<400> 44

10

15

20

25

Glu Asn Phe Leu Glu Arg Leu Lys Thr Ile Met Arg Glu Lys Tyr Ser
 1 5 10 15
 Lys Cys Ser Ser Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
 20 25 30
 Pro Cys Pro Ala Pro Glu Leu Leu Gly
 35 40

<210> 45

<211> 40

<212> PRT

5 <213> Secuencia artificial

<220>

<223> unión de fusión de Fc-GMCSF

<400> 45

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
 1 5 10 15
 Ser Pro Gly Lys Pro Ala Arg Ser Pro Ser Pro Ser Thr Gln Pro Trp
 20 25 30
 Glu His Val Asn Ala Ile Gln Glu
 35 40

10 <210> 46

<211> 13

<212> PRT

<213> Secuencia artificial

<220>

15 <223> epítipo potencial de células T en la unión de fusión de IL4-Fc

<400> 46

Glu Lys Tyr Ser Lys Cys Ser Ser Glu Pro Lys Ser Cys
 1 5 10

<210> 47

<211> 41

20 <212> PRT

<213> Secuencia artificial

<220>

<223> fusión de IL4-Fc modificada

<400> 47

Glu Asn Phe Leu Glu Arg Leu Lys Thr Ile Met Arg Glu Lys Tyr Ser
 1 5 10 15

Lys Cys Ser Ser Thr Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
20 25 30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
35 40

<210> 48

<211> 40

<212> PRT

5 <213> Secuencia artificial

<220>

<223> unión de fusión de Fc-GMCSF desinmunizada

<400> 48

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Ala Thr Ala
1 5 10 15

Thr Pro Gly Lys Pro Ala Arg Ser Pro Ser Pro Ser Thr Gln Pro Trp
20 25 30

Glu His Val Asn Ala Ile Gln Glu
35 40

10 <210> 49

<211> 35

<212> PRT

<213> Secuencia artificial

<220>

15 <223> unión de fusión bisagra de IgG2CH1-IgG1

<400> 49

Gln Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val
1 5 10 15

Asp Lys Thr Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
20 25 30

Pro Cys Pro
35

<210> 50

<211> 35

20 <212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión de IgG1bisagra-IgG2CH2

<400> 50

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
1 5 10 15

Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
20 25 30

Asp Thr Leu
35

<210> 51

<211> 13

5 <212> PRT

<213> Secuencia artificial

<220>

<223> epítopo de células T en la unión de fusión bisagra de IgG2CH1-IgG1

<400> 51

10 Thr Lys Val Asp Lys Thr Val Glu Pro Lys Ser Cys Asp
1 5 10

<210> 52

<211> 13

<212> PRT

<213> Secuencia artificial

15 <220>

<223> epítopo de células T en la unión de fusión bisagra de IgG2CH1-IgG1

<400> 52

Lys Thr Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr
1 5 10

<210> 53

20 <211> 35

<212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión bisagra de IgG2CH1-IgG1 modificada

25 <400> 53

Gln Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Ala
1 5 10 15

Asp Lys Thr Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
20 25 30

Pro Cys Pro
35

<210> 54

<211> 35

<212> PRT

<213> Secuencia artificial

5

<220>

<223> unión de fusión bisagra de IgG2CH1-IgG1 modificada

<400> 54

Gln Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val

1 5 10 15

Asp Lys Thr Val Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro
20 25 30

Pro Cys Pro
35

10

<210> 55

<211> 35

<212> PRT

<213> Secuencia artificial

<220>

15

<223> unión de fusión de IgG1bisagra-IgG2CH2 modificada

<400> 55

Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
1 5 10 15

Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
20 25 30

Asp Thr Leu
35

<210> 56

<211> 166

20

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia de EPO mutante

<400> 56

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu Gly
20 25 30

Pro Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
65 70 75 80

Leu Val Asn Ser Ser Gln Pro Cys Glu Gly Leu Gln Leu His Val Asp
85 90 95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
100 105 110

Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
115 120 125

Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
130 135 140

Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
145 150 155 160

Cys Arg Thr Gly Asp Arg
165

<210> 57

<211> 17

5 <212> PRT

<213> Secuencia artificial

<220>

<223> unión de fusión de CH3-EPO

<400> 57

Thr Gln Lys Ser Ala Thr Ala Thr Pro Gly Ala Ala Pro Pro Arg Leu
1 5 10 15

Ile

10

<210> 58

<211> 8

<212> PRT

<213> Secuencia artificial

15

<220>

<223> secuencia de IgG2 CH3

<400> 58

Lys Ser Leu Ser Leu Ser Pro Gly
1 5

<210> 59

<211> 8

5

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia de IgG2CH3 modificada

<400> 59

Lys Ser Ala Thr Ala Thr Pro Gly
1 5

10

REIVINDICACIONES

1. Método para la reducción de la inmunogenicidad mediante la eliminación de epítomos de células T no propias de una proteína de fusión, que comprende una primera proteína y una segunda proteína enlazada a dicha primera proteína mediante una unión de fusión; el método comprende una modificación de la secuencia de aminoácidos de la región de unión que rodea dicha unión de fusión, en donde la modificación se realiza mediante uno de los siguientes pasos:
- (i) introducción de un sitio de glicosilación N-ligada u O-ligada,
 - (ii) reemplazo de una Leu, Val, Ile, Met, Pe, Tir o Trp en los ocho aminoácidos más cercados al extremo C-terminal del compañero de fusión del extremo N-terminal en dicha proteína de fusión con una Tr, Ala o Pro, o
 - (iii) reemplazo de una secuencia de aminoácidos Leu - Ser - Leu - Ser por la secuencia de aminoácidos Ala - Tr - Ala - Tr si la primera proteína es una molécula IgG o un fragmento de la misma.
2. Método según la reivindicación 1, en donde el paso (i) se logra mediante la introducción del sitio de glicosilación dentro de 10, 5 ó 2 aminoácidos de la unión de fusión.
3. Método según la reivindicación 1, en donde el paso (i) se logra mediante la introducción de una secuencia de aminoácido Asn-X-Ser/Tr, en donde X es cualquier aminoácido.
4. Método según la reivindicación 1, en donde la primera proteína es una molécula Ig o un fragmento de la misma.
5. Método según la reivindicación 1, en donde la primera proteína es una Molécula Ig o un fragmento de la misma y la segunda proteína es una citocina.
6. Método según la reivindicación 5, en donde – a fin de aumentar la vida media en suero – se realiza una mutación adicional entre la IgG- y la fracción citocina de la proteína de fusión mediante el reemplazo de la lisina del extremo C-terminal de la fracción del anticuerpo con una alanina o una leucina.
7. Método según cualquiera de las reivindicaciones 1 a 6, en donde la región de unión comprende entre 1 y 15 aminoácidos.
8. Método según la reivindicación 7, en donde la región de unión comprende entre 1 y 9 aminoácidos.
9. Proteína de fusión con inmunogenicidad reducida obtenida mediante el método según cualquiera de las reivindicaciones 1 a 8, en donde dicha proteína de fusión se selecciona entre el grupo que consiste en:
- (i) huKS-IL2 que comprende Ala - Tr - Ala - Tr en la secuencia IgG dentro de la región de unión en lugar de Leu - Ser - Leu - Ser, en donde huKS es un anticuerpo dirigido a la molécula de adhesión celular epitelial humana KSA (EP-CAM);
 - (ii) Fc-IL12-IL2 que comprende Ala - Tr - Ala - Tr en la secuencia IgG dentro de la región de unión en lugar de Leu - Ser - Leu - Ser, en donde Fc es dicha primera proteína y la proteína de fusión IL12-IL2 es dicha segunda proteína;
 - (iii) proteína de fusión Fc-IL12-L2 que comprende un sitio de glicosilación Asn-Gli en las primeras posiciones dentro de la molécula IL2; en donde la proteína de fusión Fc-IL12 es dicha primera proteína e IL2 es dicha segunda proteína;
 - (vi) Fc-EPO que comprende Ala - Tr-Ala -Tr en la secuencia IgG dentro de la región de unión en lugar de Leu - Ser- Leu - Ser;
 - (v) Fc-EPO que comprende Ala - Tr - Ala -Tr en la secuencia IgG dentro de la región de unión en lugar de Leu - Ser- Leu - Ser, en donde la cadena Ig es la IgG2 que comprende una región bisagra IgG1.
10. Composición farmacéutica que comprende en una cantidad terapéuticamente efectiva una proteína de fusión según la reivindicación 9 opcionalmente junto con un portador farmacéuticamente aceptable.
11. Ácido nucleico que codifica una proteína de fusión según la reivindicación 9.