



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 367 475**

51 Int. Cl.:
C12P 21/06 (2006.01)
C12N 9/00 (2006.01)
C12N 15/00 (2006.01)
C07H 21/02 (2006.01)
C07H 21/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **00973574 .7**
96 Fecha de presentación : **16.10.2000**
97 Número de publicación de la solicitud: **1282722**
97 Fecha de publicación de la solicitud: **12.02.2003**

54 Título: **Proteínas estabilizadas.**

30 Prioridad: **15.10.1999 US 159763 P**

45 Fecha de publicación de la mención BOPI:
03.11.2011

45 Fecha de la publicación del folleto de la patente:
03.11.2011

73 Titular/es: **AVATAR MEDICAL, L.L.C.**
140 58Th Street, Bldg A, Unit 8J
Brooklyn, New York 11220, US

72 Inventor/es: **Marshall, Christopher, P.;**
Hoffman, Alexander;
Errico, Joseph, P. y
Marshall, Paul, B.

74 Agente: **Pons Ariño, Ángel**

ES 2 367 475 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

Proteínas estabilizadas

1. Campo de la invención

5 La presente invención se refiere a métodos de entrecruzamiento para estabilizar polipéptidos y complejos de polipéptido para usos comerciales (farmacéutico, terapéutico, y industrial), y a polipéptidos y complejos de polipéptido así entrecruzados.

2. Antecedentes de la invención

2.1. Estructura y función de los polipéptidos y complejos de polipéptido

10 Una molécula de proteína consiste en una cadena de polipéptido lineal de aminoácidos que se pliega en forma compleja en tres dimensiones para formar, por ejemplo, superficies de interacción, bolsillos de unión y sitios activos. De manera general se requiere un plegado tridimensional específico para la función de la proteína, en donde el plegado en sí mismo se especifica por la secuencia lineal de aminoácidos (es decir, la estructura principal de la proteína). Es notable, sin embargo, que estructuras principales diferentes puedan tener plegados tridimensionales casi idénticos. La evolución ha conservado plegados específicos a un grado mayor que las estructuras principales
15 específicas. En los procesos de plegado de proteína sigue abierto un campo activo de estudio. Se sabe, sin embargo, que los elementos de estructura secundarios tal como hélices alfa, láminas beta y giros beta contribuyen al ensamble de la estructura terciaria de un polipéptido. Se dice que una entidad de proteína biológica hecha de varios polipéptidos tiene estructura cuaternaria.

20 El plegado de una proteína finalmente resulta de la interacción de fuerzas intra e intermoleculares. Como tal, una proteína plegada tiene una estabilidad finita que se traduce en una "vida útil" estructural y funcional finita en un ambiente de solvente dado. Por ejemplo, en un ambiente acuoso, las proteínas logran estabilidad en parte al agrupar los residuos hidrófobos en el núcleo de proteína y los residuos hidrófilos en la interfaz de proteína-solvente. De acuerdo con lo anterior, la vida útil de actividad para una proteína dada es en parte una función de las propiedades del solvente. Adicionalmente, los enlaces químicos tales como disulfuros ocurren en la naturaleza para
25 fijar la coordinación de las cadenas laterales no vecinas en proximidad cercada a una proteína plegada, estabilizando por lo tanto su estructura y su función.

30 En muchos sistemas biológicos, las proteínas se asocian una a la otra para formar dímeros o multímeros de orden mayor (es decir estructuras cuaternarias), y solo como tales llevan a cabo sus funciones específicas. La formación de tales complejos es frecuentemente un evento importante en la regulación de la actividad de las proteínas. Se han encontrado diversos mecanismos para regular la formación del complejo de proteína, tales como la unión de ligando o la modificación post-translacionales. Las funciones de los complejos de proteína pueden variar desde proveer estructura a la matriz intracelular, en donde, por ejemplo, la actina forma una red estructural, hasta factores de transcripción.

35 Las proteínas consisten de dominios funcionales discretos. Los dominios de la función análoga o similar en diferentes proteínas muestran usualmente las similitudes de la secuencia de aminoácidos y se relacionan con la evolución. El "barajado de dominios" ha jugado una función principal en la evolución (así como también en la ingeniería genética) de las proteínas con funcionalidades altamente diversas. Los dominios de interacción, por ejemplo, se pueden encontrar en las proteínas de muchas diferentes funciones; sin embargo, las similitudes de secuencia revelan su presencia. Los estudios cristalográficos han mostrado que los dominios relacionados son aún
40 más conservados en la estructura secundaria, terciaria y cuaternaria en la secuencia de aminoácidos primaria, de tal manera que las inferencias estructurales se pueden hacer acerca de un dominio particular si los datos estructurales están disponibles en uno o preferiblemente múltiples dominios relacionados (ver por ejemplo, Hofmann K., Cell Mol. Life Sci. vol. 55(8-9): pp. 1113-28, 1999; Chou J.J. et al., Cell vol. 94(2): pp. 171-80, 1998).

2.2. Enzimas biocatalíticas

45 Existen numerosas aplicaciones comerciales concebibles de las proteínas estabilizadas, complejos de proteína e interacciones proteína-proteína. Como un ejemplo de una clase de proteínas para las que es deseable la estabilización, se consideran en esta sección enzimas y otras proteínas que se han utilizado como biocatalizadores en aplicaciones industriales. también se considera la evaluación del mercado de enzimas biocatalíticas.

50 Los procesos biocatalíticos industriales tienen uso en muchos sectores industriales, que incluyen las industrias química, de detergentes, farmacéutica, agrícola, de alimentos, cosméticos, textiles, procesamiento de materiales, y de papel. Dentro de estas industrias, los biocatalizadores tienen muchas aplicaciones, que varían desde la síntesis de producto (por ejemplo, fabricación de aminoácidos), uso como agentes activos en ciertos productos (por ejemplo,

polvos de lavado biológico), uso en equipos de pruebas de diagnóstico, y uso como agentes terapéuticos. Las ventas totales de biocatalizadores industriales en 1999 fueron de alrededor de \$1.4 miles de millones. Se espera que esta figura crezca significativamente durante la siguiente década cuando se habiliten aplicaciones de biocatalizadores mediante tecnologías novedosas tal como la invención descrita aquí.

5 Se cree que los sectores comerciales que tienen potencial de crecimiento e innovación tecnológica incluyen enzimas diseñadas por ingeniería (por ejemplo, para proporcionar rendimiento más rápido, producción más barata, y/o la capacidad de producir productos novedosos), sistemas de control de la polución (por ejemplo, para biorremediación), y sistemas biocatalíticos no acuosos (por ejemplo, para bioprocesamiento de grasas y aceites y fabricación de fármacos) (véase Business Intelligence Center, Explorer: "BIC Explorer"; Business Opportunities in Technology Commercialization).

10 Históricamente, solo un puñado de las compañías de productos químicos finos tal como DSM, Lonza y Avecia Ltd., han abordado e invertido en los procesos biocatalíticos. Más recientemente, sin embargo, ha habido varias inversiones corporativas significativas en el campo de los biocatalizadores. Un ejemplo de tales inversiones es el anuncio reciente de Bayer de que utilizará 6-7% de las ventas de productos químicos para desarrollar procesos basados en enzimas para ciertas moléculas.

15 Los clientes principales de las compañías de química fina tienden a preferir proveedores con un amplio rango de desarrollo de procesos. Esta consideración sugiere que aquellos con experiencia en biocatalítica pueden obtener una ventaja más competitiva en el mercado. Algunas firmas han reconocido esto y tratan rápidamente de cerrar la brecha por medio de adquisiciones (por ejemplo la adquisición por Great Lakes de NSC Technologies y la compra por Cambrex de Celgene). Otros reconocen que van a perder nuevas oportunidades de negocio si no hacen algo para acceder a los conocimientos básico necesarios para la biocatálisis (Joe Blanchard, Altus Biologics Inc., 1999).

20 Los principales fabricantes de enzimas (por ejemplo Novo, Genencor, Roche, etc.) tienden a enfocarse en una producción de enzimas a gran escala para los mercados industriales principales (tal como detergentes y textiles) y no en la aplicación de enzimas para el desarrollo de productos químicos finos (Joe Blanchard, Altus Biologics Inc., 1999).

25 El crecimiento continuo en interés por el uso comercial de los biocatalizadores y la fragmentación de la industria de los biocatalizadores permitirá a las compañías grandes y pequeñas explotar biocatalizadores innovadores y los productos y procesos que los utilizan (BIC Explorer: Business Opportunities in Technology Commercialization, 1999).

30 Las aplicaciones de biorremediación pueden, en el futuro, convertirse en una de las aplicaciones de mayor importancia económica de las enzimas biocatalíticas. Por ejemplo, aproximadamente 2.3 millones de millones de galones de aguas residuales y 4.9 miles de millones de galones de residuos industriales pasan a las aguas en los Estados Unidos cada año, y aproximadamente 1 millón de galones de hidrocarburos ingresan a nuestro ambiente por día. La limpieza de hidrocarburos es un requerimiento de rutina para diversas operaciones comerciales (por ejemplo, buques petroleros, sentinas marinas, tanques de almacenamiento, combustible y carrotaques).

35 Actualmente, existen varios procesos en desarrollo que utilizan biocatalizadores para descontaminación/descomposición de hidrocarburos y aguas de desecho. No solo son estos procesos los sistemas más prometedores comercialmente debido a su eficiencia y bajos costes, sino que también son más limpios.

40 Adicionalmente, la desulfurización biocatalítica es una tecnología no costosa y atractiva para el mercado de producción de aceite crudo, en donde el aceite crudo bajo en azufre tiene un precio premium sobre el aceite crudo alto en azufre. Subsiste una necesidad creciente alrededor del mundo para el manejo de azufre y desulfurización con racionalización de costes debido a un nivel incrementado de azufre en combustibles fósiles y las regulaciones crecientemente exigentes que exigen menores emisiones de azufre. Se espera que el cumplimiento de estas regulaciones le cueste solo a la industria de refinación europea más de \$50 miles de millones en capital y \$10 miles de millones anualmente en gastos de operación.

45 Toda la fabricación de catalizadores en 1997 representa un mercado de más de \$10 miles de millones en los Estados Unidos, una cifra citada por la Sociedad Química Americana (véase también, "Catalyst Industry Stresses Need for Partners as Key to Future Success," C&E News, Julio 11, 1994; CatCon '96 presentations by T. Ludermann of CONDEA Chemie GmbH, Paul Lamb of Englehard Corporation, y J. Ohmer y K. Herbert de Degussa Corporation). De acuerdo con Maxigen, el mercado de enzimas industriales total (un segmento del mercado de fabricación de catalizadores) se estima en \$1.4 miles de millones hoy en día, con un crecimiento de alrededor de 10% anualmente.

2.3. Estrategias de estabilización

Se conocen en la técnica y se han descrito previamente varias estrategias de estabilización de la proteína, como se resalta más adelante.

2.3.1. Estabilización de enzimas biocatalíticas

Se han abordado métodos para mejorar la estabilidad de los biocatalizadores. Al nivel de las proteínas, los métodos más promisorios incluyen el descubrimiento de biocatalizadores estables a partir de la investigación de organismos termófilos, evolución directa, e ingeniería de proteínas y computacional, como se describe adelante.

- 5 Se buscan organismos termófilos, o 'extremófilos', en ambientes extremos tal como respiraderos en aguas profundas y géiseres en Yellowstone. Aunque se han identificado entre ellos las enzimas de relevancia comercial, esta metodología de 'descubrimiento' está limitado por lo que se pueda encontrar en la naturaleza. Este método no ha producido muchos biocatalizadores termoestables comercialmente relevantes como se esperaba y/o se proyectó inicialmente.
- 10 Las técnicas de 'evolución dirigida' son métodos poderosos capaces de generar enzimas estabilizadas, frecuentemente también con especificidades funcionales alteradas/mejoradas. Sin embargo, el método está limitado por la factibilidad del procedimiento de selección.

15 Los algoritmos que calculan las fuerzas intramoleculares dentro de las proteínas están siendo utilizados para diseñar y/o hacer evolucionar enzimas con termoestabilidad mayor *in silico*. Este método está aún severamente obstaculizado por el entendimiento limitado de las fuerzas intramoleculares y los procesos involucrados en el plegamiento de las proteínas.

20 La adición de modificaciones químicas que pueden mantener las proteínas en su conformación correcta se denomina frecuentemente ingeniería de proteínas. Tales métodos de ingeniería de proteínas incluyen derivación (por ejemplo PEGilación, adición de sacarosa polimérica y/o dextrano, metoxipolietilenglicol, etc.) y métodos antiguos de entrecruzamiento de proteína (por ejemplo producción de cristales de enzima entrecruzados o CLEC). Desafortunadamente, estos métodos son frecuentemente ineficaces o provocan enormes pérdidas en la actividad.

25 Las estrategias para la estabilización operacional de los biocatalizadores que se han probado exitosamente en algunos aspectos incluyen (a) inmovilización del catalizador y (b) el uso de solventes orgánicos en el medio de reacción (denominado ingeniería del medio). La estabilidad térmica luego de la inmovilización es el resultado de la rigidez molecular y la creación de un microambiente protegido. Los métodos incluyen adhesión covalente multipunto y atrapamiento en gel. La inmovilización de los biocatalizadores es la estrategia más utilizada puesto que se obtienen beneficios adicionales, tales como flexibilidad del diseño del reactor, y recuperación de producto facilitada sin contaminación del catalizador. Sin embargo, a pesar de su gran potencial tecnológico, pocos procesos a gran escala utilizan enzimas inmovilizadas. Surgen frecuentemente restricciones severas en escala ascendente debido a

30 costes adicionales, pérdidas de actividad, y asuntos que se relacionan con la difusión.

El propósito principal de la ingeniería del medio en la biocatálisis era originalmente utilizar enzimas hidrolíticas comerciales robustas en síntesis orgánica. Sin embargo, la termoestabilidad mejorada en el medio orgánico ha demostrado un beneficio significativo y adicional. Se parte de la hipótesis de que la sustitución parcial o sustitución casi total del agua es beneficiosa debido a que el agua se involucra en la inactivación de la enzima. Cualquiera que sea el mecanismo, se han reportado recientemente numerosos casos en donde la estabilización notable de la enzima se ha obtenido en un medio orgánico tal como poliglicoles y glicinas. A pesar de este avance, no es probable que la ingeniería del medio resuelva todos los problemas de estabilidad de la biocatálisis.

35

40 Algunas de las soluciones más promisorias para los problemas de la biocatálisis han combinado métodos evolutivos con técnicas de estabilización operacional, tal como la utilización de la evolución directa para generar enzimas con índices de reacción mayores en solventes orgánicos. Tales métodos combinados pueden proporcionar sinergias significativas que se mejoran al máximo y facilitan procesos biocatalíticos comercialmente relevantes. En principio, la invención descrita de aquí adelante se puede aplicar en combinación con cualquiera de los métodos de estabilización conocidos anteriormente mencionados.

2.3.2. Estabilización de otras proteínas

45 Las técnicas de biología molecular han hecho posible estabilizar algunas proteínas mediante, por ejemplo, ingeniería de proteínas de fusión. Tales proteínas de fusión han exhibido incluso funcionalidades novedosas. Para hacer una proteína de fusión, se crea una construcción única de ácido nucleico que dirige la expresión de los dominios modulares derivados de por lo menos dos proteínas como una proteína. Debido a la fusión, se pueden mantener dos dominios en proximidad muy cercana uno del otro, haciendo por lo tanto la concentración local de cada dominio muy

50 alta con respecto a la otra. En esta forma, se estabiliza un complejo funcional. Por ejemplo, homo- y heterodímeros de la familia de la interleucina 8 se han estabilizado de esta forma, manteniendo una funcionalidad similar al tipo natural (Leong S.R. et al. Protein Sci.; vol. 6(3): pp: 609-17, 1997). Otro ejemplo de los complejos de proteína estabilizados en esta forma es el método para estabilizar los fragmentos Fv de inmunoglobulina, que consisten en los dominios variables de las cadenas liviana y pesada de inmunoglobulina, que carecen de efecto estabilizante de

los enlaces disulfuro intercadena. Es necesario para estabilizar el complejo por otros medios mantener la afinidad del complejo de inmunoglobulina, y uno de los métodos utilizados es la expresión de ambos polipéptidos como una cadena única (Pluckthun y P. Pack. *Immunotechnology*; vol. 3(2): pp. 83-105, 1997).

5 Sin embargo, en el diseño de los reactivos farmacológicos, es frecuentemente desventajoso crear proteínas de fusión que requieran una secuencia ligadora para estabilizarlas. Por ejemplo, tales ligadores introducen no-auto epítomos que son reconocidos frecuentemente por el organismo como respuestas inmunes provocadas o externas. Esto reduce la eficiencia de tales terapias y/o diagnósticos debido a que los reactivos son eliminados luego por el sistema inmune (ver, por ejemplo, Raag R. y Whitlow M. *FASEB*; vol. 9: pp. 73-80, 1995).

10 En el caso de fragmentos Fv de cadena sencilla, el ligador, que se selecciona más frecuentemente por ser una estructura más flexible, le permite al complejo disociarse, debido a que es baja la afinidad de los dos polipéptidos para cada uno. Los fragmentos Fv de cadena única luego se agregan, o aglomeran, y por lo tanto pierden su funcionalidad (Webber K.O. et al. *Mol. Immunol.*; vol. 32(4): pp. 249-258, 1995). Los ligadores más rígidos que le prestan al complejo más estabilidad, y pudieran por lo tanto reducir el nivel o velocidad de agregación y pérdida de la funcionalidad, se asocian con inmunogenicidad incrementada (Raag R. and Whitlow M. *FASEB*; vol. 9: pp. 73-80, 15 1995).

El entrecruzamiento de los dominios en sitios de contacto cercano eludirían estos problemas, donde es posible dirigir el entrecruzamiento entre las dos proteínas para tales superficies de la proteínas en donde después de la reacción se detiene el entrecruzamiento. Un tal medio es estabilizar los complejos al introducir un enlace de disulfuro entre dos polipéptidos al introducir mutaciones puntuales a cistina en ambas cadenas de polipéptido. Las mutaciones se introducen en posiciones que permiten la formación de tales enlaces (véase, por ejemplo, Reiter Y. et al. *Nat Biotech.*; vol. 14: pp. 1239-1245, 1996; Pastan et al. Patente Estadounidense No. 5,747,654, presentada el 5 de mayo de 1998).

20 Los enlaces disulfuro son, sin embargo, inestables bajo muchas condiciones fisiológicas (Klinman J.P. (ed). *Methods in Enzymology*; vol. 258, 1995). Las condiciones fisiológicas varían ampliamente, por ejemplo con respecto al potencial redox (oxidación vs. reducción) y acidez (pH alto vs. bajo) de los diversos medios fisiológicos (intracelular, extracelular, vesículas pinocitosis, lumen gastro-intestinal, etc.). Los enlaces disulfuro se encuentran en la naturaleza solo en las proteínas extracelulares, y se sabe que se deshacen en ambientes reductores, tales como el medio intracelular. Pero aún en el medio ambiente extracelular, muchos enlaces disulfuro construidos por ingeniería son inestables.

30 Varias otras metodologías de entrecruzamiento químico permiten la formación de enlaces que son inestables bajo un amplio rango de pH fisiológico y no fisiológico y condiciones redox. Sin embargo, con el fin de mantener la actividad y especificidad del complejo, es necesario que el entrecruzamiento se dirija y controle específicamente de tal manera que, primero, la estructura de la proteína se interrumpa mínimamente, y segundo, que el entrecruzamiento se haga en el complejo de proteína de manera que no sea inmunogénica. Pero con la mayoría de metodologías de entrecruzamiento, el grado en el que es posible dirigir el enlace a un sitio específico es muy limitado como para permitirles ser utilizadas para la mayor parte de aplicaciones biofarmacéuticas y/o diagnósticas. Ejemplos de tales metodologías de entrecruzamiento incluyen entrecruzamiento por UV, y tratamiento de proteína con formamida o glutaraldehído.

2.3.3. Fragmentos Fv

40 Los fragmentos Fv de inmunoglobulina comprenden otro ejemplo de una clase de proteínas para la cual es deseable la estabilización. Los fragmentos Fv de inmunoglobulina son los fragmentos más pequeños de los complejos de inmunoglobulina mostrados para unirse a un antígeno. Los fragmentos Fv consisten en las regiones variables de las cadenas pesada y liviana de inmunoglobulina y tienen aplicabilidad amplia en las configuraciones farmacéuticas e industriales.

45 Valor del Mercado de Fragmentos Fv

Un análisis reciente estima que 20 a 40 por ciento de todos las terapias y diagnósticos biotecnológicos actualmente en desarrollo se basan en la inmunoglobulina (Pharmaceutical Research and Manufacturers of America. *New Medicines in Development, Survey*. 1998). Adicionalmente, una porción significativa, y la mayoría de las terapias y diagnósticos del "estado de la técnica" actual con base en Ig en desarrollo se basan en el fragmento Fv (Price Waterhouse: *Survey of Biopharmaceutical Industry*, 1998). Para las revisiones de la utilidad de la inmunoglobulina como un agente farmacológico, ver Penichet M.L. et al., *Hum Antibodies*; vol. 8(3): pp. 106-18, 1997; Sensel M.G. et al. *Chem. Immunol.*; vol. 65: pp. 129-58, 1997; Reiter Y. y Pastan I. *TIBTECH*; vol. 16(12): pp. 513-520, 1998; Reiter Y. et al. *Nat Biotech.*; vol. 14: pp. 1239-1245, 1996; Pluckthun and P. Pack. *Immunotechnology*; vol. 3(2): pp. 83-105, 1997; Wright A. and Morrison S.L. *Trends Biotechnol.*; vol. 15 (1): pp. 26-32, 1997; Schwartz M.A. et al. *Cancer*

Chemother. Biol. Response Modif.; vol. 13:pp. 156-74, 1992; Houghton A.N. and Scheinberg D.A. Semin Oncol.; vol. 13(2): pp. 165-79, 1986; y Cao Y. and Suresh M.R. Bioconjugate Chemistry; vol. 9(6): pp. 635-644, 1998.

5 Luego de la introducción exitosa del primer fármaco biotecnológico basado en Ig, ReoPro por Centocor, en 1994, se aprobaron seis fármacos más con base en Ig en 1997 y 1998 y seis más están en ensayos clínicos fase III al final de 1998. Las ventas de un producto individual con base en inmunoglobulina clínicamente exitoso, puede dar como resultado ingresos anuales del orden de varios cientos de millones de dólares (Pharmaceutical Research and Manufacturers of America. New Medicines in Development, Survey, 1998). Juntos, estos hechos dan evidencia del valor clínico y comercial de estos tipos de productos.

10 El coste de desarrollar, producir y probar clínicamente tales productos es, sin embargo, inmenso y el riesgo de fallo es frecuentemente grande. Debido a esto, cualquier tecnología que pueda incrementar la efectividad del producto, ampliar su rango de aplicaciones o incrementar sus oportunidades de éxito en ensayos clínicos sumará enormemente al Valor Presente Neto de un producto en desarrollo (Boston Consulting Group: The Contribution of Pharmaceutical Companies: What's at stake for America, 1993).

Métodos de estabilización del fragmento Fv

15 Hasta la fecha, se ha empleado una variedad de metodologías para estabilizar los anticuerpos construidos por ingeniería. Primero, la introducción de enlaces disulfuro adicionales se ha desarrollado a través de la manipulación biológica molecular de la construcción que expresa el anticuerpo (Reiter Y. and Pastan I. TIBTECH; vol. 16(12): pp. 513-520, 1998). Segundo, se ha empleado la introducción de un ligador que permite que ambos fragmentos sean expresados como una cadena única (fragmentos Fv de cadena sencilla) (Pluckthun and P. Pack. Immunotechnology; vol. 3(2): pp. 83-105, 1997; Cao Y. y Suresh M.R. Bioconjugate Chemistry; vol. 9(6): pp. 635-644, 1998). Finalmente, se ha llevado a cabo la fusión de un dominio de di u oligomerización exógeno a cada una de las cadenas de fragmento Fv (Pluckthun and P. Pack. Immunotechnology; vol. 3(2): pp. 83-105, 1997; Cao Y. y Suresh M.R. Bioconjugate Chemistry; vol. 9(6): pp. 635-644, 1998; ver también Antibody Engineering Page, IMT, University of Marburg, FRG: http://aximtl.imt.uni-marburg.de/_rek/indexfenster.html).

25 Sin embargo, todas estas tecnologías tienen inconvenientes significativos. Los enlaces disulfuro son un enlace adecuado en el contexto de los fragmentos Fab (ver Figura 1D), y muchas otras proteínas extra-celulares, para estabilizar los complejos de proteína. Adicionalmente la introducción de enlaces disulfuro evita la necesidad de introducir péptidos externos, y los complejos estabilizados resultantes son mínimamente inmunogénicos. No obstante, la introducción de enlaces disulfuro en los fragmentos Fv mediante medios biológicos moleculares da como resultado complejos que son insuficientemente estables bajo muchas condiciones fisiológicas, comercialmente relevantes, tales como el medio intracelular y algunas veces aún suero. Como tales tienen utilidad limitada en el contexto farmacéutico.

35 Con los fragmentos Fv de cadena individual existe una compensación entre la estabilidad del complejo y su inmunogenicidad en un contexto terapéutico o diagnóstico in vivo. Los ligadores que resultan en conjugados estables son estructuras más rígidas, y provocan respuestas inmunes, que a su vez dan como resultado una utilidad reducida. Los ligadores que no son inmunogénicos son de manera general los ligadores más flexibles que proporcionan estabilidad insuficiente (ver anterior, Raag R. and Whitlow M. FASEB; vol. 9: pp. 73-80, 1995).

40 Los fragmentos Fv estabilizados mediante fusión a dominios de multimerización son significativamente inmunogénicos, y carecen de las ventajas más significativas de los fragmentos Fv en primer lugar: tamaño reducido y penetración de tejido incrementada resultante.

Otros métodos de entrecruzamiento químico actualmente disponibles, tales como entrecruzamiento por UV (véase más arriba), se limitan severamente en el grado en el que es posible dirigir el enlace a un sitio específico. Como las aplicaciones biofarmacéuticas y/o diagnósticas requieren el mantenimiento de la función de los polipéptidos, la especificidad en la reacción de entrecruzamiento es de suma importancia.

45 2.4. El entrecruzamiento oxidativo tirosilo-tirosilo

50 Se ha demostrado que las reacciones de entrecruzamiento oxidativas entre las cadenas laterales tirosilo ocurren de forma natural. Por ejemplo, se ha demostrado que el compuesto I citocromo c peroxidasa forma enlaces di-tirosina durante la reducción endógena de su sitio activo (Spangler B.D. and Erman J.E. Biochim. Biophys. Acta; vol. 872(1-2): pp. 155-7, 1986), y se informa que los dímeros ligados a di-tirosina de gama B-cristalina se asocian con cataractogenia de los lentes del ojo. In vitro, los enlaces proteína-proteína de di-tirosina se forman fácilmente por vía fotodinámica en la presencia de sensibilizadores (Kanwar R. and Balasubramanian D. Exp. Eye Res.; vol. 68(6): pp. 773-84, 1999). Adicionalmente, el entrecruzamiento de la proteína a través de la formación de enlaces di-tirosina se puede catalizar, por ejemplo, mediante peroxidasa (Gmeiner B. and Seelos C. FEBS Lett ; vol. 255(2): pp. 395-7, 1989), o mediante complejos metalo-ión (Campbell et al. Bioorganic and Medicinal Chemistry, vol. 6: pp. 1301-1037,

1998; Brown K.C. et al. Biochem.; vol. 34(14): pp. 4733-4739, 1995), y mediante oxidantes activados por luz (Fancy D.A. and Kodadek T. Proc. Natl. Acad. Sci., U.S.A.; vol. 96: pp. 6020-24, 1999).

5 Como describe Campbell et al., en la presencia de un catalizador apropiado y un reactivo de oxidación apropiado, puede ocurrir una reacción de entrecruzamiento oxidativa entre las cadenas laterales tirosilo de proteínas que se espacian apropiadamente. En esta reacción, los grupos hidroxilo de las cadenas laterales tirosilo reaccionan una con la otra, se libera una molécula de H₂O, y las cadenas laterales se unen mediante un enlace covalente. Se considera que esta reacción procede a través de un complejo metalo-oxo de alta valencia que sustrae un electrón de una cadena lateral tirosilo accesible, seguido por el acoplamiento covalente del radical tirosilo resultante con otra cadena lateral tirosilo que está en proximidad suficiente.

10 Esta metodología de entrecruzamiento se desarrolla originalmente para entrecruzar las proteínas que interactúan en los lisados celulares, como un *proxy* para la situación in vivo, para permitir el estudio de la funcionalidad de las proteínas al identificar otras proteínas que interactúan. La reacción solo ocurre con cadenas laterales tirosina que están en proximidad muy cercana una a la otra. Adicionalmente, el enlace formado entre las cadenas laterales tirosilo es irreversible y estable bajo un rango muy amplio de condiciones fisiológicas.

15 Ninguna de las referencias citadas anteriormente describe o sugiere métodos que utilicen el entrecruzamiento de ditirosilo para la formación de reticulaciones químicas en profundidad para estabilizar un complejo de proteína mientras que mantiene las actividades y especificidades del complejo. De acuerdo con lo anterior, subsiste una necesidad para tales métodos en donde el producto sea funcional bajo un amplio rango de condiciones fisiológicas y no fisiológicas, y en donde se mantenga la estructura, función, y especificidad del complejo de proteína entrecruzado.

20 La citación o identificación de cualquier referencia en la Sección 2 o en cualquier otra sección de esta aplicación no se debe entender como una admisión de que tal referencia está disponible como técnica anterior para la presente invención.

3. Resumen de la invención

25 Esta invención proporciona un método para la estabilización de un polipéptido o complejo de polipéptido, mediante la introducción de enlaces di-tirosina intra-polipéptido y/o inter-polipéptido, que mantienen simultáneamente la estructura y función del polipéptido o complejo de polipéptido. Adicionalmente, esta invención se relaciona con diversos métodos para optimizar la estabilización de la proteína. Tales métodos incluyen análisis estadístico de las secuencias de aminoácidos primarias de las proteínas relacionadas (análisis de datos bidimensionales) y se considera que los análisis estadísticos de las coordenadas tridimensionales de las proteínas se relaciona con la estructura tridimensional (análisis de datos tridimensionales).

35 Adicionalmente, esta invención se relaciona con polipéptidos estabilizados y complejos de polipéptido. Para lograr la estabilización, la reacción de entrecruzamiento se controla cuidadosamente de tal manera que los polipéptidos y los complejos de polipéptido mantienen su funcionalidad original. En un aspecto, la invención se relaciona con un método para la identificación de los residuos de aminoácidos que, cuando se entrecruzan, son menos problemáticos para la estructura y función del polipéptido o complejo de polipéptido. En otro aspecto, la invención se relaciona con un método para mutagénesis de los residuos identificados para controlar adicionalmente la reacción de entrecruzamiento. Los polipéptidos y complejos de polipéptido así estabilizados se pueden utilizar bajo una amplia variedad de condiciones fisiológicas y no fisiológicas. Adicionalmente, la metodología de entrecruzamiento descrita aquí puede eliminar la necesidad de la adición de estructuras exógenas para proteínas y complejos construidos por ingeniería, tal como ligadores de péptidos. La invención también se relaciona con un método para el análisis estadístico de las bases de datos de información de secuencia y/o estructurales disponibles para los polipéptidos y complejos de polipéptido que van a ser estabilizados. El análisis estadístico identifica los pares de residuos adecuados que son probablemente menos problemáticos para la estructura y función cuando se entrecruzan.

40 Adicionalmente, en una cadena o cadenas de polipéptido que se van a entrecruzar, se pueden alterar cadenas laterales reactivas potencialmente indeseables utilizando mutagénesis dirigida al sitio, por ejemplo, para introducir una mutación puntual máximamente conservadora que no soportará la reacción de entrecruzamiento. Las condiciones de reacción de entrecruzamiento también se pueden ajustar para evitar entrecruzamientos indeseados. En los residuos identificados como posiciones deseables para el entrecruzamiento, se pueden introducir cadenas laterales reactivas mediante mutagénesis dirigida a sitio, y se lleva a cabo la reacción de entrecruzamiento utilizando las condiciones identificadas anteriormente.

4. Breve descripción de las figuras

La presente invención se puede entender más completamente por referencia a la siguiente descripción detallada, ejemplos ilustrativos de las realizaciones específicas y las figuras adjuntas.

FIGURA 1. El entrecruzamiento de ditirosilo y las proteínas de ejemplo se puede estabilizar de acuerdo con los métodos de la invención. A. Representación esquemática de un entrecruzamiento ditirosilo. La adición de un catalizador de entrecruzamiento y un reactivo de oxidación en una preparación de proteína o complejo de proteína en donde por lo menos dos residuos tirosina ocurren en proximidad cercana y en la orientación apropiada resulta en un entrecruzamiento ditirosilo y una molécula de agua. B. Representación esquemática del plegado canónico de las hidrolasas a/b, un grupo de enzimas que incluye lipasas. Las posiciones topológicas de los residuos de sitio activo se indican como círculos sólidos. De K.-E. Jaeger et al., 1999, Ann. Rev. Microbiol. 53, 315-351. C. Representación esquemática de la estructura secundaria de la lipasa B de Candida antarctica. Las posiciones topológicas de los residuos de sitio activo se indican como los residuos S105, D 187, y H224. De J. Uppenberg et al., 1994, Structure 2, 293-308. D. Representación esquemática de una molécula de inmunoglobulina (IgG). El hetero-tetrámero de inmunoglobulina comprende dos cadenas ligeras idénticas, y dos cadenas pesadas idénticas. El complejo se estabiliza mediante enlaces disulfuro inter-cadena; los enlaces disulfuro se indican por los enlaces "S-S" en la representación esquemática. Ambos dominios de unión a antígeno, uno en el extremo del "tenedor", consiste de un par de regiones variables de cadena ligera y pesada, y se denominan como los "fragmentos Fv". El dominio de unión a antígeno es el fragmento Fv, consiste de la región variable de la cadena ligera y pesada de cuatro Regiones de Estructura Principal relativamente conservadas que proporcionan la estructura general, y de tres Regiones Determinantes de Complementariedad que prestan al fragmento Fv su especificidad para un antígeno específico. El fragmento Fab, que comprende las regiones variables de cadena pesada y ligera (V1 & Vh), la región constante de cadena ligera (Cl), y la primera región constante de la cadena pesada (Ch1), se estabilizan mediante un enlace disulfuro inter-cadena. En el fragmento Fv ninguno de los enlaces disulfuro intra-cadena de inmunoglobulina están presentes, como se indica, dando como resultado que el requerimiento para este complejo de proteína sea estabilizado artificialmente.

FIGURA 2. A. Representación esquemática de una cadena lateral tirosilo, que consiste en un carbono alfa (A) que es aún parte de la estructura principal del polipéptido, un carbono beta (B), el primer átomo en la cadena lateral que no es parte de la estructura principal, un anillo aromático, que, a su vez, consiste de seis átomos de carbono, y un grupo hidroxilo (OH). Se indica el ángulo β en el carbono beta entre el eje de oxígeno de beta-hidroxilo del carbono y el enlace de carbono de carbono alfa-beta. B. Representación esquemática de un enlace tirosilo-tirosilo indica adicionalmente el ángulo ω , que es el ángulo entre el enlace ditirosilo y el enlace carbono-carbono en el anillo aromático de la cadena lateral tirosilo entrecruzada que está próxima al carbono beta de la misma cadena lateral, se proyecta en el plano dos de los dos anillos aromáticos. También se indican el ángulo α , el ángulo entre todos los residuos de carbono en el plano de los anillos aromáticos (120°), y los grados de libertad rotacional (1) en el enlace ditirosina en sí mismo, y (2), del carbono alfa alrededor del eje de carbono beta-carbono gamma (más próximo al átomo de carbono en el anillo aromático). C. Ángulos tridimensionales formados por los ejes de carbono alfa-carbono alfa, el carbono beta (ψ y ϕ), y los dos planos (χ) descritos por el eje carbono alfa-carbono alfa y (1) el enlace carbono alfa-carbono beta de la primera cadena (A1-B1), y (2) el enlace de carbono alfa-carbono beta de la segunda cadena (A2-B2).

FIGURA 3. El ángulo ω , indicado en la Figura 2B, es $+120^\circ$. Para esta configuración, las distancias del carbono alfa, los ángulos ψ y ϕ , y las diferencias de la distancia alfa-beta (ver texto) se representan geoméricamente para configuraciones máximas y mínimas (que caen dentro de un plano), da este ángulo ω . Se asume que el ángulo β es 109.5° , el ángulo tetrahédrico de los átomos de carbono, y se asume libertad rotacional completa del carbono alfa alrededor del eje de carbono beta-carbono gama. En A, la longitud c es la distancia entre los dos átomos de carbono de un enlace carbono-carbono; la longitud v es $\cos((180^\circ-\alpha)/2) \times c$, la longitud h es $\sin((180^\circ-\alpha)/2) \times c$, la longitud a es la mitad de la raíz cuadrada de la suma de $7v$ cuadrado y h cuadrado, y la longitud b es la raíz cuadrada de la suma del cuadrado de $(a+v)$ y h cuadrado. En B, v es el $\cos(180^\circ-(\beta-(180^\circ-\alpha)/2+\arctan(h/7v))) \times c$, h es el $\sin(180^\circ-(\beta-(180^\circ-\alpha)/2+\arctan(h/7v))) \times c$, y, análogamente, la longitud a es la mitad de la raíz cuadrada de la suma de $7v$ cuadrado y h cuadrado, y la longitud b es la raíz cuadrada de la suma del cuadrado de $(a+v)$ y h cuadrado. En la configuración descrita en A, en la cual la distancia de carbono alfa es máxima, los ángulos ψ y ϕ son $(180^\circ-\alpha)/2-\arctan(h/7v)$; en la configuración en B, en la que la distancia de carbono alfa es mínima para un ángulo ω de $+120^\circ$, ψ y ϕ son $\beta-(180^\circ-\alpha)/2-\arctan(h/7v)$.

FIGURA 4. El ángulo ω , indicado en la Figura 2B, es -120° . En la Figura 4, las distancias, ángulos ψ y ϕ , del carbono alfa y las diferencias de la distancia alfa-beta (ver texto) se representan geoméricamente para las configuraciones máxima y mínima (que caen dentro de un plano), dado este ángulo ω . El ángulo β se mantiene constante a 109.5° , el ángulo tetrahédrico de los átomos de carbono, y se asume libertad rotacional completa del carbono alfa alrededor del eje de carbono beta-carbono gama. En A, la longitud x es $4v$, la longitud y es la raíz cuadrada de la suma de h cuadrado y $3v$ cuadrado, la longitud z es el $\cos(180^\circ-120^\circ+\arctan(h/3v)) \times y$, la longitud a es la mitad de la raíz cuadrada de la suma de $(x+z)$ cuadrado y y cuadrado, la longitud v es el $\cos(120^\circ-\beta) \times c$, y la longitud b es la suma de las longitudes a y v . En B, la longitud v es el $\cos(\beta-2 \times (180^\circ-\alpha)/2) \times c$, y la longitud b es la diferencia de las longitudes a y v . En la configuración descrita en A, en la que la distancia de carbono alfa es máxima para un ángulo ω de $+120^\circ$, ψ y ϕ son $\alpha-\beta$; en la configuración en B, en la que la distancia del carbono alfa es mínima, ψ y ϕ son $180^\circ-(\beta-2 \times (180^\circ-\alpha)/2)$.

FIGURA 5. Datos Coordinados Estructurales, datos primarios (o entrada) de una base de datos 3-D. Primero dos residuos de aminoácidos de una cadena ligera (L) y cadena pesada (H) del fragmento Fv representativo, en Angstroms; los datos de cada átomo se representan en filas, los átomos se enumeran en las columnas. Los datos coordinados se representan para todos los átomos de residuo diferentes de los átomos de hidrógeno, que incluyen aquellos involucrados en la estructura principal del polipéptido y aquellos en la cadena lateral del aminoácido. En la columna a mano izquierda, bajo el encabezado "Cadena", se enumera la identidad de la cadena de polipéptidos, con la cual se asocian las coordenadas del átomo. Un fragmento Fv consiste en dos polipéptidos: una cadena pesada (H; por debajo) y una cadena ligera (L; por encima.). El número bajo el encabezado "K&W" indica la posición del residuo del átomo dentro del sistema de alineación Kabat & Wu (K&W). Bajo el encabezado "Átomo", se indica la identidad de un átomo del aminoácido específico presente en el polipéptido representativo en este residuo particular (identificado bajo el encabezado "Aminoácido" en código de tres letras). Los coordinados tridimensionales x, y, y z de cada átomo se representan en las columnas a mano derecha, según se indica.

FIGURA 6. Representación esquemática de 3 entradas de fragmento Fv actual en una base de datos 3-D. Las disposiciones de los datos de coordenadas carbono alfa de los residuos de cadena ligera y pesada de los fragmentos Fv, y, como un ejemplo de los datos derivados relevantes, se calculan distancias inter-atómicas inter-cadena. Los datos de carbono alfa de cadena pesada se representan en filas, como se describe en la descripción de la Figura 5, y se transponen los datos de carbono alfa de cadena ligera, y los datos de cadena ligera descritos en la Figura 5 se representan en las columnas. Los datos derivados que describen las relaciones 3-D inter-cadena de los átomos en ambas cadenas se representan en la intersección de cada fila de cadena pesada y columna de cadena ligera.

FIGURA 7. Mediciones estadísticas en una base de datos 3-D de las distancias de carbono alfa entre los pares de residuo de cadena ligera y pesada del fragmento Fv, como un ejemplo de datos derivados relevantes. A. Mediciones estadísticas ilustrativas de las distancias de carbono alfa entre los pares de residuo de las tres cadenas pesada ligera y pesada del fragmento Fv representativo en la descripción de la Figura 6 (es decir datos mostrados para n=3). B. Mediciones estadísticas reales de las distancias carbono alfa entre los pares de residuos de todas las cadenas ligera y pesada del fragmento Fv en la muestra de los fragmentos Fv utilizados para la selección (datos mostrados para n=17).

FIGURA 8. Representación esquemática de una entrada de fragmento Fv (Fragmento Fv 1 de la Figura 6) dentro de una base de datos 3-D. Disposiciones de los datos de coordenadas carbono beta de los residuos de cadena ligera y pesada del fragmento Fv, y, como un ejemplo de los datos derivados relevantes, se calculan distancias inter-atómicas inter-cadena. Los datos carbono beta de cadena pesada se representan en filas, y los datos carbono beta de cadena ligera se transponen y se representan en columnas, como se describe en la descripción de la Figura 5. Los datos derivados descritos inter-cadena, las relaciones 3-D de los átomos en ambas cadenas se representan en la intersección de cada fila de cadena pesada y cada columna de cadena ligera.

FIGURA 9. Representación esquemática del método usado para calcular las diferencias entre las distancias interatómicas, carbono alfa y carbono beta inter-cadena del par de residuos ('diferencias de distancia alfa-beta') para un fragmento individual Fv en la base de datos 3-D (Fragmento Fv 1 de la Figura 6 y 8). Los datos de carbono beta (medio) y alfa- de cadena pesada (superior) se representan en filas, y se transponen los datos carbono alfa y beta de cadena ligera, y se representan en columnas, como se describe en la descripción de la Figura 5. Los datos derivados que describen las distancias inter-atómicas inter-cadena en los paneles medio y superior, y las diferencias de la distancia alfa-beta en el panel de fondo, se representan en la intersección de cada fila de cadena pesada y columna de cadena ligera.

FIGURA 10. Datos de diferencia de distancia alfa-beta, derivados como se describe en la Figura 9, de los fragmentos Fv representativos (fragmentos Fv 1, 2, y 3 de la Figura 6) en una base de datos 3-D. Los residuos de cadena ligera y pesada se representan en disposiciones, en donde los residuos de cadena pesada se enumeran verticalmente, y los residuos de cadena liviana se enumeran horizontalmente. Los datos correlacionados con los residuos de cadena pesada y ligera se representan en la intersección de cada fila de cadena pesada y columna de cadena ligera.

FIGURA 11. Mediciones estadísticas en una base de datos en 3-D de las diferencias de distancia alfa-beta de los pares de residuo de cadena ligera y pesada del fragmento Fv, como un ejemplo de datos derivados relevantes. A. Mediciones estadísticas ilustrativas de las diferencias de la distancia alfa-beta de los pares entre las tres cadenas ligera y pesada de fragmento Fv representativo en la Figura 6 (es decir los datos mostrados para n=3). B. Mediciones estadísticas reales de las diferencias de la distancias alfa-beta de los pares entre todas las cadenas ligera y pesada del fragmento Fv en las muestras de fragmentos Fv utilizados para selección (datos mostrados para n=17).

FIGURA 12. Cuantificación de propiedades físicas de la cadena lateral de aminoácidos, como un ejemplo de datos derivados relevantes, en los residuos (los primeros cuatro, representativos) de la cadena pesada del fragmento Fv,

con base en los datos de secuencia de polipéptido del fragmento Fv, compilados en una base de datos en 2-D. A. Datos de la Secuencia de Aminoácidos. Representación de datos primarios compilados en una base de datos 2-D. Los aminoácidos (AA) que se presentan en cada residuo se clasifican mediante la frecuencia (F) de su ocurrencia en este residuo específico. B. Tablas de Cuantificación de Cadena Lateral de Aminoácidos. Representación de valores numéricos utilizados en una base de datos 2-D para obtener datos derivados relevantes al cuantificar las propiedades físicas de los aminoácidos: por ejemplo volumen de van der Waals [A3] (Richards, F.M.) y los valores de hidrofobicidad numérica (Eisenberg, D.). C. Cuantificación de las propiedades físicas, ejemplificadas aquí por volúmenes de van der Waals, de las cadenas laterales de aminoácidos en cada residuo en la muestra de las secuencias de fragmento Fv en la base de datos en 2-D.

FIGURA 13. Mediciones estadísticas en una base de datos en 2-D de propiedades físicas de la cadena lateral en cada residuo de las cadenas pesadas del fragmento Fv presentes en la base de datos 2-D (muestra), como un ejemplo de los datos derivados relevantes, cuantificadas como se describe en la descripción de la Figura 12. En la tercera columna de la izquierda, bajo el encabezado "Cons", se enumera el consenso, o el aminoácido de ocurrencia más frecuente para cada residuo representado. Como mediciones estadísticas representativas, se muestran las desviaciones estándar y promedio, ambas ponderadas y no ponderadas por la frecuencia de cada ocurrencia del aminoácido en la muestra en cada residuo representado en esta figura. A. Se muestran las desviaciones estándar y promedio para los volúmenes van der Waals del residuo, ambas ponderadas y no ponderadas mediante la frecuencia de cada ocurrencia de aminoácido en la muestra en cada residuo representado en esta figura. B. Se muestran las desviaciones estándar y promedio para las cantidades de hidrofobicidad del residuo, ambas ponderadas y no ponderadas mediante la frecuencia de cada ocurrencia de aminoácido en la muestra en cada residuo representado, en esta figura.

FIGURA 14. Ilustración esquemática de una disposición sucesiva y una disposición paralela de filtros diseñados para automatización utilizando un sistema de computador y software para el proceso de selección de par de residuo. Los filtros mostrados son un conjunto ilustrativo de filtros tomados a partir de los filtros descritos anteriormente (ver la Identificación de residuos Adecuados para la Reacción). En esta ilustración, el número de residuos seleccionados que "pasan" cada filtro, en sucesión (izquierda) o en paralelo (derecha), se deriva de un análisis de 106 aminoácidos de la cadena ligera de fragmento Fv, los 120 aminoácidos de la cadena pesada del fragmento Fv, y los 12720 pares de residuos posibles resultantes en un fragmento Fv dado. Los porcentajes que indican la permisividad de cada filtro también son ilustraciones del ejemplo del fragmento Fv. Ver texto para discusión adicional (Software para el Proceso de Selección).

5. Descripción detallada de la invención

La invención descrita aquí comprende métodos para estabilizar polipéptidos y complejos de polipéptidos. También se describen polipéptidos y complejos de polipéptidos estabilizados utilizando los métodos descritos. La reacción de estabilización se controla de tal manera que los polipéptidos y complejos de polipéptidos mantienen su funcionalidad original al proporcionar cadenas laterales reactivas específicamente localizadas. Los polipéptidos y complejos de polipéptido estabilizados se pueden mantener y utilizar bajo una amplia variedad de condiciones fisiológicas y no fisiológicas sin estructuras químicas exógenas que pueden ser inmunogénicas y/o reducir significativamente su eficacia.

Al tomar un método estadístico para analizar las bases de datos de la formación de secuencia y estructural para los dominios de proteínas, se pueden identificar los pares de residuos adecuados en los que la reacción de entrecruzamiento es probable que sea menos perturbadora para la estructura general.

En estos residuos, las cadenas laterales reactivas se colocan por medio de mutaciones puntuales dirigidas a sitio. En las cadenas de polipéptidos que se entrecruzan, los codones de cadenas laterales potencialmente reactivas en otras posiciones también se alteran para introducir una mutación puntual máximamente conservadora que no soportará la reacción.

5.1. POLIPÉPTIDOS Y COMPLEJOS DE POLIPÉPTIDOS ADECUADOS PARA LA APLICACIÓN DE LA INVENCION

Los polipéptidos y complejos de polipéptidos que se pueden estabilizar por los métodos descritos aquí son polipéptidos sencillos o complejos que consisten de dos o más polipéptidos y que permanecen funcionalmente activos luego de la aplicación de la invención actual. También se describen los ácidos nucleicos que codifican los polipéptidos anteriores. El término material "funcionalmente activo", como se utiliza aquí, se refiere a aquel material que exhibe una o más actividades funcionales o funcionalidades asociadas con uno o más de los polipéptidos del complejo. Tales actividades o funcionalidades pueden ser las actividades o funcionalidades originales, naturales o tipo natural de complejos de polipéptidos, o se pueden diseñar y/o construir por ingeniería. Tal diseño y/o construcción por ingeniería se puede lograr, por ejemplo, al eliminar los aminoácidos, o agregar aminoácidos a partes de uno, cualquiera, ambos, varios, o todos los polipéptidos, al fusionar los polipéptidos de diferentes

polipéptidos o complejos de polipéptido, al agregar o eliminar las modificaciones post-translacionales, al agregar modificaciones o apéndices químicos, o al introducir cualquier otra mutación mediante cualquiera de los métodos conocidos en la técnica para este fin como se establece en detalle adelante.

5 Las composiciones pueden consistir esencialmente de los polipéptidos de un complejo, y fragmentos, análogos, y derivados de las mismas. Alternativamente, las proteínas y fragmentos y derivados de las mismas pueden ser un componente de una composición que comprende otros componentes, por ejemplo, un diluyente, tal como solución salina, un portador y excipiente farmacéuticamente aceptable, un medio de cultivo, etc.

10 En aspectos específicos de la descripción, se describen los fragmentos de un polipéptido estabilizado que consiste de por lo menos 3 aminoácidos o de un complejo de polipéptidos estabilizado que consiste de por lo menos 6 aminoácidos, 10 aminoácidos, 20 aminoácidos, 50 aminoácidos, 100 aminoácidos, 200 aminoácidos, 500 aminoácidos, 1000 aminoácidos, 2000 aminoácidos, o de por lo menos 5000 aminoácidos.

5.1.1. Derivados y análogos de polipéptido

15 Los derivados o análogos de proteínas incluyen aquellas moléculas que comprenden regiones que son sustancialmente homólogas a una proteína o fragmento de la misma (por ejemplo, en varias realizaciones, por lo menos 40% o 50% o 60% o 70% o 80% o 90% o 95% de identidad sobre una secuencia de aminoácidos o ácido nucleico de tamaño idéntico o cuando se compara con una secuencia alineada en la que se hace la alineación, por ejemplo, mediante un programa de homología de computador conocido en la técnica) o cuyo ácido nucleico codificado es capaz de hibridar en una secuencia de gen codificante, bajo condiciones de alta exigencia, de exigencia moderada, o de baja exigencia.

20 Adicionalmente, se pueden sustituir uno o más residuos de aminoácidos dentro de la secuencia por otro aminoácido de una polaridad similar que actúa como un equivalente funcional, lo que da como resultado una alteración inactiva. Las sustituciones para un aminoácido dentro de la secuencia se pueden seleccionar de otros miembros de la clase a la que pertenece el aminoácido. Por ejemplo, los aminoácidos no polares (hidrófobos) incluyen alanina, leucina, isoleucina, valina, prolina, fenilalanina, triptófano y metionina. Los aminoácidos neutros polares incluyen glicina, serina, treonina, cisteína, tirosina, asparagina, y glutamina. Los aminoácidos cargados positivamente (básicos) incluyen arginina, lisina y histidina. Los aminoácidos cargados negativamente (ácidos) incluyen ácido aspártico y ácido glutámico. Tales sustituciones se entienden de manera general como sustituciones conservadoras.

30 Los derivados y análogos de los polipéptidos del complejo que va ser estabilizado mediante aplicación de la invención actual se pueden producir mediante varios métodos conocidos en la técnica. Las manipulaciones que resultan en su producción pueden ocurrir al nivel del gen o de la proteína. Por ejemplo, se puede modificar una secuencia de un gen clonado mediante cualquiera de las numerosas estrategias conocidas en la técnica.

35 Se pueden hacer polipéptidos quiméricos que comprenden uno o varios de los polipéptidos de un complejo que va a ser estabilizado por la invención actual, o fragmento, derivado, análogo del mismo (que consiste preferiblemente de por lo menos un dominio de un complejo de proteína que va a ser estabilizado, o por lo menos 6, y preferiblemente por lo menos 10 aminoácidos de la proteína) unidos en su terminal amino- o carboxi por medio de un enlace de péptido a una secuencia de aminoácidos de una proteína diferente.

40 Tal polipéptido quimérico se puede producir mediante cualquier método conocido, que incluyen: la expresión recombinante de un ácido nucleico que codifica el polipéptido (que comprende una secuencia codificante de polipéptido unida en marco a una secuencia codificante para un polipéptido diferente); ligar las secuencias de ácidos nucleicos apropiadas que codifican las secuencias de aminoácido deseadas una a la otra en una estructura codificante apropiada, y que expresa el producto quimérico; y las técnicas sintéticas de proteínas, por ejemplo, mediante el uso de un sintetizador de péptidos.

5.1.2. Manipulaciones de una secuencia de proteína en el nivel de proteína

45 Se incluyen dentro del alcance de la descripción relacionada con la invención los polipéptidos, fragmentos de polipéptido, u otros derivados o análogos, que se modifican diferencialmente durante o después de la traducción o síntesis, por ejemplo, mediante glucosilación, acetilación, fosforilación, amidación, derivación mediante grupos protectores/bloqueadores conocidos, división proteolítica, etc.

50 Se puede llevar a cabo numerosas modificaciones químicas mediante técnicas conocidas, que incluyen pero no se limitan a, división química específica mediante bromuro cianógeno, tripsina, quimiotripsina, papaína, proteasa V8, NaBH₄, acetilación, formilación, oxidación, reducción, síntesis metabólica en la presencia de tunicamicina, etc.

Adicionalmente, se pueden sintetizar químicamente los polipéptidos, fragmentos de polipéptido, u otros derivados o análogos que se pueden estabilizar utilizando los métodos de la invención actual. Por ejemplo, se puede sintetizar un péptido que corresponde a una porción de una proteína mediante el uso de un sintetizador de péptidos. Adicionalmente, si se desea, los aminoácidos no clásicos o los análogos de aminoácidos químicos se pueden introducir como sustituciones y/o adiciones en la secuencia de uno, cualquiera, ambos, varios o todos los polipéptidos del complejo.

Los aminoácidos no clásicos incluyen, pero no se limitan a, los isómeros D de los aminoácidos comunes, fluoroaminoácidos, aminoácidos diseñados tales como β -metil aminoácidos, C y γ -metil aminoácidos, N y γ -metil aminoácidos, y los análogos de aminoácido en general.

- 10 Ejemplos de aminoácidos no clásicos incluyen: ácido α -aminocaprílico, Acpa; (S)-2-aminoetil-L-cisteína MHCl, Aecys; aminofenilacetato, Afa; ácido 6-amino hexanoico, Ahx; ácido γ -amino isobutírico y ácido α -aminoisobutírico, Aiba; aloisoleucina, Aile; L-alilglicina, Alg; ácido 2-amino butírico, ácido 4-aminobutírico, y ácido α -aminobutírico, Aba; paminofenilalanina, Aphe; b-alanina, Bal; p-bromofenilalanina, Brphe; ciclohexilalanina, Cha; citrulina, Cit; β -cloroalanina, Clala; cicloleucina, Cle; p-colorfenilalanina, Clphe; ácido cisteico, Cya; ácido 2,4-diaminobutírico, Dab; 15 ácido 3-amino propiónico y ácido 2,3-diaminopropiónico, Dap; 3,4-dehidroprolina, Dhp; 3,4-dihidroxiifenilalanina, Dhphe; p-fluorofenilalanina, Fphe; ácido D-glucoseamínico, Gaa; homoarginina, Hag; δ -hidroxilisina MHCl, Hlys; DL- β -hidroxinorvalina, Hnvl; homoglutamina, Hog; homofenilalanina, Hoph; homoserina, Hos; hidroxiprolina, Hpr; p-yodofenilalanina, Iphe; isoserina, Ise; α -metilileucina, Mle; DL-metionina-S-metilsulfonipcloruro, Msmet; 3-(1-naftil) alanina, 1Nala; 3-(2-naftil) alanina, 2Nala; norleucina, Nle; N-metilalanina, Nmala; Norvalina, Nva; O-bencilserina, Obser; O-benciltirosina, Obtyr; O-etiltirosina, Oetyr; O-metilserina, Omser; O-metiltreonina, Omthr; O-metiltirosina, Omtyr; Ornitina, Orn; fenilglicina, penicillamina, Pen; ácido piroglutámico, Pga; ácido pipercolico, Pip; sarcosina, Sar; t-butilglicina; t-butilalanina; 3,3,3-trifluoroalanina, Tfa; 6-hidroxidopa, Thphe; L-vinilglicina, Vig; dihidroxocloruro de ácido (-)-(2R)-2-amino- 3-(2-aminoetilsulfonil) propanoico, Aaspa; ácido (2S)-2-amino-9-hidroxi-4,7-dioxanonanoico, Ahdna; ácido (2S)-2-amino-6-hidroxi-4-oxahexanoico, Ahoaha; ácido (-)-(2R)-2-amino-3-(2-hidroxietilsulfonil) propanoico, Ahsopa; ácido (-)-(2R)-2-amino-3-(2-hidroxietilsulfonil) propanoico, Ahspa; ácido (2S)-2-amino-12-hidroxi-4,7,10-trioxadodecanoico, Ahtda; ácido (2S)-2,9-diamino-4,7-dioxanonanoico, Dadna; ácido (2S)-2,12-diamino-4,7,10-trioxadodecanoico, Datda; (S)-5,5-difluoronorleucina, Dfnl; (S)-4,4-difluoronorvalina, Dfnv; ácido (3R)-1-1-dioxo-[1,4]tiaziana-3-carboxílico, Dtca; (S)-4,4,5,5,6,6,6-heptafluoronorleucina, Hfnl; (S)-5,5,6,6,6-pentafluoronorleucina, Pfnl; (S)-4,4,5,5,5-pentafluoronorvalina, Pfnv; y ácido (3R)-1,4-tiaziana-3-carboxílico, Tca. 25
- 30 Adicionalmente, el aminoácido puede ser D (dextrorrotario) o L (levorotario). Para una revisión de aminoácidos clásicos y no clásicos, ver Sandberg et al. (Sandberg M. et al. J. Med. Chem.; vol. 41(14): pp. 2481-91, 1998).

5.1.3. Métodos de biología molecular

Se describen los ácidos nucleicos que codifican uno o más polipéptidos estabilizados por la metodología de la invención actual. Los polipéptidos, sus derivados, análogos, y/o quimeras, del complejo se pueden hacer al expresar 35 las secuencias de ADN que los codifican in vitro o in vivo mediante cualquiera de los métodos conocidos en la técnica. Los ácidos nucleicos que codifican uno, cualquiera, ambos, varios, o todos los derivados, análogos, y/o quimeras del complejo que va a ser estabilizado mediante la metodología de la invención actual se pueden hacer al alterar la secuencia de ácido nucleico que codifica el polipéptido o los polipéptidos mediante sustituciones, adiciones (por ejemplo, inserciones) o eliminaciones que proporcionan moléculas funcionalmente activas. Las secuencias se 40 pueden dividir en sitios apropiados con las endonucleasas de restricción, seguido por modificación enzimática adicional si se desea, se aísla, y se liga in vivo o in vitro. Adicionalmente, se puede mutar una secuencia de ácido nucleico in vitro o in vivo, para crear y/o destruir las secuencias de traducción, iniciación, y/o terminación, o para crear variaciones en las regiones codificantes y/o para formar nuevos, o destruir sitios de endonucleasa de restricción preexistentes para facilitar la modificación in vitro adicional.

- 45 Debido a la degeneración de las secuencias codificantes de nucleótido, muchas secuencias de ácido nucleico diferentes que codifican sustancialmente la misma secuencia de aminoácidos como uno, cualquiera, ambos, varios, o todos los polipéptidos de complejo que va a ser estabilizado se pueden utilizar en la práctica de la presente invención. Pueden incluir las secuencias de nucleótido que comprenden todo o porciones de un dominio que se altera mediante la sustitución de diferentes codones que codifican el mismo aminoácido, o un residuo de aminoácido 50 funcionalmente equivalente dentro de la secuencia, produciendo así un cambio "inactivo" (funcionalmente o fenotípicamente irrelevante).

Se puede utilizar cualquier técnica para mutagénesis conocida en el arte, incluyendo pero no limitándose a, mutagénesis química, mutagénesis dirigida a sitio in vitro, utilizando, por ejemplo, el kit de Mutagénesis Dirigida a Sitio QuikChange (Stratagene), etc.

5.2. Aplicaciones de la tecnología de estabilización

Los métodos de estabilización de polipéptido y complejo de polipéptido de la invención tienen aplicación amplia. A continuación se establecen algunos ejemplos no limitantes.

5.2.1. General

5 Los complejos de polipéptido que se mantienen juntos en la naturaleza mediante los dominios que median las interacciones proteína-proteína se pueden estabilizar utilizando los métodos de la invención. Adicionalmente, las cadenas sencillas de polipéptido se pueden estabilizar utilizando los métodos de la invención para construir por ingeniería los entrecruzamientos di-tirosina intra-cadena. Como se utilizan aquí, los términos "complejo", "proteína o complejo de polipéptido", o "complejo estabilizado" significan la inclusión de una cadena sencilla, o más de una de las cadenas de polipéptido. Por ejemplo, hormonas (por ejemplo insulina, eritropoyetina, hormona de crecimiento humana o hormona de crecimiento bovina), otros factores de crecimiento (por ejemplo factores de crecimiento similares a insulina, factores neurotróficos) se pueden estabilizar, solos o juntos como un complejo como un receptor u otro patrón de unión a proteína (McInnes C. and Sykes B.D. *Biopolymers*; vol. 43(5): pp. 339-66, 1997). Ejemplos de los dominios de interacción proteína-proteína que se pueden estabilizar utilizando los métodos de la invención incluyen, pero no se limitan a, dominios leucina-zipper (Alber T. *Curr. Opin. Genet. Dev.*; vol. 2(2): pp. 205-10, 1992), dominios SH2 y SH3 (Pawson T. *Princess Takamatsu Symp.*; vol. 24: pp. 303-22, 1994), dominios PTB y PDZ (Cowburn D. *Curr. Opin. Struct. Biol.*; vol. 7(6): pp. 835-8, 1997; Bockaert J. y Pin J.P. *EMBO J.*; vol. 18(7): pp. 1723-9, 1999), dominios WD40 (Royet J. et al. *EMBO J.*; vol. 17(24): pp. 7351-60, 1998), dominios efectores de muerte y muerte (Strasser A. and Newton K. *Int. J. Biochem. Cell. Biol.*; vol. 31(5): pp. 533-7, 1999), dominios de desintegrina (Black R.A. and White J.M. *Curr Opin Cell Biol.*; vol. 10(5): pp. 654-9, 1998), y dominios CARD (Chou J.J. et al. *Cell*; vol. 94(2): pp. 171-80, 1998).

25 Las proteínas que dimerizan o multimerizan para funcionar se pueden estabilizar utilizando los métodos de la invención. Tales proteínas incluyen la mayor parte de complejos de inmunoglobulina, que incluyen los fragmentos que retienen la funcionalidad de la inmunoglobulina, tales como, por ejemplo, los fragmentos Fab, F(ab)₂, Fc, y Fv (Penuche M.L. et al. *Hum Antibodies*; vol. 8(3): pp. 106-18, 1997; Sensel M.G. et al. *Chem. Immunol.*; vol. 65: pp. 129-58, 1997). La mayor parte de los receptores de superficie celular que transmiten las señales extracelulares a sistemas de señalización intracelular dimerizan y contienen algunos de los dominios mencionados anteriormente que median las interacciones proteína-proteína (McInnes C. and Sykes B.D. *Biopolymers*; vol. 43(5): pp. 339-66, 1997; Guogiang J. et al.; *Nature*; vol. 401: pp.606-610, 1999). Ejemplos adicionales son complejos de proteína intracelulares, tal como, por ejemplo, las caspasas (Chou J.J. et al. *Cell*; vol. 94(2): pp. 171-80, 1998).

30 Los factores de crecimiento que se pueden estabilizar utilizando los métodos de la invención incluyen, pero no se limitan a, aquellos que dimerizan para funcionar, tal como interleucina-8 (Leong S.R. et al. *Protein Sci.*; vol. 6(3): pp. 609-17, 1997) y miembros de la familia NGF/TGF. Estas proteínas se caracterizan de manera general porque tienen 110-120 residuos de aminoácido, hasta 50% de homología uno al otro, y se utilizan para el tratamiento de una variedad de trastornos de la salud, tal como cáncer, osteoporosis, lesión de médula espinal y regeneración neuronal. Ejemplos de la familia NGF incluyen, pero no se limitan a, los polipéptidos NGF, BDNF, NT-3, NT-4/5, y NT-6, TRAIL, OPG, y FasL (Lotz M. et al. *J. Leukoc. Biol.*; vol. 60(1): pp. 1-7, 1996; Casaccia-Bonnel P. et al. *Microsc Res Tech.*; vol. 45(4-5): pp. 217-24, 1999; Natoli G. et al. *Biochem. Pharmacol.*; vol. 56(8): pp. 915-20, 1998). El TRAIL está actualmente en ensayos clínicos, y pueden ser útiles para inducir la apoptosis en células de cáncer. El OPG también está en ensayos clínicos y puede ser útil para fortalecer el tejido óseo y evitar la pérdida ósea durante la menopausia (Wickelgren I. *Science*; vol. 285(5430): pp. 998-1001, 1999).

Los factores de crecimiento que no dimerizan para funcionar en la naturaleza también se pueden estabilizar utilizando los métodos de la invención. Tales factores de crecimiento incluyen, pero no se limitan a, insulina, eritropoyetina, y cualquiera de los factores que estimulan la colonia (CSF) y PDGF. Estos polipéptidos se pueden estabilizar al introducir los enlaces di-tirosina intra-cadena de acuerdo con la invención.

45 Los biocatalizadores que se pueden estabilizar utilizando los métodos de la invención incluyen, pero no se limitan a, enzimas con aplicaciones de investigación industrial, aplicada o básica, o sectores de la industria, que incluyen, por ejemplo, industrias de productos químicos, detergentes, farmacéuticos, agrícolas, de alimentos, cosméticos, textiles, procesamiento de materiales, y de papel. Dentro de tales sectores de la industria, los biocatalizadores estabilizados se pueden utilizar, por ejemplo, para la síntesis del producto, como agentes activos en productos, en equipos de pruebas diagnósticas, o en cualquiera de otras aplicaciones conocidas en la técnica. Tales aplicaciones incluyen, pero no se limitan a, agua de desperdicio y tratamiento agrícola del suelo, y refinamiento de aceite crudo. Ejemplos de aplicaciones sintéticas incluyen, pero no se limitan a, fabricación de aminoácidos y síntesis de productos químicos finos. Un ejemplo de uso de un biocatalizador estabilizado como agente activo en un producto está en polvos de lavado biológico.

Los biocatalizadores que van a ser estabilizados se pueden seleccionar de enzimas de cualquier clase, familia, u otra categorización de las enzimas conocidas en la técnica. Tales enzimas incluyen, pero no se limitan a, proteasas, hidrolasas, ligasas, y polimerasas. Cada biocatalizador estabilizado puede comprender una única cadena de polipéptido o dos o más cadenas de polipéptido de un complejo de polipéptido.

5 5.2.2. Fragmentos Fv de inmunoglobulina

Los anticuerpos o moléculas de inmunoglobulina (Ig) están entre las moléculas más útiles terapéuticamente. Su utilidad resulta de su capacidad de unir a las moléculas objetivo dadas con especificidad y afinidad extremadamente altas. Su función en el sistema inmune es para unir a las moléculas externas (tal como las presentes en la superficie de los patógenos) y para activar la eliminación de estas moléculas externas del cuerpo utilizando una variedad de mecanismos efectores.

Con el advenimiento de la tecnología de hibridoma, con base en el trabajo de G. Kohler and C. Milstein al inicio de la década de los 80, ha llegado ser posible construir por ingeniería clones puros de células que expresan un anticuerpo único. Es vasta la utilidad de tales anticuerpos monoclonales (MAb), cuya especificidad de unión única se puede caracterizar en detalle. A partir de una población monoclonal de las células que producen anticuerpos es posible aislar los genes que codifican las cadenas de polipéptidos que constituyen el anticuerpo. La producción eficiente a gran escala de la inmunoglobulina recombinante en los sistemas de expresión bacteriana o de levadura es un interés activo de la industria de la biotecnología. De manera más importante, sin embargo, las técnicas de biología molecular permiten manipular estos genes y por lo tanto producir las proteínas derivadas de anticuerpos hechas a la medida para aplicaciones individuales, tal como las descritas adelante.

Una de las limitaciones principales para la efectividad clínica de los anticuerpos es su tamaño. Las moléculas de inmunoglobulina de longitud completa son efectivas como agentes humorales, pero su tamaño las hace difíciles de penetrar tejidos tales como tumores sólidos. Como resultado, se han diseñado versiones construidas por ingeniería, más pequeñas de los anticuerpos. Tales anticuerpos construidos por ingeniería se diseñan para retener la especificidad funcional normal con respecto a la unión a antígeno de una molécula mucho más pequeña, aunque al mismo tiempo se desacopla esta función de unión de la molécula de inmunoglobulina de otras funciones efectoras biológicas (por ejemplo activación de complementos o unión de macrófagos, Figura 1D).

Se ha mostrado que los fragmentos Fv son los fragmentos derivados de Ig más pequeños que retienen especificidad de unión completa (Figura 1D). El fragmento Fv comprende esencialmente sólo aquellas secuencias de aminoácidos de la molécula de anticuerpos que constituyen el "dominio variable" responsable de la unión de antígenos. Debido a su tamaño mínimo, los fragmentos Fv muestran penetración de tejido significativamente mejor y por lo tanto se pueden utilizar en un rango más amplio de contextos (por ejemplo terapia de tumores sólidos). Como se utiliza aquí, los fragmentos Fv deben incluir la región variable de las moléculas de inmunoglobulina o la región equivalente u homóloga de un receptor de células T.

Las comparaciones de la secuencia de aminoácidos de las regiones VH y VL de 110-120 residuos de longitud revelan que cada una está constituida por cuatro segmentos de secuencia relativamente conservados, llamados "Regiones de Estructura Principal" (FR), y tres segmentos de secuencia altamente variable, llamados "Regiones Determinantes de Complementariedad" (CDR I, II, & III), que determinan mayormente la especificidad del anticuerpo (Figura 1D, "brazo derecho").

Los polipéptidos del fragmento Fv de cadena ligera y pesada se asocian uno al otro mayormente en los sitios dentro de los FR conservados. Los fragmentos Fv, sin embargo, carecen de enlaces disulfuro intracadena estabilizantes estructurales en las regiones constantes Ig. Con el fin de mantener las cadenas ligera y pesada Fv recombinantes asociadas y alcanzar la estabilidad funcional y afinidad, las dos cadenas de la molécula se pueden "estabilizar" mediante algunos otros medios.

5.3. Biocatalizadores

Los biocatalizadores son una clase preferida de catalizadores para el desarrollo de procesos industriales, debido a su alta especificidad y los rendimientos del proceso. Específicamente, permiten el uso de menos energía y menos materias primas costosas (materiales de partida), reducen el número de etapas individuales que conducen a un producto, y reducen los productos de desecho. Su uso comercial, sin embargo, aún se limita por su inestabilidad, que restringe las aplicaciones clave. Esta invención proporciona métodos para estabilizar tales enzimas, mejorando su desempeño como catalizadores industriales, y prolongando sus vidas útiles y vida de estantería. La aplicación de la presente invención también permite el uso industrial de biocatalizadores, previamente inestables, novedosos, y por lo tanto también acorta los tiempos del ciclo de innovación del proceso industrial.

Específicamente, la aplicación de la presente invención estabiliza los biocatalizadores, por ejemplo, al evitar el despliegue de la proteína. Esto incrementa su capacidad de catalizar las reacciones químicas bajo condiciones de

reacción adversas, prolonga su vida útil y vida de estantería, y maximiza su actividad a temperaturas de proceso reales, más suaves.

5.4. Polipéptidos obtenidos a ser estabilizados

5 Se puede utilizar cualquier método conocido por un experto en la técnica para obtener un polipéptido o complejo de polipéptido que va a ser estabilizado de acuerdo con los métodos de la invención.

5.4.1. Purificación de polipéptidos

10 Se puede obtener un polipéptido o complejo de polipéptido que va a ser estabilizado utilizando los métodos de la presente invención, por ejemplo, mediante cualquier método de purificación de proteínas conocido en la técnica. Tales métodos incluyen, pero no se limitan a, cromatografía (por ejemplo cromatografía de intercambio iónico, de afinidad, y/o de separación por tamaño en columna), precipitación de sulfato de amonio, centrifugación, solubilidad diferencial, o mediante cualquier otra técnica estándar para la purificación de las proteínas. Se puede purificar un polipéptido a partir de cualquier fuente que la produzca. Por ejemplo, se pueden purificar los polipéptidos a partir de fuentes que incluyen, fuentes procariótica, eucariótica, mono- celular, multi-celular, animal, planta, hongo, vertebrado, mamífero, humana, porcina, bovina, felina, equina, canina, aviar, células de cultivo de tejido, y cualquier otra fuente natural, modificada, construida por ingeniería, o de cualquier otra forma de fuente de origen no natural. El grado de pureza puede variar, pero en varias realizaciones, la proteína purificada es mayor de 50%, 75%, 85%, 95%, 99%, o 99.9% de los miligramos de proteína totales. Así, un lisado celular crudo no comprenderá una proteína purificada.

20 Cuando es necesario introducir uno o más residuos de tirosina que se van a entrecruzar en un polipéptido o complejo de polipéptido purificado, los polipéptidos se pueden microsecuenciar para determinar una secuencia de aminoácidos parcial. La secuencia de aminoácidos parcial luego se puede utilizar junto con métodos selección de bibliotecas y de ácidos nucleicos recombinantes bien conocidos en la técnica para aislar los clones necesarios para introducir las tirosinas.

5.4.2. Expresión de ADN que codifica un polipéptido

25 Fuente de ADN

Cualquier célula procariótica o eucariótica puede servir como fuente de ácido nucleico para clonación molecular. Se puede aislar una secuencia de ácido nucleico que codifica una proteína o dominio que va a ser entrecruzado o estabilizado a partir de fuentes que incluyen procarióticos, eucarióticos, mono- celular, multi-celular, animal, planta, hongos, vertebrados, mamíferos, humano, porcino, bovino, felino, equino, canino, aviar, etc.

30 El ADN se puede obtener mediante procedimientos estándar conocidos en la técnica a partir de ADN clonado (por ejemplo, una "biblioteca" de ADN), mediante síntesis química, mediante clonación de cADN, mediante la clonación del ADN genómico, o fragmentos de los mismos, purificado de la célula deseada (ver por ejemplo, Sambrook et al.; Glover (ed.). MRL Press, Ltd., Oxford, U.K.; vol. I, II, 1985). El ADN también se puede obtener mediante transcripción inversa de ARN celular, preparado mediante cualquiera de los métodos conocidos en la técnica, tales como transcripción inversa aleatoria o cebada con poli A. Tal ADN se puede amplificar utilizando cualquiera de los métodos conocidos en la técnica, que incluyen técnicas de PCR y 5' RACE (Weis J.H. et al. Trends Genet. 8(8): pp. 263-4, 1992; Frohman M.A. PCR Methods Appl. 4(1): pp. S40-58, 1994).

40 Cualquiera que sea la fuente, el gen se debe clonar molecularmente en un vector adecuado para la propagación del gen. Adicionalmente, el ADN se puede dividir en sitios específicos utilizando varias enzimas de restricción, se puede utilizar ADNsa en la presencia de manganeso, o el ADN se puede cortar físicamente, como por ejemplo, mediante sonicación. Los fragmentos de ADN lineal luego se pueden separar de acuerdo con el tamaño mediante técnicas estándar, tal como cromatografía de columna y electroforesis en gel de poliacrilamida y agarosa.

Clonación

45 Una vez se generan los fragmentos de ADN, la identificación del fragmento de ADN específico que contiene el gen deseado se puede llevar a cabo en un sinnúmero de formas. Por ejemplo, se pueden utilizar los clones al utilizar técnicas PCR que pueden utilizar dos oligonucleótidos específicos para la secuencia deseada, o un oligonucleótido único específico para la secuencia deseada, utilizando, por ejemplo, el sistema 5' RACE (Cale J.M. et al. Methods Mol. Biol.; vol.105: pp. 351-71, 1998; Frohman M.A. PCR Methods Appl.; vol. 4(1): pp. S40-58, 1994). Los oligonucleótidos pueden o no contener residuos de nucleótidos degenerados. Alternativamente, si una porción de un gen o su ARN específico o un fragmento del mismo está disponible y se puede purificar y marcar, los fragmentos de ADN generados se pueden detectar mediante la hibridación de ácido nucleico para la sonda marcada (por ejemplo

Benton and Davis. Science; vol. 196(4286): pp. 180-2, 1977). Hibridarán aquellos fragmentos de ADN con homología sustancial a la sonda. También es posible identificar el fragmento apropiado mediante la digestión de la enzima de restricción y comparación de los tamaños de fragmento con aquellos esperados de acuerdo con un mapa de restricción conocido si está disponible. La selección adicional se puede llevar a cabo con base en las propiedades del gen.

La presencia del gen deseado también se puede detectar mediante ensayos basados en las propiedades físicas, químicas, o inmunológicas de su producto expresado. Por ejemplo, los clones de cADN, o los clones de ADN que hibridan seleccionan los mRNA apropiados, se pueden seleccionar y expresar para producir una proteína que tiene, por ejemplo, migración electroforética idéntica o similar, comportamiento enfocado isoeléctrico, mapas de digestión proteolítica, actividad hormonal u otra actividad biológica, actividad de unión, o propiedades antigénicas tal como se conoce para una proteína.

Utilizando un anticuerpo para una proteína conocida, se pueden identificar otras proteínas mediante la unión del anticuerpo marcado para expresar las proteínas putativas, por ejemplo, en un procedimiento tipo ELISA (ensayo de inmunoabsorción ligado a enzima). Adicionalmente, utilizando una proteína de unión específica a una proteína conocida, se pueden identificar otras proteínas mediante unión a tal una proteína in vitro o un sistema de célula adecuado, tal como el sistema de dos híbridos de levadura (ver por ejemplo Clemmons D.R. Mol. Reprod. Dev.; vol. 35: pp. 368-374, 1993; Loddick S.A. et al. Proc. Natl. Acad. Sci., U.S.A.; vol. 95: pp. 1894-1898, 1998).

También se puede identificar un gen mediante selección de mRNA utilizando hibridación de ácido nucleico seguido por traducción in vitro. En este procedimiento, se utilizan fragmentos para aislar los mRNA complementarios mediante hibridación. Tales fragmentos de ADN pueden representar ADN purificado, disponible de otras especies (por ejemplo, Drosophila, ratón, humano). El análisis de inmunoprecipitación o los ensayos funcionales (por ejemplo capacidad de agregación in vitro, unión al receptor, etc.) de los productos de traducción in vitro de los productos aislados de los mRNA aislados identifican el mRNA y, por lo tanto, los fragmentos de ADN de complementariedad que contienen las secuencias deseadas.

Adicionalmente, se pueden seleccionar mRNA específicos mediante adsorción de polisomas aislados de células para los anticuerpos inmovilizados dirigidos específicamente contra la proteína. Un cADN radiomarcado se puede sintetizar utilizando el mRNA seleccionado (de los polisomas adsorbidos) como una plantilla. El mRNA o cADN radiomarcados se pueden utilizar luego como una sonda para identificar los fragmentos de ADN de entre otros fragmentos de ADN genómico.

Las alternativas para aislar el ADN genómico incluyen sintetizar químicamente la secuencia de gen en sí misma de una secuencia conocida o hacer el cADN al mRNA que codifica la proteína. Por ejemplo, el ARN para la clonación del cADN del gen se puede aislar de las células que expresan el gen.

Vectores

El gen identificado y aislado se puede insertar luego en un vector de expresión o clonación apropiado. Se puede utilizar un gran número de sistemas de vector-anfitrión conocidos en la técnica. Los posibles vectores incluyen plásmidos o virus modificados, pero el sistema de vector puede ser compatible con la célula anfitriona utilizada. Tales vectores incluyen bacteriófagos tales como derivados lambda, o plásmidos tal como PBR322 o derivados de plásmido pUC o el vector Bluescript (Stratagene).

La inserción dentro de un vector de clonación, por ejemplo, se puede llevar a cabo al ligar el fragmento de ADN en un vector de clonación que tenga un terminal cohesivo complementario. Sin embargo, si no están presentes los sitios de restricción de complementariedad para el fragmento de ADN en el vector de clonación, los extremos de las moléculas de ADN se pueden modificar enzimáticamente. Alternativamente, se puede producir cualquier sitio deseado al ligar las secuencias de nucleótido (ligadores) en el terminal de ADN; estos ligadores enlazados pueden comprender oligonucleótidos químicamente sintetizados específicos que codifican las secuencias de reconocimiento de endonucleasa de restricción. Adicionalmente, el gen y/o el vector se pueden amplificar utilizando técnicas de PCR y oligonucleótidos específicos para el terminal del gen y/o el vector que contiene los nucleótidos adicionales que proporcionan el terminal cohesivo de complementariedad deseado. En los métodos alternativos, el vector dividido y un gen se pueden modificar mediante terminación homopolimérica (Cale J.M. et al. Methods Mol. Biol.; vol. 105: pp. 351-71, 1998). Las moléculas recombinantes se pueden introducir en células anfitrionas por medio de transformación, transfección, infección, electroporación, etc., de tal manera que se generan muchas copias de la secuencia de gen.

Preparación de ADN

En realizaciones específicas, la transformación de las células anfitrionas con moléculas de ADN recombinante que incorporan un gen aislado, cADN, o secuencia de ADN sintetizada permite la generación de múltiples copias del gen.

Así, el gen se puede obtener en grandes cantidades al hacer crecer los transformantes, aislar las moléculas de ADN recombinante de los transformantes y, cuando sea necesario, recuperar el gen insertado del ADN recombinante aislado.

- 5 Las secuencias relacionadas con la presente invención incluyen aquellas secuencias de nucleótidos que codifican sustancialmente las mismas secuencias de aminoácidos como se encuentran en las proteínas nativas, y aquellas codificadas por las secuencias de aminoácidos con aminoácidos funcionalmente equivalentes, así como también aquellos que codifican otros derivados o análogos, como se describe adelante para los derivados y análogos.

Estructura de Genes y Proteínas

- 10 La secuencia de aminoácidos de una proteína se puede derivar mediante deducción de la secuencia de ADN, o alternativamente, mediante secuenciamiento directo de la proteína, por ejemplo, con un secuenciador de aminoácidos automatizado.

- 15 Una secuencia de proteína se puede caracterizar adicionalmente mediante un análisis de hidrofiliidad (Hopp T.P. y Woods K.R. Proc. Natl. Acad. Sci., U.S.A.; vol. 78: pp. 3824, 981). Se puede utilizar un perfil de hidrofiliidad para identificar las regiones hidrófobas e hidrófilas de la proteína y las regiones correspondientes de la secuencia del gen que codifica tales regiones.

- 20 Segundo, también se puede hacer un análisis estructural (Chou P.Y. y Fasman G.D. Biochemistry; vol. 13(2): pp. 222-45, 1974), para identificar las regiones de una proteína que asumen estructuras secundarias específicas. La manipulación, traducción, y predicción de estructura secundaria, la predicción de estructura de lectura abierta y gráfica, así como también la determinación de las homologías de secuencia, también se puede llevar a cabo utilizando programas de software de ordenador disponibles en la técnica. Otros métodos de análisis estructural incluyen cristalografía de rayos X, espectroscopía de resonancia magnética nuclear y modelamiento por ordenador.

5.5. Residuos adecuados para una reacción de entrecruzamiento

La identificación y/o construcción por ingeniería de los residuos adecuados para una reacción de entrecruzamiento puede involucrar una o más de las varias etapas establecidas adelante.

- 25 5.5.1. Introducción de mutaciones puntuales para controlar la reacción de entrecruzamiento

- 30 La estructura general construida por ingeniería y la función de un polipéptido o complejo de polipéptido estabilizado se logran al controlar la disponibilidad de las cadenas laterales tirosilo para la reacción de entrecruzamiento, por ejemplo, pero no limitándose a, mutagénesis. La funcionalidad de un polipéptido o complejo de polipéptido se puede comprometer o alterar mediante una reacción de entrecruzamiento tirosina-tirosina. En este caso, un grupo hidroxilo indeseable de una cadena lateral tirosilo se puede retirar mediante la mutación de tales residuos a fenilalanina, o enmascararse para inhibir su participación en tal reacción. En esta forma, se retira un residuo tirosilo disponible para la reacción de entrecruzamiento pero que puede conducir a distorsión de la estructura y comprometer la funcionalidad y/o especificidad del polipéptido o complejo de polipéptido. Más aún, las mutaciones puntuales a la tirosina se pueden introducir en las posiciones en donde las cadenas laterales tirosilo reaccionarán una con la otra para formar un enlace que origina la distorsión de la estructura y función; estas posiciones se identifican como se describe en detalle adelante. Por lo tanto, se mantiene la estructura general y funcionalidad del polipéptido o complejo de polipéptido.

5.5.2. Retirar cadenas laterales reactivas indeseables

- 40 Se identifica que las cadenas laterales reactivas identificadas en una cadena de polipéptido o en las cadenas de polipéptido de un complejo que se someten a condiciones de entrecruzamiento oxidativo descrito anteriormente dan como resultado un enlace que distorsionará la estructura del complejo. Estos residuos se identifican mediante comparación de las secuencias de aminoácidos de los polipéptidos para información estructural disponible en tal o complejos similares (ver adelante). Se puede formar tal enlace entre las dos cadenas de polipéptido del complejo (enlace inter-cadena) o entre dos residuos de una y la misma cadena de polipéptidos (enlace intra-cadena). El efecto de formación de un enlace se determina mediante ambas cadenas laterales reactivas involucradas en la formación de tal enlace, y por lo tanto estos residuos se identificarían en pares.

- 50 Para neutralizar este efecto dañino de la reacción de entrecruzamiento, se pueden utilizar los reactivos de enmascaramiento que protegen las cadenas laterales aromáticas (Pollitt S. y Schultz P. Agnew. Chem. Int. Ed.; vol. 37(15): pp. 2104-2107, 1998), o las sustituciones de aminoácidos por fenilalanina, o cualquier otro aminoácido, se puede introducir por lo menos en uno de los residuos involucrado, por ejemplo, al introducir una mutación puntual en el cADN del gen que dirige la expresión del polipéptido.

5.5.3. Introducir cadenas laterales reactivas

Para lograr un polipéptido o complejo de polipéptido estabilizado sin perturbar su estructura y/o función, las posiciones dentro de cada polipéptido se identifican con una cadena lateral reactiva que sería capaz de formar un enlace con una cadena lateral reactiva en, o una de las, otras cadenas de polipéptido. Tales posiciones se seleccionan con respecto hacia mantener la estructura general del mismo polipéptido, y con respecto hacia la adecuabilidad de una posición en el otro polipéptido involucrado en el enlace, y las posiciones por lo tanto se seleccionan en pares (ver adelante para la descripción detallada del proceso de selección).

Cuando en un residuo seleccionado de ambos, o cualquiera, de los polipéptidos, la cadena lateral tirosilo reactiva ya no están presente, se puede introducir un punto de mutación, por ejemplo, pero no limitado a, al utilizar los métodos de biología molecular para introducir tal una mutación puntual en el cadn del gen que dirige su expresión, de tal manera que está presente una cadena lateral reactiva y está disponible para la reacción.

5.6. Dominios conservados estructuralmente

5.6.1. Relación entre estructura y función

Esta es la estructura tridimensional, o terciaria, de cada proteína, y la estructura cuaternaria de cada complejo de proteína que les presta la funcionalidad que les ha permitido mantenerse y desarrollarse a través del proceso evolutivo con el tiempo. Una mutación puntual en el gen de polipéptido o complejo de polipéptido que conduce a una sustitución de aminoácido en cualquier residuo dado alterará la estructura del polipéptido y/o del complejo general a un mayor o menor grado. El grado de tal efecto de sustitución de aminoácidos en la estructura del polipéptido o complejo de polipéptido es dependiente del contexto estructural del residuo, y de la naturaleza de la cadena lateral de aminoácido resultante.

Los dominios de proteína que muestran la similitud extensa en sus secuencias de aminoácidos en los dominios en otras proteínas se denominan como "dominios conservados". Dentro de los dominios conservados los residuos individuales son más conservados que otros; algunos pueden ser 100% conservados, y otros no en todos. La mayor parte de los dominios conservados no son solo similares en sus secuencias de aminoácido, sino también en sus estructuras tridimensionales, y también en sus funciones. En la ausencia de las presiones evolutivas que requieren un residuo de un dominio que va a ser conservado, es evidente que el aminoácido presente en un residuo variará ampliamente debido al índice de mutación que conduce la diversificación evolutiva. Por lo tanto, los residuos dentro de un dominio conservado que son altamente conservados se considera que son importantes contribuidores en la estructura general, o la arquitectura, del dominio. Entre los residuos que son menos conservados están aquellos que contribuyen a la especificidad del dominio individual del grupo.

Los dominios conservados, sin embargo, también pueden mostrar muy poca homología de secuencia y aún tienen estructuras conservadas, tal como, por ejemplo, cremalleras de leucina (Alber T. Curr. Opin. Genet. Dev.; vol. 2(2): pp. 205-10, 1992). Debido a que una estructura conservada también produce residuos estructuralmente conservados, la distinción entre los residuos 'arquitectónicos' y 'determinantes de especificidad' descritos anteriormente también se puede hacer en la ausencia de la conservación de la secuencia. Para los propósitos de la presente invención, se define un dominio conservado, dependiente de la disponibilidad de los datos, mediante homología de secuencia, que puede ser tan baja como 5% de identidad o similitud, o mediante el grupo de funcionalidad o estructura de los dominios.

5.6.2. Alineación de los residuos conservados

La alineación de las secuencias bidimensionales de los dominios conservados revela adicionalmente que entre los residuos conservados que se intercalan frecuentemente mediante las cadenas de varias longitudes, es decir existen números que varían de los residuos de aminoácidos entre los residuos conservados importantes para la estructura general del dominio. Con el fin de ser capaz de comparar las secuencias de los dominios individuales para determinar en donde se dirige la reacción de entrecruzamiento, es esencial que las secuencias se alineen en tal forma que se comparan los aminoácidos que corresponden estructuralmente uno al otro. Para los residuos identificados en los análisis de secuencia de nucleótidos y de aminoácidos como altamente conservados, se lleva a cabo fácilmente.

5.7. Método de selección estadística

La selección de los pares de residuo a los que el entrecruzamiento se puede dirigir para estabilizar un polipéptido o funcionalidad de complejo se lleva a cabo preferiblemente al analizar los datos en varias estructuras de un grupo de dominios conservados y estadísticamente sus ligandos y seleccionar los pares de residuo con base en el criterio de selección desarrollado como se describe adelante. A partir de las mediciones que se hacen en un conjunto de datos, es posible hacer predicciones para la adecuabilidad de las selecciones de par de residuo de tales dominios que

están, y que no están, presentes en la muestra. Un residuo que es altamente conservado en la muestra tiene una alta probabilidad de tener similitud en todos los dominios individuales, que incluyen aquellos que no están presentes en la muestra. Por lo tanto, utilizando análisis estadísticos para determinar a qué pares de residuo la reacción de entrecruzamiento se debe dirigir esta tecnología en un mayor grado de aplicabilidad general.

5 Cuando no es posible o es inconveniente obtener los datos requeridos para tales análisis, los pares de residuo también se pueden seleccionar en cualquier otra forma, que incluye, por ejemplo, ensayo y error. Tales procesos de selección producen pares de residuo al cual se puede dirigir el entrecruzamiento mientras se mantiene la funcionalidad del polipéptido o complejo de polipéptido.

10 Dada la disponibilidad de los datos relevantes, es frecuentemente posible asignar valores cuantitativos para ciertas características de una cadena lateral de aminoácido presente en cada residuo de un dominio. Adicionalmente, dados los datos relevantes en los polipéptidos o complejos de polipéptido, es posible dar pares de los valores de aminoácidos que describan su relación.

15 Estos valores se pueden comparar entre los dominios individuales al alinear los datos en tal una forma que los conjuntos de valores que van a ser comparados se relacionan funcionalmente y estructuralmente (ver anterior). Si existe un número suficientemente grande de dominios individuales para los que están disponibles tales datos, también es posible analizar estos conjuntos de datos estadísticamente.

20 El análisis estadístico de tales conjuntos de datos proporciona información que se relaciona con el grado de una conservación y/o variabilidad estructural del residuo en la muestra, y una indicación hasta qué grado un residuo se involucra en proporcionar la arquitectura destacada de un dominio. Esta información se deriva de mediciones estadísticas tales como varianza, desviación estándar, rango, máximo y mínimo, y otros. Valores altos de tal medición implican alta variabilidad de los valores de residuo, y así un bajo grado de conservación, y viceversa.

5.8. Criterios de selección para las sustituciones del residuo de aminoácidos

5.8.1. Adecuabilidad estructural

25 El objeto de los análisis es determinar qué pares de residuos serán más adecuados para la reacción de entrecruzamiento con el fin de mantener la estructura, función, y especificidad de un polipéptido o complejo de polipéptido. Por lo tanto, se seleccionan muchos de los criterios de los pares de residuo para relacionar el potencial de los pares para acomodar las dos cadenas laterales reactivas entrecruzadas sin distorsionar la estructura principal del enlace de péptido y alterar la estructura del polipéptido o del complejo en las posiciones que permiten y definen su función y especificidad.

30 Las mediciones que se pueden hacer para obtener la información que se relaciona con este potencial se relaciona con los determinantes del espacio disponible para las cadenas laterales reactivas y el enlace. Tales mediciones incluyen la distancia entre los carbonos alfa de los pares de residuos, que son los átomos de carbono que son una parte de la "estructura principal" formada por los enlaces de péptido entre todos los aminoácidos del polipéptido. Los pares de residuos seleccionados deben tener una distancia de carbono alfa promedio cercana a la distancia en que los carbonos alfa de las cadenas laterales tirosilo entrecruzadas estarían una de la otra si se introducen mutaciones puntuales, y la reacción de entrecruzamiento se dirige a aquella del par de residuo. Los pares de residuo seleccionados deben estar tan cerca a la distancia de los carbonos alfa de las cadenas laterales tirosilo entrecruzadas para asegurar que se mantiene la funcionalidad del polipéptido o complejo de polipéptido. Los criterios para esta selección se describen en detalle adelante (Proceso de Selección: Determinación de la Distancia de Carbono Alfa en el Enlace tirosilo-tirosilo, Los Filtros). Debido a que la variabilidad de una característica estructural de los pares de residuo también es un criterio importante en la selección de pares de residuo adecuados para la reacción de entrecruzamiento (ver adelante), la proximidad requerida en la distancia óptima se calcula para cada par de residuo, dependiente de la variabilidad de sus distancias de carbono alfa en la muestra. El cálculo de este requerimiento también se describe en detalle adelante (Proceso de Selección: Los Filtros).

45 También se pueden hacer mediciones para determinar si la proteína se plegará en una forma tal que las cadenas laterales reactivas se dirigirán una hacia otra. Los criterios de selección se pueden desarrollar con base en los ángulos de las cadenas laterales reactivas y del entrecruzamiento, la libertad rotacional de las cadenas laterales reactivas, y las mediciones relacionadas con la relación geométrica tridimensional entre los carbonos alfa y los carbonos beta de cada par de residuos. El carbono beta es el primer átomo de carbono de las cadenas laterales de aminoácido que no es parte de la estructura principal. Tales criterios de selección se describen en detalle adelante (Proceso de Selección: Cálculos de los Ángulos de Cadena Lateral en el Enlace Tirosilo, Los Filtros). El aminoácido más pequeño, glicina, no tiene un carbono beta, y por lo tanto los pares de residuos de los cuales uno o ambos aminoácidos es una glicina conservada no se pueden analizar en esta forma. Debido a que la mutación de una glicina conservada conduciría probablemente a una distorsión estructural significativa, se eliminan los pares de

residuos de los que uno o ambos residuos son una glicina conservada. Este criterio de selección también se describe en detalle adelante (Proceso de Selección: Los Filtros).

Adicionalmente, el contexto estructural del par de residuos se considera preferiblemente para determinar la disponibilidad del espacio tridimensional para las cadenas laterales reactivas y el enlace. Las características de cadena lateral de aminoácido relevantes de los residuos próximos se tienen por lo tanto preferiblemente en cuenta, para justificar adicionalmente que las cadenas laterales reactivas serán capaces de rotar de tal manera que el enlace se puede formar sin distorsionar la estructura principal del polipéptido. Si el contexto es de tal manera que las cadenas laterales reactivas introducidas mediante mutación puntual no son capaces de rotar libremente en la posición deseada, el enlace no se formará fácilmente, u ocurrirán distorsiones que pueden deteriorar potencialmente o alterar la función y/o especificidad del polipéptido o complejo de polipéptido. Por lo tanto, se desarrollan criterios de selección para permitir mutaciones puntuales más conservadoras por ser introducidas que serán menos capaces de provocar distorsiones estructurales. Tales criterios se basan en los aminoácidos presentes en, y que rodean, los residuos de un par, y se cuantifican con base en valores numéricos de las propiedades físicas de aquellas cadenas laterales de aminoácidos. El cálculo de tales requerimientos se describe en detalle adelante (Proceso de Selección: Los Filtros).

Si se identifica un par de residuos adecuado que es un aminoácido reactivo apropiado en ambas cadenas en alguna frecuencia en la muestra, este par sería una selección ideal. Sin embargo, las cadenas laterales reactivas presentes en el polipéptido o los polipéptidos del complejo que va a ser entrecruzado que provocaría las distorsiones estructurales al formar enlaces inter- o intra-cadena se debe neutralizar, mediante medios para enmascararlos/protegerlos (Pollitt S. y Schultz P. Agnew. Chem. Int. Ed.; vol. 37(15): pp. 2104-2107, 1998) o al introducir mutaciones puntuales máximamente conservadoras. Tales pares de residuos reactivos se identifican utilizando el mismo criterio como para la selección positiva de los pares de residuos adecuados para entrecruzamiento. Sin embargo, la presencia de cadenas laterales indeseadas sólo se puede determinar al analizar la secuencia específica de un dominio individual, y al compararla con la información estructural utilizada para la selección positiva.

5.8.2. Variabilidad

La especificidad de cada dominio individual y su contraparte en el mismo polipéptido o en otro polipéptido de un complejo se determina de manera general mediante los residuos que son menos, o no son conservados. Por lo tanto, considerando la especificidad de un dominio individual, un residuo con alta variabilidad puede ser una elección menos deseable en la cual se dirige la reacción de entrecruzamiento. Sin embargo, considerando la estructura general y la arquitectura de un dominio, la arquitectura del dominio puede acomodar más fácilmente una mutación en un residuo que exhibe un alto grado de variabilidad. Así, desde esta perspectiva, la alta variabilidad indica que un residuo es un mejor candidato en el cual introducir una mutación puntual, y colocar una cadena lateral reactiva.

Dependiendo de la confiabilidad y exactitud de estos análisis, que, a su vez, dependen de la confiabilidad de las entradas en los análisis (ver adelante), es posible variar el requerimiento para una variabilidad de par o de una posición (que indica un cierto grado de flexibilidad y/o robustez). Así, si las entradas son altamente exactas, y los datos suficientes están presentes en la muestra, es posible determinar que un par de residuos es altamente adecuado para la reacción aunque su variabilidad es baja. Sin embargo, en casos en donde existen datos insuficientes o una exactitud insuficiente en las entradas en los análisis permite la baja variabilidad, se puede seleccionar un residuo que es importante para la especificidad, pero no para la arquitectura general del dominio.

En la ausencia de los datos funcionales es muy difícil determinar una contribución del residuo a la especificidad del dominio.

5.9. Generación y uso de las bases de datos

5.9.1. Datos generados relevantes para los criterios de selección

La disponibilidad incrementada de los datos que se relacionan con los genes, proteínas, y otras biomoléculas de muchas especies vivas, hace posible compilar una cantidad significativa de datos en varios dominios/módulos de proteína para los análisis estadísticos para hacer las predicciones descritas anteriormente. Estos datos se pueden transformar en datos que se pueden utilizar para tales análisis directamente.

Tales transformaciones, por ejemplo, se pueden hacer al convertir los datos de nucleótidos en los datos de secuencia de aminoácidos, y al convertir adicionalmente los datos de secuencia de aminoácidos en datos numéricos que se relacionan con la propiedades físicas de las cadenas laterales de los aminoácidos de un residuo dado. Tales propiedades, por ejemplo, pueden ser la carga o el grado de hidrofobicidad de las cadenas laterales de un residuo.

Adicionalmente, los datos estructurales de un polipéptido o de dos o más polipéptidos en un complejo se pueden transformar en datos numéricos que describen las relaciones estructurales entre los residuos individuales de un polipéptido único y/o entre los varios polipéptidos de un complejo. Un ejemplo para tal transformación sería el cálculo de las distancias entre los carbonos alfa de un par de residuos utilizando datos de coordenadas tridimensionales de resolución cristalográfica de una estructura utilizando geometría tridimensional de Pitágoras.

Es posible generar muchos conjuntos diferentes de datos relevantes para la estabilización de acuerdo con el procedimiento de esta invención que relacionan muchas de las características estructurales de los residuos y los pares de residuos de un dominio o un complejo. Como se requieren juicios frecuentemente más cualitativos para determinar la confiabilidad de las entradas de selección, también llega a ser una decisión más cualitativa cómo se deben utilizar diferentes conjuntos de datos en la selección de los pares de residuos que deben formar el enlace estabilizante. Cuanto menos confiables sean las entradas, es más útil implementar la información adicional en la selección.

5.9.2. Fuentes de datos

Datos de Secuencia

La forma más directa de acumular secuencias es mediante clonación y secuenciación de cADN de proteínas que contiene 722 dominios/módulos de interés. Los datos de secuencia llegan a ser y son más disponibles a través de los proyectos de genoma. Muchos de los datos de secuencia están disponibles en las bases de datos a las que se puede tener acceso a través de internet, o de otra forma, y adicionalmente existen varias fuentes publicadas que tienen secuencias acumuladas de dominios/módulos específicos. Una tal colección de datos de secuencia específicos es la Base de datos Kabat de las Secuencias de Proteínas de Interés Inmunológico (<http://immuno.bme.nwu.edu>; Johnson, G. et al. Weir's Handbook of Experimental Immunology I. Immunochemistry and Molecular Immunology, Quinta Edición, Ed. L. A. Herzenberg, W. M. Weir, y C. Blackwell, Blackwell Science Inc., Cambridge, MA, Chapter 6.1-6.21, 1996) que contiene, entre otras cosas, las secuencias de las moléculas de inmunoglobulina (ver Sección VI, Ejemplos). Tales datos de secuencia también están disponibles en Genebank (<http://www.ncbi.nlm.nih.gov>).

Datos Estructurales

Las estructuras tridimensionales, como se describen mediante datos de coordenadas atómicas, de un polipéptido o un complejo de dos o más polipéptidos se puede obtener en varias formas.

El primer método es la base de datos de las coordenadas estructurales existentes de las proteínas de interés. Los datos de las estructuras resueltas está disponible frecuentemente en la bases de datos a las que se accede fácilmente en la forma de coordenadas tridimensionales (x, y, y z) en unidades Angstrom (Å o 10^{-10} metros). Frecuentemente estos datos también son accesibles a través de internet (por ejemplo la base de datos de estructura de proteína en línea de National Brookhaven Laboratory: www.nsl.pdb.gov).

El segundo utiliza patrones de difracción (mediante, por ejemplo, pero no limitándose, a rayos X o electrones) de disposiciones regulares en 2- o 3- dimensiones de las proteínas, por ejemplo como se utiliza en el campo de cristalografía de rayos X. Se utilizan métodos computacionales para transformar tales datos en coordenadas atómicas tridimensionales en espacio real.

El tercero utiliza Resonancia Magnética Nuclear (RMN) para determinar las distancias inter-atómicas de las moléculas en la solución. Los métodos de RMN multi-dimensionales combinados con métodos computacionales han sido exitosos en la determinación de las coordenadas atómicas de los polipéptidos de tamaño incrementado.

Un cuarto método consiste completamente en modelamiento computacional. Los algoritmos se pueden basar en la naturaleza fisicoquímica conocida de los aminoácidos y los enlaces encontrados en las proteínas, o en métodos iterativos que se limitan experimentalmente, o ambos. Un ejemplo de software es el programa CNS desarrollado por Axel Brunger y colegas en HHMI en la Universidad de Yale (Adams P.D. et al. Acta Crystallogr. D. Biol. Crystallogr.; vol. 55 (Pt 1): pp. 181-90, 1999).

Datos funcionales

Los datos funcionales no se utilizan fácilmente, cuando no existe una manera uniforme de estandarizarlos y compilarlos, tal como los datos de secuencia de aminoácidos o nucleótidos, o coordenadas de los datos estructurales. Se genera de muchas formas diferentes, tal como análisis genéticos, bioquímicos, y mutacionales, disección en biología molecular y construcción de dominios quiméricos. En muchos casos los datos disponibles no son siempre interpretables claramente y por lo tanto su uso llega a ser delineado con menor claridad. Pero cuando

están disponibles, los datos de función proporcionan información valiosa que se relaciona con la especificidad y la funcionalidad de un dominio/módulo, y cuando sea posible se incorpora preferiblemente en el Proceso de Selección.

Los datos funcionales también se generan preferiblemente después de la reacción de entrecruzamiento de acuerdo con la presente invención para asegurar que las predicciones hechas son exactas para la aplicación específica, y que el polipéptido o complejo de polipéptido retiene actualmente su funcionalidad y especificidad.

5.9.3. Construcción de la base de datos

Base de datos 3-D

Se compila una base de datos de datos relevantes de información estructural que incluye los datos de coordenadas atómicas de polipéptidos resueltos cristalográficamente y los complejos de polipéptido de un grupo de polipéptidos o dominios conservados y su ligandos, y derivados. Los datos de entrada se derivan de sus archivos de datos de coordenadas estructurales. Los datos relevantes para el Proceso de Selección en esta base de datos se deriva de los datos de coordenadas al aplicar geometría de coordenadas en tres dimensiones. Esta base de datos contiene preferiblemente, por ejemplo, además de los datos de coordenadas estructurales, los siguientes datos relevantes junto con las mediciones estadísticas (por ejemplo media, mediana, modo, desviación estándar, máximo, y mínimo) en cada una de las siguientes características para cada par de residuos, por lo cual los polipéptidos de muestra y complejos de polipéptido se alinean como se describió anteriormente.

1. Distancias intercadena de Carbono alfa a carbono alfa de los pares de polipéptidos de un complejo de polipéptido, con el fin de encontrar los pares de residuo que se separan apropiadamente para un enlace tirosilo-tirosilo que va a ser formado. Estas distancias se calculan mediante, por ejemplo, pero no limitándose a, la aplicación de geometría de Pitágoras a las coordenadas 3D de los carbonos alfa. Para cada par de residuos se calculan las mediciones, tal como el promedio, la desviación estándar, rango y mediana de las distancias carbono alfa - carbono alfa correspondientes.

2. Los tres ángulos, Θ , χ y ψ (Figura 2c) en relación a los cuales las cadenas laterales en cada par de residuos se orientan uno hacia otra con relación a los ejes carbono alfa - carbono alfa intercadena, se calculan a partir de las coordenadas de los carbonos alfa y beta de cada par para cada polipéptido o complejo de polipéptido en la muestra. Los ángulos se calculan definiendo dos planos, cada uno de los cuales se define mediante ambas posiciones de los carbonos alfa y una de las posiciones de los carbonos beta. Al aplicar la geometría analítica, se calcula cada uno de los ángulos en los carbonos alfa (productos escalares), y el ángulo formado por los planos (productos vectoriales). Las mediciones estadísticas también se hacen a partir de este conjunto de datos, como se describe para el espaciamiento carbono alfa.

La diferencia entre la distancia del carbono alfa (es decir la distancia de carbono de la estructura principal) y la distancia del carbono beta (es decir la distancia entre los primeros carbonos en cada cadena lateral) de cada par de residuos también se puede calcular como un proxy de la orientación de las cadenas laterales una con respecto a la otra (ver adelante).

35 Base de Datos 2-D

Se compilan los datos relevantes en una base de datos de ADN o las secuencias de aminoácidos de un polipéptido o polipéptidos involucradas en complejos de un tipo, incluyendo el uso de la cadena lateral de residuos de datos de secuencia y derivados. Los datos relevantes en el Proceso de Selección en esta base de datos se deriva de los datos de secuencia al aplicar un valor numérico que representa las propiedades físicas de ocurrencia de cada cadena lateral de aminoácidos en cada residuo, por lo cual los polipéptidos y complejos de polipéptido de la muestra se alinean como se describió anteriormente. Esta base de datos contiene, por ejemplo, además de los datos de secuencia, los siguientes datos relevantes con mediciones estadísticas (por ejemplo media, mediana, modo, desviación estándar, máximo, y mínimo) en cada una de las siguientes características para cada par de residuo. Las mediciones estadísticas se pueden hacer y almacenar con relación a los aminoácidos presentes en cada residuo ponderadas y no se ponderadas mediante la frecuencia con la cual aparece la cadena lateral específica en este residuo.

1. Datos numéricos que se relacionan con masa/volumen de las cadenas laterales de los residuos, tal como, pero no limitándose a, la composición química, el peso molecular y los volúmenes de van der Waals (Xia X. y Li W.H.; Richards, F.M.).

50 2. Datos numéricos que se relacionan con la polaridad de las cadenas laterales de los residuos, tal como, pero no limitándose a, carga, punto isoelectrico, e hidrofobicidad (Xia X. y Li W.H.; Eisenberg, D.).

Ejemplos de otras mediciones apropiadas de cadena lateral de aminoácido que se pueden incorporar en tal una base de datos y que se pueden analizar son aromaticidad, alifaticidad, hidrogenación, y hidroxitiolación (Xia X. y Li W.H.).

Base de Datos Funcionales

- 5 Cuando es posible obtener datos funcionales que indican la importancia de un par residuo/residuo para un polipéptido o complejo de polipéptido de estructura y/o especificidad general, se incorpora preferiblemente dentro del Proceso de Selección, puesto que este mejora la exactitud de las predicciones estadísticas hechas. Tales datos se cuantifican preferiblemente, en cualquier grado posible, con respecto a residuos individuales y/o pares de residuos de un polipéptido o complejo, o con respecto a los sub-dominios o dominios que median el plegado de la proteína o las interacciones proteína-proteína, y se compilan en una base de datos adecuada.

5.9.4. Tamaño de la muestra requerida (N)

- 15 Frecuentemente la disponibilidad de los datos es limitante para este método. Sin embargo, para hacer mediciones estadísticas en una muestra de polipéptidos o complejos de polipéptido con el fin de predecir en qué pares de residuos se debe dirigir el entrecruzamiento, por supuesto, es mejor utilizar una muestra grande, puesto que producirá predicciones más exactas. Pero frecuentemente es muy laborioso acumular y/o alinear los datos de tal forma que las mediciones lleguen a ser significativas (véase más arriba). Debido a que siempre existe un rango limitado de valores, y debido a que por lo tanto su variabilidad también se limita, se pueden hacer también predicciones exactas a partir de conjuntos más pequeños de datos. Es preferible una muestra con más de 15 estructuras individuales, secuencias o unidades funcionales.

- 20 Sin embargo, se han utilizado los métodos previamente descritos para posicionar otros entrecruzamientos, tal como enlaces disulfuro, al examinar solo un polipéptido o complejo en el que se hacen las mutaciones puntuales, y esto ha dado como resultado los complejos funcionales (Pastan et al., Patente de los Estados Unidos No. 5,747,654 presentada en Mayo 5, 1998). Por lo tanto es posible hacer predicciones que pueden ser exactas en una muestra pequeña. Sin embargo, con el fin de hacer predicciones con base en estadísticas que incluyan tales mediciones como desviaciones estándar, no es significativo utilizar un tamaño de muestra menor de tres (una desviación estándar en 2 puntos de datos no es una medición significativa). Por lo tanto el mínimo de un tamaño de muestra es tres para cualquier análisis estadístico.

5.10. Proceso de selección

5.10.1. Determinación de la distancia de carbono alfa en el enlace tirosilo-tirosilo

- 30 Como se estableció anteriormente, los pares de residuos seleccionados deben tener una distancia de carbono alfa promedio cercana a la distancia de los carbonos alfa de las cadenas laterales tirosilo entrecruzadas. El rango de distancias que es posible entre los carbonos alfa de dos tirosinas entrecruzadas se calcula para la isoforma unida épsilon-épsilon del entrecruzamiento al aplicar geometría estándar, geometría de Pitágoras, y trigonometría. Los cálculos se basan en todos los enlaces carbono-carbono, enlaces de di-tirosina que forman ángulos de 120 grados debido a la estructura plana del anillo aromático con la excepción del ángulo en el carbono beta, que forma el ángulo tetrahédrico de 109.5 grados (Figura 2A).

- 40 Adicionalmente, estos cálculos toman en consideración que la estructura de la di-tirosina tiene grados significativos de libertad rotacional, y que por lo tanto la distancia entre los carbonos alfa de las dos tirosinas puede ser bastante diferente dependiendo de su conformación. Específicamente, se consideran las libertades rotacionales en los enlaces carbono beta – carbono gama, y la libertad rotacional en el enlace que conecta los anillos aromáticos. Sin embargo, son posibles otras isoformas de entrecruzamiento, que permitirían distancias aún más cercanas entre los carbonos alfa de la di-tirosina, que adicionalmente se toma en consideración en la configuración de los rangos posibles en el proceso de selección de los pares de residuos, como se describe adelante en los “Filtros”.

- 45 El ángulo χ en la figura 2C es el ángulo formado por los dos planos, cada uno definido por el eje carbono alfa – carbono alfa e individualmente por las posiciones de cada una los carbonos beta de las dos cadenas laterales tirosilo involucradas en el enlace. El ángulo ω , determinado por la libertad rotacional en el enlace di-tirosina propiamente dicho es 120° en la figura 3, y -120° en la figura 4.

- 50 Las representaciones esquemáticas de posibles configuraciones de enlaces para un ángulo ω de 120° en la figura 3 representa un ángulo χ de 180°, en ambos ángulos máximos y mínimos están en el plano proyectado. Las descripciones esquemáticas de posibles configuraciones de enlace para un ángulo ω de 120° en la figura 4 representa un ángulo χ de 0°, en el que ambos ángulos máximos y mínimos están en el plano proyectado.

Para un ángulo ω de 120° y un ángulo χ de 180° , y en la configuración en la cual la distancia del carbono alfa está en un mínimo (figura 3A), la distancia de carbono alfa es 11.74\AA : en la configuración, en la cual la distancia del carbono alfa está en un máximo (Figura 4B), la distancia del carbono alfa es 9.56\AA .

5 Para un ángulo ω de -120° y un ángulo χ de 180° , y en la configuración en la cual la distancia del carbono alfa está en un mínimo (figura 4A), la distancia de carbono alfa es 10.73\AA : en la configuración, en la cual la distancia del carbono alfa está en un máximo (Figura 4B), la distancia del carbono alfa es 5.70\AA .

5.10.2. Cálculos de ángulos de cadena lateral en el enlace tirosilo

10 Los ángulos ϕ y ψ (figura 2C) son los ángulos en cada uno de los átomos de carbono alfa entre el eje de carbono alfa – carbono alfa y el enlace de carbono alfa – carbono beta. Se calculan para las distancias máximas y mínimas entre los átomos de carbono alfa con base en la flexibilidad rotacional de los enlaces carbono – carbono en el átomo de carbono. Las descripciones esquemáticas de las posibles configuraciones de enlaces para un ángulo ω de 120° en la figura 3 representan un ángulo χ de 180° , en el que ambos ángulos máximos y mínimos están en el plano proyectado. Las descripciones esquemáticas de las posibles configuraciones de enlace para un ángulo ω de 120° en la figura 4 representan un ángulo χ de 0° , en el que ambos ángulos máximos y mínimos están en el plano proyectado.

15 Para un ángulo ω de 120° y un ángulo χ de 180° , y en la configuración en la cual la distancia del carbono alfa está en un mínimo (figura 3A), los ángulos ϕ y ψ son máximos e iguales a aproximadamente 77.1° ; en la configuración, en la cual la distancia del carbono alfa está en un máximo (figura 3B), los ángulos ϕ y ψ son mínimos e iguales, a aproximadamente 34.5° .

20 Para un ángulo ω de -120° y un ángulo χ de 0° , en el cual la distancia del carbono alfa está en un mínimo (figura 3A), los ángulos ϕ y ψ son máximos e iguales a aproximadamente 130.5° ; en la configuración, en la cual la distancia del carbono alfa está en un máximo (figura 3B), los ángulos ϕ y ψ son mínimos e iguales, a aproximadamente 10.5° .

Diferencias en las distancias alfa-alfa y beta-beta

25 Como un proxy para la orientación de las cadenas laterales, las diferencias en las distancias alfa-alfa y beta-beta (“diferencia de distancia alfa-beta”) y su rango se calculan de nuevo con base en los extremos de separación del carbono alfa para los ángulos ω de 120° y -120° figuras 3 y 4). El máximo y mínimo de la diferencia de la distancia alfa-beta se calcula para ambos ángulos ω en el que ambos anillos aromáticos de las cadenas laterales tirosilo están en el mismo plano, y en el que la diferencia alfa-beta está en sus extremos. Esta diferencia se calcula al restar el doble de la longitud del doble de la longitud b en las Figuras 3 y 4.

30 Para un ángulo ω de 120° (figura 3), y en la configuración, en la que la distancia del carbono alfa es máxima, la diferencia de la distancia alfa-beta es 2.37\AA ; en la configuración, en la que la distancia del carbono alfa es mínima, la diferencia de la distancia alfa-beta es 0.19\AA . Para un ángulo ω de -120° (figura 4), y en la configuración, en la que la distancia del carbono alfa es máxima, la diferencia de la distancia alfa-beta es 3.03\AA ; en la configuración, en la distancia del carbono alfa es mínima, la diferencia de la distancia alfa-beta es $+2.00\text{\AA}$.

35 5.11. Los filtros

En los casos donde hay datos suficientes disponibles, el proceso de selección consiste preferiblemente en una serie de pruebas estadísticas o “filtros” destinados estrechar sucesivamente los pares de residuos que más probablemente darán como resultado un par de tirosina intercadena entrecruzado de un polipéptido o complejo de polipéptido que altera de forma mínima las características estructurales del polipéptido o del complejo de polipéptido.

40 Los filtros son los siguientes:

1. Selección con base en la separación de carbono alfa de par residuo, con base en (1) las distancias máximas y mínimas en un par de tirosina entrecruzado (véase anteriormente), y (2) las distancias se miden y compilan en una base de datos en 3D. La selección se lleva a cabo sobre el promedio, la mediana, el modo, o cualquier otro valor estadístico adecuado para determinar si el par se separa probablemente en una forma tal que el entrecruzamiento distorsiona mínimamente la estructura completa. El rango óptimo de las distancias de carbono alfa de par residuo que van a ser seleccionadas se determina al promediar primero las distancias mínimas en un par de tirosina entrecruzada de la isoforma representada en la figura 2B para los ángulos ω de 120° y -120° , y luego, de forma análoga, promediar las distancias máximas, como se calculó anteriormente. Estos cálculos dan como resultado el siguiente rango óptimo: Min: 7.63\AA , Max: 11.24\AA .

45

Puesto que las distancias son posibles en un rango más grande y debido a que también son posibles otras isoformas que deberían permitir configuraciones con distancia cero, el promedio entre una distancia cero y la distancia mínima entre carbonos alfa para cualquier ángulo ω proporciona el límite inferior y la distancia máxima entre carbonos alfa para cualquier ángulo ω proporciona el límite superior del rango preferido. Por lo tanto, el rango preferido es: Min: 2.85Å, Max: 11.74Å

Adicionalmente, se ha demostrado en varios casos que la estructura de la proteína puede a menudo absorber cierta cantidad de cambios estructurales, y que no obstante se mantiene la especificidad y funcionalidad. Por lo tanto también es posible, aunque es menos preferido, introducir las cadenas laterales reactivas en pares de residuos que se separan aún más allá del el rango preferido. Dado este grado de flexibilidad estructural el rango más grande posible es: Min: 0Å, Max: 13.74Å. 2.

2. Se lleva a cabo la selección con base en la flexibilidad posicional, como ejemplo, sobre las desviaciones o rangos estándar medidos/calculados de las distancias de carbono alfa en la muestra o cualquier otra medida estadística que cuantifica la variabilidad de las distancias de los pares medidas/calculadas y compiladas en una base de datos 3D. El rango para esta selección se establece preferiblemente en una forma tal que la distancia de carbono alfa medida promedio de los pares de residuo seleccionados está dentro de menos de una desviación estándar del rango preferido. Sin embargo, 2 desviaciones estándar también son posibles como criterios de selección.

3. Selección con base en la orientación de la cadena lateral, determinada ya sea al calcular los ángulos tridimensionales relacionados con el eje del carbono alfa-carbono alfa (ángulos ψ , ϕ , y χ , como se describe en la Figura 2C), o al calcular un proxy, por ejemplo un estimativo de la orientación con base en la diferencia de la distancia alfa - beta descrita anteriormente. Se lleva a cabo la selección sobre el promedio, mediana, modo, o cualquier otro valor estadístico de los ángulos, o el proxy, adecuados para determinar si las cadenas laterales del par se orientan probablemente de tal manera que el entrecruzamiento distorsiona mínimamente la estructura completa.

El ángulo χ puede variar en 360°, y el enlace aún es posible sin ninguna distorsión de la estructura, mientras que los ángulos ψ y ϕ se ajustan de manera correspondiente. Por lo tanto, el rango de selección con base en el ángulo χ se debe establecer mediante una métrica conducida por los ángulos ψ , ϕ , y χ con un grado de flexibilidad similar a aquel para los ángulos ψ y ϕ , o para la diferencia de distancia alfa- beta, el rango para el que se describe adelante.

El rango para los ángulos ψ y ϕ es análogo al rango óptimo de las distancias de carbono alfa en el Filtro 1, óptimamente entre los promedios de los valores extremos calculados para la isoforma del par de di-tirosina representado en la Figura 2B, y para los ángulos ω de 120° y 120°. Este rango óptimo está así entre:

Min: 22.49°, Max: 103.80°.

Debido a que son posibles estos ángulos en un rango más grande aún dentro de esta isoforma del enlace di-tirosina, y ya que el rango óptimo anterior es a menudo muy restrictivo, el rango mínimo para cualquier ángulo ω proporciona el límite inferior y el ángulo máximo para cualquier ángulo ω proporciona el límite superior del rango preferido. Por lo tanto, el rango preferido es:

Min: 10.5°, Max: 130.5°.

Adicionalmente, se ha demostrado en varios casos que una estructura de proteína puede a menudo absorber una cierta cantidad de cambios estructurales, y que no obstante se mantienen la especificidad y funcionalidad. También por lo tanto es posible, aunque es menos preferido, introducir las cadenas laterales reactivas en pares residuo que tienen ángulos ψ y ϕ aún más allá del rango preferido. Dado este grado de flexibilidad estructural el rango más grande posible es:

Min: 0°, Max: 140°.

El rango óptimo de las distancias de par residuo de carbono alfa por ser seleccionado se determina al promediar primero la diferencia de distancia alfa- beta mínima en un par de tirosina entrecruzada de la isoforma representada en la Figura 2B, y para los ángulos ω de 120° y 120°, y luego, de forma análoga, promediar la diferencia de distancia alfa- beta máxima, como se calculó anteriormente. Estos cálculos resultan en el siguiente rango óptimo:

Min: 0.90Å, Max: 2.70 Å.

Debido a que son posibles diferencias de distancias en un rango más grande, y ya que el rango óptimo anterior es a menudo muy restrictivo, la diferencia de distancia alfa- beta mínima para cualquier ángulo ω proporciona el límite

inferior y la diferencia de distancia alfa-beta máxima para cualquier ángulo ω proporciona el límite superior del rango preferido. Por lo tanto, el rango preferido es:

Min: -2.00Å, Max: 3.03Å.

5 Adicionalmente, se ha demostrado en varios casos que una estructura de proteína puede a menudo absorber una cierta cantidad de cambios estructurales, y que no obstante se mantienen la especificidad y funcionalidad. Adicionalmente, son posibles otras isoformas del enlace di-tirosina. También por lo tanto es posible, aunque es menos preferido, introducir las cadenas laterales reactivas en los pares de residuo que tienen diferencia de distancia alfa- beta aún más allá del rango preferido. Dado este grado de flexibilidad estructural el rango más grande posible es: Min: -2.75Å, Max: 3.08Å.

10 4. La flexibilidad de la orientación de las cadenas laterales una hacia la otra se mide sobre el rango o desviación estándar de la muestra, como ejemplo, o cualquier otra medida estadística que cuantifique la variabilidad de las cadenas laterales de los pares medidos y compilados en una base de datos 3D. El rango para esta selección se establece preferiblemente en una forma tal que la diferencia de distancia alfa- beta medida promedio de los pares de residuo seleccionados está dentro de menos de una desviación estándar del rango preferido. Sin embargo, también son posibles 2 desviaciones estándar como un criterio de selección.

15 5. Se eliminan los pares que contienen uno o ambos residuos que están por lo menos 95% o más, preferiblemente 80% o más, posiblemente también 50% o más conservados entre los dominios en la muestra, ya que es probable que sean importantes para la arquitectura general del dominio, por ejemplo las cisteínas en la formación de enlaces disulfuro, las leucinas en la formación de cremalleras de leucina, etc.

20 6. Propiedades físicas de la cadena lateral, por ejemplo carga, hidrofobicidad, volúmenes de van der Waals, peso molecular, etc. Se lleva a cabo la selección sobre el promedio, mediana, modo, o cualquier otro valor estadístico de estas propiedades, de forma individual o combinada, adecuados para determinar si las mutaciones para tirosina y el entrecruzamiento entre un par de residuos distorsionará mínimamente la estructura completa. El grado en el cual se conserva un residuo, se mide por el rango o desviación estándar, como ejemplo, o cualquier otra medida estadística de la muestra que cuantifique la variabilidad de las propiedades físicas de las cadenas laterales que se miden y compilan en una base de datos 2-D.

25 Se puede establecer el rango, como un ejemplo, de la siguiente manera: el valor de una propiedad física para un par de tirosina (2 x valor de tirosina) se compara con el valor combinado de ambos residuos de un par, y se obtiene la diferencia mediante sustracción. La diferencia luego se compara con las desviaciones estándar combinadas del par de residuo. Debe constituirse un múltiplo más pequeño que 2 de las desviaciones estándar combinadas para la diferencia entre el valor de par de tirosina y los promedios combinados del par de residuo. Sin embargo, también se pueden utilizar medidas más directas o intuitivas, así como también medidas más sofisticadas y exactas, para clasificar y seleccionar las propiedades físicas de los pares de residuos.

30 7. Eliminación de pares de los que uno o ambos residuos están en un mínimo de 90% o más, de glicinas conservadas, preferiblemente 60% o más. La glicina es el más pequeño de los aminoácidos y no tiene carbono beta. La glicina a menudo se asocia con cambios en las estructuras de las proteínas, y la sustitución de una glicina con uno de los aminoácidos más grandes, tirosina, probablemente tendría un impacto muy grande sobre la estructura completa.

35 8. Las propiedades mencionadas estructurales y/o de conservación de la cadena lateral de aminoácidos y/o físicas del residuo/pares de residuo próximos a cada residuo/par de residuo. Se puede determinar la proximidad con respecto a las secuencias de polipéptido (2-D) y la estructura completa del polipéptido o complejo de polipéptido (3-D).

40 9. Propiedades funcionales relacionadas con el efecto de un residuo/par de residuos sobre la funcionalidad y/o especificidad del polipéptido o complejo de polipéptido.

45 5.11.1. Incorporación de datos derivados de modelado

Particularmente en las realizaciones de la presente invención, en las que se estabiliza un polipéptido único, tal como, por ejemplo, un factor de crecimiento de péptido o un biocatalizador, se pueden emplear cualquiera de los métodos para calcular y/o computar los efectos de las mutaciones y/o el entrecruzamiento sobre la estructura, estabilidad, actividad, o especificidad del polipéptido resultante. Un ejemplo de tal paquete de software es el CNS mencionado anteriormente (Adams P.D. et al. Acta Crystallogr. D. Biol. Crystallogr.; vol. 55 (Pt 1): pp. 181-90, 1999) utilizando un módulo de minimización de energía CHARMM. Los datos derivados de tales análisis se pueden utilizar para estrechar la selección o pares de residuo, y también se puede utilizar para informar los ajustes de los parámetros de selección anteriores, tales como, por ejemplo, los rangos de selección.

5.11.2. Filtros mínimamente requeridos para la selección

Dependiendo de la naturaleza del polipéptido o complejo de polipéptido, y de la disponibilidad de los datos, un subconjunto de filtros puede, sin embargo, ser suficiente para seleccionar un par adecuado para la reacción de entrecruzamiento. Por ejemplo, se puede utilizar solo un filtro con base en el promedio de la separación del carbono alfa residuo (filtro 1, anterior). También es posible hacer una selección utilizando los filtros 6 y 7 anteriores, ambos con base en el grado en el que se conservan los residuos, si están disponibles los datos estructurales para por lo menos una estructura de tal polipéptido o complejo de polipéptido. Se puede utilizar para la selección cualquiera o más de uno de los filtros anteriores y cualquier combinación de los mismos.

El orden de los filtros no es de importancia. Adicionalmente, donde se podría sumar a la calidad de la selección, los filtros anteriores se pueden dividir en dos o más filtros para poner de relieve ciertos aspectos del filtro. Los filtros se pueden combinar adicionalmente al designar métricas que cuantifiquen varios criterios simultáneamente. Por lo cual, por ejemplo, se puede refinar la selección al seleccionar un criterio que toma en cuenta el valor de otro criterio.

5.12. Construcciones del vector de ADN

La secuencia de nucleótido que codifica para el polipéptido, o para uno, cualquiera, ambos, varios o todos los polipéptidos de un complejo, o análogos funcionalmente activos o fragmentos u otros derivados de los mismos se puede insertar en unos vectores de expansión o expresión apropiados, es decir, un vector que contiene los elementos necesarios para solo la transcripción o la transcripción y traducción, de la(s) secuencia(s) que codifica(n) proteínas insertadas. Los genes naturales y/o sus secuencias de flaqueo también pueden suministrar las señales de transcripción y/o traducción necesarias.

La expresión de una secuencia de ácido nucleico que codifica un polipéptido o fragmento de péptido se puede regular mediante una segunda secuencia de ácido nucleico de tal manera que el polipéptido se expresa en un anfitrión transformado con la molécula de ADN recombinante. Por ejemplo, la expresión de un polipéptido se puede controlar mediante cualquier elemento promotor/mejorador conocido en la técnica.

Los promotores que se pueden utilizar para controlar la expresión de gen incluyen, como ejemplo, la región promotora temprana de SV40, el promotor contenido en la repetición terminal larga 3' del sarcoma de Rous, el promotor timidina quinasa del herpes, las secuencias reguladoras del gen de metalotioneina; vectores de expresión procariótica tales como el promotor β -lactamasa, o el promotor lac; vectores de expresión en plantas que comprenden el promotor de nopalina sintasa o el promotor de ARN 35S del virus mosaico de la coliflor, y el promotor de la enzima fotosintética ribulosa bifosfato carboxilasa; los elementos promotores de la levadura u otros hongos tales como el promotor Gal 4, el promotor de alcohol deshidrogenasa, promotor fosfoglicerol quinasa, promotor de fosfatasa alcalina, y las siguientes regiones de control transcripcional de animales, que exhiben especificidad de tejido y se han utilizado en animales transgénicos: región de control del gen elastasa I que está activa en las células acinares pancreáticas (Swift et al. Cell; vol. 38: pp. 639-646, 1984); una región de control de gen que está activa en las células beta pancreáticas (Hanahan D., Nature; vol. 315: pp. 115-122, 1985), una región de control de gen de inmunoglobulina que está activa en las células linfoides (Grosschedl R. et al. Cell; vol. 38: pp. 647-658, 1984), región de control de virus de tumor mamario en ratón que está activa en células testiculares, de mama, linfoides y mastocitos (Leder A. et al. Cell; vol. 45: pp. 485-495, 1986), región de control de gen de albúmina que está activa en el hígado (Pinkert C.A. et al. Genes Dev.; vol. 1: pp. 268-276, 1987), región de control de gen de alfa-fetoproteína que está activa en el hígado (Krumlauf R. et al. Mol. Cell. Biol.; vol. 5: pp. 1639-1648, 1985); región de control de gen de alfa 1-antitripsina que está activa en el hígado (Kelsey G.D. et al. Genes Dev.; vol. 1: pp. 161-171, 1987), región de control de gen de beta-globina que está activa en células mieloides (Magram J. et al. Nature; vol. 315: pp. 338-340, 1985); región de control de gen de proteína básica de mielina que está activa en las células de los oligodendrocitos en el cerebro (Readhead C. et al. Cell; vol. 48: pp. 703-712, 1987); región de control de gen de cadena ligera de miosina 2 que está activa en el sistema musculoesquelético (Shani M. Nature; vol. 314: pp. 283-286, 1985), y región de control de gen de hormona de liberación gonadotrópica que está activa en el hipotálamo (Mason A.J. et al. Science; vol. 234: pp. 1372-1378, 1986).

En una realización específica, se utiliza un vector que comprende un promotor enlazado de forma operativa a un ácido nucleico de un gen, uno o más orígenes de replicación, y, opcionalmente, uno o más marcadores seleccionables (por ejemplo, un gen resistente a antibióticos). En las bacterias, el sistema de expresión puede comprender el sistema de respuesta lac para la selección de bacterias que contiene el vector. Se pueden hacer construcciones de expresión, por ejemplo, al subclonar una secuencia codificadora en uno de los sitios de restricción de cada uno o cualquiera de los vectores pGEX (Pharmacia, Smith D.B. y Johnson K.S. Gene; vol. 67: pp. 31-40, 1988). Esto permite la expresión del producto de proteína.

Se pueden identificar vectores que contienen insertos de gen mediante tres métodos generales: (a) identificación de uno o varios atributos específicos del ADN en sí mismo, tal como, por ejemplo, longitudes de fragmento producidas por el tratamiento de restricción de endonucleasa, secuenciación directa, PCR, o hibridación de ácido nucleico; (b)

5 presencia o ausencia de funciones de gen "marcadoras"; y, cuando el vector es un vector de expresión, (c) expresión de secuencias insertadas. En el primer método, se puede detectar la presencia de un gen insertado en un vector, por ejemplo, mediante secuenciación, PCR o hibridación de ácido nucleico utilizando sondas que comprenden secuencias que son homólogas a un gen insertado. En el segundo método, el vector recombinate/sistema de anfitrión se puede identificar y seleccionar con base en la presencia o ausencia de ciertas funciones de gen "marcadoras" (por ejemplo, actividad de timidina quinasa, resistencia a antibióticos, fenotipo de transformación, formación de cuerpo de oclusión en baculovirus, etc.) originadas por la inserción de un gen en el vector. Por ejemplo, si se inserta el gen dentro de la secuencia de gen marcadora del vector, los recombinantes que contiene el inserto se identifican por la ausencia de la función de gen marcadora. En el tercer método, se pueden identificar los vectores de expresión recombinantes al ensayar el producto expresado por los vectores de expresión recombinantes que contienen las secuencias insertadas. Se pueden basar tales ensayos, por ejemplo, en las propiedades físicas o funcionales de la proteína en sistemas de ensayo in vitro, por ejemplo, unión con anticuerpo anti-proteína.

15 Una vez se identifica y aísla una molécula de ADN recombinante particular, se pueden utilizar varios métodos conocidos en la técnica para propagarla. Una vez se establece un sistema anfitrión adecuado y condiciones de crecimiento, los vectores de expresión recombinante se pueden propagar y preparar en cantidad. Algunos de los vectores de expresión que se pueden utilizar incluyen virus humanos o animales tales como virus vaccinia o adenovirus; los virus de insectos tales como baculovirus; vectores de levadura; vectores bacteriófagos (por ejemplo, fago lambda), y vectores de ADN plásmidos y cósmidos.

20 Una vez se identifica el vector recombinante que dirige la expresión de una secuencia deseada, se puede analizar el producto de gen. Esto se logra mediante ensayos con base en las propiedades físicas y funcionales del producto, que incluyen el etiquetado radioactivo del producto seguido por electroforesis en gel, inmunoensayo, etc.

5.13. Sistemas de expresión de gen y purificación de proteínas

25 Se puede utilizar una variedad de sistemas de hospedaje de vector para expresar la las secuencias que codifican proteína. Esto incluye, como ejemplo, sistemas de células de mamíferos infectados con virus (por ejemplo, virus vaccinia, adenovirus, etc.); sistemas de células de insectos infectados con virus (por ejemplo, baculovirus); microorganismos tales como levaduras que contienen vectores de levaduras, o bacterias transformadas con bacteriófago, ADN, ADN de plásmido, o ADN de cósmido. Los elementos de expresión de los vectores varían en sus resistencias y especificidades. Dependiendo del sistema de anfitrión-vector utilizado, se puede utilizar uno cualquiera de un número de elementos de transcripción y traducción adecuados.

30 En una realización específica, se puede expresar el gen en las bacterias que son deficientes en proteasa y que tienen bajos niveles constitutivos y altos niveles de expresión inducidos donde se utiliza un vector de expresión que es inducible, por ejemplo, mediante la adición de IPTG al medio.

35 En aún otra realización específica, el polipéptido, o uno, cualquiera, ambos, varios todos los polipéptidos de un complejo se pueden expresar con péptidos de señal, tales como, por ejemplo, péptido de señal bacteriana pelB, que dirige la proteína al periplasma bacteriano (Lei et al. J. Bacterol., vol. 169: pp. 4379, 1987). Alternativamente, se le puede permitir a la proteína formar cuerpos de inclusión, y posteriormente ser resolubilizada y replegada (Kim S.H. et al. Mo Immunol, vol. 34: pp. 891, 1997).

40 En aún otra realización específica, se expresa un fragmento del polipéptido, o uno, cualquiera, ambos, varios o todos los polipéptidos de un complejo que comprende uno o más dominios de la proteína. Cualquiera de los métodos descritos previamente para la inserción de fragmentos de ADN en un vector se pueden utilizar para construir vectores de expresión que contienen un gen quimérico que consiste de señales de control de transcripción/traducción apropiadas y secuencias que codifican proteínas. Estos métodos pueden incluir técnicas sintéticas y de ADN recombinante in vitro y recombinantes in vivo (recombinación genética).

45 Adicionalmente, se puede seleccionar una cepa celular anfitriona que modula la expresión de las secuencias insertadas o modifica o procesa el producto de gen en la forma específica deseada. La expresión de ciertos promotores se puede elevar en la presencia de ciertos inductores; así, se puede controlar la expresión de los polipéptidos creados por ingeniería genética. Adicionalmente, las células anfitrionas diferentes tienen características y mecanismos específicos para el procesamiento de traducción y de post-traducción y modificación (por ejemplo, glucosilación, fosforilación de proteínas). Se pueden seleccionar las estirpes celulares o sistemas de anfitrión apropiados para asegurar la modificación y procesamiento deseado para los polipéptidos extraños expresados. Por ejemplo, la expresión en un sistema bacteriano se puede utilizar para producir un producto de proteína con núcleo no glucosilado. La expresión en la levadura producirá un producto glucosilado. La expresión en las células de mamífero se puede utilizar para asegurar la glucosilación "natural" de una proteína heteróloga. Adicionalmente, diferentes sistemas de expresión de vector/anfitrión pueden efectuar las reacciones de procesamiento a diferentes grados.

5 En otras realizaciones de la invención, el polipéptido, o uno, cualquiera, ambos, varios o todos los polipéptidos de un complejo, y/o fragmentos, análogos, o derivados de los mismos se pueden expresar como un producto de proteína de fusión, quimérico (que comprende la proteína, fragmento, análogo, o derivado unido a través de un enlace de péptido a una secuencia de proteína heteróloga de una proteína diferente). Tal producto quimérico se puede hacer al ligar una a otra las secuencias de ácido nucleico apropiadas que codifican las secuencias de aminoácidos deseadas mediante métodos conocidos en la técnica, en el marco de codificación apropiado, y que expresan el producto quimérico mediante métodos habitualmente conocidos en la técnica. Alternativamente, se puede hacer tal producto quimérico mediante técnicas de síntesis de proteína, por ejemplo, mediante el uso de un sintetizador de péptidos.

10 Los polipéptidos de un complejo se pueden expresar juntos en las mismas células ya sea en el mismo vector, conducir por las mismas señales de transcripción y/o traducción o señales de transcripción y/o traducción independientes, o en vectores de expresión separados por ejemplo mediante cotransfección o cotransformación y selección, por ejemplo, puede ser con base en ambos marcadores de selección individual de los vectores. Alternativamente, uno, cualquiera, ambos, varios o todos los polipéptidos de un complejo se pueden expresar separadamente; se pueden expresar en el mismo sistema de expresión o en sistemas de expresión diferentes, y se pueden expresar individualmente y colectivamente como fragmentos, derivados o análogos del polipéptido original.

5.14. Reacción de entrecruzamiento

5.14.1. Introducción de mutaciones puntuales a la fenilalanina

20 Uno de los codones de cada par de residuos tirosina que puede reaccionar uno con otro y originar distorsiones estructurales y/o funcionales indeseables se muta de forma puntual preferiblemente a codones que dirigen la expresión de fenilalanina.

25 Se pueden introducir mutaciones puntuales en el ADN que codifica el polipéptido, o uno, cualquiera, ambos, varios o todos los polipéptidos de un complejo mediante cualquier método conocido en la técnica, tal como mutagénesis dirigida a sitio mediada por oligonucleótidos. Tales métodos pueden utilizar oligonucleótidos que son homólogos a las secuencias de flanqueo de tales codones, pero que codifican la tirosina en el sitio o sitios seleccionados. Con estos oligonucleótidos, los fragmentos de ADN que contienen la mutación o mutaciones puntuales se amplifican e insertan en el gen o genes, por ejemplo, al subclonar. Un ejemplo de tales métodos es la aplicación del Equipo de Mutagénesis Dirigida a Sitio de QuikChange™ (Stratagene, Catalog # 200518); este equipo utiliza la enzima Pfu que no tiene acción de desplazamiento en cadena sencilla en ninguna mutación de plásmido de doble cadena en las reacciones PCR. Otros métodos pueden utilizar otras enzimas tales como polimerasas de ADN, o fragmentos y/o análogos de los mismos.

30 El plásmido o plásmidos que contienen la mutación puntual o mutaciones puntuales, por ejemplo, se transforman en bacterias para expansión, y se prepara el ADN como se describió anteriormente. Se puede examinar el ADN aislado, expandido y preparado para verificar que codifica el polipéptido o polipéptidos del complejo, y que se logra la mutación o mutaciones correctas. Esto, por ejemplo, se puede verificar mediante técnicas de secuenciación directa de ADN, hibridación de ADN, o cualquier otro método conocido en la técnica.

5.14.2. Purificación de productos de gen

40 Se puede aislar y purificar el producto de gen mediante métodos estándar que incluyen cromatografía (por ejemplo, cromatografía de intercambio iónico, de afinidad, de separación por tamaño en columna) precipitación con sulfato de amonio, centrifugación, solubilidad diferencial, o mediante cualquier otra técnica estándar para la purificación de proteínas.

45 Se pueden evaluar las propiedades funcionales utilizando cualquier ensayo adecuado. La secuencia de aminoácidos de la proteína se puede deducir a partir de la secuencia de nucleótidos del gen quimérico contenido en el vector recombinante. Como resultado se puede sintetizar la proteína mediante métodos químicos estándar conocidos en la técnica (por ejemplo, véase Hunkapiller m. Et al. Nature; vol. 310(5973); pp. 105-11, 1984).

5.14.3. Reacción

50 La reacción de entrecruzamiento puede utilizar cualquier reacción química o física conocida en la técnica que introduce específicamente entrecruzamientos de ditirosina, tales como entrecruzamiento catalizado por peroxidasa, o fotodinámicamente en la presencia o ausencia de sintetizadores (véase Sección II). Preferiblemente, sin embargo, la reacción se cataliza mediante un complejo metalo- ión, como se describe en detalle adelante.

Los polipéptidos parcialmente purificados que contienen residuos de tirosina apropiados se pueden equilibrar mediante diálisis en un regulador, tal como solución salina regulada con fosfato (PBS), juntos o en forma separada antes de mezclarlos. Luego se agrega el catalizador (sobre hielo o de otra forma). El catalizador de la reacción es cualquier compuesto que resultará en la reacción de entrecruzamiento anterior. El catalizador puede tener los componentes estructurales que transmiten la especificidad de la reacción, proporcionados de manera general por una estructura que compleja un ión de metal, y la capacidad de abstraer un electrón del sustrato en la presencia de un reactivo de oxidación, proporcionado de manera general por el ión de metal. Un metal activo está encerrado en un ligando estable que bloquea la unión no específica a los sitios de quelación sobre las superficies de la proteína. Por ejemplo, ya sea una metaloporfirina, tal como, pero no limitándose a, cloruro de 20-tetrakis (4-sulfonatofenil)-21H,23H- porfina manganeso (III) (MnTPPS) o cloruro de protoporfirina IX de hemina de hierro (III) (Campbell L.A. et al. *Bioorganic and Medicinal Chemistry*, vol. 6: pp. 1301-1037, 1998), o un complejo de péptido- ión de metal, tal como el tripéptido NH₂-Gly-Gly-His-COOH que compleja Ni⁺⁺ puede servir como catalizador de la reacción. Las metaloporfirinas son una clase de complejos de metal de ligandos oxidativos para los cuales existen pocos, o ningún, sitios de afinidad en las proteínas eucarióticas de origen natural. La reacción también se puede catalizar mediante complejos de péptido Ni⁺⁺ intramoleculares, tales como aminoácidos N y C terminales que consisten de 3 o más residuos de histidina (his-tag), o de los tripéptidos GGH anteriores. La reacción se inicia mediante la adición del reactivo de oxidación a temperatura ambiente o de otra forma. Los reactivos de oxidación incluyen, pero no se limitan a, peróxido de hidrógeno, oxona, y hexahidrato de ácido monoperóxifáltico de magnesio (MMPP) (Brown K.C. et al. *Biochem.*; vol. 34(14): pp. 4733-4739, 1995). Se puede alcanzar mayor especificidad al utilizar un oxidante fotogenerado, tal como el oxidante utilizado en el proceso descrito por Fancy D. y Thomas Kodadek, que involucra breve fotólisis del dicatión de tris-bipiridilrutenio (II) con luz visible en la presencia de un aceptor de electrones, tal como persulfato de amonio (Fancy D.A. y Kodadek T. *Proc. Natl. Acad. Sci., U.S.A.*; vol. 96: pp. 6020-24, 1999). El periodo de reacción óptimo se determina preferiblemente para cada aplicación; sin embargo, en los casos en donde no es posible un proceso de optimización, la reacción preferiblemente se debe detener después de un minuto. Utilizando un oxidante fotogenerado, tal como se describió anteriormente, la exposición a la luz puede ser menor de un segundo. La reacción se detiene mediante la adición de una cantidad suficiente de un agente de reducción, tal como b-mercaptoetanol, para contraatacar y/o neutralizar el agente de oxidación.

Alternativamente, la reacción se puede detener mediante la adición de un agente de quelación, tal como, por ejemplo, EDTA o EGTA. La solución se equilibra de nuevo mediante diálisis en un regulador, tal como solución salina regulada por fosfato (PBS), para retirar los reactivos requeridos para la reacción de entrecruzamiento, tales como el reactivo de oxidación, el catalizador, o el ión de metal, los agentes de reducción, reactivos de quelación, etc. Las condiciones de reacción de entrecruzamiento se juntan preferiblemente de tal manera que los polipéptidos o polipéptidos de un complejo que han mutado para eliminar las cadenas laterales tirosilo indeseadas ya no forman un enlace. Estas condiciones se ajustan al variar la temperatura de reacción, pH, o condiciones de osmolaridad, o al variar la concentración de los polipéptidos, el catalizador, el agente de oxidación, o cualquier otro reactivo que se aplique a tal reacción. El catalizador es una molécula pequeña que se difunde fácilmente, y se puede utilizar en concentraciones variantes. Los núcleos hidrófobos de polipéptido cuidadosamente empacados tienen un grado de accesibilidad de solvente. Esto se puede modular mediante cualquier método conocido en la técnica, que incluye, pero no se limita a, alteración de la temperatura de reacción, o la adición de sales, detergentes, desoxicolato, o guanidinio.

5.15. Alcanzar un polipéptido o complejo estabilizado

5.15.1. Mutación puntual a tirosina y purificación de producto genético

Los codones de los residuos identificados como un par adecuado al que se debe dirigir el entrecruzamiento, como se describió anteriormente, y seleccionar para una realización particular de la presente invención, mutan puntualmente de tal manera que los pares de residuo resultantes dirigen la expresión de cadenas laterales de tirosilo. Se introducen mutaciones puntuales como se describió anteriormente.

Los productos genéticos se purifican de nuevo como se describió anteriormente.

5.15.2. Entrecruzamiento del polipéptido o complejo

Los polipéptidos que ahora contienen cadenas laterales de tirosilo en los residuos a los cuales se puede dirigir la reacción de entrecruzamiento se someten a la reacción de entrecruzamiento bajo las condiciones determinadas como se describió anteriormente y se llevan a cabo, también como se describió anteriormente. La eficiencia de la reacción se puede examinar, por ejemplo, mediante experimentos de inmunotransferencia Western, en los que se puede ejecutar un complejo entrecruzado en aproximadamente el peso molecular de ambos o todos los polipéptidos del complejo. Si se forma el enlace fácilmente bajo las condiciones anteriores, la resistencia de la reacción aún se ajusta adicionalmente a la resistencia mínimamente requerida.

En las realizaciones de la invención en donde el entrecruzamiento se dirige a pares de residuos que están ocultos y/o no son fácilmente accesibles al catalizador o reactivos de oxidación, se puede disociar temporalmente la estructura de polipéptido de orden secundario y mayor para permitir el acceso del reactivo. Por ejemplo, tal método puede ser necesario al dirigir el entrecruzamiento al núcleo hidrófobo de un polipéptido único o a un par de residuos ocultos del complejo de polipéptido que tiene muy alta afinidad entre las subunidades. Se puede utilizar cualquier medio conocido en la técnica para desnaturalizar de forma reversible la estructura de polipéptido para permitir el acceso del reactivo a los pares de residuo ocultos. Tales medios incluyen, pero no se limitan a, manipular (incrementar o reducir) la concentración de sal o la temperatura de reacción, o emplear detergentes, o agentes tales como guanidina HCl. A medida que se eliminan las condiciones de desnaturalización (por ejemplo, mediante diálisis) y el polipéptido o complejo comienza a replegarse/reasociarse, se pueden agregar el catalizador y los reactivos de oxidación, como se describió anteriormente.

5.16. Purificación de complejos entrecruzados

El polipéptido o complejo entrecruzado se puede aislar y purificar a partir de proteínas en la reacción que falla en entrecruzarse, o de cualquier otro producto secundario indeseable, mediante métodos estándar que incluyen cromatografía (por ejemplo, cromatografía de separación por tamaño en columna, gradientes de glicerol, afinidad), centrifugación, o mediante otras técnicas estándar para la purificación de proteínas. En realizaciones específicas puede ser necesario separar los polipéptidos que no se entrecruzan, pero que se homo- o heterodimerizan con otros polipéptidos debido a la alta unión de afinidad. Se puede lograr la separación por cualquier medio conocido en la técnica, que incluye, por ejemplo, la adición de detergente y/o agentes de reducción.

Se puede determinar el rendimiento de los polipéptidos o complejos entrecruzados funcionalmente mediante cualesquier medios conocidos en la técnica, por ejemplo, al comparar la cantidad de complejo estabilizado, purificado como se describió anteriormente, con el material de partida. Las concentraciones de proteína se determinan mediante procedimientos estándar, tales como, por ejemplo, ensayos de proteína Bradford o Lowry. El ensayo Bradford es compatible con agentes de reducción y agentes de desnaturalización (Bradford, M. Anal. Biochem.; vol. 72: pp. 248, 1976), el ensayo Lowry tiene más compatibilidad con detergentes y la reacción es más lineal con respecto a las concentraciones de proteína y lectura (Lowry, O. J. Biol. Chem.; vol. 193: pp. 265, 1951).

5.17. Ensayo de un polipéptido o complejo entrecruzado

5.17.1. Función retenida

Funcionalidad

Dependiendo de la naturaleza del polipéptido o complejo de polipéptido, se puede probar la funcionalidad retenida, por ejemplo, al comparar la funcionalidad del complejo entrecruzado, entrecruzado como se describió anteriormente, con aquel del polipéptido o complejo antes de estabilización, entrecruzado o estabilizado mediante otro método, o estabilizado naturalmente mediante una modificación post-translacional que, por ejemplo, regula la asociación de ciertos polipéptidos. Los ensayos para la funcionalidad retenida se pueden basar, por ejemplo, en las propiedades bioquímicas en sistemas in vitro. Alternativamente, el polipéptido o complejo se puede probar para la funcionalidad al utilizar sistemas de ensayo biológico. Por ejemplo, la actividad de una quinasa se puede probar en ensayos de quinasa in vitro, y un factor de crecimiento, tal como un miembro de la familia IL-8, se puede probar para actividad en ensayos de migración celular quimiotácticos o ensayos de liberación de beta-glucuronidasa (Leong S.R. et al. Protein Sci.; vol. 6(3): pp. 609-17, 1997). Como otro ejemplo, se puede determinar la actividad enzimática retenida de un biocatalizador mediante cualquier método conocido por un experto en la técnica. La actividad de una enzima preferiblemente se mide directamente al comparar la actividad de la enzima sobre un sustrato antes y después de estabilización, y la cuantificación del producto de la reacción. Como ejemplo, tales ensayos incluyen, pero no se limitan a, visualización luego de separación cromatográfica de los compuestos en la reacción, análisis espectrofotométricos y fluorométricos de los productos de reacción, análisis de marcadores detectables incorporados o liberados, tales como, por ejemplo, isótopos radioactivos. Los métodos indirectos, que incluyen, pero no se limitan a, análisis computacional, estructural u otros análisis termodinámicos, también se pueden utilizar para la determinación de la actividad del biocatalizador estabilizado. Más específicamente, como ejemplo de un biocatalizador, la actividad de una lipasa, o específicamente la actividad de carboxilesterasas que catalizan la hidrólisis de acilglicérols de cadena larga, se determina mediante cualquier método conocido en la técnica, que incluye, pero no se limita a la medición de la hidrólisis de p-nitrofenilésteres de ácidos grasos con varias longitudes de cadena (\geq C-10) en solución mediante detección espectrofotométrica de p-nitrofenol a 410 nm. Cuando sea necesario distinguir entre lipasas y esterasas, el derivado triglicérido éster de resorufina de ácido 1,2-O-dilauril-rac-glicero-3-glutámico (disponible de Boehringer Mannheim Roche GmbH, Alemania), también se puede utilizar como un sustrato, proporcionando resorufina, que se puede determinar espectrofotométricamente a 572 nm, o fluorométricamente a 583 nm (Jaeger K-E et al. Annu. Rev. Microbiol. 1999. 53: pp. 315-51).

Especificidad

Dependiendo de la naturaleza del polipéptido o complejo de polipéptido, se puede probar la especificidad retenida, por ejemplo, al comparar la especificidad del polipéptido o complejo entrecruzado con la del polipéptido o complejo antes de estabilización, entrecruzado o estabilizado por otro métodos, o estabilizado naturalmente por una modificación post-translacional. Los ensayos para la especificidad retenida se pueden basar, por ejemplo, en la especificidad del sustrato enzimático, o procedimientos tipo ELISA. Por ejemplo, la especificidad retenida o resultante de una lipasa (carboxilesterasa) se puede determinar mediante cualquier método conocido por un experto en la técnica. Ejemplos no limitantes de tales métodos incluyen utilizar un número de alquildiacilglicerol fluorogénicos como sustratos para un análisis de la electroselectividad del biocatalizador. Para una descripción detallada de tales métodos y de ciertos tales compuestos, véase el artículo "New fluorescent glycerolipids for a dual wavelength assay of lipase activity and stereoselectivity" (Zandonella G. et al., 1997, J. Mol. Catal. B: Enzym. 3: pp. 127-30).

5.17.2. Estabilidad

In vitro

Se puede probar la estabilidad del polipéptido o complejo in vitro en, por ejemplo pero no limitándose a, experimentos en función del tiempo que incuban el polipéptido o complejo a concentraciones y temperaturas variadas. La estabilidad del polipéptido o complejo se puede probar en diversos niveles de pH y bajo diversas condiciones de oxidorreducción. Para todas las condiciones anteriores, los niveles restantes de polipéptidos o complejos de polipéptido funcionales se determinan al ensayar como se describió anteriormente (Funcionalidad). En el anterior ejemplo de un biocatalizador, se puede determinar la estabilidad mejorada o alterada de un polipéptido o complejo estabilizado mediante cualquier método conocido por un experto en la técnica. Tales métodos incluyen, pero no se limitan a, análisis calorimétricos y/o estructurales, cálculos termodinámicos y análisis, y comparación de las actividades de las enzimas estabilizadas y no estabilizadas bajo sus condiciones óptimas y bajo condiciones de reacción subóptimas o adversas, tales como mayor o menor temperatura, presión, pH, concentración de sal, compuesto inhibidor, o concentración de enzima y/o sustrato. Cualquiera de los anteriores análisis también puede incluir experimentos en función del tiempo dirigidos a la determinación de la vida media y/o vida útil del biocatalizador estabilizado. La estabilización de un biocatalizador de acuerdo con la invención también se puede evaluar en el contexto de otros métodos de estabilización de biocatalizador. Como ejemplos no limitantes, se pueden probar las anteriores actividades enzimáticas en los geles de inmovilización u otras matrices o en solventes orgánicos parciales o puros. Adicionalmente, un biocatalizador estabilizado mediante cualquiera de los métodos conocidos en la técnica (tal como evolución dirigida o mutagénesis designada, véase Antecedentes) también se pueden someter a los métodos de la presente invención para lograr estabilización adicional.

In vivo

Se prueban mejor las aplicaciones farmacéuticas y terapéuticas in vivo o bajo condiciones que se asemejan a las condiciones fisiológicas (véase también, adelante). La estabilidad del polipéptido o complejo se puede probar en, por ejemplo pero no limitándose a, suero, que incubaba el polipéptido o complejo en experimentos en función del tiempo a varias temperaturas (por ejemplo 37, 38, 39, 40, 42, y 45° C), y en diferentes concentraciones de suero, y que se ensayan para los niveles restantes de los polipéptidos o complejos funcionales. Adicionalmente, la estabilidad de un polipéptido o complejo en el citoplasma se puede probar en los experimentos en función del tiempo en los lisados celulares, sometidos a lisis bajo diversas condiciones (por ejemplo diversas concentraciones de diversos detergentes) a diferentes temperaturas (por ejemplo 37, 38, 39, 40, 42, y 45° C), y ensayando en los niveles restantes de polipéptidos o complejos funcionales. Más directamente, se puede probar la estabilidad en el citoplasma en experimentos en función del tiempo al raspar las células de cultivo de tejido cargadas con polipéptido o complejo estabilizado y ensayando en los niveles restantes de función. La estabilidad del polipéptido o complejo también se puede probar al inyectarlo en un animal experimental y probar la actividad específica. Alternativamente, se puede recuperar el compuesto del animal en un punto de tiempo o varios puntos de tiempo y se prueban la actividad y estabilidad, como se describió anteriormente.

5.17.3. Biodistribución

Para determinar la utilidad de un polipéptido o complejo de polipéptido estabilizado más directamente, se puede determinar la biodistribución y/u otros atributos farmacocinéticos. En una realización específica, un polipéptido o complejo de polipéptido estabilizado se puede inyectar en un organismo modelo y se puede ensayar al rastrear un marcador, tal como, pero no limitándose a, radiomarcadores ¹²⁵I o ¹⁸F (Choi C.W. et al. Cancer Research, vol. 55: pp. 5323-5329, 1995), y/o al rastrear la actividad como se describió anteriormente (Colcher D. et al. Q.J. Nucl. Med. vol. 44(4): pp. 225-241, 1998). Se puede obtener información importante, por ejemplo, al determinar la cantidad de polipéptido o complejo de polipéptido funcional que se puede esperar que sea farmacéuticamente activo debido a su penetración del tejido específicamente objetivado, tal como, por ejemplo, un tumor. También se pueden determinar

en este contexto la vida media en la circulación y en el tejido específicamente buscado, depuración renal, inmunogenicidad, y velocidad de penetración.

5.17.4. Estudios animales y clínicos

5 La utilidad de un polipéptido o complejo se puede determinar directamente al medir su actividad farmacológica ya sea en estudios en animales o clínicos. En una realización específica, tales mediciones pueden incluir, por ejemplo, mediciones con pro- o regresión de tumor que se monitorizan luego de tratamiento de un modelo animal p de uno o varios pacientes con un polipéptido o complejo estabilizado designado como un agente farmacológico anti-neoplásico. En otra realización, tales mediciones pueden incluir, por ejemplo, mediciones de masa ósea, tales como mediciones con rayos x, luego de tratamiento de un modelo de animal o uno o varios pacientes compuesto un polipéptido o complejo estabilizado designado como un agente farmacológico anti-pérdida ósea por menopausia.

5.18. Solución de problemas

5.18.1. Polipéptido o complejo no entrecruzado

15 Si un polipéptido o complejo no llega a ser entrecruzado o estabilizado por la reacción anteriormente descrita (como se determina, por ejemplo, mediante electroforesis en gel de poliacrilamida en dodecil sulfato de sodio no reductor o SDS-PAGE), pueden existir varias explicaciones y soluciones al problema.

Ajuste de Condiciones de Concentración de sal/ Osmolaridad y/o pH del Polipéptido

20 Para la estabilización de un complejo de polipéptido, un problema que se puede encontrar es que los polipéptidos, cuando no son estabilizados, no forman un complejo suficientemente estable en solución para el entrecruzamiento para formarse bajo las presentes condiciones de la reacción. En este caso, se indica que se deben variar las condiciones de reacción. Se puede determinar muy floja una asociación entre los miembros del complejo, por ejemplo, al inmunoprecipitar uno de los polipéptidos y probar la presencia y cantidad relativa de los otros polipéptidos en el precipitado (por ejemplo, mediante transferencia Western). También puede ser posible incrementar la resistencia de asociación entre polipéptidos utilizando cualesquier medios conocidos en la técnica. Tales medios pueden incluir, pero no se limitan a, ajustar ciertas condiciones de la reacción, tal como concentración de sal, Tris, polipéptido, o pH. Si con ello la resistencia de la asociación de polipéptidos se incrementa, por ejemplo, como se determina mediante SDS PAGE no reductor, la reacción de entrecruzamiento se debe juzgar de nuevo bajo estas condiciones.

30 También puede ocurrir el problema opuesto. Esto es, los polipéptidos de un complejo (o la estructura de orden secundario o mayor de un polipéptido único), se pueden asociar uno con el otro muy firmemente de tal manera que las cadenas laterales de tirosilo no se exponen al catalizador o a los reactivos de oxidación y no se forma el enlace di-tirosina. En tales casos, la estructura de orden secundario o mayor debe ser "aflojada" o parcialmente desnaturalizada. Esto se puede lograr utilizando cualesquier medios conocidos en la técnica, como se describió anteriormente, al ajustar, por ejemplo, la concentración de sal, detergente, guanidina HCl, y/o cualquier otro agente que origine desnaturalización reversible (por ejemplo temperatura, presión, y/o tiempo de reacción). También puede ser posible agregar el agente de oxidación y el catalizador en un punto de tiempo temprano o posterior. Por ejemplo, ya que se invierten las condiciones de desnaturalización, como se describió anteriormente, y el polipéptido o complejo de polipéptido comienza a replegarse/reasociarse.

Resistencia de incrementada de las condiciones de reacción

40 Si el entrecruzamiento no se forma a pesar del plegamiento de polipéptido apropiado de o formación de complejo adecuada bajo las condiciones de la reacción, la siguientes solución sería incrementar la resistencia de las condiciones de la reacción, por ejemplo al incrementar la concentración del reactivo de oxidación y/o del catalizador. Un método preferido aún puede utilizar la resistencia mínima de la reacción requerida para formar el entrecruzamiento.

Reacción de entrecruzamiento directa para un par de residuo alternativo par de residuos

45 Una reacción de entrecruzamiento puede fallar algunas veces debido a que el entrecruzamiento se dirige a un par de tirosinas que no se pueden entrecruzar debido a las consideraciones estructurales no capturadas en el proceso de selección. Si los métodos anteriores no permiten formar el entrecruzamiento entre los residuos seleccionados de un par, se puede seleccionar otro par de residuo, y se juzga de nuevo la reacción de entrecruzamiento, cuando sea necesario ajustar las condiciones de reacción, como se describió anteriormente.

Método combinado

Puede ser necesario emplear varios de los anteriores métodos para la solución de los problemas para lograr la estabilización deseada del enlace de di-tirosina.

5.18.2. Funcionalidad comprometida de polipéptido o complejo

Resistencia reducida de las condiciones de reacción

- 5 Reducir la resistencia de la reacción al ajustar, por ejemplo, pero no limitándose a, la concentración de ya sea el catalizador o el reactivo de oxidación, la temperatura, presión, y/o tiempo de reacción, puede dar como resultado un polipéptido o complejo estabilizado de polipéptido con mejor funcionalidad retenida.

Ajustes de las condiciones de concentraciones de proteína, sal/osmolaridad y/o pH

- 10 Las reacciones entrecruzadas no específicas pueden comprometer la funcionalidad del polipéptido o complejo de polipéptido, que puede ocurrir bajo ciertas condiciones de reacción, tales como, pero no limitadas a, altas concentraciones de proteína relacionadas con las concentraciones óptimas, ciertos niveles de pH, o sal, detergente, desnaturalización, y/o cualesquier otras concentraciones de los componentes en la reacción. Estas condiciones se pueden ajustar para minimizar o eliminar la formación de enlaces de ditirosina comprometidos no específicos.

Reacción directa del entrecruzamiento para un par de residuos alternativo

- 15 Si los datos de entrada para el proceso de selección son menores que la exactitud completa, o por cualquier otra razón, el par de residuos seleccionado podría producir un entrecruzamiento que distorsiona la estructura completa del polipéptido o complejo de polipéptido, y por lo tanto compromete o altera su funcionalidad. Si este fuera el caso, se debe seleccionar y mutar otro par de residuos de tal manera que ambos residuos codifiquen la tirosina, y se debe juzgar de nuevo la reacción de entrecruzamiento

20 Método combinado

Puede ser necesario emplear varios o todos los métodos anteriores para la lograr la solución de problemas deseada.

5.19. Software para el proceso de selección

- 25 Esta invención proporciona un software que permite la solución automática de los pares de residuo adecuados en los que un se puede colocar un enlace de di-tirosina. Se puede utilizar tal software de acuerdo con los criterios geométricos físicos y químicos descritos anteriormente (véase especialmente identificación de pares de residuo adecuados para la reacción), y un diagrama de flujo de selección de par de residuos tal como se establece en la sección 6 adelante. Como se describió anteriormente, se implementa una matriz sucesiva de filtros y los pares de residuo que "pasan" a través de los filtros comprenden los pares de residuo seleccionados (figura 14, lado izquierdo). Alternativamente, se puede implementar los filtros para procesar todos los pares de residuo en una matriz paralela (figura 14, lado derecho). Los pares de residuo que "pasan" a través de un filtro definen ese conjunto de filtro de pares de paso. En una realización preferida, los pares de residuo que están en todos los conjuntos de paso de filtros (es decir pares de residuo que forman la intersección de todos los conjuntos de filtro) son los pares seleccionados. Los requerimientos de filtro son como se describió anteriormente (identificación de pares de residuo adecuados para la reacción).

35 5.20. Composición farmacéutica

- 40 En una realización, esta invención se relaciona con una composición farmacéutica que comprende una cantidad efectiva de un polipéptido o complejo estabilizado de polipéptido, y un portador farmacéuticamente aceptable. Como se utiliza aquí, "una cantidad efectiva" significa una cantidad requerida para lograr un resultado final deseado. La cantidad requerida para lograr un resultado final deseado dependerá de la naturaleza de la enfermedad o trastorno que se va a tratar, y se puede determinar mediante técnicas clínicas estándar. Adicionalmente, se pueden emplear opcionalmente en ensayos in vitro para ayudar a identificar los rangos de dosificación óptimos. La dosis precisa que se va a emplear también dependerá de la ruta de administración y la gravedad de la enfermedad o trastorno, y se debe decidir de acuerdo con el juicio del médico y cada una de las circunstancias del sujeto. Las dosis efectivas se pueden extrapolar de las curvas de respuesta de dosis derivadas de sistemas de prueba de modelo animal o in vitro.

- 45 Se conocen diversos sistemas de suministro y se pueden utilizar para administrar una composición farmacéutica relacionada con la presente invención. Los métodos de introducción incluyen pero no se limitan a rutas intradérmicas, intramusculares, intraperitoneales, intravenosas, subcutáneas, intranasales, epidurales, y orales. Se

5 pueden administrar los compuestos mediante cualquier ruta convencional, por ejemplo, mediante infusión o inyección por bolo, mediante absorción a través de los revestimientos epiteliales o mucocutáneos (por ejemplo, mucosa oral, rectal y mucosa intestinal, etc.) y se pueden administrar con otros agentes biológicamente activos. La administración puede ser sistémica o local. Adicionalmente, puede ser deseable introducir las composiciones farmacéuticas que se relacionan con la invención en el sistema nervioso central mediante cualquier ruta adecuada, que incluye inyección intraventricular e intratecal; la inyección intraventricular se puede facilitar mediante un catéter intraventricular, por ejemplo, adherido a un reservorio, tal como un reservorio Ommaya. También se puede emplear la administración pulmonar, por ejemplo, mediante el uso de un inhalador o nebulizador, y formulación con un agente de formación de aerosol.

10 En un aspecto específico, puede ser deseable administrar las composiciones farmacéuticas que se relacionan con la invención localmente en el área que requiere tratamiento; esto se puede lograr, por ejemplo, y no en forma de limitación, mediante infusión local durante la cirugía, mediante inyección, por medio de un catéter, o por medio de un implante, siendo dicho implante de un material gelatinoso, poroso o no poroso, que incluye membranas, tales como membranas sialásticas, o fibras. En un aspecto, la administración puede ser inyección directa en el sitio (o sitio anterior) de un tumor maligno o tejido neoplásico o pre-neoplásico.

15 En otro aspecto, las composiciones farmacéuticas que se relacionan con la invención se pueden suministrar en un sistema de liberación controlada. En un aspecto, se puede utilizar una bomba (véase Langer, supra; Sefton, CRC Crit. Ref. Biomed. Eng.; vol. 14: pp. 201, 1987; Buchwald et al., Surgery; vol. 88: pp. 507, 1980; Saudek et al., N. Engl. J. Med.; vol. 321: pp. 574, 1989). En otro aspecto, se pueden utilizar materiales poliméricos (véase Medical Applications of Controlled Release, Langer y Wise (eds.), CRC Pres., Boca Raton, Florida, 1974; Controlled Drug Bioavailability, Drug Product Design and Performance, Smolen and Ball (eds.), Wiley, New York, 1984; Ranger y Peppas, J. Macromol. Sci. Rev. Macromol. Chem.; vol. 23: pp. 61, 1983; véase también Levy et al. Science; vol. 228: pp. 190, 1985; During et al. Ann. Neurol.; vol. 25: pp. 351, 1989; Howard et al. J. Neurosurg; vol. 71: pp. 105, 1989). En aun otro aspecto, se puede colocar un sistema de liberación controlada en la proximidad del objetivo terapéutico, es decir, el cerebro, requiriendo así solo una fracción de la dosis sistémica (véase, por ejemplo, Goodson, in Medical Applications of Controlled Release, supra, vol. 2, pp. 115-138, 1984).

Se describen otros sistemas de liberación en la revisión de Langer (Science; vol. 249: pp. 527-1533, 1990).

20 En un aspecto preferido, la composición se formula de acuerdo con procedimientos de rutina como una composición farmacéutica adaptada para administración intravenosa a seres humanos. Típicamente, las composiciones para la administración intravenosa son soluciones en un regulador acuoso isotónico estéril. Cuando sea necesario, la composición también puede incluir un agente de solubilización y un anestésico local tal como lidocaína para aliviar el dolor en el sitio de la inyección. Generalmente, los ingredientes se suministran ya sea de forma separada o mezclados juntos en una forma de dosificación unitaria, por ejemplo, como un polvo liofilizado seco o concentrado libre de agua en un contenedor cerrado herméticamente tal como una ampolla o bolsita que indica la cantidad del agente activo. Cuando la composición es para ser administrada mediante infusión, se puede suministrar con una botella de infusión que contiene solución salina o agua grado farmacéutico estéril. Cuando la composición se administra mediante inyección, se puede proporcionar una ampolla de agua estéril para la inyección o solución salina de tal manera que los ingredientes se pueden mezclar antes de la administración.

5.21. Consideración para composiciones farmacéuticas

40 Los polipéptidos o complejos de polipéptido estabilizados que se relacionan con la invención se pueden administrar en un portador que es farmacéuticamente aceptable. El término "farmacéuticamente aceptable" significa aprobado por una agencia reguladora del gobierno Federal o estatal o enumerado en la Farmacopea Estadounidense u otra farmacopea reconocida en general o que recibe aprobación individual o específica de una o más agencias reguladoras reconocidas en general para uso en animales y más particularmente en humanos. El término "portador" se refiere a un diluyente, adyuvante, excipiente, o vehículo con los cuales se administra el compuesto terapéutico. Tales portadores farmacéuticos pueden ser líquidos estériles, tales como agua, solventes orgánicos, tales como ciertos alcoholes, y aceites, que incluyen aquellos de origen de petróleo, animal, vegetal o sintético, tal como aceite de cacahuete, aceite de soja, aceite mineral, aceite de sésamo y similares. La solución salina regulada es un portador adecuado cuando se administra la composición farmacéutica por vía intravenosa. Las soluciones salinas y la dextrosa acuosa y las soluciones de glicerol también se pueden emplear como portadores líquidos, particularmente para soluciones inyectables. La composición, si se desea, también puede contener menos cantidades de agentes humectantes o emulsificantes, agentes reguladores de pH. Estas composiciones pueden tomar la forma de soluciones, suspensiones, emulsiones y similares. Ejemplos de portadores farmacéuticos adecuados se describen en "Remington's Pharmaceutical Sciences" por E.W. Martin. Tales composiciones contendrán una cantidad terapéuticamente efectiva del compuesto terapéutico, preferiblemente en forma purificada, junto con una cantidad adecuada de portador con el fin de proporcionar la forma para administración apropiada al paciente. La formulación debe adecuarse al modo de administración. En un aspecto preferido, la composición se formula de acuerdo con procedimientos de rutina como una composición farmacéutica adaptada para administración

intravenosa a seres humanos. Típicamente, las composiciones para administración intravenosa son soluciones en regulador acuoso estéril.

6. Ejemplos

- 5 Los siguientes ejemplos ilustran ciertas variaciones de los métodos de la invención para la estabilización de proteína o complejo de proteína. Estos ejemplos se presentan a manera de ilustración y no a manera de limitación del alcance de la invención.

6.1. Introducción

- 10 En los años recientes se han identificado varios polipéptidos y complejos de polipéptido con valor comercial significativo, y adicionalmente, se han identificado varios dominios modulares que median las interacciones proteína-proteína. Para muchos de estos dominios, también se pueden mapear los sitios de interacción con otras proteínas.

- 15 En la siguiente sección, se describen en detalle los métodos para estabilizar un tal complejo, un complejo de fragmento Fv, para el cual hay abundancia disponible de datos. Específicamente, se describen más adelante el montaje de las bases de datos importantes para el proceso de selección, el proceso de selección en sí mismo, la introducción de mutaciones puntuales, la expresión bacteriana de los polipéptidos y su purificación, el ajuste de las condiciones de reacción de entrecruzamiento, la reacción de entrecruzamiento en sí misma, y el análisis del complejo estabilizado resultante.

- 20 Los datos de entrada para la base de datos 2-D se obtienen de Weir's Handbook of Experimental Immunology I. Immunochemistry and Molecular Immunology, Fifth Edition. Los datos de entrada para la base de datos 3-D se obtienen de The Brookhaven National Laboratory Protein Database. Los datos derivados importantes para el proceso de selección en ambas bases de datos se calcula como se describió. El proceso de selección se lleva a cabo utilizando un conjunto de filtros que es conveniente y apropiado para esta aplicación de la presente invención.

- 25 Las mutaciones puntuales para la tirosina (que dirigen la reacción de entrecruzamiento) se introducen de acuerdo con la selección final del proceso de selección, y las mutaciones puntuales a la fenilalanina (que limitan la reacción de entrecruzamiento) de acuerdo con la secuencia específica de cada fragmento Fv y la información correspondiente e importante contenida en la base de datos 3-D. Los polipéptidos del complejo se expresan en términos bacterianos como las proteínas de fusión GST, y se purifican en una columna de afinidad GT. Los polipéptidos del complejo purificados se escinden proteolíticamente de las partes GST de las proteínas de fusión, y el polipéptido GST se elimina, de nuevo utilizando una columna de afinidad GT.

- 30 Las condiciones de reacción mínimamente requeridas se ajustan utilizando una construcción con las mutaciones para fenilalanina, pero carecen de las mutaciones para tirosina, y la reacción de entrecruzamiento luego se lleva a cabo con las construcciones que contienen ambos conjuntos de mutaciones puntuales. Se prueba la eficiencia de la reacción, y los fragmentos Fv estabilizados, resultantes se prueban luego en cuanto a las características de afinidad retenida, estabilidad, inmunogenicidad, y biodistribución.

- 35 6.2. Ventajas del entrecruzamiento tirosil- tirosil para los fragmentos Fv

- 40 La química subyacente de la tecnología cubierta por la presente invención origina un entrecruzamiento oxidativo para formación entre las cadenas laterales reactivas de las proteínas que forman complejos estables. Debido a que se cataliza la reacción de entrecruzamiento una vez establecida, el entrecruzamiento es estable en la ausencia del catalizador bajo un amplio rango de pH y condiciones de oxidorreducción. La reacción de entrecruzamiento requiere proximidad muy cercana entre las moléculas que se entrecruzarán y por lo tanto solo ocurre entre moléculas que interactúan normalmente y se asocian cercanamente en solución y por lo tanto se limitan a moléculas que tienen interacciones funcionales legítimas.

- 45 Así, la presente invención describe una nueva tecnología que permitirá la estabilización de los conjugados derivados de inunoglobulina y dará como resultado un muy alto grado de estabilidad y mínima inmunogenicidad en contextos terapéuticos. Esta tecnología se diseña para mejorar las tecnologías precedentes, y compatibles en complemento.

Los fragmentos Fv estabilizados resultantes tendrán las siguientes características:

1. Los conjugados serán estables bajo un amplio rango de pH y condiciones de oxidorreducción y a altas concentraciones de proteína.

2. El complejo entrecruzado resultante será mínimamente inmunogénico debido a que no se alteran los residuos expuestos .

Esta tecnología de estabilización de fragmento Fv es bien adecuada para el desarrollo de nuevos productos con aplicaciones novedosas, la mejora de los productos basados en inmunoglobulina existentes, y la complementación de tecnologías existentes para el desarrollo de aplicaciones de inmunoglobulina.

6.3. Aplicaciones del fragmento Fv

Existe un amplio espectro de aplicaciones potenciales de los productos basados en inmunoglobulina, los límites de los cuales se determinan por los siguientes factores:

El objetivo debe ser un ambiente que es accesible a los productos derivados de inmunoglobulina, tales como, por ejemplo, suero, la matriz extracelular, el cerebro, o el espacio intracelular por medio de liposomas (Hoffman R.M. J. Drug Target.; vol. 5(2): pp. 67-74, 1998) o absorción celular inducida por péptido (Schwarze S.R. et al. Science; vol. 285: pp. 21565-72, 1999). Para aplicaciones intracelulares de inmunoglobulina, véase Bosilevac J.M. et al. J. Biol. Chem.; vol. 273(27): pp. 16874-79, 1998; Graus-Porta D. et al. Mol. Cell Biol.; vol 15: pp. 1182-91, 1995; Richardson J.H. et al. Proc. Nat. Acad. Sci., USA; vol. 92: pp. 3137-41, 1995; Maciejewski J.P. et al. Nat. Med.; vol. 1: pp. 667-73, 1995; Marasco W.A. et al. Proc. Nat. Acad. Sci., USA; vol. 90: pp. 7889-93, 1993; Levy Mintz P. et al. J. Virol.; vol. 70: pp. 8821-32, 1996; Duan L. et al. Hum. Gene Ther.; vol. 6(12): pp. 1561-73, 1995; y Kim S.H. et al. Mol. Immunol.; vol. 34(12-13): pp. 891-906, 1997. Un ambiente favorable está presente en todos los tejidos y órganos que son alcanzados por el suministro sanguíneo, y cuando está presente la molécula objetivo sobre la superficie celular o en la matriz extracelular. Ya que la funcionalidad de los fragmentos Fv derivados de inmunoglobulina es principalmente unirse a moléculas objetivo, la unión al objetivo debe ser suficiente para lograr el efecto terapéutico y diagnóstico deseado. La funcionalidad catalítica es, sin embargo, también conocida para la inmunoglobulina, y por lo tanto también se puede lograr en los contextos farmacológicos y/o industriales (Pluckthun A. et al. Ciba Found. Symp.; vol. 159: pp. 103-12; discusión 112-7, 1991; Kim S.H. et al. Mol. Immunol, vol. 34: pp. 891-906, 1997).

Hay una multitud de aplicaciones potenciales basadas en inmunoglobulina que cumplen estos criterios, y es el propósito de los siguientes párrafos solo señalar ciertas aplicaciones importantes, como ejemplo.

6.3.1. Suministro de fármaco/selección de tejido objetivo

Muchas aplicaciones existentes de terapia de inmunoglobulina hacen uso de la capacidad de los anticuerpos para dirigir los agentes terapéuticos a los tejidos objetivados. Tales agentes terapéuticos han sido hasta ahora toxinas y radioisótopos dirigidos a tumores mediante la vinculación a antígeno asociado a anti-tumor o anticuerpos específicos a anti-tumor, por un lado, y los agentes diagnósticos, es decir anticuerpos enlazados a un agente de formación de imagen, por otro lado.

6.3.2. Modulación de procesos bioquímicos extracelulares

Existe una multitud de procesos bioquímicos que son de importancia terapéutica así como comercial que ocurren en ambientes extracelulares, tales como suero sanguíneo. Un ejemplo de tales procesos es el proceso de coagulación sanguínea. En este ejemplo, la inmunoglobulina se une a las proteínas involucradas en la cascada bioquímica de las reacciones que llevan a la formación de coágulos de sangre, e interrumpe esta cascada, bloqueando de este modo la formación de coágulos de sangre. El valor terapéutico de ser capaz de inhibir la formación de coágulos de sangre, de hecho, impulsó el desarrollo de uno de los primeros productos farmacéuticos basados en inmunoglobulina para que entrara en el mercado.

6.4. Selección de residuos óptimos para entrecruzamiento tirosilo-tirosilo

El proceso de selección consiste de una serie de pruebas estadísticas o 'filtros' dirigidos a estrechar sucesivamente los pares de residuo que más probablemente den como resultado un par de tirosina de cadena ligera-cadena pesada entrecruzada que altera mínimamente el Fv.

6.4.1. Datos utilizados para el análisis

Los datos de uso de aminoácidos de residuo son datos compilados en los aminoácidos codificados y expresados en cada residuo de los fragmentos Fv conocidos y secuenciados. Están recogidos en, y se obtienen de la publicación "Proteins of Immunological Interest", Kabat y Wu, Government Printing Office, NIH Publicación 91-3242, 1991 ("K&W"). Las secuencias de aminoácidos es esta publicación se ordenan de acuerdo con un sistema de numeración estandarizado que tiene en cuenta la estructura de gen de las regiones variables de cadena pesada y ligera. En las regiones variables de las cadenas pesada y ligera y similares, cuatro segmentos de Región de Estructura (FRs) –

que se conservan relativamente – se intercalan por tres – Regiones Determinantes de Complementariedad-altamente variables. Las CDR contienen los aminoácidos que determinan la especificidad del anticuerpo, y que ponen en contacto físicamente el antígeno. Alinear todas las secuencias de acuerdo con el sistema de numeración K&W es muy importante para el propósito de realizar un análisis estadístico como se describe en este ejemplo ya que los residuos correspondientes de los FRs por lo tanto siempre están alineados, independientemente de las longitudes de secuencia variantes de los CDR intercalados. Esto asegura que las mediciones estadísticas se hacen con conjuntos de datos que contienen puntos de datos apropiados y comparables. Los datos de coordenadas para los cálculos de distancia de todos los átomos diferentes de hidrógeno de 17 fragmentos Fv de estructuras de inmunoglobulina resueltas cristalográficamente se han descargado de la base de datos de estructuras de proteínas Brookhaven National Laboratory (www.bnl.pdb.gov; Figura 5). Estos datos proporcionan las coordenadas tridimensionales (x, y, y z) para cada átomo en una estructura resuelta, expresada en unidades métricas, es decir Angströms (10⁻¹⁰m, Å). Con estos datos es posible calcular las distancias tridimensionales entre cualesquier átomos deseados (por ejemplo átomos de carbono amino alfa y beta) y calcular las mediciones estadísticas de la variabilidad de tal distancia entre los diferentes fragmentos Fv en la muestra que se analiza (Figuras 5, 6, y 7).

6.4.2. Metodología de selección

Los residuos óptimos, a los cuales se dirige la reacción de entrecruzamiento, se seleccionan mediante una serie de filtros con base en las mediciones estadísticas de los valores en las bases de datos compiladas para los propósitos de esta selección. Estas bases de datos contienen mediciones numéricas de (i) separación de carbono alfa, (2) separación de carbono beta y la diferencia entre las distancias alfa y beta, y (3) uso de aminoácidos de residuo (véase adelante).

6.5. Filtro 1: Eliminación de pares de residuo con glicinas

La glicina es el más pequeño de los aminoácidos y no tiene carbono beta y a menudo se asocia con la flexibilidad posicional de las estructuras de proteína. La sustitución de una glicina con uno de los aminoácidos más grandes, tirosina, puede tener probablemente muy grande impacto sobre la estructura completa del complejo de proteína, y por lo tanto sobre las características de unión a antígeno del fragmento Fv entrecruzado. Por lo tanto, como un primer corte, de entre todos los pares de residuo candidatos de las Regiones de Estructura, aquellos pares, de los cuales uno de los residuos es más frecuentemente una glicina (como se determina mediante comparación con los datos K&W) se eliminan a priori. Para los propósitos de este análisis ocurrencia 'más frecuente' de un aminoácido particular en un residuo dado se define como ocurrencia en más del 75% de la muestra.

Tabla 1. Pares candidatos de cadena pesada- cadena ligera con mediciones de distancia promedio de carbono alfa mx, dentro del rango de 5.70Å a 11.74Å (clasificados por la numeración K&W, primero sobre las posiciones de la cadena ligera, segundo sobre las pociones de la cadena pesada).

Ligera	Pesada	PROMEDIO	STDEV	Ligera	Pesada	PROMEDIO	STDEV
36	45	10.38	0.23	44	91	9.33	0.33
36	103	10.99	0.31	44	92	10.91	0.40
37	45	11.49	0.36	44	93	9.74	0.29
38	39	11.49	0.18	44	103	6.92	0.30
38	45	10.17	0.43	44	105	8.95	0.55
38	103	11.26	0.41	45	93	10.43	0.41
40	41	11.27	1.50	45	103	7.40	0.41
40	43	11.68	1.34	45	105	10.95	0.45
42	39	11.04	0.84	46	93	10.78	0.40
42	89	10.28	0.99	46	94	11.19	0.25

Ligera	Pesada	PROMEDIO	STDEV	Ligera	Pesada	PROMEDIO	STDEV
42	90	11.72	0.88	46	103	8.98	0.33
42	91	10.5	0.66	85	43	11.04	0.49
42	103	10.13	0.34	85	45	10.93	0.37
42	105	7.14	0.40	86	45	10.63	0.35
42	107	11.18	0.82	87	43	11.64	0.32
43	4	11.50	0.56	87	45	8.19	0.25
43	37	10.94	0.87	87	46	10.90	0.33
43	38	10.97	0.98	88	45	10.04	0.10
43	39	10.34	0.79	88	46	11.69	0.21
43	45	10.78	0.71	98	37	10.24	0.31
43	89	9.95	0.71	98	38	11.25	0.25
43	90	10.23	0.72	98	39	11.17	0.20
43	91	8.04	0.71	98	43	11.60	0.39
43	92	10.21	0.59	98	45	6.49	0.18
43	93	10.14	0.65	98	46	6.66	0.29
43	103	6.74	0.51	98	48	7.65	0.57
43	105	5.74	0.44	98	49	11.37	0.58
43	107	10.66	0.62	100	39	11.42	0.29
44	37	10.58	0.39	100	43	8.27	0.41
44	38	11.31	0.50	100	45	7.82	0.27
44	39	10.73	0.48	100	46	9.56	0.46
44	45	9.43	0.48	102	43	11.47	0.36

6.6. Filtro 2: identificación de pares de residuo separados apropiadamente

5 Para encontrar pares de residuo apropiadamente separados para un enlace tirosilo-tirosilo, se calculan las distancias de el carbono alfa a carbono alfa de cada residuo en la cadena ligera para cada residuo en la cadena pesada en los fragmentos Fv representados en The Brookhaven National Protein Structure Database en una base de datos 3D. Este cálculo se realiza al aplicar la geometría de Pitágoras a las coordenadas 3D de los carbonos alfa (Figura 6). Para cada combinación de residuos de cadena pesada y ligera, el promedio, desviación estándar, rango y mediana de la distancia carbono alfa - carbono alfa se calcula sobre los fragmentos Fv en la muestra (Figura 7). Con base en los cálculos anteriores, como un segundo corte, todos los residuos de par seleccionados de aquellos carbonos alfa se separan en un promedio, m, dentro del rango de selección. El rango que se selecciona es el siguiente: Min 5.70Å, Max 11.74 Å.

10

La distancia óptima (T) se calcula al promediar el máximo y el mínimo del rango. Por lo tanto,

$$T = (5.70\text{Å} + 11.74\text{Å}) / 2 = 8.72 \text{ Å}.$$

En este ejemplo, 64 pares de residuo cumplen este criterio, enumerados en la Tabla 1.

6.7. Filtro 3: identificación de pares de residuo con flexibilidad posicional suficiente

- 5 Con el fin de identificar pares de residuo en los que la sustitución para la tirosina se interrumpe mínimamente, se seleccionan pares de residuos con flexibilidad posicional significativa. Por lo tanto, se eliminan los pares de residuo de entre aquellos en la Tabla 1 en los que la distancia óptima, 8.72Å, no entra dentro de 2 tiempos de aquella desviación estándar del par de residuos específica de su promedio. En este ejemplo, 36 pares de residuo cumplen este criterio.
- 10 Adicionalmente, la flexibilidad posicional relativa de los 12 pares de residuo candidatos restantes se clasifican de acuerdo con la siguiente fórmula:

Clasificación $I = a_x^2 / \sigma_x$.

$a = T - \mu_x + 2\sigma_x$, para todos los $\mu_x \geq T$

$a_x = \mu_x + 2\sigma_x - T$, para todos los $\mu_x < T$

- 15 $x T$ = distancia óptima

μ_x = la distancia promedio para cualquier par de residuos dado

σ_x = desviación estándar de la distancia de cualquier par de residuos dado

- 20 Así, los residuos que se clasifican altamente bajo esta métrica son aquellos que (i) tienen una separación promedio cercana a la distancia óptima, y/o (ii) tienen una desviación estándar grande. Los 12 pares de residuo estantes se enumeran, clasificados por la clasificación I en la Tabla 2.

Tabla 2. Pares de residuo de la Tabla 1 seleccionados y clasificados de la Clasificación I².

Pesada	Ligera	Clasificación I	AVG	STDEV
44	105	1.35	8.95	0.55
43	91	0.76	8.04	0.71
46	103	0.49	8.98	0.33
100	43	0.33	8.27	0.41
43	37	0.26	10.9	0.87
42	89	0.17	10.3	0.99
40	41	0.14	11.3	1.50
44	45	0.13	9.43	0.48
43	89	0.06	9.95	0.71
100	46	0.01	9.56	0.46

98	48	0.01	7.56	0.57
44	91	0.01	9.33	0.33

¹ Criterio de selección: la distancia óptima (T) debe caer dentro del rango de la distancia promedio específica del par de residuos (μ_x) +/- 2 veces la desviación estándar específica del par de residuos (σ_x).

² Fórmula de clasificación I: a_x^2 / σ_x^2

donde T es la distancia óptima, y $a_x = T - \mu_x + 2\sigma_x$, para todos los $\mu_x \geq T$, y $a_x = \mu_x + 2\sigma_x - T$, para todos los $\mu_x \leq T$.

6.8. Filtro 4: orientación de la cadena lateral

5 En el espacio que ocupan las cadenas pesada y ligera, las cadenas laterales de tirosina se deben orientar una hacia la otra para un entrecruzamiento para formarse con mínima distorsión estructural. La diferencia entre la distancia del carbono alfa (es decir la distancia del carbono de la estructura principal; Figura 6) y la distancia del carbono beta (es decir la distancia entre los primeros carbonos en cada cadena lateral; Figura 8) de cada par de residuos se calcula como un proxy, es decir un estimado de la orientación de las cadenas relativas una a la otra (Figura 9).

El rango que se selecciona es el siguiente:

Min -0.5Å, Max 2.0Å.

10 La diferencia de distancia óptima (D) se calcula al promediar el máximo y el mínimo del rango. Por lo tanto,

$D = (-0.5\text{Å} + 2.0\text{Å}) / 2 = 0.75\text{Å}.$

15 De nuevo, con base en la geometría de coordenadas 3D, para cada par de residuos, se calcula la distancia entre los carbonos beta (Figura 8). La distancia beta se sustrae luego de la distancia alfa del par de residuos (Figura 9). Este filtro es con base en si la diferencia promedio en las distancias alfa y beta de un par de residuos (Figures 10 y 11) cae dentro del rango óptimo estimado. En este ejemplo, 12 pares de residuo cumplen este criterio, enumerado en la Tabla 3.

Tabla 3. Pares de residuo de la Tabla 2 seleccionados por la diferencia de distancia alfa- beta promedio

Pesada	Ligera	Clasificación	AVG	STDEV	AVG	STDEV
91	43	0.76	8.04	0.71	1.33	0.70
45	43	0.56	10.78	0.71	-0.04	0.31
103	46	0.49	8.98	0.33	0.81	0.18
39	42	0.48	11.04	0.84	0.21	0.14
91	42	0.30	10.5	0.66	-0.14	0.17
37	43	0.26	10.94	0.87	0.81	0.59
89	42	0.17	10.28	0.99	0.01	0.06

92	43	0.15	10.21	0.59	-0.23	0.61
89	43	0.06	9.95	0.71	0.71	0.36
93	43	0.02	10.14	0.65	1.07	0.73
48	98	0.01	7.65	0.57	0.87	0.17
30	43	0.00	10.34	0.79	0.41	0.28

Adicionalmente, de forma análoga a la selección con base en las distancias de carbono alfa, se eliminan aquellos pares para los cuales la diferencia de la distancia promedio óptima, 0.75Å, no cae dentro de 2 veces aquella desviación estándar específica del par de residuos de su promedio.

5 Clasificación = a_x^2 / σ_x

$a_x = D - \mu_x + 2\sigma_x$, para todos los $\mu_x \geq D$

$a = \mu_x + 2\sigma - D$, para todos los $\mu_x \geq D$

D= diferencia de distancias óptimas

μ_x = la diferencia de distancia promedio de cualquier par de residuos dado

10 σ_x = desviación estándar de la diferencia de distancia para cualquier par de residuos dado

Del conjunto de pares de residuo potenciales enumerados en la Tabla 4, cinco pares cumplen estos criterios. Este conjunto de pares de residuo potenciales se enumera en la Tabla 5.

Tabla 4 Pares de residuo de la Tabla 5 seleccionados¹ y clasificados de acuerdo con la Clasificación II²

Diferencia entre las distancias de C alfa y C beta					distancia de Carbono alfa		
Pesada	Ligera	Clasificación II	Promedio	Stdev	Clasificación I	Promedio	Stdev
92	43	0.10	-0.23	0.61	0.15	10.21	0.59
39	43	0.17	0.41	0.28	0.00	10.34	0.79
48	98	0.30	0.87	0.17	0.01	7.65	0.57
103	46	0.49	0.81	0.18	0.49	8.98	0.33
91	43	0.96	1.33	0.70	0.76	8.04	0.71
89	43	1.27	0.71	0.36	0.06	9.95	0.71
93	43	1.79	1.07	0.73	0.02	10.14	0.65
37	43	2.10	0.81	0.59	0.26	10.94	0.87

¹ Criterio de selección: Diferencia óptima en distancias alfa y beta (D) debe caer dentro del rango de la diferencia de distancia alfa- beta promedio del par de residuos ($\bar{\delta}_x$) 2 x la desviación estándar específica del par de residuos (σ_x).

² Fórmula de Clasificación II: = a_x^2 / σ_x , por lo cual D es la diferencia de distancia óptima, y $a_x = D - \bar{\delta}_x + 2\sigma_x$, para todas

las $\bar{d}_x \geq D$, y $a_x = \bar{d}_x + 2\sigma_x - \Delta$, para todas las $\bar{d}_x \leq D$.

5 Nótese que la diferencia entre la distancia alfa-alfa óptima y la distancia alfa- beta (Objetivo) también cae cómodamente dentro del rango de valores medidos realmente de la mayoría de los pares de residuo seleccionados, como se muestra en la Tabla 5. Esto es importante, ya que resalta adicionalmente la probabilidad de que los pares candidatos seleccionados darán como resultado cadenas laterales de tirosina entrecruzadas que interrumpen mínimamente la estructura y función del fragmento Fv.

Tabla 5. Promedio, mediana, desviación estándar, y rango de las diferencias de las distancias alfa- alfa medidas realmente y distancias alfa-beta. Los pares de residuo restantes se identifican en las dos primeras filas por sus número de residuo K&W de cadena pesada y ligera.

	Pesada	37	39	89	91	92	93	103	48
	Ligera	43	43	43	43	43	43	46	98
	Promedio	10.94	10.34	9.95	8.04	10.21	10.14	8.98	7.65
Distancia de carbono alfa	Stdev	0.87	0.79	0.71	0.71	0.59	0.65	0.33	0.57
	Max	13.23	12.37	11.75	9.82	11.81	11.81	9.63	8.68
	Min	9.94	9.63	9.05	7.32	9.56	9.42	8.39	6.78
	Mediana	10.81	10.10	9.80	7.92	9.99	9.95	8.95	7.89
	Promedio	0.81	0.41	0.71	1.33	-0.23	1.07	0.81	0.87
Diferencia de Ca-Cb	Stdev	0.59	0.28	0.36	0.70	0.61	0.73	0.18	0.17
	Max	1.42	0.84	1.17	2.02	0.33	1.74	1.09	1.37
	Min	-0.64	-0.10	-0.08	-0.25	-1.86	-0.69	0.40	0.63
	Mediana	1.03	0.45	0.75	1.65	0.05	1.29	0.77	0.81

10

6.9. Filtro 5: uso de cadena lateral de aminoácido

15 Debido a que los pares de residuo van a ser sustituidos con tirosina de tal manera que las sustituciones perturban mínimamente la estructura y función del complejo entrecruzado resultante, se seleccionan pares de residuo de entre los en las Tablas 4 y 5 de tal manera que las propiedades de las cadenas laterales de aminoácido originales son posiblemente las más similares a la tirosina. Las propiedades de la cadena principal que se miden son (i) volumen de van der Waals y (ii) hidrofobicidad. Estas mediciones se utilizan como proxies para el tamaño y carga de las cadenas laterales de aminoácidos, respectivamente.

20 En cada residuo, cada cadena lateral de aminoácido que se presenta se da en un valor numérico que representa su volumen de van der Waals y su hidrofobicidad (Figura 12). Con base en los datos de uso de aminoácido para estos residuos (Kabat & Wu), se calculan el promedio y la desviación estándar del volumen de van der Waals y la hidrofobicidad del residuo, ponderados y no ponderados por la frecuencia en la que se presenta la cadena lateral específica en este residuo. Se calcula una medición estadística ponderada en cada valor presente en la muestra (n = número de secuencias en la base de datos 2D), y se calcula una medición estadística no ponderada sobre el valor de cada aminoácido de ocurrencia (n = 20 máximamente) (Figura 13).

25 Por ejemplo, dadas 10 secuencias en una base de datos, por lo cual en un residuo dado se presenta la alanina 8 veces, y la leucina dos veces, el promedio ponderado de los volúmenes de van der Waals sería:

$$(8 \times \text{valor ala} + 2 \times \text{valor leu})/2$$

$$=(8 \times 67 + 2 \times 124)/10 = 78.4$$

En el mismo ejemplo, el promedio no ponderado sería

$$(\text{valor ala} + \text{valor leu})/2$$

$$5 = (67 + 124)/2 = 95.5$$

Los valores numéricos para todos los 20 aminoácidos de volumen de van der Waals y la hidrofobicidad utilizados para la selección se enumeran en la Tabla 6.

10 Cada uno de los 6 pares de residuo identificados en el análisis estructural se examina en cuanto a su capacidad de ser sustituido “de forma conservadora” con dos residuos de tirosina, al comparar los valores de van der Waals e hidrofobicidad promedio del par y sus desviaciones estándar con los de un par de tirosina.

Tabla 6. Valores numéricos de volúmenes de van der Waals (Richards, F.M._ J. Mol. Biol. 82, 1-14, 1974) y hidrofobicidad (Eisenberg, D._ Ann. Rev. Biochem. 53, 595-623, 1984) de cadena lateral de aminoácido.

Aminoácido	Volúmenes de Van der Walls [Å ³]	Hidrofobicidad
Ala	67	0.62
Arg	148	-2.50
Asn	96	-0.78
Asp	91	-0.90
Cys	86	0.29
Gln	114	-0.85
Glu	109	-0.79
Gly	48	0.48
His	118	-0.40
Ile	124	1.40
Leu	124	1.10
Lys	135	-1.50
Met	124	0.64
Phe	135	1.20
Pro	90	0.12
Ser	73	-0.18
Thr	93	-0.05
Trp	163	0.81
Tyr	141	0.26

Val	105	1.10
-----	-----	------

Para cada uno de los residuos enumerados en la Tabla 5, los valores de volúmenes de van der Waals e hidrofobicidad promedio y sus desviaciones estándar, ponderadas y no ponderadas, se enumeran en la Tabla 7 y 8, respectivamente.

5 Tabla 7. Valores de Van der Waals para los pares de residuo y comparación con un par tyr-tyr.

Pesada		37	39	89	91	92	93	103	48
Consensus		VAL	GLN	VAL	TYR	CYS	ALA	TRP	VA
Pesada		37	39	89	91	92	93	103	48
no ponderado	Promedio	109	113	110	141	86	69	160	110
	Stdev	8	12	12	1	-	9	11	9
	Promedio	116	103	122	138	86	78	136	116
	Stdev	10	51	18	4	-	26	27	10
	Ligera	43	43	43	43	43	43	46	98
ponderado	Consensus	ALA	ALA	ALA	ALA	ALA	ALA	LEU	PHE
	Promedio	72	72	72	72	72	72	124	135
	Stdev	14	14	14	14	14	14	3	2
no ponderado	Promedio	94	94	94	94	94	94	118	128
	Stdev	24	24	24	24	24	24	11	6
	Pesada	37	39	89	91	92	93	103	48
	Ligera	43	43	43	43	43	43	46	98
ponderado	2 x valor tyr	282	282	282	282	282	282	282	282
	Valor comb. ¹	181	185	182	213	158	141	283	245
	Diferencia ²	101	97	100	69	124	141	1	38
	Stdev Comb. ³	22	26	26	15	14	23	14	11
	Clasificación III ⁴	0.21	0.27	0.26	0.21	0.11	0.16	10.39	0.28
no ponderado	2 x valor tyr	282	282	282	282	282	282	282	282
	Valor comb. ¹	210	197	216	232	180	172	253	244
	Diferencia ²	72	85	66	50	102	110	29	39
	Stdev Comb. ³	35	75	43	29	24	50	38	17

Clasificación IV ⁴	0.49	0.89	0.64	0.57	0.24	0.46	1.32	0.43
-------------------------------	------	------	------	------	------	------	------	------

¹ Suma de los valores de van der Waals promedio del par de residuo

² Tamaño de la diferencia (raíz cuadrada de la diferencia cuadrada) entre la suma del valor para dos residuos de tirosina (282) y la suma de los valores promedio de pares de residuo (1)

³ Suma de la desviación estándar de ambos residuos

⁴ Fórmula utilizada: $\text{Stdev}/\text{Diferencia}^{(3/2)}$

Tabla 8. Valores de hidrofobicidad para pares de residuo y comparación con un par tyr-tyr.

Pesada		37	39	89	91	92	93	103	48
Ponderado	Consensus	VAL	GLN	- VAL	TYR	CYS 0.29	ALA	TRP	VAL
	Promedio Stdev	1.14 0.14	0.86 0.35	- 0.90 0.66	0.30 0.20	-	0.58 0.19	0.79 0.30	1.14 0.11
No ponderado	Promedio Stdev	1.07 0.27	0.96 1.49	0.41 1.37	0.73 0.66	0.29	0.54 0.47	0.41 1.05	1.25 0.17
						-			
Ligera		43	43	43	43	43	43	46	98
Ponderado	Consensus	ALA	ALA	ALA	ALA	ALA	ALA	LEU	PHE
	Promedio Stdev	0.50 0.33	0.50 0.33	0.50 0.33	0.50 0.33	0.50 0.33	0.50 0.33	1.08 0.09	1.20 0.03
No ponderado	Promedio Stdev	0.47 0.59	0.47 0.59	0.47 0.59	0.47 0.59	0.47 0.59	0.47 0.59	0.95 0.27	1.23 0.15

Pesada		37	39	89	91	92	93	103	48
Ligera		43	43	43	43	43	43	46	98
Ponderado	2 x valor tyr	0.52	0.52	0.52	0.52	0.52	0.52	0.52	2.34
	Valor comb.1	1.64	-0.36	1.40	0.80	0.79	1.08	1.87	1.82
	Diferencia 2	1.12	0.88	0.88	0.28	0.27	0.56	1.35	0.13
	Stdev Comb..3	0.46	0.69	1.00	0.53	0.33	0.53	0.38	0.07
	Clasificación V4	0.42	0.78	1.13	1.89	1.24	0.97	0.28	0.06
No ponderado	2 x valor tyr	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52
	Valor comb.1	1.54	-0.49	0.88	1.20	0.76	1.01	1.35	2.48
	Diferencia 2	1.02	1.01	0.36	0.68	0.24	0.49	0.83	1.96
	Stdev Comb..3	0.87	2.09	1.97	1.26	0.59	1.07	1.32	0.33

Clasificación IV4 | 0.85 2.07 5.44 1.86 2.49 2.20 1.58 0.17

¹ Suma del par de valores de hidrofobicidad promedio del residuo

² Tamaño de la diferencia (raíz cuadrada de la diferencia cuadrada) entre la suma del valor para dos residuos de tirosina (0.52) y a suma de los valores promedio de los pares de residuos (¹) ³ Suma de la desviación estándar de ambos residuos

⁴ Fórmula utilizada : $Stdev/Diferencia \left(\frac{3}{1} \right)^2$

6.10. Filtro 6: eliminación parcial de pares con residuos altamente conservados

5 Todos los residuos bajo consideración están dentro de las Regiones de Estructura de ya sea la cadena pesada o la ligera de los fragmentos Fv, y se puede esperar por lo tanto que se conserven. Por lo tanto, para el propósito de este análisis, los residuos que se conservan más del 80% (véase Tabla 9) se eliminan, con la excepción de pares en los que se conserva un residuo aromático (véase adelante).

Tabla 9. Conservación de identidad de residuos de aminoácidos

	Consensus ¹	Ocurrencia de consensus ²	Tamaño de muestra, N ³	No. de ocurrencia AAs ⁴	Conservación de identidad de AA ⁵
Cadena Pesada					
37	VAL	31	40	4	78%
39	GLN	35	37	3	95%
48	VAL	30	42	4	71%
89	VAL	25	40	7	63%
91	TYR	42	44	2	95%
92	CYS	44	44	1	100%
93	ALA	37	42	4	88%
103	TRP	30	33	3	91%
Cadena Ligera					
43	ALA	49	65	6	75%
46	LEU	54	57	3	95%
98	PHE	66	68	3	97%

¹ Más frecuentemente ocurrencia de aminoácido del residuo indicado

² Número de las ocurrencias de aminoácido consensus (1) en el residuo indicado

³ Número de aminoácidos conocidos para un fragmento Fv en el residuo indicado

⁴ Número de aminoácidos diferentes (AAs) que se producen en el residuo indicado

5 Ocurrencia del aminoácido consensus (2) dividido por el tamaño de muestra, $N(3)$.

De los residuos del par de residuos de las tablas 4, 5, 6, 8, y 9, cuatro pares no contienen un aminoácido aromático conservado que conserve más del 80%, y por lo tanto se eliminan.

5 Los pares de residuo restantes, que se predicen por ser las posiciones óptimas para el entrecruzamiento, se enumeran en la Tabla 10 con todas las clasificaciones descritas anteriormente.

Tabla 10. Pares de residuo potenciales seleccionados para el entrecruzamiento tyr- tyr al cual se dirigen

Pares de residuo (H/L)	Clasificación I	Clasificación II	Clasificación III/IV	Clasificación V/VI
103/46	0.49	0.49	10.39/1.32	0.28/1.58
89/43	0.06	1.27	0.26/0.64	1.13/5.44
37/43	0.26	2.10	0.21/0.49	0.42/0.85
48/98	0.01	0.30	0.28/0.43	0.06/0.17

6.11. Diagrama de flujo de selección de par de residuos para software

Montaje de base de datos

10 Material de partida

Importación de base de datos 2D y Clasificación de los datos

Datos de Secuencia

- Importación de datos de secuencia de polipéptido 2D

Definir:

15 s = tamaño de muestra (número) de secuencias de las cadenas de polipéptido individual del complejo de proteína (preferiblemente en pares de polipéptido de un complejo)

- Alineación de los datos de acuerdo con conservación funcional (por ejemplo sistema de numeración Kabat & Wu para Ig)

Definir:

20 i (subíndice) = posición del aminoácido dentro del sistema de alineación a la cual cualquier átomo dado pertenece

- Compilación de identidad (código de tres letras) y frecuencia de ocurrencia de aminoácidos en cada residuo

Definir:

f_i = frecuencia de la ocurrencia de un aminoácido particular en un residuo dado,

25 n_i = número de aminoácidos que ocurren en un residuo dado,

- Definir y marcar residuos de ambos polipéptidos dentro de las regiones conservadas de ambos polipéptidos (Regiones de Estructura para los fragmentos Fv)

Asignar:

con = residuos conservados

non = residuos variables

Asignación de consensus

5 Definir:

El consensus es el aminoácido que se presenta con más frecuencia en cualquier residuo dado de cualquier polipéptido.

Asignar:

Para cada residuo, i ,

10 Asignar el consensus utilizando, por ejemplo, código de aminoácidos de una letra. Para los residuos en los cuales se producen dos o más aminoácidos más frecuentemente, se asignan todos los aminoácidos que se presentan más frecuentemente.

Datos sobre las propiedades físicas de las cadenas laterales de aminoácido

15 • Compilación de tablas de consulta con aminoácidos y valores numéricos correspondientes. Los valores numéricos corresponden a las propiedades físicas más importantes de las cadenas laterales de aminoácido ya que estas influyen la estructura completa de los complejos de polipéptido (por ejemplo volumen de cadena lateral, carga, hidrofobicidad, y grados de libertad de rotación, etc.)

Definir:

p (subíndice): propiedad física de cadena lateral de aminoácido escogida para el proceso de selección

20 N_{pi} = valor numérico de una propiedad física que corresponde a un aminoácido de ocurrencia en un residuo dado, i

Importación de base de datos 3D y Clasificación de los datos

15

Clasificación mediante Secuencia (2-D)

• Importación de datos ordenados 3D de los polipéptidos (de la estructura del complejo como un todo)

25 Definir:

m (subíndice) = tamaño de muestra (número) de diferentes archivos de estructuras impostados (para ambos polipéptidos de un complejo)

• Alineación de los datos de acuerdo con la conservación funcional (por ejemplo sistema de numeración Kabat & Wu para Ig)

30 Clasificación mediante posición atómica 3-D

• Clasificación de datos de coordenadas por residuo de aminoácido y posición de aminoácido

Seleccionar carbonos alfa y beta

Definir:

$Ca1_i$ = carbono alfa que pertenece al primero de dos polipéptidos

35 $Ca2_i$ = carbono alfa que pertenece al segundo de dos polipéptidos

Cb1_i = carbono beta que pertenece al primero de dos polipéptidos

Cβ2_i = carbono beta que pertenece al segundo de dos polipéptidos

Coordenadas de Ca1_i: x_{Ali}, y_{Ali}, z_{Ali}

Coordenadas de Ca2_i: x_{A2i}, y_{A2i}, z_{A2i}

5 Coordenadas de CB1_i: x_{Bli}, y_{Bli}, z_{Bli}

Coordenadas de CB2_i: x_{B2i}, y_{B2i}, z_{B2i}

Montaje de los pares de residuos

- Montaje de todos los posibles pares inter- cadena de residuos

Definir

10 j (subíndice) = par de aminoácidos cuando caen dentro del sistema de alineación anterior de ambas cadenas de polipéptido

Compilación de Mediciones Importantes; Datos Secundarios, Derivados

Datos derivados 2-D

Cómputo de las características del residuo para cada propiedad física

15 • Recuperación de valores numéricos de cada propiedad física de la cadena lateral para cada aminoácido que se presenta en cada residuo. Empareja cada identidad de aminoácido en cada residuo en la tabla de consulta, y recupera valores numéricos correspondientes

- Cálculo de mediciones estadísticas ponderadas para cada residuo

Definir:

20 μ_{pi} = promedio ponderado de la muestra, s, de valores numéricos de una propiedad física en cada residuo, i, ponderado por cada aminoácido de ocurrencia s frecuencia de ocurrencia, f_i

σ_{pi} = desviación estándar ponderada de la muestra, s, de valores numéricos de una propiedad física en cualquier residuo, i, ponderado por cada aminoácido de ocurrencia s frecuencia de ocurrencia, f_i

Calcular:

25 para la muestra de secuencias en la base de datos, s, para cada residuo, h, y para cada propiedad física, p

$$w\mu_{pi} = \sum (N_{pi} * f_{pi}) / \sum f_{pi}$$

$$w\sigma_{pi} = \text{SQRT}((\sum_{pi} * \sum (f_{pi} * N_{pi}^2) - \sum (f_{pi} * N_{pi})^2) / \sum f_{pi} * \sum (f_{pi} - 1))$$

- Cálculo de mediciones estadísticas no ponderadas para cada residuo

Definir:

30 μ_{pi} = promedio no ponderado de la muestra, s, de valores numéricos de una propiedad física en cualquier residuo, i, no ponderado por cada frecuencia de ocurrencia de los aminoácidos de ocurrencia, f_i

$u\sigma_{pi}$ = desviación estándar no ponderada de la muestra, s, de los valores numéricos de una propiedad física en cualquier residuo, i, no ponderado por cada frecuencia de ocurrencia de los aminoácidos de ocurrencia, f_i

Calcular:

para la muestra de secuencias en la base de datos, s, para cada residuo, i, y para cada propiedad física, p:

$$u\mu_{pi} = (\sum n_{pi}) / n_i$$

5

$$u\sigma_{pi} = \text{SQRT} \left((n_i * \sum n_{pi}^2 - \sum (n_i * N_{pi})^2) / n_i * (n_i - 1) \right)$$

- Cálculo de cada promedio combinado de par y desviación estándar

Para ambos residuos de cada par la suma de los valores promedio y desviación estándar se calculan para cada propiedad física.

Calcular:

10 Para cada par de residuo, j:

$$w\mu_{pj} = w\mu_{pi} + w\mu_{pi}$$

$$u\mu_{pj} = u\mu_{pi} + u\mu_{pi}$$

$$w\sigma_{pj} = w\sigma_{pi} + w\sigma_{pi}$$

$$u\sigma_{pj} = u\sigma_{pi} + u\sigma_{pi}$$

Datos derivados 3/D

Cálculo de distancias de carbono alfa atómicas del par de residuo, D_{aj}

- Aplicación de la geometría de Pitágoras para las coordenadas de carbono alfa de cada par de residuo, j

15 Calcular:

Para cada par de residuo, j :

$$D_{aj} = \text{Sqrt}((x_{A1i} - x_{A2i})^2 + (y_{a1i} - y_{a2i})^2 + (z_{A1i} - z_{A2i})^2)$$

Y para la muestra de estructuras en la base de datos, m

μ_{aj} = Promedio de todos los D_{aj}

20 v_{aj} = Mediana de todos los D_{aj}

σ_{aj} = Desviación estándar de todos los D_{aj}

$\text{Max}_{\alpha_j} = \text{Máximo de todos los } D_{\alpha_j}$ $\text{Min}_{\alpha_j} = \text{Mínimo de todos los } D_{\alpha_j}$

Cálculo de diferencia entre las distancias de carbono alfa - y beta del par de residuo, Δ_j

- Aplicación de la geometría de Pitágoras a las coordenadas de carbono beta de par de residuo, y sustracción

5 Calcular:

Para cada par de residuo, j:

D_{β_j} : fórmula como se describe para la medición de distancia de carbono alfa con medición de distancia de carbono beta con coordenadas de carbono beta X_{B1} y Z_{B1}

$$\Delta_j = D_{\alpha_j} - D_{\beta_j}$$

10 y para la muestra de estructuras en la base de datos, m

P_{Δ_j} = Promedio de todos los Δ_j

v_{Δ_j} = Mediana de m Δ_j

σ_{Δ_j} = Desviación estándar de todos los Δ_j

Max_{Δ_j} = Máximo de todos los Δ_j

15 Min_{Δ_j} = Mínimo de todos los Δ_j

Cálculo de ángulos 3D, ϕ_j y ψ_j

Definir:

ϕ_j = ángulo descrito por los átomos (puntos) $C\beta_{1i}$ - $C\alpha_{1i}$ - $C\alpha_{2i}$

ψ_j = ángulo descrito por los puntos $C\beta_{2i}$ - $C\alpha_{2i}$ - $C\alpha_{1i}$

20 va_{1j} = vector de $C\alpha_{1i}$ a $C\alpha_{2i}$,

va_{2j} = vector de $C\alpha_{2i}$ a $C\alpha_{1i}$,

vb_{1j} = vector de $C\alpha_{1i}$ a $C\beta_{1i}$,

vb_{2j} = vector de $C\alpha_{2i}$ a $C\beta_{2i}$,

Calcular:

25 coordenadas de vector, para cada par de residuo, j :

va_{1j}	va_{2j}	vb_{1j}	vb_{2j}
$X_{va1j} = X_{A2i} - X_{A1i}$	$X_{va2j} = X_{A1i} - X_{A2i}$	$X_{vb1j} = X_{B1i} - X_{A1i}$	$X_{vb2j} = X_{B1i} - X_{A2i}$
$Y_{va1j} = Y_{A2i} - Y_{A1i}$	$Y_{va2j} = Y_{A1i} - Y_{A2i}$	$Y_{vb1j} = Y_{B1i} - Y_{A1i}$	$Y_{vb2j} = Y_{B1i} - Y_{A2i}$
$Z_{va1j} = Z_{A2i} - Z_{A1i}$	$Z_{va2j} = Z_{A1i} - Z_{A2i}$	$Z_{vb1j} = Z_{B1i} - Z_{A1i}$	$Z_{vb2j} = Z_{B1i} - Z_{A2i}$

Calcular:

Ángulo φ_j (con base en productos escalares), para cada par de residuo, j

$$\varphi_j = \arccos \left(\frac{(x_{valj} * x_{vblj} + y_{valj} * y_{vblj} + z_{valj} * z_{vblj})}{\sqrt{x_{valj}^2 + y_{valj}^2 + z_{valj}^2} * \sqrt{x_{vblj}^2 + y_{vblj}^2 + z_{vblj}^2}} \right)$$

Y para la muestra de estructuras en la base de datos, m

5 μ_{φ_j} = Promedio de todos los φ_j

v_{φ_j} = Mediana de todos los φ_j

σ_{φ_j} = Desviación estándar de todos los φ_j

Max_{φ_j} = Máximo de todos los φ_j

Min_{φ_j} = Mínimo de todos los φ_j

10 Calcular:

Ángulo ψ_j (con base en productos escalares), para cada par de residuo, j

$$\Psi_j = \arccos \left(\frac{(x_{va2j} * x_{vb2j} + y_{va2j} * y_{vb2j} + z_{va2j} * z_{vb2j})}{\sqrt{x_{va2j}^2 + y_{va2j}^2 + z_{va2j}^2} * \sqrt{x_{vb2j}^2 + y_{vb2j}^2 + z_{vb2j}^2}} \right)$$

Y para la muestra de estructuras en la base de datos, m

15 P_{ψ_j} = Promedio de todos los ψ_j

v_{ψ_j} = Mediana de todos los ψ_j

σ_{ψ_j} = Desviación estándar de todos los ψ_j

Max_{ψ_j} = Máximo de todos los ψ_j

Min_{ψ_j} = Mínimo de todos los ψ_j

20 Cálculo del tercer ángulo 3D

Definir:

Vector g_{1_j} (vg_{1_j}): $A_{1_i} - B_{2_i}$

Plano E_{1_j} , descrito por los vectores va_{1_j} y vb_{1_j}

Plano E_{2_j} , descrito por los vectores va_{1_j} y vb_{1_j}

Vector n_{1j} (vn_{1j}), perpendicular a E_{1j} , el producto de vector de va_{1j} y vb_{1j}

Vector n_{2j} (vn_{2j}), perpendicular a E_{2j} , el producto de vector de va_{1j} y vb_{1j}

Calcular:

Coordenadas vg_{1j} , para cada par de residuo, j

vg_{1j}
$X_{vg1j} = X_{B2j} - X_{A1j}$
$Y_{vg1j} = Y_{B2j} - Y_{A1j}$
$Z_{vg1j} = Z_{B2j} - Z_{A1j}$

5

Calcular:

vn_1 y vn_2 coordinan (los productos de vector), para cada par de residuo, j

vn_{1j} = producto de vector de va_{1j} y va_{2j}

vn_{2j} = producto de vector de va_{1j} y vg_{1j}

vn_{1j}	vn_{2j}
$X_{vn2j} = Y_{va1j} * Z_{vb1j} - Y_{vb1j} * Z_{va1j}$	$X_{vn2j} = Y_{va1j} * Z_{vb1j} - Y_{vb1j} * Z_{vg1j}$
$Y_{vn2j} = Z_{va1j} * X_{vb1j} - Z_{vb1j} * X_{va1j}$	$Y_{vn2j} = Z_{va1j} * X_{vb1j} - Z_{vb1j} * X_{vg1j}$
$Z_{vn2j} = X_{va1j} * Y_{vb1j} - X_{vb1j} * Y_{va1j}$	$Z_{vn2j} = X_{va1j} * Y_{vb1j} - X_{vb1j} * Y_{vg1j}$

10

Calcular:

Ángulo entre vn_{1j} y vn_{2j} , ángulo χ_j , para cada par de residuo, j

$$\chi_j = \arccos \left(\frac{(x_{vn1j} * x_{vn2j} + y_{vn1j} * y_{vn2j} + z_{vn1j} * z_{vn2j})}{\text{Sqrt}(x_{vn1j}^2 + y_{vn1j}^2 + z_{vn1j}^2) * \text{sqrt}(x_{vn2j}^2 + y_{vn2j}^2 + z_{vn2j}^2)} \right)$$

Y para la muestra de estructuras en la base de datos, m

15 μ_{χ_j} = Promedio de todos los χ_j

v_{χ_j} = Promedio de todos los χ_j

σ_{χ_j} = Desviación estándar de todos los χ_j

Max_{χ_j} = Máximo de todos los χ_j

Min_{χ_j} = Mínimo de todos los χ_j

20

Compilación de Clasificaciones de Par de residuo; Datos Terciarios, Derivados

Clasificaciones de Par de residuos con base en Base de datos 2-D

Para cada propiedad física escogida para el proceso de selección

Definir:

T_p = suma de los valores numéricos de las propiedades físicas de los aminoácidos que se van a sustituir con cadenas de polipéptido (2 x valor de tirosina para el entrecruzamiento oxidativo de tirosina)

- 5 v_p = múltiplos permisibles de las desviaciones estándar ponderadas y no ponderadas de unos valores de la propiedad física, $u\sigma_{pj}$.

- Clasificación (R) con base en valores numéricos de una propiedad física, p, que corresponde a aminoácidos de ocurrencia, ponderada por la frecuencia de cada ocurrencia de aminoácido.

Calcular:

- 10 Para cada par de residuo, j

$$wR_{pj} = v_p * w\sigma_{pj} / (\text{abs}(T_p - w\mu_{pj} - v_p * w\sigma_{pj}))$$

Clasificación con base en valores numéricos de una propiedad física, p, que corresponde a aminoácidos de ocurrencia.

Calcular:

- 15 Para cada par de residuo, j

$$uR_{pj} = v_p * u\sigma_{pj} / (\text{abs}(T_p - u\mu_{pj} - v_p * u\sigma_{pj}))$$

Clasificaciones de Par de residuos con base en base de datos 3D

Separación de carbono alfa

Definir:

- 20 $v_{R\alpha}$ múltiplos permitidos de la desviación estándar de distancias de carbono alfa intercadena, $\sigma_{\alpha j}$

$vMax_{\alpha}$: valor máximo permitido para $\mu_{\alpha j}$ in el proceso de selección $vMin_{\alpha}$:

valor mínimo permitido para $\mu_{\alpha j}$ en el proceso de selección

T_{α} : Valor objetivo para la separación de carbono alfa

- 25 $R_{\alpha j}$: Clasificación con base en separación de carbono alfa intercadena, clasificaciones altas para pares de residuo, j, con valor objetivo $\mu_{\alpha j}$, T_{α} , y/o con valores $\sigma_{\alpha j}$ altos (flexibilidad)

Calcular:

T_{α} = promedio de $vMax_{\alpha}$ y $vMin_{\alpha}$

Para todos los pares de residuo, j

Para todos los $\mu_{\alpha j} < T_{\alpha}$:	Para todos los $\mu_{\alpha j} < T_{\alpha}$:
$R_{\alpha j} = (T_{\alpha} - \mu_{\alpha j} + V_{ra} * \sigma_{\alpha j})^2 / \sigma_{\alpha j}$	$R_{\alpha j} = (\mu_{\alpha j} + v_{ra} * \sigma_{\alpha j} - T_{\alpha})^2 / \sigma_{\alpha j}$

Ángulos Φ y ψ

Definir:

5 $V_{R\phi,\psi}$: múltiplos permitidos de la desviación estándar de ángulos ϕ_j y ψ_j , σ_{ϕ_j} y σ_{ψ_j} , $vMax_{\phi,\psi}$: valor máximo permitido para $\mu_{\alpha j}$ en el proceso de selección (mismo valor para ambos ángulos)

$vMin_{\phi,\psi}$: valor mínimo permitido para $\mu_{\alpha j}$ en el proceso de selección (mismo valor para ambos ángulos)

$T_{\phi,\psi}$: valor objetivo de ángulos ϕ y ψ (mismo valor para ambos ángulos)

R_{ϕ,ψ_j} : Clasificación con base en los ángulos ϕ y ψ ; clasificaciones altas para pares de residuo, j, con valores μ_{ϕ_j} y μ_{ψ_j} cercanos la valor objetivo, $T_{\phi,\psi}$, y/o con altos valores σ_{ϕ_j} y σ_{ψ_j} (flexibilidad)

10 r_{ϕ} : sub-clasificación con base en el ángulo ϕ

r_{ψ} : sub-clasificación con base en el ángulo ψ

Calcular:

$T_{\phi,\psi}$: promedio de $vMax_{\phi,\psi}$ y $vMin_{\phi,\psi}$

Para cada par de residuo, j

Para todos los $\mu_{\phi j} < T_{\phi,\psi}$:	Para todos los $\mu_{\psi j} < T_{\phi,\psi}$:
$r_{\phi j} = (T_{\phi,\psi} - \mu_{\phi j} + V_{R\phi,\psi} * \sigma_{\phi j})^2 / \sigma_{\phi j}$	$r_{\psi j} = (\mu_{\psi j} + V_{R\phi,\psi} * \sigma_{\psi j} - T_{\phi,\psi})^2 / \sigma_{\psi j}$
$r_{\phi j} = (T_{\phi,\psi} - \mu_{\phi j} + V_{R\phi,\psi} * \sigma_{\phi j})^2 / \sigma_{\phi j}$	$r_{\psi j} = (\mu_{\psi j} + V_{R\phi,\psi} * \sigma_{\psi j} - T_{\phi,\psi})^2 / \sigma_{\psi j}$
$R_{\phi,\psi j} = \text{promedio de } r_{\phi j} \text{ and } r_{\psi j}$	

15

Diferencia entre separación de carbono alfa- y beta

Definir:

vR_{Δ} : múltiplos permitidos de la desviación estándar para cada par de residuo, j, de diferencias m entre las distancias de carbono alfa y beta σ_{Δ_j}

20 $vMax_{\Delta}$: valor máximo permitido para μ_{Δ_j} en el proceso de selección

$vMin_{\Delta}$: valor mínimo permitido para μ_{Δ_j} en el proceso de selección

T_{Δ} : Valor objetivo para la diferencia entre la separación de carbono alfa beta

25 R_{Δ} : Clasificación con base en las diferencias entre las distancias de carbono alfa- y beta inter-cadena, clasificaciones altas para pares de residuo j, con calores μ_{Δ_j} , cercanos al valor objetivo T_{Δ_j} , y/o con altos valores σ_{Δ_j} (flexibilidad)

Calcular:

T_{Δ} = promedio de $v_{Max_{\Delta}}$ y $v_{Min_{\Delta}}$

Para todos los pares de residuo, j

Para todos los $\mu_{\Delta_j} < T_{\Delta}$	Para todos los $\mu_{\Delta_j} > T_{\Delta}$
$R_{\Delta_j} = (T_{\Delta} - \mu_{\Delta_j} + v_{R_{\Delta}} \cdot \sigma_{\Delta_j})^2 / \sigma_{\Delta_j}$	$R_{\Delta_j} = (\mu_{\Delta_j} + v_{R_{\Delta}} \cdot \sigma_{\Delta_j} - T_{\Delta})^2 / \sigma_{\Delta_j}$

Procesos de selección

- 5 La secuencia de filtros no es significativa

I 2D Procesos de selección

Filtro 1.1: Selección para los residuos conservados

Para todos los pares de residuo

Si los aminoácidos del par de residuos J reciben la asignación de marca 'con' (conservados), seleccionar

- 10 Si cualquier aminoácido de un par de residuos j se asigna 'no' (variable), descartar

Filtro L2: Selección contra residuos que tienen glicina como consenso

La selección de pares de los cuales ningún residuo es más frecuentemente glicina,

para todos los pares de residuo:

si el consenso (más a menudo aminoácido de ocurrencia) de ningún residuo de un par j es glicina, seleccionar

- 15 si el consenso (más a menudo aminoácido de ocurrencia) de cualquier residuo de un par j es glicina, descartar

Filtro I.3: Selección con base en mediciones estadísticas ponderadas

La selección utilizando mediciones estadísticas de una propiedad física, p, de aminoácidos de ocurrencia en cada residuo, i, de cada par de residuo, j, ponderado por la frecuencia de ocurrencia del aminoácido de ocurrencia.

Definir:

- 20 Max_{wRp} : límite máximo para la selección de una propiedad física de cadena lateral de aminoácido, p, con base en mediciones estadísticas ponderadas

Min_{wRp} : límite mínimo para la selección de una propiedad física de cadena lateral de aminoácido, p, con base en mediciones estadísticas ponderadas

Calcular:

- 25 SI $[Min_{wRp} < wR_{pj} < Max_{wRp}]$ es Verdadero, seleccionar

SI $[Min_{wRp} < wR_{pj} < Max_{wRp}]$ es Falso, descartar

Filtro I.4: Selección con base en mediciones estadísticas no ponderadas

La selección utilizando mediciones estadísticas de una propiedad física, p, de aminoácidos de ocurrencia en cada residuo, i, de cada par, j, no ponderado by la frecuencia de ocurrencia del aminoácido de ocurrencia

Definir:

Max_{uRp} : límite máximo para la selección de una propiedad física de cadena lateral de aminoácidos, p, con base en mediciones estadísticas ponderadas

Min_{uRp} : límite mínimo para la selección de una propiedad física de cadena lateral de aminoácidos, p, con base en mediciones estadísticas ponderadas

5 Calcular:

SI $[Min_{uRp} < uR_{pj} < Max_{uRp}]$ es Verdadero, seleccionar

SI $[Min_{uRp} < uR_{pj} < Max_{uRp}]$ es Falso, descartar

II Proceso de selección 3D

Filtro II.1: Selección para promedio distancias de carbono alfa dentro del rango de selección

10 Cálculo:

Para todos los pares de residuo:

SI $[vMin_{\alpha} < \mu_{\alpha j} < vMax_{\alpha}]$ es Verdadero, seleccionar

SI $[vMin_{\alpha} < \mu_{\alpha j} < vMax_{\alpha}]$ es Falso, descartar

Filtro II.2: Selección para flexibilidad suficiente de separación de carbono alfa

15 Cálculo:

Para todos los pares de residuos:

Para todos los $\mu_{\alpha j} < T_{\alpha}$

SI $[\mu_{\alpha j} + v_{R\alpha} * \sigma_{\alpha jm} > T_{\alpha}] = \text{Verdadero}$, seleccionar

SI $[\mu_{\alpha j} + v_{R\alpha} * \sigma_{\alpha j} > T_{\alpha}] = \text{Falso}$, descartar

20 Para todos los $\mu_{\alpha} > T_{\alpha}$

SI $[\mu_{\alpha j} - v_{R\alpha} * \sigma_{\alpha j} > T_{\alpha}] = \text{Verdadero}$, seleccionar

SI $[\mu_{\alpha j} + v_{R\alpha} * \sigma_{\alpha j} < T_{\alpha}] = \text{Falso}$, descartar

Filtro II.3 : Selección para pares con ángulos ϕ y ψ dentro del rango de selección

Cálculo:

25 SI $[vMin_{\phi,\psi} < \mu_{\phi j} < vMax_{\phi,\psi}]$ y $[vMin_{\phi,\psi} < \mu_{\psi j} < vMax_{\phi,\psi}]$ es Verdadero, seleccionar

SI $[vMin_{\phi,\psi} < J_{\phi j} < vMax_{\phi,\psi}]$ y $[vMin_{\phi,\psi} < \mu_{\psi j} < vMax_{\phi,\psi}]$ es Falso, descartar

Filtro II.4: Selección para diferencias promedio entre distancias de carbono alfa y beta dentro del rango de selección

$\mu_{\Delta j}$ = diferencia promedio entre residuo de distancias de carbono alfa y beta

Cálculo:

Para todos los pares de residuo

SI $[vMin_{\Delta} < \mu_{\Delta j} < vMax_{\Delta}]$ es Verdadero, seleccionar

SI $[vMin_{\Delta} < \mu_{\Delta j} < vMax_{\Delta}]$ es Falso, descartar

Filtro II.5: Selección para flexibilidad suficiente de la diferencia de pares entre distancias de carbono alfa y beta

5 Cálculo:

Para todos los pares de residuos:

Para todos los $\mu_{\Delta j} < T_{\Delta}$

SI $[\mu_{\Delta j} + v_{R\Delta} * \sigma_{\Delta j} > T_{\Delta}] =$ Verdadero, seleccionar

SI $[\mu_{\Delta j} + v_{R\Delta} * \sigma_{\Delta j} > T_{\Delta}] =$ Falso, descartar

10 Para todos los $\mu_{\alpha} > T_{\Delta}$

SI $[\mu_{\Delta j} - v_{R\Delta} * \sigma_{\Delta j} > T_{\Delta}] =$ Verdadero, seleccionar

SI $[\mu_{\Delta j} - v_{R\Delta} * \sigma_{\Delta j} > T_{\Delta}] =$ Falso, descartar

Selección final

Pares de aminoácidos seleccionados

15 Todos los pares de residuos, j, que se seleccionan en todos los Filtros (I.1-4 y II. 1-6) se compilan y enumeran.

Orden y Selección por Clasificación

Todos los residuos pares enumerados se comparan mediante su clasificación, y el par con las más altas clasificaciones en la SELECCIÓN FINAL.

6.12 Mutagénesis puntual y subclonación en vectores de expresión

20 6.12.1. Sustituciones conservadoras para residuos tirosina indeseados

Los fragmentos de cADN que codifican las cadenas ligera y pesada del fragmento Fv del anticuerpo anti- $\alpha 5$ -integrina monoclonal (ejemplo 1), o el anticuerpo anti- $\beta 1$ -integrina (ejemplo 2) se aíslan de los hibridomas que los producen de acuerdo con los procedimientos estándar conocidos en la técnica. Por ejemplo, el ARN se aísla del sedimento de un cultivo de suspensión de las células de hibridomas, el ARN se transcribe en forma inversa utilizando una mezcla de cebadores aleatorios y poli-A, y el cADN de las cadenas ligera y pesada se aíslan mediante el método RACE. Las secuencias de las cadenas ligera y pesada, que se van a entrecruzar de acuerdo con los procedimientos de la presente invención, se identifican mediante procedimientos estándar, y se alinean con el sistema de numeración K&W. Los residuos tirosina identificados se examinan por su proximidad predicha y la flexibilidad posicional de uno hacia el otro. Se identifican los pares de residuos en los que las cadenas laterales reactivas se encuentran en la secuencia que está dentro de un promedio de 15Å o menos en la muestra, o que tienen un promedio o desviación estándar, tal que el promedio menos una desviación estándar es 15Å o menos en la muestra. De estos pares, el residuo del par en el que ocurre la tirosina en la secuencia más baja en la base de datos 2-D, es un punto mutado a fenilalanina. Las mutaciones puntuales se introducen al utilizar el equipo de Mutagénesis Dirigida al Sitio QuikChange™ (Stratagene, Catalog #200518).

35 6.12.2. Sustitución de residuos de un par seleccionado con tirosina

En los residuos de par seleccionado, como se describió anteriormente, se introducen sustituciones de aminoácidos mediante mutación puntual, en la medida en que la tirosina ya no esté presente en los residuos seleccionados del par en las secuencias de las cadenas ligera y pesada de fragmento fv que se va a estabilizar. Las mutaciones puntuales se introducen utilizando el equipo de mutagénesis dirigida al sitio quikchange™ (véase arriba).

6.12.3. Sistema y vector de expresión

Se aíslan los fragmentos de ADN que codifican las cadenas ligera y pesada del fragmento Fv, conteniendo todas las sustituciones conservadoras de los residuos tirosina indeseados, identificados como se describió anteriormente, con y sin las sustituciones de residuos del par seleccionado con tirosina. Los fragmentos aislados (insertos) se subclonan en un vector de expresión que contiene el sitio de división pGEX que contiene el sitio de división TEV-proteasa. Para los propósitos de medición de los fragmentos Fv que retienen la afinidad por su antígeno, el inserto que codifica la cadena pesada también se fusiona con una secuencia de nucleótido que codifica una etiqueta de hemaglutinina (HA) en el extremo 3' (terminal C de la proteína), para lo cual un anticuerpo secundario está disponible comercialmente. Para los propósitos de utilizar el fragmento Fv en aplicaciones diagnósticas, terapéuticas o cualquier otra aplicación comercial, sin embargo, se debe eliminar de nuevo la etiqueta HA. La subclonación se lleva a cabo mediante procedimientos estándar conocidos en la técnica.

6.13 purificación y expresión bacteriana del fragmento Fv

Los plásmidos de expresión descritos anteriormente que codifican los fragmentos Fv ligeros y pesados modificados son componentes BL21 transformados o bacterias XA90. Se preparan (0.5 ml) de patrón de glicerol congelado a partir de clones individuales resistentes a la ampicilina, con el que se inoculan los cultivos de expresión (por ejemplo 1000 ml de Caldo de cultivo Loria: 10gm triptona, 5 gm extracto de levadura, 5 gm de NaCl) que contienen 100µg/ml de ampicilina). Se cultivan las células a 30° C sobre un agitador giratorio (300 rpm), y se induce la expresión de proteína con 1mM IPTG en un OD600 de 0.6. Luego de una incubación de tres horas, se cosechan bacterias mediante centrifugación a 4000g a 4° C. El sedimento se resuspende con 50 ml de Regulador de Lisis enfriado en hielo (20mM Tris.Cl pH 7.9, NaCl 500mM, 10% glicerol, β-mercaptoetanol 20mM, PMSF 1mM, 20µg/ml de leupeptina, 20µg/ml de pepstatina, aprotinina 1%) y luego se somete a sonicación sobre hielo hasta que se completa >90% de lisis. Se retira la materia insoluble mediante centrifugación a 20.000g a 4° C durante 20 min. Luego se incuban los sobrenadantes con 2 ml de sefarsa glutationa (Pharmacia) durante 2 horas a 4° C. Los glóbulos se peletizan mediante centrifugación a 4000g, y se lavan dos veces (se resuspenden y peletizan) en 10 ml de Regulador de Lisis y dos veces en 10 ml de Regulador de División TEV-proteasa (Novagen). Posteriormente se incuban los glóbulos con 1µg de proteasa TEV His etiquetada (Novagen) a 30° C durante 1 hora en 2 ml de Regulador de División. Posteriormente se retira la proteasa al agregar 0.1ml de suspensión NTA-agarosa equilibrada (Quiagen) a la suspensión. Se presentan fragmentos FvH y FvL parcialmente purificados en el sobrenadante luego de centrifugación a 4000g.

6.14. Introducción del entrecruzamiento tirosilo-tirosilo oxidativo

Los productos genéticos de cadena ligera y de cadena pesada de fragmentos Fv que solo contienen las mutaciones de los residuos tirosina indeseados reactivos a fenilalanina, sin las mutaciones del par de residuos seleccionados para tirosina se purifican parcialmente y se equilibran mediante diálisis en solución salina regulada con fosfato (PBS) antes de mezclarlas en una molaridad igual (0.1-1000µM). El catalizador, cloruro de metaloporfirin 20-tetrakis (4-sulfonatofenil)-21H,23H-porfino manganeso (III) (MnTPPS) se agrega después en hielo a concentraciones de 1µM, 5µM, 10µM, 50µM y 100µM a la reacción. La reacción se inicia después mediante la adición del oxidante monopersulfato de potasio a una concentración de 1-100µM, a temperatura ambiente o de otra forma, para cada una de las concentraciones del catalizador, y en diversas concentraciones de proteína. Después de 45 segundos se detiene la reacción mediante la adición de Tris.Cl pH 7.9 a 50 mM y β-mercaptoetanol a 10mM, y la solución se dializa de nuevo contra PBS para retirar el catalizador, los agentes de oxidación y reducción. Los monómeros y heterodímeros entrecruzados y no entrecruzados se aíslan mediante filtración por gel FPLC. La eficiencia de la reacción de entrecruzamiento se prueba mediante tinción con PAGE sin reducción y azul de Coomassie.

Se observa en cada concentración de proteína, la concentración máxima de agente de oxidación y catalizador en la que no se forma un entrecruzamiento entre los polipéptidos de la reacción. Estas condiciones se utilizan para catalizar la reacción entre los productos genéticos de cadena ligera y cadena pesada del fragmento Fv que contienen ambas mutaciones de los residuos tirosina indeseados reactivos a fenilalanina, y las mutaciones del par de residuos seleccionados a tirosina. Los monómeros y heterodímeros entrecruzados y no entrecruzados se aíslan mediante filtración por gel FPLC. La eficiencia de la reacción de entrecruzamiento se prueba mediante tinción con PAGE sin reducción y azul de Coomassie.

6.15. Prueba del complejo estabilizado

6.15.1. Rendimiento del complejo de fragmento Fv estabilizado funcionalmente

Se prueba el rendimiento de los fragmentos Fv funcionalmente entrecruzados al pasar una cantidad cuidadosamente determinada de una proteína de fragmento Fv purificada de gradiente de glicerol y entrecruzada en una columna de antígeno inmovilizada, y comparar el flujo pasante con el material de partida y la elución de la columna. Las

mediciones de concentración de proteína se llevan a cabo mediante procedimientos estándar, tal como ensayos Bradfords y Lowrie (Bradford, 1976, and Lowrie, 1954), tinción de Coomassie o con plata, o transferencia Western.

6.15.2. Afinidad retenida

5 Los fragmentos Fv que se entrecruzan exitosamente bajo las diversas condiciones descritas anteriormente se prueban en cuanto a su afinidad retenida en los procedimientos del tipo ELISA. Utilizando placas de 96 pozos, las superficies internas de los pozos de la placa de ensayo ELISA se cubren con antígeno, por ejemplo integrina $\alpha 5$ (ejemplo 1) e integrina $\beta 1$ (ejemplo 2). Los pozos se lavan, y uno con respecto al otro, la mitad de la concentración del anticuerpo de longitud completa y una concentración molar igual del fragmento F(ab) del anticuerpo (véase adelante) como controles positivos, y el fragmento Fv del anticuerpo, entrecruzado como se describió anteriormente, se incuban en PBS durante dos horas a 37° c en diluciones en serie en los pozos cubiertos con el antígeno respectivo sobre una placa. Los fragmentos F(ab) se derivan mediante digestión de pepsina del anticuerpo de longitud completa y la posterior purificación al retirar primero los fragmentos Fc al correr la solución de anticuerpo/proteasa a través de una columna de proteína A, y segundo mediante fraccionamiento del flujo pasante de la columna de proteína A mediante FPLC de intercambio iónico para eliminar la proteasa. Los pozos se lavan cuatro veces con 200 μ l de PBS y marcador anti-ha y anticuerpo secundario acoplado a fosfatasa alcalina y se incuban secuencialmente en PBS durante una hora adicional a 37° c. Los pozos se lavan de nuevo cuatro veces con 200 μ l de PBS. Las concentraciones de IgG enlazada, fragmento F(ab), y fragmento Fv se determinan mediante procedimientos estándar con un lector de prueba ELISA.

6.15.3 Estabilidad en suero, lisado, y citoplasma

20 La estabilidad del complejo en suero se prueba en experimentos en función del tiempo al incubar el complejo en suero humano a 37°C, 38°C, 39°C, 40°C, 42°C, y 45°C durante hasta dos semanas, y se prueban en cuanto a los restantes niveles de complejos de fragmentos Fv funcionales. Como controles, se compara la estabilidad del Fab, scFv y/o ds/Fv, todas etiquetadas con el mismo marcador.

25 Se prueba la estabilidad del complejo en el citoplasma, también en experimentos en función del tiempo, en forma análoga a la incubación en el suero, al incubar el complejo en los lisados celulares. Más directamente, la estabilidad del complejo en el citoplasma se prueba al raspar-cargar las células de cultivo de tejido con fragmentos Fv y ensayar los restantes niveles de complejos funcionales. Como controles, se compara la estabilidad del scFv y dsFv de la misma molécula de inmunoglobulina original, ambas etiquetadas con el mismo marcador como el fragmento Fv entrecruzado.

30 En todos estos experimentos, los niveles restantes de los complejos funcionales se determinarán en ensayos ELISA con el mismo anticuerpo secundario, como se describió anteriormente.

6.15.4. Inmunogenicidad

35 Se inyectan ratones con varias dosis, que varían de 1 μ g a 10 mg, de complejo estabilizado. El complejo estabilizado se inyecta en la presencia y ausencia del adyuvante (Completo) de Freund. Se dan inyecciones adicionales a los ratones como refuerzos cada cinco días (en la presencia y ausencia del Adyuvante Incompleto). Los ratones reciben un total de tres o cuatro refuerzos de inmunizaciones.

Se toman muestras de sangre de la vena de la cola antes de cada inyección, y una semana después del refuerzo final. Las muestras sanguíneas se centrifugan a 3000g durante 30 min a 4° C.

40 Las placas ELISA se cubre con el complejo estabilizado y una mezcla de fragmento Fv no estabilizado pesado y cadenas ligeras, y se realizan ensayos ELISA de acuerdo con los procedimientos estándar, utilizando un anticuerpo secundario anti ratón etiquetado.

La inmunogenicidad de los complejos estabilizados por los métodos de la presente invención se compara con las construcciones dsFc y scFv de la misma molécula de inmunoglobulina original como controles.

6.15.5. Biodistribución

45 Se inyectan en ratones fragmentos Fv estabilizados radiomarcados ^{18}F , marcados de acuerdo con los procedimientos publicados por Lang L. y Eckelmann U., 1994. Cada ratón se inyecta con 3 μ g de menos de 4.5 MBq/ μ g de complejo de fragmento Fv. Los animales se sacrifican en 15, 45, 90, 360 min. y 24 h. e inmediatamente se desangran mediante punción cardíaca. Los tejidos se separan, se secan y pesan en una balanza analítica, y se cuentan en un contador de radiación gamma utilizando una configuración de alta energía (para ^{18}F). También se

secan y cuentan las alícuotas de sangre. Los conteos se corrigen por deterioro. Tejido: índice de sangre, y se calcula para cada tejido el porcentaje de dosis inyectada por gramo de tejido.

5 Se realizan estudios de depuración sanguínea en fase temprana en ratones inyectados con la misma cantidad de los anteriormente descritos fragmentos Fv estabilizados radiomarcados Fv ¹⁸F. Se toman muestras de sangre en serie en la vena de la cola en 1, 2, 5, 10, 15, y 30 min. Las muestras se secan y se cuentan como se describió anteriormente, y la vida útil de los fragmentos Fv en la sangre se calcula de acuerdo con procedimientos estándar (Choi C.W. et al. Cancer Research; vol. 55: pp. 5323-5329, 1995).

Como controles para los anteriores estudios, se comparan las construcciones de fragmentos Fv de disulfuro y cadena sencilla de la misma molécula de inmunoglobulina original.

10 7. EJEMPLO II

Los siguientes ejemplos ilustran ciertas variaciones de los métodos de la invención para la estabilización de proteínas. Este ejemplo se presenta a manera de ilustración y no a manera de limitación del alcance de la invención.

7.1 Introducción

15 En años recientes se han identificado diversos polipéptidos con significativo valor comercial y, adicionalmente, para muchos de estos polipéptidos los datos estructurales están disponibles.

20 En la siguiente sección, se describen los métodos para estabilización de un polipéptido, un biocatalizador, para el que solo están disponibles los datos del polipéptido propiamente dichos, pero no para los otros polipéptidos estructural o funcionalmente relacionados. Específicamente, se describe más adelante el proceso de selección del par de residuos y la reacción de entrecruzamiento propiamente dicha. Para la descripción de la introducción de mutaciones puntuales, expresión bacteriana de lo polipéptidos y su purificación, y ajuste de las condiciones de reacción de entrecruzamiento, vea la Sección 6, arriba.

25 El biocatalizador estabilizado en el ejemplo más adelante es la lipasa B de Candida antarctica ("CALB", Figura 1C), una enzima para la que son posibles múltiples aplicaciones comercialmente muy importantes debido a su exquisita enantioselectividad, algunas de las cuales siguen siendo poco rentables gracias a su falta de estabilidad bajo ciertas condiciones de reacción adversas.

30 El archivo de estructuras que contienen las coordenadas atómicas tridimensionales de la estructura de cristal del polipéptido se obtiene de la base de datos de proteínas del laboratorio nacional de Brookhaven (Brookhaven National Laboratory Protein Database). Los datos derivados relacionados con el proceso de selección se calculan como se describió. El proceso de selección se lleva a cabo utilizando un grupo de filtros que es conveniente y apropiados para esta aplicación de la presente invención.

35 Se introducen mutaciones puntuales para tirosina (dirigiendo la reacción de entrecruzamiento) de acuerdo con la selección final del proceso de selección, y las mutaciones puntuales para la fenilalanina (limitando la reacción de entrecruzamiento) como se describe en la Sección 6, arriba. El polipéptido se expresa desde el punto de vista bacteriano como una proteína de fusión GST, y se purifica y divide proteolíticamente, también como se describe en la Sección 6, arriba. El núcleo hidrófobo de la proteína, a la que se dirige el entrecruzamiento, se expone al desnaturalizar con guanidina, en una concentración en la que la proteína se repliega y forma el enlace. Se ajustan las condiciones de reacción mínimas requeridas utilizando una construcción con las mutaciones para fenilalanina, pero faltando las mutaciones a tirosina, y la reacción de entrecruzamiento se lleva a cabo después con las construcciones que contienen ambos grupos de mutaciones puntuales. La eficiencia de la reacción se prueba en 40 cuanto al biocatalizador estabilizado, y el biocatalizador estabilizado resultante se prueba luego en cuanto a especificidad y actividad retenida, estabilidad mejorada en tiempo, y bajo condiciones más adversas.

7.2. Ventajas del entrecruzamiento tirosilo-tirosilo para biocatalizadores

45 La química fundamental de la tecnología cubierta por la presente invención provoca un entrecruzamiento oxidativo para conformar entre cadenas laterales reactivas de polipéptidos que forman complejos estables. Debido a que la reacción de entrecruzamiento se cataliza, el entrecruzamiento, una vez establecido, es estable en la ausencia del catalizador bajo un amplio rango de pH y condiciones de oxidorreducción. La reacción de entrecruzamiento requiere una proximidad muy cercana entre las cadenas laterales reactivas que se entrecruzarán.

50 Así, la presente invención describe una nueva tecnología que permite la estabilización de los biocatalizadores y permite su uso en un más amplio rango de aplicaciones industriales. Esta tecnología se diseña para mejorar las tecnologías precedentes, y complementar las tecnologías compatibles.

Los biocatalizadores estabilizados resultantes tendrán las siguientes características:

1. Las enzimas serán más estables bajo un amplio rango de condiciones de reacción, que incluyen, pero no se limitan a, temperatura, pH, presión, salinidad, o concentración de otros compuestos en la reacción, tal como un agente de reducción, que es frecuentemente un componente de la reacción química para la que se requiere el catalizador.

2. El biocatalizador estabilizado y entrecruzado resultante retendrá su actividad y especificidad debido a la especificidad de la reacción de entrecruzamiento y al proceso de selección.

Esta tecnología de estabilización es apta para el desarrollo de nuevos productos con aplicaciones novedosas, la mejora de los biocatalizadores industriales existentes, y la complementación de las tecnologías existentes para el desarrollo de biocatalizadores novedosos.

7.3 Aplicaciones de los biocatalizadores

Las enzimas biocatalíticas constituyen la clase preferida de catalizadores para procesos industriales debido a sus altos índices de especificidad y rotación, y sus bajos costes de desarrollo y tiempos de ciclo. Sin embargo, su utilidad está limitada por la inestabilidad relativa y vida útil limitada de moléculas de proteína que se exacerba por reacciones adversas y/o condiciones de almacenamiento.

La aplicación de la presente invención estabiliza las enzimas con entrecruzamientos internos colocados específicamente, y por lo tanto aumenta la estabilidad de las enzimas sin deteriorar su actividad en las condiciones de reacción deseadas. El aumento resultante en la estabilidad de las enzimas no solo supera de esta manera las limitaciones de la vida útil sino que también aumenta los índices de reacción de las enzimas y los rendimientos de proceso.

Los procesos biocatalíticos se utilizan en muchos sectores de la industria, que incluyen las industrias químicas, de detergentes, farmacéuticas, agrícolas, alimenticias, cosméticas, textiles, procesamiento de materiales, y del papel. Dentro de estas industrias, los biocatalizadores tienen muchas aplicaciones, que varían desde síntesis de producto (por ejemplo, manufactura de aminoácidos, y síntesis por química fina de fármacos de moléculas pequeñas) mediante el uso como agentes activos en productos (por ejemplo, en polvos para lavado biológico) para uso en equipos de pruebas de diagnóstico. Los biocatalizadores también tienen aplicaciones industriales que varían desde tratamiento de aguas residuales y suelos agrícolas, hasta refinación de petróleo crudo (por ejemplo, desulfurización).

7.4. Selección de residuos óptimos para entrecruzamiento tirosilo-tirosilo

El proceso de selección consiste de una serie de pruebas o "filtros" dirigidos a reducir sucesivamente los pares de residuos que den como resultado más probablemente un par de tirosina entrecruzado que altere mínimamente la actividad o especificidad de la enzima, al mismo tiempo que proporciona máxima estabilidad.

7.4.1. Datos utilizados para el análisis

Coordinar los datos para los cálculos de distancia de todos los átomos diferentes de hidrógeno de la estructura resuelta cristalográficamente que se descargó de la base de datos de estructuras de proteína del Brookhaven National Laboratory (www.bnl.pdb.gov; Figura 5). Estos datos proporcionan las coordenadas tridimensionales (x, y, y z) para cada átomo en la estructura resuelta, expresadas en unidades métricas, es decir Angströms (10-10m, Å). Estos datos también contienen la secuencia y/o el uso de los aminoácidos del polipéptido. Con estos datos es posible calcular las distancias tridimensionales entre cualquiera de los átomos deseados (por ejemplo átomos de carbono alfa y beta amino).

7.4.2. Metodología de selección

Los residuos óptimos, a los que se dirige la reacción de entrecruzamiento, se seleccionan mediante una serie de filtros basados en las mediciones de valores en una base de datos compilada para los propósitos de esta selección. Esta base de datos contiene mediciones numéricas de (1) separación de carbono alfa, (2) separación de carbono beta y la diferencia entre las distancias alfa y beta, y (3) uso del residuo de aminoácido (véase adelante).

7.5. Filtro 1: Selección de residuos aromáticos suficientemente separados

Debido a que hay un número significativo de residuos aromáticos disponibles en la secuencia de CALB, y debido a la mutación de un residuo aromático (diferente a tirosina, es decir triptófano, fenilalanina, o histidina) a tirosina sería

máximamente conservadora, para el proceso de selección de este ejemplo, solo se analizan los pares de residuos aromáticos.

Adicionalmente para maximizar el grado al cual la aplicación de la presente invención estabiliza la enzima, solo se seleccionan los pares que se separan más de 40 aminoácidos en la secuencia de aminoácidos bidimensional

Par de residuos CALB		Distancia de carbono alfa	Diferencia de distancia C α -C β
Phe9	Tyr82	9.29	-0.20
Phe48	Trp104	8.85	1.53
Trp52	Tyr234	8.71	0.02
Phe131	Tyr183	6.19	-1.31
Trp104	His224	9.33	0.33
Tyr135	Tyr203	7.58	0.10
Tyr183	His224	8.20	-1.09
Phe117	Tyr300	7.7	2.07

5

7.6. Filtro 2: Identificación de pares de residuos separados apropiadamente

Para encontrar pares de residuos separados apropiadamente para un enlace tirosilo-tirosilo, la distancia de carbono alfa a carbono alfa entre cada par de residuos en el polipéptido se calcula en una base de datos 3D. Este cálculo se realiza al aplicar la geometría de Pitágoras a las coordenadas 3D de los carbonos alfa (Figura 6). Con base en los anteriores cálculos, como un segundo corte, se seleccionan todos los pares de residuo cuyos carbonos alfa se separan dentro del rango de selección.

10

Debido a la falta de mediciones estadísticas que dan idea de la flexibilidad posicional, se reduce el rango de selección a 2 Å, pero solo en el límite superior.

El rango que se seleccionó es el siguiente:

15

Min 5.70 Å, Max 9.74 Å.

7.7. Filtro 3: Orientación de cadena lateral

Las cadenas laterales de tirosina se deben orientar una hacia la otra para que se forme un entrecruzamiento con mínima distorsión estructural. La diferencia entre la distancia del carbono alfa (es decir la distancia del carbono de estructura principal; Figura 6) y la distancia del carbono beta (es decir la distancia entre los primeros carbonos en cada cadena lateral; Figura 8) de cada par de residuos se calcula como un proxy, es decir un estimado de la orientación de las cadenas laterales en relación una con la otra (Figura 9).

20

El rango que se seleccionó es el siguiente:

Min-2Å, Max 3.0Å.

De nuevo, con base en la geometría de coordenadas 3D, para cada par de residuo, se calcula la distancia entre los carbonos beta (Figura 8). Después se resta la distancia beta de la distancia alfa del par de residuos (Figura 9). Este filtro se basa en si la diferencia en las distancias alfa y beta de un par de residuos cae dentro del rango óptimo estimado. En este ejemplo, todos los pares de residuo en la Tabla 11 cumplen con este criterio.

25

7.8. Filtro 3: Eliminación parcial de pares con residuos en proximidad al sitio activo de la enzima

La funcionalidad de una enzima como un biocatalizador radica en su capacidad para catalizar una reacción química. La actividad y la selectividad de un catalizador es más sensible en aquellos sitios en donde el catalizador y los reactivos entran en contacto físicamente uno con el otro. Por lo tanto, las mutaciones y/o entrecruzamientos son por lo menos deseables en el sitio activo, y se excluyen los residuos en o próximos al sitio activo.

- 5 La His224 está en el sitio activo y por lo tanto se excluye. Debido a que la Tyr183 está muy próxima a la His224, los residuos seleccionados adelante se deben mutar para generar polipéptidos con los pares de tirosina, con y sin la mutación de Tyr183 a Phe183. Adicionalmente, debido a que la His224 también está muy cerca del Trp104, y debido a que el Trp104 está muy cerca de la Phe28, también se excluyen los pares de residuo que contienen los anteriores residuos. Los restantes pares de residuos se enumeran en la Tabla 12.

- 10 Tabla 12. Lista de pares de residuos restantes con mediciones de distancia relevantes.

Par de residuos CALB	Diferencia de distancia C α -C β	Diferencia de distancia C α -C ϵ
Phe117 Tyr300	7.7	2.07
Trp52 Tyr234	8.71	0.02
Tyr135 Tyr203	7.58	0.10
Phe9 Tyr82	9.29	-0.20

7.8. Análisis de distancia de carbono épsilon

- 15 Debido a que el isómero más probable del enlace di-tirosina se considera que es el enlace épsilon-épsilon, y debido a que están disponibles los datos de las coordenadas para un átomo en la posición épsilon de todos los aminoácidos seleccionados, se analizan las distancias entre las posiciones épsilon de los pares de residuos seleccionados anteriormente en la Tabla 12.

Los pares en la Tabla 12 se clasifican de acuerdo con sus distancias de carbono épsilon. Sin embargo, debido a que se requieren tres de los cuatro pares en una mutación puntual para generar un par tirosina, las distancias se pueden alterar, y se generan y examinan todos los pares.

- 20 7.9 mutaciones puntuales y sistema de expresión

Véase Sección 6, arriba, para una descripción de la sustitución de residuos de los pares seleccionados con tirosina, del sistema y el vector de expresión, y la purificación y expresión bacteriana.

7.10 Introducción del entrecruzamiento tirosilo-tirosilo oxidativa

- 25 Los polipéptidos que sólo contienen las mutaciones de los residuos tirosina indeseados reactivos a fenilalanina, y las mutaciones del par de residuos seleccionados a tirosina se purifican parcialmente y equilibran mediante diálisis en solución salina regulada con fosfato (PBS) a 0.1-1000mM. El núcleo hidrófobo de la proteína se expone mediante desnaturalización con guanidina 6M. El catalizador, cloruro de metaloporfirina 20- tetrakis (4-sulfonatofenil)-21H,23H-porfirina manganeso (III) (MnTPPS) se agrega posteriormente en hielo a una concentración de 1mM, 5mM, 10mM, 50mM y 100mM a la reacción. La guanidina se diluye después a 1:2, 1:5, 1:10, 1:20, 1:50, 1:100, 1:200, 1:500, y 1:1000 para permitir que la proteína se repliegue, y la reacción se inicia mediante la adición del oxidante monopersulfato de potasio. El oxidante se agrega a una concentración de 1-100mM, a temperatura ambiente o de otra forma, para cada una de las concentraciones del catalizador, y en diversas concentraciones de proteína. Después de 45 segundos se detiene la reacción mediante la adición de TrisCl PH7.9 50 mM y b-mercaptoetanol 10 mM. Y la solución se dializa nuevamente contra PBS para retirar la guanidina restante, el catalizador, agentes de oxidación y reducción. Se aíslan los monómeros y los biocatalizadores entrecruzados y no entrecruzados mediante filtración por gel FPLC. Se prueba la eficiencia de la reacción de entrecruzamiento mediante tinción con PAGE y azul de Coomassie.
- 30
- 35

7.11 Ensayo de estabilidad, actividad y especificidad de calb

Mediante cualquiera de los métodos conocidos en la técnica, se analiza el CALB entrecruzado resultante en cuanto a su actividad y especificidad retenidas, y para mejoras en su estabilidad bajo condiciones adversas. La actividad de lipasa, por ejemplo, se determina mediante la hidrólisis de p-nitrofenilésteres de ácidos grasos con varias longitudes de cadena (>10) y detección espectrofotométrica de p-nitrofenol a 410 nm. Alternativamente, se puede hidrolizar el ácido 1,2-O-dilauril-rac-glicero-3-glutárico (Boeringer Mannheim Roche GmbH, Germany), produciendo resorufina, que se puede detectar espectrofotométricamente a 572 nm, o fluorométricamente a 583 nm. Adicionalmente, se puede hidrolizar un número de alquildiacilgliceroles novedosos sintetizados fluorogénicamente con longitud de onda dual de estereoselectividad y actividad de lipasa (Jaeger K-E. et al. *Annu. Rev. Microbiol.* vol. 53: pp. 315-51,1999 and Zandonella G. et al. *J. Mol. Catal. B: Enzym.*, vol. 3: pp. 127-130). La estabilidad mejorada se determina mediante protocolos estándar, que incluyen análisis calorimétrico y/u otros análisis termodinámicos, y experimentos en función del tiempo bajo condiciones fisiológicas o condiciones más adversas, tal como presiones, temperaturas mayores, condiciones de pH más rigurosas, salinidad aumentada o reducida, la presencia y ausencia de reactivos de oxidación y reducción, y la inclusión de varias concentraciones de diferentes solventes orgánicos.

8. Referencias

- 15 Campbell L.A. et al. Protein Cross-linking Mediated by Metalloporphyrins. *Bioorganic and Medicinal Chemistry*, vol. 6: pp. 1301-1037, 1998
- Brown K.C. et al. Highly Specific Oxidative Cross-link of Proteins Mediated by a Nickel-peptide Complex. *Biochem.*; vol. 34(14): pp. 4733-4739, 1995
- Pollitt S. and Schultz P. *Agnew. Chem. Int. Ed.*; vol. 37(15): pp. 2104-2107, 1998
- 20 Spangler B.D. and Erman J.E. Cytochrome c Peroxidase Compound I: Formation of Covalent Protein Crosslinks During the Endogenous Reduction of the Active Site. *Biochim. Biophys. Acta*; vol. 872(1-2): pp. 155-7,1986
- Gmeiner B. and Seelos C. Phosphorylation of Tyrosine Prevents Dityrosine Formation in vitro. *FEBS Lett* ; vol. 255 (2): pp. 395-7,1989
- 25 Kanwar R. and Balasubramanian D. Structure and Stability of the Dityrosine-linked Dimer of GammaB-crystallin. *Exp. Eye Res.*; vol. 68(6): pp. 773-84, 1999
- Fancy D.A. and Kodadek T. Chemistry for the Analysis of Protein-protein Interactions: Rapid and Efficient Crosslinking Triggered by Long Wavelength Light. *Proc. Natl. Acad. Sci., U.S.A.*; vol. 96: pp. 6020-24, 1999
- Klinman J.P. (ed.). *Redox-active Amino Acids in Biology. Methods in Enzymology*; vol. 258, 1995
- 30 Richards, F.M. The Interpretation of Protein Structures: Total Volume, Group Volume Distributions and Packing Density. *J. Mol. Biol.*; vol. 82: pp. 1-14, 1974
- Eisenberg, D. Three-dimensional Structure of Membrane and Surface Proteins. *Ann. Rev. Biochem.*; vol. 53: pp. 595-623,1984
- National Brookhaven Laboratory Protein Database (on-line at www.nbl.pdb.gov)
- 35 Pastan et al. Recombinant Disulfide Stabilized Polypeptide Fragments Having Binding-specificity. Patente de los Estados Unidos No. 5,747,654, emitida el 5 de mayo de 1998
- Hofmann K. The Modular Nature of Apoptotic Signaling Proteins. *Cell Mol. Life Sci.*; vol. 55(8-9): pp. 1113-28, 1999
- Johnson, G. et al. *Weir's Handbook of Experimental Immunology I. Immunochemistry and Molecular Immunology*, Fifth Edition, Ed. L. A. Herzenberg, W. M. Weir, and C. Blackwell, Blackwell Science Inc., Cambridge, MA, Chapter 6.1-6.21, 1996 Wickelgren I. Mining the genome for drugs. *Science*; vol. 285(5430): pp. 998-1001, 1999
- 40 Leong S.R. et al. IL-8 single-chain homodimers and heterodimers: interactions with chemokine receptors CXCR1, CXCR2, and DARC. *Protein Sci.*; vol. 6(3): pp: 609-17, 1997
- Pawson T. Tyrosine Kinase Signalling Pathways. *Princess Takamatsu Symp.*; vol. 24: pp. 303-22, 1994

- Cowburn D. Peptide Recognition by PTB and PDZ Domains. *Curr. Opin. Struct. Biol.*; vol. 7(6): pp. 835-8, 1997
- Bockaert J. and Pin J.P. Molecular Tinkering of G Protein-coupled Receptors: an Evolutionary Success. *EMBO J.*; vol. 18(7): pp. 1723-9, 1999
- Royet J. et al. Notchless Encodes a Novel WD40-repeat-containing Protein that Modulates Notch Signaling Activity. *EMBO J.* ; vol. 17(24): pp. 7351-60, 1998
- Chou J.J. et al. Solution Structure of the RAIDD CARD and Model for CARD/CARD Interaction in Caspase-2 and Caspase-9 Recruitment. *Cell*; vol. 94(2): pp. 171-80, 1998
- Black R.A. and White J.M. ADAMs: Focus on the Protease Domain. *Curr Opin Cell Biol.*; vol. 10(5): pp. 654-9, 1998
- Strasser A. and Newton K. FADD/MORT1, a Signal Transducer that Can Promote Cell Death or Cell Growth. *Int. J. Biochem. Cell. Biol.*; vol. 31(5): pp. 533-7, 1999
- McInnes C. and Sykes B.D. Growth Factor Receptors: Structure, Mechanism, and Drug Discovery. *Biopolymers*; vol. 43(5): pp. 339-66, 1997
- Lotz M. et al. The Nerve Growth Factor/Tumor Necrosis Factor Receptor Family. *J. Leukoc. Biol.*; vol. 60(1): pp. 1-7, 1996
- Casaccia-Bonnel P. et al. p75 Neurotrophin Receptor as a Modulator of Survival and Death Decisions. *Microsc Res Tech.*; vol. 45(4-5): pp. 217-24, 1999
- Natoli G. et al. Apoptotic, Non-apoptotic, and Anti-apoptotic Pathways of Tumor Necrosis Factor Signalling. *Biochem. Pharmacol.*; vol. 56(8): pp. 915-20, 1998
- Alber T. Structure of the Leucine Zipper. *Curr. Opin. Genet. Dev.*; vol. 2(2): pp. 205-10, 1992
- Griffith T.S. et al. Functional Analysis of TRAIL Receptors Using Monoclonal Antibodies. *J. Immunol.*; vol. 162(5): pp. 2597-605, 1999
- Yasuda H. et al. Identity of Osteoclastogenesis Inhibitory Factor (OCIF) and Osteoprotegerin (OPG): a Mechanism by which OPG/OCIF Inhibits Osteoclastogenesis in vitro. *Endocrinology*; vol. 139(3): pp. 1329-37, 1998
- Ortiz A. et al. New Kids in the Block: the Role of FasL and Fas in Kidney Damage. *J. Nephrol.*; vol. 12(3): pp. 150-8, 1999
- Price Waterhouse: Survey of Biopharmaceutical Industry, 1998 Boston Consulting Group: The Contribution of Pharmaceutical Companies: What's at stake for America, 1993 Pharmaceutical Research and Manufacturers of America. *New Medicines in Development, Survey.* www.phrma.org/publications/industry/profile99/chap2.html, 1998
- Penuche M.L. et al. Antibody-IL-2 Fusion Proteins: a Novel Strategy for Immune Protection. *Hum Antibodies*; vol. 8(3): pp. 106-18, 1997
- Sensel M.G. et al. Engineering Novel Antibody Molecules. *Chem. Immunol.*; vol. 65: pp. 129-58, 1997
- Reiter Y. and Pastan I. Recombinant Fv Immunotoxins and Fv Fragments as Novel Agents for Cancer Therapy and Diagnosis. *TIBTECH*; vol. 16(12): pp. 513-520, 1998
- Reiter Y. et al. Engineering Antibody Fv Fragments for Cancer Detection and Therapy: Disulfide-stabilized Fv Fragments. *Nat Biotech.*; vol. 14: pp. 1239-1245, 1996
- Pluckthun A. and P. Pack. New Protein Engineering Approaches to Multi-valent and Bi-specific Antibody Fragments. *Immunotechnology*; vol. 3(2): pp. 83-105, 1997
- Wright A. and Morrison S.L. Effect of Glycosylation on Antibody Function: Implications for Genetic Engineering. *Trends Biotechnol.*; vol. 15(1): pp. 26-32, 1997

- Schwartz M.A. et al. Monoclonal Antibody Therapy. *Cancer Chemother. Biol. Response Modif.*; vol. 13:pp. 156-74, 1992
- Houghton A.N. and Scheinberg D.A. Monoclonal Antibodies: Potential Applications to the Treatment of Cancer. *Semin Oncol.*; vol. 13(2): pp. 165-79, 1986
- 5 Cao Y. and Suresh M.R. Bi-specific Antibodies as Novel Bio-conjugates. *Bioconjugate Chemistry*; vol. 9(6): pp. 635-644, 1998
- Raag R. and Whitlow M. Single-chain Fvs. *FASEB*; vol. 9: pp. 73-80, 1995
- Webber K.O. et al. Preparation and Characterization of a Disulfide-stabilized Fv Fragment of the Anti-Tac Antibody: Comparison with its Single-chain Analog. *Mol. Immunol.*; vol. 32(4): pp. 249-258, 1995
- 10 Klinman J.P. (ed.). Redox-active Amino Acids in Biology. *Methods in Enzymology*, vol. 258, 1995
- Bosilevac J.M. et al. Inhibition of Activating Transcription Factor 1- and cAMP-responsive Element-binding Protein-activated Transcription by an Intracellular Single-chain Fv fragment. *J. Biol. Chem.*; vol. 273(27): pp. 16874-16879, 1998
- 15 Graus-Porta D. et al. Single Chain Mediated Intracellular Retention of ErbB-2 Impairs Neu Differentiation Factor and Epidermal Growth Factor Signaling. *Mol. Cell Biol.*; vol 15: pp. 1182-1191, 1995
- Richardson J.H. et al. Phenotypic Knockout of the High-affinity Interleukin 2 Receptor by Intracellular Single Chain Antibodies against the Alpha Subunit of the Receptor. *Proc. Nat. Acad. Sci., USA*; vol. 92: pp. 3137-3141, 1995
- Maciejewski J.P. et al. Intracellular Expression of Antibody Fragments Directed against Human Immunodeficiency Virus Reverse Transcriptase Prevents HIV Infection in vitro. *Nat. Med.*; vol. 1: pp. 667-673, 1995
- 20 Marasco W.A. et al. Design, Intracellular Expression, and Activity of a Human Anti-human Immunodeficiency Virus Type I gp120 Single Chain Antibody. *Proc. Nat. Acad. Sci., USA*; vol. 90: pp. 7889-7893, 1993
- Levy Mintz P. et al. Intracellular Expression of Single Chain Variable Fragment to Inhibit Early Stages of the Vira Life Cycle by Targeting Human Immunodeficiency Virus Type 1 Integrase. *J. Virol.*; vol. 70: pp. 8821-8832, 1996
- 25 Duan L. et al. Intracellular Immunization Against Human Immunodeficiency Virus Type I Infection of Human T Lymphocytes: Utility of Anti-rev Single Chain Variable Fragment. *Hum. Gene Ther.*; vol. 6(12): pp. 1561-1573, 1995
- Kim S.H. et al. Expression and Characterization of Recombinant Single-chain Fv and Fv Fragments Derived from a Set of Catalytic Antibodies. *Mol. Immunol.*; vol. 34(12-13): pp. 891-906, 1997
- Choi C.W. et al. Biodistribution of 18F- and 125I-labelled Anti-Tac Disulfide-stabilized Fv Fragments in Nude Mice with Interleukin 2 a Receptor-positive Tumor Xenografts. *Cancer Research*; vol. 55: pp. 5323-5329, 1995
- 30 Colcher D. et al. Pharmacokinetics and Biodistribution of Genetically-engineered Antibodies. *Q J Nucl Med.*; vol. 42(4): pp. 225-41, 1998
- Pavlinkova G. et al. Pharmacokinetics and Biodistribution of Engineered Single-chain Antibody Constructs of MAb CC49 in Colon Carcinoma Xenografts. *J. Nucl. Med.*; vol. 40(9): pp. 1536-46, 1999
- Antibody Engineering Page, IMT, University of Marburg, FRG: <http://aximt1.imt.uni-marburg.de/rek/indexfenster.html>
- 35 Hunkapiller M. et al. A Microchemical Facility for the Analysis and Synthesis of Genes and Proteins. *Nature*; vol. 310(5973): pp. 105-11, 1984
- Xia X and Li WH. What Amino Acid Properties Affect Protein Evolution, *J. Mol. Evol.*; vol. 47(5): pp. 557-64, 1998
- Sandberg M, et al. New Chemical Descriptors Relevant for the Design of Biologically Active Peptides. A Multivariate Characterization of 87 Amino Acids. *J. Med. Chem.*; vol. 41(14): pp. 2481-91, 1998
- 40 Hopp T.P. and Woods K.R. Prediction of Protein Antigenic Determinants from Amino Acid Sequences. *Proc. Natl. Acad. Sci., U.S.A.*; vol. 78: pp. 3824, 1981

- Bradford, M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-dye Binding. *Anal. Biochem.*; vol. 72: pp. 248-54, 1976
- Lowry, O. J. *Biol. Chem.*; vol. 193, pp. 265, 1951
- 5 Lei S.P. et al. Characterization of the *Erwinia Carotovora pelB* Gene and its Product Pectate Lyase. *J. Bacteriol.*; vol. 169: pp. 4379-83, 1987
- Chou P.Y. and Fasman G.D. Prediction of Protein Conformation. *Biochemistry*; vol. 13(2): pp. 222-45, 1974
- Lang L. and Eckelmann W.C. One-step Synthesis of ¹⁸F labeled [18F]-N-succinimidyl 4-(fluoromethyl) benzoate for Protein Labeling. *Appl. Radiat. Isot.*; vol. 45: pp. 1155-63, 1994
- Sambrook et al.; Glover (ed.). *DNA Cloning: A Practical Approach*. MRL Press, Ltd., Oxford, U.K.; vol. I, II, 1985
- 10 Benton and Davis. Screening *Lambd*gt Recombinant Clones by Hybridization to Single Plaques in situ. *Science*; vol. 196(4286): pp. 180-2, 1977
- Clemmons D.R. IGF Binding Proteins and their Functions. *Mol. Reprod. Dev.*; vol. 35: pp. 368-374, 1993
- Loddick S.A. et al. Displacement of Insulin-like Growth Factors from their Binding Proteins as a Potential Treatment for Stroke. *Proc. Natl. Acad. Sci., U.S.A.*; vol. 95: pp. 1894-1898, 1998
- 15 Swift G.H. et al. Tissue-specific expression of the rat pancreatic elastase I gene in transgenic mice. *Cell*; vol. 38: pp. 639-646, 1984
- Hanahan D. Heritable formation of pancreatic beta-cell tumours in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. *Nature*; vol. 315: pp. 115-122, 1985
- 20 Grosschedl R. et al. Introduction of a mu immunoglobulin gene into the mouse germ line: specific expression in lymphoid cells and synthesis of functional antibody. *Cell*; vol. 38: pp. 647-658, 1984
- Leder A et al. Consequences of widespread deregulation of the c-myc gene in transgenic mice: multiple neoplasms and normal development. *Cell*; vol. 45: pp. 485-495, 1986
- Pinkert C.A. et al. An albumin enhancer located 10 kb upstream functions along with its promoter to direct efficient, liver-specific expression in transgenic mice. *Genes Dev.*; vol. 1: pp. 268-276, 1987
- 25 Krumlauf R. et al. Developmental regulation of alpha-fetoprotein genes in transgenic mice. *Mol. Cell. Biol.*; vol. 5: pp. 1639-1648, 1985
- Kelsey G.D. et al. Species- and tissue-specific expression of human alpha 1-antitrypsin in transgenic mice. *Genes Dev.*; vol. 1: pp. 161-171, 1987
- 30 Magram J. et al. Developmental regulation of a cloned adult beta-globin gene in transgenic mice. *Nature*; vol. 315: pp. 338-340, 1985
- Readhead C. et al. Expression of a myelin basic protein gene in transgenic shiverer mice: correction of the dysmyelinating phenotype. *Cell*; vol. 48: pp. 703-712, 1987
- Shani M. Tissue-specific expression of rat myosin light-chain 2 gene in transgenic mice. *Nature*; vol. 314: pp. 283-286, 1985
- 35 Mason A.J. et al. The hypogonadal mouse: reproductive functions restored by gene therapy. *Science*; vol. 234: pp. 1372-1378, 1986
- Smith D.B. and Johnson K.S. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene*; vol. 67: pp. 31-40, 1988
- 40 Lei S.P. et al. Characterization of the *Erwinia carotovora pelB* gene and its product pectate lyase. *J. Bacteriol.*, vol. 169: pp. 4379, 1987

- Kim S.H. et al. Expression and characterization of recombinant single-chain Fv and Fv fragments derived from a set of catalytic antibodies. *Mol. Immunol*, vol. 34: pp. 891-906, 1997
- Cale J.M. et al. Optimization of a reverse transcription-polymerase chain reaction (RT-PCR) mass assay for lowabundance mRNA. *Methods Mol. Biol.*; vol. 105: pp. 351-71, 1998
- 5 Weis J.H. et al. Detection of rare mRNAs via quantitative RT-PCR. *Trends Genet.*; vol. 8(8): pp. 263-4, 1992
- Frohman M.A. On beyond classic RACE (rapid amplification of cDNA ends). *PCR Methods Appl.*; vol. 4(1): pp. S40-58, 1994
- Adams P.D. et al. Extending the limits of molecular replacement through combined simulated annealing and maximum-likelihood refinement. *Acta Crystallogr. D. Biol. Crystallogr.*; vol. 55 (Pt 1): pp. 181-90, 1999
- 10 Schwarze S.R. et al. In Vivo Protein Transduction: Delivery of a Biologically Active Protein into the Mouse. *Science*; vol. 285: pp. 1565-72, 1999
- Hoffman R.M. Topical liposome targeting of dyes, melanins, genes, and proteins selectively to hair follicles. *J. Drug Target.*; vol. 5(2): pp. 67-74, 1998
- 15 Pluckthun A. et al. Catalytic antibodies: contributions from engineering and expression in *Escherichia coli*. *Ciba Found. Symp.*; vol. 159: pp. 103-12; discussion 112-7, 1991
- Guogiang J. et al. Dimerization Inhibits the Activity of Receptor-like Protein-tyrosine Phosphatase alpha. *Nature*; vol. 401: pp.606-610, 1999
- BIC, Explorer, Business Opportunities in Technology Commercialization. Illanes A. Stability of biocatalysts. *Elec.J.Biotech.*, vol. 2(1): pp. 7-15, 1999
- 20 DeSantis G. and Jones J.B. Chemical modification of enzymes for enhanced functionality. *Curr. Op. Biotech.*, vol. 10(4): pp. 324-340, 1999
- Govardhan C.P. Crosslinking of enzymes for improved stability and performance. *Curr Opin Biotechnol*. Aug; vol 10(4):331-5, 1999
- Beguín P. Hybrid enzymes. *Curr. Op. Biotech.*, vol. 10(4): pp. 336-340, 1999
- 25 Haring D. and Schreier P. Cross-linked enzyme crystals. *Curr Opin Chem Biol.*; vol. 3(1): pp. 35-8, 1999
- Moreno-Hagelsieb G. and Soberon X. Protein engineering as a powerful tool for the chemical modification of enzymes. *Biol Res.*; vol. 29(1): pp. 127-40, 1996 Jaeger K-E. et al. Bacterial Biocatalysts: Molecular Biology, Three-Dimensional Structures, and Biotechnological Applications of Lipases. *Annu. Rev. Microbiol.* vol. 53: pp. 315-51, 1999.

REIVINDICACIONES

1. Un método para elaborar una proteína estabilizada que comprende:
- (a) seleccionar uno o más pares de residuos en una cadena o cadenas de polipéptidos para entrecruzamiento;
 - (b) sustituir por lo menos uno de los residuos seleccionados con tirosina dando como resultado que ambos residuos seleccionados son tirosina; y
 - (c) entrecruzar los pares de residuos;
- 5
- en donde las proteínas entrecruzadas di-tirosina retienen por lo menos una actividad funcional desplegada por la proteína en la ausencia de entrecruzamiento de di-tirosina, y donde al menos una tirosina de un entrecruzamiento de di-tirosina es el resultado de una sustitución de aminoácido a tirosina.
- 10
2. Un método como se reivindica en la reivindicación 1 que comprende adicionalmente la sustitución de por lo menos un residuo tirosina por un residuo de aminoácido que no se entrecruza bajo condiciones de entrecruzamiento.
- 15
3. Un método como se reivindica en la reivindicación 1 o reivindicación 2, en donde la función retenida se selecciona del grupo que consiste en actividad catalítica y especificidad de unión.
- 20
4. Un método como se reivindica en cualquiera de las reivindicaciones 1 a 3, en donde la proteína se selecciona del grupo que consiste en una enzima, un factor de crecimiento, una hormona y un anticuerpo o fragmento funcional del mismo.
- 25
5. Un método como se reivindica en cualquiera de las reivindicaciones 1 a 4, en donde ocurre la reacción de entrecruzamiento en la presencia de un oxidante seleccionado del grupo que consiste en peróxido de hidrógeno, oxona, hexahidrato de ácido magnesio monoperoxiftálico (MMPP), un oxidante fotogenerado, y persulfato de amonio.
- 30
6. Un método como se reivindica en cualquiera de las reivindicaciones 1 a 4, en donde el entrecruzamiento se cataliza mediante un catalizador seleccionado del grupo que consiste en polihistidina, Gly-Gly-His, una metaloporfirina, una peroxidasa, o cualquier combinación de las mismas.

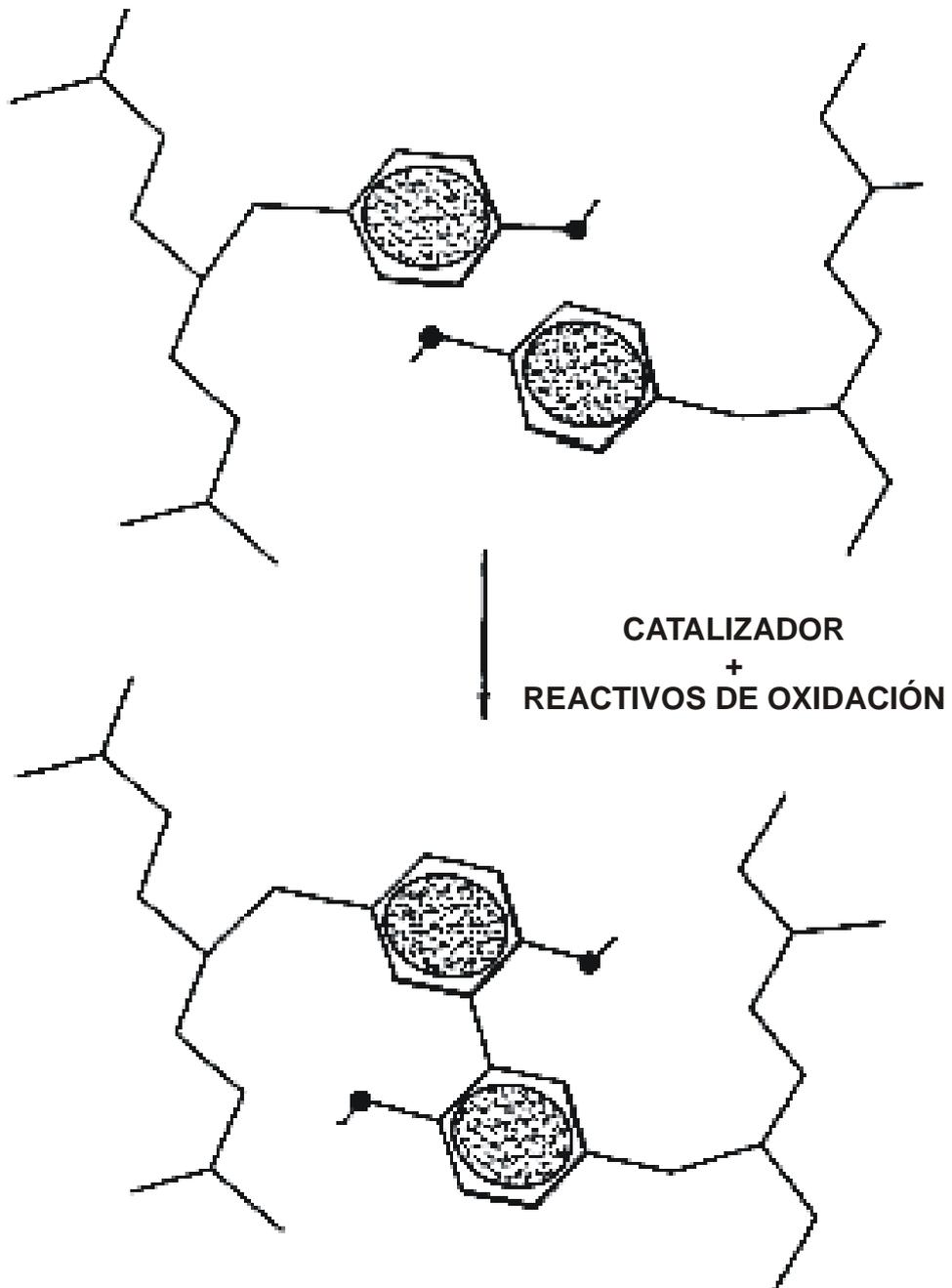


FIG.1A

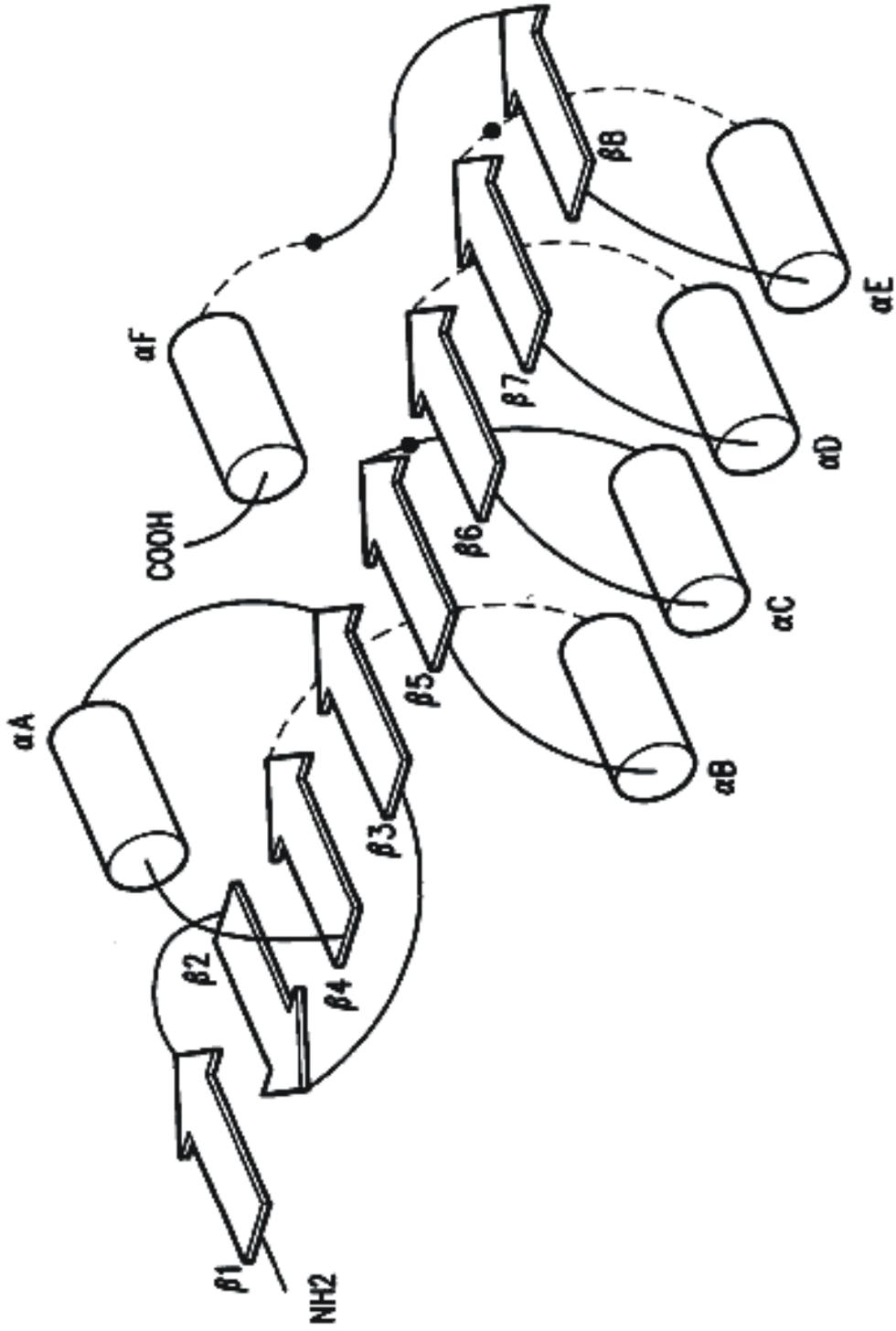


FIG.1B

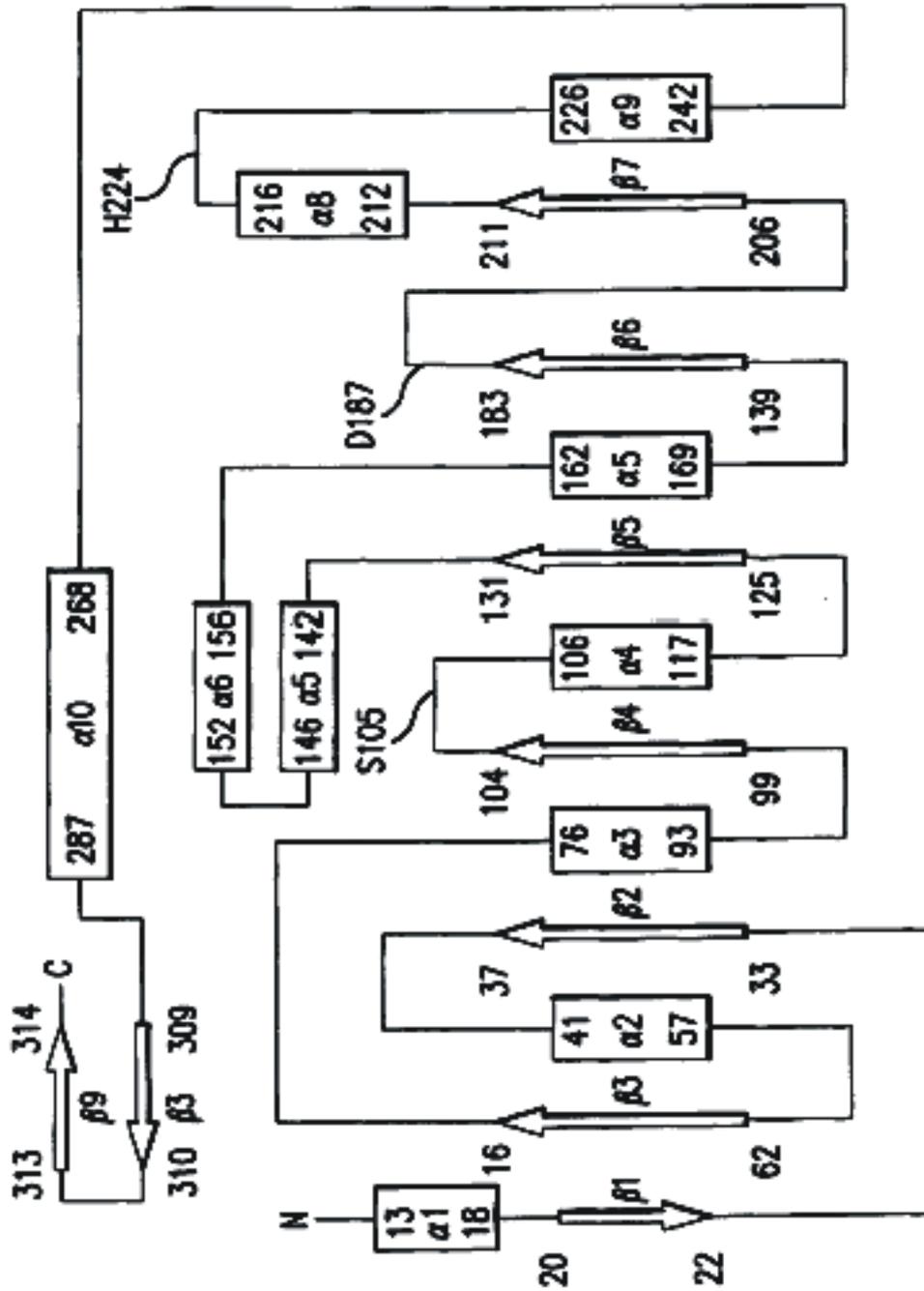


FIG.1C

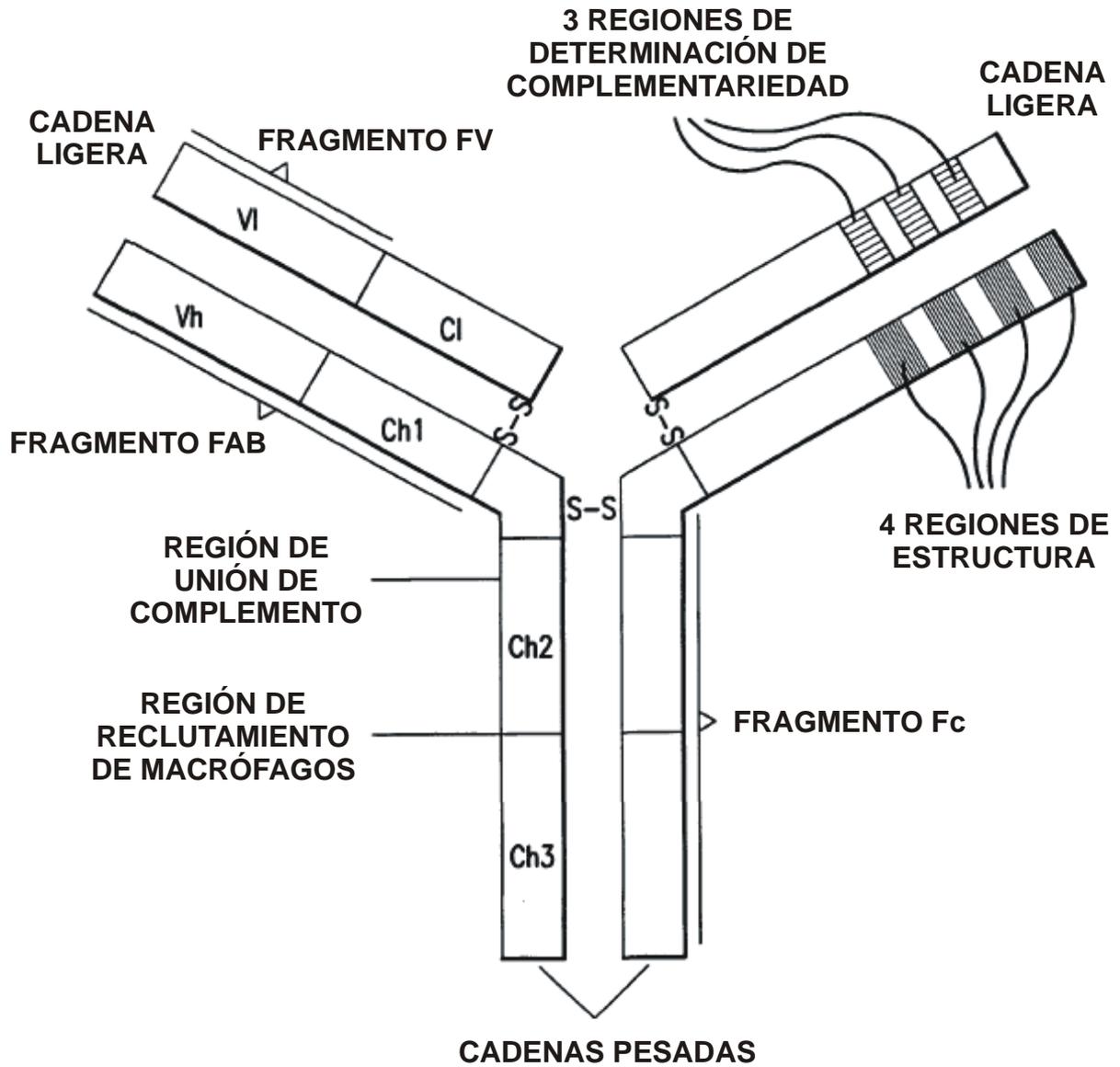


FIG.1D



FIG.2A

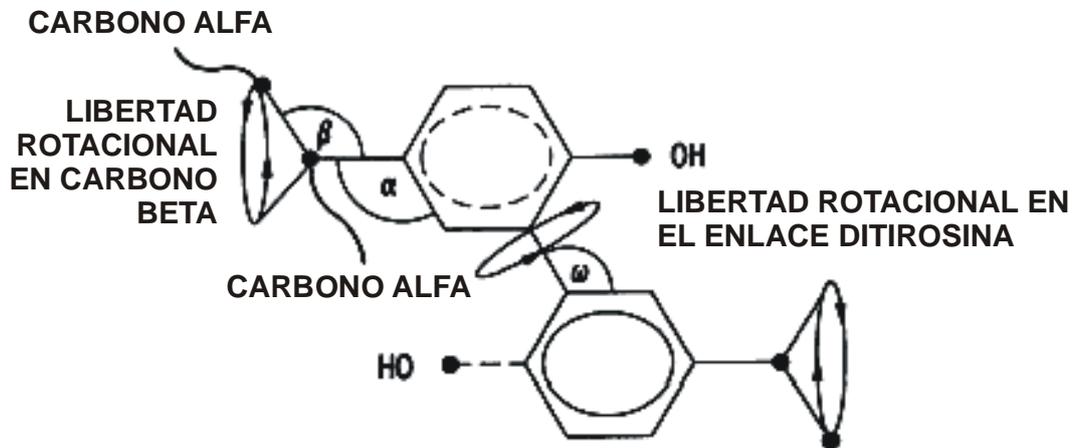


FIG.2B

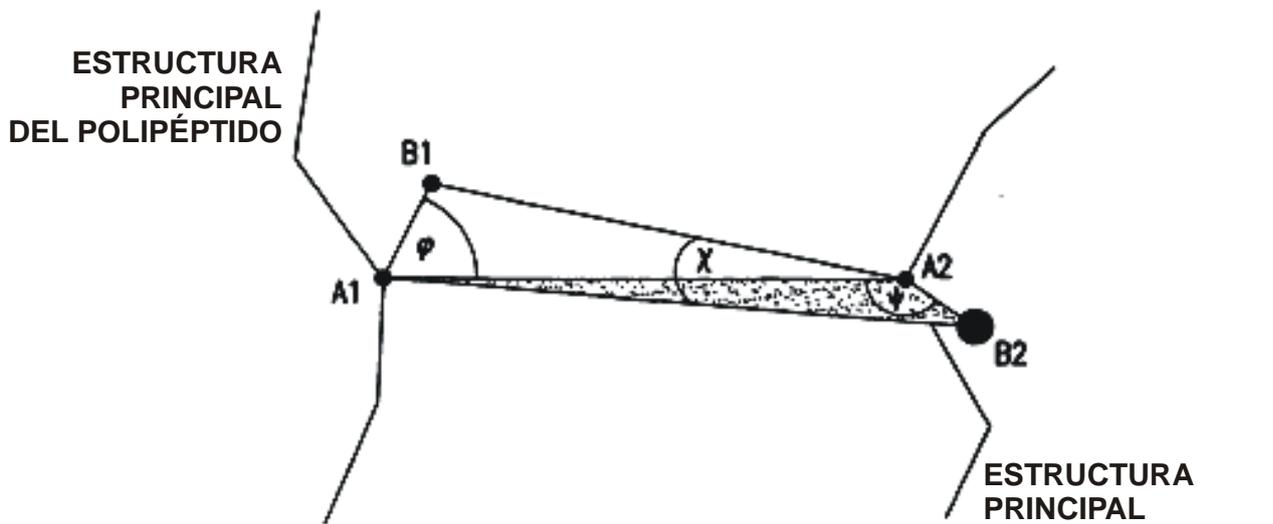


FIG.2C

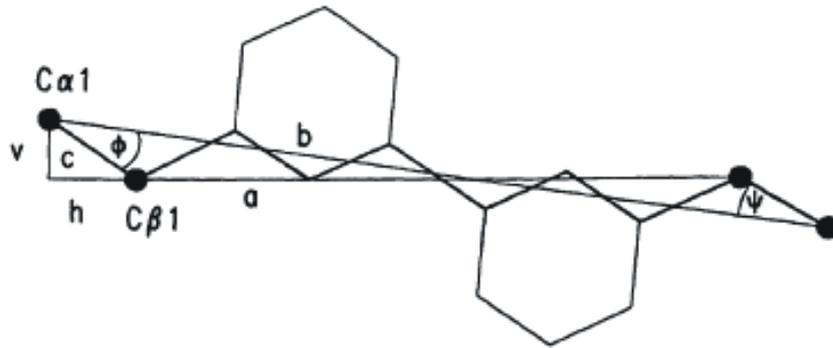


FIG.3A

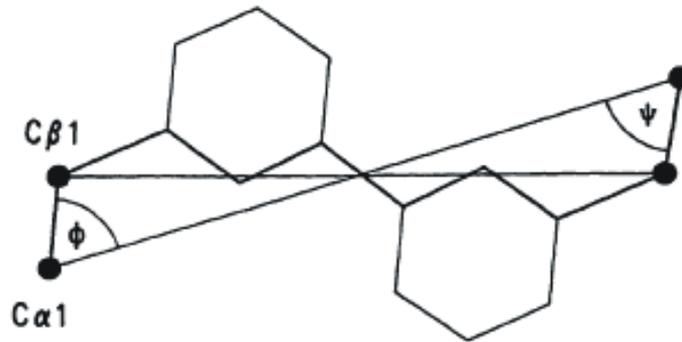


FIG.3B

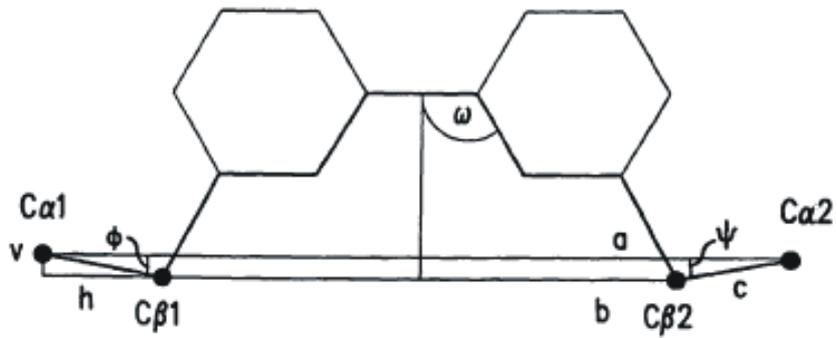


FIG. 4A

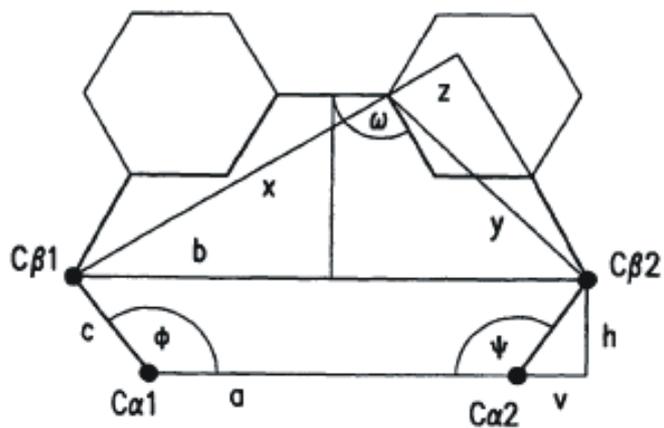


FIG. 4B

CADENA LIGERA (L)

CADENA	K&W	ÁTOMO	AMINOÁCIDO	COORD. x	COORD. y	COORD. z
L	1	N	ASP	2.37	-5.00	-27.24
L	1	C α	ASP	2.98	-3.78	-26.64
L	1	C	ASP	1.91	-2.70	-26.52
L	1	O	ASP	1.33	-2.29	-27.53
L	1	C β	ASP	4.14	-3.29	-27.53
L	1	C γ	ASP	5.18	-2.49	-26.76
L	1	O δ 1	ASP	4.86	-1.38	-26.27
L	1	O δ 2	ASP	6.34	-2.97	-26.65
L	2	N	ILE	1.63	-2.26	-25.30
L	2	C α	ILE	0.60	-1.24	-25.07
L	2	C	ILE	1.19	0.15	-24.94
L	2	O	ILE	2.14	0.35	-24.94
L	2	C β	ILE	-0.21	-1.52	-23.78
L	2	C γ 1	ILE	-0.90	-2.88	-23.86
L	2	C γ 2	ILE	-1.24	-0.43	-23.58
L	2	C δ 1	ILE	-1.66	-3.26	-22.59
L	3
.
.

FIG.5A

CADENA	K&W	ÁTOMO	AMINOÁCIDO	COORD. x	COORD. y	COORD. z
H	1	N	GLU	11.12	-2.19	9.00
H	1	C α	GLU	11.43	-1.08	8.05
H	1	C	GLU	11.93	-1.63	6.71
H	1	O	GLU	13.10	-1.98	6.56
H	1	C β	GLU	12.47	-0.12	8.66
H	1	C γ	GLU	13.82	-0.75	9.05
H	1	C δ	GLU	13.70	-1.77	10.17
H	1	O ϵ 1	GLU	13.38	-1.36	11.31
H	1	O ϵ 2	GLU	13.94	-2.97	9.92
H	2	N	ILE	11.02	-1.70	5.74
H	2	C α	ILE	11.36	-2.24	4.42
H	2	C	ILE	12.10	-1.22	3.59
H	2	O	ILE	11.77	-0.04	3.64
H	2	C β	ILE	10.11	-2.68	3.62
H	2	C γ 1	ILE	9.31	-3.73	4.39
H	2	C γ 2	ILE	10.52	-3.22	2.28
H	3	C δ 1	ILE	8.49	-3.17	5.55
H	3
.
.

FIG.5B

FRAGMENTO Fv 1

Ch	K&W	At	AA	x	y	z	Ch	L	L	L	L	L
							K&W	1	2	3	4	5
							At	Cα	Cα	Cα	Cα	Cα
							AA	Asp	Ile	.	.	.
							x	2.98	0.60	.	.	.
							y	-3.78	-1.24	.	.	.
							z	-26.64	-25.07	.	.	.
H	1	Cα	Glu	11.43	-1.08	8.05		35.80	34.84	.	.	.
H	2	Cα	Ile	11.36	-2.24	4.42		32.21	31.42	.	.	.
H	3	Cα
H	4	Cα
H	5	Cα

FIG.6A

FRAGMENTO Fv 2

Ch	K&W	At	AA	x	y	z	Ch	L	L	L	L	L
							K&W	1	2	3	4	5
							At	Cα	Cα	Cα	Cα	Cα
							AA	Glu	Ser	.	.	.
							x	35.61	31.94	.	.	.
							y	83.10	83.89	.	.	.
							z	56.99	56.85	.	.	.
H	1	Cα	Glu	10.23	61.09	64.74		34.48	32.46	.	.	.
H	2	Cα	Val	13.63	62.72	65.19		31.07	29.20	.	.	.
H	3	Cα
H	4	Cα
H	5	Cα

FIG.6B

FRAGMENTO Fv 3

Ch	K&W	At	AA	x	y	z	Ch	L	L	L	L	L
							K&W	1	2	3	4	5
							At	C α				
							AA	Glu	Ser	.	.	.
							x	19.56	19.09	.	.	.
							y	-13.02	-15.06	.	.	.
							z	-15.86	-12.67	.	.	.
H	1	C α	GLN	26.71	9.76	10.88		35.84	35.05	.	.	.
H	2	C α	Val	27.45	8.61	7.34		32.69	32.11	.	.	.
H	3	C α
H	4	C α
H	5	C α

FIG.6C

RESIDUO	PARES	PROMEDIO	DESV. EST.	MAX	MIN	MEDIANA
H1	L1	35.38	0.78	35.84	34.48	35.80
H1	L2	34.12	1.44	35.05	32.46	34.84
H1	L3
H1	L4
.
H1	L106
H2	L1	31.99	0.83	32.69	31.07	32.21
H2	L2	30.91	1.52	32.11	29.20	31.41
H2	L3
H2	L4
.
H2	L106
H3	L1

FIG.7A

RESIDUO	PARES	PROMEDIO	DESV. EST.	MAX	MIN	MEDIANA
H1	L1	35.09	1.56	37.37	31.23	35.54
H1	L2	34.00	1.87	37.36	29.92	34.38
H1	L3
H1	L4
.
H1	L106
H2	L1	32.26	1.57	36.71	30.34	32.14
H2	L2	31.32	1.99	36.77	29.20	31.11
H2	L3
H2	L4
.
H2	L106
H3	L1

FIG.7B

Ch	K&W	At	AA	x	y	z	Ch	L	L	L	L	L
							K&W	1	2	3	4	5
							At	C β				
							AA	ASP	ILE	.	.	.
							x	4.14	-0.21	.	.	.
							y	-3.29	-1.52	.	.	.
							z	-27.53	-23.78	.	.	.
H	1	C β	GLU	12.47	-0.12	8.66		37.27	34.85	.	.	.
H	2	C β	ILE	10.11	-2.68	3.62		31.73	29.30	.	.	.
H	3	C β
H	4	C β
H	5	C β

FIG.8

DISTANCIAS ALFA							Ch	L	L	L	L	L
							K&W	1	2	3	4	5
							At	C α				
							AA	ASP	ILE	.	.	.
							x	2.98	0.60	.	.	.
							y	-3.78	-1.24	.	.	.
							z	-26.64	-25.07	.	.	.
Ch	K&W	At	AA	x	y	z						
H	1	C α	GLU	11.43	-1.08	8.05	35.80	34.84	.	.	.	
H	2	C α	ILE	11.36	-2.24	4.42	32.21	31.42	.	.	.	
H	3	C α	
H	4	C α	
H	5	C α	

FIG.9A

DISTANCIAS BETA							Ch	L	L	L	L	L
							K&W	1	2	3	4	5
							At	C β				
							AA	ASP	ILE	.	.	.
							x	4.14	-0.21	.	.	.
							y	-3.29	-1.52	.	.	.
							z	-27.53	-23.78	.	.	.
Ch	K&W	At	AA	x	y	z						
H	1	C β	GLU	12.47	-0.12	8.66	37.27	34.85	.	.	.	
H	2	C β	ILE	10.11	-2.68	3.62	31.73	29.30	.	.	.	
H	3	C β	
H	4	C β	
H	5	C β	

FIG.9B

DIFERENCIA			Ch	L	L	L	L	L	
			K&W	1	2	3	4	5	
DIFERENCIAS ENTRE LAS DISTANCIAS DE CARBONO ALFA Y BETA DEL PAR DE RESIDUOS			AA	ASP	ILE	.	.	.	
						.	.	.	
					
					
Ch	K&W	AA							
H	1	GLU	-1.47	-0.01	.	.	.		
H	2	ILE	0.48	2.10	.	.	.		
H	3		
H	4		
H	5		

FIG.9C

		L1	L2	L3	L4	L5	.	.
FRAGMENTO Fv 1	H1	-1.47	-0.01
	H2	0.48	2.10
	H3
	H4

	.	L1	L2	L3	L4	L5	.	.
FRAGMENTO Fv 2	H1	-1.61	0.46
	H2	0.18	2.04
	H3
	H3

	.	L1	L2	L3	L4	L5	.	.
FRAGMENTO Fv 3	H1	0.92	1.59
	H2	0.69	1.31
	H3
	H3

	.	L1	L2	L3	L4	L5	.	.
FRAGMENTO Fv 4	H1

FIG.10

RESIDUO	PARES	PROMEDIO	DESV. EST.	MAX	MIN	MEDIANA
H1	L1	-0.72	1.42	0.92	-1.61	-1.47
H1	L2	0.68	0.82	1.59	-0.01	0.46
H1	L3
H1	L4
.
.
H1	L106
H2	L1	0.45	0.26	0.69	0.18	0.48
H2	L2	0.68	0.82	1.59	-0.01	0.46
H2	L3
H2	L4
.
.
H2	L106
H3	L1

FIG.11A

RESIDUO	PARES	PROMEDIO	DESV. EST.	MAX	MIN	MEDIANA
H1	L1	-0.68	1.04	0.92	-2.20	0.83
H1	L2	0.34	0.82	2.37	-0.54	0.09
H1	L3
H1	L4
.
.
H1	L106
H2	L1	0.74	0.69	1.83	-0.18	0.59
H2	L2	1.78	0.50	2.55	0.75	1.94
H2	L3
H2	L4
.
.
H2	L106
H3	L1

FIG.11B

Res.	AA	F	AA	F	AA	F	AA	F	AA	F	AA	F
1	Glu	58	Glu	24	Asp	3	Glu	3	Gly	2	Ala	1
2	Val	99	Ile	2	Ala	1	Glu	1	Met	1	-	-
3	Gln	90	Thr	5	Glu	3	His	2	Leu	2	Lys	2
4	Leu	101	Val	3	-	-	-	-	-	-	-	-

FIG.12A

Aminoácido	van der Waals volúmenes [A ³]	Hidrofobicidad
Ala	67	0.62
Arg	148	-2.50
Asn	96	-0.78
Asp	91	-0.90
Cys	86	0.29
Gln	114	-0.85
Glu	109	-0.79
Gly	48	0.30
His	118	-0.40
Ile	124	1.40
Leu	124	1.10
Lys	135	-1.50
Met	124	0.64
Phe	135	1.20
Pro	90	0.12
Ser	73	-0.18
Thr	93	-0.05
Trp	163	0.81
Tyr	141	0.26
Val	105	1.10

FIG.12B

Res.	AA	F	AA	F	AA	F	AA	F	AA	F	AA	F
1	109	61	109	24	91	3	48	2	67	1	-	-
2	105	99	124	2	67	1	109	1	124	1	-	-
3	114	90	93	5	109	3	118	2	124	2	135	2
4	124	101	105	3	-	-	-	-	-	-	-	-

FIG.12C

VOLÚMENES VAN DER WAALS

CADENA	K&W	CONS.	PONDERADO PROMEDIO	DES. EST.	NO PONDERADO PROMEDIO	DES. EST.
H	1	Glu	108	11	90	27
H	2	Val	105	5	106	23
H	3	Gln	114	6	116	14
H	4	Leu	123	3	115	13

FIG.13A

HIDROFOBICIDAD

CADENA	K&W	CONS.	PONDERADO PROMEDIO	DES. EST.	NO PONDERADO PROMEDIO	DES. EST.
H	1	Glu	-0.77	0.24	-0.37	0.72
H	2	Val	1.08	0.20	0.59	0.84
H	3	Gln	-0.78	0.33	-0.42	0.89
H	4	Leu	1.10	0.00	1.10	0.00

FIG.13B

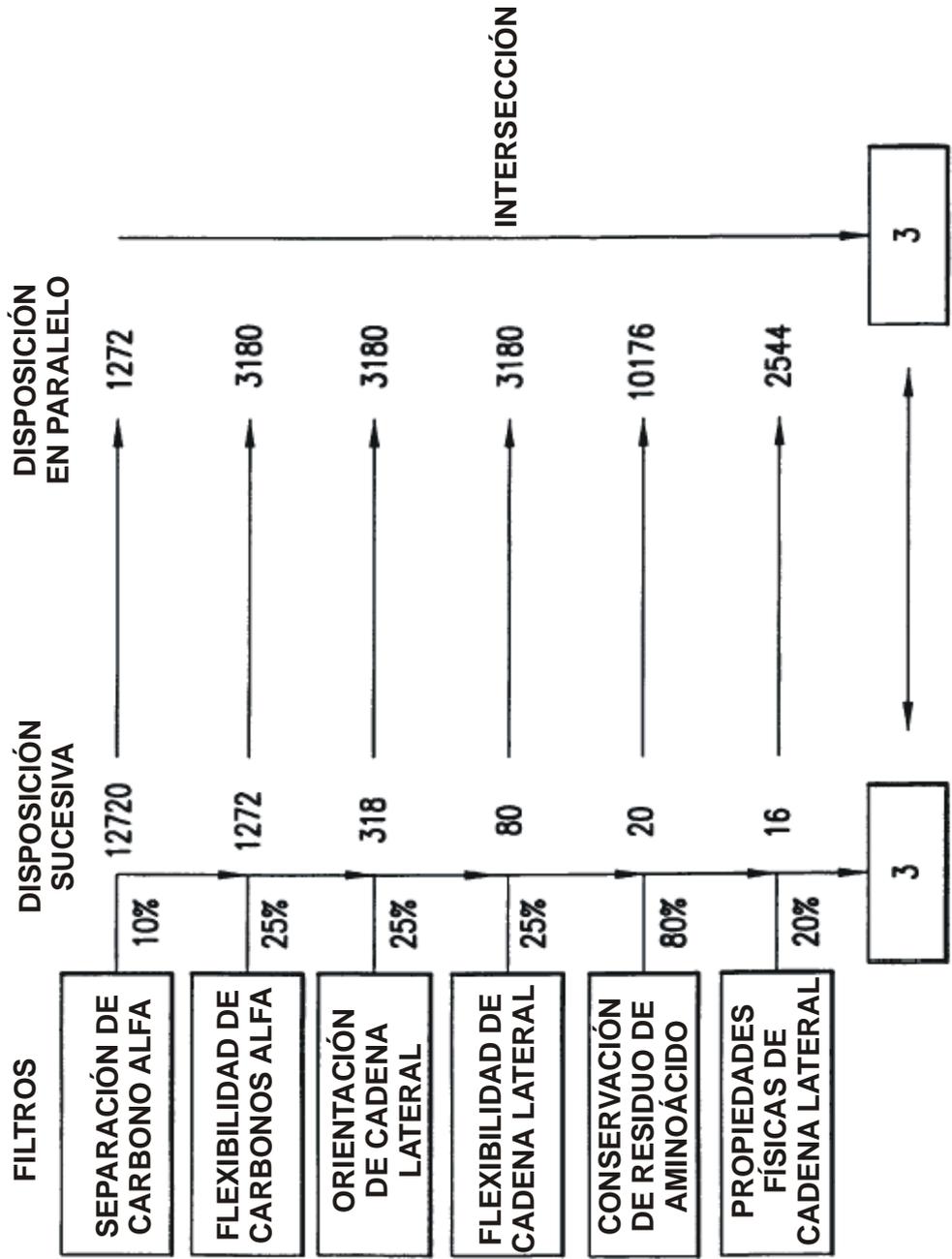


FIG.14