

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 691 642**

51 Int. Cl.:

C07K 14/00 (2006.01)

C07K 7/00 (2006.01)

A61K 38/00 (2006.01)

A61K 47/42 (2007.01)

A61K 39/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.05.2011 PCT/EP2011/058307**

87 Fecha y número de publicación internacional: **24.11.2011 WO11144756**

96 Fecha de presentación y número de la solicitud europea: **20.05.2011 E 11725656 (0)**

97 Fecha y número de publicación de la concesión europea: **08.08.2018 EP 2571510**

54 Título: **Polipéptidos de enrollamiento al azar de prolina/alanina biosintéticos y sus usos**

30 Prioridad:

29.12.2010 US 201061428016 P

21.05.2010 EP 10163564

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

28.11.2018

73 Titular/es:

XL-PROTEIN GMBH (50.0%)

Lise-Meitner-Strasse 30

85354 Freising, DE y

TECHNISCHE UNIVERSITÄT MÜNCHEN (50.0%)

72 Inventor/es:

SKERRA, ARNE;

BINDER, ULI y

SCHLAPSCHY, MARTIN

74 Agente/Representante:

ARIAS SANZ, Juan

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 691 642 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

Polipéptidos de enrollamiento al azar de prolina/alanina biosintéticos y sus usos

La presente invención se refiere a un polipéptido de enrollamiento al azar biosintético o a un segmento de polipéptido de enrollamiento al azar biosintético o a un conjugado, comprendiendo dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido de enrollamiento al azar biosintético o conjugado una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en los que dicha secuencia de aminoácidos consiste en al menos aproximadamente 50 residuos de aminoácido de prolina (Pro) y alanina (Ala). Dichos al menos aproximadamente 50 residuos de aminoácido de prolina (Pro) y alanina (Ala) pueden ser (a) constituyente(s) de un polipéptido heterólogo o un constructo de polipéptido heterólogo. También se describen usos y métodos de uso de estos polipéptidos de enrollamiento al azar biosintéticos, dichos segmentos de polipéptido o dichos conjugados. Los usos pueden comprender, entre otros, usos médicos, usos de diagnóstico o usos en la industria alimentaria así como otras aplicaciones industriales, tales como en la industria del papel, en la recuperación de petróleo y similares. La presente invención también se refiere a (a) uso(s) específico(s) del polipéptido de enrollamiento al azar biosintético o segmento de polipéptido de enrollamiento al azar biosintético o conjugados proporcionados en el presente documento, comprendiendo dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido de enrollamiento al azar biosintético o conjugados una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina. La secuencia de aminoácidos del polipéptido de enrollamiento al azar biosintético o segmento de polipéptido de enrollamiento al azar biosintético proporcionado en el presente documento consiste en al menos aproximadamente 50, en al menos aproximadamente 100, en al menos aproximadamente 150, en al menos aproximadamente 200, en al menos aproximadamente 250, en al menos aproximadamente 300, en al menos aproximadamente 350 o en al menos aproximadamente 400 residuos de aminoácido de prolina (Pro) y alanina (Ala). Dichos al menos aproximadamente 50, al menos aproximadamente 100, al menos aproximadamente 150, al menos aproximadamente 200, al menos aproximadamente 250, al menos aproximadamente 300, al menos aproximadamente 350 o al menos aproximadamente 400 residuos de aminoácido de prolina (Pro) y alanina (Ala) son un constituyente de un conjugado de fármaco, tal como un conjugado con un compuesto biológicamente activo. En particular, en el presente documento se proporcionan proteínas heterólogas mediante lo cual estas proteínas comprenden al menos dos dominios, en las que un primer dominio de dichos al menos dos dominios comprende una secuencia de aminoácidos que tiene y/o que media en una actividad, tal como una actividad biológica, y un segundo dominio de dichos al menos dos dominios que comprenden el polipéptido de prolina/alanina de enrollamiento al azar biosintético o segmento de polipéptido de prolina/alanina de la presente invención. La presente invención se refiere en particular a un conjugado de fármaco que comprende (i) un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala), y (ii) un fármaco seleccionado del grupo que consiste en (a) una proteína biológicamente activa o un polipéptido que comprende o que es una secuencia de aminoácidos que tiene o media en una actividad biológica y (b) un fármaco de molécula pequeña. Un objeto adicional de la presente invención es un conjugado de fármaco que comprende el polipéptido de prolina/alanina de enrollamiento al azar biosintético o segmento de polipéptido de prolina/alanina según se proporcionan en el presente documento y, adicionalmente, (a) molécula(s) pequeña(s) farmacéuticamente útil(es) o proteína(s) unidas y/o acopladas a dicho polipéptido de prolina/alanina de enrollamiento al azar biosintético o segmento de polipéptido de prolina/alanina. Además, se dan a conocer moléculas de ácido nucleico que codifican para el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido y/o las proteínas heterólogas biológicamente activas así como vectores y células que comprenden dichas moléculas de ácido nucleico. Además, se dan a conocer métodos para la producción de los polipéptidos de enrollamiento al azar biosintéticos o segmentos de polipéptido de la invención descritos en el presente documento y conjugados de fármaco o alimento correspondientes, es decir, conjugados que comprenden los polipéptidos de enrollamiento al azar biosintéticos o segmentos de polipéptido definidos en el presente documento y un ingrediente alimenticio o un aditivo alimenticio. También se dan a conocer conjugados correspondientes (que comprenden como constituyente el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido dado a conocer en el presente documento) que comprenden, entre otras cosas, un componente o aditivo cosmético o un compuesto biológica o espectroscópicamente activo. Además, la presente invención proporciona composiciones que comprenden los conjugados de fármaco o las moléculas de ácido nucleico, vectores y células de la invención. También se proporcionan métodos de producción y/u obtención de los polipéptidos de enrollamiento al azar biosintéticos o segmentos de polipéptido de la invención así como de producción y/u obtención de las proteínas heterólogas biológicamente activas y/o constructos de polipéptido o conjugados de fármaco de la invención. Además, en el presente documento se proporcionan usos médicos, farmacéuticos así como de diagnóstico para el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina (o para moléculas y conjugados que comprenden los mismos) tal como se define en el presente documento. Tal uso médico o farmacéutico puede comprender el uso de dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido como expansor del plasma y similares. Sin embargo, los medios y métodos proporcionados en el presente documento no se limitan a usos farmacéuticos, médicos y biológicos, sino que también pueden emplearse en otros sectores industriales, tales como la industria del papel, en la recuperación de petróleo, etc. Un aclaramiento rápido de la circulación sanguínea mediante filtración renal es una propiedad típica de moléculas pequeñas (incluyendo proteínas

pequeñas y péptidos). Sin embargo, expandiendo las dimensiones moleculares aparentes más allá del tamaño de poro de los glomérulos renales puede prolongarse la semivida en plasma de proteínas terapéuticas hasta un intervalo médicamente útil de varios días. Una estrategia para lograr un efecto de este tipo es la conjugación química del componente biológico con el polímero sintético polietilenglicol (PEG). Esto ha conducido a varios fármacos aprobados, por ejemplo PEG-interferón alfa2a (Pegasys®), PEG-G-CSF (Neulasta®) y, recientemente, un alfaTNF-fragmento Fab pegilado (Cimzia®). No obstante, la tecnología de la “pegilación” tiene varios inconvenientes: los derivados de PEG de calidad clínica son caros y su acoplamiento covalente a una proteína recombinante requiere etapas de procesamiento y purificación posteriores adicionales, reduciendo por tanto el rendimiento y aumentando los costes. Además, el PEG no es biodegradable, lo cual puede provocar efectos secundarios tales como vacuolización del epitelio renal tras el tratamiento continuo; véase, por ejemplo, Gaberc-Porekar (2008) *Curr Opin Drug Discov Devel* 11:242-50; Knop (2010) *Angew Chem Int Ed Engl* 49:6288-308 o Armstrong en: Veronese (Ed.), “PEGylated Protein Drugs: Basic Science and Clinical Applications”; Birkhäuser Verlag, Basilea 2009.

Con el fin de superar algunos de los inconvenientes de la tecnología de PEG, se han proporcionado determinados compuestos miméticos de polipéptidos recombinantes en la técnica, algunos de los cuales se basan en secuencias de aminoácidos que se producen de manera natural o tramos de aminoácido sintéticos.

La mayoría de las secuencias de aminoácidos naturales no se comportan como una cadena al azar ideal en disolución fisiológica porque o bien tienden a adoptar una conformación plegada (estructura secundaria) o bien, si están desplegadas, habitualmente son insolubles y forman agregados. De hecho, la mayoría de los experimentos clásicos para investigar el comportamiento de cadenas al azar de polipéptidos se llevaron a cabo en condiciones desnaturizantes, es decir en presencia de desnaturizantes químicos tales como urea o cloruro de guanidinio (véase, por ejemplo, Cantor (1980) *Biophysical Chemistry*. W.H. Freeman and Company, Nueva York). Por tanto, tales tecnologías se basan generalmente en secuencias de aminoácidos peculiares que resisten al plegamiento, la agregación así como la adsorción inespecífica y, por tanto, proporcionan cadenas al azar estables en condiciones y temperatura de tampón fisiológico aunque se fusionen genéticamente con un dominio de proteína terapéutica plegado. En estas circunstancias, tales compuestos miméticos de PEG recombinantes pueden conferir un aumento de tamaño mucho mayor de lo que se esperaría normalmente basándose en su masa molecular sola, retardando eventualmente la filtración renal y prolongando eficazmente la semivida en plasma del componente unido en factores considerables.

Sin embargo, muchas de estas tecnologías tienen desventajas y advertencias adicionales.

Por ejemplo, se han sometido a prueba secuencias de aminoácidos repetitivas que se producen de manera natural para determinar su utilidad en las ciencias médicas y en enfoques farmacéuticos. Uno de estos enfoques se refiere a la trans-sialidasa de *Trypanosoma cruzi*. Contiene un dominio catalítico de 680 residuos de aminoácido seguido por un dominio repetitivo C-terminal, denominado “antígeno secretado de fase aguda” (SAPA), que comprende un número variable de repeticiones de aminoácidos de 12 meros. Estudios de farmacocinética (PK) en ratones de la trans-sialidasa que contiene 13 repeticiones de aminoácidos correspondientes hidrófilas y (a pH fisiológico) cargadas negativamente que tienen la secuencia natural DSSAHSTPSTPA revelaron una semivida en plasma cinco veces más larga en comparación con la enzima recombinante de la que se había delecionado la secuencia repetitiva C-terminal (Buscaglia (1999) *Blood* 93 :2025-32). Se observó un efecto de prolongación de la semivida similar tras la fusión de la misma trans-sialidasa, es decir su dominio catalítico de 76 kDa, con 13 repeticiones de aminoácidos cargadas de la secuencia EPKSA que se encontraron en el antígeno proteico 13 de *Trypanosoma cruzi*. Tanto las repeticiones de SAPA como del antígeno 13 pudieron prolongar la semivida en plasma de la proteína heteróloga glutatión S-transferasa (GST) de *Schistosoma japonicum* en un factor de 7-8 tras la fusión genética a ambos extremos C-terminales de esta enzima homodimérica (véase Buscaglia, citado anteriormente). Sin embargo, aunque estas secuencias de aminoácidos repetitivas que se producen de manera natural de patógenos humanos en principio pueden parecer atractivas para optimizar la farmacocinética de proteínas terapéuticas, se encontró que eran altamente inmunogénicas (véase Affranchino (1989), *Mol Biochem Parasitol* 34:221-8 o Buscaglia (1998), *J Infect Dis* 1998; 177:431-6).

Otro enfoque se refiere al uso de gelatina. La gelatina, colágeno animal hidrolizado y desnaturizado, contiene largos tramos de repeticiones de Gly-Xaa-Yaa, en los que Xaa y Yaa constituyen principalmente prolina y 4-hidroxiprolina, respectivamente. La succinilación de gelatina, principalmente mediante los grupos ϵ -amino de cadenas laterales de lisina intercaladas de manera natural, aumenta la hidrofilia de este biopolímero y reduce su punto isoeléctrico (pI). La repulsión electrostática intramolecular entre los grupos carboxilato cargados negativamente de las cadenas laterales modificadas extiende supuestamente la molécula para dar una conformación más o menos extendida. El volumen expandido resultante hace que la gelatina succinilada sea una macromolécula para su uso como expansor del plasma en seres humanos y se comercializa, entre otras cosas, como Volplex® (Beacon Pharmaceuticals Ltd) o Gelofusine® (B. Braun Melsungen AG). Además, se logró un efecto de prolongación de la semivida mediante fusión genética de factor estimulante de colonias de granulocitos (G-CSF) a un polipéptido similar a gelatina artificial (Huang (2010) *Eur J Pharm Biopharm* 74:435-41). Para ello, se intercambiaron todas las cadenas laterales hidrófobas en una gelatina natural por residuos hidrófilos, dando como resultado una proteína similar a gelatina de 116 aminoácidos (GLK) que comprendía los aminoácidos G, P, E, Q, N, S y K en un orden variable. Se fusionó G-CSF en su extremo N-terminal con 4 copias de esta secuencia de GLK y se secretó en *Pichia pastoris*. *Pichia pastoris* pareció ser un organismo de producción favorable para proteínas de

fusión de GLK; sin embargo, queda por determinar si también pueden producirse GLK en otros organismos ya que se conoce que fragmentos de gelatina recombinantes sólo pueden expresarse con bajo rendimiento en *E. coli*, por ejemplo, tal como se ilustra en Olsen (2003), *Adv Drug Deliv Rev* 55:1547-67.

5 La elastina es un componente de la matriz extracelular en muchos tejidos. Se forma a partir del precursor soluble tropoelastina, que consiste en un dominio hidrófilo rico en Lys/Ala y un dominio elastomérico hidrófobo con
 10 secuencia repetitiva. La reticulación enzimática de cadenas laterales de lisina dentro del dominio hidrófilo conduce a la formación de elastina insoluble. Los polipéptidos similares a elastina (ELP) son secuencias de aminoácidos repetitivas, diseñadas artificialmente, derivadas del dominio hidrófobo de tropoelastina. El motivo de secuencia de
 15 repetición más común de ELP es V-P-G-X-G, en el que "X" puede ser cualquier aminoácido excepto Pro (MacEwan (2010) *Biopolymers* 94:60-77; Kim (2010) *Adv Drug Deliv Rev* 62:1468-78). Pueden fusionarse ELP adecuados con proteínas terapéuticas y producirse en *E. coli*. Por consiguiente, la capacidad de los ELP de formar depósitos
 similares a gel tras la inyección puede prolongar significativamente la semivida *in vivo* de un compuesto biológico unido, aunque mediante un mecanismo diferente de los demás polipéptidos no estructurados. Sin embargo, la unión
 a ELP puede obstaculizar la bioactividad de la pareja de fusión tal como se demuestra para el antagonista de receptor de interleucina-1 en un bioensayo de proliferación de linfocitos inducida por IL-1 (Shamji (2007) *Arthritis Rheum.* 11:3650- 3661). Además, los ELP son propensos a degradación mediante proteasas endógenas tales como
 colagenasa. Además, las proteínas agregadas son generalmente más susceptibles a la inmunogenicidad.

Enfoques adicionales se refieren al uso de polímeros polianiónicos. Por ejemplo, se ha acoplado químicamente
 20 poliglutamato (PG) a fármacos de molécula pequeña citotóxicos escasamente solubles para el tratamiento de cáncer. Un producto correspondiente es Opaxio™, un conjugado de fármaco de paclitaxel actualmente en estudios clínicos de fase III. La semivida de un conjugado de paclitaxel-PG se prolongó en un factor de 3 a 14 en
 comparación con el compuesto sin modificar (Singer (2005) *J Control Release* 109:120-6). Se produjeron proteínas de fusión adicionales, por ejemplo G-CSF fusionado en su extremo N-terminal con un tramo de 175 residuos de Glu consecutivos o IFN-alfa2 que porta en su extremo C-terminal una cola de PG de 84 residuos, en un estado soluble
 25 en el citoplasma de *E. coli* (véase el documento WO2002/077036). Para una traducción eficaz, la fusión N-terminal requirió un péptido líder, que posteriormente se eliminó mediante escisión con proteasa de virus de grabado del tabaco (TEV). Las fusiones con poliglutamato de G-CSF e INF α 2 mostraron bioactividad en ensayos de cultivo celular. Sin embargo, hasta la fecha no se han notificado datos farmacocinéticos de estas fusiones con PG. Además,
 la carga altamente negativa de las fusiones con PG es una desventaja general con respecto a interacciones biomoleculares (por ejemplo, unión del factor soluble o receptor diana) debido a efectos de atracción o repulsión electrostáticas artificiales.

El documento WO 2006/081249 describe una secuencia de polipéptido con aproximadamente de 2 a 500 unidades
 de repetición de 3 a 6 aminoácidos, en la que G, N o Q representan los principales constituyentes mientras que
 35 constituyentes minoritarios pueden ser A, S, T, D o E. Esta composición de aminoácidos permite la integración del *sequon* de glicosilación Asn-Xaa-Ser/Thr (en el que Xaa es cualquier aminoácido excepto Pro) para la glicosilación unida en N de la cadena lateral de Asn en sistemas de expresión eucariotas. El tamaño macromolecular aumentado de una proteína de fusión resultante, incluyendo la modificación tras la traducción con estructuras de hidratos de carbono solvatadas voluminosas, puede prolongar la farmacocinética de la proteína genéticamente conjugada. Tales uniones de oligosacáridos ("glicoingeniería") en general pueden tanto reducir la propensión a proteólisis como
 40 aumentar el volumen hidrodinámico (Sinclair (2005) *J Pharm Sci* 94:1626-35). Una desventaja es la heterogeneidad molecular intrínseca de la biomacromolécula glicosilada, lo que provoca un esfuerzo adicional durante la producción por biotecnología y el control de calidad.

El documento WO 2010/091122 (y el documento WO 2007/103515) y Schellenberger (2009) *Nat Biotechnol* 27:1186-90 dan a conocer polímeros de aminoácidos no repetitivos, no estructurados, que abarcan y que
 45 comprenden los residuos P, E, S, T, A y G. Este conjunto de aminoácidos, que muestra una composición similar a la repetición de PSTAD descrita adicionalmente con anterioridad, se examinó sistemáticamente para determinar secuencias para producir un polipéptido solvatado con gran tamaño molecular, adecuado para el desarrollo biofarmacéutico, evitando cadenas laterales hidrófobas (en particular F, I, L, M, V y W) que pueden dar lugar a agregación y pueden provocar una respuesta inmunitaria mediada por HLA/MHC-II. Además, se excluyeron residuos
 50 de Cys posiblemente reticulantes, los aminoácidos catiónicos K, R y H, que podían interactuar con membranas celulares cargadas negativamente, y las cadenas laterales de amida de N y Q, que eran posiblemente propensas a hidrólisis (véase Schellenberger (2009), citado anteriormente). Se examinaron bibliotecas génicas sintéticas que codificaban para secuencias no repetitivas que comprendían el conjunto de residuos de PESTAG, que se fusionaron con la proteína verde fluorescente (GFP), con respecto a niveles de expresión solubles en *E. coli*, y se investigó
 55 adicionalmente un subconjunto resultante para determinar la estabilidad genética, solubilidad de proteína, termoestabilidad, tendencia a la agregación y perfil de contaminantes. Finalmente, se sometió a pruebas adicionales una secuencia de 864 aminoácidos que contenía 216 residuos de Ser (el 25,0% en moles), 72 residuos de Ala (el 8,3% en moles) y 144 aminoácidos (el 16,7% en moles) de cada uno de Pro, Thr, Glu y Gly, para determinar su fusión al agonista de receptor de GLP-1 exendina-4 (E-XTEN) y algunos otros compuestos biológicos. Las proteínas de fusión (que portaban normalmente un dominio de unión a celulosa, que posteriormente se eliminó por escisión) se produjeron en un estado soluble en el citoplasma de *E. coli* y se aislaron. La investigación de E-XTEN mediante espectroscopía de dicroísmo circular (CD) reveló una ausencia de estructura secundaria mientras que durante la
 60

cromatografía de exclusión molecular (SEC) la proteína de fusión mostró sustancialmente menos retención de la esperada para una proteína de 84 kDa, demostrando por tanto un volumen hidrodinámico aumentado (Schellenberg (2009), citado anteriormente). La estructura desordenada del polipéptido PESTAG y el aumento asociado del radio hidrodinámico pueden verse favorecidos por la repulsión electrostática entre aminoácidos que portan una alta carga negativa neta que se distribuyen a lo largo de la secuencia de XTEN (véase el documento WO 2010/091122). Sin embargo, un estudio adicional, Geeting (2010) PLoS One 2010; 5:e10175, demostró que XTEN disminuye la potencia de su pareja de fusión terapéutica. En un ensayo de cultivo celular, una fusión de glucagón-XTEN mostró una bioactividad de simplemente el 15% del péptido no modificado. Se describió una pérdida aún mayor de afinidad de receptor (CE_{50} aumentada 17 veces) para una fusión con XTEN de hormona del crecimiento humana (hGH); véase el documento WO 2010/144502. También se ha considerado glicina, el aminoácido más pequeño y estructuralmente más sencillo, como el conformacionalmente más flexible basándose en razonamientos teóricos; véase, por ejemplo, Schulz GE, Schirmer RH. Principles of Protein Structure. Springer, Nueva York 1979. Además, simulaciones informáticas han indicado que polímeros de Gly carecen de estructura secundaria y es probable que formen un enrollamiento al azar en disolución; véase Shental-Bechor (2005) Biophys J 88:2391-402. Desde un punto de vista químico, la poliglicina es una poliamida no ramificada, lineal, que muestra cierto parecido con el poliéter PEG en la medida en que ambos son macromoléculas esencialmente unidimensionales con muchos grados de libertad rotacionales a lo largo de la cadena, que están compuestas por unidades de hidrocarburos cortas repetidas que están interrumpidas de manera regular por enlaces de hidrógeno y grupos polares altamente solvatados. Por consiguiente, la poliglicina debe constituir el compuesto mimético de PEG genéticamente codificable más sencillo con la perspectiva de prolongar la semivida en plasma de proteínas terapéuticas. Este concepto se empleó en forma de "polímero de homocisteína (HAP)" o como secuencia rica en glicina (GRS), respectivamente; véase, Schlapschy (2007) Protein Eng Des Sel 20:273-84; documento WO 2007/103515. Sin embargo, se ha sabido desde hace mucho que los polímeros de Gly puros químicamente sintetizados muestran una escasa solubilidad en agua; véase, entre otros, en Bamford CH *et al.* Synthetic Polypeptides - Preparation, Structure, and Properties. Academic Press, Nueva York 1956. Por tanto, se realizaron diferentes intentos por aumentar la hidrofilia, o bien introduciendo cadenas laterales de alcohol de serina formadoras de enlaces de hidrógeno (documento WO 2007/103515 así como Schlapschy (2007), citado anteriormente) o bien, adicionalmente, residuos de glutamato cargados negativamente (documento WO 2007/103515). Debe observarse que ya se han descrito espaciadores peptídicos con la composición $(Gly_4Ser)_n$ en la técnica con el fin de unir dominios en proteínas de fusión de una manera flexible. Se detectó un volumen hidrodinámico significativamente aumentado para estas proteínas de fusión en SEC analítica. En el caso de la versión de HAP de 200 residuos, el aumento de tamaño aparente fue del 120% en comparación con el fragmento Fab sin fusionar, mientras que la masa auténtica sólo fue mayor en un 29%, revelando por tanto el efecto de un volumen hidrodinámico potenciado debido a la estructura de enrollamiento al azar solvatada de la cola de poliglicina. Además, los espectros de diferencia de CD fueron característicos para la estructura secundaria desordenada para el resto de HAP. Finalmente, la semivida en plasma terminal del fragmento Fab que portaba el HAP de 200 residuos en ratones se prolongó en un factor de aproximadamente 3. Aunque moderado, este efecto puede resultar apropiado para determinadas aplicaciones de diagnóstico (especializadas), tales como obtención de imágenes *in vivo*; véase Schlapschy (2007); citado anteriormente. Desafortunadamente, la producción de proteínas de fusión con secuencias de repetición de $(Gly_4Ser)_n$ más largas pareció ser menos viable debido a un aumento de la tendencia a formar agregados, planteando por tanto una limitación natural al uso de polímeros de glicina (más o menos puros) como compuestos miméticos de PEG.

El documento WO 2008/155134 da a conocer que secuencias con una mezcla apropiada de residuos de Pro, Ala y Ser (es decir PAS) conducen a una cancelación mutua de sus preferencias de estructura secundaria diferenciadas y, por tanto, dan como resultado un polipéptido desordenado de manera estable. Sin embargo, el documento WO 2008/155134 también documenta que las proteínas de fusión con un dominio compuesto únicamente por residuos de serina y alanina (SA), es decir un dominio que comprende sólo dos tipos de aminoácidos, no forman un enrollamiento al azar, sino en su lugar una estructura de lámina β . El documento US 2009/0192087 da a conocer una secuencia de poli-Ala que sirve como grupo de unión para fusionar al menos un subdominio extracelular de la proteína Klotho con un polipéptido que comprende una hormona del crecimiento de fibroblastos. El documento EP-A1 1 739 167 da a conocer un método para la generación de virus y su uso en la preparación de vacunas. El documento EP-A1 1 739 167 también da a conocer una secuencia de poli-A de un polipéptido que puede usarse para preparar virus. Sin embargo, estas secuencias de poli-A no consisten únicamente en residuos de prolina y alanina.

El documento WO 2008/14500 da a conocer proteínas de parvovirus con determinadas inserciones y su uso médico. Sin embargo, el documento WO 2008/14500 no da a conocer una proteína que tiene una secuencia de aminoácidos que consiste únicamente en al menos 50 residuos de prolina y alanina ni un conjugado de fármaco que comprende la misma.

El documento WO 01/78503 da a conocer proteínas de fusión con GFP para su expresión en células vegetales, en las que las proteínas de fusión pueden comprender una secuencia de (hidroxi)prolina-alanina. Sin embargo, GFP no es un fármaco.

La síntesis química de polipéptidos se conoce bien y se ha descrito en la técnica. Izuka da a conocer la síntesis química de polipéptidos que contienen prolina (véase Izuka (1993), Bull. Chem. Soc. Jpn 66, 1269-1272). Estos

5 copolipéptidos contienen secuencias al azar de prolina y o glicina, L-alanina, ácido L- α -aminobutírico (Abu), L-norvalina (Nva) o L-leucina, respectivamente, y se sintetizan mediante copolimerización química. Izuka da a conocer que tales copolipéptidos tienen principalmente una conformación similar a colágeno definida. Además, en esta publicación se describe que los copolipéptidos de prolina y alanina (o prolina y ácido L- α -aminobutírico) son
parcialmente solubles en agua, mientras que otros copolipéptidos eran completamente insolubles. En Izuka se
especula que los copolipéptidos de prolina/alanina pueden tener una conformación desordenada parcial. Izuka
enfatisa que los polipéptidos químicamente sintetizados con una secuencia de prolina/alanina al azar se producen
de manera predominante en una conformación similar a colágeno, es decir en una conformación estructurada.

10 Por tanto, el problema técnico subyacente a la presente invención es proporcionar polipéptidos grandes con una auténtica conformación de enrollamiento al azar. El problema técnico se resuelve proporcionando las realizaciones caracterizadas en las reivindicaciones y tal como se proporciona en el presente documento.

15 Por consiguiente, la invención se refiere a un conjugado de fármaco que comprende (i) un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala), y (ii) un fármaco seleccionado del grupo que consiste en (a) una proteína biológicamente activa o un polipéptido que comprende o que es una
20 secuencia de aminoácidos que tiene o media en una actividad biológica y (b) un fármaco de molécula pequeña. Los polipéptidos con auténtica conformación de enrollamiento al azar y segmentos de polipéptido con auténtica conformación de enrollamiento al azar tal como se proporciona en el presente documento también son útiles en el contexto de usos cosméticos así como usos en la industria alimentaria y la producción de bebidas. Los polipéptidos grandes proporcionados en el presente documento que muestran auténtica conformación de enrollamiento al azar consisten única o simplemente en residuos de prolina (P, Pro) y alanina (A, Ala) y comprenden más de al menos 50 aminoácidos, en particular de al menos aproximadamente 100, en particular de al menos aproximadamente 150,
25 en particular de al menos aproximadamente 200, en particular de al menos aproximadamente 250, en particular de al menos aproximadamente 300, en particular de al menos aproximadamente 350, en particular de al menos aproximadamente 400 residuos de aminoácido de prolina y alanina. Ambos aminoácidos, P y A, necesitan estar presentes en los polipéptidos grandes con auténtica conformación de enrollamiento al azar y segmentos de polipéptido con auténtica conformación de enrollamiento al azar proporcionados el presente documento. En el presente documento también se proporcionan moléculas de ácido nucleico que codifican para los polipéptidos de
30 enrollamiento al azar biosintéticos o segmentos de polipéptido dados a conocer en el presente documento así como para conjugados de fármaco o alimento que comprenden dichos polipéptidos de enrollamiento al azar biosintéticos o segmentos de polipéptido y una proteína de interés (covalentemente unida), tal como una proteína biológicamente activa.

35 El polipéptido de enrollamiento al azar biosintético o segmento de polipéptido de enrollamiento al azar biosintético tal como se describe en el presente documento y que va a usarse en conjugados de fármaco o alimento tal como se proporciona en el presente documento y que comprende una secuencia de aminoácidos que consiste en al menos aproximadamente 50, en al menos aproximadamente 100, en al menos aproximadamente 150, en al menos aproximadamente 200, en al menos aproximadamente 250, en al menos aproximadamente 300, en al menos aproximadamente 350, en al menos aproximadamente 400 residuos de aminoácido de prolina (P) y alanina (A) va a usarse, entre otras cosas, en un contexto heterólogo, es decir en una proteína heteróloga biológicamente activa,
40 constructo de proteína y/o en un conjugado de fármaco que comprende dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido y moléculas farmacéutica o medicamento útiles, tales como moléculas pequeñas, péptidos o biomacromoléculas tales como proteínas, ácidos nucleicos, hidratos de carbono, vesículas lipídicas y similares. Tal como se ilustra en los ejemplos adjuntos, los inventores pudieron proporcionar satisfactoriamente conjugados de fármaco que consisten en los polipéptidos de enrollamiento al azar auténtico tal como se define en el presente documento y proteínas biológicamente activas o tramos de proteína así como conjugados de fármaco que consisten en moléculas pequeñas o fármacos de molécula pequeña que comprenden y/o están unidos a los polipéptidos de enrollamiento al azar descritos en el presente documento, que consisten únicamente en residuos de aminoácido de prolina y alanina (es decir en aminoácidos tanto P como A).

50 Por consiguiente, la presente invención proporciona, entre otras cosas, una proteína heteróloga biológicamente activa que comprende al menos dos dominios en la que (a) un primer dominio de dichos al menos dos dominios comprende una secuencia de aminoácidos que tiene y/o que media en dicha actividad biológica; y (b) un segundo dominio de dichos al menos dos dominios comprende el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que consiste en una secuencia de aminoácidos que consiste en al menos aproximadamente 50, en al menos aproximadamente 100, en al menos aproximadamente 150, en al menos aproximadamente 200, en al menos aproximadamente 250, en al menos aproximadamente 300, en al menos aproximadamente 350, en al menos aproximadamente 400 residuos de aminoácido de prolina y alanina. Según esta invención, dicho "primer dominio" y dicho "segundo dominio" no están comprendidos ni en una proteína natural (es decir, que se produce en la naturaleza) ni en una proteína hipotética según se deduce a partir de secuencias de ácido nucleico codificantes que se producen de manera natural, tales como marcos de lectura abiertos, etc.

60 Además, esta invención proporciona un conjugado de fármaco que consiste en el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste en al

menos aproximadamente 50, en al menos aproximadamente 100, en al menos aproximadamente 150, en al menos aproximadamente 200, en al menos aproximadamente 250, en al menos aproximadamente 300, en al menos aproximadamente 350, en al menos aproximadamente 400 residuos de aminoácido de prolina y alanina y (a) molécula(s) farmacéuticamente útil(es), tal(es) como una(s) molécula(s) pequeña(s) o una(s) proteína(s) que se conjuga(n) con dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido. De nuevo, debe observarse que el término "biológicamente activo" en el contexto de los conjugados dados a conocer en el presente documento no se limita a moléculas biológicas puras sino que también comprende moléculas médicamente activas, terapéuticamente activas, farmacéuticamente activas y similares. Resulta evidente para el experto en la técnica que los medios y métodos proporcionados en el presente documento no se limitan a usos farmacéuticos y médicos, sino que pueden emplearse en una amplia variedad de tecnologías, incluyendo, pero sin limitarse a, tecnologías cosméticas, alimentarias, de bebidas y nutrición, industria del petróleo, industria del papel y similares.

A diferencia de copolipéptidos químicamente sintetizados (tal como en Izuka, citado anteriormente), los polipéptidos de enrollamiento al azar proporcionados en el presente documento se producen de manera biosintética. El término "biosintético" tal como se usa en el presente documento se refiere a la síntesis por medio de medios biotecnológicos (a diferencia de la síntesis química). Tales métodos biotecnológicos se conocen bien en la técnica y también se describen en el presente documento a continuación. La biosíntesis de los polipéptidos de enrollamiento al azar de la presente invención permite la producción de polipéptidos con una secuencia de residuos de prolina y alanina definida, una longitud definida y/o una razón definida de residuos de prolina y alanina. Además, los polipéptidos proporcionados según la presente invención son sustancialmente puros, es decir, los polipéptidos producidos son esencialmente uniformes y comparten las características anteriores (es decir secuencia definida, longitud definida y/o razón de aminoácidos definida). Los polipéptidos de enrollamiento al azar que consisten en al menos aproximadamente 50, en particular en al menos aproximadamente 100, en particular en al menos aproximadamente 150, en particular en al menos aproximadamente 200, en particular en al menos aproximadamente 250, en particular en al menos aproximadamente 300, en particular en al menos aproximadamente 350, en particular en al menos aproximadamente 400 residuos de aminoácido de prolina y alanina están comprendidos según esta invención en conjugados de fármaco.

En conjunto, las características anteriores de los polipéptidos de la presente invención permiten la formación de un enrollamiento al azar estable de los polipéptidos y estos polipéptidos de enrollamiento al azar tienen propiedades sorprendentes y ventajosas. Por ejemplo, los polipéptidos de la presente invención son completamente solubles en disolución acuosa y tienen un volumen hidrodinámico aumentado. De manera inesperada, los polipéptidos de enrollamiento al azar tal como se define en el presente documento también pueden conferir una estabilidad *in vivo/in vitro* aumentada. Esto es particularmente importante para aplicaciones médicas, por ejemplo, para proteínas biológicamente activas o conjugados de fármaco que comprenden el polipéptido de enrollamiento al azar de esta invención. Sin embargo, las numerosas propiedades ventajosas de los polipéptidos de enrollamiento al azar de la presente invención no sólo permiten su uso en el campo médico sino también en otros campos, tales como en cosmética/tratamientos cosméticos o en los campos de nutrición y tecnología de los alimentos, tal como en la industria láctea o en el procesamiento de la carne. Ejemplos de conjugados útiles en industria alimentaria y similares son conjugados que comprenden el polipéptido de enrollamiento al azar o segmento de polipéptido dado a conocer en el presente documento que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina y compuestos que son útiles en estas tecnologías, tales como, por ejemplo, polímeros de polioxipropileno o polioxietileno, que son tensioactivos no iónicos usados como emulsionantes. En el presente documento también se prevé el uso del polipéptido de enrollamiento al azar biosintético tal como se define en el presente documento en métodos bioquímicos y en procedimientos técnicos, tales como producción de papel, recuperación de petróleo y similares. Las características sorprendentes y ventajosas de los polipéptidos de enrollamiento al azar biosintéticos que consisten simplemente en residuos de prolina y alanina tal como se proporciona en el presente documento (y también de los conjugados y constructos dados a conocer en el presente documento, tales como conjugados/constructos de fármaco o alimento, que comprenden dichos polipéptidos de enrollamiento al azar auténticos biosintéticos) se describen a continuación con más detalle. Además, a continuación se proporcionan usos y medios y métodos ilustrativos que emplean estos polipéptidos de enrollamiento al azar biosintéticos de la invención. En el presente documento también se proporcionan medios y métodos para la producción de tales polipéptidos de enrollamiento al azar biosintéticos así como polipéptidos heterólogos o constructos de polipéptido biológicamente activos y de los conjugados y constructos dados a conocer en el presente documento, tales como constructos de fármaco, que comprenden dichos polipéptidos de enrollamiento al azar.

En el contexto de esta invención, se ha encontrado sorprendentemente que los polímeros/polipéptidos de prolina-alanina con una composición uniforme forman una conformación de enrollamiento al azar estable. Esto también se demuestra en los ejemplos adjuntos, en los que la estructura de enrollamiento al azar de (co)polímeros/polipéptidos de prolina-alanina biosintéticos se confirma mediante espectroscopía de dicroísmo circular (CD). La obtención y el empleo de tales polipéptidos/polímeros auténticamente de enrollamiento al azar biosintéticos fueron sorprendentes dado que el método de Chou-Fasman establecido (Chou y Fasman (1974), *Biochemistry* 13, 223-245) predice una estructura secundaria en hélices α al 100% de polímeros/polipéptidos (o segmentos de los mismos) compuestos por prolina y alanina, tal como se muestra en la figura 7. Sin embargo, en el presente documento se ha encontrado sorprendentemente y se ha mostrado experimentalmente que los polímeros/polipéptidos de prolina-alanina con una composición uniforme forman una conformación de enrollamiento al azar estable. Esto también se demuestra en los

ejemplos adjuntos, en los que la estructura de enrollamiento al azar de (co)polímeros/polipéptidos de prolina/alanina se confirma mediante técnicas experimentales tales como espectroscopía de dicroísmo circular (CD) y cromatografía de exclusión molecular (SEC).

5 A diferencia de los polipéptidos/polímeros de la presente invención, los polipéptidos químicamente sintetizados descritos, por ejemplo, en Izuka (1993), citado anteriormente, tienen una secuencia arbitraria/no definida y estocástica y una longitud diversa. Por tanto, los polipéptidos químicamente sintetizados comprenden una mezcla de péptidos completamente diferentes con diversas razones de prolina/alanina, longitudes, y así sucesivamente. Tal como se menciona en Izuka, los polipéptidos químicamente sintetizados de una mezcla de este tipo no forman (o sólo forman parcialmente) un enrollamiento al azar y, por consiguiente, no tienen ninguna de las propiedades ventajosas de los polipéptidos biosintéticos proporcionados y descritos a continuación en el presente documento. Por consiguiente, la presente divulgación comprende y se refiere a composiciones que comprenden los polipéptidos/polímeros de enrollamiento al azar biosintéticos de la invención tal como se da a conocer en el presente documento, mediante lo cual dichos polipéptidos/polímeros de enrollamiento al azar biosintéticos se definen, entre otras cosas, por su secuencia que comprende únicamente residuos de prolina y alanina. La invención se define por las reivindicaciones adjuntas. La invención se refiere a conjugados de fármaco que comprenden, como constituyente, estos polipéptidos/polímeros de enrollamiento al azar dados a conocer en el presente documento. Estos polipéptidos/polímeros de enrollamiento al azar biosintéticos de la invención comprendidos en dichas composiciones son, en una realización, de longitud uniforme.

20 Tal como se mencionó anteriormente, los polipéptidos de enrollamiento al azar biosintéticos (o segmentos de polipéptido de enrollamiento al azar) de esta invención que consisten únicamente en residuos de prolina y alanina forman de manera inesperada una conformación de enrollamiento al azar estable. El término "enrollamiento al azar" tal como se usa en el presente documento se refiere de manera general a cualquier conformación de una molécula polimérica, incluyendo polímeros de aminoácidos/secuencias de aminoácidos/polipéptidos, en la que los elementos monoméricos individuales que forman dicha estructura polimérica están orientados esencialmente al azar hacia los elementos monoméricos adyacentes mientras que todavía están químicamente unidos a dichos elementos monoméricos adyacentes. En particular, un polipéptido, secuencia de aminoácidos o polímero de aminoácidos que adopta/tiene/forma una "conformación de enrollamiento al azar" carece sustancialmente de una estructura secundaria y terciaria definida. En el contexto de los polipéptidos de la presente invención, los elementos monoméricos que forman la estructura polimérica (es decir el polipéptido/secuencia de aminoácidos) son o bien aminoácidos individuales tales como prolina y alanina en sí mismos o bien tramos peptídicos tales como las "repeticiones de aminoácidos"/"casetes de aminoácidos"/"repeticiones de casete"/"bloques estructurales"/"módulos" (o fragmentos de los mismos) que se describen y definen adicionalmente a continuación.

35 La naturaleza de enrollamientos al azar de polipéptidos y sus métodos de identificación experimental los conoce el experto en la técnica y se han descrito en la bibliografía científica (Cantor (1980) Biophysical Chemistry, 2ª ed., W. H. Freeman and Company, Nueva York; Creighton (1993) Proteins – Structures and Molecular Properties, 2ª ed., W. H. Freeman and Company, Nueva York; Smith (1996) Fold Des 1:R95-R106). El término "segmento" tal como se usa en el presente documento se refiere a una parte del polipéptido de enrollamiento al azar biosintético definido en el presente documento, mediante el cual una parte de este tipo puede ser una parte interna del polipéptido de enrollamiento al azar biosintético descrito en el presente documento. Un "segmento" de este tipo puede ser, por ejemplo, un polipéptido de enrollamiento al azar biosintético tal como se define en el presente documento en el que uno (o más) aminoácidos se han delecionado, por ejemplo desde el comienzo y/o desde el final del polipéptido de la invención. Además, un "segmento" de este tipo puede usarse como, o formar parte de, una proteína o polipéptido más grande, por ejemplo, de una proteína de fusión con una proteína biológicamente activa. Una "proteína de fusión" de este tipo también será un ejemplo de un polipéptido/proteína/construido de polipéptido biológicamente activo, heterólogo, de la presente invención. El término "heterólogo" tal como se usa en el presente documento se define a continuación en el presente documento.

50 El polipéptido de enrollamiento al azar (o segmento de enrollamiento al azar del mismo), tal como se proporciona en la presente invención y que va a emplearse en el contexto de esta invención, adopta/forma una conformación de enrollamiento al azar, por ejemplo, en disolución acuosa o en condiciones fisiológicas. El término "condiciones fisiológicas" se conoce en la técnica y se refiere a las condiciones en las que las proteínas adoptan habitualmente su conformación plegada nativa. Más específicamente, el término "condiciones fisiológicas" se refiere a los parámetros biofísicos tal como son normalmente válidos para formas de vida superiores y, particularmente, en mamíferos, lo más preferiblemente seres humanos. El término "condiciones fisiológicas" puede referirse a los parámetros bioquímicos y biofísicos tal como se encuentran normalmente en el organismo (en particular en líquidos corporales) de mamíferos y en particular en seres humanos. Dichas "condiciones fisiológicas" pueden referirse a los parámetros correspondientes encontrados en el organismo sano así como los parámetros encontrados en estados patológicos o en pacientes humanos. Por ejemplo, un paciente humano o mamífero enfermo puede tener una condición de temperatura superior, pero "fisiológica" cuando dicho mamífero o dicho ser humano padece fiebre. Con respecto a "condiciones fisiológicas" en las que las proteínas adoptan su conformación/estado nativo, los parámetros más importantes son la temperatura (37°C para el cuerpo humano), pH (7,35 - 7,45 para la sangre humana), osmolaridad (280 - 300 mmol/kg de H₂O), y, si es necesario, contenido en proteínas (66 - 85 g/l de suero). Sin embargo, el experto en la técnica es consciente de que en condiciones fisiológicas estos parámetros pueden variar, por ejemplo la temperatura, pH, osmolaridad y contenido en proteínas pueden ser diferentes en líquidos corporales o tisulares

datos tales como sangre, líquido cefalorraquídeo, líquido peritoneal y linfa (Klinke (2005) Physiologie, 5ª ed., Georg Thieme Verlag, Stuttgart). Por ejemplo, en el líquido cefalorraquídeo la osmolaridad puede ser de aproximadamente 290 mmol/kg de H₂O y la concentración de proteínas puede ser de entre 0,15 g/l y 0,45 g/l, mientras que en la linfa el pH puede ser de aproximadamente 7,4 y el contenido en proteínas puede ser de entre 3 g/l y 5 g/l. Cuando se determina si un polipéptido (o segmento del mismo)/secuencia de aminoácidos forma/adopta una conformación de enrollamiento al azar en condiciones experimentales usando los métodos tal como se describe a continuación en el presente documento, los parámetros biofísicos tales como temperatura, pH, osmolaridad y contenido en proteínas pueden ser diferentes de las condiciones fisiológicas normalmente encontradas *in vivo*. Temperaturas de entre 1°C y 42°C o preferiblemente de 4°C a 25°C pueden considerarse útiles para someter a prueba y/o verificar las propiedades biofísicas y la actividad biológica de una proteína en condiciones fisiológicas *in vitro*.

Se considera que varios tampones, en particular en entornos experimentales (por ejemplo en la determinación de estructuras de proteínas, en particular en mediciones de CD y otros métodos que permiten al experto en la técnica determinar las propiedades estructurales de una proteína/tramo de aminoácidos) o en tampones, disolventes y/o excipientes para composiciones farmacéuticas, representan “disoluciones fisiológicas”/“condiciones fisiológicas” *in vitro*. Ejemplos de tales tampones son, por ejemplo, solución salina tamponada con fosfato (PBS: NaCl 115 mM, KH₂PO₄ 4 mM, Na₂HPO₄ 16 mM, pH 7,4), tampones Tris, tampones acetato, tampones citrato o tampones similares tales como los usados en los ejemplos adjuntos. Generalmente, el pH de un tampón que representa “condiciones de disolución fisiológica” debe encontrarse en un intervalo de desde 6,5 hasta 8,5, preferiblemente en un intervalo de desde 7,0 hasta 8,0, lo más preferiblemente en un intervalo de desde 7,2 hasta 7,7 y la osmolaridad debe encontrarse en un intervalo de desde 10 hasta 1000 mmol/kg de H₂O, más preferiblemente en un intervalo de desde 50 hasta 500 mmol/kg de H₂O y lo más preferiblemente en un intervalo de desde 200 hasta 350 mmol/kg de H₂O. Opcionalmente, el contenido en proteínas de un tampón que representa condiciones de disolución fisiológica puede encontrarse en un intervalo de desde 0 hasta 100 g/l, despreciando la propia proteína con actividad biológica, mediante lo cual pueden usarse proteínas estabilizantes típicas, por ejemplo albúmina sérica humana o bovina.

En el presente documento se ha encontrado que los polipéptidos (o segmentos de los mismos) no sólo forman una conformación de enrollamiento al azar en condiciones fisiológicas sino, más generalmente, en disolución acuosa. El término “disolución acuosa” se conoce bien en la técnica. Una “disolución acuosa” puede ser una disolución con un contenido en agua (H₂O) de al menos aproximadamente el 20%, de al menos aproximadamente el 30%, de al menos aproximadamente el 40%, de al menos aproximadamente el 50%, de al menos aproximadamente el 60%, de al menos aproximadamente el 70%, de al menos aproximadamente el 80% o de al menos aproximadamente el 90% de H₂O (peso/peso). Por consiguiente, el polipéptido (o segmento del mismo) de la presente invención puede formar una conformación de enrollamiento al azar en disolución acuosa, que contiene posiblemente otros disolventes miscibles, o en dispersiones acuosas con un intervalo más amplio de temperaturas, valores de pH, osmolaridades o contenido en proteínas. Esto es particularmente relevante para aplicaciones del polipéptido de enrollamiento al azar (o segmento del mismo) fuera de la terapia médica o diagnóstico *in vivo*, por ejemplo en cosmética, nutrición o tecnología de los alimentos.

Por consiguiente, también se prevé en el contexto de esta invención que la conformación de enrollamiento al azar del polipéptido biosintético de prolina/alanina (o segmento del mismo) de la presente invención se mantiene y/o se usa en el contexto de composiciones farmacéuticas, tales como productos farmacéuticos/biológicos líquidos o composiciones farmacéuticas liofilizadas. Esto es particularmente importante en el contexto de los conjugados de fármaco proporcionados en el presente documento que comprenden, entre otras cosas, el polipéptido de enrollamiento al azar (o segmento de polipéptido) de la invención. Preferiblemente, las “condiciones fisiológicas” van a usarse en sistemas de tampón, disolventes y/o excipientes correspondientes. Sin embargo, por ejemplo en composiciones liofilizadas o secadas (tales como, por ejemplo, composiciones farmacéuticas/productos biológicos), se prevé que la conformación de enrollamiento al azar del polipéptido de enrollamiento al azar (o segmento de polipéptido) proporcionado en el presente documento temporalmente no está presente y/o no puede detectarse. Sin embargo, dicho polipéptido de enrollamiento al azar (o segmento de polipéptido) adoptará/formará de nuevo su enrollamiento al azar tras la reconstitución en tampones/disoluciones/excipientes/disolventes correspondientes o tras la administración al organismo. En la técnica se conocen métodos para determinar si un polipéptido (o segmento del mismo) forma/adopta una conformación de enrollamiento al azar (Cantor (1980), citado anteriormente; Creighton (1993), citado anteriormente; Smith (1996), citado anteriormente). Tales métodos incluyen espectroscopía de dicroísmo circular (CD) tal como se muestra a modo de ejemplo a continuación en el presente documento. La espectroscopía de CD representa un método de espectroscopía de absorción de luz en el que se mide la diferencia en la absorbancia de luz polarizada de manera circular a la derecha y a la izquierda por una sustancia. La estructura secundaria de una proteína puede determinarse mediante espectroscopía de CD usando espectros en el ultravioleta lejano con una longitud de onda de entre aproximadamente 190 y 250 nm. A estas longitudes de onda, pueden analizarse las diferentes estructuras secundarias encontradas habitualmente en polipéptidos, dado que las conformaciones de hélices α , de láminas β paralelas y antiparalelas, y de enrollamiento al azar dan cada una lugar a una forma y magnitud características del espectro de CD. Por consiguiente, usando espectrometría de CD el experto en la técnica puede determinar fácilmente si el polipéptido (o segmento del mismo) forma/adopta una conformación de enrollamiento al azar en disolución acuosa o en condiciones fisiológicas. Otros métodos biofísicos establecidos incluyen espectroscopía de resonancia magnética nuclear (RMN), espectrometría de absorción, espectroscopía de infrarrojos y de Raman, medición del volumen hidrodinámico mediante cromatografía de exclusión molecular,

ultracentrifugación analítica o dispersión de la luz dinámica/estática así como mediciones del coeficiente de fricción o la viscosidad intrínseca (Cantor (1980), citado anteriormente; Creighton (1993), citado anteriormente; Smith (1996), citado anteriormente).

Además de los métodos experimentales anteriores, se han descrito métodos teóricos para la predicción de estructuras secundarias en proteínas. Un ejemplo de un método teórico de este tipo es el método de Chou-Fasman (Chou y Fasman, citado anteriormente) que se basa en un análisis de las frecuencias relativas de cada aminoácido en hélices α , láminas β y giros basándose en estructuras de proteína conocidas resueltas, por ejemplo, con cristalografía de rayos X. Sin embargo, se sabe que la predicción teórica de la estructura secundaria de proteínas no es fiable. Tal como se muestra a modo de ejemplo a continuación en el presente documento, se encontró que secuencias de aminoácidos que se esperaba que adoptasen una estructura secundaria en hélices α según el método de Chou-Fasman formaban un enrollamiento al azar. Por consiguiente, métodos teóricos tales como el algoritmo de Chou-Fasman pueden tener tan sólo un valor predictivo limitado de si un polipéptido dado adopta una conformación de enrollamiento al azar, tal como también se ilustra en los ejemplos y figuras adjuntos. No obstante, la predicción teórica descrita anteriormente es con frecuencia el primer enfoque en la evaluación de una supuesta estructura secundaria de un polipéptido/secuencia de aminoácidos dado. Una predicción teórica de una estructura de enrollamiento al azar también indica con frecuencia que puede merecer la pena verificar mediante los medios experimentales anteriores si un polipéptido/secuencia de aminoácidos dado tiene de hecho una conformación de enrollamiento al azar.

Los homopolímeros de la mayoría de los aminoácidos, en particular los aminoácidos hidrófobos, son habitualmente insolubles en disolución acuosa (Bamford (1956) *Synthetic Polypeptides - Preparation, Structure, and Properties*, 2ª ed., Academic Press, Nueva York). Se sabe que los homopolímeros de varios aminoácidos hidrófilos forman estructuras secundarias, por ejemplo hélices α en el caso de Ala (Shental-Bechor (2005) *Biophys J* 88:2391-2402) y láminas β en el caso de Ser (Quadrioglio (1968) *J Am Chem Soc* 90:2760-2765), mientras que poliprolina, el homooligopéptido más rígido (Schimmel (1967) *Proc Natl Acad Sci USA* 58:52-59), forma una hélice trans de tipo II en disolución acuosa (Cowan (1955) *Nature* 176:501-503).

Usando los principios teóricos de biofísica de polímeros, el diámetro de enrollamiento al azar de una cadena de 200 residuos de aminoácido ascenderá en el caso de poliglicina, por ejemplo, hasta cerca de 75 Å (calculado como

media cuadrática promedio de la distancia de extremo a extremo de $\sqrt{\langle r^2 \rangle_0} = l \cdot \sqrt{n \cdot C_\infty}$, con $n = 200$ enlaces rotatorios de longitud $l = 3,8$ Å para cada distancia $C_\alpha-C_\alpha$ y la "razón característica" $C_\infty \approx 2,0$ para poli(Gly) (Brant (1967) *J Mol Biol* 23:47-65; Creighton, (1993), citado anteriormente)). Esta relación muestra que el experto en la técnica esperará que el volumen hidrodinámico de un polímero de aminoácidos de cadena al azar pueda extenderse o bien (a) usando una longitud de cadena l más larga o bien (b) usando aminoácidos que muestran una razón característica, C_∞ , mayor. C_∞ es una medida de la rigidez inherente de la cadena al azar molecular y tiene un valor general de 9 para la mayoría de los aminoácidos (Brant (1967), citado anteriormente). Sólo Gly, que carece de una cadena lateral, y también el iminoácido Pro, muestran valores significativamente menores. Por tanto, se espera que Gly y Pro (en condiciones desnaturizantes) contribuyan a reducir las dimensiones de proteínas de enrollamiento al azar (Miller (1968) *Biochemistry* 7:3925-3935). Por consiguiente, se espera que secuencias de aminoácidos que comprenden residuos de prolina tengan un volumen hidrodinámico relativamente compacto. Sin embargo, a diferencia de estas enseñanzas, en el presente documento se muestra que el volumen hidrodinámico de los polímeros de aminoácidos/polipéptidos de la invención que comprenden una mezcla de residuos de prolina y alanina tienen un volumen hidrodinámico drásticamente aumentado tal como se determina mediante permeación en gel analítica/cromatografía de exclusión molecular en comparación con el volumen hidrodinámico esperado. De hecho, resulta sorprendente que polipéptidos que comprenden mezclas de estos dos aminoácidos (prolina y alanina), de los que cada uno por sí solo tiende a formar un homooligopéptido con estructura secundaria definida, adopten una conformación de enrollamiento al azar en condiciones fisiológicas. Tales polipéptidos de prolina/alanina de la invención tienen un radio hidrodinámico mayor que homopolímeros que comprenden el mismo número de residuos de Gly, por ejemplo, y confieren mejor solubilidad a las proteínas biológicamente activas o constructos, es decir proteínas heterólogas biológicamente activas o conjugados de fármaco, según la invención.

Tal como se mencionó anteriormente, los polipéptidos de prolina/alanina de enrollamiento al azar biosintéticos de la presente invención difieren de polipéptidos químicamente sintetizados en que pueden adoptar una longitud uniforme definida mediante medios y métodos fáciles. Mientras que la técnica anterior proporciona mezclas/composiciones de polipéptidos con enormes variaciones en cuanto a la longitud de los péptidos, la presente invención puede proporcionar mezclas/composiciones de polipéptidos de enrollamiento al azar biosintéticos con una longitud definida. De manera preferible, esencialmente todos los polipéptidos de la invención comprendidos en una mezcla/composición de este tipo tienen la misma longitud definida y, por tanto, comparten las mismas características bioquímicas. Una composición uniforme de este tipo es más ventajosa en las diversas aplicaciones médicas, cosméticas, nutricionales, en las que pueden emplearse los polipéptidos de enrollamiento al azar biosintéticos. Además, en particular en un contexto médico o farmacéutico, el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido definido en el presente documento que comprende una secuencia de aminoácidos que consiste en al menos aproximadamente 50, en particular en al menos aproximadamente 100, en particular en al menos aproximadamente 150, en particular en al menos aproximadamente 200, en particular en al menos

aproximadamente 250, en particular en al menos aproximadamente 300, en particular en al menos aproximadamente 350, en particular en al menos aproximadamente 400 residuos de aminoácido de prolina y alanina también puede usarse en la prevención, mejora y/o tratamiento de trastornos asociados y/o afiliados con una situación de plasma sanguíneo alterada, por ejemplo tras lesiones, quemaduras, cirugía y similares. Por consiguiente, un uso médico de dichos polipéptidos de enrollamiento al azar biosintéticos o segmentos de polipéptido es el uso como expansor del plasma. Sin embargo, debe observarse que según esta invención los conjugados de fármaco descritos en el presente documento, tales como polipéptidos heterólogos o constructos de polipéptido heterólogo, también pueden emplearse en el contexto de la intervención médica o farmacéutica de un trastorno relacionado con una cantidad de plasma sanguíneo alterada o contenido en plasma sanguíneo o de un trastorno relacionado con un volumen de sangre alterado.

Por consiguiente, la presente invención se refiere en una realización a un polipéptido al azar biosintético (o segmento del mismo) que comprende una secuencia de aminoácidos que consiste únicamente en al menos aproximadamente 50 residuos de aminoácido de prolina y alanina, en al menos aproximadamente 100 residuos de aminoácido de prolina y alanina, en al menos aproximadamente 150 residuos de aminoácido de prolina y alanina o en al menos aproximadamente 200 residuos de prolina y alanina cuando está comprendido en un conjugado de fármaco, tal como una proteína heteróloga/polipéptido/constructo de polipéptido. La presente invención también se refiere a polipéptidos de enrollamiento al azar biosintéticos que comprenden una secuencia de aminoácidos que consiste únicamente en al menos aproximadamente 200 residuos de aminoácido de prolina y alanina, incluso más preferiblemente en al menos aproximadamente 300 residuos de aminoácido de prolina y alanina, de manera particularmente preferible en al menos aproximadamente 400 residuos de aminoácido de prolina y alanina, de manera más particularmente preferible en al menos aproximadamente 500 residuos de aminoácido de prolina y alanina y lo más preferiblemente en al menos aproximadamente 600 residuos de aminoácido de prolina y alanina que van a usarse en el presente documento. La secuencia de aminoácidos que forma una conformación de enrollamiento al azar puede consistir como máximo en aproximadamente 3000 residuos de aminoácido de prolina y alanina, como máximo en aproximadamente 2000 residuos de aminoácido de prolina y alanina, como máximo en aproximadamente 1500 residuos de aminoácido de prolina y alanina, como máximo en aproximadamente 1200 residuos de aminoácido de prolina y alanina, como máximo en aproximadamente 800 residuos de aminoácido de prolina y alanina. Por consiguiente, el tramo de secuencia de aminoácidos de prolina/alanina puede consistir en aproximadamente 50, en aproximadamente 100, en aproximadamente 150, en aproximadamente 200, en aproximadamente 250, en aproximadamente 300, en aproximadamente 350, en aproximadamente 400, en aproximadamente 500, en aproximadamente 600, en aproximadamente 700, en aproximadamente 800, en de aproximadamente 900 a aproximadamente 3000 residuos de aminoácido de prolina y alanina. En determinadas realizaciones, la secuencia de aminoácidos biosintética de la invención que va a usarse en el presente documento comprende de aproximadamente 200 a aproximadamente 3000 residuos de prolina y alanina, de aproximadamente 200 a aproximadamente 2500 residuos de prolina y alanina, de aproximadamente 200 a aproximadamente 2000 residuos de prolina y alanina, de aproximadamente 200 a aproximadamente 1500 residuos de prolina y alanina, de aproximadamente 200 a aproximadamente 1000 residuos de prolina y alanina, de aproximadamente 300 a aproximadamente 3000 residuos de prolina y alanina, de aproximadamente 300 a aproximadamente 2500 residuos de prolina y alanina, de aproximadamente 300 a aproximadamente 2000 residuos de prolina y alanina, de aproximadamente 300 a aproximadamente 1500 residuos de prolina y alanina, de aproximadamente 300 a aproximadamente 1000 residuos de prolina y alanina, de aproximadamente 400 a aproximadamente 3000 residuos de prolina y alanina, de aproximadamente 400 a aproximadamente 2500 residuos de prolina y alanina, de aproximadamente 400 a aproximadamente 2000 residuos de prolina y alanina, de aproximadamente 400 a aproximadamente 1500 residuos de prolina y alanina, de aproximadamente 400 a aproximadamente 1000 residuos de prolina y alanina, de aproximadamente 500 a aproximadamente 3000 residuos de prolina y alanina, de aproximadamente 500 a aproximadamente 2500 residuos de prolina y alanina, de aproximadamente 500 a aproximadamente 2000 residuos de prolina y alanina, de aproximadamente 500 a aproximadamente 1500 residuos de prolina y alanina, de aproximadamente 500 a aproximadamente 1000 residuos de prolina y alanina, de aproximadamente 600 a aproximadamente 3000 residuos de prolina y alanina, de aproximadamente 600 a aproximadamente 2500 residuos de prolina y alanina, de aproximadamente 600 a aproximadamente 2000 residuos de prolina y alanina, de aproximadamente 600 a aproximadamente 1500 residuos de prolina y alanina, de aproximadamente 600 a aproximadamente 1000 residuos de prolina y alanina, de aproximadamente 700 a aproximadamente 3000 residuos de prolina y alanina, de aproximadamente 700 a aproximadamente 2500 residuos de prolina y alanina, de aproximadamente 700 a aproximadamente 2000 residuos de prolina y alanina, de aproximadamente 700 a aproximadamente 1500 residuos de prolina y alanina, de aproximadamente 700 a aproximadamente 1000 residuos de prolina y alanina, de aproximadamente 800 a aproximadamente 3000 residuos de prolina y alanina, de aproximadamente 800 a aproximadamente 2500 residuos de prolina y alanina, de aproximadamente 800 a aproximadamente 2000 residuos de prolina y alanina, de aproximadamente 800 a aproximadamente 1500 residuos de prolina y alanina, de aproximadamente 800 a aproximadamente 1000 residuos de prolina y alanina. Tal como resulta evidente a partir del contenido de esta invención, secuencias de aminoácidos biosintéticas más grandes (que consisten esencialmente en prolina y alanina) también están dentro del alcance de esta invención y pueden emplearse fácilmente en las proteínas biológicamente activas o constructos de proteína definidos en el presente documento que comprenden como un dominio de al menos dos dominios una secuencia de aminoácidos que tiene y/o que media en dicha actividad biológica y como otro dominio de al menos dos dominios el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que consiste en al menos

- aproximadamente 50 residuos de aminoácido de prolina y alanina, en al menos aproximadamente 100 residuos de aminoácido de prolina y alanina, en al menos aproximadamente 150 residuos de aminoácido de prolina y alanina, en al menos aproximadamente 200, en al menos aproximadamente 250, en al menos aproximadamente 300, en al menos aproximadamente 350, en al menos aproximadamente 400 residuos de aminoácido de prolina y alanina. Un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido de este tipo corresponde a la parte de enrollamiento al azar biosintética de una proteína heteróloga/constructo de proteína. Estos tramos de prolina/alanina biosintéticos consisten como máximo en aproximadamente 3000 residuos de aminoácido de prolina y alanina. Estas secuencias de aminoácidos (tramos de prolina/alanina) comprenden prolina y alanina como residuos principales o únicos tal como se explica adicionalmente a continuación.
- Se prevé que es la secuencia de aminoácidos biosintética definida en el presente documento que consiste únicamente en residuos de aminoácido de prolina (P) y alanina (A) la que forma/adopta/tiene una conformación de enrollamiento al azar. En el caso más sencillo, el polipéptido biosintético o segmento de polipéptido consiste en la secuencia de aminoácidos que tiene una conformación de enrollamiento al azar tal como se define en el presente documento. Sin embargo, el polipéptido biosintético (o segmento del mismo) puede comprender, además de la secuencia de aminoácidos descrita en el presente documento que forma/adopta/tiene una conformación de enrollamiento al azar, secuencias de aminoácidos/residuos de aminoácido adicionales que no contribuyen a la formación de la conformación de enrollamiento al azar o que no pueden formar/adoptar/tener una conformación de enrollamiento al azar por sí solos. Sin apartarse de la esencia de la invención, tales polipéptidos biosintéticos (o segmentos de los mismos) también son polipéptidos de "enrollamiento al azar" biosintéticos o segmentos de polipéptido. Las secuencias de aminoácidos/residuos de aminoácido adicionales pueden ser útiles, por ejemplo, como grupos de unión. Entre otras cosas, también se prevén dímeros, trímeros, es decir multímeros en general del polipéptido de enrollamiento al azar biosintético, en el contexto de la presente invención y tales multímeros pueden unirse mediante secuencias de aminoácidos/residuos que no forman una conformación de enrollamiento al azar. Un ejemplo de una proteína que puede comprender un polipéptido de enrollamiento al azar de este tipo es la proteína biológicamente activa proporcionada en el presente documento, que puede comprender adicionalmente, además del polipéptido de enrollamiento al azar que consiste en residuos de aminoácido de prolina y alanina tal como se define en el presente documento, otro polipéptido que tiene/media en una actividad biológica. De nuevo, un constructo de este tipo puede ser un constructo de polipéptido o proteína heterólogo biológicamente activo tal como se describe en el presente documento.
- El término "al menos aproximadamente 50/100/150/200/300/400/500/600/700/800/etc. residuos de aminoácido" no está limitado al número específico de residuos de aminoácido sino que comprende tramos de aminoácidos que comprenden o bien aproximadamente el 1-20%, tal como del 10% al 20% más residuos o bien aproximadamente el 1-20%, tal como aproximadamente del 10% al 20% menos residuos. Por ejemplo "al menos aproximadamente 100 residuos de aminoácido" también puede comprender aproximadamente de 80 a 100 y aproximadamente de 100 a 120 residuos de aminoácido sin apartarse de la esencia de la presente invención. Por ejemplo "al menos aproximadamente 200 residuos de aminoácido" también puede comprender aproximadamente de 160 a 200 y aproximadamente de 200 a 240 residuos de aminoácido sin apartarse de la esencia de esta invención. La definición y explicaciones facilitadas anteriormente en el presente documento también se aplican, cambiando lo que sea necesario, al término "como máximo aproximadamente 3000/2000/1500/1200/800 residuos de aminoácido" etc. Por consiguiente, el término "aproximadamente" no se limita o restringe al número específico de residuos de aminoácido en el contexto de secuencias de aminoácidos más largas (por ejemplo secuencias de aminoácidos que comprenden o consisten como máximo en 3000 residuos de aminoácido). Por tanto, el término "como máximo aproximadamente 3000/2000/1500/1200/800 residuos de aminoácido" también puede comprender tramos de aminoácidos que comprenden del 10% al 20% más o del 10% al 20% menos residuos sin apartarse de esta invención.
- Además, los polipéptidos de enrollamiento al azar biosintéticos (o segmentos de los mismos) se caracterizan por un contenido definido o razón de residuos de aminoácido, en particular de los constituyentes principales prolina y alanina. Tal como se mencionó anteriormente, la presente invención se refiere a un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos aproximadamente 50, en al menos aproximadamente 100, en al menos aproximadamente 150, en al menos aproximadamente 200 en al menos aproximadamente 250, en al menos aproximadamente 300, en al menos aproximadamente 350, en al menos aproximadamente 400 residuos de aminoácido de prolina (Pro) y alanina (Ala) en particular cuando está comprendida en una proteína biológicamente activa heteróloga/constructo de proteína/polipéptido o conjugado de fármaco.
- Por consiguiente, la presente invención se refiere en una realización a un polipéptido de enrollamiento al azar biosintético (o segmento del mismo) mediante el cual la secuencia de aminoácidos consiste principalmente en prolina y alanina, y en el que los residuos de prolina constituyen más de aproximadamente el 10% y menos del 75% de la secuencia de aminoácidos. Los residuos de alanina comprenden el al menos del 25% al 90% restante de dicha secuencia de aminoácidos (o el polipéptido de enrollamiento al azar o segmento de polipéptido si consiste en la secuencia de aminoácidos).

Preferiblemente, la secuencia de aminoácidos comprende más de aproximadamente el 10%, preferiblemente más de aproximadamente el 12%, incluso más preferiblemente más de aproximadamente el 14%, de manera

particularmente preferible más de aproximadamente el 18%, de manera más particularmente preferible más de aproximadamente el 20%, de manera incluso más particularmente preferible más de aproximadamente el 22%, el 23% o el 24% y lo más preferiblemente más de aproximadamente el 25% de residuos de prolina. La secuencia de aminoácidos comprende preferiblemente menos de aproximadamente el 75%, más preferiblemente menos del 70%, el 65%, el 60%, el 55% o el 50% de residuos de prolina, en la que se prefieren los valores inferiores. Incluso más preferiblemente, la secuencia de aminoácidos comprende menos de aproximadamente el 48%, el 46%, el 44%, el 42% de residuos de prolina. Se prefieren particularmente secuencias de aminoácidos que comprenden menos de aproximadamente el 41%, el 40%, el 39% el 38%, el 37% o el 36% de residuos de prolina, mediante lo cual se prefieren los valores inferiores. Lo más preferiblemente, la secuencia de aminoácidos comprende menos de aproximadamente el 35% de residuos de prolina; véase también los constructos de PA proporcionados a continuación en el presente documento.

A la inversa, la secuencia de aminoácidos comprende preferiblemente menos de aproximadamente el 90%, más preferiblemente menos del 88%, el 86%, el 84%, el 82% o el 80% de residuos de alanina, en la que se prefieren los valores inferiores. Incluso más preferiblemente, la secuencia de aminoácidos comprende menos de aproximadamente el 79%, el 78%, el 77%, el 76% de residuos de alanina, mediante la cual se prefieren los valores inferiores. Lo más preferiblemente, la secuencia de aminoácidos comprende menos de aproximadamente el 75% de residuos de alanina.

También se prefiere en el presente documento una secuencia de aminoácidos que comprende más de aproximadamente el 25%, preferiblemente más de aproximadamente el 30%, incluso más preferiblemente más de aproximadamente el 35%, de manera particularmente preferible más de aproximadamente el 40%, de manera más particularmente preferible más de aproximadamente el 45% o el 50%, de manera incluso más particularmente preferible más de aproximadamente el 52%, el 54%, el 56%, el 58% o el 59% de residuos de alanina, en la que se prefieren los valores superiores. Incluso más preferiblemente, la secuencia de aminoácidos comprende más de aproximadamente el 60%, el 61%, el 62%, el 63% o el 64% de residuos de alanina y lo más preferiblemente más de aproximadamente el 65% de residuos de alanina.

Por consiguiente, el polipéptido de enrollamiento al azar (o segmento del mismo) puede comprender una secuencia de aminoácidos que consiste en aproximadamente el 25% de residuos de prolina, y aproximadamente el 75% de residuos de alanina. Alternativamente, el polipéptido de enrollamiento al azar (o segmento del mismo) puede comprender una secuencia de aminoácidos que consiste en aproximadamente el 35% de residuos de prolina y aproximadamente el 65% de residuos de alanina. El término "aproximadamente el X%" tal como se usó anteriormente en el presente documento no está limitado al número específico del porcentaje, sino que también comprende valores del 10% al 20% más o del 10% al 20% menos residuos. Por ejemplo, el término el 10% también puede referirse al 11% o al 12% y al 9% y al 8%, respectivamente.

Según la invención, un polipéptido de enrollamiento al azar biosintético (o segmento del mismo)/la secuencia de aminoácidos consiste exclusivamente en residuos de aminoácido de prolina y alanina, es decir, no hay ningún otro residuo de aminoácido presente en el polipéptido de enrollamiento al azar o en la secuencia de aminoácidos.

Mientras que lo anterior se refiere a la longitud y contenido en prolina/alanina globales de la secuencia de aminoácidos comprendida en el polipéptido de enrollamiento al azar (o segmento del mismo), lo siguiente se refiere en más detalle a las secuencias de aminoácidos a modo de ejemplo específicas (o fragmentos de las mismas).

En una realización, las secuencias de aminoácidos/polipéptidos que adoptan una conformación de enrollamiento al azar (el polipéptido de enrollamiento al azar o segmento del mismo tal como se define en el presente documento), por ejemplo, en disolución acuosa o en condiciones fisiológicas, pueden comprender una pluralidad de "repeticiones de aminoácidos"/"casetes de aminoácidos"/"repeticiones de casete", en las que dichas "repeticiones de aminoácidos"/"casetes de aminoácidos"/"repeticiones de casete"/"bloque estructural"/"módulos" (estos términos se usan en el presente documento de manera intercambiable) consisten exclusivamente en residuos de aminoácido de prolina (Pro, P) y alanina (Ala, A) (representados en el presente documento como "PA" o como "AP"), en las que no más de 6 residuos de aminoácido consecutivos son idénticos. Un "bloque estructural" ilustrativo es, por ejemplo, "AP" y esto también se ha proporcionado en los ejemplos ilustrativos adjuntos como dominio de enrollamiento al azar biosintético funcional de la presente invención. Este ejemplo ilustrativo es la secuencia "P1A1" tal como también se proporciona en forma de APAPAPAPAPAPAPAPAP (SEQ ID NO: 51), es decir, una "repetición de aminoácidos"/"aminoácido casete"/"repetición de casete" de "poli-PA". En una realización preferida, la secuencia de aminoácidos/polipéptido que comprende las "repeticiones de aminoácidos"/"casetes de aminoácidos"/"repeticiones de casete" definidas anteriormente y similares no comprende más de 5 residuos de aminoácido consecutivos idénticos. A continuación en el presente documento se proporcionan otras realizaciones alternativas en el contexto de bloques estructurales individuales mostrados a modo de ejemplo.

Dentro de un polipéptido de enrollamiento al azar (o segmento del mismo) según esta invención, las repeticiones de aminoácidos pueden ser idénticas o no idénticas. A continuación en el presente documento se proporcionan ejemplos no limitativos de "repeticiones de aminoácidos", "bloques estructurales", "módulos", "repeticiones", "casetes de aminoácidos", etc., que consisten en residuos de prolina y alanina; véanse, por ejemplo SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 y SEQ ID NO: 51 (la lista de secuencias adjunta

también comprende secuencias de ácido nucleico ilustrativas que codifican para tales “repeticiones”/“módulos”, etc. Las secuencias adjuntas en dicha lista de secuencias tal como se presenta con el presente documento constituyen parte de esta memoria descriptiva y descripción). En el presente documento también se prevé el uso de fragmentos (idénticos y/o no idénticos) de estas secuencias, mediante lo cual un “fragmento” comprende al menos 2 aminoácidos y comprende al menos una prolina y/o alanina, preferiblemente al menos una prolina y una alanina. Los “fragmentos” de estas secuencias que van a emplearse según esta invención para la generación del polipéptido de enrollamiento al azar (o segmento del mismo) pueden consistir en al menos 3, preferiblemente en al menos 4, más preferiblemente en al menos 5, incluso más preferiblemente en al menos 6, todavía más preferiblemente en al menos 8, de manera particularmente preferible en al menos 10, de manera más particularmente preferible en al menos 12, de manera incluso más particularmente preferible en al menos 14, de manera todavía más particularmente preferible en al menos 16, y lo más preferiblemente en al menos 18 aminoácidos consecutivos de la secuencia de aminoácidos seleccionada del grupo que consiste en dichas SEQ ID NO: 1, 2, 3, 4, 5, 6 y 51 (debe observarse en este caso que SEQ ID NO: 51 consiste en una repetición de “AP” o “PA” ilustrativa).

Basándose en las enseñanzas facilitadas en el presente documento, el experto en la técnica está fácilmente en una posición para generar secuencias de aminoácidos/polipéptidos adicionales que forman una conformación de enrollamiento al azar, por ejemplo en condiciones acuosas o fisiológicas y que están constituidas principalmente por prolina y alanina tal como se define en el presente documento. Ejemplos adicionales de secuencias de aminoácidos/polipéptidos que forman una conformación de enrollamiento al azar que van a usarse como bloques estructurales o módulos del polipéptido de enrollamiento al azar (o segmento del mismo) definido en el presente documento pueden comprender, entre otras cosas, combinaciones y/o fragmentos o versiones permutadas de manera circular de los “bloques estructurales”, “casetes de polímeros” o “repeticiones de polímeros” específicos mostrados anteriormente. Por consiguiente, los módulos/unidades de secuencia/repeticiones de polímeros/casetes de polímeros mostrados a modo de ejemplo del polipéptido de enrollamiento al azar/secuencia de aminoácidos también pueden proporcionar fragmentos individuales que pueden combinarse de una manera nueva para formar módulos/unidades de secuencia/repeticiones de polímeros/casetes de polímeros adicionales según esta invención.

Los términos “módulo(s)”, “unidad(es) de secuencia”, “repetición/repeticiones de polímeros”, “casete(s) de polímeros” y “bloque(s) estructural(es)” se usan como sinónimos en el presente documento y se refieren a tramos de aminoácidos individuales que pueden usarse para formar el polipéptido de enrollamiento al azar (o segmento del mismo)/secuencia de aminoácidos definido en el presente documento.

Una repetición de aminoácidos (usada como “bloque estructural”, etc., de un polipéptido de enrollamiento al azar biosintético de la presente invención) puede consistir en al menos 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 o más residuos de aminoácido, en la que cada repetición comprende un(os) residuo(s) de prolina y alanina. Sin embargo, tal como se ilustra en la SEQ ID NO: 51 adjunta, dicho “bloque estructural” también puede consistir simplemente en los 2 residuos de aminoácido P y A proporcionados en el presente documento, concretamente en forma de “PA” o “AP”. En una realización, la repetición de aminoácidos según la presente invención no comprende más de 50 residuos de aminoácido. Sin embargo, resulta evidente para el experto en la técnica que una “repetición” de este tipo puede comprender incluso más de 50 residuos de aminoácido, por ejemplo, en casos en los que dicho polipéptido/polímero de enrollamiento al azar biosintético de la invención comprende más de aproximadamente, por ejemplo, 100 aminoácidos, más de aproximadamente 150 aminoácidos, más de aproximadamente 200 aminoácidos, etc. Por consiguiente, la cantidad máxima de residuos de aminoácido comprendidos en una “repetición” de este tipo está condicionada por la longitud global del polipéptido biosintético (o segmento del mismo)/polímero tal como se proporciona en el presente documento.

Sin embargo, debe observarse que los polipéptidos de enrollamiento al azar biosintéticos/secuencias de aminoácidos que comprenden las repeticiones, etc., anteriores deben tener preferiblemente la longitud y/o contenido en prolina/alanina globales tal como se definió y explicó anteriormente en el presente documento, es decir, consistir en aproximadamente 50, en aproximadamente 100, en aproximadamente 150, en aproximadamente 200, en aproximadamente 250, en aproximadamente 300, en aproximadamente 350, en de aproximadamente 400 a aproximadamente 3000 aminoácidos y/o comprender más de aproximadamente el 10% y menos de aproximadamente el 75% de residuos de prolina. Todas las definiciones facilitadas anteriormente en el presente documento en este contexto también se aplican aquí, cambiando lo que sea necesario.

Tal como se comenta en detalle en el presente documento y tal como se proporcionó anteriormente en el presente documento, la presente invención proporciona una(s) proteína(s) heteróloga(s) biológicamente activa(s) o un(os) constructo(s) de proteína que es/son particularmente útiles en un entorno farmacéutico, médico y/o medicinal. Estas proteínas heterólogas biológicamente activas/constructos de proteína comprenden como al menos un dominio de dichos al menos dos dominios el polipéptido de enrollamiento al azar o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de prolina y alanina, en las que dicha secuencia de aminoácidos consiste en aproximadamente 50, en aproximadamente 100, en aproximadamente 150, en aproximadamente 200, en aproximadamente 250, en aproximadamente 300, en aproximadamente 350, en de aproximadamente 400 a aproximadamente 3000 residuos de prolina (Pro) y alanina (Ala).

En el contexto de las proteínas heterólogas biológicamente activas, polipéptidos o constructos de proteína tal como

se da a conocer en el presente documento, el término “heterólogo” se refiere a al menos dos dominios dentro de dichas proteínas, polipéptidos o constructos de proteína en las que un primero de dichos al menos dos dominios confiere, tiene y/o media en una actividad biológica definida y en las que un segundo de dichos al menos dos dominios comprende el polipéptido de enrollamiento al azar biosintético que consiste únicamente en residuos de aminoácido de prolina y alanina y mediante lo cual dichos al menos dos dominios no se encuentran operativamente unidos entre sí en la naturaleza o no se codifican mediante una única secuencia de ácido nucleico codificante (tal como un marco de lectura abierto) que existe en la naturaleza. El polipéptido de enrollamiento al azar biosintético/segmento de polipéptido que consiste únicamente en residuos de aminoácido de prolina y alanina tal como se proporciona en el presente documento y tal como se emplea en las proteínas heterólogas biológicamente activas/constructos de proteína de esta invención preferiblemente no están (químicamente) modificados de manera adicional, por ejemplo preferiblemente no están glicosilados ni hidroxilados.

Debe observarse que determinadas proteínas que se producen de manera natural o proteínas hipotéticas tal como se deducen a partir de tramos de secuencias de ácido nucleico encontradas en la naturaleza se describen como que comprenden un contenido relativamente alto (es decir, por encima de la media) de prolina y alanina. Por ejemplo, se ha descrito una proteína hipotética homóloga para la cepa Friedlin de *Leishmania major* (Ivens (2005) Science 309, 436-442.). El marco de lectura dado a conocer que comprende 1514 tripletes de codones incluye un tramo de 412 tripletes compuesto por 240 codones de Ala, 132 de Pro, 34 de Lys y 4 de Val. Los residuos de Lys, que están cargados positivamente en condiciones de tampón fisiológico, están distribuidos de manera casi uniforme entre esta secuencia, lo que sugiere un efecto solubilizante. Sin embargo, tal como resulta evidente a partir de la divulgación en el presente documento, una proteína hipotética homóloga de este tipo tal como se deduce a partir de un marco de lectura abierto o molécula de ácido nucleico que se produce de manera natural, que comprende un alto contenido en prolina y alanina por encima de la media no forma parte de esta invención. La invención se basa en el hecho de que se proporciona un polipéptido de enrollamiento al azar o segmento de polipéptido bastante grande que no se produce en la naturaleza de una manera aislada y que comprende una secuencia de aminoácidos que consiste únicamente en residuos de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en aproximadamente 50, en aproximadamente 100, en aproximadamente 150, en aproximadamente 200, en aproximadamente 250, en aproximadamente 300, en aproximadamente 350, en de aproximadamente 400 a aproximadamente 3000 residuos en prolina (Pro) y alanina (Ala), que es particularmente útil en el contexto médico/farmacéutico. Los polipéptidos de enrollamiento al azar biosintéticos o segmentos de polipéptido aislados descritos en el presente documento que no se producen en la naturaleza de una manera aislada también están comprendidos en la(s) proteína(s) heteróloga(s) biológicamente activa(s) o el/los constructo(s) de proteína dados a conocer en el presente documento que son particularmente útiles en un entorno farmacéutico, médico y/o medicinal. Estas proteínas heterólogas biológicamente activas/constructos de proteína comprenden como al menos un dominio de dichos al menos dos dominios el polipéptido de enrollamiento al azar o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en aproximadamente 50, en aproximadamente 100, en aproximadamente 150, en aproximadamente 200, en aproximadamente 250, en aproximadamente 300, en aproximadamente 350, en de aproximadamente 400 a aproximadamente 3000 residuos en prolina (Pro) y alanina (Ala).

Además, las proteínas de arabinogalactano (AGP), proteínas ricas en Pro y extensinas pertenecen a un gran grupo de glicoproteínas, conocidas como glicoproteínas ricas en hidroxiprolina (Hip) (HRGP), que se expresan a lo largo de todo el reino vegetal. Un motivo de AGP de este tipo que comprende una repetición de Ala-Pro (AP)₅₁ se expresó como péptido de glicomódulo sintético con secuencia señal N-terminal y proteína verde fluorescente C-terminal en *Arabidopsis thaliana* transgénica y se investigó como sustrato para prolil hidroxilasas y posterior O-glicosilación de los residuos de hidroxiprolina (Estévez (2006) Plant Physiol. 142, 458-470). De nuevo, las cadenas laterales de Pro hidroxiladas y/o glicosiladas dadas a conocer, que pueden formar enlaces de hidrógeno con moléculas de agua, parecen tener un efecto solubilizante.

Debe observarse que las “proteínas biológicamente activas o constructos de proteína que comprenden como (al menos) un dominio un polipéptido de enrollamiento al azar biosintético o segmento de péptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de prolina y alanina” descritos en el presente documento se refieren a proteínas o constructos de proteína que no se producen normalmente en la naturaleza y, por tanto, son “heterólogos”. Además, y a diferencia de secuencias ricas en prolina descritas en el reino vegetal, los polipéptidos de enrollamiento al azar biosintéticos/segmentos de polipéptido descritos en el presente documento preferiblemente no están químicamente modificados, es decir, preferiblemente no están glicosilados o hidroxilados.

El polipéptido de enrollamiento al azar biosintético/secuencias de aminoácidos de la presente invención puede comprender concatenados de bloques individuales que comprenden tramos de prolina/alanina combinadas de la secuencia (Pro)_x-(Ala)_y, mediante lo cual x puede tener un valor de número entero de desde 1 hasta preferiblemente 15, más preferiblemente de 1 a 10, incluso más preferiblemente de 1 a 5, e y puede tener un valor de número entero de desde 1 hasta preferiblemente 15, más preferiblemente de 1 a 10, incluso más preferiblemente de 1 a 5, y x e y pueden variar entre bloques posteriores. Dichos x e y también pueden ser un número entero de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 ó 15.

Las secuencias de aminoácidos/polipéptidos que forman una conformación de enrollamiento al azar en disolución acuosa o en condiciones fisiológicas pueden tener la fórmula (I):

[Pro_xAla_y]_n

5 en la que x se selecciona independientemente de un número entero de 1 a 5. Además, para cada n, y se selecciona independientemente de un número entero de 1 a 5. Finalmente, n es cualquier número entero siempre que el polipéptido de enrollamiento al azar (o segmento del mismo)/secuencia de aminoácidos consista preferiblemente en al menos aproximadamente 50, en al menos aproximadamente 100, en al menos aproximadamente 150, en al menos aproximadamente 200, en al menos aproximadamente 250, en al menos aproximadamente 300, en al menos aproximadamente 350, en al menos aproximadamente 400 residuos de aminoácido y hasta aproximadamente 3000 residuos de aminoácido. También en este contexto debe observarse que los polipéptidos/secuencias de aminoácidos que comprenden los concatémicos anteriores o que tienen la fórmula (I) anterior deben tener preferiblemente la longitud y/o contenido en prolina/alanina globales tal como se definió y explicó anteriormente en el presente documento, es decir, consistir en aproximadamente 50, en aproximadamente 100, en aproximadamente 150, en aproximadamente 200, en aproximadamente 250, en aproximadamente 300, en aproximadamente 350, en de aproximadamente 400 a aproximadamente 3000 aminoácidos y/o comprender más de aproximadamente el 10% y menos de aproximadamente el 75% de residuos de prolina. De nuevo, todas las definiciones facilitadas anteriormente en el presente documento en este contexto también se aplican aquí, cambiando lo que sea necesario.

10 La presente invención también se refiere a polipéptidos de enrollamiento al azar (un(os) segmento(s) de polipéptido)/secuencias de aminoácidos que comprenden un tramo de aminoácidos seleccionado del grupo que consiste en AAPAAPAAPAAPAAPAAP (SEQ ID NO: 1); AAPAAAPAAPAAPAAP (SEQ ID NO: 2); AAAPAAAPAAAPAAAP (SEQ ID NO: 3 que es un ejemplo para [Pro₁Ala₃]₅); AAPAAPAAPAAPAAPAAP (SEQ ID NO: 4); APAAAPAPAAAPAPAAAPAPAAAP (SEQ ID NO: 5); AAAPAAPAAPAAAPAAAP (SEQ ID NO: 6) y APAPAPAPAPAPAPAPAP (SEQ ID NO: 51 que es un ejemplo para [Pro₁Ala₁]₁₀) o versiones permutadas circulares o un(os) múltiplo(s) de estas secuencias en su conjunto o partes de estas secuencias. Por consiguiente, el polipéptido de enrollamiento al azar (un(os) segmento(s) de polipéptido del mismo)/secuencia de aminoácidos puede comprender el tramo de aminoácidos

20 AAPAAPAAPAAPAAPAAP (SEQ ID NO: 1), AAPAAPAAPAAPAAPAAP (SEQ ID NO: 1); AAPAAPAAPAAPAAPAAP (SEQ ID NO: 2); AAAPAAAPAAAPAAAP (SEQ ID NO: 3); AAPAAPAAPAAPAAPAAP (SEQ ID NO: 4); APAAAPAPAAAPAPAAAPAPAAAP (SEQ ID NO: 5); AAAPAAPAAPAAAPAAAP (SEQ ID NO: 6) y APAPAPAPAPAPAPAPAP (SEQ ID NO: 51), así como combinaciones de estos motivos o combinaciones de fragmentos y partes de estos motivos siempre que el polipéptido de enrollamiento al azar biosintético resultante consista únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala).

Según la presente invención también pueden usarse versiones permutadas circulares de las secuencias de aminoácidos anteriores. Pueden generarse fácilmente versiones permutadas circulares a modo de ejemplo de, por ejemplo, AAPAAPAAPAAPAAPAAP (SEQ ID NO: 1), por ejemplo retirando la primera alanina y añadiendo otra alanina al final de la secuencia anterior. Una versión permutada circular de este tipo de SEQ ID NO: 1 será entonces AAPAAPAAPAAPAAPAAP (SEQ ID NO: 7). Además, ejemplos no limitativos de versiones permutadas circulares de SEQ ID NO: 1 son:

PAAPAAPAAPAAPAAP (SEQ ID NO: 8),

40 AAPAAPAAPAAPAAPAAP (SEQ ID NO: 9),

APAAPAAPAAPAAPAAP (SEQ ID NO: 10),

PAPAAPAAPAAPAAPAAP (SEQ ID NO: 11),

APAAPAAPAAPAAPAAP (SEQ ID NO: 12),

PAAPAAPAAPAAPAAP (SEQ ID NO: 13),

45 AAPAAPAAPAAPAAPAAP (SEQ ID NO: 14),

APAAPAAPAAPAAPAAP (SEQ ID NO: 15),

PAAPAAPAAPAAPAAP (SEQ ID NO: 16),

y similares. Basándose en las enseñanzas de la presente invención, un experto está fácilmente en posición de generar versiones permutadas circulares correspondientes de los tramos de aminoácidos tal como se muestra en SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 y SEQ ID NO: 51 (basándose dicha SEQ ID NO: 51 totalmente en repeticiones de "AP" y una versión permutada circular puede basarse totalmente en repeticiones/bloques estructurales de "PA" o "AP").

Tales versiones permutadas circulares también pueden considerarse ejemplos de un "módulo"/"bloque estructural", etc., adicional de los polipéptidos/secuencias de aminoácidos proporcionados en el presente documento que pueden

usarse por consiguiente en el presente documento.

Resulta evidente para el experto en la técnica que "módulos" y fragmentos (más cortos) o versiones permutadas circulares de los tramos de aminoácidos proporcionados en el presente documento también pueden usarse como "módulos", "repeticiones" y/o bloques estructurales para el polipéptido de enrollamiento al azar (o segmento del mismo)/secuencia de aminoácidos definido en el presente documento.

Según lo anterior, el polipéptido de enrollamiento al azar/secuencia de aminoácidos que forma una conformación de enrollamiento al azar puede comprender un multímero de cualquiera de los tramos de aminoácidos anteriores (o versiones permutadas circulares o fragmentos de los mismos), preferiblemente los mostrados en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 y SEQ ID NO: 51. Debe observarse que estas secuencias no son de ningún modo limitativas en el contexto de esta invención.

De nuevo, los polipéptidos/secuencias de aminoácidos que comprenden los tramos de aminoácidos anteriores (o fragmentos de los mismos), versiones mutadas circulares (o fragmentos de las mismas) deben tener preferiblemente la longitud y/o contenido en prolina/alanina globales tal como se definió y explicó anteriormente en el presente documento, es decir consistir en aproximadamente 50, en aproximadamente 100, en aproximadamente 150, en aproximadamente 200, en aproximadamente 250, en aproximadamente 300, en aproximadamente 350, en de aproximadamente 400 a aproximadamente 3000 aminoácidos y/o comprender más de aproximadamente el 10% y menos de aproximadamente el 75% de residuos de prolina. Todas las definiciones facilitadas anteriormente en el presente documento en este contexto también se aplican aquí, cambiando lo que sea necesario. El término "fragmento" también se definió anteriormente.

Tal como se mencionó anteriormente, en el contexto de esta invención se encontró sorprendentemente que los polipéptidos de enrollamiento al azar biosintéticos (o segmento de polipéptido)/polímeros tal como se proporciona en el presente documento se caracterizan por un volumen hidrodinámico relativamente grande. Este volumen hidrodinámico, también denominado tamaño aparente, puede determinarse fácilmente mediante filtración en gel analítica (también conocida como cromatografía de exclusión molecular, SEC). Preferiblemente, el polipéptido de enrollamiento al azar (o segmento del mismo) tiene un tamaño aparente de al menos 10 kDa, preferiblemente de al menos 25 kDa, más preferiblemente de al menos 50 kDa, incluso más preferiblemente de al menos 100 kDa, de manera particularmente preferible de al menos 200 kDa y lo más preferiblemente de al menos 400 kDa. El experto en la técnica puede determinar fácilmente el volumen hidrodinámico de proteínas específicas. Tales métodos pueden incluir dispersión de la luz dinámica/estática, ultracentrifugación analítica o filtración en gel analítica tal como se muestra a modo de ejemplo a continuación en el presente documento. La filtración en gel analítica representa un método conocido en la técnica para medir el volumen hidrodinámico de macromoléculas. Alternativamente, el volumen hidrodinámico de un polipéptido globular puede estimarse mediante su peso molecular (Creighton (1993), citado anteriormente). Tal como se describe en el presente documento, el volumen hidrodinámico de los polipéptidos de la invención que consisten preferiblemente en al menos aproximadamente 50, en al menos aproximadamente 100, en al menos aproximadamente 150, en al menos aproximadamente 200, en al menos aproximadamente 250, en al menos aproximadamente 300, en al menos aproximadamente 350, en de al menos aproximadamente 400 a aproximadamente 3000 residuos de aminoácido de prolina y alanina y que tienen una conformación de enrollamiento al azar muestra valores inesperadamente altos con respecto al volumen hidrodinámico que se estimaría para una proteína globular plegada correspondiente basándose en el peso molecular. Lo siguiente se refiere a proteínas heterólogas biológicamente activas o constructos de proteína que comprenden, entre otras cosas, el polipéptido de enrollamiento al azar biosintético (o segmento del mismo) tal como se describió y se definió anteriormente en el presente documento que representan una realización preferida de la presente invención. Sin limitarse a la teoría, se encontró sorprendentemente en el contexto de la presente invención que los tramos de polipéptido de enrollamiento al azar biosintético tal como se proporciona en el presente documento y que consisten únicamente en prolina y alanina pueden incluso proporcionar un volumen hidrodinámico superior a un tramo de enrollamiento al azar biosintético correspondiente que tiene el mismo número total de residuos de aminoácido pero que consiste únicamente en prolina, alanina y serina (tal como se proporciona en el documento WO 2008/155134).

Las proteínas de plasma humanas tales como albúmina sérica (HSA) e inmunoglobulinas (Ig), incluyendo anticuerpos humanizados, muestran semividas largas, normalmente de 2 a 3 semanas, lo cual puede atribuirse a su interacción específica con el receptor de Fc neonatal (FcRn), conduciendo a recirculación endosómica (Ghetie (2002) Immunol Res, 25:97-113). En cambio, la mayoría de las demás proteínas de interés farmacéutico, en particular fragmentos de anticuerpo recombinantes, hormonas, interferones, etc., experimentan un aclaramiento (de la sangre) rápido. Esto es particularmente cierto para proteínas cuyo tamaño es inferior al valor de umbral para la filtración renal de aproximadamente 70 kDa (Caliceti (2003) Adv Drug Deliv Rev 55:1261-1277). En estos casos la semivida en plasma de una proteína farmacéutica sin modificar puede ser considerablemente menos de una hora, haciendo por tanto que sea esencialmente inútil para la mayoría de las aplicaciones terapéuticas. Con el fin de lograr una acción farmacológica sostenida y también mejorar el cumplimiento del paciente (al requerir intervalos de dosificación que se prolongan a días o incluso semanas), anteriormente se establecieron varias estrategias con fines de desarrollo de biofármacos.

En primer lugar, se ha empleado el mecanismo de recirculación de proteínas de plasma naturales produciendo proteínas de fusión con la parte Fc de Ig, por ejemplo Enbrel[®], un híbrido entre el dominio extracelular de receptor de

TNF α e IgG1 humana (Goldenberg (1999) Clin Ther 21:75-87) o con albúmina sérica, por ejemplo Albuferon® (albinterferón alfa-2b, ZALBIN™, JOULFERON®), una fusión correspondiente de IFN α con HSA (Osborn (2002) J Pharmacol Exp Ther 303:540-548). También se ha usado albúmina, con su alta concentración en plasma de 600 μ M, de una manera indirecta, sirviendo como vehículo portador para productos biofarmacéuticos que están equipados con una función de unión a albúmina, por ejemplo mediante fusión con un dominio de unión a albúmina (ABD) bacteriano de proteína G de estreptococos (Makrides (1996) J Pharmacol Exp Ther 277:534-542) o con un péptido seleccionado contra HSA a partir de una biblioteca de presentación en fagos (Dennis (2002) J Biol Chem, 277:35035-35043; Nguyen (2006) Protein Eng Des Sel 19:291-297).

En segundo lugar, una metodología fundamentalmente diferente para prolongar la semivida en plasma de productos biofarmacéuticos es la conjugación con polímeros químicos altamente solvatados y fisiológicamente inertes, agrandando por tanto eficazmente el radio hidrodinámico de la proteína terapéutica más allá del tamaño de poro glomerular de aproximadamente 3-5 nm (Caliceti (2003), citado anteriormente). El acoplamiento covalente en condiciones bioquímicamente leves con derivados activados de polietilenglicol (PEG), o bien al azar a través de cadenas laterales de Lys (Clark (1996) J Biol Chem 271:21969-21977) o bien por medio de residuos de Cys específicamente introducidos (Rosendahl (2005) BioProcess International:52-60) ha sido moderadamente satisfactorio y está aplicándose actualmente en varios fármacos aprobados. Se han logrado ventajas correspondientes especialmente en relación con proteínas pequeñas que presentan actividad farmacológica específica, por ejemplo Pegasys®, un IFN α -2a recombinante químicamente pegilado (Harris (2003) Nat Rev Drug Discov, 2:214-221; Walsh (2003) Nat Biotechnol 21:865-870).

Sin embargo, el acoplamiento químico de una proteína biológicamente activa con polímeros sintéticos tiene desventajas con respecto al desarrollo y la producción biofarmacéuticos. Los derivados de PEG adecuados son caros, especialmente ya que se necesita una alta pureza, y su conjugación con una proteína recombinante requiere etapas de procesamiento y purificación *in vitro* adicionales, lo cual reduce el rendimiento y aumenta los costes de fabricación. De hecho, el PEG con frecuencia está contaminado con aldehídos y peróxidos (Ray (1985) Anal Biochem 146:307-312) y es intrínsecamente propenso a degradación química con el almacenamiento en presencia de oxígeno. Además, la función farmacéutica de una proteína terapéutica puede verse dificultada si cadenas laterales de aminoácido en la proximidad de su sitio activo bioquímico se modifican mediante el procedimiento de pegilación. Además, el acoplamiento químico con polímeros sintéticos da habitualmente como resultado una mezcla heterogénea de moléculas que pueden mostrar una varianza sustancial de actividad *in vivo*.

En tercer lugar, se ha propuesto el uso de análogos de glicosilación de proteínas biológicamente activas en los que se introducen nuevas secuencias consenso de glicosilación unidas en N para prolongar la semivida en plasma; véase el documento WO 02/02597; Perlman (2003) J Clin Endocrinol Metab 88:2327-2335; o Elliott (2003) Nat Biotechnol 21:414-420. Sin embargo, las proteínas modificadas por glicoingeniería descritas presentaron una actividad *in vivo* alterada, lo cual indica que las nuevas cadenas laterales de hidratos de carbono influyen en la actividad biológica de la proteína modificada por ingeniería. Además, es probable que las cadenas laterales de hidratos de carbono adicionales aumenten la antigenicidad de las moléculas activas biológicas resultantes, lo cual plantea preocupaciones de seguridad sustanciales. Además, se ha notificado que proteínas de fusión que comprenden la secuencia repetitiva artificial PSTAD derivada de *Trypanosoma cruzi* inducen una semivida en plasma prolongada de trans-sialidasa (Alvarez (2004) JBC 279:3375-3381). Sin embargo, se ha notificado que tales repeticiones derivadas de *Trypanosoma cruzi* inducen una respuesta inmunitaria humoral (Alvarez (2004), citado anteriormente). Por consiguiente, se desean estrategias alternativas para prolongar la acción de proteínas biológicamente activas.

Se encontró sorprendentemente que las secuencias de aminoácidos/polipéptidos biosintéticos tal como se da a conocer en el presente documento y que consisten únicamente en prolina y alanina según la invención adoptan una conformación de enrollamiento al azar en particular en condiciones fisiológicas. Por tanto, son moléculas ventajosas para proporcionar el "segundo dominio" definido a continuación en el presente documento de la(s) proteína(s) biológicamente activa(s)/polipéptido(s), es decir que comprende un tramo de polipéptido que forma en condiciones fisiológicas una conformación de enrollamiento al azar y de ese modo media en una estabilidad *in vivo* y/o *in vitro* aumentada a proteína(s) o polipéptido(s) biológicamente activo(s) ("funcional/funcionales"), en particular, una semivida en plasma aumentada. El volumen hidrodinámico de una proteína funcional que está fusionada a dicho dominio de enrollamiento al azar se aumenta drásticamente tal como puede estimarse usando métodos convencionales mencionados en el presente documento. Dado que se piensa que el dominio de enrollamiento al azar no interfiere con la actividad biológica del primer dominio de la proteína biológicamente activa, se conserva esencialmente la actividad biológica mediada por la proteína de interés funcional a la que se fusiona. Además, se piensa que los polímeros de aminoácidos/polipéptidos que forman un dominio de enrollamiento al azar tal como se da a conocer en el presente documento son en gran medida biológicamente inertes, especialmente con respecto a la proteólisis en plasma sanguíneo, inmunogenicidad, punto isoelectrico/comportamiento electrostático, unión a receptores de superficie celular así como internalización, pero todavía son biodegradables, lo cual proporciona claras ventajas con respecto a polímeros sintéticos tales como PEG.

Según lo anterior, la presente invención se refiere a una proteína biológicamente activa que comprende el polipéptido de enrollamiento al azar biosintético descrito en el presente documento. Una proteína biológicamente activa/constructo de proteína de este tipo que comprende el polipéptido de enrollamiento al azar biosintético descrito

en el presente documento es una proteína heteróloga biológica activa/construido de proteína. En particular, en el presente documento también se da(n) a conocer una(s) proteína(s) heteróloga(s) biológicamente activa(s) que comprende(n) o que consiste(n) en al menos dos dominios en la que

5 (a) un primer dominio de dichos al menos dos dominios comprende o consiste en una secuencia de aminoácidos que tiene y/o que media en dicha actividad biológica; y

(b) un segundo dominio de dichos al menos dos dominios comprende o consiste en el polipéptido de enrollamiento al azar o segmento de polipéptido descrito y definido en el presente documento.

10 Debe observarse que según la presente invención que "primer dominio" y dicho "segundo dominio" se refieren a tramos de proteína que no se producen de manera natural dentro de la misma proteína o que no se espera que formen parte de la misma proteína hipotética según se codifica por una secuencia de ácido nucleico codificante (tal como un marco de lectura abierto) tal como se encuentra en la naturaleza.

15 Las definiciones y explicaciones facilitadas anteriormente en el presente documento en el contexto del polipéptido de enrollamiento al azar o segmento de polipéptido del mismo se aplican, cambiando lo que sea necesario, en el contexto de proteínas biológicamente activas que comprenden dicho polipéptido de enrollamiento al azar (o un(os) segmento(s) de polipéptido del mismo).

Preferiblemente, dicha conformación de enrollamiento al azar media en una estabilidad *in vivo* y/o *in vitro* aumentada de dicha proteína biológicamente activa, tal como la estabilidad *in vivo* y/o *in vitro* en muestras biológicas o en entornos fisiológicos.

20 Por ejemplo, en el presente documento se prevé que proteínas que comprenden un "segundo dominio" adicional definido en el presente documento que adopta una conformación de enrollamiento al azar en disolución acuosa o en condiciones fisiológicas (por ejemplo polímeros que consisten en aproximadamente 200 o aproximadamente 400 o aproximadamente 600 residuos de aminoácido y que comprenden PA#1/SEQ ID NO: 1, PA#2/SEQ ID NO: 2, PA#3/SEQ ID NO: 3, PA#4/SEQ ID NO: 4, PA#5/SEQ ID NO: 5, PA#6/SEQ ID NO: 6 y/o P1A1/SEQ ID NO: 51 como "bloques estructurales") tienen una semivida en plasma o estabilidad en suero ventajosa, incluso *in vivo* (en particular si se administran por vía intravenosa) en comparación con un control que carece de dicha conformación de enrollamiento al azar.

25 En el documento WO 2008/155134 (tal como se comentó anteriormente en el presente documento) se ha mostrado que proteínas biológicamente activas que comprenden un dominio con una secuencia de aminoácidos que adopta una conformación de enrollamiento al azar tienen una estabilidad *in vivo* y/o *in vitro* aumentada. Los dominios de enrollamiento al azar dados a conocer en el documento WO 2008/155134 consisten, en particular, en residuos de prolina, alanina y serina (PAS). La presencia de estos tres residuos se describe en este documento de la técnica anterior como requisito esencial para la formación de un enrollamiento al azar estable y soluble en disolución acuosa.

30 Tal como se comentó en la introducción anteriormente en el presente documento, el documento WO 2007/103515 describe polímeros recombinantes no estructurados que comprenden como constituyentes principales una gran variedad de aminoácidos, entre otras cosas, glicina, aspartato, alanina, serina, treonina, glutamato y prolina. Sin embargo, el término "polímero recombinante no estructurado" no tiene ningún significado claro reconocido, a diferencia de los términos "biosintético" y "enrollamiento al azar".

35 Anteriormente en el presente documento también se mencionó el documento WO 2006/081249. Este documento describe conjugados de proteína que comprenden una proteína biológicamente activa acoplada a un polipéptido que comprende de 2 a 500 unidades de una repetición de aminoácidos que tiene Gly, Asn y Gln como constituyente principal y Ser, Thr, Asp, Gln, Glu, His y Asn como constituyente minoritario. Se describe que dichos conjugados de proteína tienen una semivida en plasma o bien aumentada o bien reducida en comparación con la proteína biológicamente activa sin conjugar. Sin embargo, el documento WO 2006/081249 no proporciona ninguna enseñanza para predecir si una repetición de aminoácidos específica reduce o aumenta la semivida en plasma del conjugado. Además, el documento WO 2006/081249 no enseña ni sugiere que la semivida en plasma de proteínas pueda aumentarse cuando la proteína conjugada comprende una repetición de aminoácidos que forma una conformación de enrollamiento al azar tal como se muestra en la presente invención. Además, las repeticiones de aminoácidos dadas a conocer en el documento WO 2006/081249 comprenden al menos dos residuos seleccionados de Gly, Asn y Gln, lo cual contrasta claramente con los polipéptidos de enrollamiento al azar biosintéticos de la presente invención que comprenden una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina.

40 Sorprendentemente, en el presente documento se ha encontrado que secuencias de aminoácidos de enrollamiento al azar biosintéticas tal como se proporciona en el presente documento que, a diferencia de la técnica anterior, comprenden únicamente residuos de prolina y alanina (es decir que preferiblemente no comprenden una cantidad sustancial de ningún otro aminoácido, tampoco una cantidad sustancial de serina o nada de serina en absoluto) también forman una estructura de enrollamiento al azar útil. Esto es particularmente inesperado dada la divulgación en el documento WO 2008/155134 de proteínas de fusión con un dominio compuesto únicamente por residuos de

serina y alanina (SA), es decir en las que se omitieron residuos de prolina, demostrando que un dominio de este tipo que comprende únicamente dos tipos de aminoácidos no formaba un enrollamiento al azar, sino una estructura de lámina β . Estos dominios de serina-alanina tampoco mostraron un volumen hidrodinámico aumentado de este tipo tal como se observa con "PAS" o, en particular, con las secuencias de "P/A" tal como se proporciona en el presente documento.

Tal como se usa en el presente documento, el término "actividad biológica" describe el efecto biológico de una sustancia sobre materia viva. Por consiguiente, el término "proteína biológicamente activa" tal como se usa en el presente documento se refiere a proteínas que pueden inducir un efecto biológico sobre células vivas/organismos que se exponen a dicha proteína o polipéptido. Sin embargo, debe observarse que, en el contexto de la presente invención, el término "proteína biológicamente activa" se refiere a la proteína completa de la invención que comprende tanto una secuencia de aminoácidos que tiene y/o que media en dicha actividad biológica (dicho primer dominio) como la secuencia de aminoácidos de la invención que adopta/forma una conformación de enrollamiento al azar y que consiste únicamente en prolina y alanina (dicho segundo dominio).

Por consiguiente, los términos "secuencia de aminoácidos que tiene y/o que media en una actividad biológica" o "secuencia de aminoácidos con actividad biológica" tal como se usan en el presente documento para el "primer dominio" anteriormente definido de la proteína biológicamente activa de la invención, que media en o tiene o puede mediar en o tener la "actividad biológica" anteriormente definida. En los términos "secuencia de aminoácidos que tiene y/o que media en una actividad biológica" o "secuencia de aminoácidos con actividad biológica" también están comprendidas cualquier proteína de interés (y fragmentos funcionales de las mismas, tales como fragmentos de anticuerpo, fragmentos que comprenden dominio(s) extracelular(es) o intracelular(es) de un receptor de membrana, formas truncadas de un factor de crecimiento o citocina y similares), cuya la semivida, o bien *in vivo* o bien *in vitro*, necesita prolongarse. En una realización de esta invención, la secuencia de aminoácidos que tiene y/o que media en una actividad biológica según la presente invención puede deducirse de cualquier "proteína de interés", es decir cualquier proteína de interés farmacéutico o biológico o cualquier proteína que es útil como agente terapéutico/de diagnóstico.

Por consiguiente, las proteínas biológicamente activas pueden comprender un primer dominio que comprende una secuencia de aminoácidos biológicamente activa que se deriva de polipéptidos producidos de manera natural o polipéptidos producidos mediante tecnología de ADN recombinante. En una realización preferida, la proteína de interés puede seleccionarse del grupo que consiste en proteínas de unión/moléculas de unión, inmunoglobulinas, fragmentos de anticuerpo, proteínas de transporte, receptores de membrana, proteínas/péptidos de señalización tales como citocinas, factores de crecimiento, hormonas o enzimas y similares.

Tal como se explicó anteriormente en el presente documento, el polipéptido de enrollamiento al azar (o segmento de polipéptido) comprendido en el segundo dominio de la proteína biológicamente activa forma la conformación de enrollamiento al azar en particular en condiciones fisiológicas. Esto es particularmente relevante en el contexto de proteínas biológicamente activas que pueden formar parte de una composición farmacéutica que va a administrarse a un sujeto o paciente.

Debe observarse que el dominio de enrollamiento al azar biosintético de la invención (dicho "segundo dominio") de la proteína biológicamente activa adopta/forma/tiene de manera nativa (es decir, en condiciones fisiológicas) una conformación de enrollamiento al azar, en particular *in vivo* y cuando se administra a mamíferos o pacientes humanos que necesitan intervención médica. En cambio, en la técnica se conoce que proteínas que tienen una estructura secundaria y/o terciaria no al azar como conformación nativa tienden a adoptar una conformación de enrollamiento al azar en condiciones no fisiológicas (es decir, en condiciones de desnaturalización). Sin embargo, tales proteínas desnaturalizadas tienen características completamente diferentes en comparación con la proteína biológicamente activa que comprende el polipéptido de enrollamiento al azar de la presente invención. Por tanto, la esencia de esta invención es que la "proteína biológicamente activa" y la parte biológicamente activa de las proteínas de fusión/constructos de fusión tal como se proporciona en el presente documento también conservan su función biológica cuando se combinan y/o unen con el polipéptido de enrollamiento al azar biosintético (o segmento de polipéptido) de esta invención.

Además, el polipéptido de enrollamiento al azar (o segmento de polipéptido) conserva solubilidad en condiciones fisiológicas. Por consiguiente, también se prevé que el constructo de proteína de la presente invención (que comprende el "primer dominio" y "segundo dominio" definidos anteriormente) puede comprender el "segundo" dominio que forma/adopta un enrollamiento al azar de manera transitoria o temporal no en conformación de enrollamiento al azar, por ejemplo, cuando está en forma de una composición específica, tal como un liofilizado o una composición secada. Sin embargo, es importante que un "segundo dominio" de este tipo del constructo de proteína de la invención adopte de nuevo, por ejemplo, tras la reconstitución en tampones correspondientes (preferiblemente tampones/excipientes "fisiológicos" y/o disolventes), la conformación de enrollamiento al azar definida en el presente documento. Dicho "segundo dominio" puede mediar (si es necesario, tras una reconstitución correspondiente) en una estabilidad *in vivo* y/o *in vitro* aumentada de la proteína biológicamente activa de la invención. En el presente documento se prefiere que el "segundo dominio" tal como se define en el presente documento consista en el polipéptido de enrollamiento al azar (o segmento de polipéptido) de la presente invención.

Tal como se usa en el presente documento, el término “dominio” se refiere a cualquier región/parte de una secuencia de aminoácidos que puede adoptar de manera autónoma una estructura y/o función específica. Por consiguiente, en el contexto de la presente invención, un “dominio” puede representar un dominio funcional o un dominio estructural. Tal como se describe en el presente documento, las proteínas de la presente invención comprenden al menos un dominio/parte que tiene y/o media en una actividad biológica y al menos un dominio/parte que forma una conformación de enrollamiento al azar. Sin embargo, las proteínas de la invención también pueden consistir en más de dos dominios y pueden comprender, por ejemplo, una estructura de grupo de unión o espaciador adicional entre los dos dominios/partes definidos en el presente documento u otro dominio/parte tal como, por ejemplo, un sitio de escisión sensible a proteasa, una cola de afinidad tal como la cola de His₆ o la cola de estreptavidina, un péptido señal, péptido de retención, un péptido de direccionamiento tal como péptido de translocación de membrana o dominios efectores adicionales tales como fragmentos de anticuerpo para direccionamiento tumoral asociados con una toxina antitumoral o una enzima para activación de profármacos, etc.

En otra realización, la proteína biológicamente activa de la invención tiene un volumen hidrodinámico tal como se determina mediante filtración en gel analítica (también conocida como cromatografía de exclusión molecular, SEC) de al menos 50 kDa, preferiblemente de al menos 70 kDa, más preferiblemente de al menos 80 kDa, incluso más preferiblemente de al menos 100 kDa, de manera particularmente preferible de al menos 125 kDa y lo más preferiblemente de al menos 150 kDa. El experto en la técnica puede determinar fácilmente el volumen hidrodinámico de proteínas específicas. Anteriormente en el presente documento se han descrito métodos a modo de ejemplo en el contexto del polipéptido de enrollamiento al azar. Un experto está fácilmente en posición de adaptar tales métodos también en el contexto de la proteína biológicamente activa de la presente invención. Tal como se describe a continuación en el presente documento, el volumen hidrodinámico de las proteínas biológicamente activas de la invención que comprenden el segundo dominio anteriormente definido, es decir el dominio que comprende o consiste en el polipéptido de enrollamiento al azar proporcionado en el presente documento (o segmento del mismo), se muestra que tienen un volumen hidrodinámico inesperadamente grande con respecto al volumen hidrodinámico estimado para una proteína globular plegada correspondiente basándose en su peso molecular o número/composición de residuos de aminoácido.

Debe observarse que el primer dominio que comprende una “secuencia de aminoácidos que tiene y/o que media en una actividad biológica” también puede adoptar su actividad biológica en el contexto de o tras la asociación con otro polipéptido o secuencia de aminoácidos. Por ejemplo, el fragmento Fab de un anticuerpo tal como el del anticuerpo antitumoral Herceptin (Eigenbrot (1993) J. Mol. Biol. 229:969-995) consiste en dos cadenas de polipéptido diferentes, la cadena ligera de inmunoglobulina y un fragmento de la cadena pesada de inmunoglobulina, que pueden unirse además a través de un(os) enlace(s) disulfuro entre cadenas. Según la presente invención, puede ser suficiente con unir una de esas cadenas (por ejemplo, mediante fusión génica) al polipéptido de enrollamiento al azar (o segmento de polipéptido) mientras que la proteína biológicamente activa completa se reconstituye por medio de asociación con la otra cadena. Tal reconstitución puede lograrse, por ejemplo, mediante coexpresión de los diferentes polipéptidos (por un lado una proteína de fusión de una cadena y el polipéptido de enrollamiento al azar, por el otro lado la otra cadena) en la misma célula huésped, tal como se describe en los ejemplos adjuntos, o mediante reconstitución *in vitro*, por ejemplo, como parte de un protocolo de replegamiento.

Por consiguiente, tales proteínas (que comprenden dos cadenas de polipéptido independientes) también se consideran proteínas biológicamente activas según la presente invención. En tal caso, el primer dominio tal como se define en el presente documento puede comprender dos cadenas de polipéptido independientes que sólo están unidas de manera no covalente. Además, las cadenas independientes de la proteína biológicamente activa/dominio pueden unirse cada una al polipéptido de enrollamiento al azar (o segmento de polipéptido). Además de fragmentos de anticuerpo, hay muchas otras proteínas de interés como heterooligoméricas (por ejemplo, insulina, hemoglobina y similares) que están compuestas por varias cadenas de polipéptido asociadas y que son objeto de esta invención.

Tal como se usa en el presente documento, el término “proteína de unión” se refiere a una molécula que puede interactuar específicamente con una(s) posible(s) pareja(s) de unión de modo que puede distinguir entre dicha(s) posible(s) pareja(s) de unión y una pluralidad de diferentes moléculas como dicha(s) posible(s) pareja(s) de unión hasta tal punto que, a partir de una combinación de dicha pluralidad de diferentes moléculas como posible(s) pareja(s) de unión, sólo se une(n), o se une(n) significativamente, dicha(s) posible(s) pareja(s) de unión. En la técnica se conocen métodos para la medición de la unión entre una proteína de unión y una posible pareja de unión y pueden realizarse de manera rutinaria, por ejemplo, usando ELISA, calorimetría de titulación isotérmica, diálisis en equilibrio, ensayos de precipitación, resonancia de plasmón superficial o un aparato de Biacore. Las proteínas de unión/moléculas de unión a modo de ejemplo que son útiles en el contexto de la presente invención incluyen, pero no se limitan a, anticuerpos, fragmentos de anticuerpo tales como fragmentos Fab, fragmentos F(ab')₂, fragmentos variables de cadena sencilla (scFv), regiones variables aisladas de anticuerpos (regiones VL y/o VH), CDR, anticuerpos de dominio individual/inmunoglobulinas, compuestos peptidomiméticos derivados de CDR, lectinas, dominios de inmunoglobulina, dominios de fibronectina, dominios de proteína A, dominios SH3, dominios de repetición de anquirina, lipocalinas o diversos tipos de proteínas de unión derivadas de armazón tal como se describe, por ejemplo, en Skerra (2000) J Mol Recognit 13:167-187, Gebauer (2009) Curr Opin Chem Biol 13:245-255, Binz (2005) Nat Biotechnol 23:1257-1268 o Nelson (2009) Nat Biotechnol 27:331-337.

Otras proteínas biológicamente activas de interés a modo de ejemplo (en particular proteínas comprendidas en el primer dominio o que constituyen/son el primer dominio de la proteína biológicamente activa) que son útiles en el contexto de la presente invención incluyen, pero no se limitan a, factor estimulante de colonias de granulocitos, hormona del crecimiento humana, alfa-interferón, beta-interferón, gamma-interferón, lambda-interferón, factor de necrosis tumoral, eritropoyetina, factores de coagulación tales como factor de coagulación VIII, factor de coagulación VIIa, factor de coagulación IX, gp120/gp160, receptor de factor de necrosis tumoral I y II soluble, compuestos trombolíticos tales como reteplasa, péptidos con efectos metabólicos tales como GLP-1 o exendina-4, proteínas inmunosupresoras/inmunorreguladoras tales como antagonistas de receptor de interleucina-1 o anakinra, interleucina-2 y lipocalina asociada con gelatinasa de neutrófilos u otras lipocalinas naturales o modificadas por ingeniería o aquellas proteínas o compuestos indicados, por ejemplo, en Walsh (2003) Nat Biotechnol 21:865-870 o Walsh (2004) Eur J Pharm Biopharm 58:185-196 o indicados en bases de datos en línea tales como <http://www.biopharma.com/approvals.html> o <http://www.drugbank.ca>. Proteínas biológicamente activas adicionales (en particular proteínas comprendidas en el primer dominio o que constituyen/son el primer dominio de la proteína biológicamente activa) que pueden emplearse en el contexto de la presente invención son, entre otras cosas, hormona estimulante del folículo, glucocerebrosidasa, timosina alfa 1, glucagón, somatostatina, adenosina desaminasa, interleucina 11, Hematide, leptina, interleucina-20, subunidad alfa de receptor de interleucina-22 (IL-22ra), interleucina-22, hialuronidasa, factor de crecimiento de fibroblastos 18, factor de crecimiento de fibroblastos 21, péptido similar a glucagón 1, osteoprotegerina, proteína de unión a IL-18, factor de liberación de la hormona del crecimiento, receptor de TACI soluble, trombospondina-1, receptor de VEGF soluble Flt-1, α -galactosidasa A, antagonista de miostatina, polipéptido inhibidor gástrico, antitripsina alfa-1, muteína de IL-4, y similares. Tal como resultará evidente a partir de la divulgación en el presente documento, la presente invención también se refiere a comprender el polipéptido de prolina/alanina de enrollamiento al azar biosintético o segmento de polipéptido de prolina/alanina y moléculas farmacéutica o medicamento útiles, tales como moléculas pequeñas, péptidos o biomacromoléculas tales como proteínas, ácidos nucleicos, hidratos de carbono, vesículas lipídicas y similares, en particular proteínas farmacéutica o medicamento útiles, tales como (pero sin limitarse a) proteínas de unión/moléculas de unión, inmunoglobulinas, fragmentos de anticuerpo, proteínas de transporte, receptores de membrana, proteínas/péptidos de señalización, citocinas, factores de crecimiento, hormonas o enzimas y similares, pueden comprenderse en los constructos de fármaco definidos en el presente documento pero también pueden formar parte de la proteína heteróloga biológicamente activa definida en el presente documento que comprende o consiste en dichos al menos dos dominios definidos. En tal caso, dichas proteínas farmacéutica o medicamento útiles particulares (o fragmentos funcionales de las mismas) pueden ser el "primer dominio" de dichos al menos dos dominios que comprenden o consisten en una secuencia de aminoácidos que tiene y/o que media en dicha actividad biológica. Los fragmentos funcionales, en este contexto, son fragmentos de dichas proteínas farmacéutica o medicamento útiles que todavía pueden producir la respuesta biológica o farmacéutica deseada *in vivo* y/o *in vitro* y/o todavía tienen o median en la actividad biológica deseada.

El grupo de unión/espaciador de polipéptido anteriormente mencionado, insertado entre dichos dominios primero y segundo, comprende preferiblemente varios aminoácidos hidrófilos, unidos a péptido, que están unidos de manera covalente a ambos dominios. En aún otra realización dicho grupo de unión/espaciador de polipéptido comprende un sitio de escisión por proteasa plasmática que permite la liberación controlada de dicho primer dominio que comprende un polipéptido que tiene y/o media en una actividad biológica. Pueden identificarse grupos de unión de diferentes tipos o longitudes sin carga excesiva para obtener una actividad biológica óptima de proteínas específicas.

En una realización preferida, las proteínas biológicamente activas de la presente invención son proteínas de fusión. Una proteína de fusión tal como se describe en el presente documento puede comprender al menos un dominio que puede mediar en una actividad biológica y al menos otro dominio que comprende el polipéptido de enrollamiento al azar biosintético (o segmento de polipéptido) tal como se describe en el presente documento en un único polipéptido de múltiples dominios. De nuevo, debe observarse que la presente invención no está limitada a proteínas de fusión en las que un dominio media en una actividad biológica. Además, en el presente documento se proporcionan otras "proteínas de fusión"/"constructos de fusión" en los que una parte/dominio es o comprende el polipéptido de enrollamiento al azar/polímero de prolina/alanina de la invención y la otra parte/dominio comprende otro tramo de proteína/estructura.

En particular, en el caso de proteínas de fusión el polipéptido de enrollamiento al azar (o segmento de polipéptido) según esta invención no porta necesariamente residuos de Pro o Ala en su extremo amino o carboxilo-terminal. En una realización alternativa, la proteína biológicamente activa según la presente invención puede representar un conjugado de proteína en el que una proteína de interés o un polipéptido/tramo de polipéptido/péptido/secuencia de aminoácidos que tiene y/o que media en una actividad biológica se conjuga mediante un enlace no peptídico a una secuencia de aminoácidos que forma/adopta una conformación de enrollamiento al azar, en particular, el polipéptido de enrollamiento al azar (o segmento de polipéptido) tal como se proporciona en el presente documento y que consiste únicamente en residuos de prolina y alanina. En la técnica se conocen enlaces no peptídicos que son útiles para reticular proteínas, comprendiendo el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala) tal como se proporciona en el presente documento. Tales enlaces no peptídicos pueden incluir enlaces

disulfuro, por ejemplo entre cadenas laterales de Cys, enlaces tioéter o enlaces covalentes no peptídicos inducidos por grupos de reticulación químicos, tales como suberato de disuccinimidilo (DSS) o 4-[p-maleimidofenil]butirato de sulfosuccinimidilo (Sulfo-SMPB), grupos complejantes/quelantes de metales, así como interacciones proteína-proteína no covalentes.

5 Debe observarse que la "proteína biológicamente activa" de la presente invención también puede comprender más de una "secuencia de aminoácidos que tiene y/o que media en una actividad biológica". Además, la proteína biológicamente activa también puede comprender más de un polipéptido de enrollamiento al azar biosintético (o segmento del mismo). En el caso más sencillo, la proteína biológicamente activa consiste en dos dominios, es decir un primer dominio que comprende una secuencia de aminoácidos que tiene y/o que media en una actividad biológica y un segundo dominio que comprende el polipéptido biosintético (o segmento del mismo). Debe observarse que la presente invención no se limita a "proteínas biológica o terapéuticamente activas" unidas al polipéptido de enrollamiento al azar biosintético o segmento de polipéptido dado a conocer en el presente documento que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala). También pueden fabricarse otras proteínas o moléculas de interés, según sea relevante para otras industrias, tales como industria alimentaria o de bebidas, industria cosmética y similares, mediante los medios y métodos proporcionados en el presente documento.

El experto en la técnica es consciente de que el "dominio que comprende una secuencia de aminoácidos que tiene y/o que media en una actividad biológica" y el "segundo dominio que comprende el polipéptido de enrollamiento al azar (o segmento del mismo)" según está comprendido en las proteínas biológicamente activas de la invención pueden organizarse en un orden específico.

Por consiguiente, y en el contexto de la invención, el orden de los dominios "primero" y "segundo" definidos en el presente documento del polipéptido biológicamente activo de la invención puede disponerse en cualquier orden, mediante lo cual dicho "primer dominio" (es decir proteína de interés; "secuencia de aminoácidos que tiene y/o que media en dicha actividad biológica") está ubicado en el extremo amino (N)-terminal y dicho "segundo dominio" (es decir el dominio que comprende el polipéptido de enrollamiento al azar proporcionado en el presente documento (o segmento del mismo)) está ubicado en el extremo carboxilo (C)-terminal de la proteína biológicamente activa. Sin embargo, este orden puede invertirse, por ejemplo dicho "primer dominio" (es decir proteína de interés; "secuencia de aminoácidos que tiene y/o que media en dicha actividad biológica") está ubicado en el extremo carboxilo (C)-terminal y dicho "segundo dominio" (es decir el dominio que comprende el polipéptido de enrollamiento al azar proporcionado en el presente documento (o segmento del mismo)) está ubicado en el extremo amino (N)-terminal de la proteína biológicamente activa. Si la proteína biológicamente activa consiste únicamente en un primer dominio y un segundo dominio, el orden de dominios puede ser por consiguiente (desde el extremo N-terminal hasta el extremo C-terminal): primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)). A la inversa, el orden de dominios puede ser (desde el extremo N-terminal hasta el extremo C-terminal): segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica).

También se prevé que va a usarse más de un dominio que comprende o consiste en una secuencia de aminoácidos que tiene y/o que media en dicha actividad biológica en el contexto del constructo de proteína de la invención. Por ejemplo, la proteína biológicamente activa puede comprender dos "primeros dominios", es decir dos secuencias de aminoácidos específicas que tienen y/o que median en una actividad biológica, mediante lo cual esta actividad biológica puede ser la misma o una actividad diferente. Si la proteína biológicamente activa consiste en dos "primeros dominios" de este tipo, es decir dos secuencias de aminoácidos específicas que tienen y/o que median en una actividad biológica, y un "segundo dominio" (que comprende el polipéptido de enrollamiento al azar biosintético (o segmento del mismo)), el orden de dominios puede ser (desde el extremo N-terminal hasta el extremo C-terminal): primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica (opcionalmente diferente)).

Las mismas explicaciones se aplican en casos en los que la proteína biológicamente activa comprende más de un "segundo dominio" (es decir la proteína biológicamente activa comprende más de un polipéptido de enrollamiento al azar (o segmento del mismo)). Si la proteína biológicamente activa consiste en dos "segundos dominios" de este tipo, es decir, dos dominios que comprenden el polipéptido de enrollamiento al azar biosintético (o segmento del mismo), y un "primer dominio" (que comprende una secuencia de aminoácidos que tiene y/o que media en una actividad biológica), el orden de dominios puede ser (desde el extremo N-terminal hasta el extremo C-terminal): segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)). Si la proteína biológicamente activa comprende más de un "segundo dominio" en el presente documento se prevé que estos "segundos dominios" pueden ser idénticos o pueden ser diferentes.

Tal como se mencionó anteriormente, la proteína biológicamente activa puede comprender más de un "primer

- dominio”, es decir más de una secuencia de aminoácidos específica que tiene y/o que media en una actividad biológica y más de un “segundo dominio” (polipéptido de enrollamiento al azar biosintético (o segmento del mismo)), mediante lo cual estos “primeros dominios” pueden ser idénticos o diferentes y/o mediante lo cual dichos “segundos dominios” pueden ser idénticos o diferentes. En tales casos, pueden concebirse los siguientes órdenes de dominios a modo de ejemplo (desde el extremo N-terminal hasta el extremo C-terminal):
- 5 - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo));
- 10 - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) -- primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo));
- 15 - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica);
- 20 - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica);
- 25 - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica); o
- primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - primer dominio (secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)) - segundo dominio (polipéptido de enrollamiento al azar (o segmento del mismo)).
- 30 Para un experto en la técnica, órdenes de dominios correspondientes adicionales (en particular en casos en los que más de dos “primeros dominios” o más de dos “segundos dominios” están comprendidos en la proteína biológicamente activa) son fácilmente concebibles.
- Como con todas las realizaciones del presente polipéptido/proteína biológicamente activa de la invención, dicho(s) dominio(s) que comprende(n) una secuencia de aminoácidos que tiene y/o que media en dicha actividad biológica también puede(n) ser un fragmento biológicamente activo de una proteína dada con una función biológica deseada. Por tanto, el “segundo dominio” definido en el presente documento (preferiblemente que comprende el polipéptido de enrollamiento al azar proporcionado en el presente documento (o segmento del mismo)) también puede estar ubicado entre dos fragmentos biológicamente activos de una proteína de interés o entre fragmentos biológicamente activos de dos proteínas de interés. Todas las explicaciones y definiciones facilitadas anteriormente en el presente documento en el contexto de proteínas/polipéptidos de interés de “longitud completa” (es decir, cuando las secuencias de aminoácidos tienen/median en una determinada actividad biológica por sí solas) se aplican, cambiando lo que sea necesario, en el contexto de tales fragmentos.
- De nuevo, la invención anterior no se limita a los constructos que comprenden un “dominio” con una “función biológica activa”. Los constructos de la presente invención también pueden comprender dominios con otras funciones y no se limitan a actividades biológicas. Estas son simplemente realizaciones de la presente invención y resulta evidente para el experto en la técnica que otros constructos pueden realizarse fácilmente y usarse sin apartarse de la esencia de la presente invención. Por consiguiente, lo dicho en el presente documento en el contexto de “secuencia de aminoácidos que tiene y/o que media en una actividad biológica específica” se aplica, cambiando lo que sea necesario, para otros constructos, por ejemplo constructos que van a usarse en otros campos técnicos, tales como la cosmética, procesamiento de alimentos, productos lácteos, producción de papel, etc. Tal como se mencionó anteriormente en el presente documento, los polipéptidos/polímeros biosintéticos de la presente invención también puede usarse para unirse por ejemplo a moléculas pequeñas y similares.
- De nuevo, debe indicarse que el término “secuencia de aminoácidos que tiene y/o que media en una primera actividad biológica” no se limita a polipéptidos de longitud completa que tienen y/o median en dicha actividad o función biológica, sino también a fragmentos biológica y/o farmacológicamente activos de los mismos. Especialmente, pero no únicamente, en un contexto en el que dos o más “primeros dominios” tal como se define en el presente documento están comprendidos en la “proteína biológicamente activa” de la invención, también se prevé que estos “primeros dominios” son o representan diferentes partes de un complejo de proteína o fragmentos de tales

partes de complejo de proteína.

Tal como se muestra a modo de ejemplo a continuación en el presente documento, las proteínas biológicamente activas de la invención que se modifican para comprender un polipéptido de enrollamiento al azar muestran sorprendentemente una estabilidad *in vivo* y/o *in vitro* aumentada en comparación con proteínas biológicamente activas sin modificar que carecen de dicho dominio de enrollamiento al azar. Tal como se usa en el presente documento, el término “estabilidad *in vivo*” se refiere a la capacidad de una sustancia específica que se administra al organismo vivo de permanecer biológicamente disponible y biológicamente activa. *In vivo*, una sustancia puede retirarse y/o inactivarse debido a excreción, filtración renal, captación hepática, agregación, degradación y/u otros procesos metabólicos. Por consiguiente, en el contexto de la presente invención las proteínas biológicamente activas que tienen una estabilidad *in vivo* aumentada pueden excretarse menos rápidamente a través de los riñones (orina) o a través de las heces y/o pueden ser más estables frente a proteólisis, en particular frente a proteólisis *in vivo* en líquidos biológicos, tales como sangre, líquido cefalorraquídeo, líquido peritoneal y linfa. En una realización, la estabilidad *in vivo* aumentada de una proteína biológicamente activa se manifiesta en una semivida en plasma prolongada de dicha proteína biológicamente activa. En particular, la estabilidad *in vivo* aumentada de la proteína biológicamente activa es una semivida en plasma prolongada de dicha proteína biológicamente activa que comprende dicho segundo dominio en comparación con la proteína biológicamente activa que carece del segundo dominio.

En la técnica se conocen métodos para medir la estabilidad *in vivo* de proteínas biológicamente activas. Tal como se muestra a modo de ejemplo a continuación en el presente documento, proteínas biológicamente activas pueden detectarse específicamente en el plasma sanguíneo usando técnicas de inmunotransferencia de tipo Western o ensayo de inmunoabsorción asociado a enzimas (ELISA). Sin embargo, el experto en la técnica es consciente de que pueden emplearse otros métodos para medir específicamente la semivida en plasma de una proteína de interés. Tales métodos incluyen, pero no se limitan a, la detección física de una proteína de interés marcada radiactivamente. En la técnica se conocen métodos para el marcaje radiactivo de proteínas, por ejemplo, mediante radiyodación.

El término “estabilidad *in vitro* aumentada” tal como se usa en el presente documento se refiere a la capacidad de una proteína biológicamente activa de resistir la degradación y/o agregación y de conservar su actividad biológica original en un entorno *in vitro*. En la técnica se conocen bien métodos para medir la actividad biológica de proteínas biológicamente activas.

Además, se proporciona un conjugado de fármaco que comprende el polipéptido de enrollamiento al azar o segmento de polipéptido descrito y definido en el presente documento y un fármaco de molécula pequeña que se conjuga a dicho polipéptido de enrollamiento al azar o segmento de polipéptido. Ejemplos no limitativos de las moléculas pequeñas son doxorubicina, caliqueamicina, camptotecina, fumagilina, dexametasona, geldanamicina, paclitaxel, docetaxel, irinotecán, ciclosporina, buprenorfina, naltrexona, naloxona, vindesina, vancomicina, risperidona, aripiprazol, palonosetrón, granisetron, citarabina, NX1838, leuprolida, goserelina, busserelina, octreotida, teduglutida, cilengitida, abarelix, enfuvirtida, ghrelina y derivados, tubulinas, derivados de platino, inhibidores de integrina alfa 4, ácidos nucleicos antisentido, ARN de interferencia pequeños, microARN, esteroides, aptámeros de ADN o ARN, péptidos, compuestos peptidomiméticos. En general, la presente invención también se refiere a constructos de fármaco que comprenden el polipéptido de enrollamiento al azar o segmento de polipéptido definido en el presente documento y en particular moléculas farmacéutica o medicamento útiles, tales como moléculas pequeñas, péptidos o biomacromoléculas tales como proteínas, ácidos nucleicos, hidratos de carbono, vesículas lipídicas y similares. En la parte experimental ilustrativa adjunta (véase, por ejemplo, el ejemplo 22) se proporciona la generación satisfactoria de constructos/conjugados de la presente invención, también constructos en los que se han conjugado “moléculas químicas pequeñas” al polipéptido de enrollamiento al azar dado a conocer en el presente documento. Por tanto, las presentes figuras e información experimental en las leyendas de figuras correspondientes proporcionan ejemplos ilustrativos, en los que los conjugados de fármaco dados a conocer en el presente documento comprenden (i) un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala), y (ii) una molécula pequeña, que, como ilustración, se selecciona de digoxigenina y fluoresceína. Debe observarse que esto no son sólo ejemplos teóricos. La fluoresceína o derivados de fluoresceína se usan habitualmente como compuestos de diagnóstico, y se comercializan disoluciones de fluoresceína médicas con los nombres comerciales Fluorescite®, AK-FLUOR® o Fluress®. Tales compuestos pueden beneficiarse ciertamente de los medios y métodos proporcionados en el presente documento. La digoxigenina forma la parte esteroide de digoxina, un metabolito vegetal secundario bien conocido con función cardioactiva que además contiene tres azúcares digitoxosa. La digoxina, y en menor medida el compuesto estrechamente relacionado digitoxina, se usan ampliamente para el tratamiento de taquiarritmias ventriculares e insuficiencia cardiaca congestiva (Hauptman (1999) *Circulation* 99: 1265-1270). Todos los esteroides cardioactivos son inhibidores potentes y altamente específicos de la Na⁺/K⁺-ATPasa ubicada en la membrana plasmática celular, ejerciendo por tanto efectos inotrópicos positivos o simpaticolíticos.

Las definiciones y explicaciones facilitadas anteriormente en el presente documento en el contexto del polipéptido de enrollamiento al azar o segmento de polipéptido del mismo se aplican, cambiando lo que sea necesario, en el

contexto de conjugado de fármaco que comprende el polipéptido de enrollamiento al azar (o segmento de polipéptido del mismo) y un fármaco seleccionado del grupo que consiste en (a) una proteína biológicamente activa o un polipéptido que comprende o que es una secuencia de aminoácidos que tiene o media en una actividad biológica y (b) un fármaco de molécula pequeña.

5 El polímero de aminoácidos que forma una conformación de enrollamiento al azar/el polipéptido de enrollamiento al azar (o segmento del mismo) tal como se define y se proporciona en el presente documento puede conjugarse a una molécula pequeña/fármaco de molécula pequeña. Por medio de esto, puede aumentarse la semivida en plasma y/o solubilidad de la molécula pequeña/fármaco de molécula pequeña, puede reducirse la toxicidad inespecífica, y la exposición prolongada de fármaco activo a células o estructuras diana en el organismo puede dar como resultado farmacodinamia potenciada.

10 Es posible una conjugación específica del sitio del extremo N-terminal del polipéptido de enrollamiento al azar con un derivado de fármaco activado, por ejemplo tal como derivado de éster de N-hidroxisuccinimida (NHS) (Hermanson (1996) Bioconjugate Techniques, Academic Press, San Diego, CA). Generalmente, el grupo amino N-terminal puede acoplarse químicamente con una amplia variedad de grupos funcionales tales como aldehídos y cetonas (para formar bases de Schiff, que pueden reducirse para dar aminas usando borohidruro de sodio o cianoborohidruro de sodio, por ejemplo) o para dar derivados de ácido carbónico activados (anhídridos, cloruros, ésteres y similares, para formar amidas) o para dar otros productos químicos reactivos tales como isocianatos, isotiocianatos, cloruros de sulfonilo, etc. Además, el extremo N-terminal del polímero de aminoácidos/polipéptido puede modificarse en primer lugar con un grupo protector adecuado, por ejemplo un grupo acetilo, un grupo BOC o un grupo FMOC (Jakubke (1996) Peptide. Spektrum Akademischer Verlag, Heidelberg, Alemania). Además, el extremo amino-terminal puede protegerse mediante un grupo piroglutamilo, que puede formarse a partir de un residuo de aminoácido de Gln codificado que precede al polipéptido de Pro/Ala o segmento de polipéptido. Tras la activación del grupo carboxilato C-terminal, por ejemplo usando los reactivos comunes EDC (N-(3-dimetilaminopropil)-N'-etilcarbodiimida) y NHS, puede lograrse el acoplamiento específico del sitio al extremo C-terminal del polipéptido de enrollamiento al azar protegido si el fármaco porta un grupo amino libre, por ejemplo.

15 Alternativamente, el extremo N-terminal o el extremo C-terminal del polímero de aminoácidos que forma una conformación de enrollamiento al azar/el polipéptido de enrollamiento al azar puede modificarse con un reactivo de unión comercialmente disponible que proporciona un grupo maleimida, permitiendo por tanto el acoplamiento químico a un grupo tiol como parte de la molécula de fármaco. De esta manera pueden obtenerse fácilmente conjugados de fármaco uniformes. Pueden usarse técnicas similares, que se conocen bien en la técnica (Hermanson (1996), citado anteriormente), para acoplar el polipéptido de enrollamiento al azar a un péptido o incluso a un fármaco de proteína. Tales péptidos o proteínas pueden prepararse fácilmente portando una cadena lateral de Lys o Cys, que permite su acoplamiento *in vitro* al polímero de aminoácidos que forma una conformación de enrollamiento al azar a través de grupos activos de maleimida o éster de NHS. Generalmente, pueden prepararse conjugados de fármaco similares con proteínas de fusión que comprenden el polipéptido de enrollamiento al azar (o segmento del mismo). Sin embargo, y tal como se ilustra en los ejemplos y figuras adjuntos, la presente invención también proporciona la preparación de polipéptidos de enrollamiento al azar o segmentos de polipéptido de enrollamiento al azar según está comprendido en los conjugados innovadores de la presente invención.

20 Como alternativa a una única conjugación específica del sitio, el polipéptido de enrollamiento al azar puede equiparse con cadenas laterales adicionales, en el extremo N o C-terminal o de manera interna, adecuadas para su modificación química, tales como residuos de lisina con sus grupos ϵ -amino, residuos de cisteína con sus grupos tiol, o incluso aminoácidos no naturales, permitiendo la conjugación de múltiples moléculas pequeñas usando, por ejemplo, grupos activos de maleimida o éster de NHS.

25 Aparte de la conjugación estable, puede unirse transitoriamente un profármaco al polipéptido de enrollamiento al azar. La unión puede diseñarse para escindirse *in vivo*, de una manera predecible, o bien mediante un mecanismo enzimático o bien mediante hidrólisis lenta iniciada a pH fisiológico de manera similar, por ejemplo, al agente antitumoral escasamente soluble camptotecina que se conjugó a un polímero de PEG, logrando por tanto una biodistribución aumentada, toxicidad reducida, eficacia potenciada y acumulación tumoral (Conover (1998) Cancer Chemother Pharmacol, 42:407-414). Ejemplos de profármacos adicionales son agentes quimioterápicos tales como docetaxel (Liu (2008) J Pharm Sci. 97:3274-3290), doxorubicina (Veronese (2005) Bioconjugate Chem. 16: 775-784) o paclitaxel (Greenwald (2001) J Control Release 74:159-171).

30 Además, la molécula pequeña puede acoplarse a una proteína de fusión que comprende el polímero de aminoácidos/polipéptido genéticamente fusionada a un dominio de direccionamiento, por ejemplo un fragmento de anticuerpo, dando por tanto como resultado un suministro específico del fármaco de molécula pequeña. En este último caso, pueden generarse fácilmente inmunotoxinas mediante conjugación con una molécula pequeña citotóxica si se dirige el dominio de direccionamiento contra un receptor de superficie celular que experimenta internalización, por ejemplo.

35 Por tanto, según lo anterior, la presente invención también se refiere a proporcionar el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido dado a conocer en el presente documento que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina para su

acoplamiento suplementario y adicional con otros compuestos de elección. Dicho acoplamiento suplementario y/o adicional puede ser y/o puede comprender el primer acoplamiento de dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido de enrollamiento al azar biosintético con o a otro compuesto.

5 En otra realización, la presente invención se refiere a moléculas de ácido nucleico que codifican para los polipéptidos de enrollamiento al azar (o segmentos de los mismos) o proteínas biológicamente activas tal como se describe en el presente documento. Por consiguiente, dicha molécula de ácido nucleico puede comprender una secuencia de ácido nucleico que codifica para un polipéptido que tiene actividad biológica y una secuencia de ácido nucleico que codifica para el polipéptido de enrollamiento al azar (o segmento del mismo). Se pretende que el término "molécula de ácido nucleico", tal como se usa en el presente documento, incluya moléculas de ácido nucleico tales como moléculas de ADN y moléculas de ARN. Dicha molécula de ácido nucleico puede ser monocatenaria o bicatenaria, pero preferiblemente es ADN bicatenario. Preferiblemente, dicha molécula de ácido nucleico puede estar comprendida en un vector.

15 Por consiguiente, la presente divulgación también se refiere a una molécula de ácido nucleico que codifica para el polipéptido de enrollamiento al azar o segmento de polipéptido según está comprendido en los conjugados proporcionados en el presente documento, tal como un conjugado de fármaco tal como se define en el presente documento. La presente invención se refiere a una molécula de ácido nucleico que codifica para un conjugado de proteína que comprende una proteína biológicamente activa tal como se definió anteriormente y que comprende, adicionalmente, un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala).

20 En una realización se proporciona una molécula de ácido nucleico que codifica para un conjugado de fármaco tal como se definió anteriormente, comprendiendo dicha molécula de ácido nucleico

(i) una secuencia de ácido nucleico que codifica para una secuencia de aminoácidos y/o líder traducida;

25 (ii) una secuencia de ácido nucleico que codifica para un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala);

(iii) una secuencia de ácido nucleico que codifica para una proteína biológicamente activa o dicho polipéptido que comprende o que es una secuencia de aminoácidos que tiene o que media en una actividad biológica; y

30 (iv) una secuencia de ácido nucleico que representa o es un codón de terminación de la traducción.

La "secuencia de aminoácidos y/o líder traducida" anteriormente mencionada en (i) puede ser, por ejemplo, la "M" de iniciación, es decir una metionina derivada de un codón de iniciación correspondiente, también puede comprender secuencias no traducidas de un ARNm tal como la secuencia en 5' hasta un codón de iniciación que comprende, por ejemplo, un sitio de unión a ribosoma. Sin embargo, una secuencia de este tipo también puede comprender secuencias líder y/o señal clásicas, por ejemplo, para la secreción de una proteína expresada en el periplasma o en un medio de cultivo. Péptidos señal procariotas son, por ejemplo, OmpA, MalE, PhoA, DsbA, pelB, Afa, npr, STII. Péptidos señal eucariotas son, por ejemplo, secuencia señal de melitina de abejas, secuencia señal de glicoproteína ácida gp67, secuencia señal de IgM de ratón, secuencia señal de hGH.

40 Anteriormente en el presente documento se han proporcionado proteínas biológicamente activas o polipéptidos que comprenden o que son una secuencia de aminoácidos que tiene o que media en una actividad biológica así como otras proteínas de interés, tales como una proteína que va a emplearse en otros sectores industriales. Dichas realizaciones se aplican, cambiando lo que sea necesario, para la molécula de ácido nucleico (partes/segmentos (iii)) tal como se ilustró anteriormente en el presente documento.

45 En la técnica se conocen bien codones de terminación de la traducción que van a emplearse en la molécula de ácido nucleico proporcionada en el presente documento y son, por ejemplo codones UAA, UAG o UGA.

En una realización de la molécula de ácido nucleico tal como se proporcionó anteriormente en el presente documento dichas partes/segmentos de molécula de ácido nucleico (ii) y (iii) intercambian su posición en dicha molécula de ácido nucleico que codifica para un conjugado de fármaco. Una molécula de ácido nucleico de este tipo comprenderá el siguiente orden de partes/segmentos:

50 (i) una secuencia de ácido nucleico que codifica para una secuencia de aminoácidos y/o líder traducida;

(ii) una secuencia de ácido nucleico que codifica para una proteína biológicamente activa o dicho polipéptido que comprende o que es una secuencia de aminoácidos que tiene o que media en una actividad biológica;

(iii) una secuencia de ácido nucleico que codifica para un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de

aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala); y

(iv) una secuencia de ácido nucleico que representa o es un codón de terminación de la traducción.

5 Las moléculas de ácido nucleico tal como se proporcionó anteriormente en el presente documento también pueden comprender, opcionalmente, entre las partes/segmentos (i) y (ii) y/o entre las partes/segmentos (ii) y (iii), un sitio de escisión por proteasa y/o química y/o un sitio de reconocimiento. Tales sitios de escisión química se conocen bien en la técnica, y comprenden por ejemplo secuencias de aminoácidos individuales específicas (véase, por ejemplo Lottspeich y Engels (Hrsg.) (2006) Bioanalytik. 2. Auflage. Spektrum Akademischer Verlag, Elsevier, Múnich, Alemania). Por ejemplo, bromuro de cianógeno o cloruro de cianógeno escinde el enlace peptídico tras un residuo de Met; hidroxilamina escinde el enlace asparaginil-glicilo; ácido fórmico escinde Asp-Pro; 2-(2'-nitrofenilsulfenil)-3-
10 metil-3-bromoinolenina, ácido 2-yodosobenzoico o N-clorosuccinimida tras Trp; ácido 2-nitro-5-tiocianatobenzoico tras Cys. También se prevé y es posible que el residuo que precede al polipéptido de Pro/Ala o segmento de polipéptido puede sustituirse por Met mediante mutagénesis dirigida al sitio y entonces puede escindirse la proteína de fusión resultante mediante BrCN. De manera similar, pueden introducirse otras secuencias de aminoácidos que comprenden sitio de escisión en la proteína de fusión recombinante o su ácido nucleico codificante mediante mutagénesis dirigida al sitio.

En la técnica también se conocen sitios de reconocimiento/escisión por proteasa útiles. Estos comprenden, pero no se limitan a: tripsina, quimotripsina, enterocinasa, proteasa de grabado del tabaco (TEV), proteasa PreScission, proteasa 3C de HRV, SUMO proteasa, sortasa A, granzima B, furina, trombina, factor Xa o inteína autoescindible. El factor Xa hidroliza el enlace peptídico en el extremo C-terminal de la secuencia de aminoácidos IleGluGlyArg, que puede insertarse entre la pareja de fusión N-terminal y el polipéptido de Pro/Ala o segmento de polipéptido. Un método particularmente sencillo para lograr la escisión proteolítica será mediante inserción o sustitución de una cadena lateral de Lys o Arg en el extremo N-terminal del polipéptido de Pro/Ala o segmento de polipéptido seguido por digestión con tripsina, que no escinde dentro del polipéptido de Pro/Ala o segmento de polipéptido siempre que se eviten cadenas laterales de Lys o Arg internas. Sitios de reconocimiento ilustrativos son, sin ser limitativos, D-D-D-D-K (enterocinasa), ENLYFQ(G/S) (proteasa de TEV), I-(E/D)-G-R (factor Xa), L-E-V-L-F-Q-G-P (3C de HRV), R-X-(K/R)-R (furina), LPXTG (sortasa A), L-V-P-R-G (trombina) o I-E-X-D-X-G (granzima B).

30 Tal como resulta evidente a partir de la divulgación anteriormente en el presente documento, la presente invención proporciona polipéptidos de enrollamiento al azar biosintéticos producidos de manera recombinante y segmentos de polipéptido que pueden conjugarse con polipéptidos o moléculas pequeñas farmacéuticamente activos.

Se describe un ácido nucleico que codifica para un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala), comprendiendo dicha molécula de ácido nucleico

35 (i) una secuencia de ácido nucleico que codifica para una secuencia de aminoácidos y/o líder traducida;

(ii) una secuencia de ácido nucleico que codifica para dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina; y

(iii) una secuencia de ácido nucleico que representa o es un codón de terminación de la traducción.

40 Una molécula de ácido nucleico de este tipo puede comprender, opcionalmente, entre (i) y (ii) un sitio de escisión por proteasa y/o química y/o un sitio de reconocimiento.

También para esta molécula de ácido nucleico, las realizaciones proporcionadas anteriormente en el presente documento en el contexto de las dos primeras moléculas de ácido nucleico descritas (es decir un sitio de escisión por proteasa y/o química y/o un reconocimiento) se aplican en este caso cambiando lo que sea necesario.

45 Las secuencias señal útiles e ilustrativas que van a emplearse en el contexto de esta invención comprenden, pero no se limitan a, secuencias procariotas tales como: OmpA, MalE, PhoA, DsbA, pelB, Afa, npr, STII o secuencias eucariotas tales como: secuencia señal de melitina de abejas, secuencia señal de glicoproteína ácida gp67, secuencia señal de IgM de ratón, secuencia señal de hGH.

50 En particular la molécula de ácido nucleico que codifica para el polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina de la presente invención es útil en métodos tal como también se proporciona a continuación en el presente documento y tal como se ilustra en los ejemplos y figuras adjuntos. Un polipéptido de enrollamiento al azar o segmento de polipéptido de enrollamiento al azar expresado de este tipo puede aislarse, por ejemplo, de células huésped que expresan un polipéptido de enrollamiento al azar o segmento de polipéptido de enrollamiento al azar de este tipo. Tales células huésped pueden ser células transfectadas, por ejemplo con un vector tal como se proporciona en el presente documento.

55

Por tanto, se prevé transfectar células con la molécula de ácido nucleico o vectores tal como se describe en el presente documento. En una realización adicional, la presente invención se refiere a moléculas de ácido nucleico que, tras la expresión, codifican para el polipéptido de enrollamiento al azar (o segmento del mismo) o proteínas biológicamente activas de la invención. Sin embargo, en una realización adicional, la presente invención se refiere a moléculas de ácido nucleico que, tras la expresión, codifican para los polipéptidos dados a conocer en el presente documento que, en su totalidad o en parte, forman/adoptan una conformación de enrollamiento al azar en disolución acuosa o en condiciones fisiológicas. Dichas moléculas de ácido nucleico pueden fusionarse a secuencias de control de la expresión adecuadas conocidas en la técnica para garantizar una transcripción y traducción apropiadas del polipéptido así como secuencias señal para garantizar una secreción celular o direccionamiento a orgánulos. Tales vectores pueden comprender genes adicionales tales como genes marcadores que permiten la selección de dicho vector en una célula huésped adecuada y en condiciones adecuadas.

Preferiblemente, la molécula de ácido nucleico de la invención está comprendida en un vector recombinante en el que una molécula de ácido nucleico que codifica para la proteína biológicamente activa descrita en el presente documento está operativamente unida a las secuencias de control de la expresión permitiendo la expresión en células procariotas o eucariotas. La expresión de dicha molécula de ácido nucleico comprende la transcripción de la molécula de ácido nucleico para dar un ARNm que puede traducirse. Los elementos reguladores que permiten la expresión en células huésped procariotas comprenden, por ejemplo, los promotores lambda PL, lac, trp, tac, ara, phoA, tet o T7 en *E. coli*. Los expertos en la técnica conocen bien posibles elementos reguladores que garantizan la expresión en células eucariotas, preferiblemente células de mamífero o levadura. Habitualmente comprenden secuencias reguladoras que garantizan el inicio de la transcripción y opcionalmente señales de poli-A que realizan la terminación de la transcripción y estabilización del transcrito. Los elementos reguladores adicionales pueden incluir potenciadores de la transcripción así como de la traducción, y/o regiones de promotor heterólogas o asociadas de manera natural. Ejemplos de elementos reguladores que permiten la expresión en células huésped eucariotas son los promotores AOX1 o GAL1 en levadura o los promotores de CMV, SV40, RSV (virus del sarcoma de Rous), potenciador de CMV, potenciador de SV40 o un intrón de globina en células de mamífero y de otros animales. Aparte de los elementos que son responsables del inicio de la transcripción, tales elementos reguladores también pueden comprender señales de terminación de la transcripción, tales como el sitio de poli-A de SV40 o el sitio de poli-A de tk, en el sentido de 3' desde la región codificante.

Pueden usarse métodos que conocen bien los expertos en la técnica para construir vectores recombinantes (véanse, por ejemplo, las técnicas descritas en Sambrook (1989), *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory NY y Ausubel (1989), *Current Protocols in Molecular Biology*, Green Publishing Associates and Wiley Interscience, NY). En este contexto, en la técnica se conocen vectores de expresión adecuados tales como vector de expresión de ADNC de Okayama-Berg pcDV1 (Pharmacia), pCDM8, pRc/CMV, pcDNA1, pcDNA3, pPICZalfa A (Invitrogen) o pSPORT1 (GIBCO BRL). Además, dependiendo del sistema de expresión que se usa, pueden añadirse secuencias líder, que pueden dirigir el polipéptido a un compartimento celular o secretarlo en el medio de cultivo, a la secuencia codificante de la molécula de ácido nucleico de la invención.

La presente invención también se refiere a vectores, particularmente plásmidos, cósmidos, virus y bacteriófagos, que se emplean convencionalmente en ingeniería genética, que comprenden una molécula de ácido nucleico que codifica para el polipéptido de enrollamiento al azar (o segmento del mismo) o la proteína biológicamente activa de la invención. Preferiblemente, dicho vector es un vector de expresión y/o un vector de direccionamiento o transferencia génica. Pueden usarse vectores de expresión derivados de virus tales como retrovirus, virus vaccinia, virus adenoasociado, virus del herpes o virus del papiloma bovino, para el suministro de los polinucleótidos o vector de la invención al interior de poblaciones celulares seleccionadas como diana.

Los vectores que contienen las moléculas de ácido nucleico de la invención pueden transfectarse al interior de la célula huésped mediante métodos bien conocidos, que varían dependiendo del tipo de célula. Por consiguiente, la invención se refiere además a una célula que comprende dicha molécula de ácido nucleico o dicho vector. Tales métodos incluyen, por ejemplo, las técnicas descritas en Sambrook (1989), citado anteriormente y Ausubel (1989), citado anteriormente. Por consiguiente, la electroporación o la transfección con cloruro de calcio se usan habitualmente para células procariotas, mientras que puede usarse la electroporación o el tratamiento con fosfato de calcio para otros huéspedes celulares (Sambrook (1989), citado anteriormente). Como alternativa adicional, las moléculas de ácido nucleico y vectores de la invención pueden reconstituirse para dar liposomas para su suministro a células diana. La molécula de ácido nucleico o vector de la invención que está presente en la célula huésped puede o bien integrarse en el genoma de la célula huésped o bien puede mantenerse de manera extracromosómica. Por consiguiente, la presente invención también se refiere a una célula huésped que comprende la molécula de ácido nucleico y/o el vector de esta invención. En la técnica se conocen bien células huésped para la expresión de polipéptidos y comprenden células procariotas así como células eucariotas, por ejemplo células de *E. coli*, células de levadura, células de invertebrados, células CHO, células CHO-K1, células HEK 293, células Hela, células de mono COS-1, células de melanoma tales como células de Bowes, células L-929 de ratón, líneas celulares 3T3 derivadas de ratones suizos, Balb-c o NIH, líneas celulares de hámster BHK o HaK y similares.

En un aspecto adicional, la presente invención comprende métodos para la preparación de los conjugados de la presente invención así como el polipéptido de enrollamiento al azar biosintético (o segmento del mismo) o proteínas biológicamente activas proporcionadas en el presente documento y que comprenden cultivar la célula (huésped) de

esta invención y aislar dicho polipéptido de enrollamiento al azar (o segmento del mismo) o el conjugado o una proteína biológicamente activa a partir del cultivo tal como se describe en el presente documento. En general, el polipéptido de enrollamiento al azar de la invención (o segmento del mismo), el conjugado o proteína biológicamente activa que comprende un dominio de enrollamiento al azar puede producirse mediante tecnología de ADN recombinante, por ejemplo cultivando una célula que comprende la molécula de ácido nucleico descrita o vectores que codifican para la proteína biológicamente activa de la invención o polipéptido de enrollamiento al azar (o segmento del mismo) y aislando dicha proteína/polipéptido a partir del cultivo. La proteína biológicamente activa de la invención o polipéptido de enrollamiento al azar (o segmento del mismo) puede producirse en cualquier sistema de cultivo celular adecuado incluyendo células procariotas, por ejemplo *E. coli* BL21, KS272 o JM83, o células eucariotas, por ejemplo cepa X-33 de levadura *Pichia pastoris* o células CHO. Pueden obtenerse líneas celulares adecuadas adicionales conocidas en la técnica a partir de depositarios de líneas celulares tales como la Colección americana de cultivos tipo (ATCC).

Se pretende que el término "procariota" incluya células bacterianas mientras que se pretende que el término "eucariota" incluya células de levadura, plantas superiores, insectos y mamíferos. Los huéspedes transformados pueden hacerse crecer en fermentadores y cultivarse según técnicas conocidas en la técnica para lograr un crecimiento celular óptimo. En una realización adicional, la presente invención se refiere a un procedimiento para la preparación de un polipéptido de enrollamiento al azar (o segmento del mismo) o una proteína biológicamente activa descrita anteriormente que comprende cultivar una célula de la invención en condiciones adecuadas para la expresión de la proteína biológicamente activa o polipéptido de enrollamiento al azar (o segmento del mismo) y aislar dicha proteína/polipéptido a partir de la célula o del medio de cultivo.

Preferiblemente el polipéptido de enrollamiento al azar (o segmento del mismo) en sí mismo de la presente invención no comprende ningún grupo químicamente reactivo, excepto, posiblemente, por un grupo amino primario (o, en el caso de prolina, secundario) N-terminal y un grupo carboxilato en el extremo C-terminal del polímero. Sin embargo, resulta evidente para el experto en la técnica que el polipéptido/polímero de enrollamiento al azar biosintético tal como se proporciona en el presente documento puede comprender un grupo químicamente reactivo, por ejemplo cuando dicho polipéptido de enrollamiento al azar/polímero forma parte de una "proteína de fusión"/"constructo de fusión". Tal como también se describió anteriormente, el polipéptido de enrollamiento al azar biosintético (o segmento del mismo) puede prepararse mediante expresión recombinante en una célula transformada de varias maneras según métodos bien conocidos por el experto en la técnica, por ejemplo: (i) expresión directa en el citoplasma con la ayuda de un codón de iniciación/residuo de Met N-terminal; (ii) secreción mediante un péptido señal N-terminal, por ejemplo OmpA, PhoA (Monteilhet (1993) *Gene*. 1993 125:223-228), melitina (Tessier (1991) *Gene* 98: 177-183), interleucina 2 (Zhang (2005) *J Gene Med* 7: 354-365), hGH (Pecceu (1991) *Gene* 97(2):253-258) y similares, seguido por escisión intracelular dando como resultado el extremo N-terminal maduro, tal como Ala o Pro; (iii) expresión como proteína de fusión con otra proteína soluble, por ejemplo, proteína de unión a maltosa en el extremo N-terminal y con un sitio de escisión por proteasa intercalado (Kapust y Waugh (2000) *Protein Expr. Purif.* 19:312-318), seguido por escisión con proteasa específica *in vitro* o *in vivo*, liberando por tanto el polímero de aminoácidos/polipéptido con su extremo N-terminal maduro, tal como Ala o Pro. Otra pareja de fusión adecuada es la proteína SUMO, que puede escindirse mediante SUMO proteasa, tal como se describe en los ejemplos 20 y 21. Las parejas de fusión adicionales incluyen, sin limitación, glutatión-S-transferasa, tiorredoxina, un dominio de unión a celulosa, un dominio de unión a albúmina, una proteína fluorescente (tal como GFP), proteína A, proteína G, una inteína y similares (Malhotra (2009) *Methods Enzymol.* 463:239-258).

Tal como se explicó anteriormente, los polipéptidos de enrollamiento al azar (o segmentos de polipéptido)/polímeros descritos consisten predominantemente en residuos de alanina y prolina, mientras que serina, treonina o asparagina, que se requieren para la O o N-glicosilación, están preferiblemente ausentes. Por tanto, la producción del propio polipéptido o de una proteína biológicamente activa que comprende el polipéptido de enrollamiento al azar (o segmento de polipéptido del mismo) o, generalmente, una proteína de fusión que comprende el polipéptido de enrollamiento al azar (o segmento de polipéptido del mismo) puede dar sorprendentemente como resultado un producto monodispersado preferiblemente carente de modificaciones tras la traducción dentro de la secuencia de Pro-Ala. Esto es una ventaja para la producción de proteínas recombinantes en células eucariotas, tales como células de ovario de hámster chino (CHO) o levadura, que se eligen con frecuencia para la biosíntesis de proteínas complejas. Por ejemplo, se ha usado levadura para la producción de proteínas terapéuticas aprobadas tales como insulina, factor estimulante de colonias de granulocitos-macrófagos, factor de crecimiento derivado de plaquetas o hirudina (Gerngross (2004) *Nat. Biotechnol.* 22:1409-1414). Las células CHO han servido para la producción de proteínas terapéuticas tales como factor de coagulación IX, interferón β -1a, tenecteplasa (Chu (2001) *Curr. Opin. Biotechnol.* 12:180-187) o gonadotropinas, en las que el glicocomponente puede influir positivamente sobre varios aspectos tales como actividad funcional, plegamiento, dimerización, secreción así como interacción con receptor, transducción de señales y aclaramiento metabólico (Walsh (2006) *Nat. Biotechnol.* 24:-1241-1252). Por consiguiente, la preparación de los constructos, polipéptidos de enrollamiento al azar y conjugados de la invención en sistemas de expresión eucariotas también se da a conocer en el contexto de la presente invención.

Con los medios y métodos proporcionados en el presente documento ahora es posible fabricar y proporcionar los conjugados y moléculas dados a conocer en el presente documento que comprenden (i) un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que

consiste únicamente en residuos de aminoácido de prolina y alanina y (ii) una molécula de interés adicional, tal como una proteína útil, un segmento de proteína o una molécula pequeña. Por tanto, la presente invención también proporciona métodos para la preparación o fabricación de tales conjugados así como de polipéptidos de enrollamiento al azar biosintéticos y/o moléculas o conjugados que comprenden los mismos. Por consiguiente, la presente invención también proporciona un método para la preparación y/o fabricación de un polipéptido de enrollamiento al azar o un segmento de polipéptido de enrollamiento al azar según está comprendido en los conjugados de fármaco. También se proporcionan métodos para la preparación y/o fabricación de la proteína biológicamente activa o conjugado que comprende el polipéptido de enrollamiento al azar o el segmento de polipéptido de enrollamiento al azar. Además, se proporcionan métodos para la preparación y/o fabricación y/o para la preparación y/o fabricación de un polipéptido que comprende o que es una secuencia de aminoácidos que tiene o que media en una actividad biológica y que comprende adicionalmente dicho polipéptido de enrollamiento al azar o segmento de polipéptido de enrollamiento al azar. Estos métodos comprenden en particular (como una etapa) el cultivo de la célula (huésped) tal como se proporcionó anteriormente en el presente documento y (como etapa adicional) el aislamiento de dicho polipéptido de enrollamiento al azar o proteína biológicamente activa y/o dicha proteína biológicamente activa y/o dicho polipéptido conjugado a partir del cultivo o a partir de dicha célula. Entonces, este enrollamiento al azar aislado, polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina así como el conjugado aislado pueden procesarse adicionalmente. Por ejemplo, dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina puede unirse o acoplarse químicamente a una molécula de interés, tal como también se muestra en los ejemplos adjuntos. Además, y como alternativa, la molécula de interés puede conjugarse enzimáticamente, por ejemplo, mediante transglutaminasa (Besheer (2009) J Pharm Sci. 98:4420-8) u otras enzimas (Subul (2009) Org. Biomol. Chem. 7:3361-3371) a dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste en residuos de aminoácido de prolina y alanina.

El polipéptido de enrollamiento al azar (o segmento del mismo) y/o un conjugado de proteína que comprende polipéptido de enrollamiento al azar (o segmento del mismo) y una proteína de interés, tal como una proteína biológica o terapéuticamente activa o una proteína que va a usarse, por ejemplo, en métodos de diagnóstico, puede aislarse (entre otras cosas) a partir del medio de crecimiento, lisados celulares, periplasma o fracciones de membrana celular. (De nuevo, la presente invención no se limita a conjugados (de proteína) que son útiles en un entorno médico o farmacéutico. Los medios y métodos proporcionados en el presente documento también tienen utilidad en otros sectores industriales, tales como, pero sin limitarse a, industria alimentaria y de bebidas, industria de nutrientes, industria del papel, industria de biorreactivos, industria de reactivos y herramientas de investigación, industrias en las que van a usarse enzimas, industria cosmética, procesamiento de petróleo y recuperación de petróleo, y similares). El aislamiento y la purificación de los polipéptidos expresados de la invención pueden realizarse mediante cualquier medio convencional (Scopes (1982) "Protein Purification", Springer, Nueva York, NY), incluyendo precipitación con sulfato de amonio, purificación por afinidad, cromatografía en columna, electroforesis en gel y similares, y pueden implicar el uso de anticuerpos monoclonales o policlonales dirigidos, por ejemplo, contra una cola fusionada con la proteína biológicamente activa de la invención. Por ejemplo, la proteína puede purificarse mediante *Strep-tag II* usando cromatografía de afinidad por estreptavidina (Skerra (2000) Methods Enzymol. 326:271-304) tal como se describe en los ejemplos adjuntos. Se prefieren polipéptidos sustancialmente puros con una homogeneidad de al menos aproximadamente el 90 al 95% (a nivel de proteína), y lo más preferido es una homogeneidad del 98 al 99% o más, en particular para uso/aplicaciones farmacéuticas. Dependiendo de la célula huésped/organismo empleado en el procedimiento de producción, los polipéptidos de la presente invención pueden estar glicosilados o no glicosilados.

La invención se refiere además al uso de la proteína biológicamente activa, el polipéptido de enrollamiento al azar (o segmento del mismo) o los conjugados, tales como los conjugados de fármaco de la invención, la molécula de ácido nucleico de la invención, el vector de la invención o la célula (huésped) de la invención para la preparación de un medicamento, en los que dicha proteína biológicamente activa o fármaco (o cualquier otra molécula pequeña o proteína de interés) tiene una estabilidad *in vivo* y/o *in vitro* aumentada en comparación con una molécula de control que no comprende o que no está unida a un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala).

En aún otro aspecto, la presente divulgación se refiere a un método para el tratamiento de enfermedades y/o trastornos que se benefician de la estabilidad mejorada de dicha proteína biológicamente activa o fármaco, que comprende administrar la proteína biológicamente activa o conjugado de fármaco tal como se describe en el presente documento a un mamífero que necesita tal tratamiento. Dependiendo de la actividad biológica de la proteína o conjugado de fármaco de la invención, el experto puede determinar fácilmente qué enfermedad/trastorno va a tratarse con una proteína biológicamente activa o conjugado de fármaco de la invención específico. En la siguiente tabla se indican algunos ejemplos no limitativos:

Proteína	biológicamente	activa	(o	Trastorno/enfermedad que va a tratarse
----------	----------------	--------	----	----------------------------------------

ES 2 691 642 T3

componente/fragmento biológicamente activo de la misma) o fármaco	
factor estimulante de colonias de granulocitos	neutropenia relacionada con quimioterapia y/o cáncer
hormona del crecimiento humana	retraso del crecimiento y/o hipoglucemia relacionado con carencia de hormona del crecimiento
interferón α	cáncer, infección viral, hepatitis C
interferón β	enfermedad autoinmunitaria, esclerosis múltiple
interferón γ	infección viral
factor de necrosis tumoral	cáncer
interleucina-20	psoriasis
α -galactosidasa A	enfermedad de Fabry
antagonista de miostatina	sarcopenia
polipéptido inhibidor gástrico	diabetes tipo 2
antitripsina alfa-1	terapia de reposición de enzimas, fibrosis quística, enfermedades pulmonares obstructivas crónicas, síndrome respiratorio agudo, asma grave
eritropoyetina	anemia
factor de coagulación VIII	hemofilia
gp120/gp160	VIH
receptor de factor de necrosis tumoral I y II soluble	enfermedad inflamatoria
reteplasa	trombosis, infarto de miocardio
exendina-4	diabetes
antagonista de receptor de interleucina-1 (IL-1ra; anakinra)	enfermedad autoinmunitaria, artritis reumatoide
interleucina-2	cáncer
insulina	diabetes
asparaginasa	leucemia linfoblástica aguda, linfoma no Hodgkin
onconasa	mesotelioma maligno y otros tipos de cáncer
estreptocinasa	trastornos trombóticos
lipocalina asociada a gelatinasa de neutrófilos	infección microbiana, lesión por reperfusión renal
anticuerpos y sus fragmentos, incluyendo anticuerpos de dominio individual, cadena sencilla y otros fragmentos modificados por ingeniería incluyendo péptidos miméticos de CDR y las CDR	enfermedades inmunológicas, oncológicas, neovasculares e infecciosas, etc.
factor estimulante de colonias de granulocitos-macrófagos	neutropenia relacionada con quimioterapia
hormona estimulante del folículo	esterilidad
glucocerebrosidasa	enfermedad de Gaucher
timosina alfa 1	hepatitis B crónica, cáncer
glucagón	hipoglucemia
somatostatina	acromegalia
adenosina desaminasa	carencia de adenosina desaminasa
interleucina-11	trombocitopenia
factor de coagulación VIIa	hemofilia
factor de coagulación IX	hemofilia
Hematide	anemia
interferón λ	hepatitis C
leptina	lipodistrofia, obesidad, enfermedad de Alzheimer, diabetes tipo I
subunidad alfa de receptor de interleucina-22 (IL-22ra)	psoriasis
interleucina 22	melanoma metastásico
hialuronidasa	tumores sólidos
factor de crecimiento de fibroblastos 18	osteoartritis
factor de crecimiento de fibroblastos 21	diabetes tipo II, obesidad, dislipidemia, trastornos metabólicos
péptido similar a glucagón 1	diabetes
osteoprotegerina	cáncer, osteoporosis, artritis reumatoide
proteína de unión a IL-18	artritis reumatoide
factor de liberación de la hormona del crecimiento	lipodistrofia asociada con VIH
receptor de TACI soluble	lupus eritematoso sistémico, esclerosis múltiple, artritis reumatoide

trombospondina-1	cáncer
receptor de VEGF soluble Flt-1	cáncer
muteína de IL-4 (antagonista de receptor de IL-4)	asma
ciclosporina	rechazo de órgano
fumagilina	cáncer
naltrexona	adicción al alcohol
octreotida	acromegalia, tumores carcinoides
teduglutida	síndrome del intestino corto, enfermedad de Crohn
goserelina	cáncer de próstata avanzado, cáncer de mama
camptotecina	cáncer
vancomicina	neumonías Gram-positivas

Según lo anterior, la proteína biológicamente activa, el polipéptido de enrollamiento al azar (o segmento del mismo), el conjugado de fármaco, el ácido nucleico, el vector o la célula pueden usarse para la preparación de un medicamento que preferiblemente tiene o confiere una estabilidad *in vivo* y/o *in vitro* aumentada, en particular para el componente de proteína biológicamente activa y/o fármaco, para el tratamiento de carencias hormonales o trastornos relacionados, enfermedad autoinmunitaria, cáncer, anemia, enfermedades neovasculares, enfermedades infecciosas/inflamatorias, trombosis, infarto de miocardio, diabetes, esterilidad, enfermedad de Gaucher, hepatitis, hipoglucemia, acromegalia, carencia de adenosina desaminasa, trombocitopenia, hemofilia, anemia, obesidad, enfermedad de Alzheimer, lipodistrofia, psoriasis, melanoma metastásico, osteoartritis, dislipidemia, artritis reumatoide, lupus eritematoso sistémico, esclerosis múltiple, asma, osteoporosis, y lesión por reperfusión u otras enfermedades renales, por ejemplo. En una realización, la proteína biológicamente activa, el conjugado de fármaco, el ácido nucleico, el vector o la célula es para el uso como medicamento que tiene una estabilidad *in vivo* y/o *in vitro* aumentada de dicha proteína biológicamente activa/conjugado de fármaco. De manera similar, la proteína biológicamente activa, el polipéptido de enrollamiento al azar (o segmento del mismo), el conjugado de fármaco, el ácido nucleico, el vector o la célula son para su uso en el tratamiento de para el tratamiento de carencias hormonales o trastornos relacionados, enfermedad autoinmunitaria, trastornos proliferativos, tales como cáncer, anemia, enfermedades neovasculares, enfermedades infecciosas y/o inflamatorias, trombosis, infarto de miocardio, accidente cerebrovascular, diabetes, esterilidad, disfunción eréctil, enfermedad de Gaucher, enfermedad de Fabry, sarcopenia, fibrosis quística, enfermedades pulmonares obstructivas, síndrome respiratorio agudo, hepatitis, hipoglucemia, acromegalia, carencia de adenosina desaminasa, trombocitopenia, hemofilia, anemia, obesidad, enfermedad de Alzheimer, lipodistrofia, psoriasis, melanoma metastásico, osteoartritis, dislipidemia, artritis reumatoide, lupus eritematoso sistémico, esclerosis múltiple, asma, osteoporosis, y lesión por reperfusión u otras enfermedades renales, por ejemplo.

La presente invención también se refiere al uso de las moléculas de ácido nucleico, vectores así como células transfectadas tal como se proporciona en el presente documento y que comprenden las moléculas de ácido nucleico o vectores de la presente invención en enfoques médicos, tales como, por ejemplo, enfoques de terapia génica basada en células o enfoques de terapia génica basada en ácido nucleico.

En una realización adicional, el polipéptido de enrollamiento al azar (o segmento de polipéptido del mismo) tal como se proporciona en el presente documento, el conjugado de fármaco que comprende el polipéptido de enrollamiento al azar biosintético (o segmento de polipéptido del mismo) y/o la molécula de ácido nucleico o el vector o la célula huésped de la presente invención, forman parte de una composición. Dicha composición puede comprender uno o más de los conjugados de fármaco o moléculas de ácido nucleico, vectores y/o células huésped que codifican para y/o expresan los mismos. Dicha composición puede ser una composición farmacéutica, opcionalmente que comprende además un portador y/o diluyente farmacéuticamente aceptable. En una realización adicional, la presente invención se refiere al uso de la proteína biológicamente activa descrita en el presente documento, el polipéptido de enrollamiento al azar (o segmento del mismo) o el conjugado de fármaco para la preparación de una composición farmacéutica para la prevención, tratamiento o mejora de enfermedades que requieren la captación de una composición farmacéutica de este tipo.

Tal como se mencionó anteriormente en el presente documento, no sólo los conjugados dados a conocer en el presente documento, tales como conjugados de fármaco o conjugados de diagnóstico, y/o proteínas heterólogas biológicamente activas/constructos de proteína (que comprenden el polipéptido de enrollamiento al azar de la invención o segmento de polipéptido del mismo) tienen utilidad en particular médica o farmacéutica. Además, dicho polipéptido de enrollamiento al azar o segmento de polipéptido puede emplearse en sí mismo en un contexto médico de este tipo, por ejemplo como "expansor del plasma" o como sucedáneo de la sangre, en la mejora, prevención y/o tratamiento de un trastorno relacionado con una cantidad de plasma sanguíneo o contenido de plasma sanguíneo alterado o en la mejora, prevención y/o tratamiento de un trastorno relacionado con un volumen de sangre alterado. Trastornos que se tratan con expansores del plasma son, pero no se limitan a, trastornos afiliados con pérdida de sangre, tales como lesiones, cirugías, quemaduras, traumatismo o urgencias abdominales, infecciones, deshidrataciones, etc. Sin embargo, un uso médico de este tipo no se limita al polipéptido de enrollamiento al azar o segmento de polipéptido de esta invención sino que también puede extenderse a determinados conjugados de fármaco tal como se da a conocer en el presente documento o incluso a determinadas proteínas heterólogas

biológicamente activas/constructos de proteína.

En un aspecto, la composición tal como se describe en el presente documento puede ser una composición de diagnóstico, por ejemplo un reactivo de obtención de imágenes, opcionalmente que comprende además medios adecuados para detección, en la que dicha composición de diagnóstico tiene una estabilidad *in vivo* y/o *in vitro* aumentada.

Las composiciones de la invención pueden estar en forma sólida o líquida y pueden estar, entre otras cosas, en forma de un(os) polvo(s), un(os) comprimido(s), una(s) disolución/disoluciones o un(os) aerosol(es). Además, se prevé que el medicamento de la invención puede comprender agentes biológicamente activos adicionales, dependiendo del uso previsto de la composición farmacéutica.

La administración de las composiciones (farmacéuticas) adecuadas puede realizarse de diferentes maneras, por ejemplo, mediante administración parenteral, subcutánea, intravenosa, intraarterial, intraperitoneal, tópica, intrabronquial, intrapulmonar e intranasal y, si se desea para tratamiento local, administración intralesional. Las administraciones parenterales incluyen administración intraperitoneal, intramuscular, intradérmica, subcutánea, intravenosa o intraarterial. Las composiciones de la invención también puede administrarse directamente al sitio diana, por ejemplo, mediante administración biolística a un sitio diana externo o interno, tal como un órgano específicamente afectado.

En la técnica se conocen bien ejemplos de portadores, excipientes y/o diluyentes farmacéuticos adecuados e incluyen soluciones salinas tamponadas con fosfato u otras disoluciones de tampón, agua, emulsiones, tales como emulsiones de aceite/agua, diversos tipos de agentes humectantes, disoluciones estériles, etc.

Pueden formularse composiciones que comprenden tales portadores mediante métodos convencionales bien conocidos. Los portadores adecuados pueden comprender cualquier material que, cuando se combina con la proteína biológicamente activa/conjugado de fármaco de la invención, conserva su actividad biológica y/o farmacéutica (véase Remington's Pharmaceutical Sciences (1980) 16ª edición, Osol, A. Ed, Mack Publishing Company, Easton, PA). Las preparaciones para administración parenteral pueden incluir disoluciones, suspensiones y emulsiones estériles acuosas o no acuosas. Los tampones, disolventes y/o excipientes tal como se emplean en el contexto de la composición farmacéutica son preferiblemente "fisiológicos" tal como se definió anteriormente en el presente documento. Ejemplos de disolventes no acuosos son propilenglicol, polietilenglicol, aceites vegetales tales como aceite de oliva, y ésteres orgánicos inyectables tales como oleato de etilo. Los portadores acuosos incluyen agua, disoluciones alcohólicas/acuosas, emulsiones o suspensiones, incluyendo solución salina y medios tamponados. Los vehículos parenterales pueden incluir disolución de cloruro de sodio, dextrosa de Ringer, dextrosa y cloruro de sodio, solución de Ringer con lactato o aceites fijos. Los vehículos intravenosos pueden incluir reponedores de líquidos y nutrientes, reponedores de electrolitos (tales como los basados en dextrosa de Ringer), y similares. También pueden estar presentes conservantes y otros aditivos, incluyendo agentes antimicrobianos, antioxidantes, agentes quelantes y/o gases inertes y similares. Además, una composición farmacéutica de la presente invención puede comprender portadores proteicos, tales como, por ejemplo, albúmina sérica o inmunoglobulina, preferiblemente de origen humano.

Estas composiciones farmacéuticas pueden administrarse al sujeto a una dosis adecuada. El régimen de dosificación lo determinará el médico encargado y mediante factores clínicos. Tal como se conoce bien en la ciencia médica, las dosificaciones para cualquier paciente dependen de muchos factores, incluyendo el tamaño del paciente, área de superficie corporal, edad, el compuesto particular que va a administrarse, sexo, tiempo y vía de administración, salud general, y otros fármacos que están administrándose de manera concurrente. La materia farmacéuticamente activa puede estar presente en cantidades de entre 1 µg y 20 mg/kg de peso corporal por dosis, por ejemplo entre 0,1 mg y 10 mg/kg de peso corporal, por ejemplo entre 0,5 mg y 5 mg/kg de peso corporal. Si el régimen es una infusión continua, también debe estar en el intervalo de 1 µg a 10 mg por kilogramo de peso corporal por minuto. Sin embargo, también pueden preverse dosis inferiores o superiores a los intervalos indicados a modo de ejemplo, especialmente teniendo en cuenta los factores anteriormente mencionados.

Además, se prevé que la composición farmacéutica de la invención puede comprender agentes biológica o farmacéuticamente activos adicionales, dependiendo del uso previsto de la composición farmacéutica. Estos agentes biológica o farmacéuticamente activos adicionales pueden ser, por ejemplo, anticuerpos, fragmentos de anticuerpo, hormonas, factores de crecimiento, enzimas, moléculas de unión, citocinas, quimiocinas, moléculas de ácido nucleico y fármacos.

Debe observarse que la presente invención no se limita a composiciones farmacéuticas. También se prevén composiciones que van a usarse en investigación o como producto(s) de diagnóstico. Por ejemplo, se prevé que las proteínas biológicamente activas o conjugados de fármaco que comprenden un dominio o componente de enroillamiento al azar tal como se define en el presente documento se usan en un entorno de diagnóstico. Con tal fin, la proteína biológicamente activa de la invención o conjugado de fármaco de esta invención puede marcarse con el fin de permitir la detección. Tales etiquetas comprenden, pero no se limitan a, etiquetas radiactivas (tales como [³H]hidrógeno [¹²⁵I]yoduro o [¹²³I]yoduro), etiquetas fluorescentes (incluyendo proteínas fluorescentes, tales como proteína verde fluorescente (GFP) o fluoróforos, tales como isotiocianato de fluoresceína (FITC)) o etiquetas de

RMN (tales como quelatos de gadolinio). Las etiquetas o marcadores definidos en este caso no son de ninguna manera limitativos y representan simplemente ejemplos ilustrativos. Las composiciones de diagnóstico de esta divulgación son particularmente útiles en experimentos de seguimiento u obtención de imágenes o en un entorno médico de diagnóstico. En los ejemplos y figuras adjuntos, se proporciona la preparación de un constructo correspondiente que comprende conjugados que comprenden (i) un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala), y (ii) fluoresceína o digoxigenina; véanse las figuras 13 y 14 adjuntas y la leyenda de las figuras correspondiente así como ejemplo ilustrativo 22.

Pero no sólo los usos farmacéuticos o de diagnóstico de los medios y métodos proporcionados en el presente documento están dentro de la esencia de la presente invención. Los compuestos/conjugados proporcionados en el presente documento también son útiles en otros determinados sectores industriales, tales como en la industria alimentaria, la industria de las bebidas, la industria cosmética, la industria del petróleo, la industria del papel y similares. Por tanto, la presente divulgación también proporciona usos del polipéptido de enrollamiento al azar biosintético tal como se proporciona en el presente documento en tales sectores industriales. Por consiguiente, también forma parte de esta divulgación un método para la producción de un cosmético, de un compuesto que va a usarse en tratamientos cosméticos, de un alimento o de una bebida, comprendiendo dicho método el cultivo de la célula que comprende una molécula de ácido nucleico (o un vector) que codifica para un polipéptido de enrollamiento al azar tal como se define en el presente documento o que codifica para una proteína biológicamente activa y/o una proteína biológicamente activa y/o un polipéptido que comprende o que es una secuencia de aminoácidos que tiene o que media en una actividad. Un método de este tipo también incluye el aislamiento de dicho polipéptido de enrollamiento al azar, dicha proteína biológicamente activa y/o dicha proteína biológicamente activa o dicho polipéptido que comprende o que es una secuencia de aminoácidos que tiene o que media en una actividad, tal como una actividad biológica, y que adicionalmente comprende dicho polipéptido de enrollamiento al azar o segmento de polipéptido de enrollamiento al azar a partir del cultivo o a partir de dicha célula. En el mismo contexto, pueden producirse otros conjugados de interés, por ejemplo conjugados que son útiles en diferentes sectores de la industria, tales como la industria del petróleo o del papel. El experto en la técnica está fácilmente en posición de adaptar los medios y métodos proporcionados en el presente documento para la generación de constructos moleculares/recombinantes correspondientes así como para la generación de conjugados que comprenden un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala), y una molécula pequeña o un polipéptido de interés.

También se describe un kit que comprende el conjugado de fármaco, la molécula de ácido nucleico que codifica para dicha proteína biológicamente activa y/o que codifica para dicho polipéptido que comprende o que es una secuencia de aminoácidos que tiene o que media en una actividad (por ejemplo una actividad biológica), el vector que comprende dicha molécula de ácido nucleico o la célula que comprende dicho ácido nucleico o dicho vector tal como se describe en el presente documento. El kit de la presente invención puede comprender además, un(os) tampón/tampones, disoluciones de almacenamiento y/o reactivos o materiales adicionales requeridos para la realización de ensayos y fines médicos o científicos. Además, partes del kit pueden envasarse de manera individual en viales o botellas o en combinación en recipientes o unidades de múltiples recipientes.

El kit puede usarse ventajosamente, entre otras cosas, para llevar a cabo el método de la invención y puede emplearse en una variedad de aplicaciones a las que se hace referencia en el presente documento, por ejemplo, como herramientas de investigación o como herramientas médicas. Adicionalmente, el kit puede contener medios para la detección adecuados para fines científicos o médicos. La fabricación de los kits sigue preferiblemente procedimientos convencionales que conoce el experto en la técnica.

La invención se ilustra adicionalmente mediante las siguientes figuras y ejemplos no limitativos.

Figura 1. Diseño génico para la secuencia de polímero/polipéptido de Pro/Ala PA#1.

Secuencia de nucleótidos y de aminoácidos codificada de un bloque estructural para PA#1 (SEQ ID NO: 1) obtenida mediante hibridación de dos oligodesoxinucleótidos complementarios (oligodesoxinucleótido de cadena superior/codificante SEQ ID NO: 17, oligodesoxinucleótido de cadena inferior/no codificante SEQ ID NO: 18). El ácido nucleico resultante tiene dos extremos cohesivos (mostrados en minúsculas), correspondientes a un codón de Ala y anticodón, respectivamente, y son mutuamente compatibles. Tras la ligación repetida de un bloque estructural de este tipo, pueden obtenerse concatémeros que codifican para polipéptidos de Pro-Ala de longitudes variables y posteriormente clonarse, por ejemplo, mediante un(os) sitio(s) de restricción *SapI*.

Figura 2. Estrategias de clonación para una secuencia de polímero/polipéptido de Pro/Ala como fusión a un fragmento Fab o a IFNa2b humano.

(A) Tramo de secuencia de nucleótidos y de aminoácidos codificada (cadena superior/codificante SEQ ID NO: 19, cadena inferior/no codificante SEQ ID NO: 20, secuencia de aminoácidos codificada SEQ ID NO: 21) alrededor del

extremo C-terminal de la cadena ligera de inmunoglobulina de un fragmento Fab de anticuerpo tal como se codifica en pASK88-Fab-2xSapI (SEQ ID NO: 22), un derivado de pASK75, usado para la subclonación de secuencias de polímero/polipéptido de Pro/Ala y expresión de proteínas biológicamente activas correspondientes. La secuencia de nucleótidos porta dos sitios de reconocimiento SapI en orientación mutuamente inversa, lo que conduce, tras la digestión, a extremos de ADN sobresalientes que son compatibles con el casete génico sintético mostrado en la figura 1. Las secuencias de reconocimiento y los aminoácidos C-terminales de la cadena ligera están subrayados.

(B) Secuencia de nucleótidos y secuencia de aminoácidos codificada (cadena superior/codificante SEQ ID NO: 23, cadena inferior/no codificante SEQ ID NO: 24, secuencia de aminoácidos codificada SEQ ID NO: 25) de un polímero/polipéptido PA#1 con 20 residuos tras la inserción de un único casete tal como se muestra en la figura 1 en el plásmido pASK88-Fab-2xSapI. La ligación/inserción similar de 10 de tales casetes repetidos dio como resultado el vector de plásmido pFab-PA#1(200) (SEQ ID NO: 28) que codifica para un polímero/polipéptido con 200 residuos (SEQ ID NO: 26 y 27). Los sitios de restricción SapI que flanquean la secuencia que codifica para polímero de Pro/Ala están marcados (las secuencias de reconocimiento están subrayadas).

(C) Mapa de plásmido de pFab-PA#1(200) (SEQ ID NO: 28). Los genes estructurales para la cadena pesada (HC) y la cadena ligera (LC) de Fab-PA#1(200) están bajo el control transcripcional del promotor/operador de tetraciclina ($tet^{p/o}$) y los extremos de operón con el terminador de lipoproteína (t_{pp}). HC comprende el péptido señal OmpA bacteriano, el dominio variable (VH) y el primer dominio constante C (CH) de cadena pesada de IgG1 humana así como la cola de His6. LC comprende el péptido señal PhoA bacteriano, el dominio variable (VL) y dominio constante (CL) de cadena ligera humana, el polímero/polipéptido PA#1 con 200 residuos. La estructura principal de plásmido de pFab-PA#1(200) fuera del casete de expresión flanqueado por los sitios de restricción XbaI y HindIII es idéntica a la del vector de clonación y expresión genérico pASK75 (Skerra (1994) Gene 151:131-135). Se indican sitios de restricción individuales.

(D) Tramo de secuencia de nucleótidos y aminoácidos (cadena superior/codificante SEQ ID NO: 29, cadena inferior/no codificante SEQ ID NO: 30, secuencia de aminoácidos codificada SEQ ID NO: 31) alrededor del extremo N-terminal de IFNa2b humano tal como se clona en pASK-IFNa2b (SEQ ID NO: 32). El sitio de restricción individual SapI que puede usarse para la inserción de la secuencia que codifica para polímero de Pro/Ala está marcado (la secuencia de reconocimiento está subrayada). Los dos aminoácidos C-terminales de Strep-tag II están subrayados. El primer aminoácido de IFNa2b maduro está marcado con +1.

(E) Tramo de secuencia de nucleótidos y de aminoácidos codificada (cadena superior/codificante SEQ ID NO: 33, cadena inferior/no codificante SEQ ID NO: 34, secuencia de aminoácidos codificada SEQ ID NO: 35) del extremo N-terminal de IFNa2b tras la inserción de un casete de secuencia de polímero PA#1 tal como se muestra en la figura 1. El sitio de restricción individual SapI, que permanece tras la inserción de la secuencia que codifica para polímero de Pro/Ala, está marcado (las secuencias de reconocimiento están subrayadas). El primer aminoácido de IFNa2b como parte de la proteína de fusión está marcado (1) y los dos aminoácidos C-terminal de Strep-tag II están subrayados. La ligación/inserción similar de 10 casetes de secuencia de polímero PA#1 repetidos dio como resultado el vector de plásmido pPA#1(200)-IFNa2b que codifica para un polímero/polipéptido con 200 residuos (SEQ ID NO: 36).

(F) Mapa de plásmido de pPA#1(200)-IFNa2b (SEQ ID NO: 37). El gen estructural para la proteína biológicamente activa PA#1(200)-IFNa2b (que comprende el péptido señal OmpA bacteriano, Strep-tag II, el segmento de polímero/polipéptido PA#1 con 200 residuos, e IFNa2b humano) está bajo el control transcripcional del promotor/operador de tetraciclina ($tet^{p/o}$) y termina con el terminador de lipoproteína (t_{pp}). La estructura principal de plásmido fuera del casete de expresión flanqueado por los sitios de restricción XbaI y HindIII es idéntica a la del vector de clonación y expresión genérico pASK75 (Skerra (1994), citado anteriormente). Se indican sitios de restricción individuales.

Figura 3. Análisis del fragmento Fab recombinante purificado y del IFNa2b recombinante purificado así como sus fusiones de polipéptido/polímero de Pro/Ala mediante SDS-PAGE.

Las proteínas recombinantes se produjeron en *E. coli* KS272 (Strauch (1988) Proc. Natl. Acad. Sci. USA 85:1576-80) mediante secreción periplásmica y se purificaron por medio de la cola de His6 (Fab) o Strep-tag II (IFNa2b) usando cromatografía de metal inmovilizado o de afinidad por estreptavidina, respectivamente.

(A) Análisis del Fab recombinante purificado y su fusión con PA#1 con 200 residuos mediante SDS-PAGE al 12%. El gel muestra muestras de proteína de 2 µg de cada uno de Fab y Fab-PA#1(200). Las muestras a la izquierda se redujeron con 2-mercaptoetanol mientras que las muestras repetidas a la derecha se dejaron sin reducir. En el margen izquierdo se indican marcadores de tamaño de proteína (aplicados en condiciones reductoras). Tras la reducción del puente disulfuro entre cadenas el fragmento Fab y su fusión con PA#1 de 200 residuos aparecen como dos bandas homogéneas. En el caso del fragmento Fab reducido, las dos bandas con tamaños moleculares de cerca de 24 y 26 kDa, respectivamente, corresponden a las LC y HC separadas. En el caso de la proteína de fusión Fab-PA#1(200) reducida la banda a 24 kDa corresponde a la HC, mientras que la banda a cerca de 90 kDa corresponde a la LC fusionada con el segmento de polipéptido PA#1(200). En condiciones no reductoras, el fragmento Fab y su fusión con PA#1(200) aparecen como bandas homogéneas individuales con tamaños moleculares aparentes de cerca de 45 kDa y 100 kDa, respectivamente. Los dos valores de tamaño aparentes para

la proteína de fusión Fab-PA#1(200) son significativamente mayores que las masas calculadas de 64,3 kDa para Fab-PA#1(200) no reducida y de 39,1 kDa para LC-PA#1(200) aislada. Este efecto se debe a la adición del segmento de polímero/polipéptido de Pro/Ala porque el propio fragmento Fab, con una masa calculada de 48,0 kDa, o su cadena ligera sin fusionar muestran movilidad electroforética esencialmente normal.

- 5 (B) Análisis del IFNa2b recombinante purificado y su proteína de fusión con PA#1 con 200 residuos mediante SDS-PAGE al 12%. El gel muestra muestras de proteína de 2 µg de cada uno de IFNa2b y de PA#1(200)-IFNa2b. Las muestras a la izquierda se redujeron con 2-mercaptoetanol mientras que las muestras correspondientes a la derecha se dejaron sin reducir. En el margen izquierdo se indican marcadores de tamaño de proteína (aplicados en condiciones reductoras). Las dos proteínas aparecen como bandas homogéneas individuales con tamaños moleculares aparentes de cerca de 20 kDa y cerca de 80 kDa en la forma reducida. Este último valor es significativamente mayor que la masa calculada de 37,0 kDa para PA#1(200)-IFNa2b. Este efecto se debe a la adición del segmento de polímero/polipéptido de Pro/Ala porque el propio IFNa2b, con una masa calculada de 20,9 kDa, muestra movilidad electroforética esencialmente normal. IFNa2b en el estado no reducido tiene una movilidad electroforética ligeramente superior, indicando una forma más compacta como resultado de sus dos puentes disulfuro intramoleculares.

Figura 4. Análisis cuantitativo de los volúmenes hidrodinámicos del Fab recombinante purificado e IFNa2b así como sus fusiones con PA#1(200).

- 20 (A) Cromatografía de exclusión molecular analítica (SEC) de Fab y Fab-PA#1(200). Se aplicaron 250 µl de la proteína purificada a una concentración de 0,25 mg/ml a una columna Superdex S200 10/300 GL equilibrada con tampón PBS. Se monitorizó la absorción a 280 nm y se normalizó el pico de cada ejecución de cromatografía a un valor de 1. La flecha indica el volumen inicial de la columna (7,8 ml).

- 25 (B) Curva de calibración para los cromatogramas de (A) usando una columna Superdex S200 10/300 GL. El logaritmo del peso molecular (PM) de proteínas marcadoras (citocromo c, 12,4 kDa; anhidrasa carbónica, 29,0 kDa; ovoalbúmina, 43,0 kDa; albúmina sérica bovina, 66,3 kDa; alcohol deshidrogenasa, 150 kDa, β-amilasa, 200 kDa, apo-ferritina, 440 kDa) se representó gráficamente frente a sus volúmenes de elución (círculos negros) y se ajustó mediante una línea recta. A partir de los volúmenes de elución observados del fragmento Fab y su proteína de fusión con PA#1 (cuadrados negros) se determinaron sus tamaños moleculares aparentes de la siguiente manera. Fab: 31 kDa (masa auténtica: 48,0 kDa); Fab-PA#1(200): 237 kDa (masa auténtica: 64,3 kDa). Estos datos muestran que la fusión con el polipéptido PA#1 confiere un volumen hidrodinámico muy ampliado.

- 30 (C) Cromatografía de exclusión molecular analítica de IFNa2b y PA#1(200)-IFNa2b. Se aplicaron 250 µl de cada proteína purificada a una concentración de 0,25 mg/ml a una columna Superdex S200 10/300 GL equilibrada con solución salina tamponada con fosfato, PBS. Se monitorizó la absorción a 280 nm y se normalizó el pico de cada ejecución de cromatografía a un valor de 1. La flecha indica el volumen inicial de la columna (7,8 ml).

- 35 (D) Curva de calibración para el cromatograma de (C) usando una columna Superdex S200 10/300 GL. El logaritmo del peso molecular (PM) de proteínas marcadoras (véase B) se representó gráficamente frente a sus volúmenes de elución (círculos negros) y se ajustó mediante una línea recta. A partir de los volúmenes de elución observados de IFNa2b y su proteína de fusión con PA#1 (cuadrados negros) se determinaron sus tamaños moleculares aparentes de la siguiente manera. IFNa2b: 22,5 kDa (masa auténtica: 20,9 kDa); PA#1(200)-IFNa2b: 229,0 kDa (masa auténtica: 37,0 kDa). Estos datos muestran que la fusión con el polipéptido PA#1 confiere un volumen hidrodinámico muy ampliado.

Figura 5. Análisis de estructura secundaria experimental de proteínas recombinantes y sus fusiones con polímero/polipéptido PA#1 mediante espectroscopía de dicroísmo circular (CD).

Se registraron espectros a temperatura ambiente en K₂SO₄ 50 mM, fosfato de K 20 mM pH 7,5 y se normalizaron a la elipticidad molar, Θ_M , para cada proteína.

- 45 (A) Espectros de CD del Fab recombinante purificado y Fab-PA#1(200). El espectro de CD para el fragmento Fab muestra las características típicas de una proteína de láminas β predominante con un máximo negativo ancho alrededor de 216 nm (Sreerama en: Circular Dichroism - Principles and Applications (2000) Berova, Nakanishi y Woody (Eds.) Wiley, Nueva York, NY, págs. 601-620), lo que indica el plegamiento correcto del fragmento Fab producido de manera bacteriana. El espectro de su proteína de fusión con el polímero/polipéptido de Pro/Ala revela una banda negativa dominante por debajo de 200 nm, lo cual es indicativo de una conformación de enrollamiento al azar. Además, hay un hombro alrededor de 220 nm, que resulta de la contribución de láminas β del fragmento Fab e indica su plegamiento correcto incluso como parte de la proteína de fusión.

- 55 (B) Espectro de CD de diferencia molar para Fab-PA#1(200) obtenido mediante sustracción del espectro para el fragmento Fab. El espectro de CD de diferencia representa la estructura secundaria del segmento de polímero/polipéptido PA#1 de 200 residuos y revela un fuerte mínimo alrededor de 200 nm, lo cual es una clara indicación de una conformación de enrollamiento al azar en la disolución acuosa tamponada (Greenfield (1969) Biochemistry 8: 4108-4116; Sreerama (2000), citado anteriormente; Fändrich (2002) EMBO J. 21:5682-5690).

(C) Espectros de CD del IFNa2b recombinante purificado y PA#1(200)-IFNa2b. El espectro de CD para IFNa2b muestra las características típicas de una proteína de hélice α predominante con dos bandas negativas alrededor de 208 nm y 220 nm (Sreerama (2000), citado anteriormente), lo cual indica el plegamiento correcto del IFNa2b humano producido de manera bacteriana. El espectro de su proteína de fusión con el polímero/polipéptido de Pro/Alar revela desviaciones características con un mínimo dominante alrededor de 200 nm, lo cual es indicativo de una conformación de enrollamiento al azar. Además, hay un hombro alrededor de 220 nm, que resulta de la contribución de hélices α de IFNa2b e indica el plegamiento correcto del IFNa2b incluso como parte de la proteína de fusión.

(D) Espectro de CD de diferencia molar para PA#1(200)-IFNa2b obtenido mediante sustracción del espectro para IFNa2b. El espectro de CD de diferencia representa la estructura secundaria del segmento de polímero/polipéptido PA#1 de 200 residuos y revela un fuerte mínimo alrededor de 200 nm, esencialmente idéntico al mostrado en (B). De nuevo, esto es una clara indicación de una conformación de enrollamiento al azar en disolución acuosa tamponada para un polímero biológico que comprende residuos de Pro y Ala según la invención.

Figura 6. Producción secretora de una proteína de fusión entre hormona del crecimiento humana (hGH) y el polímero PA#1 genéticamente codificado en células CHO.

(A) Tramo de secuencia de nucleótidos y aminoácidos (cadena superior/codificante SEQ ID NO: 38, cadena inferior/no codificante SEQ ID NO: 39, secuencia de aminoácidos codificada SEQ ID NO: 40) alrededor del extremo N-terminal de hGH tal como se clona en pASK75-His6-hGH (SEQ ID NO: 41). Los sitios de restricción individuales *NheI*, que puede usarse junto con *HindIII* (no mostrado) para la subclonación, y *SapI*, que puede usarse para la inserción de la secuencia que codifica para polímero de Pro/Ala, están marcados (la secuencia de reconocimiento está subrayada). Los seis aminoácidos de la cola de His6 están subrayados. El primer aminoácido de hGH está marcado con +1.

(B) Secuencia de nucleótidos y de aminoácidos codificada (cadena superior/codificante SEQ ID NO: 42, cadena inferior/no codificante SEQ ID NO: 43, secuencia de aminoácidos codificada SEQ ID NO: 44) del extremo N-terminal de hGH tras la inserción de un casete de secuencia de polímero PA#1 tal como se muestra en la figura 1. Los sitios de restricción individuales *NheI*, que puede usarse para la subclonación, y *SapI*, que permanece tras la inserción de la secuencia que codifica para polímero de Pro/Ala, están marcados (las secuencias de reconocimiento están subrayadas). El primer aminoácido de hGH como parte de la proteína de fusión está marcado (1) y los aminoácidos de la cola de His6 están subrayados. La ligación/inserción similares de 10 casetes de secuencia de polímero PA#1 repetidos dieron como resultado el vector de plásmido pASK75-His6-PA#1(200)-hGH que codifica para la proteína de fusión madura SEQ ID NO: 45).

(C) Mapa de plásmido de pASK75-His6-PA#1(200)-hGH (SEQ ID NO: 46). El gen estructural para la proteína biológicamente activa His6-PA#1(200)-hGH (que comprende el péptido señal OmpA bacteriano, la cola de His6, el segmento de polímero/polipéptido PA#1 con 200 residuos, y GH humana) está bajo el control transcripcional del promotor/operador de tetraciclina ($tet^{p/o}$) y termina con el terminador de lipoproteína (t_{lpp}). La estructura principal de plásmido fuera del casete de expresión flanqueado por los sitios de restricción *XbaI* y *HindIII* es idéntica a la del vector de clonación y expresión genérico pASK75 (Skerra (1994), citado anteriormente). Se indican sitios de restricción individuales.

(D) Mapa de plásmido de pCHO-PA#1(200)-hGH, que codifica para una proteína de fusión His6-PA#1(200)-hGH (SEQ ID NO: 47). El gen estructural, que comprende el péptido señal de hormona del crecimiento humana (Sp), la cola de His6, la secuencia de polímero/polipéptido PA#1 con 200 residuos (PA#1(200)), la hormona del crecimiento humana (hGH), y que contiene la señal de poliadenilación de hormona del crecimiento bovina (bGH pA), está bajo el control transcripcional del promotor de citomegalovirus (CMV^P). Los sitios de restricción individuales *NheI* y *HindIII* están indicados. El gen de resistencia para neomicinofosfotransferasa (neo) está bajo el control del promotor de SV40 (SV40^P) y seguido por una señal de poliadenilación de SV40 (SV40 pA). Además, el plásmido contiene el origen de replicación de ColE1 bacteriano (ColEI-ori), el origen de replicación de bacteriófago f1 (f1-ori), y el gen de β -lactamasa (bla) para permitir la propagación y selección en *E. coli*.

(E) Análisis de inmunotransferencia de tipo Western de una proteína de fusión entre hGH y el polímero PA#1 genéticamente codificado de 200 residuos producido en células CHO en comparación con hGH recombinante. Se transfectoron células CHO-K1 o bien con pCHO-PA#1(200)-hGH (SEQ ID NO: 48) o bien con pCHO-hGH (SEQ ID NO: 49), un plásmido similar que codifica para hGH sin la secuencia de PA#1(200) (pero que también porta la cola de His6). Dos días tras la transfección, se sometió una muestra del sobrenadante de cultivo celular a SDS-PAGE e inmunotransferencia de tipo Western con un anticuerpo anti-hGH conjugado con peroxidasa del rábano. Las dos proteínas aparecen como bandas individuales indicadas por flechas, con tamaños moleculares aparentes de cerca de 23 kDa (His6-hGH) y cerca de 90 kDa (His6-PA#1-hGH). También hay una banda débil alrededor de 60 kDa que surge de proteínas séricas en el medio de cultivo. Aunque la hGH marcada con cola de His6 aparece a la masa calculada de 23,5 kDa, el tamaño molecular aparente de His6-PA#1-hGH es significativamente mayor que su masa calculada de 39,5 kDa. Este efecto se debe a la naturaleza hidrófila de enrollamiento al azar del polímero de Pro-Ala.

Figura 7. Predicción teórica de la estructura secundaria para la secuencia de polipéptido/polímero de Pro/Ala PA#1.

Esta ilustración muestra el resultado del algoritmo informático CHOFAS según el método de Chou-Fasman (Chou y Fasman (1974) *Biochemistry* 13: 222-245) tal como se implementa en el servidor de comparación de secuencias y predicción de estructura secundaria en la Universidad de Virginia (URL: http://fasta.bioch.virginia.edu/fasta_www2). Para evitar efectos de límite en los extremos amino y carboxilo-terminales, se pegó la repetición de aminoácidos de 20 meros según la figura 1 en tres copias consecutivas (dando como resultado un concatémero similar al codificado tras la ligación/inserción repetida del casete génico sintético) y sólo se consideró el resultado para el bloque de secuencia de 20 meros central (en recuadro). En el caso de la secuencia/segmento de polipéptido PA#1 (SEQ ID NO: 1), el algoritmo de Chou-Fasman predice una estructura secundaria en hélices α al 100%. Esto contrasta con la conformación de enrollamiento al azar predominante observada de manera experimental para el segmento de polipéptido/polipéptido PA#1 como parte de una proteína de fusión (véase la figura 5B/D).

Figura 8: Análisis cuantitativo de la farmacocinética del fragmento Fab recombinante purificado y sus fusiones con polímero PA#1 con 200 y 600 residuos en ratones BALB/c.

Se sometieron a ensayo muestras de plasma del ejemplo 16 para determinar concentraciones de Fab, Fab-PA#1(200) y Fab-PA#1(600) usando un ELISA de tipo sándwich. Para estimar la semivida en plasma de Fab, Fab-PA#1 (200) y Fab-PA#1(600), se representaron gráficamente los valores de concentración medidos frente al tiempo tras la inyección intravenosa y se ajustaron numéricamente suponiendo una disminución biexponencial. El fragmento Fab sin fusión mostró un aclaramiento muy rápido con una semivida de eliminación de $1,3 \pm 0,1$ h. En cambio, la fase de eliminación determinada para Fab-PA#1(200) y Fab-PA#1(600) fue significativamente más lenta, con semividas terminales de $4,1 \pm 1,8$ h y $38,8 \pm 11,2$ h, respectivamente, demostrando por tanto una circulación prolongada cerca de 3 veces y cerca de 30 veces debido a la fusión con polímero de Pro/Ala con 200 ó 600 residuos en comparación con el fragmento Fab sin fusionar.

Figura 9: Análisis del fragmento Fab recombinante purificado como fusión con el polímero P1A1 o P1A3 que tiene 200 residuos.

Las proteínas recombinantes se produjeron en *E. coli* KS272 mediante secreción periplásmica y se purificaron por medio de la cola de His6 usando cromatografía de afinidad de metal inmovilizado. Se analizaron las proteínas purificadas mediante SDS-PAGE al 12%. El gel muestra muestras de proteína de 2 μ g de cada uno de Fab-P1A1(200) y Fab-P1A3(200) así como, para comparación, del fragmento Fab sin fusionar (véase la figura 3A). Las muestras a la izquierda se redujeron con 2-mercaptoetanol mientras que las muestras análogas a la derecha se dejaron sin reducir. En el margen izquierdo se indican marcadores de tamaño de proteína (aplicados en condiciones reductoras). Tras la reducción de los puentes disulfuro entre cadenas, el fragmento Fab y sus fusiones con Pro/Ala de 200 residuos aparecen como dos bandas homogéneas. En el caso del fragmento Fab reducido, las dos bandas con tamaños moleculares de cerca de 24 y 26 kDa, respectivamente, corresponden al fragmento de cadena ligera (LC) y cadena pesada (HC) separadas. En el caso de la proteína de fusión Fab-P1A1(200) reducida, la banda a 24 kDa corresponde a la HC, mientras que la banda a cerca de 90 kDa corresponde a la LC fusionada con el polipéptido P1A1(200). En el caso de la proteína de fusión Fab-P1A3(200) reducida, la banda a 24 kDa corresponde a la HC, mientras que la banda a cerca de 75 kDa corresponde a la LC fusionada con el polipéptido P1A5(200). En condiciones no reductoras, el fragmento Fab, su fusión con P1A1(200) y con P1A3(200) aparecen como bandas prominentes individuales con tamaños moleculares aparentes de cerca de 45 kDa, 110 kDa y 90 kDa, respectivamente. Los tamaños aparentes para las proteínas de fusión Fab-P1A1(200) y Fab-P1A3(200) son significativamente más grandes que las masas calculadas de 65,3 kDa para Fab-P1A1(200) sin reducir y de 64,0 kDa para la Fab-P1A3(200) sin reducir. Además, los tamaños aparentes para las cadenas ligeras reducidas correspondientes son significativamente más grandes que las masas calculadas de 40,7 kDa para la LC de P1A1(200) y de 39,4 kDa para la LC de P1A3(200). Este efecto se debe a la adición del segmento de polímero/polipéptido de Pro/Ala porque el propio fragmento Fab, con una masa calculada de 48,0 kDa, o su cadena ligera sin fusionar, con una masa calculada de 23,4 kDa, muestran movilidad electroforética esencialmente normal.

Figura 10. Análisis cuantitativo de los volúmenes hidrodinámicos de las proteínas de fusión Fab recombinante purificado-P1A1(200) y Fab-P1A3(200).

Cromatografía de exclusión molecular analítica (SEC) de Fab-P1A1(200) y Fab-P1A3(200). Se aplicaron 250 μ l de la proteína purificada a una concentración de 0,25 mg/ml a una columna Superdex S200 10/300 GL equilibrada con PBS. Se monitorizó la absorción a 280 nm y se normalizó el pico de cada ejecución de cromatografía a un valor de 1. La flecha indica el volumen inicial de la columna (7,8 ml). A partir de los volúmenes de elución observados de las proteínas de fusión se determinaron sus tamaños moleculares aparentes usando una curva de calibración similar tal como se muestra en la figura 4B de la siguiente manera. Fab-P1A1(200): 180,7 kDa (masa auténtica: 65,3 kDa); Fab-P1A3(200): 160,2 kDa (masa auténtica: 64,0 kDa). Estos datos muestran que la fusión de una proteína con el polipéptido P1A1 y/o P1A5 confiere un volumen hidrodinámico muy ampliado.

Figura 11. Análisis de estructura secundaria experimental de fusiones Fab-P1A1(200) y Fab-P1A3(200) mediante espectroscopía de dicroísmo circular (CD).

Se registraron espectros a temperatura ambiente en K_2SO_4 50 mM, fosfato de K 20 mM pH 7,5 y se normalizaron a la elipticidad molar, Θ_M , para cada proteína.

5 (A) Espectros de CD de Fab recombinante purificado-P1A1(200) y Fab-P1A3(200). Los espectros de CD de las proteínas de fusión de Fab con ambos polímeros/polipéptidos de Pro/Ala revelan cada uno una banda negativa dominante por debajo de 200 nm, lo cual es indicativo de una conformación de enrollamiento al azar. Además, hay un hombro alrededor de 220 nm, que surge de la contribución de láminas β del fragmento Fab e indica su plegamiento correcto incluso como parte de la proteína de fusión.

10 (B) Espectros de CD de diferencia molar para Fab-P1A1(200) y Fab-P1A3 (200) obtenidos mediante sustracción del espectro para el fragmento Fab sin fusionar (véase la figura 5A). Los espectros de CD de diferencia representan las estructuras secundarias de los segmentos de polímero/polipéptido P1A1 (SEQ ID NO: 51) y P1A3 (SEQ ID NO: 3) de 200 residuos, respectivamente, y revelan un fuerte mínimo alrededor de 200 nm, lo cual es una clara indicación de su conformación de enrollamiento al azar en la disolución acuosa tamponada (Greenfield (1969) *Biochemistry* 8: 4108-4116; Sreerama (2000), citado anteriormente; Fändrich (2002) *EMBO J.* 21:5682-5690).

Figura 12: Preparación de un polímero/polipéptido de Pro/Ala biosintético aislado.

15 (A) Mapa de plásmido de pSUMO-PA#1(200) (SEQ ID NO: 60). El gen estructural para la proteína de fusión MKHis(6)-SUMO-PA#1(200) que comprende un codón de iniciación de metionina seguido por un codón de lisina, una cola de afinidad N-terminal de seis residuos de His consecutivos, la proteína modificador similar a ubiquitina pequeño (SUMO) escindible Smt3p (Panavas (2009) *Methods Mol Biol.* 497: 303-17), y el segmento de polímero/polipéptido PA#1 con 200 residuos (SEQ ID NO: 60) está bajo el control transcripcional del promotor de gen 10 del bacteriófago T7 y termina con el terminador $t\phi$. Los elementos de plásmido adicionales comprenden el origen de replicación (ori), el gen de resistencia a ampicilina (bla), y el origen de replicación de f1. La estructura principal de plásmido fuera del casete de expresión flanqueado por los sitios de restricción *NdeI* y *HindIII* es idéntica, excepto por un sitio de restricción *SapI* que se eliminó mediante mutación silenciosa, a la del vector de clonación y expresión genérico pRSET5a (Schoepfer (1993) 124: 83-85).

20 SEQ ID NO: 60 se proporciona en la lista de secuencias adjunta (que también forma parte de esta descripción y memoria descriptiva) y se reproduce a continuación en el presente documento.

ES 2 691 642 T3

gcacttttcg	gggaaatgtg	cgcggaaccc	ctatttgttt	atttttctaa	atacattcaa	60
atatgtatcc	gctcatgaga	caataaccct	gataaatgct	tcaataatat	tgaaaaagga	120
agagtatgag	tattcaacat	ttccgtgtcg	cccttattcc	cttttttgcg	gcattttgcc	180
ttcctgtttt	tgctcaccca	gaaacgctgg	tgaaagtaaa	agatgctgaa	gatcagttgg	240
gtgcacgagt	gggttacatc	gaactggatc	tcaacagcgg	taagatcctt	gagagttttc	300
gccccgaaga	acgtttttcca	atgatgagca	cttttaaaagt	tctgctatgt	ggcgcggtat	360
tatcccgat	tgacgccggg	caagagcaac	tccgtcgcgg	catacactat	tctcagaatg	420
acttggttga	gtactcacca	gtcacagaaa	agcatcttac	ggatggcatg	acagtaagag	480
aattatgcag	tgctgccata	accatgagtg	ataacactgc	ggccaactta	cttctgacaa	540
cgatcggagg	accgaaggag	ctaaccgctt	ttttgcacaa	catgggggat	catgtaactc	600
gccttgatcg	ttgggaaccg	gagctgaatg	aagccatacc	aaacgacgag	cgtgacacca	660
cgatgcctgt	agcaatggca	acaacgttgc	gcaaactatt	aactggcgaa	ctacttactc	720
tagcttcccg	gcaacaatta	atagactgga	tggaggcggg	taaagttgca	ggaccacttc	780
tgcgctcggc	ccttccggct	ggctggttta	ttgctgataa	atctggagcc	ggtgagcgtg	840
ggtctcgcgg	tatcattgca	gcaactgggc	cagatggtaa	gccctcccgt	atcgtagtta	900
tctacacgac	ggggagtcat	gcaactatgg	atgaacgaaa	tagacagatc	gtcagatag	960
gtgcctcact	gattaagcat	tgtaactgt	cagaccaagt	ttactcatat	atactttaga	1020
ttgatttaaa	acttcatttt	taatttaaaa	ggatctaggt	gaagatcctt	tttgataatc	1080
tcattgaccaa	aatcccttaa	cgtgagtttt	cgttccactg	agcgtcagac	cccgtagaaa	1140
agatcaaagg	atcttcttga	gatccttttt	ttctgcgcgt	aatctgctgc	ttgcaaacia	1200
aaaaaccacc	gctaccagcg	gtggttttgt	tgccggatca	agagctacca	actctttttc	1260
cgaaggtaac	tggcttcagc	agagcgcaga	taccaaatat	tgtccttcta	gtgtagccgt	1320
agttaggcca	ccacttcaag	aactctgtag	caccgcctac	atacctcget	ctgctaatcc	1380
tgttaccagt	ggctgctgcc	agtggcgata	agtcgtgtct	taccgggttg	gactcaagac	1440
gatagttacc	ggataaggcg	cagcggtcgg	gctgaacggg	gggttcgtgc	acacagccca	1500
gcttggagcg	aacgacctac	accgaactga	gatacctaca	gcgtgagcta	tgagaaagcg	1560
ccacgcttcc	cgaagggaga	aaggcggaca	ggtatccggt	aagcggcagg	gtcggaacag	1620
gagagcgcac	gagggagctt	ccagggggaa	acgcctggta	tctttatagt	cctgtcgggt	1680
ttcgccacct	ctgacttgag	cgtcgatttt	tgtgatgctc	gtcagggggg	cggagcctat	1740
ggaaaaacgc	cagcaacgcg	gcctttttac	ggttcctggc	cttttgctgg	ccttttgctc	1800
acatgttctt	tcttgcgtta	tcccctgatt	ctgtggataa	ccgtattacc	gcctttgagt	1860
gagctgatac	cgctcgcggc	agccgaacga	ccgagcgcag	cgagtcagtg	agcaggaag	1920
cggagaagcg	cccaatacgc	aaaccgcctc	tccccgcgcg	ttggccgatt	cattaatgca	1980
ggatctcgat	ccgcgaaat	taatacgat	cactataggg	agaccacaac	ggtttccctc	2040
tagaaataat	ttgttttaac	tttaagaagg	agatatacat	atgaaacatc	accaccatca	2100
ccattcggac	tcagaagtca	atcaagaagc	taagccagag	gtcaagccag	aagtcaagcc	2160
tgagactcac	atcaatttaa	aggtgtccga	tggatcttca	gaaatcttct	ttaagatcaa	2220
aaagaccact	cctttaagaa	ggctgatgga	agcgttcgct	aaaagacagg	gtaaggaaat	2280
ggactcctta	agattcttgt	acgacggtat	tagaattcaa	gctgatcaga	cccctgaaga	2340
tttggacatg	gaggataacg	atattattga	ggctcacaga	gaacagattg	gtggcgccgc	2400
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgccgc	2460
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgccgc	2520
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgccgc	2580
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgccgc	2640
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgccgc	2700
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgccgc	2760
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgccgc	2820
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgccgc	2880
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgccgc	2940
tcagctgca	cctgctccag	cagcacctgc	tgcaccagct	ccggtgctc	ctgctgcctg	3000

```

aagagcaagc ttgatccggc tgctaacaa gcccgaaagg aagctgagtt ggctgctgcc 3060
accgctgagc aataactagc ataaccctt ggggcctcta aacgggtctt gaggggtttt 3120
ttgctgaaag gaggaactat atccggatct ggcgtaatag cgaagaggcc cgcaccgatc 3180
gcccttccca acagttgctc agcctgaatg gcgaatggga cgcgcctgt agcggcgcac 3240
taagcgcggc ggggtgtggtg gttacgcgca gcgtgaccgc tacacttgcc agcgccttag 3300
cgcccgctcc tttcgctttc ttcccttctt ttctcgccac gttcgcggc tttccccgtc 3360
aagctctaaa tcgggggctc cctttagggt tccgatttag tgctttacgg cacctcgacc 3420
ccaaaaaact tgattagggg gatggttcac gtagtgggcc atcgccctga tagacggttt 3480
ttcgcccttt gacgttggag tccacgttct ttaatagtgg actcctgttc caaactggaa 3540
caaacactcaa ccctatctcg gtctattctt ttgatttata agggattttg ccgatttcgg 3600
cctattgggtt aaaaaatgag ctgatttaac aaaaatttaa cgcgaatttt aacaaaatat 3660
taacgcttac aatttaggtg 3680

```

5 (B) Análisis de la proteína de fusión His(6)-SUMO-PA#1(200) producida de manera bacteriana y su escisión mediante SDS-PAGE al 12%. El gel muestra la proteína de fusión SUMO-PAS#1(200) extraída a partir de *E. coli* y purificada mediante cromatografía de afinidad de metal inmovilizado (IMAC) y cromatografía de exclusión molecular (SEC) antes (carril 1) y después de la escisión proteolítica con proteasa específica de Ubl 1 (SUMO proteasa) (carril 2) tal como se describe en el ejemplo 21. Todas las muestras se redujeron con 2-mercaptoetanol. Los tamaños de marcadores de proteína (M), aplicados en condiciones reductoras, se indican en el margen izquierdo. La proteína de fusión His(6)-SUMO-PA#1(200) aparece como una única banda homogénea con un tamaño molecular aparente de cerca de 100 kDa. Por tanto, el tamaño aparente para la proteína de fusión His(6)-SUMO-PA#1(200) observado en 10 SDS-PAGE es significativamente más grande que la masa calculada de 28,3 kDa, lo cual se debe a la presencia del polímero/polipéptido de Pro/Ala. Tras la escisión, el polipéptido PA#1(200) hidrófilo no se tiñe de manera detectable mediante azul de Coomassie; por tanto, sólo una pequeña fracción residual de la proteína de fusión y la proteína His(6)-SUMO escindida son visibles en el gel de poli(acrilamida) con SDS (carril 2). La proteína His(6)-SUMO muestra una banda homogénea con un tamaño molecular aparente de cerca de 16 kDa (carril 2) que concuerda bien con su 15 masa molecular calculada de 12,2 kDa.

Figura 13: Conjugación de un polímero/polipéptido de Pro/Ala biosintético con compuestos químicos y/o fármacos.

(A-D) Producción de un conjugado de fluoresceína con un polímero/polipéptido PA#1(200) biosintético (SEQ ID NO: 61) monitorizada mediante cromatografía de exclusión molecular analítica (SEC). Los paneles muestran (de arriba abajo) ejecuciones de SEC de His(6)-SUMO-PA#1(200) purificada (A), His(6)-SUMO-PA#1(200) tras la reacción de escisión en presencia de SUMO proteasa (B), el lote de His(6)-SUMO-PA#1(200) escindida tras acoplamiento químico con un éster de NHS de fluoresceína (C), y el conjugado de fluoresceína-PA#1(200) tras purificación por IMAC (D). Se aplicaron 250 µl de proteína/polipéptido a una concentración de cerca de 0,5 mg/ml a una columna Superdex S200 10/300 GL equilibrada con PBS en un sistema de purificador Äkta. Se monitorizó la absorción a 225 nm, 280 nm y 494 nm usando un detector de UV/VIS UV-900 (GE Healthcare) y se normalizó un pico prominente de cada cromatograma a un valor de 1. La flecha indica el volumen inicial de la columna (7,3 ml). 20 25

(E-K) Caracterización de fluoresceína libre, el polímero/polipéptido PA#1(200) biosintético, y su conjugado de fluoresceína mediante SEC y espectroscopía UV/VIS. Los tres cromatogramas muestran (de arriba abajo) PA#1(200) purificado (E), el compuesto químico fluoresceína (F) y el conjugado de fluoresceína-PA#1(200) purificado (G). Los cuatro espectros UV/VIS muestran la proteína de fusión His(6)-SUMO-PA#1(200) purificada (H), el polímero/polipéptido PA#1(200) purificado (I), fluoresceína libre (J) y el conjugado de fluoresceína-PA#1(200) purificado (K) (todos en PBS). Las flechas indican bandas/hombros de absorción característicos de SUMO (280 nm), PA#1(200) (225 nm) y fluoresceína (494 nm). 30

(L) Curva de calibración para los cromatogramas de (A-G) usando una columna Superdex S200 10/300 GL. El logaritmo del peso molecular (PM) de proteínas marcadoras (aprotinina, 6,5 kDa; citocromo C, 12,4 kDa; anhidrasa carbónica, 29,0 kDa; albúmina sérica bovina, 66,3 kDa; alcohol deshidrogenasa, 150 kDa; β-amilasa, 200 kDa; apoferritina, 440 kDa) se representó gráficamente frente a sus volúmenes de elución (x) y se ajustó mediante una línea recta. A partir de los volúmenes de elución observados de His(6)-SUMO-PA#1(200) (10,81 ml), PA#1(200) (11,51 ml), fluoresceína-PA#1(200) (11,49 ml) y fluoresceína (27,57 ml), se determinaron sus tamaños moleculares aparentes de la siguiente manera. His(6)-SUMO-PA#1(200): 215,6 kDa, PA#1(200): 154,1 kDa (masa auténtica: 16,1 kDa), fluoresceína-PA#1(200): 155,6 kDa (masa auténtica: 16,6 kDa); SUMO: 25,7 kDa (masa auténtica: 12,2 kDa); fluoresceína: 0,09 kDa (masa auténtica: 0,33 kDa). Estos datos muestran que la fusión con el polipéptido/polímero de Pro/Ala confiere un volumen hidrodinámico muy ampliado al fármaco conjugado en comparación con el compuesto sin modificar. 35 40

(M) Caracterización del conjugado químico entre el polipéptido/polímero PA#1(200) biosintético y el compuesto esteroideo digoxigenina mediante espectrometría de masas con ionización por electropulverización (ESI-MS). Un espectro de ESI-MS simplificado de digoxigenina-PA#1(200) revela una masa de 16671,4 Da, lo que coincide esencialmente con la masa calculada para el conjugado de digoxigenina-PA#1(200) (16670,6 Da). 45

Figura 14: Ilustración de conjugados químicos entre el polipéptido/polímero PA#1(200) biosintético y fármacos de

molécula pequeña.

(A) Fluoresceína acoplada al extremo N-terminal de PA#1(200) biosintético.

(B) Digoxigenina acoplada al extremo N-terminal de PA#1(200) biosintético.

Ejemplos

- 5 La presente invención se describe adicionalmente mediante los siguientes ejemplos no limitativos ilustrativos que proporcionan una mejor comprensión de la presente invención y de sus muchas ventajas.

A menos que se indique lo contrario, se usaron métodos establecidos de tecnología génica recombinante tal como se describe, por ejemplo, en Sambrook (2001), citado anteriormente.

Ejemplo 1: Síntesis génica para polímeros/polipéptidos de aminoácidos de Pro/Ala.

- 10 Tal como se describió anteriormente en el presente documento, en el presente documento las repeticiones de aminoácidos que consisten en residuos de Pro y Ala se representan como Pro/Ala o "PA". Se obtuvieron fragmentos génicos que codificaban para una secuencia de polímero repetitiva que comprendía Pro y Ala (PA#1 que corresponde a SEQ ID NO: 1) mediante hibridación de los dos oligodesoxinucleótidos complementarios (SEQ ID NO: 17 y SEQ ID NO: 18) mostrados en la figura 1, seguido por formación de concatémeros de una manera dirigida mediante ligación de ADN de sus extremos cohesivos mutuamente compatibles pero no palindrómicos. Los oligodesoxinucleótidos se adquirieron de ThermoScientific (Ulm, Alemania) y se purificaron mediante electroforesis en gel de poliacrilamida con urea preparativa. Las secuencias de ácido nucleico de los oligodesoxinucleótidos se representan en la figura 1 (SEQ ID NO: 17 y 18) que comprenden un codón de GCC adicional para alanina, que pasa a ser parte de la siguiente repetición de secuencia de PA#1 tras la ligación de los extremos cohesivos correspondientes. Se realizó fosforilación enzimática mezclando 200 pmol de ambos oligodesoxinucleótidos en 100 μ l de Tris/HCl 50 mM pH 7,6, MgCl₂ 10 mM, DTT 5 mM, ATP 1 mM e incubación durante 30 min a 37°C en presencia de 10 u de polinucleótido cinasa (MBI Fermentas, St. Leon-Rot, Alemania). Tras la desnaturalización durante 10 min a 80°C, se enfrió la mezcla hasta temperatura ambiente durante la noche para lograr la hibridación. Después se ligaron 50 μ l de esta disolución añadiendo 1 u de ADN ligasa de T4 (MBI Fermentas) y 10 μ l de Tris/HCl 100 mM pH 7,4, MgCl₂ 50 mM, DTT 20 mM, ATP 10 mM, y en algunos casos 5 mM de cada uno de dATP, dCTP, dGTP y dTTP, en un volumen total de 100 μ l e incubación durante 55 min sobre hielo. Tras 10 min de inactivación térmica a 70°C, se separaron los productos de ligación mediante electroforesis en gel de agarosa al 1,5% (p/v) en presencia de tampón de TAE (Tris 40 mM, ácido acético 20 mM, EDTA 1 mM). Tras la tinción con bromuro de etidio se escindió y aisló la banda correspondiente al segmento génico ensamblado con una longitud de 300 pb.

- 30 Ejemplo 2: Construcción de pFab-PA#1(200) como vector de expresión para una proteína de fusión Fab-PA#1.

Para clonar una repetición de 10 meros del fragmento génico sintético que codifica para la secuencia de 20 aminoácidos de PA#1 del ejemplo 1, se empleó el vector de plásmido pASK88-Fab-2xSapI (SEQ ID NO: 22), un plásmido de expresión para un fragmento Fab (Schlupschy (2007) Protein Eng. Des. Sel. 20:273-284) que albergaba una secuencia de nucleótidos con dos sitios de restricción SapI en orientación complementaria inversa en el extremo 3' de la cadena ligera (figura 2A). Este vector, que es un derivado de pASK75 (Skerra, A. (1994) Gene 151:131-135), se cortó con SapI, se desfosforiló con fosfatasa alcalina de camarón (USB, Cleveland, OH), y se ligó con un casete de 300 pb del fragmento de ADN sintético obtenido del ejemplo 1. Se cortó de nuevo el plásmido pFab-PA#1(100) intermedio resultante con SapI, se desfosforiló con fosfatasa alcalina de camarón y se ligó con un casete de 300 pb del fragmento de ADN sintético obtenido del ejemplo 1 (tal como se muestra a modo de ejemplo en la figura 2B, sin embargo con tan sólo un casete de polímero/polipéptido PA#1(20)). El plásmido resultante se denominó pFab-PA#1(200) (SEQ ID NO: 28) (figura 2C). Debe observarse que en este plásmido la región codificante para la repetición de secuencia de PA#1 de 200 residuos estaba flanqueada por dos restricciones de SapI, lo cual permite la escisión precisa y subclonación adicional del casete de secuencia completa, que porta proyecciones de nucleótidos de GCC en 5'.

- 45 Tras la transformación de *E. coli* XL1-Blue (Bullock (1987) Biotechniques 5: 376-378), se preparó el plásmido y se confirmó la secuencia del inserto de ácido nucleico sintético clonado mediante análisis de restricción y secuenciación de ADN bicatenario (ABI-Prism™310 Genetic analyzer, Perkin-Elmer Applied Biosystems, Weiterstadt, Alemania) usando el kit de terminador BigDye™ así como cebadores de oligodesoxinucleótidos que permitieron la secuenciación desde ambos lados.

- 50 Ejemplo 3: Construcción de pASK-PA#1(200)-IFNa2b como vector de expresión para una proteína de fusión PA#1(200)-IFNa2b.

Para la construcción de un plásmido de expresión que codificaba para IFNa2b como fusión con una repetición de secuencia de PA#1 de 200 residuos, PA#1(200), se cortó pASK-IFNa2b (SEQ ID NO: 32) (figura 2D) con SapI, se desfosforiló con fosfatasa alcalina de camarón y se ligó con el fragmento génico que codificaba para el polipéptido PA#1 de 200 residuos escindido a partir del plásmido pFab-PA#1(200) anteriormente construido (ejemplo 2) mediante digestión de restricción con SapI (tal como se muestra a modo de ejemplo en la figura 2E, sin embargo con

tan sólo un casete de polímero/polipéptido PA#1(20)). Tras la transformación de *E. coli* JM83 (Yanisch-Perron. (1985) Gene 33:103-119), se preparó el plásmido y se confirmó la presencia del inserto correcto mediante análisis de restricción. El plásmido resultante se denominó pPA#1(200)-IFNa2b (SEQ ID NO: 37) (figura 2F).

5 Ejemplo 4: Producción bacteriana y purificación de proteínas de fusión entre un fragmento Fab y un polímero/polipéptido PA#1 genéticamente codificado.

Se produjeron el fragmento Fab (masa calculada: 48,0 kDa) y la fusión Fab-PA#1(200) (masa calculada: 64,3 kDa) a 22°C en *E. coli* KS272 que albergaba los plásmidos de expresión correspondientes del ejemplo 3, junto con el plásmido auxiliar de plegamiento pTUM4 (Schlappschy (2006) Protein Eng. Des. Sel. 19:385-390), usando cultivos en matraz de agitador con 2 l de medio LB que contenía ampicilina 100 mg/l y cloranfenicol 30 mg/l. Se realizó la inducción de expresión génica recombinante mediante adición de 0,4 mg de anhidrotetraciclina a $DO_{550} = 0,5$ durante la noche (dando normalmente como resultado DO_{550} de cerca de 1,0 en la recogida). Se realizó la extracción periplásmica en presencia de sacarosa 500 mM, EDTA 1 mM, Tris/HCl 100 mM pH 8,0 que contenía lisozima 50 $\mu\text{g/ml}$ tal como se describe en otra parte (Breustedt (2005) Biochim. Biophys. Acta 1764:161-173) y se siguió por purificación por medio de la cola de His6 usando cromatografía de afinidad de metal inmovilizado (Skerra (1994) Gene 141: 79-84) con un gradiente de imidazol de desde 0 hasta 200 mM en betaína 500 mM, fosfato de Na 50 mM pH 7,5.

Se obtuvieron preparaciones de proteínas homogéneas para ambos fragmentos Fab recombinantes (figura 3A) con rendimientos de $0,2 \text{ mg l}^{-1} \text{ DO}^{-1}$ para el Fab sin fusionar y $0,1 \text{ mg l}^{-1} \text{ DO}^{-1}$ para Fab-PA#1(200). Se realizó SDS-PAGE usando un sistema de tampón Tris de alta molaridad (Fling (1986) Anal. Biochem. 155: 83-88). Se determinaron las concentraciones de proteínas según la absorción a 280 nm usando coeficientes de extinción calculados (Gill (1989) Anal. Biochem. 182: 319-326) de $68290 \text{ M}^{-1} \text{ cm}^{-1}$ tanto para el Fab sin fusionar como para su fusión con polímero PA#1 ya que el polímero de Pro/Ala no contribuía a la absorción UV debido a su falta de aminoácidos aromáticos.

25 Ejemplo 5: Producción bacteriana y purificación de proteínas de fusión entre IFNa2b y un polímero/polipéptido PA#1 genéticamente codificado.

Se produjeron IFNa2b (masa calculada: 20,9 kDa) y PA#1(200)-IFNa2b (masa calculada: 37,0 kDa) a 22°C en *E. coli* KS272 que albergaba los plásmidos de expresión correspondientes del ejemplo 3, junto con el plásmido auxiliar de plegamiento pTUM4 (Schlappschy (2006), citado anteriormente), usando cultivos en matraz de agitador con 2 l de medio LB que contenía ampicilina 100 mg/l y cloranfenicol 30 mg/l. Se realizó la inducción de expresión génica recombinante mediante adición de 0,4 mg de anhidrotetraciclina a $DO_{550} = 0,5$ durante la noche (dando normalmente como resultado DO_{550} de cerca de 1,0 en la recogida). Se realizó la extracción periplásmica en presencia de sacarosa 500 mM, EDTA 1 mM, Tris/HCl 100 mM pH 8,0 que contenía lisozima 50 $\mu\text{g/ml}$ tal como se describe en otra parte (Breustedt (2005), citado anteriormente) y se siguió por purificación mediante Strep-tag II usando cromatografía de afinidad por estreptavidina (Schmidt (2007) Nat. Protoc. 2:1528-1535) en presencia de NaCl 150 mM, EDTA 1 mM, Tris/HCl 100 mM, pH 8,0.

Se obtuvieron preparaciones de proteínas homogéneas para ambas proteínas IFNa2b recombinantes (figura 3B) con rendimientos de $0,15 \text{ mg l}^{-1} \text{ DO}^{-1}$ para IFNa2b y $0,1 \text{ mg l}^{-1} \text{ DO}^{-1}$ para PA#1(200)-IFNa2b. Se realizó SDS-PAGE usando un sistema de tampón Tris de alta molaridad (Fling (1986), citado anteriormente). Se determinaron las concentraciones de proteínas según la absorción a 280 nm usando coeficientes de extinción calculados (Gill (1989), citado anteriormente) de $23590 \text{ M}^{-1} \text{ cm}^{-1}$ tanto para IFNa2b sin fusionar como para su fusión con polímero PA#1.

Ejemplo 6: Medición del volumen hidrodinámico para la proteína de fusión recombinante entre un fragmento Fab y un polímero PA#1 genéticamente codificado de 200 residuos mediante filtración en gel analítica.

Se llevó a cabo cromatografía de exclusión molecular (SEC) en una columna Superdex S200 HR 10/300 GL (GE Healthcare Europe, Friburgo, Alemania) a una velocidad de flujo de 1 ml/min usando un sistema Äkta Purifier 10 (GE Healthcare) con PBS (NaCl 115 mM, KH_2PO_4 4 mM, Na_2HPO_4 16 mM; pH 7,4) como tampón de corrida. Se aplicaron individualmente muestras de 250 μl del fragmento Fab purificado y su fusión con PA#1 de 200 residuos, obtenidas a partir de la cromatografía de afinidad por afinidad a metal tal como se describió en el ejemplo 4, a una concentración de 0,25 mg/ml en PBS. Ambas proteínas eluyeron en un único pico homogéneo tal como se muestra en la figura 4A.

50 Para la calibración de columna (figura 4B), se aplicaron 250 μl de una mezcla apropiada de las siguientes proteínas globulares (Sigma, Deisenhofen, Alemania) en PBS a concentraciones de proteína de entre 0,2 mg/ml y 0,5 mg/ml: citocromo c, 12,4 kDa; anhidrasa carbónica, 29,0 kDa; ovoalbúmina, 43,0 kDa; albúmina sérica bovina, 66,3 kDa; alcohol deshidrogenasa, 150 kDa; β -amilasa, 200 kDa; apo-ferritina, 440 kDa.

55 Como resultado, la proteína de fusión con el polímero/polipéptido PA#1 de 200 residuos mostró un tamaño significativamente más grande que las proteínas globulares correspondientes con el mismo peso molecular. El aumento de tamaño aparente para Fab-PA#1(200) fue de 7,4 veces en comparación con el fragmento Fab sin fusionar, mientras que la masa auténtica sólo era 1,3 veces mayor. Esta observación indica claramente un volumen

hidrodinámico muy aumentado conferido al fragmento Fab biológicamente activo por el segmento de polipéptido Pro/Ala según esta invención.

Ejemplo 7: Medición del volumen hidrodinámico para la proteína de fusión recombinante entre IFNa2b y un polímero PA#1 genéticamente codificado de 200 residuos mediante filtración en gel analítica.

- 5 Se llevó a cabo cromatografía de exclusión molecular con IFNa2b y PA#1(200)-IFNa2b en una columna Superdex S200 HR 10/300 GL (GE Healthcare) a una velocidad de flujo de 1 ml/min usando un sistema Äkta Purifier 10 (GE Healthcare) de manera similar tal como se describe en el ejemplo 6. Ambas proteínas eluyeron en un único pico homogéneo tal como se muestra en la figura 4C.

- 10 Como resultado, la proteína de fusión con el polímero/polipéptido PA#1 de 200 residuos mostró un tamaño significativamente más grande que las proteínas globulares correspondientes con el mismo peso molecular (figura 4D). El aumento de tamaño aparente para PA#1(200)-IFNa2b fue de 10,2 veces en comparación con la proteína IFNa2b sin fusionar mientras que la masa auténtica sólo fue 1,8 veces mayor. Esta observación indica claramente un volumen hidrodinámico muy aumentado conferido al interferón biológicamente activo por el polímero/polipéptido de Pro/Ala según esta invención.

- 15 Ejemplo 8: Detección de conformación de enrollamiento al azar para el polímero PA#1 biosintético fusionado a un fragmento Fab mediante espectroscopía de dicroísmo circular.

- 20 Se analizó la estructura secundaria usando un espectropolarímetro J-810 (Jasco, Groß-Umstadt, Alemania) equipado con una cubeta de cuarzo 106-QS (longitud de trayecto de 0,1 mm; Hellma, Müllheim, Alemania). Se registraron espectros desde 190 hasta 250 nm a temperatura ambiente acumulando 16 ejecuciones (ancho de banda de 1 nm, velocidad de barrido de 100 nm/min, respuesta de 4 s) usando disoluciones de proteína de 3,12 a 15,4 μ M obtenidas del ejemplo 4 en K_2SO_4 50 mM, fosfato de K 20 mM pH 7,5. Tras la corrección de blancos de disolución, se suavizaron los espectros usando el software del instrumento y se calculó la elipticidad molar Θ_M según la ecuación:

$$\Theta_M = \frac{\Theta_{obs}}{c \cdot d}$$

- 25 mediante la cual Θ_{obs} indica la elipticidad medida, c la concentración de proteína [mol/l], d la longitud de trayecto de la cubeta de cuarzo [cm]. Se representaron gráficamente los valores de Θ_M frente a la longitud de onda usando Kaleidagraph (Synergy Software, Reading, PA).

- 30 El espectro de dicroísmo circular (CD) medido para el Fab recombinante era según plegamiento de inmunoglobulina dominado por láminas β , mientras que el espectro para la proteína de fusión Fab-PA#1(200) reveló una contribución significativa de conformación de enrollamiento al azar (figura 5A). Para analizar la contribución espectroscópica del segmento de polipéptido Pro/Ala en más detalle, se calculó el espectro de CD de diferencia molar con respecto al fragmento Fab sin fusionar (figura 5B) mediante sustracción de este último espectro del de Fab-PA#1(200). Como resultado, se observó un fuerte mínimo alrededor de 200 nm, lo cual es característico de una conformación de enrollamiento al azar. Por tanto, la secuencia de Pro/Ala como parte de la proteína de fusión recombinante parece estar presente como polímero de enrollamiento al azar en condiciones de tampón fisiológico.

- 35 Ejemplo 9: Detección de conformación de enrollamiento al azar para el polímero PA#1 genéticamente codificado fusionado a IFNa2b mediante espectroscopía de dicroísmo circular.

- 40 Se analizó la estructura secundaria mediante mediciones de CD para IFNa2b y PA#1(200)-IFNa2b (obtenido del ejemplo 5) tal como se describió en el ejemplo 8 usando disoluciones de proteína de 3,6 a 38,7 μ M. El espectro de PA#1(200)-IFNa2b reveló contribuciones significativas de estructura secundaria de hélices α , indicativa del plegamiento en haz de hélices α conocido de interferón, así como de conformación de enrollamiento al azar (figura 5C). Para analizar las contribuciones espectroscópicas de la pareja de fusión de polímero de Pro/Ala en más detalle, se calculó el espectro de CD de diferencia molar con respecto a IFNa2b sin fusionar mediante sustracción de los dos espectros individuales (figura 5D). Como resultado, se observó un fuerte mínimo alrededor de 200 nm característico de una conformación de enrollamiento al azar. Por tanto, el segmento de polipéptido Pro/Ala como parte de la proteína de fusión recombinante parece estar presente como polímero de enrollamiento al azar en condiciones de tampón acuoso.

- 45 Ejemplo 10: Análisis cuantitativo de la estructura secundaria del fragmento Fab, de IFNa2b y de sus fusiones con polímero PA#1 de 200 residuos.

- 50 Se cuantificó individualmente el contenido en estructura secundaria del fragmento Fab, Fab-PA#1(200), IFNa2b y PA#1(200)-IFNa2b a partir de los espectros de CD correspondientes medidos en los ejemplos 8 y 9 usando el programa de deconvolución de estructura secundaria CDNN ver. 2.1 (Böhm (1992) Protein Eng. 5:191-195) con un conjunto de 33 espectros de base para la deconvolución de espectros de CD complejos. Los resultados de este análisis se proporcionan en la siguiente tabla:

	Fab	Fab-PA#1 (200)	Dif.: PA#1 (200)	IFNa2b	PA#1 (100)- IFNa2b	Dif.: PA#1 (200)
hélice α	9,5%	7,5%	2,1%	38,2%	31,0%	0,7%
lámina β antiparalela	40,4%	3,1%	0%	1,8%	0,2%	4,6%
lámina β paralela	6,9%	1,3%	0,3%	8,4%	0,7%	0,6%
giro β	6,2%	50,4%	78,6%	19,2%	75,2%	69,7%
enrollamiento al azar	37,2%	63,4%	94,8%	35,9%	64,4%	97,5%
Σ total	100,2%	125,8%	175,8%	103,5%	171,4%	170,0%
Σ de giro β y enrollamiento al azar	43,4%	113,8%	173,4%	55,1%	139,6%	169,1%

En comparación con el contenido en estructura secundaria predominantemente de láminas β del fragmento Fab recombinante, que es según su plegamiento de inmunoglobulina conocido (véase Eigenbrot (1993) J. Mol. Biol. 229:969-995), la fracción de conformación no estructurada (que comprende enrollamiento al azar y giro β) aumenta claramente si se fusiona el polímero PA#1 al fragmento Fab. El espectro de CD de diferencia para el segmento de polipéptido de Pro/Ala revela una clara conformación de enrollamiento al azar. El análisis de la estructura secundaria muestra la presencia de una alta fracción de conformaciones no estructuradas (que comprenden enrollamiento al azar y giro β) que casi comprenden el 100% de la estructura secundaria total. De manera similar, en comparación con el contenido en estructura secundaria predominantemente en hélices α de IFNa2b recombinante, que es según su estructura tridimensional conocida como proteína en haz de hélices α (Radhakrishnan (1996) Structure 4:1453-1463), la fracción de conformación no estructurada para la proteína completa aumenta claramente si se fusiona el polímero PA#1 a IFNa2b. El espectro de CD de diferencia para el segmento de polipéptido de Pro/Ala revela una clara conformación de enrollamiento al azar. El análisis de la estructura secundaria muestra la presencia de una alta fracción de conformaciones no estructuradas (que comprenden enrollamiento al azar y giro β) que casi comprenden el 100% de la estructura secundaria total.

Se obtuvieron diferentes resultados cuando se realizó un análisis teórico de la secuencia de polímero PA#1 usando el algoritmo de Chou-Fasman (Chou y Fasman (1974) Biochemistry 13: 222-245). En la figura 7 se ilustran los resultados de este análisis. Este algoritmo predice la estructura secundaria en hélices α al 100%, lo cual contrasta claramente con los datos experimentales. Por tanto, este algoritmo no es útil para predecir con confianza la conformación no estructurada para un polímero de aminoácidos según la invención.

Ejemplo 11: Construcción de pASK75-His6-PA#1(200)-hGH como vector de expresión para una proteína de fusión His6-PA#1(200)-hGH.

Para la construcción de un plásmido de expresión que codificaba para hGH como fusión con una repetición de secuencia de PA#1 de 200 residuos, PA#1(200), se cortó pASK75-His6-hGH (SEQ ID NO: 41) (figura 6A) con *SapI*, se desfosforiló con fosfatasa alcalina de camarón y se ligó con el fragmento génico que codificaba para el polipéptido PA#1 de 200 residuos escindido del plásmido anteriormente construido pFab-PA#1(200) (ejemplo 2) mediante digestión de restricción con *SapI* (tal como se muestra a modo de ejemplo en la figura 6B, con tan sólo un casete de polímero/polipéptido PA#1(20)). Tras la transformación de *E. coli* JM83 (Yanisch-Perron. (1985), citado anteriormente), se preparó el plásmido y se confirmó la presencia del inserto correcto mediante análisis de restricción. El plásmido resultante se denominó pASK75-His6-PA#1(200)-hGH (SEQ ID NO: 46) (figura 6C).

Ejemplo 12: Construcción de un vector de expresión para la producción secretora de hormona del crecimiento humana fusionada con un polímero/polipéptido PA#1 de 200 residuos en células de ovario de hámster chino.

El vector pASK75-His6-PA#1(200)-hGH (SEQ ID NO: 46), un derivado de pASK75 (Skerra (1994), citado anteriormente), que permite la producción procariota de la proteína de fusión hGH PA#1, se cortó con *NheI* y *HindIII*. Se purificó este fragmento mediante electroforesis en gel de agarosa y se ligó con el vector cortado de manera correspondiente pCHO (SEQ ID NO: 50). Tras la transformación de *E. coli* XL1-Blue (Bullock (1987), citado anteriormente), se preparó el plásmido y se verificó la correcta inserción del fragmento mediante análisis de restricción. El plásmido resultante, que codifica para el péptido señal de hGH fusionado a la cola de His6, un segmento de polipéptido PA#1(200) y la hormona del crecimiento humana (hGH), se denominó pCHO-PA#1(200)-hGH SEQ ID NO: 48) y se representa en la figura 6D.

Ejemplo 13: Producción secretora de una proteína de fusión entre hormona del crecimiento humana (hGH) y el polímero PA#1 genéticamente codificado en células CHO.

Se cultivaron células CHO-K1, ATCC n.º CCL-61, en medio Quantum 263 (PAA Laboratories, Cölbe, Alemania) en una placa de plástico de 100 mm hasta que se alcanzó una confluencia del 50%. Se transfectaron las células con 8 μ g de pCHO-PA#1(200)-hGH (SEQ ID NO: 48) o, para control, pCHO-hGH (SEQ ID NO: 49), un plásmido similar que codificaba para hGH sin la secuencia de PA#1(200), usando el kit Nanofectin (PAA Laboratories, Cölbe, Alemania). Tras 6 h, se cambió el medio de cultivo celular por 7 ml de medio de suero reducido con Opti-MEM®-I (Invitrogen, Darmstadt, Alemania) y se incubaron las células a 37°C en una atmósfera humidificada con CO₂ al 5%. Tras dos días, se tomaron 20 μ l del sobrenadante de cultivo celular y se diluyeron con 5 μ l de tampón de carga de

SDS-PAGE que contenía β -mercaptoetanol. Tras 5 min de calentamiento a 95°C, se sometieron 15 μ l de cada muestra a SDS-PAGE al 12%. Tras la electrotransferencia sobre una membrana de nitrocelulosa (Schleicher & Schuell, Dassel, Alemania) por medio de un aparato de transferencia en semiseco, se lavó la membrana 3 veces durante 15 min con 10 ml de PBST (PBS que contenía Tween 20 al 0,1% v/v). Se incubó la membrana con 10 ml de una dilución 1:1000 de anticuerpo anti-hormona del crecimiento humana ab1956 conjugado con peroxidasa del rábano (Abcam, Cambridge, R.U.). Tras la incubación durante 1 h y el lavado de la membrana dos veces durante 5 min con 20 ml de PBST y dos veces durante 5 min con PBS, se realizó la reacción cromogénica en presencia de 15 ml de disolución de 3,3-diaminobencidina SIGMAFAST™ (Sigma-Aldrich Chemie, Múnich, Alemania). Se detuvo la reacción lavando con agua y secando al aire la membrana. La transferencia reveló señales para ambas muestras de proteína recombinante (figura 6E), proporcionando por tanto la producción secretora de la proteína de fusión de hGH con el polipéptido PA#1 en células CHO.

Ejemplo 14: Producción bacteriana y purificación de proteínas de fusión entre hGH y un polímero/polipéptido PA#1 genéticamente codificado.

Se produjeron hormona del crecimiento humana (hGH) (masa calculada: 23,4 kDa), PA#1(200)-hGH (masa calculada: 39,6 kDa), PA#1(400)-hGH (masa calculada: 55,8 kDa) y PA#1(600)-hGH (masa calculada: 72,0 kDa) en *E. coli* KS272 que albergaba los plásmidos de expresión correspondientes del ejemplo 11 o sus derivados con un casete de secuencia de PA#1 doble (que codificaba para 400 residuos) o triple (600 residuos), respectivamente. Se realizó la producción bacteriana a 22°C en cultivos en matraz de agitador con 2 l de medio LB que contenía glucosa 2,5 g/l, prolina 0,5 g/l y ampicilina 100 mg/l. Se realizó la inducción de expresión génica recombinante mediante adición de 0,4 mg de anhidrotetraciclina a $DO_{550} = 0,5$ durante 3 h. Se llevó a cabo la extracción periplásmica en presencia de sacarosa 500 mM, EDTA 1 mM, Tris/HCl 100 mM pH 8,0 que contenía lisozima 50 μ g/ml tal como se describe en otra parte (Breustedt (2005), citado anteriormente) y se siguió por purificación mediante la cola de His6 usando la columna de afinidad de alto rendimiento HisTrap (GE Healthcare) con fosfato de Na 40 mM pH 7,5, NaCl 0,5 M como tampón. Se eluyeron las proteínas usando un gradiente de concentración de imidazol de desde 0 hasta 150 mM (disuelto en el tampón de corrida y ajustado con HCl a pH 7,5) y se purificaron adicionalmente mediante cromatografía de exclusión molecular usando una columna Superdex 200-HR10/30 (GE Healthcare) equilibrada con PBS (NaCl 115 mM, KH_2PO_4 4 mM, Na_2HPO_4 16 mM, pH 7,4).

Tras la cromatografía de exclusión molecular se obtuvieron preparaciones de proteínas homogéneas para todas las proteínas de fusión de hGH recombinantes sin signos de agregación y con rendimientos de 1 mg l⁻¹ DO⁻¹ para hGH, 0,3 mg l⁻¹ DO⁻¹ para PA#1(200)-hGH, 0,3 mg l⁻¹ DO⁻¹ para PA#1(400)-hGH y 0,2 mg l⁻¹ DO⁻¹ para PA#1(600)-hGH. Se realizó SDS-PAGE usando un sistema de tampón Tris de alta molaridad (Fling (1986), citado anteriormente). Se determinaron las concentraciones de proteínas según la absorción a 280 nm usando coeficientes de extinción calculados (Gill (1989), citado anteriormente) de 16050 M⁻¹ cm⁻¹ para la hGH sin fusionar y todas sus fusiones con polipéptido PA#1.

Ejemplo 15: Medición de la afinidad de unión de hormona del crecimiento humana y sus fusiones con polímero PA#1 frente al dominio extracelular de receptor de hormona del crecimiento humana usando resonancia de plasmón superficial.

Se determinó la afinidad de hGH y sus fusiones con polipéptido PA#1 frente a una proteína de fusión de receptor de hormona del crecimiento humana con Fc (hGHR-Fc; R&D Systems) mediante mediciones en tiempo real de resonancia de plasmón superficial (SPR) en un sistema Biacore 2000 (GE Healthcare). En primer lugar, se inmovilizaron 15 μ l de anticuerpo de captura de ratón anti-IgG humana-Fc (Jackson Immuno Research) a una concentración de 100 μ g/ml en acetato de Na 10 mM pH 5,0 a la superficie de dos canales de flujo de un chip de CMDP (XanTec bioanalytics) usando un kit de acoplamiento de amina (GE Healthcare). Esto dio como resultado cerca de 2700 unidades de respuesta (UR). Tras el equilibrado con PBS/T (PBS que contenía Tween 20 al 0,05% (v/v)) como tampón de flujo, se cargó un canal del chip con hGHR-Fc 2 μ g/ml a una velocidad de flujo de 5 μ l/min hasta que se alcanzó una señal adicional de cerca de 300 UR. Después, se inyectaron 75 μ l de hGH o sus fusiones con polipéptido PA#1 en PBS/T a concentraciones variables y se midieron las fases de asociación y disociación con flujo de tampón continuo de 20 μ l/min. Para la regeneración, se aplicaron tres pulsos de 6 μ l de glicina 10 mM/HCl pH 2,7. Se corrigieron los sensogramas mediante sustracción doble de las señales correspondientes medidas para el canal sin receptor inmovilizado y se determinó una referencia promedio determinada a partir de varias inyecciones de blanco de tampón (Myszka (1999) Mol. Recognit. 12: 279-284). Se realizó la evaluación de datos de cinética mediante ajuste global de los perfiles a partir de al menos siete inyecciones de muestra diferentes según el modelo de unión de Langmuir 1:1 usando el software BIAevaluation versión 3.1 (GE Healthcare). En la siguiente tabla se resumen los valores obtenidos a partir de mediciones de SPR para las constantes de equilibrio cinética y derivada de los complejos entre hGH o sus fusiones con PA#1 y el receptor de hormona del crecimiento humana:

Variante de hGH	k_{on} [$10^5 M^{-1} s^{-1}$]	k_{off} [$10^{-6} s^{-1}$]	K_D [pM]
hGH	10,2	10,6	10,4
PA#1(200)-hGH	4,75	9,18	19,3
PA#1(400)-hGH	3,26	14,0	42,9
PA#1(600)-hGH	3,29	12,5	38,0

Estos datos muestran que la fusión de hGH con polipéptidos PA#1 de diferentes longitudes no interfiere significativamente con la unión a receptor. Todas las fusiones con polipéptido PA#1 de hGH conservan actividad de unión a receptor dentro de un factor 5 en comparación con la hGH recombinante que carece de un polipéptido PA#1.

5 Ejemplo 16: Detección de semivida en plasma prolongada *in vivo* para las proteínas de fusión recombinantes entre un fragmento Fab y polímeros PA#1 genéticamente codificados.

Se realizaron inyecciones intravenosas a ratones BALB/c adultos (población reproductora SPF; TU München, Freising, Alemania) según la siguiente tabla:

Grupo	A	B	D
Elemento de prueba	Fab	Fab-PA#1(200)	Fab-PA#1(600)
Vía de administración	intravenosa		
Dosis [mg/kg de p.c.]	5,0	5,0	5,0
Concentración [mg/ml]	1,0	1,0	1,0
Volumen de aplicación [ml/kg de p.c.]	5,0		
N.º de animales/grupo	9	9	9
N.º de puntos de tiempo de toma de muestras de sangre	12	12	12
N.º de animales/punto de tiempo de toma de muestras	3	3	3
N.º de tomas de muestras de sangre/animal	4/1	4/1	4/1

10 El volumen total de elemento de prueba administrado por vía intravenosa se calculó según el peso corporal individual (p.c.) registrado en el día de administración (por ejemplo, un animal con 20 g de peso corporal recibió 100 µl de elemento de prueba 1 mg/ml). Se realizaron tomas de muestras de sangre según la siguiente tabla:

Elemento de prueba	Subgrupo	Puntos de tiempo para tomas de muestras de sangre tras la inyección											
		10 min	30 min	1 h	2 h	3 h	4 h	6 h	8 h	12 h	24 h	36 h	48 h
Fab Fab-PA#1 (200) Fab-PA#1 (600)	1	X			X			X			X		
		X			X			X			X		
		X			X			X			X		
	2		X			X			X			X	
			X			X			X			X	
			X			X			X			X	
	3			X			X			X			X
				X			X			X			X
				X			X			X			X

15 Para cada sustancia (elemento de prueba) se realizaron inyecciones en total a nueve animales (divididos en tres subgrupos 1-3 con tres animales cada uno), proporcionando cada uno cuatro muestras en diferentes puntos de tiempo. Se tomaron muestras de sangre (aproximadamente 50 µl) de la vena de la cola y se almacenaron a 4°C durante 30 min. Tras la centrifugación durante 10 min a 10.000 g y 4°C, se congeló inmediatamente el sobrenadante (plasma) y se almacenó a -20°C.

20 Para la detección cuantitativa de la proteína de fusión de Fab en un ELISA, se recubrieron los pocillos de una placa de microtitulación de 96 pocillos (Maxisorb, NUNC, Dinamarca) durante la noche a 4°C con 50 µl de una disolución 10 µg/ml de antígeno de ectodominio Her2/ErbB2 recombinante en NaHCO₃ 50 mM pH 9,6. Después, se bloquearon los pocillos con 200 µl de BSA al 3% (p/v) en PBS durante 1 h y se lavaron tres veces con PBS/T (PBS que contenía Tween 20 al 0,1% (v/v)). Se aplicaron las muestras de plasma en serie de dilución en PBS/T que contenía plasma de ratón al 0,5% (v/v) de un animal no tratado y se incubaron durante 1 h. Después se lavaron los pocillos tres veces con PBS/T y se incubaron durante 1 h con 50 µl de una disolución diluida 1:1000 de un conjugado con fosfatasa alcalina de anticuerpo anti-C_κ humana en PBS/T. Tras lavar dos veces con PBS/T y dos veces con PBS se inició la reacción cromogénica añadiendo 50 µl de fosfato de p-nitrofenilo 0,5 µg/ml en Tris/HCl 100 mM pH 8,8, NaCl 100 mM, MgCl₂ 5 mM como sustrato, y tras 15 min a 25°C se midió la absorbancia a 405 nm. Se cuantificaron las concentraciones de Fab, Fab-PA#1(200) y Fab-PA#1(600) en las muestras de plasma mediante comparación de las señales medidas con curvas patrón que se determinaron para series de dilución para las proteínas purificadas correspondientes a concentraciones definidas en PBS/T que contenía plasma de ratón al 0,5% (v/v) de animales no tratados.

30 Para estimar la semivida en plasma de Fab, Fab-PA#1(200) y Fab-PA#1(600), se determinaron los valores de concentración, c(t), para cada punto de tiempo a partir de mediciones de ELISA y se representaron gráficamente frente al tiempo tras la inyección intravenosa, t. Se ajustaron numéricamente estos datos usando el software KaleidaGraph suponiendo una disminución biexponencial según la ecuación

$$c(t) = c_{\alpha} e^{-\ln 2 \frac{t}{\tau_{1/2}^{\alpha}}} + (c_0 - c_{\alpha}) e^{-\ln 2 \frac{t}{\tau_{1/2}^{\beta}}}$$

mediante lo cual $\tau_{1/2}^{\alpha}$ y $\tau_{1/2}^{\beta}$ son los valores de semivida de la fase de distribución α y la fase de eliminación β , respectivamente. c_0 es la concentración en sangre total en el punto de tiempo cero mientras que c_{α} es la amplitud de concentración para la fase de distribución.

- 5 La figura 8 representa la farmacocinética para los tres elementos de prueba en ratones BALB/c. Mientras que el Fab recombinante muestra un aclaramiento en sangre rápido con una semivida de eliminación de tan sólo cerca de 1,3 h, las proteínas de fusión Fab-PA#1(200) y Fab-PA#1(600) tienen una semivida prolongada más de 3 veces y 29 veces con valores correspondientes de cerca de 4,1 h y 38,8 h, respectivamente. Estos datos demuestran que la semivida en plasma *in vivo* de un fragmento Fab se prolonga significativamente debido a la fusión con un polímero/polipéptido de Pro/Ala, mediante lo cual la semivida se vuelve más larga con el aumento de la longitud del polímero de aminoácidos.

Ejemplo 17: Síntesis génica para polímeros/polipéptidos de aminoácidos P1A1 y P1A3 y construcción de pFab-P1A1(200) y pFab-P1A3(200) como vectores de expresión para proteínas de fusión Fab-P1A1(200) y Fab-P1A3(200).

- 15 Se obtuvieron fragmentos génicos que codificaban para una secuencia de polímero repetitiva que comprendía los polipéptidos/polímeros de Pro/Ala P1A1 (SEQ ID NO: 51) y P1A3, también denominado PA#3 (SEQ ID NO: 3), mediante hibridación de pares de oligodesoxinucleótidos complementarios, respectivamente, SEQ ID NO: 52 y SEQ ID NO: 53 para P1A1 y SEQ ID NO: 54 y SEQ ID NO: 55 para P1A3 tal como se describió en el ejemplo 1. Se construyeron pFab-P1A1(200) (SEQ ID NO: 58) y pFab-P1A3(200) (SEQ ID NO: 59) que codificaban para fragmentos Fab con los segmentos de polímeros/polipéptidos de Pro/Ala correspondientes de 200 residuos en el extremo C-terminal de la cadena ligera (LC) (secuencia de aminoácidos de LC Fab-P1A1(200): SEQ ID NO: 56; secuencia de aminoácidos de LC Fab-P1A3(200): SEQ ID NO: 57) de una manera análoga a pFab-PA#1(200), que se describió en el ejemplo 2.

- 25 A continuación también se reproducen las siguientes SEQ ID NO: 56, 57, 58 y 59. Sin embargo, estas secuencias también están comprendidas en la lista de secuencias adjunta que es una parte específica de esta divulgación y la descripción de la presente invención.

SEQ ID NO: 56

Asp Ile Glu Leu Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Val Asn Thr Ala
 20 25 30

ES 2 691 642 T3

Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60

Ser Arg Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro
85 90 95

Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Thr Val Ala Ala
100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
195 200 205

Phe Asn Arg Gly Glu Cys Ser Ser Ala Pro Ala Pro Ala Pro Ala Pro
210 215 220

Ala Pro
225 230 235 240

Ala Pro
245 250 255

Ala Pro
260 265 270

Ala Pro
275 280 285

Ala Pro
290 295 300

Ala Pro
305 310 315 320

Ala Pro
325 330 335

Ala Pro
340 345 350

ES 2 691 642 T3

Ala Pro
 355 360 365

Ala Pro
 370 375 380

Ala Pro
 385 390 395 400

Ala Pro
 405 410 415

Ala

SEQ ID NO: 57

Asp Ile Glu Leu Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Val Asn Thr Ala
 20 25 30

Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60

Ser Arg Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Cys Ser Ser Ala Ala Ala Pro Ala Ala Ala Pro
 210 215 220

ES 2 691 642 T3

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 225 230 235 240

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 245 250 255

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 260 265 270

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 275 280 285

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 290 295 300

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 305 310 315 320

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 325 330 335

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 340 345 350

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 355 360 365

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 370 375 380

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 385 390 395 400

Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro Ala Ala Ala Pro
 405 410 415

Ala

SEQ ID NO: 58

```

accgcacacc atcgaatggc cagatgatta attcctaatt tttgttgaca ctctatcatt      60
gatagagtta ttttaccact ccctatcagt gatagagaaa agtgaaatga atagttcgac      120
aaaaatctag ataacgaggg caaaaaatga aaaagacagc tatcgcgatt gcagtggcac      180
tggctggttt cgctaccgta gcgcaggccg aagttaaact gcaggaatcc ggtgggtggtc      240
tggttcagcc aggtggttcc ctgcggtctt cgtgtgctgc ttccggtttc aacatcaaag      300
acacctacat ccaactgggtt cgtcaggctc cgggtaaagg cctggaatgg gttgctcgta      360
tctaccogac caacggttac accaggtatg cggattcagt taaaggctgt ttcaccatct      420
cggcogacac ttccaaaaaac accgcttacc tccagatgaa ctccctgctg gctgaagaca      480
cagctgttta ttattgctcc cgttgggggt gtgacggttt ctacgctatg gactactggg      540
gtcagggtag cctggtcacc gtctcctcag cctccacca gggcccatcg gtcttcccc      600
tggcaccctc ctccaagagc acctctgggg gcacagcggc cctgggctgc ctgggtcaagg      660
actacttccc cgaaccgggtg acggtgtcgt ggaactcagg cgccctgacc agcggcgtgc      720
acaccttccc ggctgtccta cagtcctcag gactctactc cctcagcagc gtgggtgactg      780
tgccctccag cagcttgggc acccagacct acatctgcaa cgttaatcac aaaccagca      840
acaccaaggt cgacaagaaa gttgagccca aatcttgcca tcaccacat caccattaat      900
aaccatggag aaaataaagt gaaacaaagc actattgcac tggcactctt accgttactg      960
tttaaccctg tgacaaaagc cgacatcgag ctcacccaat ccccgctctc cctgtccgct     1020
tccgttggcg accgtgttac catcacgtgt agggcctcgc aagacgtaa caccgccgta     1080
gcgtggatc agcagaaacc cgggaaagct ccgaaactgc tgatctatag cgcttcttc     1140
ctgtattccg gagttccgag caggttcagt ggttcccgtt ccggtaccga cttcacctg     1200
  
```

ES 2 691 642 T3

aogatatcct	ccctccagcc	ggaagacttc	gctacctact	actgtcaaca	gcactacacc	1260
accccgcgca	ccttcgggtca	gggtaccaa	ctcgagatca	aacggactgt	ggctgcacca	1320
tctgtcttca	tcttcccgcc	atctgatgag	cagttgaaat	ctggaactgc	ctctgttgtg	1380
tgctgtctga	ataacttcta	tcccagagag	gccaaagtac	agtggaaggt	ggataacgcc	1440
ctccaatcgg	gtaactccca	ggagagtgtc	acagagcagg	acagcaagga	cagcacctac	1500
agcctcagca	gcaccctgac	gctgagcaaa	gcagactacg	agaaacacaa	agtctacgcc	1560
tgcgaaagtca	cccatcaggg	cctgagttcg	cccgtcacaa	agagcttcaa	ccgcggagag	1620
tgtctctctg	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	1680
cctgcaccag	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	1740
cctgcaccag	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	1800
cctgcaccag	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	1860
cctgcaccag	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	1920
cctgcaccag	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	1980
cctgcaccag	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	2040
cctgcaccag	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	2100
cctgcaccag	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	2160
cctgcaccag	ccctgctcc	tgctccagca	cctgcaccag	cacctgctcc	agcaccagct	2220
cctgcaccag	cctgaagagc	ttaagcttga	cctgtgaagt	gaaaaatggc	gcacattgtg	2280
cgacattttt	tttgtctgcc	gtttaccgct	actgctcac	ggatctccac	gcgccctgta	2340
gcgcgccatt	aagcgcggcg	ggtgtggtgg	ttacgcgcag	cgtgaccgct	acacttgcca	2400
gogccctagc	gcccgtcct	ttcgctttct	tccttctct	tctcgccacg	ttcgccggt	2460
ttcccgcgca	agctctaaat	cgggggctcc	ctttaggggt	ccgatttagt	gctttacggc	2520
acctcgacc	caaaaaactt	gattaggggt	atggttcacg	tagtgggcca	tcgcccgtgat	2580
agacggtttt	tcgccccttg	acgttggagt	ccagttctt	taatagtgga	ctcttcttcc	2640
aaactggaac	aacactcaac	cctatctcgg	tctattcttt	tgatttataa	gggattttgc	2700
cgatttcggc	ctattgggta	aaaaatgagc	tgatttaaca	aaaatttaac	gcgaatttta	2760
acaaaatatt	aacgtttaca	atctcaggtg	gcactttctg	gggaaatgtg	cgcggaacc	2820
ctatttggtt	atcttctaa	atacattcaa	atatgtatcc	gctcatgaga	caataaccct	2880
gataaatgct	tcaataatat	tgaaaaagga	agagtatgag	tattcaacat	ttcctgtctg	2940
cccttattcc	cttttttgcg	gcattttgcc	ttcctgtttt	tgctcaccca	gaaacgctgg	3000
tgaaagtaaa	agatgctgaa	gatcagttgg	gtgcacgagt	gggttacatc	gaactggatc	3060
tcaacagcgg	taagatcctt	gagagttttc	gccccgaaga	acgttttcca	atgatgagca	3120
cttttaaagt	tctgctatgt	ggcgcggtat	tatcccgtat	tgacgcgggg	caagagcaac	3180
tcggctgcgg	catacactat	tctcagaatg	acttggttga	gtactcacca	gtcacagaaa	3240
agcattctac	ggatggcatg	acagtaagag	aattatgcag	tgctgccata	accatgagtg	3300
ataacactgc	ggccaactta	cttctgacaa	cgatcggagg	accgaaggag	ctaaccgctt	3360
ttttgcacaa	catgggggat	catgtaactc	gccttgatcg	ttgggaaccg	gagctgaatg	3420
aagccatacc	aaacgacgag	cgtgacacca	cgatgcctgt	agcaatggca	acaacgcttg	3480
gcaaaactatt	aactggcgaa	ctacttactc	tagcttcccg	gcaacaattg	atagactgga	3540
tgaggcggga	taaagttgca	ggaccacttc	tgcgctcggc	ccttccggct	ggctgggtta	3600
ttgctgataa	atctggagcc	ggtgagcgtg	gctctcggg	tatcattgca	gcactggggc	3660
cagatggtaa	gccctcccgt	atcgtagtta	tctacacgac	ggggagttag	gcaactatgg	3720
atgaacgaaa	tagacagatc	gctgagatag	gtgcctcact	gattaagcat	tggttaggaat	3780
taatgatgtc	tcgttttagat	aaaagtaaa	tgattaacag	cgcattagag	ctgcttaaatg	3840
aggctoggaat	cgaaggttta	acaaccgta	aactcgccca	gaagctaggt	gtagagcagc	3900
ctacattgta	ttggcatgta	aaaaataagc	gggctttgct	cgacgcctta	gccattgaga	3960
tgttagatag	gcaccatact	cacttttgcc	ctttagaagg	ggaaagctgg	caagattttt	4020
tacgtaataa	cgctaaaagt	tttagatgtg	ctttactaag	tcacgcgat	ggagcaaaag	4080
tacatttagg	tacacggcct	acagaaaaac	agtatgaaac	tctcgaaaat	caattagcct	4140
ttttatgcca	acaaggtttt	tcactagaga	atgcattata	tgcaactcagc	gcagtggggc	4200
atcttaactt	aggttgcgta	ttggaagatc	aagagcatca	agtcgctaaa	gaagaaagg	4260
aaacacctac	tactgatagt	atgccgccat	tattaogaca	agctatcgaa	ttatttgatc	4320
accaaggtgc	agagccagcc	ttcttattcg	gccttgaatt	gatcatatgc	ggattagaaa	4380
aacaacttaa	atgtgaaagt	gggtcttaaa	agcagcataa	cctttttccg	tgatggtaac	4440
ttcactagtt	taaaaggatc	taggtgaaga	tcctttttga	taatctcatg	acaaaaatcc	4500
cttaacgtga	gttttcgttc	cactgagcgt	cagaccccgt	agaaaaagatc	aaaggatctt	4560
cttgagatcc	ttttttctg	cgcgtaatct	gctgcttgc	aacaaaaaaa	ccaccgctac	4620
cagcgggtgg	ttggttgccg	gatcaagagc	taccaactct	ttttccgaag	gtaactggct	4680
tcagcagagc	gcagatacca	aatactgtcc	ttctagtgt	gcoctagtta	ggccaccact	4740
tcaagaactc	tgtagcaccg	cctacatacc	tcgctctgct	aatcctgtta	ccagtggctg	4800

ES 2 691 642 T3

ctgccagtgg	cgataagtgc	tgtcttaccg	ggttgactc	aagacgatag	ttaccggata	4860
aggcgcagcg	gtcgggctga	acggggggtt	cgtgcacaca	gccagcttg	gagcgaacga	4920
cctacaccga	actgagatac	ctacagcgtg	agctatgaga	aagcgccacg	cttcccgaag	4980
ggagaaaggc	ggacaggtat	ccggtaagcg	gcagggtcgg	aacaggagag	cgcacgaggg	5040
agcttccagg	gggaaacgcc	tggtatcttt	atagtctgt	cgggtttcgc	cacctctgac	5100
ttgagcgtcg	atTTTTgtga	tgctcgtcag	gggggocggag	cctatggaaa	aacgccagca	5160
acgoggcctt	tttacggttc	ctggcctttt	gctggccttt	tgctcacatg		5210

SEQ ID NO: 59

accogacacc	atcgaatggc	cagatgatta	attcctaatt	tttgttgaca	ctctatcatt	60
gatagagtta	ttttaccact	ccctatcagt	gatagagaaa	agtgaaatga	atagttcgcac	120
aaaaatctag	ataacgaggg	caaaaaatga	aaaagacagc	tatcgcgatt	gcagtggcac	180
tggctggttt	cgctaccgta	gcgcaggccg	aagttaaact	gcaggaatcc	ggtggtggtc	240
tggttcagcc	aggtggttcc	ctgcggctct	cgtgtgctgc	ttccggtttc	aacatcaaag	300
acacctacat	ccactgggtt	cgtcaggctc	cgggtaaagg	cctggaatgg	gttgctcgtg	360
tctacccgac	caacggttac	accaggtatg	ccgattcagt	taaaggctgt	ttcaccatct	420
cggccgacac	ttccaaaaac	accgcttacc	tccagatgaa	ctccctgcgt	gctgaagaca	480
cagctgttta	ttattgctcc	cgttgggggtg	gtgacggttt	ctacgctatg	gactactggg	540
gtcagggtac	cctggtcacc	gtctctcag	cctccacca	gggcccatcg	gtcttcccc	600
tggcaccctc	ctccaagagc	acctctgggg	gcacagcggc	cctgggctgc	ctgggtcaagg	660
actacttccc	ogaaccggtg	acgggtgctg	ggaactcagg	cgccctgacc	agcggcgtgc	720
acacctccc	ggctgtccta	cagtctcag	gactctactc	cctcagcagc	gtggtgactg	780
tgccctccag	cagcttgggc	accagacct	acatctgcaa	cgtaatacac	aaaccagca	840
acaccaaggt	cgacaagaaa	gttgagccca	aatcttgcca	tcaccacat	caccattaat	900
aaccatggag	aaaataaagt	gaaacaaagc	actattgcac	tggcactctt	accgttactg	960
tttaccctcg	tgacaaaagc	cgacatcgag	ctcaccat	ccccgtctc	cctgtccgct	1020
tccgttggcg	accgtgttac	catcacgtgt	agggcctcgc	aagacgtaaa	caccgccgta	1080
gcgtggatc	agcagaaacc	cgggaaagct	ccgaaactgc	tgatctatag	cgcttctctc	1140
ctgtattccg	gagttccgag	caggttcagt	ggttcccgtt	ccggtaccga	cttcaccctg	1200
acgatatcct	cctccagcc	ggaagacttc	gctacctact	actgtcaaca	gcactacacc	1260
accccgccga	ccttcgggtca	gggtaccaaa	ctcgagatca	aacggactgt	ggctgcacca	1320
tctgtcttca	tcttcccgcc	atctgatgag	cagttgaaat	ctggaactgc	ctctgttgtg	1380
tgcttctga	ataacttcta	tcccagagag	gccaaagtac	agtggaaggt	ggataacgcc	1440
ctccaatcgg	gtaactccca	ggagagtgtc	acagagcagg	acagcaagga	cagcacctac	1500
agcctcagca	gcaccctgac	gctgagcaaa	gcagactacg	agaaacacaa	agtctacgcc	1560
tgcaagtc	cccatcaggg	cctgagttcg	cccgtcacaa	agagcttcaa	ccgaggagag	1620
tgctctctg	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	1680
gcagctccag	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	1740
gcagctccag	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	1800
gcagctccag	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	1860
gcagctccag	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	1920
gcagctccag	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	1980
gcagctccag	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	2040
gcagctccag	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	2100
gcagctccag	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	2160
gcagctccag	ccgctgcacc	tgctgcagca	cctgctgcag	ctccagcagc	tgctcctgca	2220
gcagctccag	cctgaagagc	ttaagcttga	cctgtgaagt	gaaaaatggc	gcacattgtg	2280
cgacattttt	tttgtctgcc	gtttaccgct	actgcgtcac	ggatctccac	gcgcctgtga	2340
gogggcatt	aagcgcggcg	ggtgtggtgg	ttacgcgcag	cgtgaccgct	acacttgcca	2400
gogccctagc	gcccgcctct	ttcgccttct	tcccttccct	tctcgccacg	ttcggcggct	2460
ttcccgcgca	agctctaaat	cgggggctcc	ctttaggggt	ccgatttagt	gctttacggc	2520
acctcgaccc	caaaaaactt	gattaggggtg	atggttcacg	tagtgggcca	tcgcctgat	2580
agacggtttt	tcgccctttg	acgttggagt	ccacgttctt	taatagtgga	ctcttgttcc	2640
aaactggaac	aacactcaac	cctatctcgg	tctattcttt	tgatttataa	gggattttgc	2700
cgatttcggc	ctattggtta	aaaaatgagc	tgatttaaca	aaaatttaac	gcgaatttta	2760
acaaaatatt	aacgtttaca	atctcaggtg	gcacttttcg	gggaaatgtg	cgcggaacc	2820
ctatttgttt	atTTTTctaa	atacattcaa	atatgtatcc	gctcatgaga	caataaccct	2880
gataaatgct	tcaataatat	tgaaaaagga	agagatgag	tattcaacat	ttccgtgtcg	2940
cccttatctc	ctTTTTtgcg	gcattttgcc	ttcctgtttt	tgctcaccca	gaaacgctgg	3000
tgaaagtaaa	agatgctgaa	gatcagttgg	gtgcacagat	gggttacatc	gaactggatc	3060

ES 2 691 642 T3

tcaacagcgg	taagatcctt	gagagttttc	gccccgaaga	acgtttttcca	atgatgagca	3120
cttttaagt	tctgctatgt	ggcgcggtat	tatcccgtat	tgacgcggg	caagagcaac	3180
tcggtcgcg	catacactat	tctcagaatg	acttggttga	gtactcacca	gtcacagaaa	3240
agcatcttac	ggatggcatg	acagtaagag	aattatgcag	tgctgccata	accatgagtg	3300
ataaactgc	ggccaactta	cttctgacaa	cgatcggagg	accgaaggag	ctaaccgctt	3360
ttttgcacaa	catgggggat	catgtaactc	gccttgatcg	ttgggaaccg	gagctgaatg	3420
aagccatacc	aaacgacgag	cgtgacacca	cgatgcctgt	agcaatggca	acaacgttgc	3480
gcaaactatt	aactggcgaa	ctacttactc	tagcttcccg	gcaacaattg	atagactgga	3540
tggaggcgga	taaagttgca	ggaccacttc	tgcgctcggc	ccttccggct	ggctggttta	3600
ttgctgataa	atctggagcc	ggtgagcgtg	gctctcggcg	tatcattgca	gcactggggc	3660
cagatggtaa	gccctcccgt	atcgtagtta	tctacacgac	ggggagtcag	gcaactatgg	3720
atgaacgaaa	tagacagatc	gctgagatag	gtgcctcact	gattaagcat	tggtaggaat	3780
taatgatgtc	tcgttttagat	aaaagtaaag	tgattaacag	cgcattagag	ctgcttaatg	3840
aggtcggaat	cgaaggttta	acaacccgta	aactcgccca	gaagctaggt	gtagagcagc	3900
ctacattgta	ttggcatgta	aaaaataaag	gggctttgct	cgacgcctta	gccattgaga	3960
tgttagatag	gcaccatact	cacttttgcc	ctttagaag	ggaaagctgg	caagattttt	4020
tacgtaataa	cgctaaaagt	tttagatgtg	ctttactaag	tcatcgcgat	ggagcaaaag	4080
tacatttagg	tacacggcct	acagaaaaac	agtatgaaac	tctcgaaaat	caattagcct	4140
ttttatgcca	acaaggtttt	tcactagaga	atgcattata	tgactcagc	gcagtggggc	4200
attttacttt	aggttgcgta	ttggaagatc	aagagcatca	agtcgctaaa	gaagaaagg	4260
aaacacctac	tactgatagt	atgccgccat	tattacgaca	agctatcgaa	ttatttgatc	4320
accaaggtgc	agagccagcc	ttcttattcg	gccttgaatt	gatcatatgc	ggattagaaa	4380
aacaacttaa	atgtgaaagt	gggtcttaaa	agcagcataa	cctttttccg	tgatggtaac	4440
ttcactagtt	taaaaggatc	taggtgaaga	tcctttttga	taatctcatg	acaaaaatcc	4500
cttaacgtga	gttttcgttc	cactgagcgt	cagaccccgt	agaaaagatc	aaaggatcct	4560
cttgagatcc	tttttttctg	cgcgtaatct	gctgcttgca	aacaaaaaaa	ccaccgctac	4620
cagcggtggt	ttgtttgccg	gatcaagagc	taccaactct	ttttccgaag	gtaactggct	4680
tcagcagagc	gcagatacca	aatactgtcc	ttctagtgtg	gccgtagtta	ggccaccact	4740
tcaagaactc	tgtagcaccg	cctacatacc	tcgctctgct	aatcctgtta	ccagtggctg	4800
ctgccagtg	cgataagtcg	tgtcttaccg	ggttggactc	aagacgatag	ttaccggata	4860
aggcgcagcg	gtcgggctga	acggggggtt	cgtgcacaca	gccagcttg	gagcgaacga	4920
cctacaccga	actgagatac	ctacagcgtg	agctatgaga	aagcgcacg	cttcccgaag	4980
ggagaaaggc	ggacaggtat	ccgtaaacg	gcagggtcgg	aacaggagag	cgcacgagg	5040
agcttccagg	gggaaacg	tggtatcttt	atagtcctgt	cgggtttcgc	cacctctgac	5100
ttgagcgtcg	atttttgtga	tgctcgtcag	ggggcgagg	cctatggaaa	aacgccagca	5160
acgcggcctt	tttacggctt	ctggcctttt	gctggccttt	tgctcacatg		5210

Ejemplo 18: Medición del volumen hidrodinámico para la proteína de fusión recombinante entre un fragmento Fab y un polipéptido/polímero P1A1 o P1A3 genéticamente codificado mediante filtración en gel analítica.

- 5 Se llevó a cabo SEC en una columna Superdex S200 HR 10/300 GL (GE Healthcare Europe, Friburgo, Alemania) a una velocidad de flujo de 1 ml/min usando un sistema Äkta Purifier 10 (GE Healthcare) con PBS como tampón de corrida. Se aplicaron de manera individual muestras de 250 µl de las proteínas de fusión Fab-P1A1(200) y Fab-P1A3(200), que se produjeron y purificaron de manera similar (figura 9) tal como se describió para Fab-PA#1(200) en el ejemplo 4, a una concentración de 0,25 mg/ml en PBS. Ambas proteínas eluyeron en un único pico homogéneo tal como se muestra en la figura 10.
- 10 Como resultado, las proteínas de fusión con los polímeros/polipéptidos P1A1 o P1A3 de 200 residuos mostraron tamaños significativamente mayores que el fragmento Fab sin fusionar correspondiente. El aumento de tamaño aparente para Fab-P1A1(200) y Fab-P1A3(200) fue de 5,8 veces y 5,2 veces, respectivamente, en comparación con el fragmento Fab (véase la figura 4B), mientras que la masa auténtica sólo era 1,4 veces y 1,3 veces mayor. Esta observación indica claramente un volumen hidrodinámico muy aumentado conferido al fragmento Fab
- 15 biológicamente activo por los segmentos de polipéptido P1A1 y P1A3 biosintéticos según esta invención.

Ejemplo 19: Detección de conformación de enrollamiento al azar para los polímeros/polipéptidos P1A1 y P1A3 biosintéticos fusionados a un fragmento Fab mediante espectroscopía de dicroísmo circular (CD).

- 20 Se registraron espectros de CD para Fab-P1A1(200) y Fab-P1A3(200) tal como se describió en el ejemplo 8 usando disoluciones de proteína 4,2 y 6,5 µM, respectivamente, preparadas de manera similar tal como se describió en el ejemplo 4 usando K₂SO₄ 50 mM, fosfato de K 20 mM pH 7,5 como tampón.

Los espectros para las proteínas de fusión Fab-P1A1(200) y Fab-P1A3(200) revelaron una fracción significativa de conformación de enrollamiento al azar (figura 11A). Para analizar la contribución espectroscópica mediante el

segmento de polipéptido Pro/Ala en mayor detalle, se calculó el espectro de CD de diferencia molar con respecto al fragmento Fab sin fusionar (véase el ejemplo 8) (figura 11B) restando este último espectro del de Fab-P1A1(200) y Fab-P1A3(200), respectivamente, tras la normalización a la misma concentración molar. Como resultado, se observó un fuerte mínimo a una longitud de onda de aproximadamente 200 nm, lo cual es característico de la conformación de enrollamiento al azar. Por tanto, las secuencias P1A1 y P1A3 como parte de la proteína de fusión recombinante parecen estar presentes en la conformación de enrollamiento al azar en condiciones de tampón fisiológico.

Ejemplo 20: Construcción de pSUMO-PA#1(200) como vector de expresión para una proteína de fusión His(6)-SUMO-PA#1(200).

Para la construcción de un plásmido de expresión que codifica para una cola de His de seis residuos y la proteína modificador similar a ubiquitina pequeño (SUMO) (Panavas (2009) *Methods Mol. Biol.* 497: 303-17) fusionada a una repetición de secuencia PA#1 de 200 residuos, se amplificó la proteína SUMO de *Saccharomyces cerevisiae* [también conocida como Smt3p; Uniprot: Q12306] mediante reacción en cadena de la polimerasa (PCR) a partir de un ADNc clonado. El cebador en 5' introdujo un sitio de restricción *NdeI*, que contenía un codón de iniciación de Met (ATG) y un codón de Lys adicional, así como la secuencia que codificaba para cola de His6, mientras que el cebador en 3' introdujo un sitio de restricción *HindIII* y *SapI* en el producto de PCR. Se digirió el fragmento de ADN resultante con *NdeI* y *HindIII* y se ligó con un derivado digerido de manera correspondiente del plásmido pSA1 (Schmidt (1994) *J. Chromatogr.* 676: 337-345), en el que el sitio de restricción *SapI* se había eliminado mediante mutación silenciosa. Se cortó el plásmido resultante con *SapI*, se desfosforiló con fosfatasa alcalina de camarón y se ligó con el fragmento génico que codificaba para el segmento de polipéptido PA#1 de 200 residuos escindido del plásmido pFab-PA#1(200) (descrito en el ejemplo 2) mediante digestión de restricción con *SapI* (de una manera análoga tal como se muestra a modo de ejemplo en la figura 2E). El plásmido resultante se denominó pSUMO-PA#1(200) (SEQ ID NO: 60) y se representa en la figura 12A.

Ejemplo 21: Expresión bacteriana y aislamiento de un polímero/polipéptido PA#1(200) genéticamente codificado.

Inicialmente se produjo el polipéptido PA#1(200) (masa calculada: 16,1 kDa) como proteína de fusión con la proteína modificador similar a ubiquitina pequeño (SUMO) (masa calculada: 12,2 kDa) en el citoplasma de *E. coli* BLR(DE3) (NEB, Ipswich, MA, EE.UU.) que albergaba el plásmido de expresión pSUMO-PA#1(200) (descrito en el ejemplo 21) junto con el plásmido pLysE (Studier (1991) *J. Mol. Biol.* 219: 37-44), que suprime el promotor de T7. Se realizó la producción bacteriana a 30°C en cultivos en matraz de agitador con 2 l de medio LB que contenía D-glucosa 2,5 g/l, L-prolina 0,5 g/l, ampicilina 100 mg/l y cloranfenicol 30 mg/l. Se indujo expresión génica recombinante mediante adición de isopropil- β -D-tiogalactopiranosido (IPTG) hasta una concentración final de 0,5 mM. Se recogieron bacterias 3 h tras la inducción, se resuspendieron en NaCl 100 mM, fosfato de Na 40 mM pH 7,5 y se sometieron a lisis usando una célula de presión francesa (Thermo Scientific, Waltham, MA, EE.UU.). Tras la centrifugación (15 min, 15000 g) del lisado no se observaron cuerpos de inclusión.

Se incubó el sobrenadante que contenía la proteína de fusión soluble a 70°C durante 15 min y se centrifugó (15 min, 15000 g) para eliminar proteínas de células huésped térmicamente inestables. Se purificó la proteína de fusión His(6)-SUMO-PA#1(200) a partir del sobrenadante mediante IMAC (Skerra (1994) *Gene* 141: 79-84) usando una columna de alto rendimiento HisTrap cargada con Ni_2^+ de 12 ml (GE Healthcare) conectada a un sistema purificador Äkta (GE Healthcare) y se eluyó con un gradiente de imidazol de desde 0 hasta 150 mM en NaCl 500 mM, fosfato de Na 40 mM pH 7,5. Tras una etapa de SEC preparativa posterior se obtuvo una preparación homogénea de la proteína de fusión His(6)-SUMO-PA#1(200) (figura 12B) con un rendimiento de aproximadamente 5 mg por 1 l de cultivo bacteriano con $\text{DO}_{550} = 1$. Se determinó la concentración de proteína según la absorción a 280 nm usando un coeficiente de extinción calculado (Gill (1989) citado anteriormente) de $1280 \text{ M}^{-1} \text{ cm}^{-1}$ para la fusión de polipéptido His(6)-SUMO-PA#1(200). Obsérvese que el segmento de polipéptido PA#1(200) no contribuye a la absorción a 280 nm debido a su falta de cadenas laterales de aminoácidos aromáticos o que contienen azufre.

El polipéptido PA#1(200) biosintético se liberó de la proteína de fusión mediante escisión proteolítica específica del sitio (en el sentido de 3' desde un motivo Gly-Gly que precede al segmento de polipéptido de Pro/Ala) con proteasa 1 específica de Ubl 2 U/mg de *Saccharomyces cerevisiae* (Invitrogen, Carlsbad, CA, EE.UU.) durante 1 h a 30°C en tampón de escisión (Igepal al 0,2% p/v, DTT 1 mM, NaCl 150 mM, Tris-HCl 50 mM pH 8,0). Se comprobó el proceso de escisión mediante SDS-PAGE (figura 12B) usando un sistema de tampón Tris de alta molaridad (Fling (1986) *Anal. Biochem.* 155: 83-88). Con el fin de retirar la proteína His(6)-SUMO escindida, proteína de fusión sin escindir residual y también la SUMO proteasa, todas las cuales portaban la cola de His6, se sometió la mezcla de reacción a otra IMAC usando una columna de alto rendimiento HisTrap cargada con Ni_2^+ de 5 ml (GE Healthcare) y NaCl 500 M, fosfato 20 mM, pH 7,5 como tampón de corrida. Esta vez el flujo a través contenía el polipéptido PA#1(200) biosintético puro (figura 13 E). Obsérvese que el polipéptido/polímero PA#1(200) biosintético (SEQ ID NO: 61) preparado de esta manera comprende en total 201 residuos de aminoácido, que surgen del producto génico combinado codificado de 10 bloques estructurales de oligodesoxinucleótidos bicatenarios ligados, codificando cada uno para 20 residuos de aminoácido, tal como se muestra en la figura 1, y un residuo de Ala adicional codificado por la proyección de ADN triplete del sitio de restricción *SapI* en el sentido de 3' que se usó para la clonación.

Ejemplo 22: Preparación y caracterización de conjugados de molécula pequeña/fármaco con PA#1(200).

Se sometió dos veces la mezcla de reacción de escisión proteolítica sin purificar de la proteína de fusión His(6)-SUMO-PA#1(200) del ejemplo 21 a diálisis a 4°C frente a NaHCO₃ 50 mM pH 8,3 y se incubó a temperatura ambiente durante 1 h tras mezclar con un exceso molar de 10 veces de una disolución de éster de N-hidroxisuccinimida del ácido 6-[fluoresceína-5(6)-carboxamido]hexanoico (éster de fluoresceína-NHS; Sigma-Aldrich) en dimetilformamida seca (DMF). Para ello, se añadieron 200 µl de una disolución 2,5 mg/ml de la mezcla de escisión de His(6)-SUMO-PA#1(200) a 17,6 µl de una disolución 10 mM de éster de fluoresceína-NHS disuelto en DMF. Se incubó la mezcla resultante a temperatura ambiente durante 1 h y se aplicó a IMAC tal como se describió en el ejemplo 21 para retirar la proteína His(6)-SUMO escindida, proteína de fusión sin escindir residual y la SUMO proteasa y se purificó adicionalmente mediante SEC preparativa en una columna Superdex S200 10/300 GL equilibrada con PBS a una velocidad de flujo de 0,5 ml/min.

Después se analizaron muestras de las diferentes etapas mediante SEC analítica en una columna Superdex S200 10/300 GL equilibrada con PBS a una velocidad de flujo de 0,5 ml/min. Se detectó la proteína SUMO mediante sus cadenas laterales aromáticas a 280 nm y los enlaces peptídicos, incluyendo los del polipéptido de Pro/Ala o segmento de polipéptido, se detectaron a 225 nm, mientras que la fluoresceína se detectó a 494 nm (figura 13 A-G). Para comparación, se midieron espectros UV/VIS de una disolución de fluoresceína libre (Sigma-Aldrich) y de fracciones de cada pico diferenciado detectado en la SEC usando un instrumento Lambda 9 (Perkin Elmer, Waltham, MA, EE.UU.) (figura 13 H-K). Para la calibración del tamaño de la columna de cromatografía (figura 13 L), se aplicaron 250 µl de una mezcla apropiada de las siguientes proteínas globulares (Sigma-Aldrich) en PBS a concentraciones de entre 0,2 y 0,5 mg/ml: aprotinina, 6,5 kDa; citocromo c, 12,4 kDa; anhidrasa carbónica, 29,0 kDa; albúmina sérica bovina, 66,3 kDa; alcohol deshidrogenasa, 150 kDa; β-amilasa, 200 kDa; apo-ferritina, 440 kDa.

Como resultado, tras el acoplamiento del polipéptido/polímero PA#1(200) biosintético con éster de fluoresceína-NHS se aisló un conjugado macromolecular mediante IMAC y SEC que muestra esencialmente las propiedades de tamaño del polipéptido/polímero PA#1(200) y la firma espectroscópica de la molécula pequeña, es decir el grupo de fluoresceína. Esto demuestra que la molécula pequeña se acopló satisfactoriamente al polipéptido/polímero de Pro/Ala biosintético, lo cual, según esta invención, aumenta drásticamente el volumen hidrodinámico del compuesto o fármaco de molécula pequeña conjugado.

Para preparar un conjugado similar entre el polipéptido/polímero de Pro/Ala biosintético y el esteroide vegetal digoxigenina, se sometieron 0,1 mg del polipéptido PA#1(200) purificado del ejemplo 21 a diálisis frente a NaHCO₃ 50 mM pH 8,3 tal como se describió anteriormente. Se determinó la concentración de polipéptido PA#1(200) purificado según la absorción a 205 nm (Gill (1989) citado anteriormente). Se acopló el polipéptido PA#1(200) con un exceso molar de 10 veces de éster de NHS del ácido digoxigenina-3-O-metilcarbonil-ε-aminocaproico (éster de DIG-NHS; Roche Diagnostics, Mannheim, Alemania). Para ello, se añadieron 100 µl de una disolución 1 mg/ml del polipéptido PA#1(200) purificado en NaHCO₃ 50 mM pH 8,3 a 2 µl de una disolución 30 mM de éster de DIG-NHS disuelto en DMF seca y se incubó la mezcla de reacción durante 1 h a temperatura ambiente. Se purificó la disolución resultante del conjugado usando una columna de desalación de centrifugación Zeba™ con un punto de corte de 7 kDa (Thermo Scientific), se sometió dos veces a diálisis frente a tampón de acetato de amonio 10 mM pH 6,8 y se analizó mediante espectrometría de masas con ESI en un instrumento Q-Tof Ultima (Waters, Eschbronn, Alemania) usando el modo de ionización positiva. Como resultado, el espectro del conjugado de digoxigenina-PA#1(200) reveló una masa de 16671,4 Da, lo cual coincide esencialmente con la masa calculada de 16670,6 Da (figura 13M). Esto demuestra claramente que puede conjugarse eficazmente un polipéptido/polímero de Pro/Ala biosintético, en particular PA#1(200), con un fármaco de molécula pequeña.

La presente invención está relacionada con y se refiere a las siguientes secuencias mostradas a modo de ejemplo, mediante lo cual la lista de secuencias adjunta se presenta como parte de la descripción y, por consiguiente, es una parte de esta memoria descriptiva.

SEQ ID NO: 1 muestra la secuencia de aminoácidos de PA#1.

SEQ ID NO: 2 muestra la secuencia de aminoácidos de PA#2.

SEQ ID NO: 3 muestra la secuencia de aminoácidos de PA#3.

SEQ ID NO: 4 muestra la secuencia de aminoácidos de PA#4.

SEQ ID NO: 5 muestra la secuencia de aminoácidos de PA#5.

SEQ ID NO: 6 muestra la secuencia de aminoácidos de PA#6.

SEQ ID NO: 7 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.

SEQ ID NO: 8 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.

SEQ ID NO: 9 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.

SEQ ID NO: 10 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.

ES 2 691 642 T3

- SEQ ID NO: 11 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.
- SEQ ID NO: 12 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.
- SEQ ID NO: 13 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.
- SEQ ID NO: 14 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.
- 5 SEQ ID NO: 15 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.
- SEQ ID NO: 16 muestra una secuencia de aminoácidos de una versión permutada circular de SEQ ID NO: 1.
- SEQ ID NO: 17 muestra una secuencia de ácido nucleico del oligodesoxinucleótido de cadena superior/codificante usado para la generación del bloque estructural PA#1.
- 10 SEQ ID NO: 18 muestra una secuencia de ácido nucleico del oligodesoxinucleótido de cadena inferior/no codificante usado para la generación del bloque estructural para PA#1.
- SEQ ID NO: 19 muestra un tramo de secuencia de ácido nucleico (cadena superior/codificante) alrededor del extremo C-terminal de la cadena ligera de inmunoglobulina de un fragmento Fab de anticuerpo tal como se codifica en pASK88-Fab-2xSapl.
- 15 SEQ ID NO: 20 muestra un tramo de secuencia de ácido nucleico (cadena inferior/no codificante) alrededor del extremo C-terminal de la cadena ligera de inmunoglobulina de un fragmento Fab de anticuerpo tal como se codifica en pASK88-Fab-2xSapl.
- SEQ ID NO: 21 muestra una secuencia de aminoácidos del extremo C-terminal de la cadena ligera del fragmento Fab tal como se codifica en pASK88-Fab-2xSapl.
- SEQ ID NO: 22 muestra la secuencia de ácido nucleico de pASK88-Fab-2xSapl.
- 20 SEQ ID NO: 23 muestra un tramo de secuencia de ácido nucleico (cadena superior/codificante) que codifica para la secuencia de aminoácidos del extremo C-terminal de la cadena ligera de Fab tras la inserción de un polímero PA#1(20).
- SEQ ID NO: 24 muestra una secuencia de ácido nucleico (cadena inferior/no codificante) para un tramo de aminoácidos del extremo C-terminal de una cadena ligera de Fab tras la inserción de un polímero PA#1(20).
- 25 SEQ ID NO: 25 muestra un tramo de secuencia de aminoácidos del extremo C-terminal de una cadena ligera de Fab tras la inserción de un polímero PA#1(20).
- SEQ ID NO: 26 muestra la secuencia de aminoácidos de la cadena pesada de Fab tal como se codifica en pFab-PA#1(200).
- 30 SEQ ID NO: 27 muestra la secuencia de aminoácidos de la cadena ligera de Fab fusionada con el polímero PA#1(200) tal como se codifica en pFab-PA#1(200).
- SEQ ID NO: 28 muestra la secuencia de ácido nucleico de pFab-PA#1(200).
- SEQ ID NO: 29 muestra la secuencia de ácido nucleico (cadena superior/codificante) que codifica para la secuencia de aminoácidos del extremo N-terminal de INFa2b y Strep-tag II (sólo los dos últimos aminoácidos).
- 35 SEQ ID NO: 30 muestra una secuencia de ácido nucleico (cadena inferior/no codificante) que codifica para secuencia de aminoácidos del extremo N-terminal de INFa2b y Strep-tag II (sólo los dos últimos aminoácidos).
- SEQ ID NO: 31 muestra la secuencia de aminoácidos del extremo C-terminal de Strep-tag II y el extremo N-terminal de INFa2b.
- SEQ ID NO: 32 muestra la secuencia de ácido nucleico de pASK-IFNa2b.
- 40 SEQ ID NO: 33 muestra un tramo de secuencia de ácido nucleico (cadena superior/codificante) que codifica para el extremo C-terminal de Strep-tag II y el extremo N-terminal de IFNa2b tras la inserción de un casete de secuencia de polímero PA#1.
- SEQ ID NO: 34 muestra un tramo de secuencia de ácido nucleico (cadena inferior/no codificante) del extremo C-terminal de Strep-tag II y el extremo N-terminal de IFNa2b tras la inserción de un casete de secuencia de polímero PA#1.
- 45 SEQ ID NO: 35 muestra un tramo de secuencia de aminoácidos del extremo C-terminal de Strep-tag II y el extremo N-terminal de IFNa2b tras la fusión con un casete de polímero PA#1.

ES 2 691 642 T3

- SEQ ID NO: 36 muestra la secuencia de aminoácidos de IFNa2b y Strep-tag II fusionada con el polímero PA#1(200) tal como se codifica en pPA#1(200)-IFNa2b.
- SEQ ID NO: 37 muestra la secuencia de ácido nucleico de pPA#1(200)-IFNa2b.
- 5 SEQ ID NO: 38 muestra un tramo de secuencia de ácido nucleico (cadena superior/codificante) en pASK75-His6-hGH que codifica para la secuencia de aminoácidos alrededor del extremo N-terminal de His6-hGH.
- SEQ ID NO: 39 muestra un tramo de secuencia de ácido nucleico (cadena inferior/no codificante) en pASK75-His6-hGH que codifica para la secuencia de aminoácidos alrededor del extremo N-terminal de hGH.
- SEQ ID NO: 40 muestra un tramo de secuencia de aminoácidos del extremo N-terminal de His6-hGH tal como se codifica en pASK75-His6-hGH.
- 10 SEQ ID NO: 41 muestra la secuencia de ácido nucleico de pASK75-His6-hGH.
- SEQ ID NO: 42 muestra un tramo de secuencia de ácido nucleico (cadena superior/codificante) que codifica para la secuencia de aminoácidos del extremo N-terminal de His6-hGH tras la inserción del polímero PA#1(20).
- SEQ ID NO: 43 muestra una secuencia de ácido nucleico (cadena inferior/no codificante) que codifica para el extremo N-terminal de hGH tras la inserción de un casete de secuencia de polímero PA#1.
- 15 SEQ ID NO: 44 muestra la secuencia de aminoácidos del extremo N-terminal de His6-hGH tras la inserción del polímero PA#1(20).
- SEQ ID NO: 45 muestra la secuencia de aminoácidos de His6-PA#1(200)-hGH madura tal como se codifica en pASK75-His6-PA#1(200)-hGH.
- SEQ ID NO: 46 muestra la secuencia de ácido nucleico de pASK75-His6-PA#1(200)-hGH.
- 20 SEQ ID NO: 47 muestra la secuencia de aminoácidos de His6-PA#1(200)-hGH tal como se codifica en pCHO-PA#1(200)-hGH.
- SEQ ID NO: 48 muestra la secuencia de ácido nucleico de pCHO-PA#1(200)-hGH.
- SEQ ID NO: 49 muestra la secuencia de ácido nucleico de pCHO-hGH.
- SEQ ID NO: 50 muestra la secuencia de ácido nucleico de pCHO.
- 25 SEQ ID NO: 51 muestra la secuencia de aminoácidos de P1A1.
- SEQ ID NO: 52 muestra la secuencia de ácido nucleico del oligodesoxinucleótido de cadena superior/codificante usado para la generación del bloque estructural para P1A1.
- SEQ ID NO: 53 muestra la secuencia de ácido nucleico del oligodesoxinucleótido de cadena inferior/no codificante usado para la generación del bloque estructural para P1A1.
- 30 SEQ ID NO: 54 muestra la secuencia de ácido nucleico del oligodesoxinucleótido de cadena superior/codificante usado para la generación del bloque estructural para P1A3.
- SEQ ID NO: 55 muestra la secuencia de ácido nucleico del oligodesoxinucleótido de cadena inferior/no codificante usado para la generación del bloque estructural para P1A3.
- 35 SEQ ID NO: 56 muestra la secuencia de aminoácidos de la cadena ligera de Fab fusionada con el polímero P1A1(200) tal como se codifica en pFab-P1A1(200).
- SEQ ID NO: 57 muestra la secuencia de aminoácidos de la cadena ligera de Fab fusionada con el polímero P1A3(200) tal como se codifica en pFab-P1A3(200).
- SEQ ID NO: 58 muestra la secuencia de ácido nucleico de pFab-P1A1(200).
- SEQ ID NO: 59 muestra la secuencia de ácido de pFab-P1A3(200).
- 40 SEQ ID NO: 60 muestra la secuencia de ácido nucleico de pSUMO-PA#1(200).
- SEQ ID NO: 61 muestra el polipéptido/polímero PA#1(200) usado para la preparación de conjugados de fármaco (realizados mediante ligación de 10 casetes génicos codificantes de 20 meros, incluyendo un residuo de Ala C-terminal adicional resultante del sitio de ligación en el sentido de 3').

REIVINDICACIONES

1. Conjugado de fármaco que comprende
 - (i) un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala), y
 - (ii) un fármaco seleccionado del grupo que consiste en (a) una proteína biológicamente activa o un polipéptido que comprende o que es una secuencia de aminoácidos que tiene o media en una actividad biológica y (b) un fármaco de molécula pequeña.
2. Conjugado de fármaco según la reivindicación 1, en el que dicho polipéptido de enrollamiento al azar o segmento de polipéptido comprende una secuencia de aminoácidos que consiste en de aproximadamente 50 a aproximadamente 3000 residuos de aminoácido.
3. Conjugado de fármaco según la reivindicación 1 ó 2, en el que dichos residuos de prolina constituyen más de aproximadamente el 10% y menos de aproximadamente el 75% de la secuencia de aminoácidos.
4. Conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 3, en el que dicho polipéptido de enrollamiento al azar o segmento de polipéptido comprende una pluralidad de repeticiones de aminoácidos, en el que dicha repetición consiste en residuos de prolina y alanina y en el que no más de 6 residuos de aminoácido consecutivos son idénticos.
5. Conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 4, en el que dicho polipéptido de enrollamiento al azar o segmento de polipéptido comprende una secuencia de aminoácidos seleccionada del grupo que consiste en

AAPAAPAPAAPAAPAPAAPA (SEQ ID NO: 1);

AAPAAAPAPAAPAAPAPAAP (SEQ ID NO: 2);

AAAPAAAAPAAAAPAAAAP (SEQ ID NO: 3);

AAPAAPAAPAAPAAPAAPAAP (SEQ ID NO: 4);

APAAAPAPAAAAPAPAAAAP (SEQ ID NO: 5);

AAAPAAPAAPPAAAAPAAPPA (SEQ ID NO: 6); y

APAPAPAPAPAPAPAPAP (SEQ ID NO: 51)

 o versiones permutadas circulares o (a) múltiplo(s) de estas secuencias en su conjunto o partes de estas secuencias.
6. Conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 5, en el que dicho polipéptido con actividad biológica, dicha proteína biológicamente activa o dicho polipéptido que comprende o que es una secuencia de aminoácidos que tiene o que media en una actividad biológica se selecciona del grupo que consiste en proteínas de unión, fragmentos de anticuerpo, citocinas, factores de crecimiento, hormonas o enzimas.
7. Conjugado de fármaco según la reivindicación 6, en el que dicho polipéptido con actividad biológica es una proteína de unión y en el que dicha molécula de unión se selecciona del grupo que consiste en anticuerpos, fragmentos Fab, fragmentos F(ab')₂, compuestos peptidomiméticos derivados de CDR, fragmentos variables de cadena sencilla (scFv), anticuerpos de dominio, lectinas, dominios de inmunoglobulina, dominios de fibronectina, dominios de proteína A, dominios SH3, dominios de repetición de anquirina y lipocalinas.
8. Conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 7, en el que dicha proteína biológicamente activa se selecciona del grupo que consiste en factor estimulante de colonias de granulocitos, hormona del crecimiento humana, alfa-interferón, beta-interferón, gamma-interferón, factor de necrosis tumoral, eritropoyetina, factor de coagulación VIII, gp120/gp160, receptor de factor de necrosis tumoral I y II soluble, reteplasa, exendina-4, anakinra, interleucina-2, lipocalina asociada a gelatinasa de neutrófilos, hormona estimulante del folículo, glucocerebrosidasa, timosina alfa 1, glucagón, somatostatina, adenosina desaminasa, interleucina 11, factor de coagulación VIIA, factor de coagulación IX, Hematide, lambda-interferón, leptina, subunidad alfa de receptor de interleucina-22 (IL-22ra), interleucina-22, hialuronidasa, factor de crecimiento de fibroblastos 18, factor de crecimiento de fibroblastos 21, péptido similar a glucagón 1, osteoprotegerina, proteína de unión a IL-18, factor de liberación de la hormona del

crecimiento, receptor de TACI soluble, trombospondina-1, receptor de VEGF soluble Flt-1 y mteína de IL-4.

9. Conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 8, mediante el cual dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido media en una estabilidad *in vivo* y/o *in vitro* aumentada de dicho conjugado de fármaco.
- 5 10. Conjugado de fármaco según la reivindicación 9, en el que dicha estabilidad *in vivo* aumentada es una semivida en plasma prolongada de dicho conjugado de fármaco que comprende dicho polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala), en comparación con la estabilidad de un polipéptido de control o un conjugado de control que carece de dicho polipéptido de enrollamiento al azar o segmento de polipéptido.
- 10 11. Composición que comprende el conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 10.
12. Composición según la reivindicación 11, que es una composición farmacéutica.
13. Molécula de ácido nucleico que codifica para un conjugado de proteína que comprende una proteína biológicamente activa según una cualquiera de las reivindicaciones 6 a 8 y que comprende un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala).
- 15 14. Molécula de ácido nucleico que codifica para un conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 10, comprendiendo dicha molécula de ácido nucleico
 - (i) una secuencia de ácido nucleico que codifica para una secuencia de aminoácidos y/o líder traducida;
 - (ii) una secuencia de ácido nucleico que codifica para un polipéptido de enrollamiento al azar biosintético o segmento de polipéptido que comprende una secuencia de aminoácidos que consiste únicamente en residuos de aminoácido de prolina y alanina, en el que dicha secuencia de aminoácidos consiste en al menos 50 residuos de aminoácido de prolina (Pro) y alanina (Ala);
 - (iii) una secuencia de ácido nucleico que codifica para una proteína biológicamente activa o un polipéptido que comprende o que es una secuencia de aminoácidos que tiene o que media en una actividad biológica y/o terapéutica; y
 - (iv) una secuencia de ácido nucleico que representa o es un codón de terminación de la traducción.
- 20 15. Molécula de ácido nucleico según la reivindicación 14, en el que las partes o segmentos de molécula de ácido nucleico según se definen en (ii) y en (iii) intercambian su posición en dicha molécula de ácido nucleico que codifica para un conjugado de fármaco.
- 25 16. Molécula de ácido nucleico según la reivindicación 14 ó 15, que comprende opcionalmente, entre partes o segmentos según se definen en (i) y en (ii) y/o entre partes o segmentos según se definen en (ii) y (iii), un sitio de escisión por proteasa y/o química y/o un sitio de reconocimiento.
- 30 17. Vector que comprende el ácido nucleico según una cualquiera de las reivindicaciones 13 a 16.
- 35 18. Célula huésped que comprende el ácido nucleico según una cualquiera de las reivindicaciones 13 a 16 o el vector según la reivindicación 17.
- 40 19. Método para la preparación de un polipéptido de enrollamiento al azar o un segmento de polipéptido de enrollamiento al azar según está comprendido en el conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 10, para la preparación de la proteína biológicamente activa o un conjugado de fármaco que comprende dicho polipéptido de enrollamiento al azar o dicho segmento de polipéptido de enrollamiento al azar y/o para la preparación de un polipéptido que comprende o que es una secuencia de aminoácidos que tiene o que media en una actividad biológica según una cualquiera de las reivindicaciones 1 a 8 y que comprende adicionalmente dicho polipéptido de enrollamiento al azar o segmento de polipéptido de enrollamiento al azar,

comprendiendo dicho método cultivar la célula según la reivindicación 18 y aislar dicho polipéptido de enrollamiento al azar o proteína biológicamente activa y/o dicha proteína biológicamente activa o dicho polipéptido a partir del cultivo o a partir de dicha célula.
- 45 20. Conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 10, composición según la reivindicación 11 ó 12, ácido nucleico según una cualquiera de las reivindicaciones 13 a 16, vector según la reivindicación 17, célula según la reivindicación 18, o proteína biológicamente activa o polipéptido que

- 5 comprende dicho polipéptido de enrollamiento al azar o segmento de polipéptido de enrollamiento al azar y según se prepara mediante el método según la reivindicación 19, para su uso en el tratamiento de trastornos relacionados con carencia hormonal, enfermedad autoinmunitaria, trastornos proliferativos, cáncer, anemia, enfermedades neovasculares, enfermedades infecciosas/inflamatorias, trastornos alérgicos, trombosis, infarto de miocardio, reestenosis, diabetes, esterilidad, enfermedad de Gaucher, hepatitis B crónica, hepatitis C, hipoglucemia, acromegalia, carencia de adenosina desaminasa, trombocitopenia, hemofilia, anemia, obesidad, enfermedad de Alzheimer, lipodistrofia, psoriasis, melanoma metastásico, osteoartritis, dislipidemia, artritis reumatoide, lupus eritematoso sistémico, esclerosis múltiple, asma, osteoporosis y lesión por reperfusión u otras enfermedades renales.
- 10 21. Conjugado de fármaco según una cualquiera de las reivindicaciones 1 a 10, composición según la reivindicación 11 ó 12, ácido nucleico según una cualquiera de las reivindicaciones 13 a 16, vector según la reivindicación 17 o célula según la reivindicación 18, para su uso como medicamento que tiene una estabilidad *in vivo* y/o *in vitro* aumentada de dicho polipéptido de enrollamiento al azar o segmento de polipéptido, proteína biológicamente activa o conjugado de fármaco.

15

Figura 1

gccGGTCCAGCTGCACCTGCTCCAGCAGCACCTGCTGCACCAGCTCCGGCTGCTCCTGCT
 |||||
 CGAGGTCGACCGTGGACGAGGTCGTCGTGGACGACGTTGGTCGAGGCCGACGAGGACGAcgg
 AlaAlaProAlaAlaProAlaProAlaAlaProAlaProAlaProAlaProAlaProAlaProAla

Figura 2

A

SapI SapI

| |

AAGAGCTTCAACCGGGAGAGTGCTCTTCTGCCTGAAGAGCTTAAGCTT
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTCTCGAAGTTGGCCCTCTCACGAGAGACGGACTTCTCGAATTCGAA
 LysSerPheAsnArgGlyGluCysSerSerAlaEnd

B

SapI

|

GAGTGCTCTTCTGCCGCTCCAGCTGCACCTGCTCCAGCAGCACCTGCTGCACCAGCTCCGGCTGCTCCTGCCTGAAGAGCTT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CTCACGAGAAGACGGCGAGGTCGACCGTGGACGAGGTCGTCGTGGACGACGTCGGTCGAGGCCGACGAGGACGACTTCTCGAA
 GluCysSerSerAlaAlaProAlaAlaProAlaProAlaProAlaProAlaProAlaProAlaProAlaProAlaProAlaAla

Figura 2 cont.

C

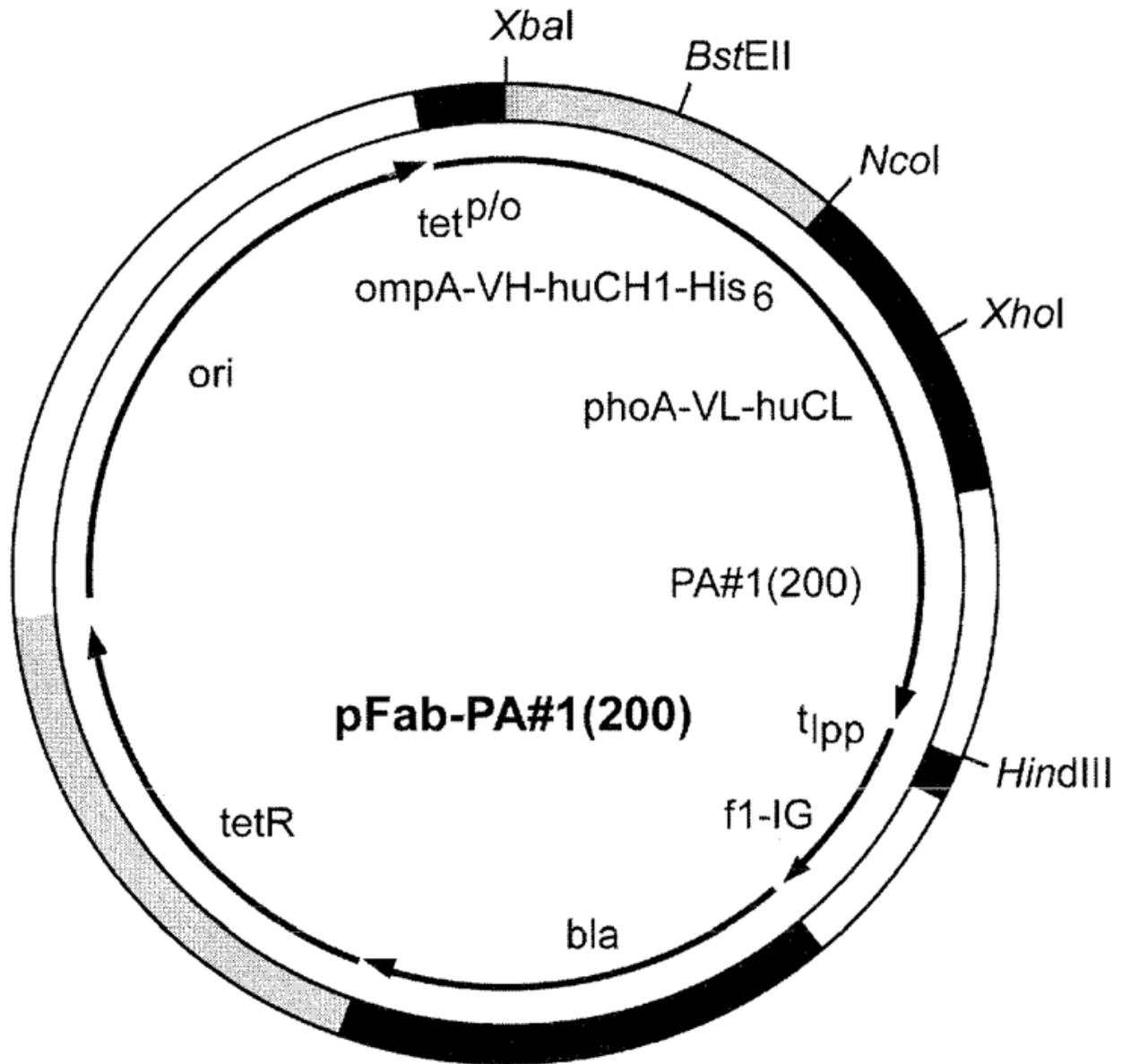


Figura 2 cont.
D

SapI
|

GAAAAAGGCGCCAGCTCTTCTGCCTGTGATCTGCCCTCAAACCCACAGCCCTGGGTAGC...
 -----+-----+-----+-----+-----+-----+-----+-----+
 CTTTTCCGGCGGTCGAGAAGACGGACACTAGACGGAGTTTGGGTGTCGGACCCATCG...
 GluLysGlyAlaSerSerAlaCysAspLeuProGlnThrHisSerLeuGlySer...
 +1

E

SapI
|

GAAAAAGGCGCCAGCTCTTCTGCCCGTCCAGCTGCACCTGCTCCAGCAGCACCTGCTGCACCAGCTCCGGGTGCTCCTGCTGCCTGTGAT...
 -----+-----+-----+-----+-----+-----+-----+-----+
 CTTTTCCGGCGGTCGAGAAGACGGCGAGGTCGACGTGGACGAGGTCGTGGACGACGTGGTCCGAGGCCGACGAGGACGACTA...
 GluLysGlyAlaSerSerAlaProAlaProAlaProAlaProAlaProAlaProAlaProAlaProAlaCysAsp...
 (1)

Figura 2 cont.
F

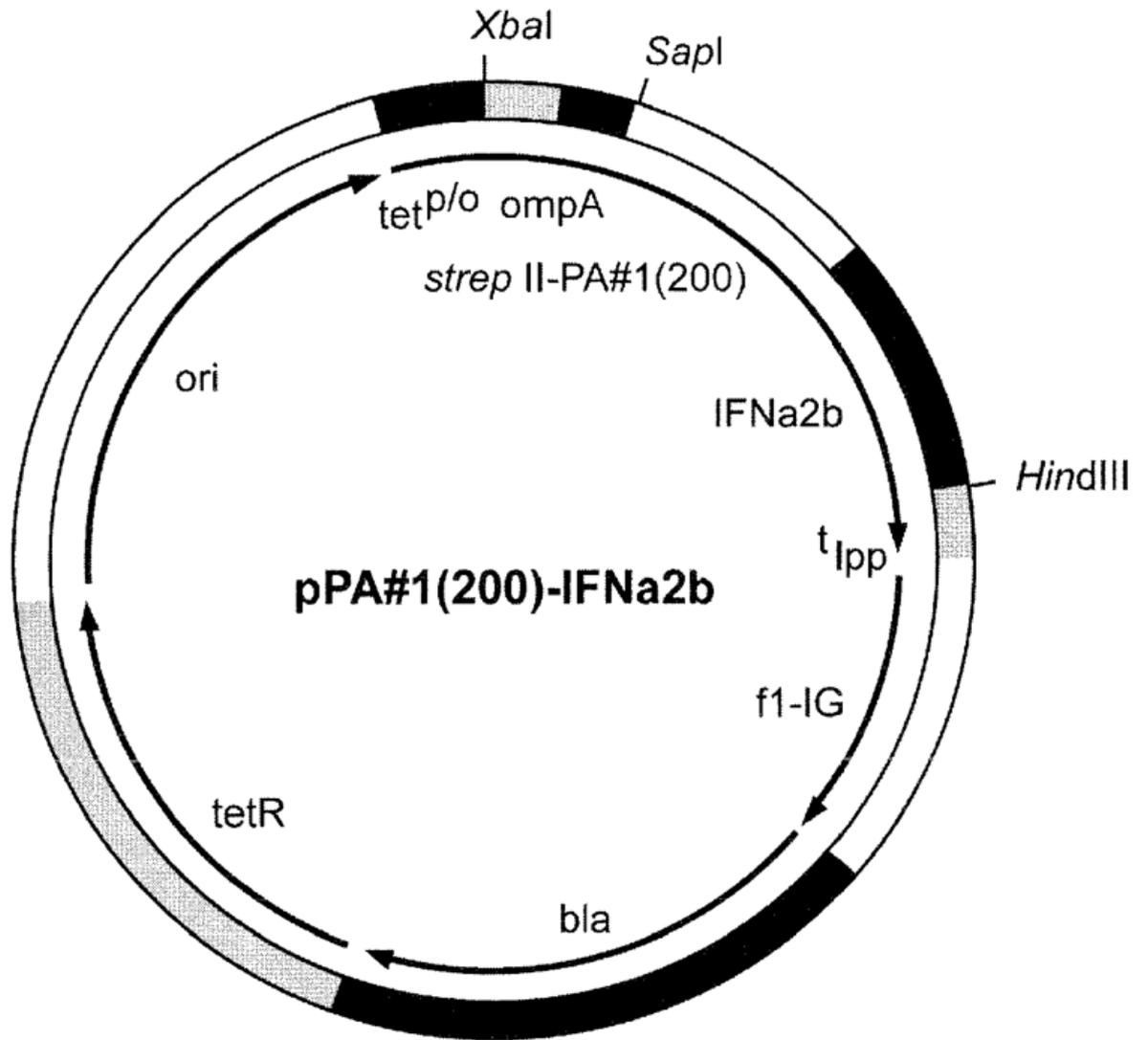
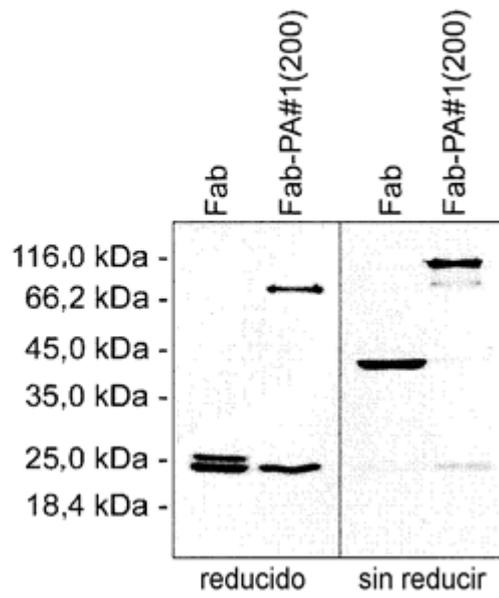


Figura 3

A



B

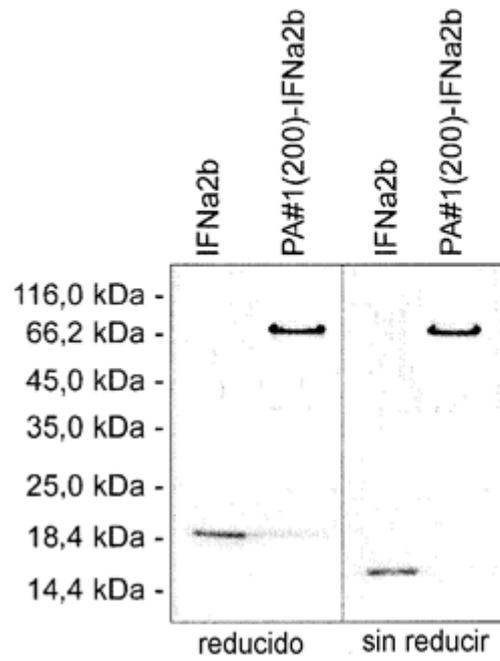
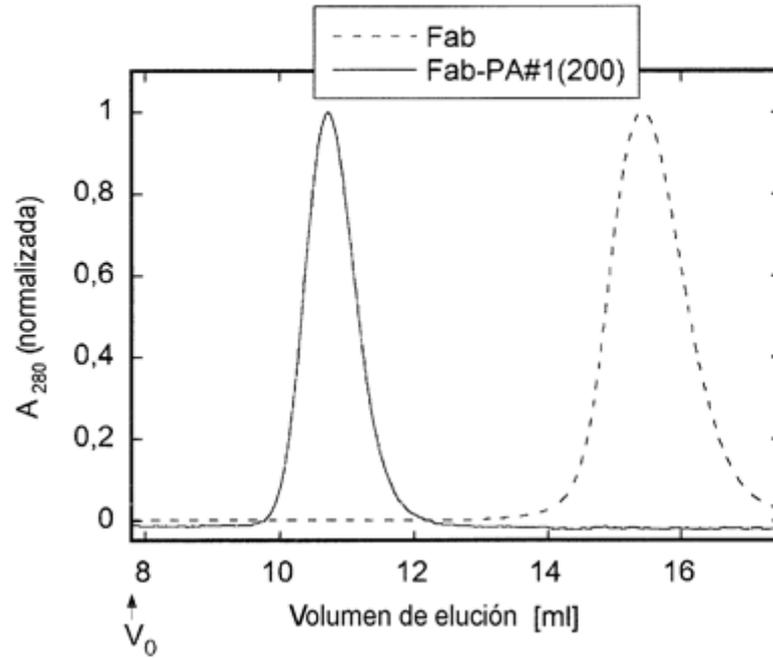


Figura 4

A



B

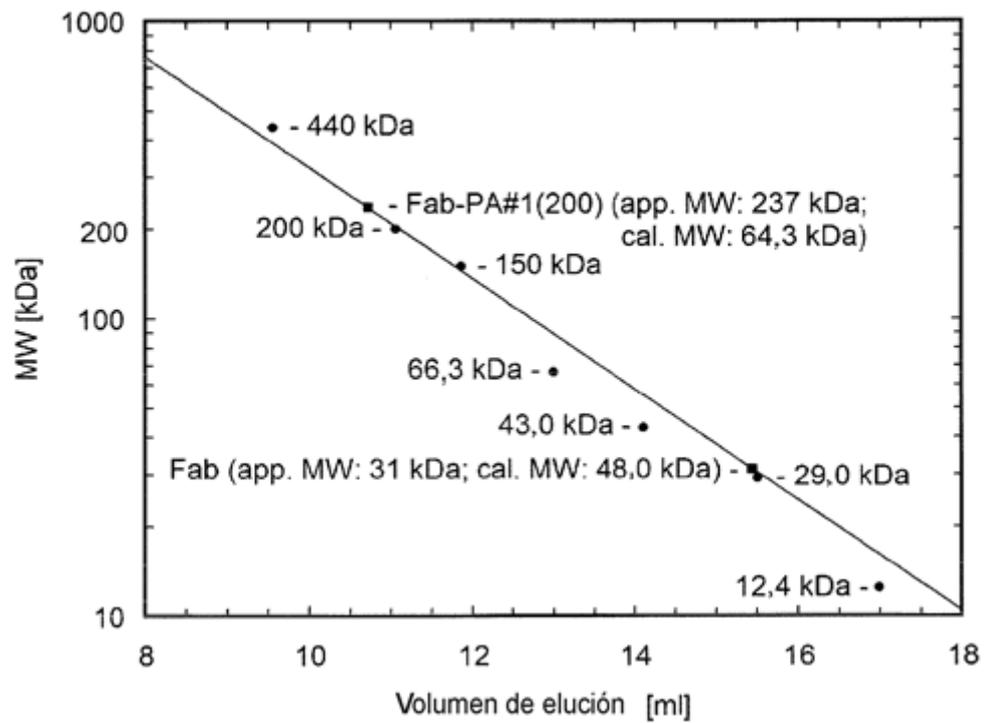
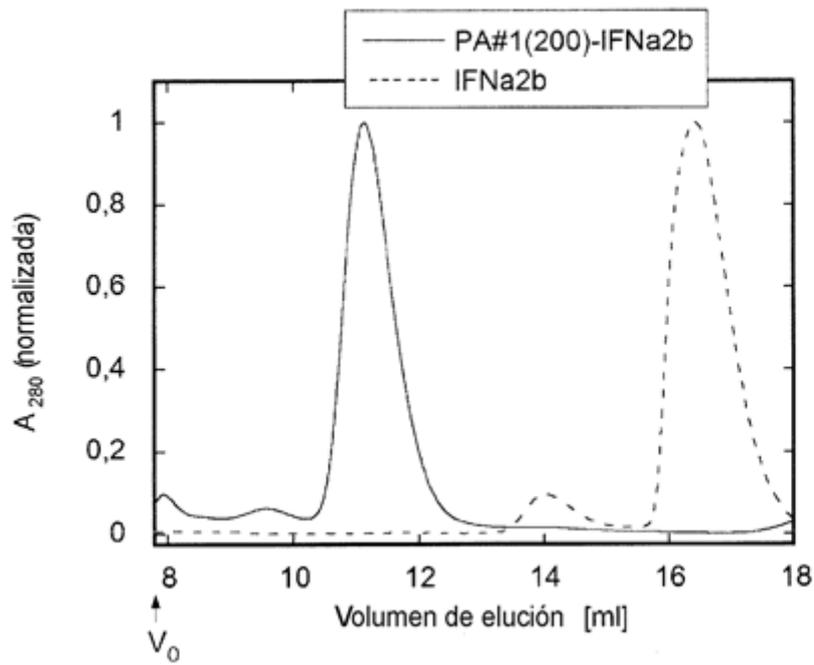


Figura 4 cont.

C



D

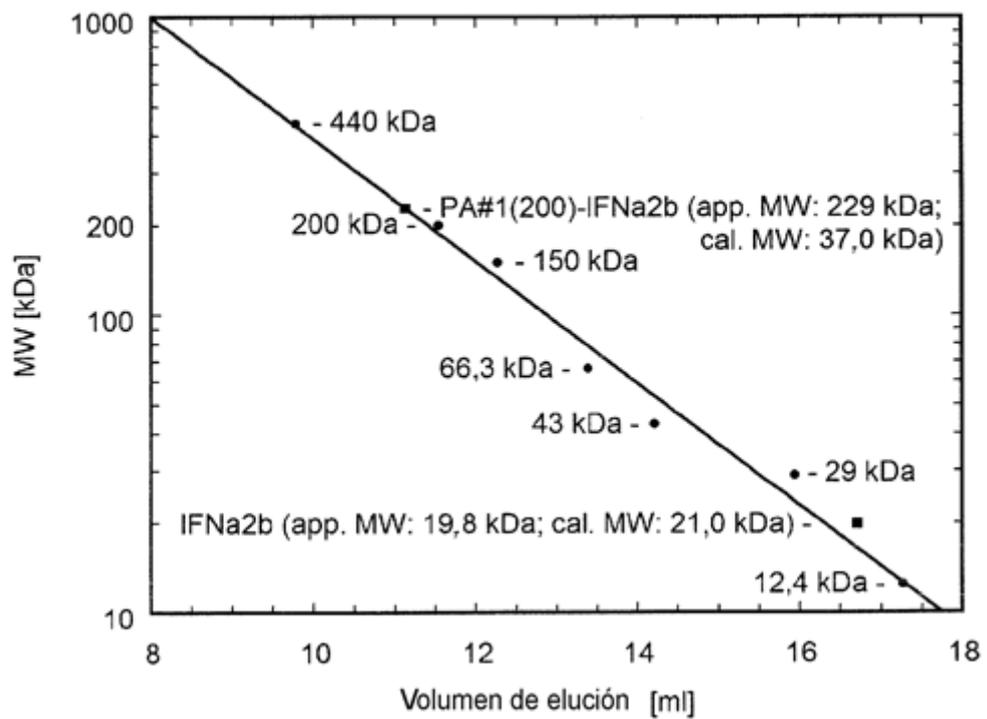
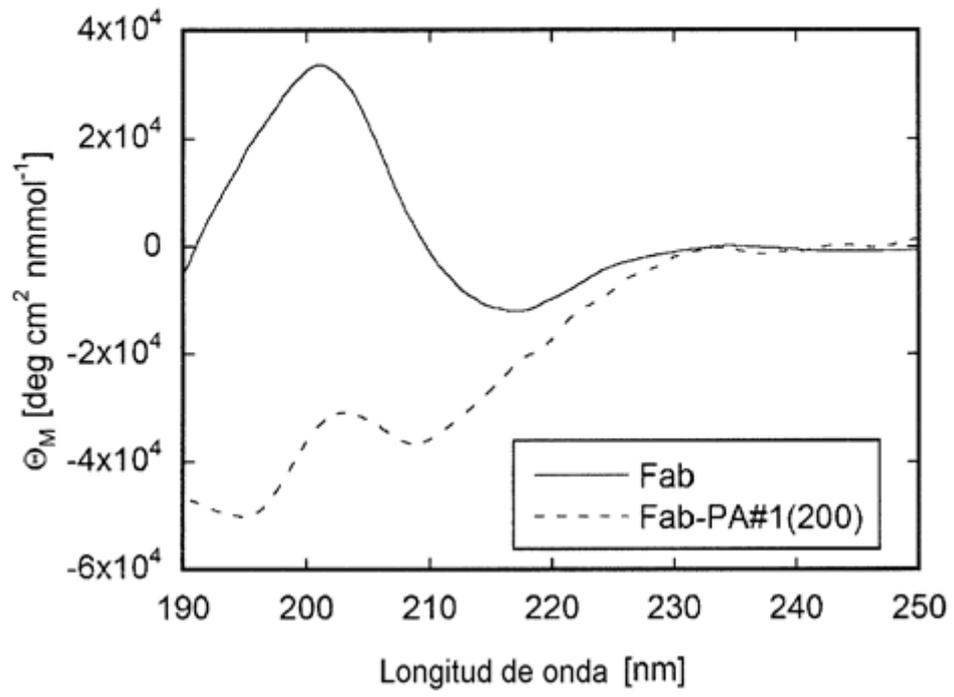


Figura 5

A



B

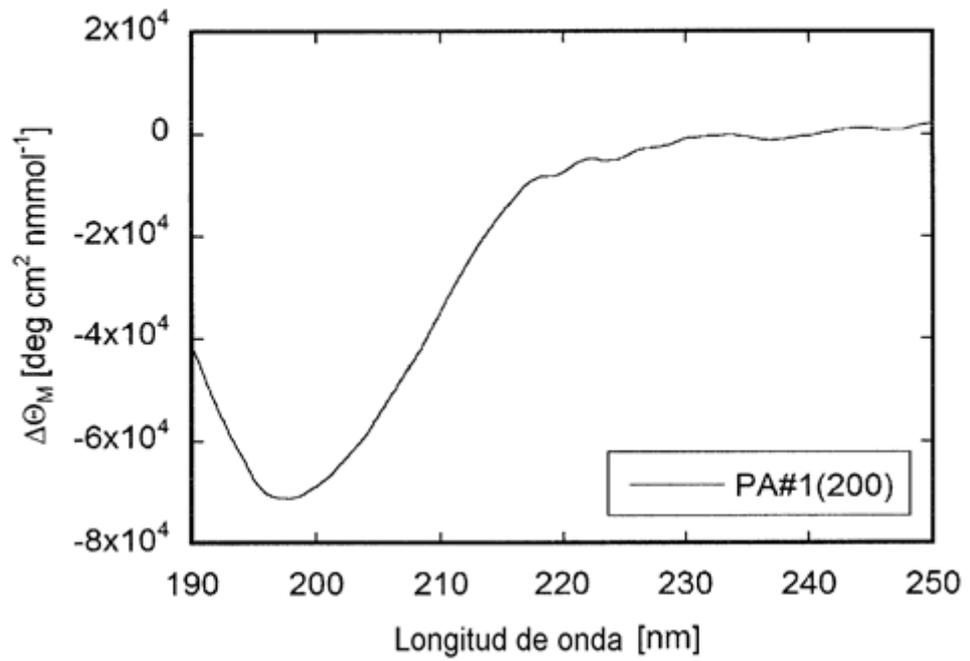
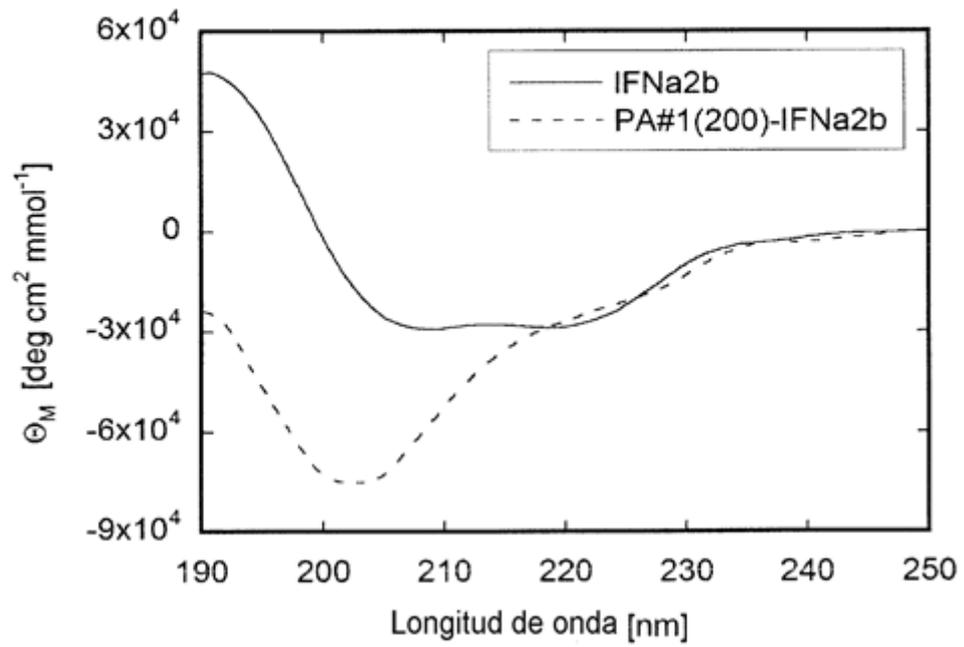


Figura 5 cont.

C



D

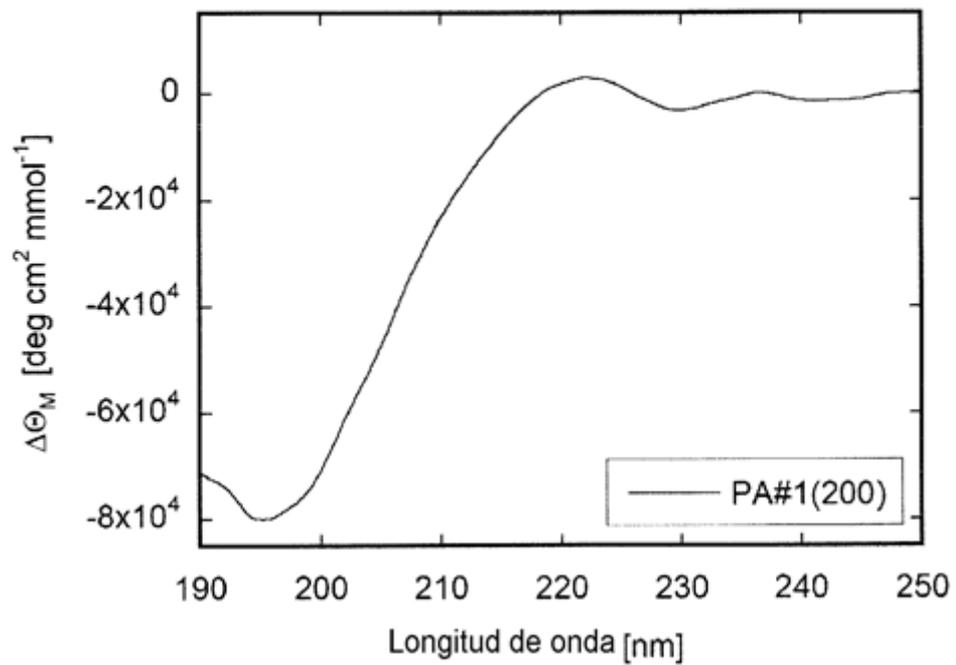


Figura 6

A

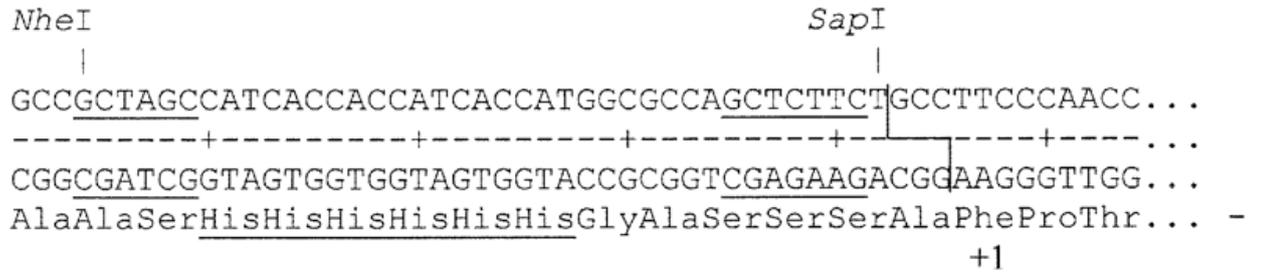


Figura 6 cont.

```

B
NheI      SapI
|          |
GCCGCTAGCCATCACCACCATCACCATGGCGCCAGCTCTTCTGCCCCTCCAGCTGCACCTGCTCCAGCAGCACCTGCTGCACCAGCTCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGCGATCGGTAGTGGTAGTGGTACCGCGGTCCGAGAACGGCCGAGGTCGACGTGGACGAGGTCCGTCGTGGACGACGTTGGTTCGAGGC
AlaAlaSerHisHisHisHisHisHisGlyAlaSerSerSerAlaAlaProAlaAlaProAlaAlaProAlaAlaProAlaAlaProAlaPro
GCTGCTCCTGCTGCCTTCCCAACC...
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGACGAGGACGACGGAGGGTTGG...
AlaAlaProAlaAlaPheProThr...
(1)

```

Figura 6 cont.

C

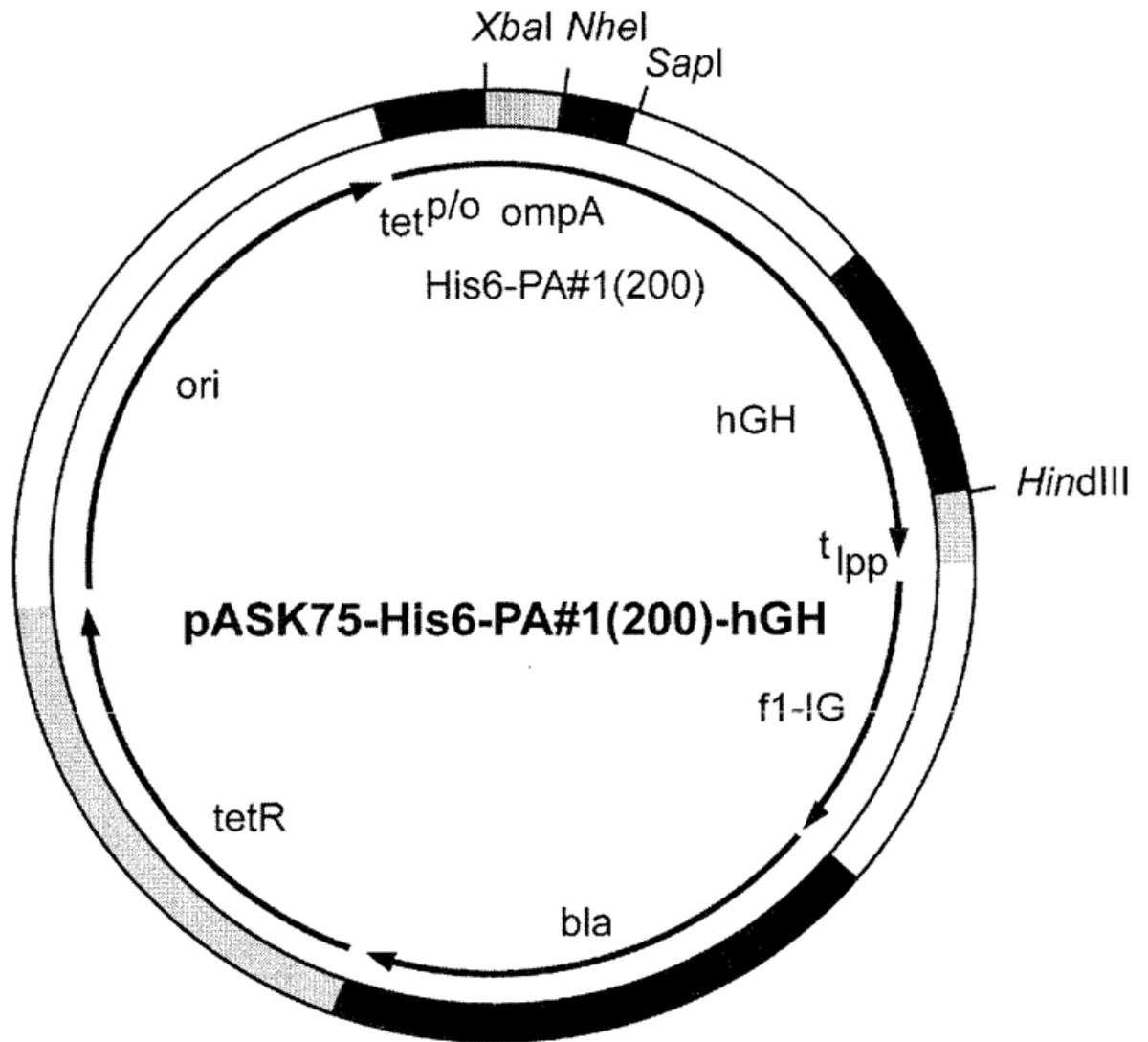


Figura 6 cont.

D

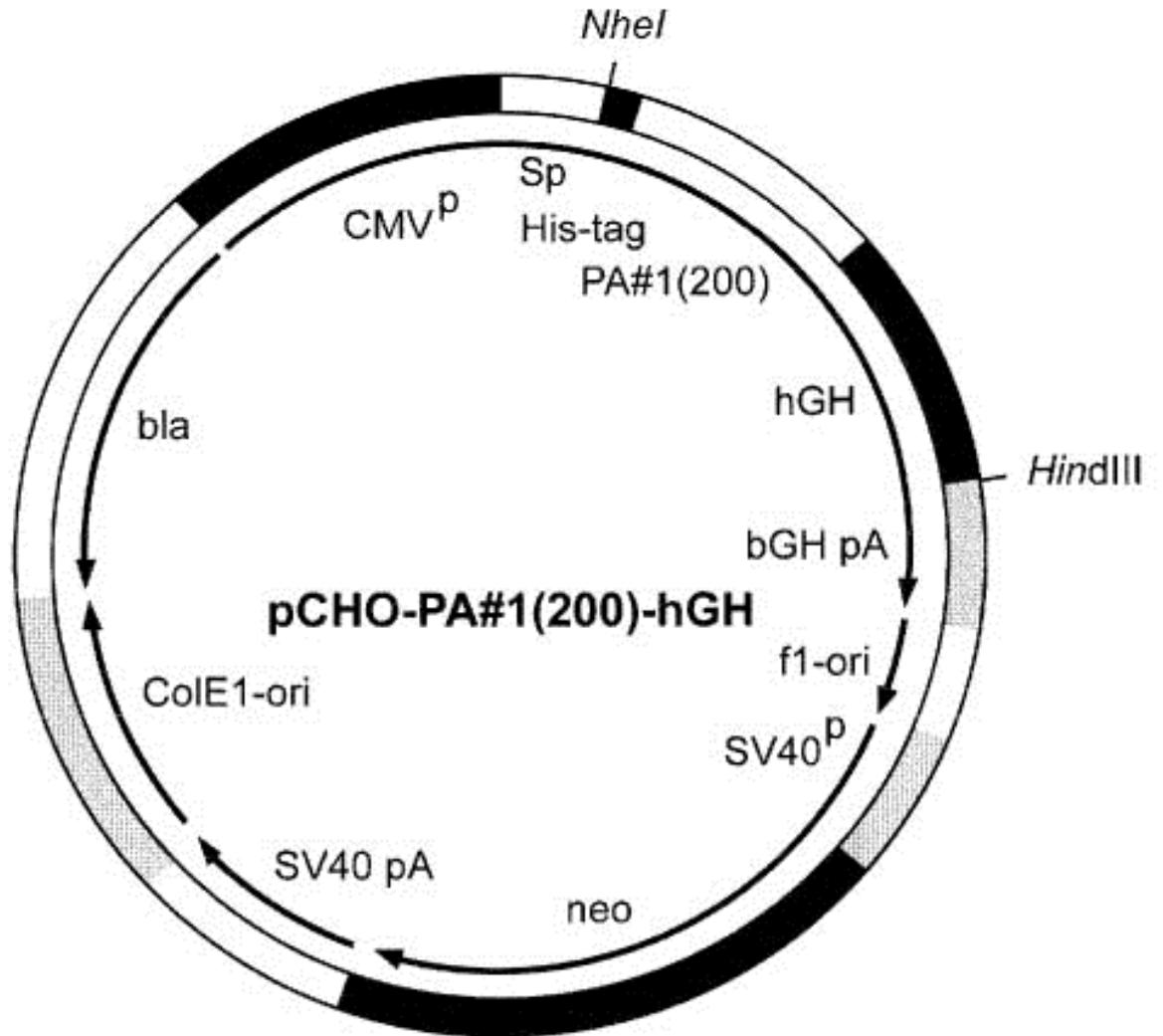


Figura 6 cont.

E

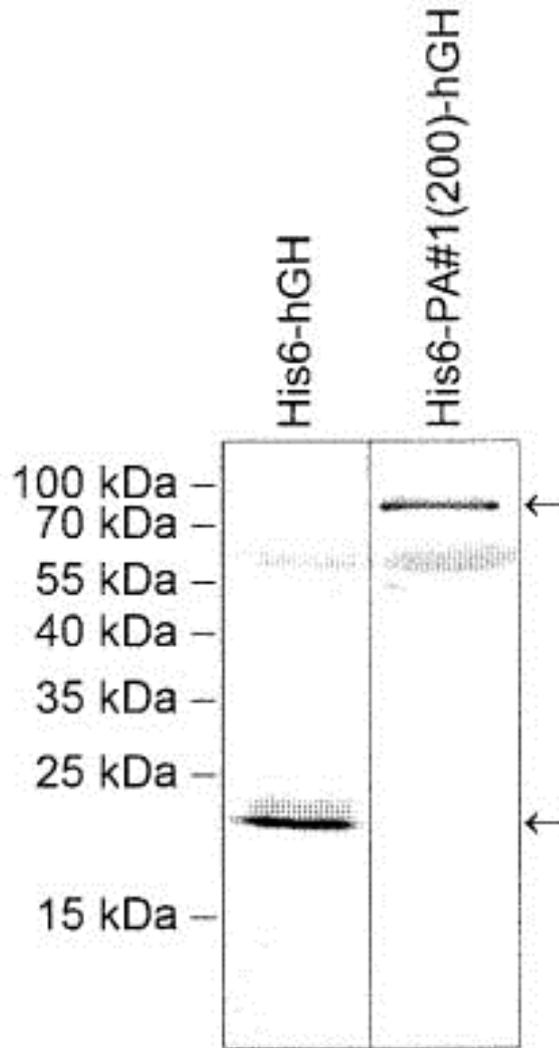


Figura 8:

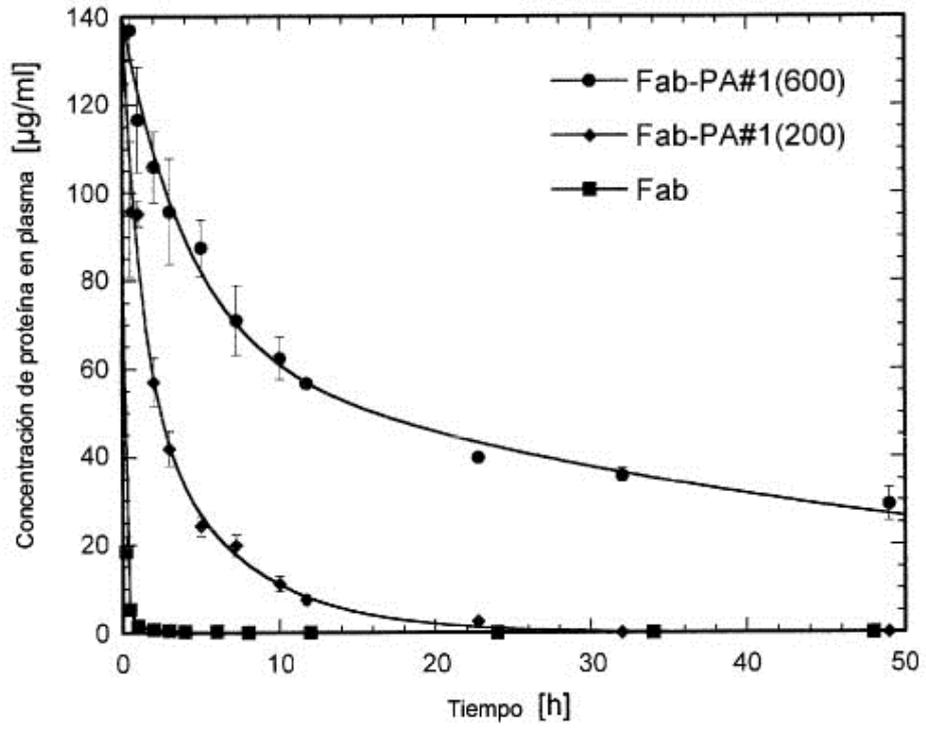


Fig.9

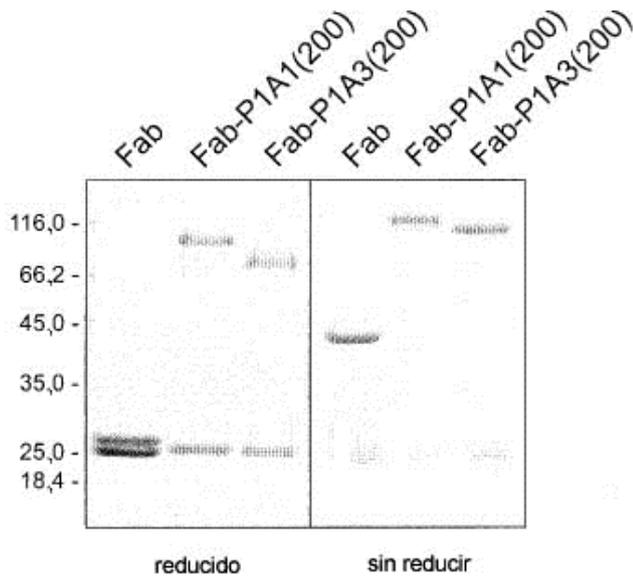


Fig. 10

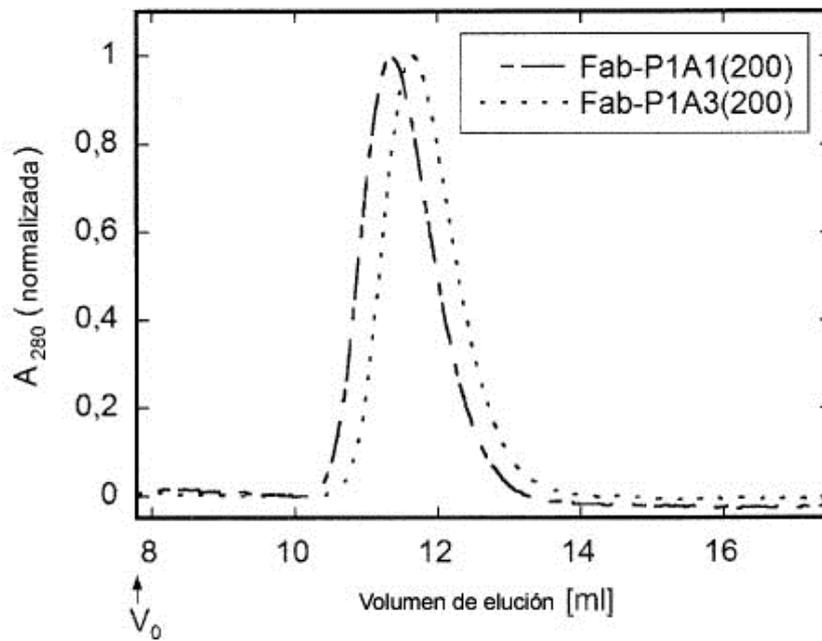
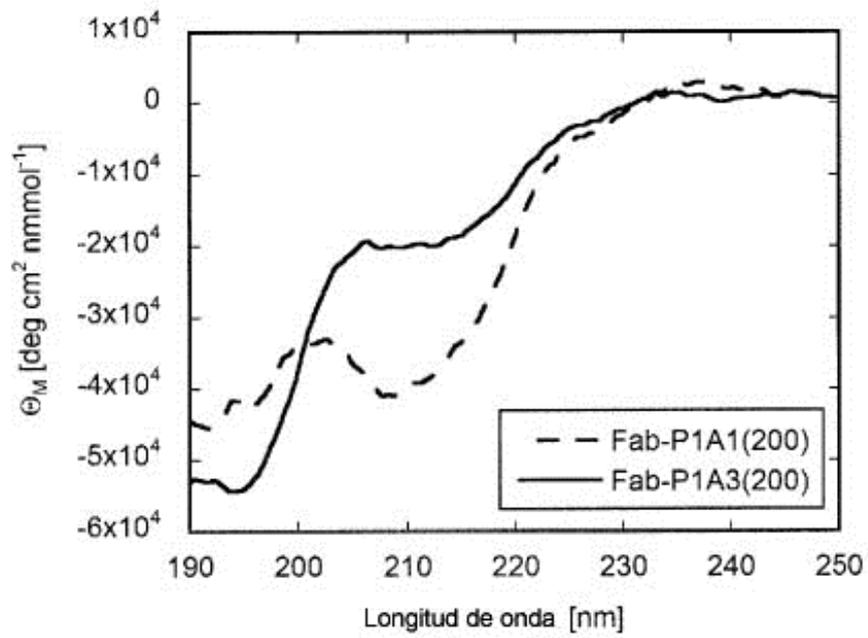


Fig. 11

A



B

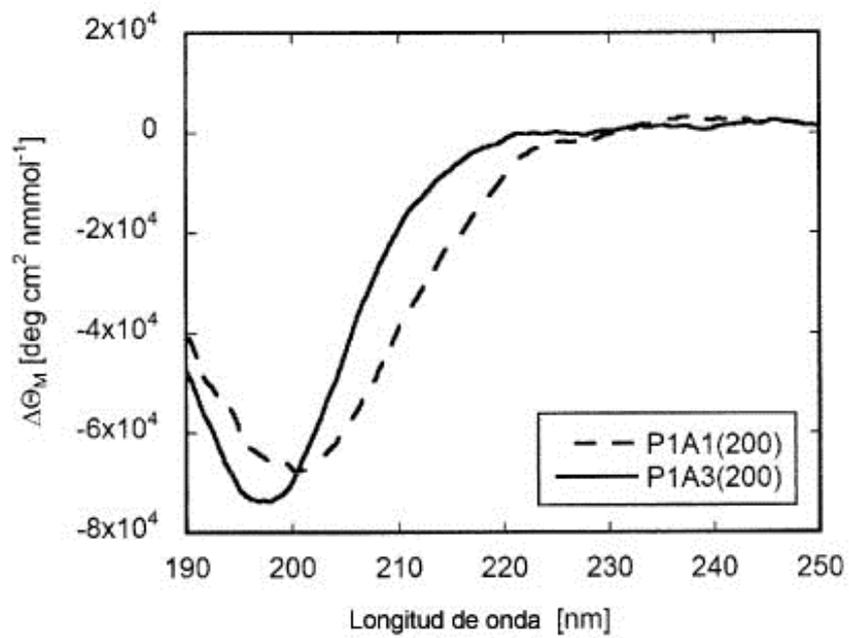
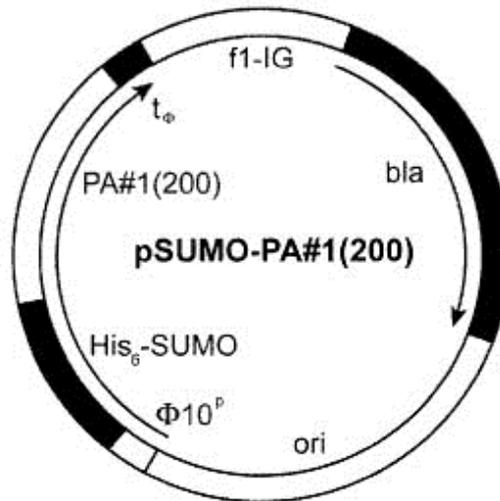
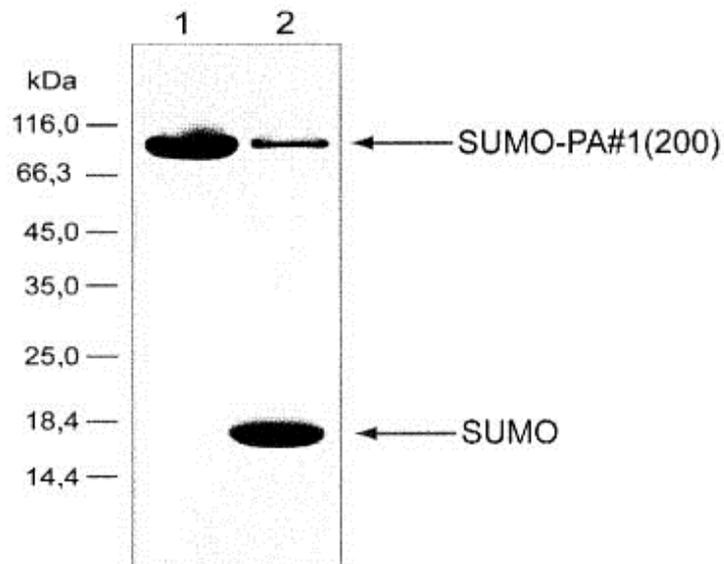


Fig. 12

A



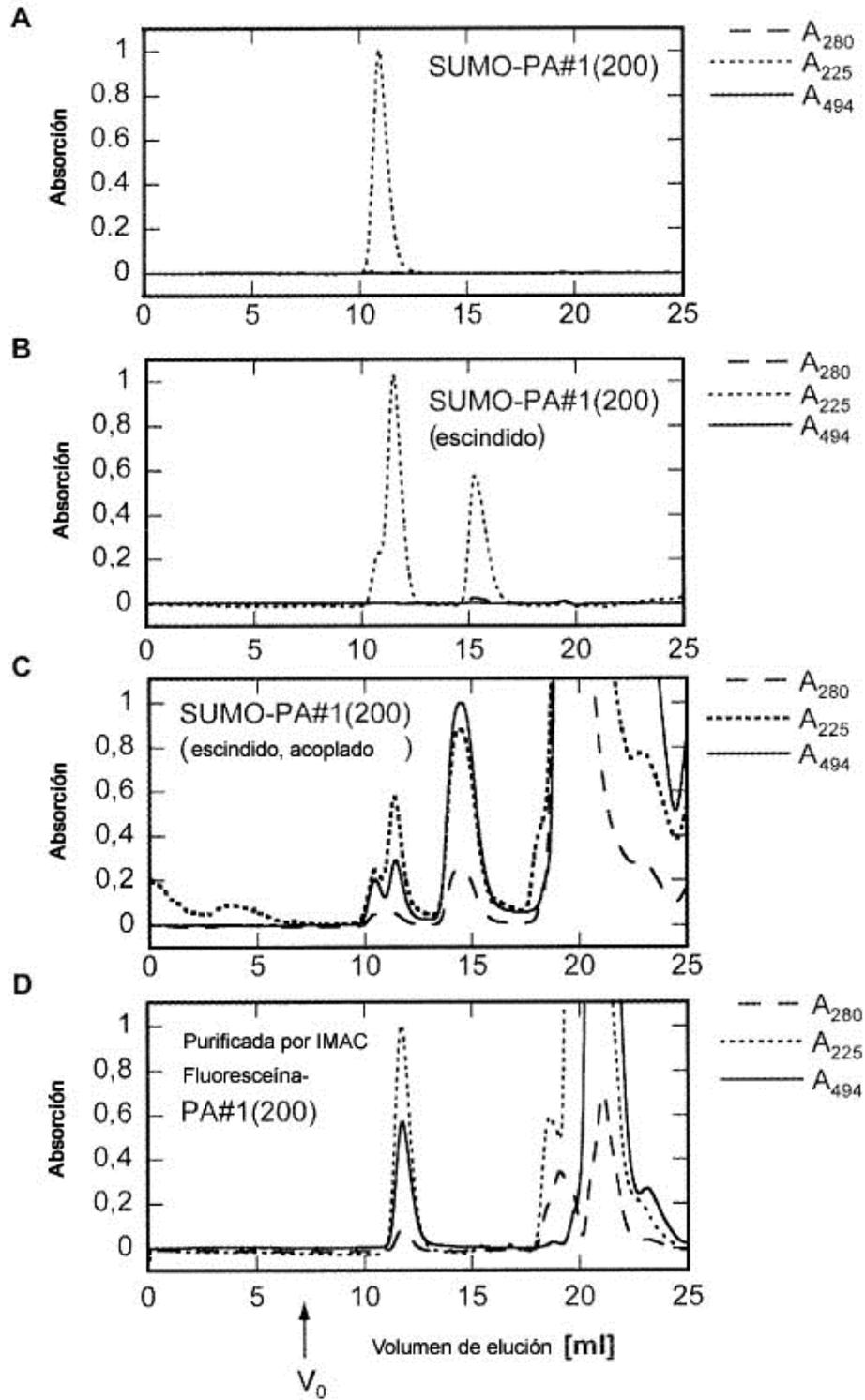
B

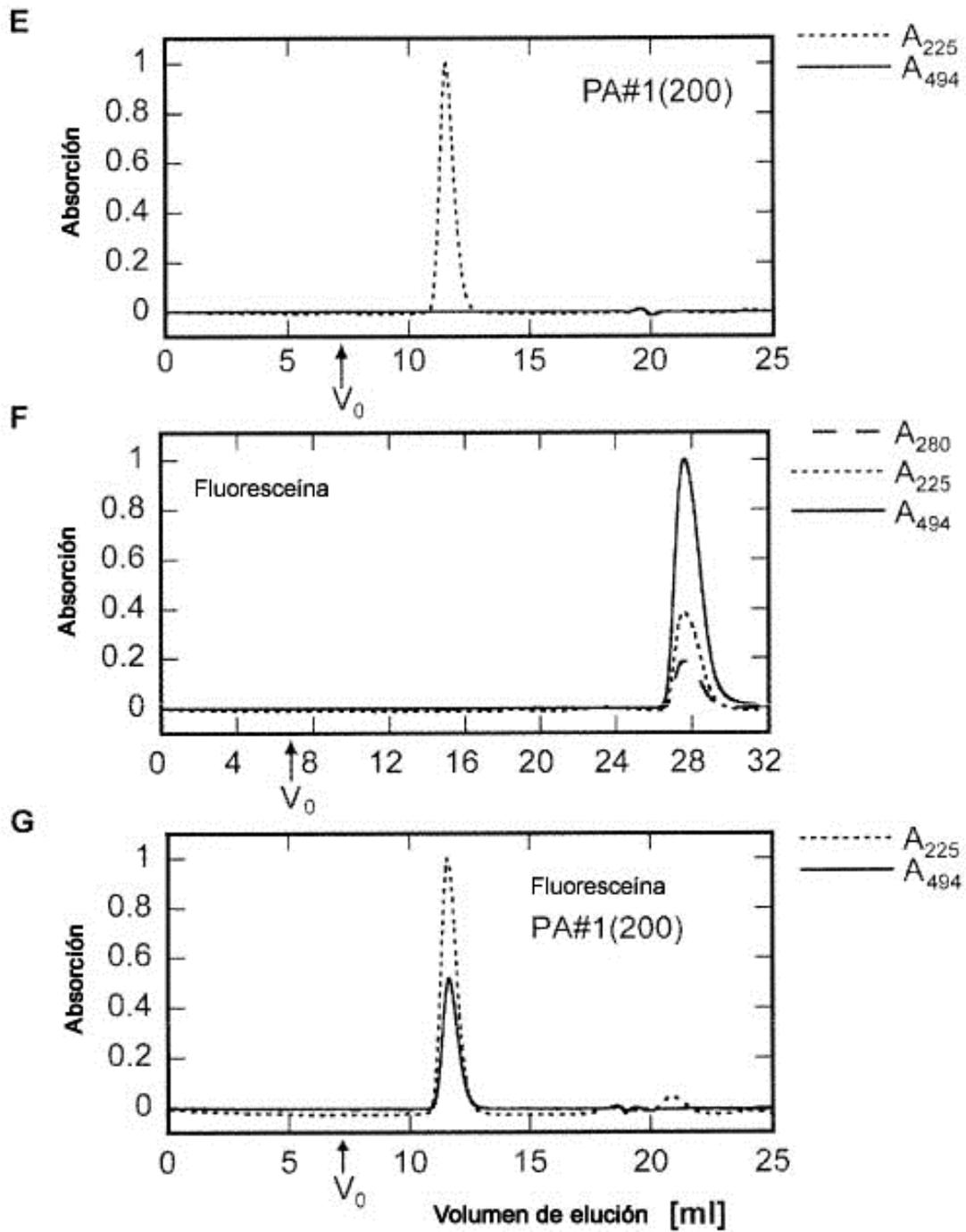


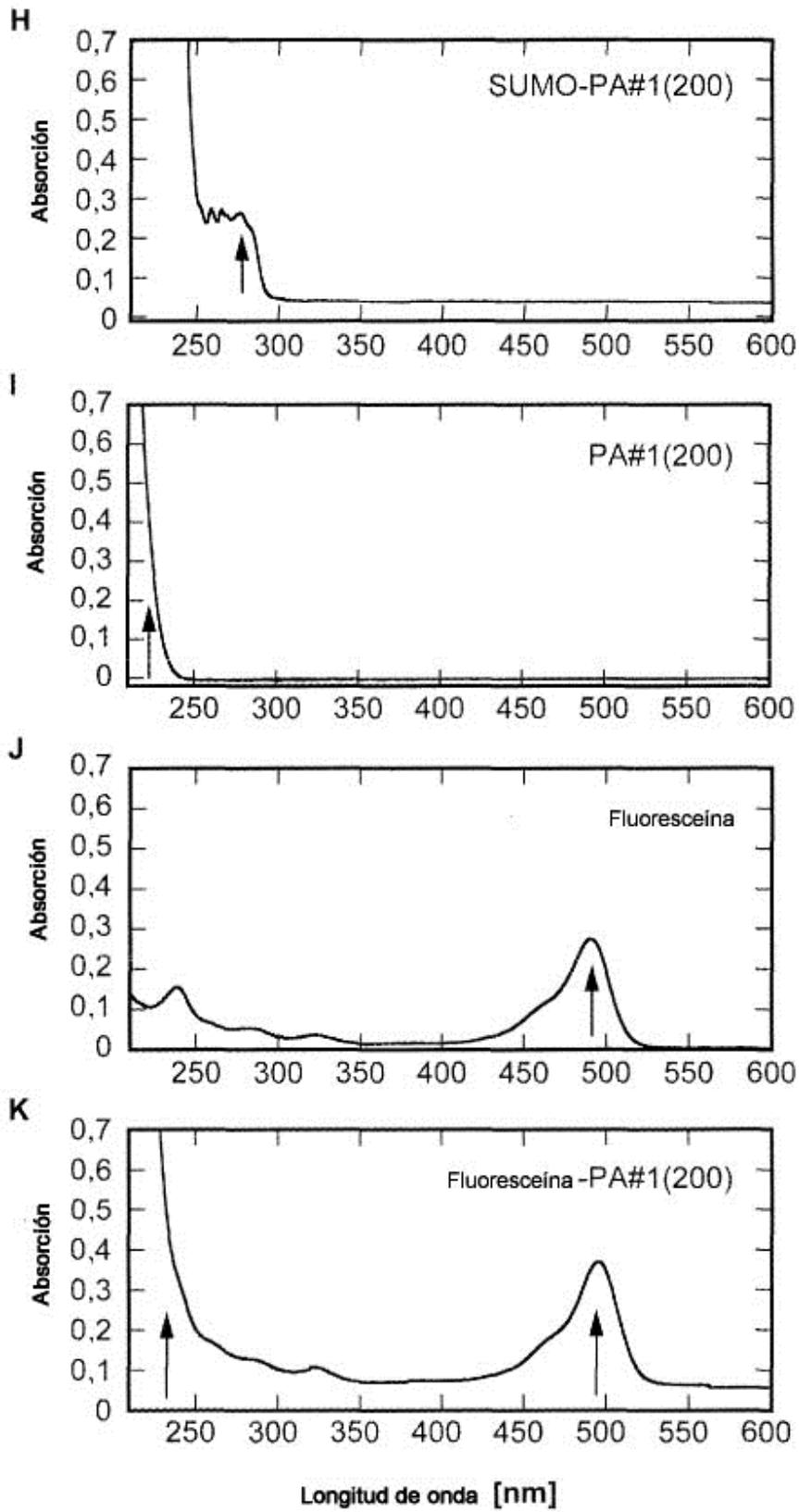
1: SUMO-PA#1(200) antes de la escisión

2: SUMO-PA#1(200) escindido con SUMO proteasa

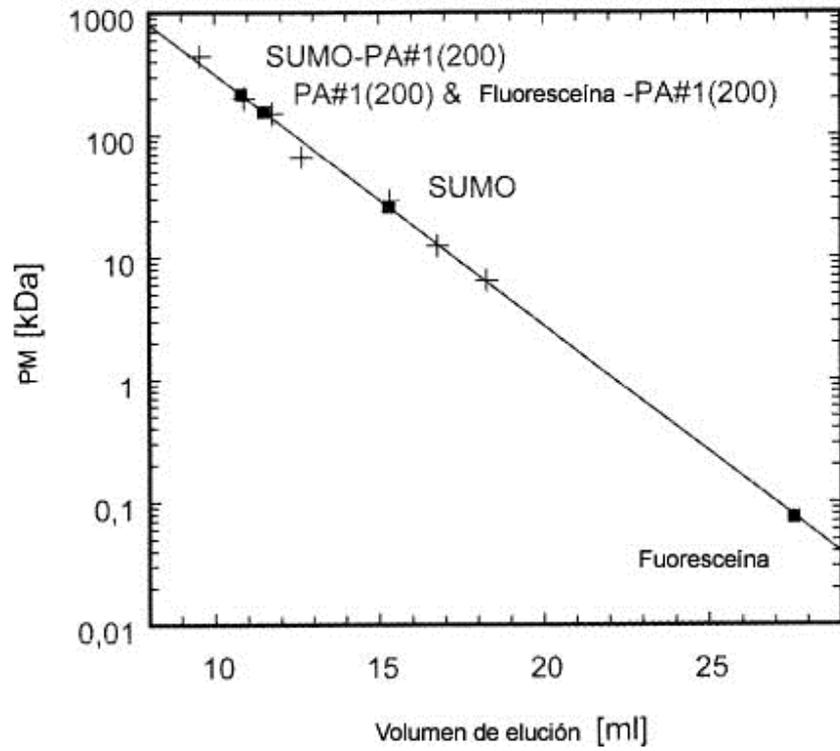
Fig. 13







L



M

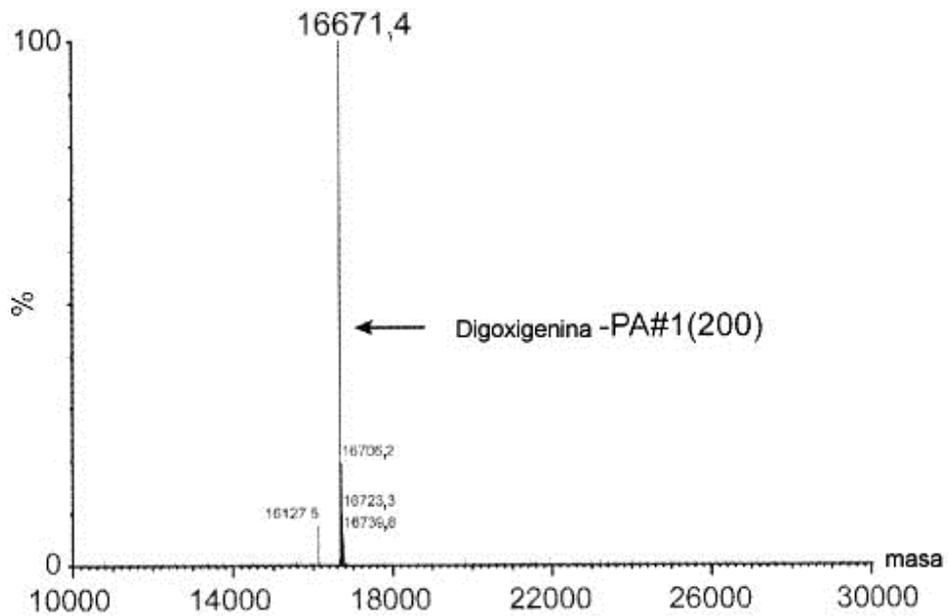
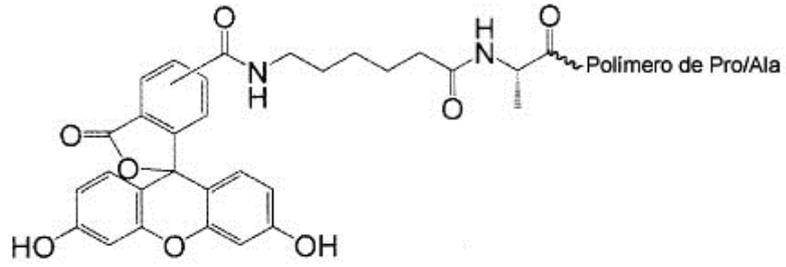


Fig. 14

A



B

