

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 799 431**

51 Int. Cl.:

C07K 1/14 (2006.01)

C07K 14/435 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.03.2011 PCT/EP2011/001605**

87 Fecha y número de publicación internacional: **06.10.2011 WO11120690**

96 Fecha de presentación y número de la solicitud europea: **30.03.2011 E 11713182 (1)**

97 Fecha y número de publicación de la concesión europea: **27.05.2020 EP 2552935**

54 Título: **Separación de proteínas objetivo insolubles**

30 Prioridad:

31.03.2010 US 319542 P

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

17.12.2020

73 Titular/es:

**AMSILK GMBH (100.0%)
Am Klopferspitz 19
82152 Planegg/Martinsried, DE**

72 Inventor/es:

**SCHMIDT, MARTIN;
LEIMER, AXEL y
RÖMER, LIN**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 799 431 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

Separación de proteínas objetivo insolubles

Campo técnico

5 La presente invención se refiere a un método para aislar una proteína de seda objetivo insoluble de una suspensión de células huésped intactas o rotas.

Antecedentes de la invención

Se han desarrollado sistemas para el aislamiento de proteínas objetivo de una suspensión de células huésped, por ejemplo, células microbianas, y/o de otras impurezas anteriormente y, por lo tanto, representan el conocimiento del estado de la técnica.

10 Como ejemplo, se hace referencia a los procedimientos estándar de purificación de proteínas para la purificación y aislamiento de proteínas objetivo solubles aplicando técnicas cromatográficas tales como cromatografía de intercambio iónico o de exclusión por tamaño (Guide to Protein Purification, Academic Press Vol. 182, 1990). Otras tecnologías aplicadas hoy en día son la separación de fases líquido-líquido y la ultrafiltración, entre otras (Guide to Protein Purification, Academic Press Vol. 182, 1990). Con variaciones, estos procesos fundamentales de purificación
15 pueden modificarse para purificar la mayoría de las proteínas necesarias para fines científicos o industriales, sin embargo, generalmente son muy costosas, complejas y requieren mucho tiempo.

Un método especial para manejar proteínas de seda en particular ha sido descrito en Huemmerich et al., Biochemistry 2004, 43, 13604-13612. En esta publicación, se aplicó con éxito una técnica que utiliza la desnaturalización térmica de las proteínas de la célula huésped después de una etapa de precipitación de la proteína objetivo, pero que carece
20 de métodos de purificación cromatográfica, para purificar proteínas de seda de araña recombinantes para aplicaciones técnicas. Como las proteínas de seda tienden a agregarse por sí mismas, este método adolece de la pérdida de una fracción significativa de la proteína objetivo, que se precipita y, por lo tanto, no está disponible para un método de purificación de proteínas solubles.

Aunque los métodos descritos funcionan bien para la mayoría de las proteínas solubles, es obvio que presentan graves desventajas cuando se manejan proteínas propensas a la agregación. Dichas proteínas propensas a la agregación
25 tienden a precipitarse en solución durante un cierto período de tiempo produciendo agregados proteicos estables, a menudo insolubles. Estos agregados de proteínas ya no pueden aislarse con los procesos de purificación de fracción de proteína soluble descritos. Por lo tanto, el tiempo de fermentación y/o el tiempo de purificación son generalmente parámetros críticos para evitar la precipitación no deseada. Se sabe que dichos agregados de proteínas se pueden
30 solubilizar usando varios detergentes, sin embargo, también se sabe que dicha solubilización afecta negativamente el rendimiento de la proteína así como la calidad de la proteína. Además, una solubilización completa de los agregados de proteínas durante un período de tiempo prolongado es casi imposible.

Por lo tanto, existe la necesidad de desarrollar un método de purificación/aislamiento para proteínas objetivo propensas a la agregación y/o proteínas objetivo ya agregadas que se enfoque en la separación de la fracción de
35 proteína objetivo propensa a la agregación y/o la fracción de proteína objetivo agregada a partir de la fracción que comprende proteínas insolubles de la célula huésped y otros residuos, sin solubilización completa de dichas proteínas objetivo. Tal método de purificación debería permitir el aislamiento de proteínas objetivo insolubles a partir de una suspensión de células huésped, por ejemplo, células microbianas, y/u otros residuos celulares, con altos rendimientos y alta calidad. Tal método de purificación también debe ser rentable, rápido, fácil y reproducible.

40 Los inventores de la presente invención han descubierto sorprendentemente que las proteínas objetivo son insolubles y permanecen insolubles bajo ciertas condiciones que son necesarias para solubilizar todas o casi todas las demás proteínas insolubles de la célula huésped, residuos de la célula huésped y/u otras impurezas potenciales relacionadas con la fermentación y que esto permite la separación y purificación de estas proteínas objetivo en solo unas pocas
45 etapas de purificación sin perder cantidades significativas de dichas proteínas objetivo debido a la solubilización no deseada de dichas proteínas objetivo o debido a otras reacciones cruzadas.

Sorprendentemente, se puede lograr una purificación de proteínas objetivo insolubles hasta preferiblemente al menos un 80% de pureza separando las proteínas objetivo insolubles de las partes insolubles de la célula huésped solubilizadas de acuerdo con el método de la presente invención. Incluso si algunas partes insolubles de la célula
50 huésped no se disuelven por completo pero permanecen en suspensión, se puede lograr una separación eficiente, ya que la densidad específica de estas partes de las células huésped suspendidas es menor que la de las proteínas objetivo insolubles, lo que permite una separación eficiente por centrifugación, sedimentación y/o filtración.

El método de la presente invención ofrece varias ventajas sorprendentes en comparación con las técnicas de purificación conocidas. Por ejemplo, el método de la presente invención reduce el número de etapas de purificación complejas y es, por lo tanto, un método de purificación rápido, confiable y fácil de realizar. Además, dicho método
55 permite la purificación/aislamiento de las proteínas objetivo insolubles con alto rendimiento y calidad. Permite una reducción drástica de costos en comparación con los métodos convencionales. Además, dicho método es amigable

con el medio ambiente.

Aunque el método de la presente invención comprende solo unas pocas etapas del método, hasta ahora era completamente imprevisible que ciertas proteínas objetivo sean y permanezcan insolubles en las condiciones aplicadas. Especialmente cuando se considera en el contexto que las concentraciones de base aplicadas (NaOH ~ 0,1 M) se usan comúnmente para lavar y limpiar biorreactores, así como tanques de reacción que entraron en contacto con proteínas y células. Por lo tanto, es poco evidente que estas condiciones también se puedan aplicar para purificar proteínas objetivo insolubles a gran escala.

Sumario de la invención

La presente invención está definida por las reivindicaciones. Cualquier tema relacionado que quede fuera del alcance de las reivindicaciones se proporciona solo a título informativo.

En un primer aspecto, la presente invención proporciona un método para aislar una proteína de seda objetivo insoluble de una suspensión de células huésped intactas o rotas que comprende las etapas de:

- a) proporcionar una suspensión de células huésped intactas o rotas que comprende una proteína objetivo insoluble y partes insolubles de la célula huésped,
- b) agregar una solución acuosa de al menos una base a dicha suspensión en una cantidad que sea suficiente para romper dichas células huésped y/o solubilizar dichas partes insolubles de la célula huésped, y
- c) separar la proteína objetivo insoluble de las partes solubilizadas de la célula huésped insoluble,

en el que en la etapa b) la proteína objetivo permanece insoluble, y en el que al menos el 80% de las partes insolubles de la célula huésped se solubilizan.

Descripción detallada de el invento

A menos que se defina lo contrario en el presente documento, todos los términos técnicos y científicos utilizados aquí tienen los mismos significados que comúnmente entiende un experto en la materia.

Preferentemente, los términos utilizados en este documento se definen como se describe en "A multilingual glossary of biotechnological terms: (IUPAC Recommendations)", Leuenberger, H.G.W, Nagel, B. and Kölbl, H. editores (1995), Helvetica Chimica Acta, CH-4010 Basilea, Suiza).

A lo largo de esta memoria descriptiva y las reivindicaciones que siguen, a menos que el contexto requiera lo contrario, se entenderá que la palabra "comprende", y variaciones tales como "que comprende", implican la inclusión de un número entero o etapa establecido o grupo de números enteros o etapas, pero no la exclusión de ningún otro número entero o etapa o grupo de números enteros o etapas.

Como se usa en esta memoria descriptiva y en las reivindicaciones adjuntas, las formas singulares "un", "uno, una" y "el, la" incluyen referentes plurales, a menos que el contenido indique claramente lo contrario.

Se dice que los residuos en dos o más polipéptidos se "corresponden" entre sí si los residuos ocupan una posición análoga en las estructuras de polipéptidos. Es bien sabido en la técnica que las posiciones análogas en dos o más polipéptidos pueden determinarse alineando las secuencias de polipéptidos basándose en la secuencia de aminoácidos o similitudes estructurales. Dichas herramientas de alineación son bien conocidas por el experto en la materia y se pueden obtener, por ejemplo, en la Internet, por ejemplo, ClustalW (www.ebi.ac.uk/clustalw) o Align (<http://www.ebi.ac.uk/emboss/align/index.html>) usando la configuración estándar, preferiblemente para Align EMBOSS::needle, Matrix: Blosum62, Apertura de hueco 10,0, Extensión de hueco 0,5.

La invención proporciona un método para aislar una proteína de seda objetivo insoluble de una suspensión de células huésped intactas o rotas que comprende las etapas de:

- a) proporcionar una suspensión de células huésped intactas o rotas que comprende una proteína objetivo insoluble y partes insolubles de la célula huésped,
- b) agregar una solución acuosa de al menos una base a dicha suspensión en una cantidad que sea suficiente para romper dichas células huésped y/o solubilizar dichas partes insolubles de la célula huésped, y
- c) separar la proteína objetivo insoluble de las partes insolubles de la célula huésped solubilizadas,

en el que en la etapa b) la proteína objetivo permanece insoluble, y en el que al menos 80%, preferiblemente al menos 85%, más preferiblemente al menos 90%, y lo más preferiblemente al menos el 95%, o incluso el 100%, de las partes insolubles de la célula huésped se solubilizan.

Por ejemplo, la invención proporciona un método para aislar una proteína objetivo insoluble de una suspensión de células huésped rotas que comprende las etapas de:

- a) proporcionar una suspensión de células huésped rotas que comprende una proteína objetivo insoluble y partes

insolubles de la célula huésped,

b) agregar una solución acuosa de al menos una base a dicha suspensión en una cantidad que sea suficiente para solubilizar dichas partes insolubles de la célula huésped, y

c) separar la proteína objetivo insoluble de las partes insolubles de la célula huésped solubilizadas,

- 5 en el que en la etapa b) la proteína objetivo permanece insoluble, y en el que al menos 80%, preferiblemente al menos 85%, más preferiblemente al menos 90%, y lo más preferiblemente al menos el 95%, o incluso el 100%, de las partes insolubles de la célula huésped se solubilizan.

Además, por ejemplo, la invención proporciona un método para aislar una proteína objetivo insoluble de una suspensión de células huésped intactas que comprende las etapas de:

- 10 a) proporcionar una suspensión de células huésped intactas que comprende una proteína objetivo insoluble y partes insolubles de la célula huésped,

b) agregar una solución acuosa de al menos una base a dicha suspensión en una cantidad que sea suficiente para romper dichas células huésped y para solubilizar dichas partes insolubles de la célula huésped, y

c) separar la proteína objetivo insoluble de las partes insolubles de la célula huésped solubilizadas,

- 15 en el que en la etapa b) la proteína objetivo permanece insoluble, y en el que al menos 80%, preferiblemente al menos 85%, más preferiblemente al menos 90%, y lo más preferiblemente a al menos el 95%, o incluso el 100%, de las partes insolubles de la célula huésped se solubilizan.

- 20 Los inventores de la presente invención han encontrado sorprendentemente que las proteínas objetivo insolubles, por ejemplo, las proteínas de seda de araña, que están presentes en una suspensión de células huésped intactas o rotas, permanecen insolubles bajo ciertas condiciones que se requieren para romper las células huésped intactas y/o para solubilizar todas o casi todas las otras partes insolubles de la célula huésped, por ejemplo, las proteínas insolubles de la célula huésped, paredes de las células huésped, residuos de las células huésped u otras impurezas potenciales
25 relacionadas con la fermentación (por ejemplo, residuos de fermentación) y que, por lo tanto, dichas proteínas objetivo insolubles, por ejemplo, proteínas de seda de araña, simplemente se pueden separar y purificar a partir de dichas suspensiones en solo unas pocas etapas de purificación sin perder cantidades significativas de proteínas.

- 30 En detalle, los inventores de la presente invención han encontrado inesperadamente que la separación de las proteínas objetivo insolubles de las partes insolubles de la célula huésped se pueden lograr mediante la adición de una solución acuosa que comprende al menos una base (por ejemplo, NaOH) en una concentración baja (por ejemplo, NaOH 0,05 M) a una suspensión de células huésped intactas o rotas. Los inventores han descubierto que la solución acuosa que comprende al menos una base altera las células huésped y/o solubiliza las proteínas insolubles de la célula huésped y los residuos celulares restantes sin afectar la proteína objetivo insoluble.

- 35 Los inventores han comprobado además que después de esta etapa, solo se requiere separar la fase sólida que comprende la proteína objetivo insoluble de la fase líquida que comprende las partes insolubles de las células huésped solubilizadas, por ejemplo, por centrifugación y/o filtración, y opcionalmente lavar el precipitado de proteína objetivo para purificar la proteína objetivo insoluble de interés. La proteína objetivo insoluble purificada resultante se puede usar directamente para aplicaciones científicas o industriales sin procesamiento adicional.

- 40 Los inventores han determinado sorprendentemente que con el método de la presente invención una proteína objetivo insoluble que tiene una pureza de al menos 80%, preferiblemente de al menos 85% o 90%, más preferiblemente de al menos 95%, 98%, o 99%, y lo más preferiblemente de al menos 99,9%, o incluso de 100%, por ejemplo al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, pueden ser aisladas. La mayoría de las proteínas objetivo insolubles, aisladas con el método de la presente invención, no requieren etapas de purificación adicionales. Por consiguiente, el método de la presente invención representa un método de purificación
45 muy efectivo y que ahorra costes.

- El término "células huésped", como se usa en el presente documento, se refiere a células que comprenden una proteína objetivo de interés, por ejemplo, una proteína de seda tal como una proteína de seda de araña o una variante de la misma. Dicha proteína objetivo, por ejemplo, una proteína de seda tal como una proteína de seda de araña o una variante de la misma, puede ser codificada por un polinucleótido, preferiblemente por un polinucleótido aislado.
50 Dicho polinucleótido se puede encontrar dentro de las células huésped (i) libremente disperso como tal, (ii) incorporado en un vector recombinante, o (iii) integrado en el genoma de las células huésped o en el ADN mitocondrial. Las células huésped pueden usarse para la amplificación y expresión de un polinucleótido que codifica una proteína objetivo de interés, por ejemplo, una proteína de seda tal como una proteína de seda de araña o una variante de la misma.

- El término "polinucleótido aislado", como se usa en el presente documento, se refiere a un polinucleótido que (i) se aisló de su entorno natural, (ii) se amplificó por reacción en cadena de la polimerasa, y/o (iii) se sintetizó total o parcialmente, y significa un polímero mono o bicatenario de bases de desoxirribonucleótidos o ribonucleótidos e incluye una molécula de ADN y ARN, tanto cadena sentido como antisentido. El término comprende ADNc, ADN genómico, ARNm y ADN recombinante. Un polinucleótido puede constar de un gen completo, o una porción del mismo.
55

En una realización preferida de la presente invención, el polinucleótido mencionado anteriormente es un polinucleótido recombinante que codifica la proteína objetivo de interés. El término "polinucleótido recombinante" se refiere a un polinucleótido sintetizado o bien manipulado *in vitro*. La proteína objetivo que está codificada por dicho polinucleótido recombinante puede designarse como una proteína objetivo recombinante. Una célula huésped que comprende dicho polinucleótido recombinante y/o dicha proteína objetivo recombinante puede designarse como una célula huésped recombinante.

El término "vectores recombinantes", como se usa en el presente documento, incluye cualquier vector conocido por la persona experta, incluidos vectores plasmídicos, vectores cósmidos, vectores fagos tales como fago lambda, vectores virales tales como vectores adenovirales o baculovirales, o vectores cromosómicos artificiales tales como cromosomas artificiales bacterianos (BAC), cromosomas artificiales de levadura (YAC) o cromosomas artificiales P1 (PAC). Dichos vectores incluyen vectores de expresión así como de clonación. Los vectores de expresión comprenden plásmidos así como vectores virales y generalmente contienen una secuencia de codificación deseada y secuencias de ADN apropiadas necesarias para la expresión de la secuencia de codificación operativamente unida en un organismo huésped particular (por ejemplo, bacteria, levadura o planta) o en sistemas expresión *in vitro*. Los vectores de clonación se usan generalmente para modificar y amplificar un cierto fragmento de ADN deseado y pueden carecer de secuencias funcionales necesarias para la expresión de los fragmentos de ADN deseados.

En una realización preferida de la presente invención, el polinucleótido que codifica la proteína objetivo de interés está comprendido en un vector de expresión y está operativamente unido a secuencias de control de expresión para regular su expresión en una célula huésped o está comprendido en un vector de clonación para permitir su amplificación en una célula huésped. El vector de clonación o el vector de expresión que comprende el polinucleótido que codifica la proteína objetivo de interés se introduce normalmente en una célula huésped mediante transformación o transfección.

En una realización preferida de la presente invención, las células huésped son células huésped microbianas tales como células huésped bacterianas o de levadura, células huésped de plantas o células huésped de insectos.

En el contexto de la presente invención, el término "células huésped microbianas", como se usa en el presente documento, se refiere a células de origen microbiano, por ejemplo, células bacterianas tales como células bacterianas Gram negativas o Gram positivas, por ejemplo, células de *Escherichia* (por ejemplo, *Escherichia coli*), *Anabaena*, *Caulobacter*, *Gluconobacter*, *Rhodobacter*, *Pseudomonas*, *Para coccus*, *Bacillus* (por ejemplo, *Bacillus subtilis*), *Brevibacterium*, *Corynebacterium*, *Rhizobium* (*Sinorhizobium*), *Flavobacterium*, *Klebsiella*, *Enterobacter*, *Lactobacillus*, *Lactococcus*, *Methylobacterium*, *Propionibacterium*, *Staphylococcus* o *Streptomyces*, o células de levadura tales como células de ascoporógeno (Endomycetales), células de basidiosporógenos, o células pertenecientes a los hongos imperfectos (Blastomicetos), por ejemplo, células de *Candida*, *Hansenula*, *Kluyveromyces*, *Saccharomyces* (por ejemplo, *Saccharomyces cerevisiae*), *Schizosaccharomyces*, *Pichia* (por ejemplo, *Pichia pastoris*) o *Yarrowia*.

En el contexto de la presente invención, el término "células huésped de insecto" se refiere a células de origen en insectos tales como células de *Spodoptera frugiperda* o *Trichoplusia ni*. Preferiblemente, las células de insecto son células SF9, células SF-21 o células High-Five. Las células SF-9 y SF-21 son células de ovario de *Spodoptera frugiperda*. Las células High-Five son células de óvulos de *Trichoplusia ni*.

En el contexto de la presente invención, el término "células huésped de plantas", como se usa en el presente documento, se refiere a células de origen vegetal tales como células de tabaco, patata o guisantes.

El término "una suspensión de células huésped intactas o rotas", como se usa en el presente documento, se refiere a células huésped intactas o rotas que contienen un líquido heterogéneo. El líquido, en el que las células intactas o rotas están suspendidas, puede ser un medio de fermentación, un medio de cultivo, una solución acuosa, por ejemplo, una solución acuosa tamponada, H₂O técnica o H₂O desionizada. La solución acuosa tamponada puede ser, por ejemplo, Tris/HCl. El pH de la solución acuosa tamponada puede estar entre pH 5,0 y pH 9,0, preferiblemente entre pH 6,0 y pH 8,0, y más preferiblemente entre pH 6,7 y pH 7,2, por ejemplo, Tris/HCl, pH 7,0, pH 7,5 o pH 8,0. Se prefiere que la solución acuosa tamponada sea una solución de Tris/HCl entre 10 y 100 mM, más preferiblemente de Tris/HCl entre 10 y 50 mM, y lo más preferiblemente de Tris/HCl entre 10 y 20 mM, por ejemplo, Tris/HCl 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 100 mM, en la que dicha solución acuosa tamponada tiene preferiblemente un pH de entre 5,0 y 9,0, más preferiblemente un pH de entre 6,0 y 8,0, y lo más preferiblemente un pH de entre 6,7 y 7,2, por ejemplo Tris/HCl, pH 7,0, pH 7,5 o pH 8,0. El medio de cultivo puede ser un medio mínimo, o un medio salino (véase, por ejemplo, Korz et al., Journal of Biotechnology 39, 1995, 59-65). La proporción de líquido con respecto a las células huésped intactas o rotas en la suspensión puede variar, por ejemplo, entre 5% y 95%. Por lo tanto, el término "suspensión de células huésped intactas o rotas" abarca suspensiones, en las que la proporción de líquido con respecto a las células huésped intactas o rotas puede ser del 5 al 95%, del 10 al 90%, del 15 al 85%, del 20 al 80%, del 25 al 75%, del 30 al 70%, del 35 al 65%, del 40 al 60%, del 45 al 55%, del 50 al 50%, del 55 al 45%, del 60 al 40%, del 65 al 35%, del 70 al 30%, del 75 al 25%, del 80 al 20%, del 85 al 15%, del 90 al 10%, o del 95 al 5%.

En el contexto de la presente invención, el término "una suspensión de células huésped intactas" abarca células huésped intactas, generalmente no lisadas, por ejemplo, células huésped microbianas, células huésped de plantas o células huésped de insectos, suspendidas en un líquido, por ejemplo, en una solución acuosa (por ejemplo, una

solución acuosa tamponada), agua técnica, agua desionizada, medio de cultivo o medio de fermentación. El término "una suspensión de células huésped rotas", como se usa en el presente documento, abarca células huésped, por ejemplo, células huésped microbianas, células huésped de plantas o células huésped de insectos, que tienen paredes celulares rotas o están lisas en gran parte y que están suspendidas en un líquido, por ejemplo, en una solución acuosa (por ejemplo, una solución acuosa tamponada), agua técnica, agua desionizada, medio de cultivo o medio de fermentación. La "suspensión de células huésped rotas" puede prepararse mediante (i) alteración no mecánica de las células (por ejemplo, lisis celular con enzimas, tratamiento celular con reactivos químicos con el fin de disolver/disolver parcialmente o para abrir/abrir parcialmente las paredes celulares y/o membranas celulares, o romper las células usando presión osmótica), (ii) alteración mecánica de las células (por ejemplo, molienda mecánica o extracción con molino de bolas), (iii) homogeneización a alta presión, o (iv) sonicación, y mediante combinaciones de los mismos.

En una realización preferida de la presente invención, las células huésped intactas, por ejemplo, las células huésped microbianas tales como las células huésped bacterianas, que se proporcionan en la etapa a) están presentes en una suspensión que tiene un contenido de humedad de entre 5 y 20%, preferiblemente de 5, 10, 15 o 20% (por ejemplo, de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20%), es decir, en un sedimento celular, preferiblemente después de la separación de dichas células del medio de cultivo, por ejemplo, por centrifugación o filtración. En otra realización preferida de la presente invención, las células huésped rotas, por ejemplo, las células microbianas, tales como las células bacterianas, están presentes en una suspensión que tiene un contenido de humedad de entre 5 y 20%, preferiblemente de 5, 10, 15 o 20% (por ejemplo, de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20%), es decir, en un sedimento celular, preferiblemente después de la separación de dichas células rotas, por ejemplo, lisadas o sonicadas, del medio de cultivo, por ejemplo, por centrifugación o filtración.

Debe observarse que la cantidad de partes solubles de la célula huésped, por ejemplo, orgánulos, proteínas solubles de células huésped o residuos solubles de células huésped, y/o de otros ingredientes solubles, por ejemplo, impurezas solubles relacionadas con la fermentación, las partes solubles del medio de fermentación, o las partes solubles del medio de cultivo, en la suspensión de células huésped intactas o rotas proporcionadas en la etapa a), pueden variar. Por ejemplo, la cantidad de partes solubles de la célula huésped y/o de otros ingredientes solubles es menor (i) en un sedimento celular de células huésped intactas o rotas que tienen un contenido de humedad de entre 5 y 10%, producido al separar las células huésped intactas o rotas del medio de cultivo o del medio de fermentación mediante centrifugación o filtración, o (ii) en una solución acuosa, en la que dicho sedimento celular se resuspende, en comparación con la cantidad de partes solubles de la célula huésped y/o de otros ingredientes solubles (i) en una suspensión de células huésped intactas en un medio de cultivo o medio de fermentación, o (ii) en una suspensión de células huésped rotas en un medio de cultivo o medio de fermentación. También es posible que no estén presentes partes solubles de células huésped y/u otros ingredientes solubles en la suspensión, por ejemplo, solución acuosa, particularmente en los casos en los que un sedimento celular de células huésped rotas se lava y se resuspende en una solución acuosa que se proporciona posteriormente en la etapa a).

Una "suspensión de células huésped", por ejemplo, células huésped microbianas, de plantas o de insectos pueden proporcionarse/producirse cultivando células huésped, por ejemplo, células huésped microbianas, de plantas o de insectos, en un cultivo o fermentación. La producción de proteínas objetivo puede realizarse en biorreactores en los que las células huésped, por ejemplo, células huésped microbianas, de plantas o de insectos, pueden cultivarse de manera eficiente. Se pueden aplicar tres estrategias de fermentación de intersección: (i) fermentación continua, (ii) fermentación discontinua o (iii) fermentación semicontinua. Estas estrategias de fermentación también se pueden combinar. Para obtener el máximo rendimiento, es decir, al proporcionar una alta biomasa, de la proteína objetivo deseada, puede ser necesario suministrar las células huésped, por ejemplo, células huésped microbianas, de plantas o de insectos, con suficientes nutrientes y combinar este tratamiento con un monitoreo continuo y la adaptación de los parámetros relevantes del proceso (por ejemplo, oxígeno disuelto, pH y temperatura). Pueden usarse sistemas de inducción de expresión de proteínas objetivo (por ejemplo, inducción de isopropil- β -D-tiogalactopiranosido (IPTG), inducción con azúcares o análogos de los mismos, inducción por temperatura usando promotores dependientes de la temperatura, inducción usando promotores controlados por estrés osmótico o inducción iniciada por cambios metabólicos dentro de la célula). Muchos parámetros pueden tener un impacto en la producción de proteínas objetivo y la calidad de las proteínas. Para aumentar el rendimiento de la proteína objetivo insoluble en particular, se puede aplicar una fase de inducción más larga, temperatura de incubación alterada, suministro de nutrientes optimizado y/o diferentes tipos de estrés celular. También puede ser posible influir en la cantidad de agregado de proteína objetivo producida durante el cultivo celular y el proceso de purificación controlando cuidadosamente el medio ambiente (por ejemplo, los componentes del medio, la temperatura durante el proceso de fermentación, la duración del proceso de fermentación) e implementando estrategias apropiadas para maximizar el grado de agregación.

Una "suspensión de células huésped", por ejemplo, las células huésped microbianas, de plantas o de insectos también se pueden proporcionar/producir resuspendiendo células huésped cultivadas, por ejemplo, células huésped microbianas, de plantas o de insectos, que se han separado del medio de cultivo, por ejemplo, por centrifugación o filtración, en una solución acuosa, por ejemplo, solución acuosa tamponada, en agua técnica o en agua desionizada. También se puede proporcionar/producir una suspensión de células de plantas suspendiendo células de plantas que se han obtenido de plantas completas, por ejemplo, extrayendo dichas células de plantas de dichas plantas o alterando plantas enteras, en una solución acuosa, por ejemplo, solución tamponada, en agua técnica o en agua desionizada.

En caso de que se pueda proporcionar una suspensión de células huésped intactas en la etapa a) del método de la

presente invención, se puede usar directamente una de las suspensiones de células huésped descritas anteriormente. En el caso de que se pueda proporcionar una suspensión de células huésped rotas en la etapa a) del método de la presente invención, las células huésped (intactas) comprendidas en una de las suspensiones descritas anteriormente se pueden romper en primer lugar, por ejemplo, utilizando métodos de alteración celular no mecánica, o métodos de alteración celular mecánica (véase más arriba).

En el contexto de la presente invención, el término "una suspensión de células huésped rotas que comprende una proteína objetivo insoluble y partes insolubles de la célula huésped" se refiere a una suspensión, por ejemplo, medio de cultivo, solución acuosa tamponada, agua técnica o agua desionizada, que comprende directamente la proteína objetivo insoluble y las partes insolubles de la célula huésped (por ejemplo, partes insolubles de la pared celular o residuos celulares insolubles) de las células huésped rotas. Por el contrario, el término "una suspensión de células huésped intactas que comprende una proteína objetivo insoluble y partes insolubles de la célula huésped", como se usa en el presente documento, se refiere a una suspensión, por ejemplo, medio de cultivo, solución acuosa (por ejemplo, solución acuosa tamponada), agua técnica o agua desionizada, que comprende células huésped intactas, en la que están comprendidas la proteína objetivo insoluble y las partes insolubles de la célula huésped (por ejemplo, la pared celular insoluble o las proteínas insolubles de la célula huésped).

El término "partes insolubles de la célula huésped", como se usa en el presente documento, se refiere a proteínas insolubles de la célula huésped, paredes celulares, membranas celulares, partes de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de células huésped y/o inclusiones citoplasmáticas insolubles (por ejemplo, cristales de oxalato de calcio o dióxido de silicio, gránulos de materiales de almacenamiento de energía tales como almidón, glucógeno o polihidroxibutirato), que son partes de las células huésped, por ejemplo, células huésped microbianas, de plantas o de insectos, o del sistema de expresión utilizado para expresar la proteína objetivo. Este término no abarca la proteína objetivo insoluble de la presente invención. El término "partes insolubles de la célula huésped" se refiere además a partes de células huésped que son insolubles en las condiciones que existen en la etapa a) y que se solubilizan en las condiciones que existen en la etapa b).

El término "partes solubles de células huésped", como se usa en este documento, se refiere a proteínas solubles de células huésped, orgánulos celulares, partes de orgánulos celulares y/u otros componentes celulares solubles, que son partes de las células huésped, por ejemplo, células huésped microbianas, de plantas o de insectos, o del sistema de expresión utilizado para expresar la proteína objetivo. El término "partes solubles de la célula huésped" se refiere además a las partes de la célula huésped que ya son solubles en las condiciones que existen en la etapa a) y que permanecen solubles en las condiciones que existen en la etapa b).

El término "proteína objetivo", como se usa en el contexto de la presente invención, se refiere a la proteína de interés, por ejemplo, una proteína de seda tal como una proteína de seda de araña o proteína de seda de insecto, colágeno, resilina o queratina, que puede aislarse de la suspensión de células huésped, por ejemplo, células huésped microbianas, de plantas o de insectos. En una realización preferida, dicha proteína objetivo es una proteína objetivo recombinante, más preferiblemente una proteína objetivo recombinante codificada por un polinucleótido recombinante. En una realización adicional, la proteína objetivo recombinante es una proteína híbrida de una proteína de seda de araña y una proteína de seda de insecto, una proteína de seda de araña y colágeno, una proteína de seda de araña y resilina o una proteína de seda de araña y queratina. Se prefiere particularmente que dicha proteína objetivo se produzca de forma recombinante en dichas células huésped.

En una realización preferida, las proteínas objetivo son proteínas que comprenden unidades/dominios repetidos que tienen la propiedad de formar agregados de proteínas objetivo en una célula huésped o en una suspensión, por ejemplo, en un medio de cultivo, en una solución acuosa (por ejemplo, una solución acuosa tamponada), en agua técnica o en agua desionizada.

En otra realización preferida, las proteínas objetivo están presentes en forma de agregados de proteínas objetivo en una célula huésped, o en una suspensión, por ejemplo, en un medio de cultivo celular, en una solución acuosa (por ejemplo, solución acuosa tamponada), en agua técnica o en agua desionizada. Dichos agregados de proteínas objetivo pueden formarse mediante la autoagregación de proteínas objetivo en una célula huésped sin la influencia del armazón celular u otros mecanismos intracelulares.

En particular, los agregados de proteínas objetivo pueden formarse mediante la autoagregación de múltiples copias/unidades de proteínas objetivo en un cuerpo o masa sólida sin la influencia del armazón celular u otros mecanismos intracelulares. Los agregados de proteínas objetivo también pueden formarse mediante autoagregación de proteínas objetivo en una suspensión, por ejemplo, en un medio de cultivo celular, en una solución acuosa (por ejemplo, solución acuosa tamponada), en agua técnica o en agua desionizada. Los agregados de proteínas objetivo pueden estar formados por varios mecanismos que pueden incluir interacciones covalentes o no covalentes entre las moléculas de proteína objetivo.

Preferiblemente, los agregados de proteína objetivo comprenden al menos 85%, más preferiblemente al menos 90%, y lo más preferiblemente al menos 95%, o incluso 100%, por ejemplo, al menos 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% o 100% de las mismas proteínas objetivo, que consisten en la agregación de múltiples copias de dichas proteínas objetivo. Fue sorprendente que dichos agregados de proteínas objetivo no solo sean una co-

localización de proteínas objetivo dentro de agregados no específicos, sino que comprendan al menos el 85%, o incluso el 100% de las mismas proteínas objetivo que muestran una preferencia de las proteínas objetivo por autoagregación.

5 El término "proteína objetivo insoluble" o el término "agregado de proteína objetivo insoluble", como se usa en el presente documento, se refiere a una proteína objetivo o a un agregado de proteína objetivo, que no es soluble en una suspensión, por ejemplo, medio de cultivo celular, medio de fermentación, solución acuosa (por ejemplo, solución acuosa tamponada tal como Tris/HCl, pH 7,5), agua técnica o agua desionizada, y que, por lo tanto, puede separarse de las partes solubles de la célula huésped presentes en dicha suspensión, por ejemplo, por centrifugación y/o filtración. Además, el término "proteína objetivo insoluble" o el término "agregado de proteína objetivo insoluble", como se usa en el presente documento, se refiere a una proteína objetivo o a un agregado de proteína objetivo, que tampoco es soluble en una suspensión, por ejemplo, medio de cultivo celular, medio de fermentación, solución acuosa (por ejemplo, solución acuosa tamponada tal como Tris/HCl, pH 7,5), agua técnica o agua desionizada, después de la adición de una solución acuosa que comprende una base (por ejemplo, NaOH 0,05 M) a dicha suspensión, mientras que otras partes de la célula huésped que son insolubles en una suspensión, por ejemplo, medio de cultivo celular, medio de fermentación, solución acuosa (por ejemplo, solución acuosa tamponada tal como Tris/HCl, pH 7,5), agua técnica o agua desionizada, se vuelven solubles después de la adición de una solución acuosa que comprende una base a dicha suspensión y que, por lo tanto, puede, separarse adicionalmente de las partes insolubles de la célula huésped solubilizadas presentes en dicha suspensión, por ejemplo por centrifugación y/o filtración.

20 Se prefiere que la proteína objetivo insoluble o el agregado de proteína objetivo insoluble sea insoluble en una solución acuosa tamponada de Tris/HCl entre 10 y 100 mM, más preferiblemente de Tris/HCl entre 10 y 50 mM, y lo más preferiblemente Tris/HCl entre 10 y 20 mM, por ejemplo, de Tris/HCl 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 100 mM Tris/HCl, en la que dicha solución acuosa tamponada tiene preferiblemente un pH de entre 5,0 y 9,0, más preferiblemente un pH de entre 6,0 y 8,0, y lo más preferiblemente un pH de entre 6,7 y 7,2, por ejemplo Tris/HCl, pH 7,0, pH 7,5 o pH 8,0.

25 Si, por ejemplo, en la etapa a) del método de la presente invención, se proporciona una suspensión de células huésped intactas, por ejemplo, células huésped microbianas, de plantas o de insectos, se agrega una solución acuosa que comprende al menos una base en la etapa b) en una cantidad que sea suficiente/requerida para romper dichas células huésped en partes insolubles de célula huésped y para solubilizar aún más dichas partes insolubles de la célula huésped en tal extensión que al menos se solubilicen 80%, preferiblemente al menos 90%, y más preferiblemente al menos 95%, o incluso 100%, por ejemplo al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9%, o 100%, de dichas partes insolubles de células huésped. Además, si, por ejemplo, en la etapa a) del método de la presente invención, se proporciona una suspensión de células huésped ya rotas, por ejemplo, células huésped microbianas, de plantas o de insectos, es decir, una suspensión que comprende las partes insolubles de las células huésped de las células rotas, se agrega una solución acuosa que comprende al menos una base en la etapa b) en una cantidad que es suficiente/requerida para solubilizar dichas partes insolubles de la célula huésped se extienden de tal manera que al menos se solubilizan el 80%, preferiblemente al menos 90%, y más preferiblemente al menos 95%, o incluso 100%, por ejemplo al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9%, o 100%, de las partes insolubles de la célula huésped.

40 La persona experta puede acceder fácilmente a la cantidad de una solución acuosa que comprende una base que es suficiente/requerida para romper las células huésped, por ejemplo, células huésped microbianas, de plantas o de insectos, y/o para solubilizar partes insolubles de la célula huésped de tal manera que al menos se solubilicen 80%, preferiblemente al menos 90%, y más preferiblemente al menos 95%, o incluso 100%, de dichas partes insolubles de células huésped, por ejemplo, (i) realizando una serie de diluciones de una base seleccionada, por ejemplo, NaOH, (ii) añadiendo la base seleccionada (por ejemplo, NaOH) a varias concentraciones a una suspensión de células huésped intactas o rotas (por ejemplo, NaOH en concentraciones finales de NaOH 0,01 M, NaOH 0,02 M, NaOH 0,03 M, NaOH 0,04 M, NaOH 0,05 M, NaOH 0,06 M, NaOH 0,07 M, NaOH 0,08 M, NaOH 0,09 M, NaOH 0,1 M, NaOH 0,2 M, NaOH 0,3 M, NaOH 0,4 M, NaOH 0,5 M, NaOH 0,6 M, NaOH 0,7 M, NaOH 0,8 M, NaOH 0,9 M, o NaOH 1 M), (iii) separando la proteína objetivo insoluble de las partes insolubles de la célula huésped completa o parcialmente solubilizadas, por ejemplo, por centrifugación o filtración, (iv) evaluando la porción de partes insolubles de la célula huésped dentro de la porción de proteína objetivo insoluble separada, por ejemplo, en el centrifugado o retenido, y (v) determinando la cantidad de la base que es suficiente/requerida para solubilizar al menos 80%, preferiblemente al menos 90%, más preferiblemente al menos 95%, o incluso 100% de las partes insolubles de la célula huésped.

55 El término "la proteína objetivo permanece insoluble", como se usa en el presente documento, se refiere a una proteína objetivo que es insoluble en una suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, y que permanece insoluble, por ejemplo al menos al 80%, preferiblemente al menos al 90%, y más preferiblemente al menos al 95%, o incluso al 100%, por ejemplo, al menos al 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9%, o 100%, en dicha suspensión, por ejemplo, en dicha solución acuosa (por ejemplo, solución acuosa tamponada), medio de cultivo, medio de fermentación, agua técnica o agua desionizada, que comprende una base, por ejemplo durante un cierto período de tiempo y/o a una temperatura específica y, por lo tanto, permite la separación y el aislamiento de dicha proteína objetivo, por ejemplo, por filtración y/o centrifugación.

En una realización preferida de la presente invención, la proteína objetivo es insoluble en una suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, y permanece insoluble, por ejemplo al menos al 80%, preferiblemente al menos al 90%, y más preferiblemente al menos al 95%, o incluso al 100%, por ejemplo, al menos al 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9%, o 100%, en dicha suspensión, por ejemplo, en dicha solución acuosa (por ejemplo, solución acuosa tamponada), medio de cultivo, medio de fermentación, agua técnica o agua desionizada, que comprende una base durante un período de tiempo de al menos 5 minutos, preferiblemente de al menos 10 minutos, más preferiblemente de al menos 90 minutos y lo más preferiblemente de al menos 180 minutos, por ejemplo, durante al menos 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 130, 140, 150, 160, 170 o 180 minutos.

En una realización preferida adicional de la presente invención, la proteína objetivo es insoluble en una suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, y permanece insoluble hasta al menos 80%, preferiblemente al menos 90%, y más preferiblemente hasta al menos 95%, o incluso hasta 100%, por ejemplo, al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9%, o 100%, en dicha suspensión, por ejemplo, en dicha solución acuosa (por ejemplo, solución acuosa tamponada), medio de cultivo, medio de fermentación, agua técnica o agua desionizada, que comprende una base durante un período de tiempo de entre 5 y 180 minutos, preferiblemente de entre 10 y 90 minutos, por ejemplo, durante 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 130, 140, 150, 160, 170 o 180 minutos, y/o a una temperatura de entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C.

En una realización más preferida de la presente invención, la proteína objetivo es insoluble en una suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, y permanece insoluble hasta al menos 90%, preferiblemente hasta al menos 95%, en dicha solución, por ejemplo en dicha solución acuosa, medio de cultivo, medio de fermentación, agua técnica o agua desionizada, que comprende una base durante un período de tiempo de 10 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C, o 40 °C, durante un período de 20 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C, durante un período de tiempo de 30 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C, durante un período de 40 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C en una solución acuosa, durante un período de 50 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C, durante un período de 60 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C, o durante un período de 90 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C.

En una realización más preferida de la presente invención, la proteína objetivo es insoluble en una suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, y permanece insoluble hasta al menos 80%, preferiblemente hasta al menos 90%, y más preferiblemente hasta al menos 95%, o incluso hasta 100%, por ejemplo, al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9%, o 100%, en dicha suspensión, por ejemplo, en dicha solución acuosa (por ejemplo, solución acuosa tamponada), medio de cultivo, medio de fermentación, agua técnica o agua desionizada, que comprende una base, por ejemplo, un hidróxido metálico y/o amoníaco tal como hidróxido de sodio (NaOH), hidróxido de potasio (KOH) y/o hidróxido de calcio (CaOH), cuya concentración final varía de 0,005 M a 1 M, preferiblemente de 0,01 M a 0,6 M, más preferiblemente de 0,02 M a 0,2 M, 0,05 M a 0,15 M o 0,04 M a 0,1 M, y lo más preferiblemente de 0,04 M a 0,06 M, durante un período de tiempo de entre 5 y 180 minutos, preferiblemente de entre 10 y 90 minutos, por ejemplo, más de 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 130, 140, 150, 160, 170 o 180 minutos, y/o a una temperatura de entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C.

Se prefiere particularmente que la proteína objetivo permanezca insoluble hasta al menos 90%, preferiblemente al menos 95%, o incluso hasta 100%, por ejemplo, hasta al menos 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, después de la adición de una solución acuosa que comprende al menos una base en la etapa b) durante un período de tiempo de entre 10 y 40 min y a una temperatura de entre 15 °C y 25 °C. Es particularmente más preferido que la proteína objetivo permanezca insoluble hasta al menos 90%, preferiblemente hasta al menos 95%, o incluso hasta 100%, por ejemplo, hasta al menos 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, después de la adición de una solución acuosa que comprende al menos una base en la etapa b) durante un período de tiempo de entre 20 y 30 minutos y a una temperatura de entre 20 °C y 25 °C.

En una realización preferida adicional de la presente invención, las proteínas objetivo, por ejemplo, las proteínas de seda de araña, que permanecen insolubles, se separan/aíslan de las partes insolubles de la célula huésped solubilizadas usando un filtro, por ejemplo, un filtro que tiene un tamaño de poro de 0,1 μM o 0,22 μM y/o por centrifugación, por ejemplo, de 3000 a 8000 x g durante 20 a 30 minutos.

En otra realización preferida, las proteínas objetivo permanecen insolubles en forma de agregados de proteínas objetivo (por ejemplo, agregados de proteínas de seda tales como agregados de proteínas de seda de araña). Por lo tanto, el término "el agregado de proteína objetivo permanece insoluble" se refiere a una formación de agregado de

5 proteína objetivo que no es reversible o que no es reversible hasta más de 0,1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20% en suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, que comprende una base, por ejemplo, durante un cierto período de tiempo y/o a una temperatura específica y, por lo tanto, permite la separación y el aislamiento de dichos agregados de proteínas objetivo, por ejemplo, por filtración y/o centrifugación.

10 El término "agregado de proteína objetivo permanece insoluble" también se refiere a un agregado de proteína objetivo que permanece insoluble, por ejemplo, hasta al menos 80%, preferiblemente hasta al menos 90%, y más preferiblemente hasta al menos 95%, o incluso hasta 100%, por ejemplo, hasta al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, en suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o agua desionizada, que comprende una base, por ejemplo durante un cierto período de tiempo y/o a una temperatura específica y, por lo tanto, permite la separación y el aislamiento de dichos agregados de proteínas objetivo, por ejemplo, por filtración y/o centrifugación.

15 En una realización más preferida de la presente invención, la formación de agregado de proteína objetivo no es reversible o no es reversible a más de 0,1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20% en una suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, que comprende una base, por ejemplo un hidróxido metálico y/o amoníaco, durante un período de tiempo de entre 5 y 180 minutos, preferiblemente de entre 10 y 90 minutos, por ejemplo, más de 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 130, 140, 150, 160, 170 o 180 minutos, y/o a una temperatura de entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C.

25 En otra realización de la presente invención, el agregado de proteína objetivo permanece insoluble hasta al menos 80%, preferiblemente hasta al menos 90%, y más preferiblemente hasta al menos 95%, o incluso hasta 100%, por ejemplo, hasta al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, en una suspensión, por ejemplo en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, que comprende una base, por ejemplo un hidróxido metálico y/o amoníaco, durante un período de tiempo de entre 5 y 180 minutos, preferiblemente de entre 10 y 90 minutos, por ejemplo, más de 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 130, 140, 150, 160, 170 o 180 minutos, y/o a una temperatura de entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C.

35 En una realización preferida particular de la presente invención, la formación de agregado de proteína objetivo no es reversible hasta más del 10% en una suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o agua desionizada, que comprende una base, por ejemplo un hidróxido metálico y/o amoníaco, durante un período de tiempo de 10 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C, durante un período de 20 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C, durante un período de tiempo de 30 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C, durante un período de 40 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C en una solución acuosa, durante un período de 50 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C, durante un período de 60 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C, o durante un período de 90 minutos a una temperatura de 15 °C, 20 °C, 25 °C, 30 °C, 35 °C o 40 °C.

45 En otra realización de la presente invención, el agregado de proteína objetivo permanece insoluble hasta al menos 80%, preferiblemente hasta al menos 90%, y más preferiblemente hasta al menos 95%, o incluso hasta 100%, por ejemplo, hasta al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, en una suspensión, por ejemplo en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, que comprende una base, por ejemplo, un hidróxido metálico y/o amoníaco como el hidróxido de sodio (NaOH), hidróxido de potasio (KOH) y/o hidróxido de calcio (CaOH), cuya concentración final varía de 0,005 M a 1 M, preferiblemente de 0,01 M a 0,6 M, más preferiblemente de 0,02 M a 0,2 M, 0,05 M a 0,15 M o 0,04 M a 0,1 M, y lo más preferiblemente de 0,04 M a 0,06 M, durante un período de tiempo de entre 5 y 180 minutos, preferiblemente de entre 10 y 90 minutos, por ejemplo, más de 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 130, 140, 150, 160, 170, o 180 minutos, y/o a una temperatura de entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C.

60 En una realización más preferida, la proteína objetivo es una proteína objetivo con dominios repetidos/unidades repetitivas (por ejemplo, una proteína de seda tal como una proteína de seda de araña) y tiene la propiedad de formar agregados de proteínas. Una vez que se forman los agregados de proteínas, dichos agregados de proteínas son preferiblemente insolubles en una suspensión, por ejemplo, en una solución acuosa (por ejemplo, solución acuosa tamponada), en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada, y permanecer insoluble hasta al menos 80%, preferiblemente hasta al menos 90%, y más preferiblemente hasta al

5 menos 95%, o incluso hasta 100%, por ejemplo, hasta al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9%, o 100%, en dicha suspensión, por ejemplo, en dicha solución acuosa (por ejemplo, solución acuosa tamponada), medio de cultivo, medio de fermentación, agua técnica o agua desionizada, que comprende una base, por ejemplo un hidróxido metálico y/o amoníaco, durante un período de tiempo de entre 5 y 180 minutos, preferiblemente de entre 10 y 90 minutos, por ejemplo, más de 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 130, 140, 150, 160, 170 o 180 minutos, y/o a una temperatura de entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C.

10 Los agregados de proteínas objetivo, por ejemplo, los agregados de proteínas de seda de araña, que permanecen insolubles, son visibles como partículas discretas por ejemplo, a simple vista, o usando un microscopio óptico o un microscopio electrónico, y/o pueden ser removidos/separados de las partes insolubles de la célula huésped solubilizadas mediante un filtro, por ejemplo un filtro que tiene un tamaño de poro de 0,1 µM o 0,22 µM.

15 Preferiblemente, la proteína objetivo insoluble forma un agregado de proteína, que comprende al menos 85%, más preferiblemente al menos 90%, y lo más preferiblemente al menos 98%, o incluso 100%, por ejemplo, al menos 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100% de dicha proteína objetivo. El término "agregado de proteína objetivo que comprende al menos 85%, más preferiblemente al menos 90%, y lo más preferiblemente al menos 98%, o incluso 100% de la misma proteína objetivo", como se usa en el contexto de la presente invención, significa que el agregado de proteína objetivo se forma mediante la agregación de múltiples copias/unidades de la misma proteína objetivo y que este agregado comprende al menos 85%, más preferiblemente al menos 90%, y lo más preferiblemente al menos 98%, o incluso 100% de dicha proteína objetivo.

20 Se prefiere que la etapa b) del método de la presente invención se lleve a cabo durante un período de tiempo durante el cual la proteína objetivo permanece insoluble.

25 Se prefiere que la etapa b) del método de la presente invención se lleve a cabo durante un período de tiempo de entre 5 y 180 minutos, preferiblemente de entre 10 y 90 minutos, y lo más preferiblemente de entre 10 y 40 minutos, por ejemplo, más de 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 130, 140, 150, 160, 170 o 180 minutos. Por ejemplo, la etapa b) del método de la presente invención se lleva a cabo durante un período de tiempo de entre 30 y 180 minutos para solubilizar partes insolubles de la célula huésped a partir de un medio de cultivo de células huésped bacterianas intactas, por ejemplo, a una temperatura de entre 20 °C y 25 °C.

30 Preferiblemente, la proteína objetivo insoluble o el agregado de proteína objetivo insoluble aislado/separado con el método de la presente invención tiene una pureza de al menos 50% o 60%, más preferiblemente de al menos 70% u 80%, y lo más preferiblemente de al menos 90%, 95% o incluso de 100%, por ejemplo, de al menos 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%.

35 Una pureza de al menos 50% o 60%, más preferiblemente de al menos 70% u 80%, y lo más preferiblemente de al menos 90%, 95%, o incluso de 100%, por ejemplo, de al menos 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9%, o 100%, de la proteína objetivo insoluble aislada o agregado de proteína objetivo significa preferiblemente que es al menos 50% o 60%, más preferiblemente al menos 70% u 80%, y lo más preferiblemente a al menos 90%, 95% o incluso hasta 100%, por ejemplo, al menos 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, libre de (i) partes insolubles de la célula huésped (por ejemplo, proteínas insolubles de la célula huésped, partes de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de la célula huésped y/o inclusiones citoplasmáticas) y/o partes solubles de la célula huésped (por ejemplo, proteínas solubles de la célula huésped, partes de orgánulos celulares y/o componentes celulares), y más preferiblemente libres de (i) partes insolubles de la célula huésped (por ejemplo, proteínas insolubles de la célula huésped, partes de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de células huésped y/o inclusiones citoplasmáticas) y/o partes solubles de células huésped (por ejemplo, proteínas solubles de la célula huésped, partes de orgánulos celulares y/o componentes celulares) y (ii) residuos de suspensión solubles o insolubles (por ejemplo, impurezas relacionadas con la fermentación/cultivo, tales como minerales y/u oligoelementos).

50 Además, una pureza de al menos 50% o 60%, más preferiblemente de al menos 70% u 80%, y lo más preferiblemente de al menos 90%, 95%, o incluso del 100%, por ejemplo, de al menos 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100% de la proteína objetivo insoluble aislada o agregado de proteína objetivo más preferiblemente significa que es al menos 50% o 60%, más preferiblemente al menos 70% u 80%, y lo más preferiblemente al menos 90%, 95% o incluso hasta 100%, por ejemplo, al menos 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, libre de (i) proteínas insolubles y/o solubles de la célula huésped, más preferiblemente libre de (i) proteínas insolubles y/o solubles de la célula huésped y (ii) partes insolubles de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de células huésped, inclusiones citoplasmáticas, partes solubles de orgánulos celulares y/o componentes celulares solubles, y aún más

preferiblemente libres de (i) proteínas insolubles y/o solubles de la célula huésped, (ii) partes insolubles de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de la célula huésped, inclusiones citoplasmáticas, partes solubles de orgánulos celulares y/o componentes celulares solubles, y (iii) residuos de suspensiones solubles y/o insolubles (por ejemplo, impurezas relacionadas con la fermentación/cultivo tales como minerales y/u oligoelementos).

Como alternativa, una pureza de al menos 50% o 60%, más preferiblemente de al menos 70% u 80%, y lo más preferiblemente de al menos 90%, 95%, o incluso de 100%, por ejemplo, de al menos 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100% de la proteína objetivo insoluble aislada o agregado de proteína objetivo significa preferiblemente que comprende no más del 50% o 40%, más preferiblemente no más del 30% o 20%, y lo más preferiblemente no más del 10%, 5% o 0%, por ejemplo, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1% o 0%, (i) partes insolubles de la célula huésped (por ejemplo, proteínas insolubles de la célula huésped, partes de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de la célula huésped y/o inclusiones citoplasmáticas) y/o partes solubles de la célula huésped (por ejemplo, proteínas solubles de la célula huésped, partes de orgánulos celulares y/o componentes celulares), y más preferiblemente (i) partes insolubles de la célula huésped (por ejemplo, proteínas insolubles de la célula huésped, partes de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de la célula huésped y/o inclusiones citoplasmáticas) y/o partes solubles de la célula huésped (por ejemplo, proteínas solubles de la célula huésped, partes de orgánulos celulares y/o componentes de la célula) y (ii) residuos de suspensión solubles y/o insolubles (por ejemplo, impurezas relacionadas con la fermentación/cultivo tales como minerales y/u oligoelementos).

Además, una pureza de al menos 50% o 60%, más preferiblemente de al menos 70% u 80%, y lo más preferiblemente de al menos 90%, 95%, o incluso de 100%, por ejemplo, de al menos 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100% de la proteína objetivo insoluble aislada o agregado de proteína objetivo más preferiblemente significa que comprende no más del 50% o 40%, más preferiblemente no más del 30% o 20%, y lo más preferiblemente no más del 10%, 5% o 0%, por ejemplo, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1% o 0%, (i) proteínas insolubles y/o solubles de la célula huésped, lo más preferiblemente (i) proteínas insolubles y/o solubles de la célula huésped y (ii) partes insolubles de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de la célula huésped, inclusiones citoplasmáticas, partes solubles de orgánulos celulares y/o componentes celulares solubles, e incluso lo más preferiblemente (i) proteínas insolubles y/o solubles de la célula huésped, (ii) partes insolubles de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de la célula huésped, inclusiones citoplasmáticas, partes solubles de orgánulos celulares y/o componentes celulares solubles, y (iii) residuos de suspensión solubles y/o insolubles (por ejemplo, impurezas relacionadas con la fermentación/cultivo, tal como minerales y/u oligoelementos).

El experto en la materia conoce técnicas para determinar la pureza de la proteína objetivo insoluble o del agregado de proteína objetivo aislado con el método de la presente invención.

La pureza de la proteína objetivo insoluble aislada o agregado de proteína objetivo se mide preferiblemente usando (i) espectrometría, preferiblemente espectrometría de masas (MS), (ii) cromatografía, preferiblemente cromatografía líquida (LC), más preferiblemente cromatografía líquida de alto rendimiento (HPLC), (iii) electroforesis en gel, preferiblemente electroforesis en gel de SDS, (iv) transferencia Western/Inmunotransferencia, o (v) combinaciones de los mismos.

Se prefiere que la cromatografía, preferiblemente la cromatografía líquida (LC), más preferiblemente la cromatografía líquida de alto rendimiento (HPLC), se combine con espectrometría, preferiblemente espectrometría de masas (MS). Por consiguiente, la pureza de la proteína objetivo insoluble aislada o agregado de proteína objetivo se mide preferiblemente usando cromatografía líquida-espectrometría de masas (LC-MS) y más preferiblemente se mide usando cromatografía líquida de alto rendimiento-espectrometría de masas (HPLC-MS).

Preferiblemente, la pureza de la proteína objetivo insoluble o agregado de proteína objetivo aislado con el método de la presente invención se calcula sobre una base de peso seco, por ejemplo, expresado como % en peso seco (p/p).

Se prefiere particularmente que el peso seco de la proteína objetivo insoluble aislada o agregado de proteína objetivo se calcule en relación con el peso seco de las (i) partes insolubles y/o solubles de la célula huésped, más preferiblemente (i) partes insolubles y/o solubles de la célula huésped y (ii) residuos solubles y/o insolubles en suspensión (véase más arriba).

Se prefiere particularmente que el peso seco de la proteína objetivo insoluble aislada o agregado de proteína objetivo se calcule en relación con el peso seco de las (i) proteínas insolubles y/o solubles de la célula huésped, más preferiblemente (i) las proteínas insolubles y/o solubles de la célula huésped y (ii) las partes insolubles de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de la célula huésped, inclusiones citoplasmáticas, partes solubles de orgánulos celulares y/o componentes celulares solubles, y más preferiblemente (i)

las proteínas insolubles y/o solubles de la célula huésped, (ii) las partes insolubles de la pared celular, partes de la membrana celular, partes del citoesqueleto, residuos de la célula huésped, inclusiones citoplasmáticas, partes solubles de orgánulos celulares y/o componentes celulares solubles, y (iii) los residuos solubles y/o insolubles en suspensión (por ejemplo, impurezas relacionadas con la fermentación/cultivo, tales como minerales y/u oligoelementos).

- 5 Se prefiere que la solución acuosa que comprende una base en la etapa b) tenga un pH de entre 7 y 14, más preferiblemente de entre 9 y 13, y lo más preferiblemente de entre 10 y 12. Por ejemplo, la solución acuosa que comprende una base en la etapa b) tiene un pH de 7, 7,5; 8, 8,5; 9, 9,5; 10, 10,5; 11, 11,5; 12, 12,5; 13, 13,5 o 14.

- 10 En una realización preferida de la presente invención, la base es un hidróxido metálico o amoníaco. En otra realización preferida de la presente invención, se usan un hidróxido metálico y amoníaco. Preferiblemente, el hidróxido metálico se selecciona del grupo que consta de hidróxido de sodio (NaOH), hidróxido de potasio (KOH) e hidróxido de calcio (CaOH), o es una combinación de los mismos. Lo más preferiblemente, la base es hidróxido de sodio (NaOH).

- 15 Preferiblemente, la concentración final de la base en la etapa b) varía de 0,005 M a 1 M, preferiblemente de 0,01 M a 0,6 M, más preferiblemente de 0,02 M a 0,2 M, 0,05 M a 0,15 M, o 0,04 M a 0,1 M, y lo más preferiblemente de 0,04 M a 0,06 M. Por ejemplo, la concentración final de la base en la etapa b) del método de la presente invención es 0,005 M, 0,01 M, 0,015 M, 0,02 M, 0,025 M, 0,03 M, 0,035 M, 0,04 M, 0,045 M, 0,05 M, 0,055 M, 0,06 M, 0,065 M, 0,07 M, 0,075 M, 0,08 M, 0,085 M, 0,09 M, 0,095 M, 0,1 M, 0,15 M, 0,2 M, 0,25 M, 0,3 M, 0,35 M, 0,4 M, 0,45 M, 0,5 M, 0,55 M, 0,6 M, 0,65 M, 0,7 M, 0,75 M, 0,8 M, 0,85 M, 0,9 M, 0,95 M, o 1 M.

- 20 En las realizaciones preferidas del método de la presente invención, la concentración final de hidróxido de sodio (NaOH), hidróxido de potasio (KOH) y/o hidróxido de calcio (CaOH) en la etapa b) varía de 0,005 M a 1 M, preferiblemente de 0,01 M a 0,6 M, más preferiblemente de 0,02 M a 0,2 M, 0,05 M a 0,15 M, o 0,04 M a 0,1 M, y lo más preferiblemente de 0,04 M a 0,06 M. Por ejemplo, la concentración final de hidróxido de sodio (NaOH), hidróxido de potasio (KOH) y/o hidróxido de calcio (CaOH) en la etapa b) de las realizaciones preferidas del método de la presente invención es 0,005 M, 0,01 M, 0,015 M, 0,02 M, 0,025 M, 0,03 M, 0,035 M, 0,04 M, 0,045 M, 0,05 M, 0,055 M, 0,06 M, 0,065 M, 0,07 M, 0,075 M, 0,08 M, 0,085 M, 0,09 M, 0,095 M, 0,1 M, 0,15 M, 0,2 M, 0,25 M, 0,3 M, 0,35 M, 0,4 M, 0,45 M, 0,5 M, 0,55 M, 0,6 M, 0,65 M, 0,7 M, 0,75 M, 0,8 M, 0,85 M, 0,9 M, 0,95 M, o 1 M.

- 25 En realizaciones más preferidas del método de la presente invención, la base en la etapa b) es hidróxido de sodio (NaOH) y la concentración final de hidróxido de sodio (NaOH) en la etapa b) varía de 0,005 M a 1 M, preferiblemente de 0,01 M a 0,6 M, más preferiblemente de 0,02 M a 0,2 M, 0,05 M a 0,15 M, o 0,04 M a 0,1 M, y lo más preferiblemente de 0,04 M a 0,06 M. Por ejemplo, la concentración final de hidróxido de sodio (NaOH) en la etapa b) de las realizaciones preferidas del método de la presente invención es 0,005 M, 0,01 M, 0,015 M, 0,02 M, 0,025 M, 0,03 M, 0,035 M, 0,04 M, 0,045 M, 0,05 M, 0,055 M, 0,06 M, 0,065 M, 0,07 M, 0,075 M, 0,08 M, 0,085 M, 0,09 M, 0,095 M, 0,1 M, 0,15 M, 0,2 M, 0,25 M, 0,3 M, 0,35 M, 0,4 M, 0,45 M, 0,5 M, 0,55 M, 0,6 M, 0,65 M, 0,7 M, 0,75 M, 0,8 M, 0,85 M, 0,9 M, 0,95 M, o 1 M.

- 30 Preferiblemente, la proporción de una suspensión de células huésped intactas o rotas, por ejemplo, un medio de cultivo celular, un medio de fermentación, una solución acuosa (por ejemplo, solución acuosa tamponada) o una suspensión que tiene un contenido de humedad de entre 5 a 20%, por ejemplo 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20%, es decir, un sedimento celular, a una solución acuosa que comprende una base que oscila entre 3:1 y 1:20, más preferiblemente entre 1:1 y 1:10, y lo más preferiblemente entre 1:2 y 1:4, por ejemplo 3:1, 2:1, 1:1, 1:2, 1:3, 1:4, 1:5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20.

- 35 Se prefiere particularmente que la etapa b) de la presente invención se lleve a cabo con una solución acuosa que comprende hidróxido de sodio (NaOH) como base a una concentración final de entre 0,05 M a 0,15 M, preferiblemente a una concentración final de 0,05 M, durante un período de tiempo de entre 10 y 30 minutos, preferiblemente durante un período de tiempo de entre 20 y 30 minutos y a una temperatura de entre 20 y 25 °C, preferiblemente a una temperatura de 20 °C.

- 40 Se prefiere que, para mejorar el rendimiento, la pureza y/o la capacidad de procesamiento adicional de la proteína objetivo insoluble aislada, el método comprenda además la adición de al menos un reactivo (i) antes de la etapa b), (ii) en la etapa b) y/o (iii) posterior a la etapa b), por ejemplo, antes de la etapa b), en la etapa b), después de la etapa b), antes de la etapa b) y en la etapa b), antes de la etapa b) y después de la etapa b), en la etapa b) y después de la etapa b), o antes de la etapa b), en la etapa b) y posterior a la etapa b).

- 45 Preferiblemente, el reactivo se selecciona del grupo que consta de un agente desnaturalizante, un agente cosmotrópico y un detergente, o es una combinación de los mismos.

Por consiguiente, en una realización preferida de la presente invención, se agrega un agente desnaturalizante, un agente cosmotrópico y/o un detergente (i) antes de la etapa b), (ii) en la etapa b) y/o (iii) posterior a la etapa b).

- 50 El agente desnaturalizante mejora la solubilización de las partes insolubles de la célula huésped, mientras que el detergente ayuda a solubilizar las paredes celulares y/o las membranas celulares de las células huésped, por ejemplo, células huésped microbianas, células huésped de plantas o células huésped de insectos. Además, el agente cosmotrópico ayuda a estabilizar los agregados de proteínas objetivo.

Por lo tanto, se prefiere agregar a una suspensión de células huésped intactas, por ejemplo, células huésped bacterianas o de plantas, antes de la etapa b), en la etapa b) y/o posterior a la etapa b) un detergente que ayuda a solubilizar las paredes celulares y/o las membranas celulares, preferiblemente en combinación con un agente desnaturizante que mejora aún más la solubilización de las partes restantes insolubles de la célula huésped. Por lo tanto, también se prefiere agregar a una suspensión de células huésped ya rotas, por ejemplo, células huésped bacterianas o de plantas, antes de la etapa b), en la etapa b) y/o posterior a la etapa b) un agente desnaturizante que mejora la solubilización de las partes restantes insolubles de la célula huésped, preferiblemente en combinación con un agente cosmotrópico que estabiliza los agregados de proteínas objetivo insolubles.

Los inventores de la presente invención han descubierto sorprendentemente que la adición de un agente desnaturizante, un agente cosmotrópico y/o detergente antes de la etapa b), en la etapa b) y/o posterior a la etapa b) puede mejorar la pureza de la proteína objetivo insoluble aislada de hasta 5 a 20%, por ejemplo, hasta 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20% y/o el rendimiento de la proteína objetivo insoluble aislada de hasta 5 a 20%, por ejemplo, hasta 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20%. Por ejemplo, una proteína objetivo insoluble aislada que tiene después de la etapa de separación c) una pureza del 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94 o 95% sin la adición de un agente desnaturizante, un agente cosmotrópico y/o un detergente antes de la etapa b), en la etapa b) y/o posterior a la etapa b), puede tener una pureza del 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o 100% después de la adición de un agente desnaturizante, un agente cosmotrópico y/o detergente antes de la etapa b), en la etapa b) y/o posterior a la etapa b).

Preferiblemente, (i) el agente desnaturizante es urea ($\text{CH}_4\text{N}_2\text{O}$) o clorhidrato de guanidinio ($\text{CH}_5\text{N}_3\text{HCl}$), (ii) el agente cosmotrópico es una sal de fosfato o una sal de sulfato, o (iii) el detergente es Tween 20, Tween 60, Tween 80, Triton X-15, Triton X-45, Triton X-100, Triton X-102, Triton X-114, Triton X-151, Triton X-165, Triton X-200, Triton X-207, Triton X-301, Triton X-305, Triton X-405, Triton X-705, SDS o Brij. El experto en la materia puede seleccionar otros agentes desnaturizantes, agentes cosmotrópicos o detergentes adecuados.

Preferiblemente, la sal de fosfato es fosfato de amonio, fosfato de potasio o fosfato de sodio. La sal de fosfato puede reducir la solubilidad de la proteína objetivo o los agregados de proteína objetivo en una suspensión de células huésped intactas o rotas, por ejemplo, en una solución acuosa, medio de fermentación, medio de cultivo, agua técnica o agua desionizada. Preferiblemente, la sal de sulfato es sulfato de amonio.

Se prefiere que la concentración final de urea varíe de 0,1 M a 10 M, preferiblemente de 1 M a 5 M, y más preferiblemente de 4 a 5 M. Por ejemplo, la concentración final de urea es 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 1,5, 2, 2,5, 3, 3,5, 4, 4,5, 5, 5,5, 6, 6,5, 7, 7,5, 8, 8,5, 9, 9,5, o 10 M. También se prefiere que la concentración final de clorhidrato de guanidinio varíe de 0,01 M a 3 M, preferiblemente de 0,1 M a 2 M. Por ejemplo, la concentración final de clorhidrato de guanidinio es 0,01, 0,02, 0,03, 0,04, 0,05, 0,06, 0,07, 0,08, 0,09, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9, 2, 2,1, 2,2, 2,3, 2,4, 2,5, 2,6, 2,7, 2,8, 2,9 o 3 M.

En una realización preferida de la presente invención, la proteína objetivo o agregado de proteína objetivo permanece insoluble al menos al 85%, preferiblemente al menos al 90%, y más preferiblemente al menos al 95%, o incluso al 100% en un solución acuosa que comprende hidróxido de sodio (NaOH) (por ejemplo, a una concentración final que varía de 0,005 M a 0,5 M) e hidrocloreuro de guanidinio ($\text{CH}_5\text{N}_3\text{HCl}$) (por ejemplo, a una concentración final que varía de 0,01 M a 3 M) o urea (por ejemplo, a una concentración final de urea que varía de 0,1 M a 10 M), durante un período de tiempo de entre 20 a 40 min y a una temperatura de entre 20 °C y 25 °C.

En otra realización preferida de la presente invención, la proteína objetivo, por ejemplo, la proteína de seda de araña tal como la proteína de seda de araña C_{16} , o el agregado de proteína objetivo permanece insoluble hasta al menos 85%, preferiblemente hasta al menos 90%, y más preferiblemente hasta al menos 95%, o incluso hasta 100% en medio acuoso que incluye hidróxido de sodio (NaOH) (por ejemplo, a una concentración final que varía de 0,005 M a 0,2 M) y clorhidrato de guanidinio ($\text{CH}_5\text{N}_3\text{HCl}$) (por ejemplo, a una concentración final que varía de 0,1 M a 2 M) o urea (por ejemplo, a una concentración final que varía de 1 M a 5 M), durante un período de tiempo de entre 20 y 40 min y a una temperatura de entre 20 °C y 25 °C.

En una realización preferida adicional de la presente invención, la proteína objetivo, por ejemplo, la proteína de seda de araña tal como la proteína de seda de araña C_{16} , o el agregado de proteína objetivo permanece insoluble hasta al menos 85%, preferiblemente hasta al menos 90%, y más preferiblemente hasta al menos 95%, o incluso hasta 100% en un medio acuoso que comprende hidróxido de sodio (NaOH) a una concentración final de 0,5 M durante un período de tiempo de 30 min y a una temperatura de entre 20 y 25 °C.

Se prefiere que la separación de la proteína objetivo insoluble o el agregado de proteína objetivo insoluble de las partes insolubles de la célula huésped solubilizadas en la etapa c) se logre por centrifugación. La etapa de centrifugación permite la separación de la fase sólida que comprende la proteína objetivo insoluble o el agregado de proteína objetivo insoluble de la fase líquida, por ejemplo, solución acuosa (por ejemplo, solución tamponada acuosa), medio de cultivo, medio de fermentación, agua técnica o agua desionizada, que comprende las partes insolubles de la célula huésped solubilizadas (por ejemplo, proteínas de la célula huésped, residuos de células huésped) y opcionalmente los residuos insolubles solubilizados (por ejemplo, residuos de fermentación), partes solubles de la célula huésped y/u otras impurezas solubles. Después de la etapa de centrifugación, la proteína objetivo insoluble o

- el agregado de proteína objetivo insoluble está presente como precipitado, mientras que el sobrenadante contiene las partes insolubles de la célula huésped solubilizadas y opcionalmente los residuos insolubles solubilizados (por ejemplo, residuos de fermentación), partes solubles de la célula huésped y/u otras impurezas solubles y se pueden desechar. Por lo tanto, debe ser claro para la persona experta que durante esta etapa de separación, la proteína objetivo insoluble se separa de las partes insolubles de la célula huésped solubilizadas y opcionalmente de las partes solubles de la célula huésped que están presentes o que todavía están presentes en la fase líquida, por ejemplo, solución acuosa o medio de cultivo.
- También se prefiere que la separación de la proteína objetivo insoluble o el agregado de proteína objetivo insoluble de las partes insolubles de la célula huésped solubilizadas en la etapa c) se logre por filtración. En este método de separación, la proteína objetivo insoluble se concentra pasando la fase líquida, por ejemplo, solución acuosa (por ejemplo, solución tamponada acuosa), medio de cultivo, medio de fermentación, agua técnica o agua desionizada, que comprende la proteína objetivo insoluble, las partes insolubles de la célula huésped solubilizadas y los residuos insolubles opcionalmente solubilizados (por ejemplo, residuos de fermentación), partes solubles de células huésped y/u otras impurezas solubles a través de una membrana filtrante que retiene la proteína objetivo insoluble y que deja pasar las partes insolubles de la célula huésped solubilizadas y opcionalmente los residuos insolubles solubilizados (por ejemplo, residuos de fermentación), partes solubles de la célula huésped y/u otras impurezas solubles. Por lo tanto, debe ser claro para la persona experta que durante esta etapa de separación, la proteína objetivo insoluble se separa de las partes insolubles de la célula huésped solubilizadas y opcionalmente de las partes solubles de la célula huésped que están presentes o que todavía están presentes en la fase líquida, por ejemplo, solución acuosa o medio de cultivo, que se pasa a través de la membrana del filtro. En realizaciones preferidas del método de la presente invención, el movimiento de la fase líquida, por ejemplo, la solución acuosa (por ejemplo, solución acuosa tamponada), el medio de cultivo, el medio de fermentación, el agua técnica o el agua desionizada, que comprende la proteína objetivo insoluble y las partes insolubles de la célula huésped solubilizadas, se acelera por presión de fuerza gravitacional, vacío y/o centrifugación.
- Se prefiere además que la separación de la proteína objetivo insoluble de las partes insolubles de la célula huésped solubilizadas en la etapa c) se logre por sedimentación. En este método de separación, la proteína objetivo insoluble o el agregado de proteína objetivo insoluble se establece fuera de la fase líquida, por ejemplo, la solución acuosa (por ejemplo, solución acuosa tamponada), el medio de cultivo, el medio de fermentación, el agua técnica o el agua desionizada, en la que está comprendida y se apoyan contra una pared. Esto se debe al movimiento de la proteína objetivo insoluble o del agregado de proteína objetivo insoluble a través del líquido en respuesta a las fuerzas que actúan sobre él, por ejemplo, la gravedad. Las partes insolubles de la célula huésped solubilizadas y opcionalmente los residuos insolubles solubilizados (por ejemplo, los residuos de fermentación), las partes solubles de la célula huésped y/u otras impurezas solubles retenidas en la solución, forman el sobrenadante y, por lo tanto, simplemente se pueden desechar.
- Se prefiere particularmente que la separación de la proteína objetivo insoluble de las partes insolubles de la célula huésped solubilizadas en la etapa c) se logre mediante una combinación de centrifugación, sedimentación y/o filtración, por ejemplo, centrifugación y filtración, sedimentación y filtración, sedimentación y centrifugación, o centrifugación, sedimentación y filtración.
- Dependiendo de las diferentes proteínas objetivo insolubles, el parámetro para la centrifugación, sedimentación o filtración puede variar. El experto en la materia puede adaptar fácilmente los parámetros de separación apropiados, por ejemplo, la fuerza de aceleración/fuerza G y/o el tiempo usando la centrifugación para la separación, el tamaño del filtro usando la filtración para la separación y/o el tiempo de sedimentación usando la sedimentación para la separación, con el fin de separar/aislar las partes insolubles de la célula huésped solubilizadas de la proteína objetivo insoluble.
- Preferiblemente, la centrifugación se realiza a tal fuerza de aceleración/fuerza G y/o durante dicho tiempo de centrifugación que es suficiente para permitir la sedimentación de al menos 80%, más preferiblemente de al menos 90%, y la mayoría preferiblemente de al menos 95%, o incluso de 100%, por ejemplo de al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, de la proteína objetivo insoluble o del agregado de proteína objetivo insoluble, mientras que las partes insolubles de la célula huésped solubilizadas o la mayoría de las partes insolubles de la célula huésped solubilizadas permanecen en solución. En una realización preferida del método de la presente invención, la centrifugación tiene lugar entre 3000 x g y 12000 x g, por ejemplo, a 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10000, 10500, 11000, 11500 o 12000 x g, y/o durante 10 a 40 min, por ejemplo, para 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40 min. Estos parámetros de centrifugación permiten la sedimentación de al menos 80%, preferiblemente de al menos 90%, y más preferiblemente de al menos 95%, o incluso del 100%, por ejemplo, de al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, de la proteína objetivo insoluble o del agregado de proteína objetivo insoluble, mientras que las partes insolubles de la célula huésped solubilizadas o la mayoría de las partes insolubles de la célula huésped solubilizadas permanecen en solución.
- Lo mismo se aplica para la sedimentación o filtración. Preferiblemente, la sedimentación se realiza durante un tiempo de sedimentación tal que sea suficiente para permitir la sedimentación de al menos 80%, más preferiblemente de al

menos 90%, y lo más preferiblemente de al menos 95%, o incluso de 100%, por ejemplo, de al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, de la proteína objetivo insoluble o del agregado de proteína objetivo insoluble, mientras que las partes insolubles de la célula huésped solubilizadas o la mayoría de las partes insolubles de la célula huésped solubilizadas permanecen en solución y pueden descartarse fácilmente. En una realización preferida del método de la presente invención, la sedimentación se realiza durante un período de tiempo de entre 1 y 24 horas, preferiblemente de entre 3 y 15 horas, más preferiblemente de entre 5 y 10 horas, por ejemplo, durante 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 o 24 horas. El tiempo de sedimentación mencionado anteriormente permite la sedimentación de al menos 80%, preferiblemente de al menos 90%, y más preferiblemente de al menos 95%, o incluso de 100%, por ejemplo, de al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, de la proteína objetivo insoluble o del agregado de proteína objetivo insoluble, mientras que las partes insolubles de la célula huésped solubilizadas o la mayoría de las partes insolubles de la célula huésped solubilizadas permanecen en solución.

Preferiblemente, la filtración se realiza con un tamaño de poro de filtro tal que sea suficiente para permitir la retención de al menos 80%, más preferiblemente de al menos 90%, y lo más preferiblemente de al menos 95%, o incluso de 100%, por ejemplo, de al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, de la proteína objetivo insoluble o del agregado de proteína objetivo insoluble, mientras que las partes insolubles de la célula huésped solubilizadas o la mayoría de las partes insolubles de la célula huésped solubilizadas pueden pasar a través de los poros del filtro. Por lo tanto, en caso de que la filtración se use como método de separación, el tamaño de poro del filtro puede ser ligeramente más pequeño que la proteína objetivo insoluble o el agregado de proteína objetivo insoluble. En una realización preferida del método de la presente invención, la filtración se realiza con filtros que tienen un tamaño de poro de entre 10 nm y 200 nm, preferiblemente de entre 20 nm y 180 nm, y más preferiblemente de entre 50 y 100 nm, por ejemplo, de 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190 o 200 nm. Estos parámetros de tamaño de poro del filtro permiten la retención de al menos 80%, preferiblemente de al menos 90%, y más preferiblemente de al menos 95%, o incluso de 100%, por ejemplo, de al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o 100%, de la proteína objetivo insoluble o del agregado de proteína objetivo insoluble, mientras que las partes insolubles de la célula huésped solubilizadas o la mayoría de las partes insolubles de la célula huésped solubilizadas pueden pasar a través de los poros del filtro.

Preferiblemente, para mejorar la mezcla de partes insolubles de la célula huésped ya solubilizadas y/o solubilizadas de forma incompleta en la suspensión y/o para mejorar la formación de fragmentos más pequeños a partir de partes insolubles de la célula huésped ya solubilizadas y/o solubilizadas de forma incompleta, para facilitar la separación de la proteína objetivo insoluble o el agregado de proteína objetivo insoluble de las partes insolubles de la célula huésped solubilizadas y/o solubilizadas de forma incompleta en la etapa c), el método comprende adicionalmente después de la etapa b) y antes de la etapa c) una etapa b') de homogeneización de dicha suspensión, preferiblemente con un homogeneizador (por ejemplo, de acuerdo con el método descrito en el ejemplo 1), o lisis celular, preferiblemente usando enzimas adecuadas. Alternativamente, la homogeneización de dicha suspensión se realiza mediante un molino de bolas o sonicación.

Se prefiere que, para eliminar impurezas adicionales tales como residuos celulares, residuos de fermentación, base, desnaturante, detergente y/o sales, el método de la presente invención comprende además, después de la etapa c) una etapa d) de lavar la proteína objetivo insoluble separada o el agregado de proteína objetivo insoluble, preferiblemente el centrifugado, sedimento o material retenido. Se prefiere más que el método de la presente invención comprenda adicionalmente después de la etapa c) una etapa d) de lavar la proteína objetivo insoluble separada o el agregado de proteína objetivo insoluble, preferiblemente el centrifugado, sedimento o retenido, con una solución acuosa, una solución orgánica y/o urea.

Esta etapa de purificación adicional puede ser requerida en casos en los que se proporciona un medio de cultivo que comprende células huésped intactas que comprenden la proteína objetivo insoluble de interés en la etapa a) del método de la presente invención. Esta etapa puede no ser necesaria en los casos en que se proporciona un sedimento de la célula huésped o un sedimento de la célula huésped resuspendido en agua o en una solución acuosa en la etapa a) del método de la presente invención. En los últimos casos, las células huésped ya están separadas o principalmente separadas del medio de cultivo o del medio de fermentación y, por lo tanto, de todos o casi todos los residuos de fermentación y/o impurezas.

Preferiblemente, la solución acuosa es una solución acuosa tamponada. Además, preferiblemente, la solución orgánica contiene etanol (EtOH), metanol (CH₃OH), hexano (C₆H₁₄), éter dietílico (C₄H₁₀O) y/o isopropanol (C₃H₈O).

Como se mencionó anteriormente, en una realización más preferida del método de la presente invención, la proteína objetivo insoluble separada o el agregado de proteína objetivo insoluble, preferiblemente el centrifugado, sedimento o material retenido, es posterior a la etapa c) lavado en la etapa d) con urea. El uso de urea permite una mayor eliminación de las proteínas de la célula huésped potencialmente restantes que no se han separado en la etapa c).

Los inventores de la presente invención han descubierto sorprendentemente que la etapa d) de lavado posterior puede mejorar la pureza de la proteína objetivo insoluble aislada de hasta 5 a 20%, por ejemplo, hasta 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20%. Por ejemplo, una proteína objetivo insoluble aislada que tiene después de la

etapa c) de separación una pureza de 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94 o 95 % sin una etapa d) de lavado posterior, puede tener una pureza de 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o 100% después de una subsiguiente etapa d) de lavado.

5 Preferiblemente, el método de la presente invención se lleva a cabo a una temperatura entre 4 °C y 60 °C, más preferiblemente entre 15 °C y 40 °C, y lo más preferiblemente entre 15 °C y 25 °C, o de entre 20 °C y 25 °C, por ejemplo, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59 o 60 °C.

10 Preferiblemente, el método de la presente invención se lleva a cabo a una presión de entre 10 kPa y 1000 kPa, preferiblemente de entre 50 kPa y 150 kPa, por ejemplo, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 o 1000 kPa.

Se prefiere que el método para aislar una proteína objetivo insoluble (por ejemplo, una proteína de seda de araña) de un medio de cultivo, sedimento celular o solución tamponada de células huésped bacterianas intactas comprenda las etapas de:

15 a) proporcionar (i) un medio de cultivo de células huésped bacterianas intactas que comprende una proteína objetivo insoluble (por ejemplo, una proteína de seda de araña) y partes insolubles de la célula huésped bacteriana, (ii) un sedimento celular de células huésped bacterianas intactas que comprende una proteína objetivo insoluble (por ejemplo, una proteína de seda de araña) y partes insolubles de la célula huésped bacteriana, que preferiblemente tienen un contenido de humedad residual del 20%, o (iii) una solución tamponada en la que dicho sedimento celular se resuspende,

20 b) agregar NaOH a una concentración final de entre 0,05 M a 0,15 M, preferiblemente de 0,05 M, a dicho (i) medio de cultivo, (ii) sedimento celular o (iii) solución tamponada, para romper dichas células huésped bacterianas y para solubilizar dichas partes insolubles de la célula huésped bacteriana, y

25 c) separar la proteína objetivo insoluble (por ejemplo, una proteína de seda de araña) de las partes insolubles de la célula huésped bacteriana solubilizadas mediante centrifugación entre 3000 xg y 12000 xg, es decir, la proteína objetivo insoluble está presente como precipitado y las partes insolubles de la célula huésped bacteriana solubilizadas están en el sobrenadante que puede descartarse,

en el que en la etapa b) la proteína objetivo permanece insoluble, y en el que al menos el 80% o el 85%, preferiblemente el 95%, de las partes insolubles de la célula huésped bacteriana se solubilizan.

30 También se prefiere que el método para aislar una proteína objetivo insoluble (por ejemplo, una proteína de seda de araña) de un medio de cultivo, sedimento celular o solución tamponada de células huésped bacterianas rotas comprenda las etapas de:

35 a) proporcionar (i) un medio de cultivo de células huésped bacterianas rotas que comprende una proteína objetivo insoluble (por ejemplo, una proteína de seda de araña) y partes insolubles de la célula huésped bacteriana, (ii) un sedimento celular de células huésped bacterianas rotas que comprende una proteína objetivo insoluble (por ejemplo, una proteína de seda de araña) y partes insolubles de la célula huésped bacteriana, que preferiblemente tienen un contenido de humedad residual del 15%, o (iii) una solución tamponada en la que dicho sedimento celular se resuspende,

40 b) agregar NaOH a una concentración final de entre 0,05 M a 0,15 M, preferiblemente de 0,15 M, a dicho (i) medio de cultivo, (ii) sedimento celular o (iii) solución tamponada para solubilizar dichas partes insolubles de la célula huésped bacteriana y

45 c) separar la proteína objetivo insoluble (por ejemplo, una proteína de seda de araña) de las partes insolubles de la célula huésped bacteriana solubilizadas mediante centrifugación entre 3000 xg y 12000 xg, es decir, la proteína objetivo insoluble está presente como precipitado y las partes insolubles de la célula huésped bacteriana solubilizadas están en el sobrenadante que puede descartarse,

en el que en la etapa b) la proteína objetivo permanece insoluble, y en el que al menos el 80% o el 85%, preferiblemente el 95%, de las partes insolubles de la célula huésped bacteriana se solubilizan.

50 Los términos "polipéptido" y "proteína" se usan indistintamente en el presente documento y significan cualquier cadena de aminoácidos unida a péptido, independientemente de la longitud o la modificación postraduccional.

55 Cualquier proteína objetivo insoluble, preferiblemente agregado de proteína objetivo insoluble, como se definió anteriormente, puede aislarse con el método de la presente invención. Sin embargo, se prefiere que la proteína objetivo insoluble, preferiblemente el agregado de proteína objetivo insoluble, se seleccione del grupo que consta de una proteína de seda, particularmente una proteína de seda que comprende al menos dos unidades repetitivas idénticas (por ejemplo, una proteína de seda de artrópodos tal como una proteína de seda de araña o una proteína de seda de insecto), colágeno, resilina y queratina. Preferiblemente, el colágeno es una proteína del biso del mejillón o colágeno humano. Preferiblemente, la queratina es queratina humana.

60 En el contexto de la presente invención, el término "proteína de seda", "colágeno", "resilina" o "queratina" puede

referirse a una proteína tal como una proteína recombinante que se expresa en una célula huésped (por ejemplo, célula huésped microbiana, insecto o de planta) tal como un sistema de expresión o célula huésped recombinante (por ejemplo, un sistema de expresión microbiano, de insecto o de planta) tal como un sistema de expresión recombinante, es decir, separado de su medio natural. Además, en el contexto de la presente invención, el término "proteína de seda aislada", "colágeno aislado", "resilina aislada" o "queratina aislada" puede referirse a una proteína tal como una proteína recombinante que se expresa en una célula huésped (por ejemplo, célula huésped microbiana, de insecto o de planta), tal como una célula huésped recombinante o sistema de expresión (por ejemplo, sistema de expresión microbiano, de insecto o de planta), tal como un sistema de expresión recombinante, es decir, separado de su medio natural y que está aislado de dicha célula huésped o dicho sistema de expresión. Los términos mencionados anteriormente "sistema de expresión" y "célula huésped" se usan indistintamente en el presente documento.

Una "proteína de seda", como se usa en el contexto de la presente invención, se refiere además a una proteína con una secuencia de aminoácidos que comprende o consta de al menos 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, preferiblemente al menos 95% y lo más preferiblemente 100% de copias múltiples de una unidad repetitiva idéntica (por ejemplo, A₂, Q₆ o C₁₆, en la que los ítems 2, 6 o 16 representan el número de unidades repetitivas) o copias múltiples de dos o más unidades repetitivas diferentes (por ejemplo, (AQ)₂₄ o (AQ)₁₂C₁₆).

Los términos "unidad repetitiva" y "unidad repetida" se pueden usar de manera intercambiable en el contexto de la presente invención.

En el contexto de la presente invención, el término "proteína de seda" también se refiere a una proteína de seda que comprende o consta de al menos dos unidades repetitivas idénticas que comprenden o consisten en copias idénticas de secuencias de aminoácidos de polipéptidos de seda natural o variaciones de secuencias de aminoácidos de polipéptidos de seda natural o de combinaciones de ambos.

En el contexto de la presente invención, una "unidad repetitiva" se refiere a una región que corresponde en la secuencia de aminoácidos a una región que comprende o consta de al menos un motivo peptídico (por ejemplo, AAAAAA (SEQ ID NO: 13) o GPGQQ (SEQ ID NO: 4)) que ocurre repetidamente dentro de un polipéptido de seda natural (por ejemplo, MaSpl, ADF-3, ADF-4 o Flag) (es decir, una secuencia de aminoácidos idéntica) o a una secuencia de aminoácidos sustancialmente similar a la misma (es decir, secuencia de aminoácidos variacional). A este respecto, "sustancialmente similar" significa un grado de identidad de aminoácidos de al menos 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o incluso 99,9%, preferiblemente en toda la longitud de la secuencia respectiva de aminoácidos natural de referencia. Una "unidad repetitiva" que tiene una secuencia de aminoácidos que es "sustancialmente similar" a una secuencia de aminoácidos correspondiente dentro de un polipéptido de seda natural (es decir, una unidad repetitiva de tipo silvestre) también es similar con respecto a sus propiedades, por ejemplo, una proteína de seda que comprende la "unidad repetitiva sustancialmente similar" todavía es insoluble y retiene su insolubilidad y, por lo tanto, todavía puede separarse de las partes insolubles de la célula huésped solubilizadas en la etapa c) del método de la presente invención. Preferiblemente, una proteína de seda que comprende la "unidad repetitiva sustancialmente similar" es insoluble en una solución acuosa, por ejemplo, solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada y permanece insoluble en más de 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o incluso hasta 100% en una solución acuosa, por ejemplo, solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o agua desionizada que comprende una base (por ejemplo, NaOH, 0,005 a 1 M) durante un período de 10 a 90 minutos, por ejemplo durante 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85 o 90 minutos y/o a una temperatura entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C. En realizaciones preferidas de la presente invención, una proteína de seda que comprende la "unidad repetitiva sustancialmente similar", que permanece/queda insoluble, se separa/aísla de las partes insolubles de la célula huésped solubilizadas mediante filtración, preferiblemente usando una membrana filtrante que tiene un tamaño de poro de 10 nm a 200 nm, y/o por centrifugación, preferiblemente de 3000 a 12000 xg y más preferiblemente de 4000 a 8000 x g, durante 10 a 40 min, preferiblemente 20 a 40 min.

Una "unidad repetitiva" que tiene una secuencia de aminoácidos que es "idéntica" a la secuencia de aminoácidos de un polipéptido de seda natural, por ejemplo, puede ser una porción de un polipéptido de seda correspondiente a uno o más motivos peptídicos de MaSp I (SEQ ID NO: 43) MaSp II (SEQ ID NO: 44), ADF-3 (SEQ ID NO: 1) y/o ADF-4 (SEQ ID NO: 2). Una "unidad repetitiva" que tiene una secuencia de aminoácidos que es "sustancialmente similar" a la secuencia de aminoácidos de un polipéptido de seda natural, por ejemplo, puede ser una porción de un polipéptido de seda correspondiente a uno o más motivos peptídicos de MaSpl (SEQ ID NO: 43) MaSpII (SEQ ID NO: 44), ADF-3 (SEQ ID NO: 1) y/o ADF-4 (SEQ ID NO: 2), pero con una o más sustituciones de aminoácidos en posiciones de aminoácidos específicos.

La "unidad repetitiva" no incluye el dominio de aminoácidos hidrófilo no repetitivo que generalmente se cree que está presente en el extremo amino y/o carboxilo de los polipéptidos de seda naturales.

Una "unidad repetitiva" de acuerdo con la presente invención se refiere además a una secuencia de aminoácidos con una longitud de 3 a 200 aminoácidos, o de 5 a 150 aminoácidos, preferiblemente con una longitud de 10 a 100

aminoácidos, o 15 a 80 aminoácidos y más preferiblemente con una longitud de 18 a 60, o de 20 a 40 aminoácidos. Por ejemplo, la unidad repetitiva de acuerdo con la presente invención puede tener una longitud de 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, o 200 aminoácidos. Lo más preferiblemente, la unidad repetitiva de acuerdo con la invención consta de 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 18, 20, 24, 27, 28, 30, 34, 35 o 39 aminoácidos.

La proteína de seda aislada con el método de la presente invención consiste preferiblemente entre 6 y 1500 aminoácidos, o entre 200 y 1300 aminoácidos y lo más preferiblemente entre 250 y 1200 aminoácidos, o entre 500 y 1000 aminoácidos.

Preferiblemente, la proteína de seda aislada con el método de la presente invención comprende o consiste de entre 2 y 80 unidades repetitivas, entre 3 y 80 unidades repetitivas, o entre 4 y 60 unidades repetitivas, más preferiblemente entre 8 y 48 unidades repetitivas, o entre 10 a 40 unidades repetitivas y lo más preferiblemente entre 16 a 32 unidades repetitivas. Por ejemplo, la proteína de seda aislada con el método de la presente invención puede comprender o consta de 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 u 80 unidades repetitivas. Lo más preferiblemente, el polipéptido de seda comprende 4, 8, 12, 16, 24, 32 o 48 unidades repetitivas. Como se mencionó anteriormente, al menos dos de las unidades repetitivas comprendidas en el polipéptido de seda aislado con el método de la presente invención son unidades repetitivas idénticas. Por lo tanto, el polipéptido de seda aislado con el método de la presente invención puede comprender múltiples copias de una unidad repetitiva idéntica (por ejemplo, A₂ o C₁₆, en la que los ítems 2 o 16 representan el número de unidades repetitivas) o copias múltiples de dos o más unidades repetitivas diferentes (por ejemplo, (AQ)₂₄ o (QAQ)₈). Por ejemplo, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 u 80 de las 80 unidades repetitivas que pueden estar comprendidas en el polipéptido de seda aislado con el método de la presente invención son unidades repetitivas idénticas.

El polipéptido de seda aislado con el método de la presente invención puede comprender o constar de una secuencia de aminoácidos de cualquier polipéptido de seda conocido por un experto en la técnica. Se prefiere que el polipéptido de seda aislado con el método de la presente invención comprenda o conste de una secuencia de aminoácidos de un polipéptido de seda de artrópodos, preferiblemente de un polipéptido de seda de araña, o un polipéptido de seda de insecto. El polipéptido de seda aislado con el método de la presente invención también puede comprender o constar de una secuencia de aminoácidos de un polipéptido de seda de mejillón.

Se prefiere que el polipéptido de seda de araña comprenda o conste de una secuencia de aminoácidos de un polipéptido de la glándula ampulácea mayor, tal como un polipéptido de seda de araña dragalina, un polipéptido de glándula ampulácea menor, un polipéptido flageliforme, un polipéptido de seda araña agregado, un polipéptido de seda de araña aciniforme o un polipéptido de seda de araña piriforme. Lo más preferiblemente, el polipéptido de seda de araña comprende o consta de una secuencia de aminoácidos de un polipéptido de seda de araña dragalina o un polipéptido de seda de araña flageliforme. En general, se prefiere seleccionar la secuencia de aminoácidos del polipéptido de dragalina o del polipéptido flageliforme de un polipéptido dragalina o polipéptido flageliforme de arañas de telaraña de Araneidae o Araneoides.

Se prefiere que el polipéptido de seda de insecto comprenda o conste de una secuencia de aminoácidos de un polipéptido de seda de lepidópteros. Más preferiblemente, el polipéptido de seda de insecto comprende o consta de una secuencia de aminoácidos de un polipéptido de seda de Bombycidae, lo más preferiblemente de Bombyx mori.

Preferiblemente, los polipéptidos de seda mencionados anteriormente se producen de forma recombinante, es decir, son polipéptidos de seda recombinantes. Por ejemplo, los polipéptidos de seda aislados con el método de la presente invención son polipéptidos de seda de araña recombinantes tales como polipéptidos de seda de araña dragalina o polipéptidos de seda de araña flageliforme, polipéptidos de seda de insecto recombinantes o polipéptidos de seda de mejillón recombinantes.

La unidad repetitiva del polipéptido de seda aislado con el método de la presente invención puede comprender o constar de una secuencia de aminoácidos de cualquier región que comprenda o conste de al menos un motivo peptídico que ocurra repetidamente dentro de un polipéptido de seda natural conocido por un experto en la materia. Preferiblemente, la unidad repetitiva del polipéptido de seda aislado con el método de la presente invención comprende o consta de una secuencia de aminoácidos de una región que comprende o consta de al menos un motivo peptídico que ocurre repetidamente dentro de un polipéptido de seda de araña, o un polipéptido de seda de insecto. La unidad repetitiva del polipéptido de seda aislado con el método de la presente invención también puede comprender o constar de una secuencia de aminoácidos de una región que comprende o consta de al menos un motivo peptídico que ocurre repetidamente dentro de un polipéptido de seda de mejillón.

5 Se prefiere que la unidad repetitiva de seda de araña comprenda o conste de una secuencia de aminoácidos de una región que comprenda o conste de al menos un motivo peptídico que ocurre repetitivamente dentro de un polipéptido de glándula ampulácea mayor de origen natural, tal como polipéptido de seda de araña dragalina, un polipéptido de la glándula ampulácea menor, un polipéptido flageliforme, un polipéptido de seda de araña agregado, un polipéptido de seda de araña aciniforme o un polipéptido de seda de araña piriforme. Lo más preferiblemente, la unidad repetitiva comprende o consta de una secuencia de aminoácidos de una región que comprende o consta de al menos un motivo peptídico que ocurre repetidamente dentro de un polipéptido de seda de araña draglina natural o un polipéptido de seda de araña flageliforme.

10 Se prefiere que la unidad repetitiva de seda de insecto comprenda o conste de una secuencia de aminoácidos de una región que comprenda o conste de al menos un motivo peptídico que ocurre repetidamente dentro de un polipéptido de seda natural de lepidópteros. Más preferiblemente, la unidad repetitiva de seda de insecto comprende o consta de una secuencia de aminoácidos de una región que comprende o consta de al menos un motivo peptídico que ocurre repetidamente dentro de un polipéptido de seda de insecto natural de Bombycidae, lo más preferiblemente de Bombyx mori.

15 El término "secuencia de consenso" como se usa en el contexto de la presente invención se refiere a una secuencia de aminoácidos que contiene aminoácidos que ocurren frecuentemente en una determinada posición (por ejemplo, "G") y en donde, otros aminoácidos que no son además determinados se sustituyen por el marcador de posición "X".

20 Preferiblemente, la proteína de seda aislada con el método de la presente invención comprende, esencialmente consta de, o consta de al menos dos unidades repetitivas idénticas, cada una de las cuales comprende al menos una, preferiblemente una secuencia consenso seleccionada del grupo que consta de:

i) GPGXX (SEQ ID NO: 3), en la que X es cualquier aminoácido, preferiblemente en cada caso independientemente seleccionado de A, S, G, Y, P y Q;

ii) GGX, en la que X es cualquier aminoácido, preferiblemente en cada caso independientemente seleccionado de Y, P, R, S, A, T, N y Q, más preferiblemente en cada caso independientemente seleccionado de Y, P y Q; y

25 iii) A_x, en la que x es un número entero de 5 a 10.

También se prefiere que la proteína de seda aislada con el método de la presente invención comprenda o conste de al menos dos unidades repetitivas idénticas, cada una de las cuales comprende al menos una, preferiblemente una, secuencia de aminoácidos seleccionada del grupo que consta de: GGRPSDTYG (SEQ ID NO: 18) y GGRPSSSYG (SEQ ID NO: 19).

30 Los motivos iterados (péptido) GPGXX (SEQ ID NO: 3) y GGX, es decir, motivos ricos en glicina, proporcionan flexibilidad al polipéptido de seda y, por lo tanto, al hilo formado a partir de la proteína de seda que contiene dichos motivos. En detalle, el motivo iterado GPGXX (SEQ ID NO: 3) forma estructuras espirales de giro β, que imparten elasticidad al polipéptido de seda. Las sedas de ampulácea mayor y flageliformes tienen ambas un motivo GPGXX (SEQ ID NO: 3). El motivo GGX iterado está asociado con una estructura helicoidal que tiene tres aminoácidos por giro y se encuentra en la mayoría de las sedas de araña. El motivo GGX puede proporcionar propiedades elásticas adicionales a la seda. El motivo iterado de polialanina A_x (péptido) forma una estructura cristalina de lámina β que proporciona resistencia al polipéptido de seda (documento WO 03/057727). Los motivos GGRPSDTYG (SEQ ID NO: 18) y GGRPSSSYG (SEQ ID NO: 19) (péptido) se han seleccionado de resilina (documento WO 2008/155304). La resilina es una proteína elastomérica que se encuentra en la mayoría de los artrópodos (*Arthropoda*). Se encuentra en regiones especializadas de la cutícula, proporcionando baja rigidez y alta resistencia (Elvin et al., Nature (473): 999-1002, 2005).

35 Por lo tanto, en una realización preferida de la presente invención, la proteína de seda comprende o consta de unidades repetitivas que comprenden cada una al menos una (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8 o 9), preferiblemente una secuencia de aminoácidos seleccionada del grupo que consta de GPGAS (SEQ ID NO: 5), GPGSG (SEQ ID NO: 6), GPGGY (SEQ ID NO: 7), GPGGP (SEQ ID NO: 8), GPGGA (SEQ ID NO: 9), GPGQQ (SEQ ID NO: 4), GPGGG (SEQ ID NO: 10), GPGQG (SEQ ID NO: 40) y GPGGS (SEQ ID NO: 11). En otra realización preferida de la presente invención, el polipéptido de seda comprende o consta de unidades repetitivas que comprenden cada una al menos una (por ejemplo, 1, 2, 3, 4, 5, 8, 7 u 8), preferiblemente una, secuencia de aminoácidos seleccionada del grupo que consta de GGY, GGP, GGA, GGR, GGS, GGT, GGN y GGQ. En una realización preferida adicional de la presente invención, el polipéptido de seda comprende o consta de unidades repetitivas que comprenden cada una al menos una (por ejemplo, 1, 2, 3, 4, 5 o 6), preferiblemente una, secuencia de aminoácidos seleccionada del grupo que consta de AAAAA (SEQ ID NO: 12), AAAAAA (SEQ ID NO: 13), AAAAAA (SEQ ID NO: 14), AAAAAA (SEQ ID NO: 15), AAAAAA (SEQ ID NO: 16) y AAAAAA (SEQ ID NO: 17).

55 En otra realización preferida de la invención, la proteína de seda comprende o consta de unidades repetitivas que comprenden cada una al menos una (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25), preferiblemente una secuencia de aminoácidos seleccionada del grupo que consta de GPGAS (SEQ ID NO: 5), GPGSG (SEQ ID NO: 6), GPGGY (SEQ ID NO: 7), GPGGP (SEQ ID NO: 8), GPGGA (SEQ ID NO: 9), GPGQQ (SEQ ID NO: 4), GPGGG (SEQ ID NO: 10), GPGQG (SEQ ID NO: 40), GPGGS (SEQ ID NO: 11), GGY,

GGP, GGA, GGR, GGS, GGT, GGN, GGQ, AAAAA (SEQ ID NO: 12), AAAAA (SEQ ID NO: 13), AAAAAA (SEQ ID NO: 14), AAAAAA (SEQ ID NO: 15), AAAAAA (SEQ ID NO: 16), AAAAAA (SEQ ID NO: 17), GGRPSDTYG (SEQ ID NO: 18) y GGRPSSSYG (SEQ ID NO: 19).

5 Lo más preferiblemente, la proteína de seda aislada con el método de la presente invención comprende, esencialmente consta de, o consta de unidades repetitivas, que comprenden o consisten en

- (i) GPGAS (SEQ ID NO: 5), AAAAA (SEQ ID NO: 13), GGY y GPGSG (SEQ ID NO: 6) como secuencia de aminoácidos, preferiblemente en este orden,
- (ii) AAAAAA (SEQ ID NO: 15), GPGGY (SEQ ID NO: 7), GPGGY (SEQ ID NO: 7) y GPGGP (SEQ ID NO: 8) como secuencia de aminoácidos, preferiblemente en este orden,
- 10 (iii) GPGQQ (SEQ ID NO: 4), GPGQQ (SEQ ID NO: 4), GPGQQ (SEQ ID NO: 4) y GPGQQ (SEQ ID NO: 4) como secuencia de aminoácidos,
- (iv) GPGGA (SEQ ID NO: 9), GGP, GPGGA (SEQ ID NO: 9), GGP, GPGGA (SEQ ID NO: 9) y GGP como secuencia de aminoácidos, preferiblemente en este orden,
- 15 (v) AAAAAA (SEQ ID NO: 15), GPGQG (SEQ ID NO: 40) y GGR como secuencia de aminoácidos, preferiblemente en este orden,
- (vi) AAAAAA (SEQ ID NO: 15), GPGGG (SEQ ID NO: 10), GGR, GGN y GGR como secuencia de aminoácidos, preferiblemente en este orden,
- (vii) GGA, GGA, GGA, GGS, GGA y GGS como secuencia de aminoácidos, preferiblemente en este orden, y/o
- 20 (viii) GPGGA (SEQ ID NO: 9), GPGGY (SEQ ID NO: 7), GPGGS (SEQ ID NO: 11), GPGGY (SEQ ID NO: 7), GPGGS (SEQ ID NO: 11) y GPGGY (SEQ ID NO: 7) como secuencia de aminoácidos, preferiblemente en este orden.

Debe observarse que al menos dos de las unidades repetitivas comprendidas en los polipéptidos de seda mencionados anteriormente son unidades repetitivas idénticas.

25 Preferiblemente, el polipéptido de seda aislado con el método de la presente invención comprende, esencialmente consta de, o consta de entre 2 a 80 unidades repetitivas, entre 3 a 80 unidades repetitivas, o entre 4 a 60 unidades repetitivas, más preferiblemente entre 8 a 48 unidades repetitivas, o entre 10 a 40 unidades repetitivas y lo más preferiblemente entre 16 a 32 unidades repetitivas, es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 u 80 unidades repetitivas, cada una de las cuales comprende al menos una, preferiblemente una, secuencia de consenso seleccionada del grupo que consta de:

- i) GPGXX (SEQ ID NO: 3), en la que X es cualquier aminoácido, preferiblemente en cada caso independientemente seleccionado de A, S, G, Y, P y Q;
- 35 ii) GGX, en la que X es cualquier aminoácido, preferiblemente en cada caso independientemente seleccionado de Y, P, R, S, A, T, N y Q, más preferiblemente en cada caso seleccionado independientemente de Y, P y Q; y
- iii) A_x, en la que x es un número entero de 5 a 10.

Como se mencionó anteriormente, al menos dos de las unidades repetitivas comprendidas en el polipéptido de seda aislado con el método de la presente invención son unidades repetitivas idénticas.

40 También se prefiere que la proteína de seda aislada con el método de la presente invención comprenda o conste de 2 a 80 unidades repetitivas, entre 3 a 80 unidades repetitivas, o entre 4 a 60 unidades repetitivas, más preferiblemente entre 8 a 48 unidades repetitivas, o entre 10 a 40 unidades repetitivas y lo más preferiblemente entre 16 a 32 unidades repetitivas, cada una de las cuales comprende al menos una, preferiblemente una, secuencia de aminoácidos seleccionada del grupo que consta de: GGRPSDTYG (SEQ ID NO: 18) y GGRPSSSYG (SEQ ID NO: 19).

45 Por lo tanto, la proteína de seda aislada en el método de la presente invención comprende preferiblemente o consiste de entre 2 a 80 unidades repetitivas, entre 3 a 80 unidades repetitivas, o entre 4 a 60 unidades repetitivas, más preferiblemente entre 8 a 48 unidades repetitivas, o entre 10 a 40 unidades repetitivas y lo más preferiblemente entre 16 a 32 unidades repetitivas, cada una de las cuales comprende al menos una (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25), preferiblemente una secuencia de aminoácidos seleccionada del grupo que consta de GPGAS (SEQ ID NO: 5), GPGSG (SEQ ID NO: 6), GPGGY (SEQ ID NO: 7), GPGGP (SEQ ID NO: 8), GPGGA (SEQ ID NO: 9), GPGQQ (SEQ ID NO: 4), GPGQG (SEQ ID NO: 40), GPGGG (SEQ ID NO: 10), GPGGS (SEQ ID NO: 11), GGY, GGP, GGA, GGR, GGS, GGT, GGN, GGQ, AAAAA (SEQ ID NO: 12), AAAAA (SEQ ID NO: 13), AAAAAA (SEQ ID NO: 14), AAAAAA (SEQ ID NO: 15), AAAAAA (SEQ ID NO: 16), AAAAAA (SEQ ID NO: 17), GGRPSDTYG (SEQ ID NO: 18) y GGRPSSSYG (SEQ ID NO: 19).

55 Más preferiblemente, la proteína de seda aislada con el método de la presente invención comprende, consiste esencialmente o consta de

- (i) unidades repetitivas que comprenden o consisten en GPGAS (SEQ ID NO: 5), AAAAA (SEQ ID NO: 13), GGY y GPGSG (SEQ ID NO: 6) como secuencia de aminoácidos, preferiblemente en este orden,
- (ii) unidades repetitivas que comprenden o consisten en AAAAAA (SEQ ID NO: 15), GPGGY (SEQ ID NO: 7),

GPGGY (SEQ ID NO: 7) y GPGGP (SEQ ID NO: 8) como secuencia de aminoácidos, preferiblemente en este orden,
 (iii) unidades repetitivas que comprenden o consisten en GPGQQ (SEQ ID NO: 4), GPGQQ (SEQ ID NO: 4), GPGQQ
 (SEQ ID NO: 4) y GPGQQ (SEQ ID NO: 4) como secuencia de aminoácidos,
 (iv) unidades repetitivas que comprenden o consisten en GPGGA (SEQ ID NO: 9), GGP, GPGGA (SEQ ID NO: 9),
 5 GGP, GPGGA (SEQ ID NO: 9) y GGP como secuencia de aminoácidos, preferiblemente en este orden,
 (v) unidades repetitivas que comprenden o consisten en AAAAAAAAA (SEQ ID NO: 15), GPGQG (SEQ ID NO: 40) y
 GGR como secuencia de aminoácidos, preferiblemente en este orden,
 (vi) unidades repetitivas que comprenden o consisten en AAAAAAAAA (SEQ ID NO: 15), GPGGG (SEQ ID NO: 10),
 10 GGR, GGN y GGR como secuencia de aminoácidos, preferiblemente en este orden,
 (vii) unidades repetitivas que comprenden o consisten en GGA, GGA, GGA, GGS, GGA y GGS como secuencia de
 aminoácidos, preferiblemente en este orden, y/o
 (viii) unidades repetitivas que comprenden o consisten en GPGGA (SEQ ID NO: 9), GPGGY (SEQ ID NO: 7), GPGGS
 (SEQ ID NO: 11), GPGGY (SEQ ID NO: 7), GPGGS (SEQ ID NO: 11) y GPGGY (SEQ ID NO: 7) como secuencia de
 aminoácidos, preferiblemente en este orden.

15 Debe observarse que al menos dos de las unidades repetitivas comprendidas en los polipéptidos de seda
 mencionados anteriormente son unidades repetitivas idénticas.

Preferiblemente, el polipéptido de seda aislado con el método de la presente invención comprende, consiste
 esencialmente o consta de

20 (i) (GPGXX)_n (SEQ ID NO: 3) como una unidad repetitiva, en la que X es cualquier aminoácido, preferiblemente en
 cada caso independientemente seleccionado de A, S, G, Y, P y Q y n es 2, 3, 4, 5, 6, 7, 8 o 9;
 ii) (GGX)_n como una unidad repetitiva, en la que X es cualquier aminoácido, preferiblemente en cada caso
 independientemente seleccionado de Y, P, R, S, A, T, N y Q, más preferiblemente en cada caso independientemente
 25 seleccionado de Y, P y Q, y n es 2, 3, 4, 5, 6, 7 u 8; y/o
 iii) (A_x)_n como una unidad repetitiva, en la que x es un número entero de 5 a 10 y n es 2, 3, 4, 5, 6, 7, 8, 9 o 10.

Como se mencionó anteriormente, al menos dos de las unidades repetitivas comprendidas en los polipéptidos de seda
 aislados con el método de la presente invención son unidades repetitivas idénticas.

30 Se prefiere que las unidades repetitivas se seleccionen independientemente del módulo A (SEQ ID NO: 20), módulo
 C (SEQ ID NO: 21), módulo Q (SEQ ID NO: 22), módulo K (SEQ ID NO: 23), módulo sp (SEQ ID NO: 24), módulo S
 (SEQ ID NO: 25), módulo R (SEQ ID NO: 26), módulo X (SEQ ID NO: 27) o módulo Y (SEQ ID NO: 28), o variantes
 de las mismas (es decir, variantes del módulo A, variantes del módulo C, variantes del módulo Q, variantes del módulo
 K, variantes del módulo sp, variantes del módulo S, variantes del módulo R, variantes del módulo X o variantes del
 módulo Y). Los módulos A (SEQ ID NO: 20) y Q (SEQ ID NO: 22) se basan en la secuencia de aminoácidos de ADF-
 35 de la araña *Araneus diadematus*. El módulo C (SEQ ID NO: 21) se basa en la secuencia de aminoácidos de ADF-4
 de la araña *Araneus diadematus*. Los módulos K (SEQ ID NO: 23), sp (SEQ ID NO: 24), X (SEQ ID NO: 27) e Y (SEQ
 ID NO: 28) se basan en la secuencia de aminoácidos de la proteína FLAG flageliforme de la araña *Nephila clavipes*
 (documento WO 2006/008163). Los módulos S (SEQ ID NO: 25) y R (SEQ ID NO: 26) se basan en resilina (*Arthropoda*)
 (documento WO 2008/155304).

40 Por lo tanto, en una realización preferida de la presente invención, las unidades repetitivas del polipéptido de seda
 consisten en el módulo A: GPYGPASAAAAAGGYGPGSGQQ (SEQ ID NO: 20), módulo C: GSSAAAAAAAASGP
 GGYGPENQPSGPGGYGPGGP: (SEQ ID NO: 21), módulo Q: GPGQQGPGQQGPGQQGPGQQ (SEQ ID NO: 22),
 módulo K: GPGGAGGPYGGAGGPYGGAGGPY (SEQ ID NO: 23), el módulo sp: GGTTIIEDLDITIDGAD
 GPITISEELTI (SEQ ID NO: 24), el módulo S: PGSSAAAAAAAASGPGQQGQGQGQGGRRPSDTYG (SEQ ID NO:
 45 25), el módulo de R: SAAAAAAGPGGGNGGRPSDTYGAPGGGNGGRPSSSYG (SEQ ID NO: 26), módulo X:
 GGAGGAGGAGGSGGAGGS (SEQ ID NO: 27), o módulo Y: GPGGAGPGGYGPGGSGPGGYGPGGSGPGGY (SEQ
 ID NO: 28), o variantes de las mismas.

50 Preferentemente, la proteína de seda aislada con el método de la presente invención comprende, esencialmente
 consta de, o consiste de entre 2 a 80 unidades repetitivas, entre 3 a 80 unidades repetitivas, o entre 4 a 60 unidades
 repetitivas, más preferiblemente entre 8 a 48 unidades repetitivas, o entre 10 a 40 unidades repetitivas y lo más
 preferiblemente entre 16 a 32 unidades repetitivas, es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50,
 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 u 80
 55 unidades repetitivas, que se seleccionan independientemente del módulo A (SEQ ID NO: 20), módulo C (SEQ ID NO:
 21), módulo Q (SEQ ID NO: 22), módulo K (SEQ ID NO: 23), módulo sp (SEQ ID NO: 24), módulo S (SEQ ID NO: 25),
 módulo R (SEQ ID NO: 26), módulo X (SEQ ID NO: 27) o módulo Y (SEQ ID NO: 28), o variantes de los mismos (es
 decir variantes del módulo A, variantes del módulo C, variantes del módulo Q, variantes del módulo K, variantes del
 módulo sp, variantes del módulo S, variantes del módulo R, variantes del módulo X o variantes del módulo Y). Debe
 observarse que al menos dos de las unidades repetitivas comprendidas en el polipéptido de seda aislado con el método
 de la presente invención son unidades repetitivas idénticas (módulos).

60 Por lo tanto, se prefiere que el polipéptido de seda aislado con el método de la presente invención comprenda,

que los módulos (unidades repetitivas) se combinan o concatenan directamente entre sí o pueden significar en el contexto de la presente invención que los módulos (unidades repetitivas) se combinan o concatenan entre sí a través de uno o más aminoácidos espaciadores. En realizaciones preferidas, los módulos (unidades repetitivas) comprendidos en el polipéptido de seda se combinan o concatenan directamente entre sí. En otras realizaciones preferidas, los módulos (unidades repetitivas) comprendidos en el polipéptido de seda se combinan o concatenan entre sí a través de 1 a 25 o de 1 a 20 aminoácidos espaciadores, más preferiblemente a través de 1 a 15 o de 1 a 10 aminoácidos espaciadores, y más preferiblemente, a través de 1 a 5 aminoácidos espaciadores, es decir, a través de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 aminoácidos espaciadores. Dichos aminoácidos espaciadores pueden ser cualquier aminoácido que se encuentre naturalmente en las proteínas. Preferiblemente, dicho aminoácido espaciador no es prolina. Los aminoácidos no cargados tales como alanina o glicina son particularmente preferidos. Se prefiere que dichos aminoácidos espaciadores sean aminoácidos que no cambian la propiedad de la proteína de seda de ser insoluble y retener su insolubilidad, de modo que dicha proteína de seda todavía puede separarse de las partes insolubles de la célula huésped solubilizadas en la etapa c) del método de la presente invención. Se prefiere además que dichos aminoácidos espaciadores sean aminoácidos que no causen impedimento estérico, por ejemplo, aminoácidos que tienen un tamaño pequeño tal como alanina o glicina. En realizaciones más preferidas, el polipéptido de seda comprende módulos que se combinan directamente entre sí y módulos que se combinan entre sí mediante 1 a 25 o 1 a 20 aminoácidos espaciadores, más preferiblemente a través de 1 a 15 o 1 a 10 aminoácidos espaciadores, y lo más preferiblemente, a través de 1 a 5 aminoácidos espaciadores, es decir, a través de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 aminoácidos espaciadores.

Una variante del módulo A, C, Q, K, sp, S, R, X o Y difiere del módulo de referencia (tipo silvestre) A, C, Q, K, sp, S, R, X o Y del que se derivan hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 cambios de aminoácidos en la secuencia de aminoácidos (es decir, sustituciones, adiciones, inserciones, eliminaciones, truncamientos del terminal N y/o truncamientos del terminal C). Tal variante de módulo puede caracterizarse de forma alternativa o adicional por un cierto grado de identidad de secuencia con el módulo de referencia (tipo silvestre) del que se deriva. Por lo tanto, una variante del módulo A, C, Q, K, sp, S, R, X o Y tiene una identidad de secuencia de al menos 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o incluso 99,9% con el respectivo módulo de referencia (tipo silvestre) A, C, Q, K, sp, S, R, X o Y. Preferiblemente, la identidad de secuencia se encuentra en un tramo continuo de al menos 10, 15, 18, 20, 24, 27, 28, 30, 34, 35 o más aminoácidos, preferiblemente sobre la longitud total del respectivo módulo de referencia (tipo silvestre) A, C, Q, K, sp, S, R, X o Y.

Se prefiere particularmente que la identidad de secuencia sea al menos del 80% en toda la longitud, sea al menos del 85% en toda la longitud, sea de al menos 90% en toda la longitud, sea de al menos 95% en toda la longitud, sea de al menos 98% en toda la longitud, o sea de al menos 99% en toda la longitud del respectivo módulo de referencia (tipo silvestre) A, C, Q, K, sp, S, R, X o Y. se prefiere particularmente que la identidad de secuencia sea al menos 80% en un tramo continuo de al menos 10, 15, 18, 20, 24, 28 o 30 aminoácidos, es al menos 85% en un tramo continuo de al menos 10, 15, 18, 20, 24, 28 o 30 aminoácidos, es al menos 90% en un tramo continuo de al menos 10, 15, 18, 20, 24, 28 o 30 aminoácidos, es al menos 95% en un tramo continuo de al menos 10, 15, 18, 20, 24, 28 o 30 aminoácidos, es al menos 98% en un tramo continuo de al menos 10, 15, 18, 20, 24, 28 o 30 aminoácidos, o es al menos 99% en un tramo continuo de al menos 10, 15, 18, 20, 24, 28 o 30 aminoácidos del respectivo módulo de referencia (tipo silvestre) A, C, Q, K, sp, S, R, X o Y.

Un fragmento (o variante de eliminación) del módulo A, C, Q, K, sp, S, R, X o Y tiene preferiblemente una eliminación de hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 aminoácidos en su terminal N y/o en su terminal C. La eliminación también puede ser internamente.

Además, la variante o fragmento del módulo A, C, Q, K, sp, S, R, X o Y solo se considera como una variante o fragmento del módulo A, C, Q, K, sp, S, R, X o Y dentro del contexto de la presente invención, si las modificaciones con respecto a la secuencia de aminoácidos en la que se basa la variante o fragmento no cambian la propiedad de una proteína de seda de ser insoluble y retener su insolubilidad, de modo que dicha proteína de seda todavía pueda separarse de las partes insolubles de la célula huésped solubilizadas en la etapa c) del método de la presente invención. Preferiblemente, una proteína de seda que comprende la variante o fragmento del módulo A, C, Q, K, sp, S, R, X o Y es insoluble en una solución acuosa, por ejemplo, solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada y permanece insoluble en más del 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o incluso hasta 100% en una solución acuosa, por ejemplo, solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o agua desionizada que comprende una base (por ejemplo, NaOH, 0,005 a 1 M) durante un período de 10 a 90 minutos, por ejemplo durante 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85 o 90 minutos y/o a una temperatura entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C. En realizaciones preferidas de la presente invención, una proteína de seda que comprende la variante o fragmento del módulo A, C, Q, K, sp, S, R, X o Y, que permanece/queda insoluble, se separa/aísla de las partes insolubles de la célula huésped solubilizadas mediante filtración, preferiblemente usando una membrana filtrante que tiene un tamaño de poro de 10 nm a 200 nm, y/o por centrifugación, preferiblemente de 3000 a 12000 x g y más preferiblemente de 4000 a 8000 x g, durante 10 a 40 min, preferiblemente 20 a 40 min.

Por lo tanto, en una realización preferida de la presente invención, las unidades repetitivas se seleccionan independientemente del módulo A^C (SEQ ID NO: 29), módulo A^K (SEQ ID NO: 30), módulo C^C (SEQ ID NO: 31), módulo C^{K1} (SEQ ID NO: 32), módulo C^{K2} (SEQ ID NO: 33) o módulo C^{KC} (SEQ ID NO: 34). Los módulos A^C (SEQ ID NO: 29), A^K (SEQ ID NO: 30), C^C (SEQ ID NO: 31), C^{K1} (SEQ ID NO: 32), C^{K2} (SEQ ID NO: 33) y C^{KC} (SEQ ID NO: 34) son variantes del módulo A que se basa en la secuencia de aminoácidos de ADF-3 de la araña *Araneus diadematus* y del módulo C que se basa en la secuencia de aminoácidos de ADF-4 de la araña *Araneus diadematus* (documento WO 2007/025719). En el módulo A^C (SEQ ID NO: 29) el aminoácido S (serina) en la posición 21 ha sido reemplazado por el aminoácido C (cisteína), en el módulo A^K (SEQ ID NO: 30) el aminoácido S en la posición 21 ha sido reemplazado por el aminoácido K (lisina), en el módulo C^C (SEQ ID NO: 31) el aminoácido S en la posición 25 ha sido reemplazado por el aminoácido C, en el módulo C^{K1} (SEQ ID NO: 32) el aminoácido S en la posición 25 ha sido reemplazado por el aminoácido K, en el módulo C^{K2} (SEQ ID NO: 33) el aminoácido E (glutamato) en la posición 20 ha sido reemplazado por el aminoácido K, y en el módulo C^{KC} (SEQ ID NO: 34) el aminoácido E en la posición 20 ha sido reemplazado por el aminoácido K y el aminoácido S en la posición 25 ha sido reemplazado por el aminoácido C (documento WO 2007/025719).

Preferentemente, las unidades repetitivas en el polipéptido de seda aislado con el método de la presente invención consisten en el módulo A^C: GPYGGASAAAAAAGGYGPGCGQQ (SEQ ID NO: 29), módulo A^K: GPYGGASAAAAAAGGYGPGKGGQ (SEQ ID NO: 30), módulo C^C: GSSAAAAAASGPGGYGPENQGPCGPGGYGPGGP (SEQ ID NO: 31), módulo C^{K1}: GSSAAAAAASGPGGYGPENQGPKGGPYGPGGP (SEQ ID NO: 32), módulo C^{K2}: GSSAAAAAASGPGGYGPKNQGPGGYGPGGP (SEQ ID NO: 33), o módulo C^{KC}: GSSAAAAAASGPGGYGPKNQGPGPGGYGPGGP (SEQ ID NO: 34).

También se prefiere que el polipéptido de seda aislado con el método de la presente invención comprenda, esencialmente consista o conste de entre 2 a 80 unidades repetitivas, entre 3 a 80 unidades repetitivas, o entre 4 a 60 unidades repetitivas, preferiblemente entre 8 y 48 unidades repetitivas, o entre 10 y 40 unidades repetitivas y lo más preferiblemente entre 16 y 32 unidades repetitivas, es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 u 80 unidades repetitivas, que se seleccionan independientemente del módulo A^C (SEQ ID NO: 29), módulo A^K (SEQ ID NO: 30), módulo C^C (SEQ ID NO: 31), módulo C^{K1} (SEQ ID NO: 32), módulo C^{K2} (SEQ ID NO: 33) o módulo C^{KC} (SEQ ID NO: 34). Por ejemplo, el polipéptido de seda aislado con el método de la presente invención puede comprender o constar de los módulos C^C₄, C^C₈, C^C₁₆, C^C₃₂, A^C₅, A^C₈ o A^C₁₀. Debe observarse que al menos dos de las unidades repetitivas comprendidas en el polipéptido de seda aislado con el método de la presente invención son unidades repetitivas idénticas (módulos).

Por ejemplo, el polipéptido de seda puede comprender o consta de los módulos C^C₄, C^C₈, C^C₁₆, C^C₃₂, A^C₅, A^C₈ o A^C₁₀.

Los módulos A^K, C^C, C^{K1}, C^{K2} y C^{KC} también se pueden combinar entre sí, es decir, el módulo (unidad repetitiva) A^K se puede combinar con módulo (unidad repetitiva) C^C (es decir, combinación A^KC^C), el módulo (unidad repetitiva) C^{K1} se puede combinar con el módulo (unidad repetitiva) C^{K2} y con el módulo (repetitivo unidad) C^{KC} (es decir, combinación C^{K1}C^{K2}C^{KC}), etc., con la condición de que el polipéptido de seda aislado con el método de la presente invención comprenda o conste de al menos dos unidades repetitivas que sean idénticas. Por lo tanto, el polipéptido de seda aislado con el método de la presente invención también puede comprender o constar de los módulos (A^K)_n, (C^C)_n, (C^{K1})_n, (C^{K2})_n, (C^{KC})_n, (A^KA^C)_n, (C^CC^C)_n, (C^{K1}C^{K2})_n, (C^{K2}C^{K1})_n, (C^{K1}C^{K2}C^{K1})_n, (C^{K2}C^{K1}C^{K2})_n, (C^{K1}C^{K2}C^{KC})_n, (C^{KC}C^{K2}C^{KC})_n, or (C^{KC}C^{K2}C^{K1})_n, en los que n es al menos 2, preferiblemente 4, 5, 6, 7, 8, 10, 12, 16 o 20. El término "combinados entre sí" se definió anteriormente.

Se prefiere además que el polipéptido de seda aislado con el método de la presente invención comprenda, esencialmente consista o conste de entre 2 a 80 unidades repetitivas, entre 3 a 80 unidades repetitivas o entre 4 a 60 unidades repetitivas, preferiblemente entre 8 y 48 unidades repetitivas, o entre 10 y 40 unidades repetitivas y lo más preferiblemente entre 16 y 32 unidades repetitivas, es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 u 80 unidades repetitivas, que se seleccionan independientemente del módulo A (SEQ ID NO: 20) o variantes del mismo, módulo C (SEQ ID NO: 21) o variantes del mismo, módulo Q (SEQ ID NO: 22) o variantes del mismo, módulo K (SEQ ID NO: 23) o variantes del mismo, módulo sp (SEQ ID NO: 24) o sus variantes, módulo S (SEQ ID NO: 25) o variantes del mismo, módulo R (SEQ ID NO: 26) o variantes del mismo, módulo X (SEQ ID NO: 27) o variantes del mismo, módulo Y (SEQ ID NO: 28) o variantes del mismo, módulo A^C (SEQ ID NO: 29), módulo A^K (SEQ ID NO: 30), módulo C^C (SEQ ID NO: 31), módulo C^{K1} (SEQ ID NO: 32), módulo C^{K2} (SEQ ID NO: 33) o módulo C^{KC} (SEQ ID NO: 34). Nuevamente, debe observarse que al menos dos de las unidades repetitivas comprendidas en el polipéptido de seda aislado con el método de la presente invención son unidades repetitivas idénticas (módulos).

Los módulos A^K, C^C, C^{K1}, C^{K2} y C^{KC} también se pueden combinar con los módulos A, C, Q, K, sp, S, R, X o Y, es decir, el módulo (unidad repetitiva) A^K se puede combinar con el módulo (unidad repetitiva) C (es decir, la combinación A^KC), o el módulo (unidad repetitiva) C^C se puede combinar con el módulo (unidad repetitiva) C (es decir, combinación C^CC), etc., con la condición de que el polipéptido de seda aislado con el método de la presente invención comprenda o conste de al menos dos unidades repetitivas que sean idénticas. Por lo tanto, el polipéptido de seda aislado con el

método de la presente invención también puede comprender o constar de los módulos (AQA^K)_n, (QA^K)_n, (QA^KQ)_n, (A^KQA)_n, (A^KQA^K)_n, (CC^C)_n, (CC^CC)_n, (C^CC^CC)_n, (CC^CC^C)_n, (C^CQ)_n, (QC^C)_n, (QC^CQ)_n, (C^CQC)_n, (CQC^C)_n, (C^CQC^C)_n, (CC^{K1})_n, (C^{K1}C)_n, (C^{K1}CC)_n, (CC^{K1}C)_n, (C^{K1}CC^{K1})_n, (C^{K1}CC^{K1}C)_n, (C^{K1}CC^{K1}C)_n, (C^{K1}CC^{K1})_n, (QC^{K1})_n, (QC^{K1}Q)_n, (A^KC^{K1}Q)_n, (QC^{K2}A^K)_n, o (C^{K1}C^{K2}C)_n, en los que n es al menos 2, preferiblemente 4, 5, 6, 7, 8, 10, 12, 16 o 20. El término "combinados entre sí" se definió anteriormente.

Por ejemplo, el polipéptido de seda aislado con el método de la presente invención comprende o consta de los módulos C₁₆C^C, C^CC₁₆, C₈C^CC₈, C₈C^CC₈, C^C₈C₈, C₄C^C₈C₄, C^C₄C₈C^C₄, C^C(AQ)₂₄ o (AQ)₂₄C^C.

El polipéptido de seda aislado con el método de la presente invención puede comprender además al menos una unidad no repetitiva (NR), es decir, 1, 2, 3, 4, 5, 6 o más unidades NR, preferiblemente una unidad NR. En el contexto de la presente invención, el término "unidad no repetitiva (NR)" se refiere a una región de aminoácidos presente en un polipéptido de seda natural que no muestra un patrón de repetición obvio (unidad no repetitiva o unidad NR). Preferiblemente, la secuencia de aminoácidos de la unidad no repetitiva corresponde a una secuencia de aminoácidos no repetitiva de polipéptidos de draglina naturales, preferiblemente de ADF-3 (SEQ ID NO: 1) o ADF-4 (SEQ ID NO: 2), o a una secuencia de aminoácidos sustancialmente similar a la misma.

Se prefiere particularmente que la secuencia de aminoácidos de la unidad no repetitiva corresponda a una secuencia de aminoácidos no repetitiva del terminal carboxilo de polipéptidos de draglina naturales, preferiblemente de ADF-3 (SEQ ID NO: 1) o ADF-4 (SEQ ID NO: 2), o a una secuencia de aminoácidos sustancialmente similar a la misma. Más preferiblemente, la secuencia de aminoácidos de la unidad no repetitiva corresponde a una secuencia de aminoácidos no repetitiva del terminal carboxilo de ADF-3 (SEQ ID NO: 1) que comprende los aminoácidos 513 a 636, o de ADF-4 (SEQ ID NO: 2) que comprende los aminoácidos 302 a 410, o una secuencia de aminoácidos sustancialmente similar a la misma.

A este respecto, "sustancialmente similar" significa un grado de identidad de aminoácidos de al menos 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o incluso 99,9%, preferiblemente más de 20, 30, 40, 50, 60, 70, 80 o más aminoácidos, más preferiblemente en toda la longitud de la secuencia de aminoácidos respectiva no repetitiva de referencia (terminal carboxilo) de polipéptidos de draglina naturales, preferiblemente de ADF-3 (SEQ ID NO: 1) o ADF-4 (SEQ ID NO: 2).

Una "unidad no repetitiva" que tiene una secuencia de aminoácidos que es "sustancialmente similar" a una secuencia de aminoácidos correspondiente no repetitiva (terminal carboxilo) dentro de un polipéptido de draglina natural (es decir, unidad no repetitiva de tipo silvestre (terminal carboxilo), preferiblemente dentro de ADF-3 (SEQ ID NO: 1) o ADF-4 (SEQ ID NO: 2), es también similar con respecto a sus propiedades, por ejemplo, una proteína de seda que comprende la "unidad no repetitiva sustancialmente similar" es todavía insoluble y retiene su insolubilidad y, por lo tanto, aún puede separarse de las partes insolubles de la célula huésped solubilizadas en la etapa c) del método de la presente invención. Preferiblemente, una proteína de seda que comprende la "unidad no repetitiva sustancialmente similar" es insoluble en una solución acuosa, por ejemplo, una solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada y permanece insoluble en más del 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o incluso hasta 100% en una solución acuosa, por ejemplo, solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o agua desionizada que comprende una base (por ejemplo, NaOH, 0,005 a 1 M) durante un periodo de 10 a 90 minutos, por ejemplo durante 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85 o 90 minutos y/o a una temperatura entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C. En realizaciones preferidas de la presente invención, una proteína de seda que comprende la "unidad no repetitiva sustancialmente similar", que permanece/queda insoluble, se separa/aísla de las partes insolubles de la célula huésped solubilizadas mediante filtración, preferiblemente usando una membrana filtrante que tiene un tamaño de poro de 10 nm a 200 nm, y/o por centrifugación, preferiblemente de 3000 a 12000 xg y más preferiblemente de 4000 a 8000 xg, durante 10 a 40 min, preferiblemente 20 a 40 min.

Más preferiblemente, la unidad no repetitiva (NR) es NR3 (SEQ ID NO: 41) o variantes de la misma, o NR4 (SEQ ID NO: 42) o variantes de la misma. La unidad NR3 (SEQ ID NO: 41) se basa en la secuencia de aminoácidos de ADF-3 de la araña *Araneus diadematus* y la unidad NR4 (SEQ ID NO: 42) se basa en la secuencia de aminoácidos de ADF-4 del araña *Araneus diadematus* (documento WO 2006/008163).

Una variante de la unidad NR3 o NR4 difiere de la unidad de referencia NR3 (SEQ ID NO: 41) o NR4 (SEQ ID NO: 42) de la cual se deriva por hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25 o 30 cambios de aminoácidos en la secuencia de aminoácidos (es decir, intercambios, inserciones, eliminaciones, truncamientos del terminal N y/o truncamientos del terminal C). Tal variante de la unidad NR3 o NR4 se puede caracterizar alternativa o adicionalmente por un cierto grado de identidad de secuencia con la unidad de referencia NR3 o NR4 de la que se deriva. Por lo tanto, una variante de unidad NR3 o NR4 tiene una identidad de secuencia de al menos 50%, 55%, 60%, 65%, 70%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o incluso 99,9% con la respectiva unidad de referencia NR3 o NR4. Preferiblemente, la identidad de secuencia está en un tramo continuo de al menos 10, 20, 30, 40, 50, 60, 70, 80, 90 o

más aminoácidos, preferiblemente en toda la longitud de la respectiva unidad de referencia NR3 o NR4.

Se prefiere particularmente que la identidad de secuencia sea al menos 80% en toda la longitud, sea al menos 85% en toda la longitud, sea al menos 90% en toda la longitud, sea al menos 95% en toda la longitud, sea al menos 98% en toda la longitud, o sea al menos 99% en toda la longitud de la respectiva unidad de referencia NR3 o NR4. Además, se prefiere particularmente que la identidad de secuencia sea al menos 80% en un tramo continuo de al menos 20, 30, 40, 50, 60, 70 u 80 aminoácidos, sea al menos 85% en un tramo continuo de al menos 20, 30, 40, 50, 60, 70 u 80 aminoácidos, sea al menos 90% en un tramo continuo de al menos 20, 30, 40, 50, 60, 70 u 80 aminoácidos, sea al menos 95% en un tramo continuo de al menos 20, 30, 40, 50, 60, 70 u 80 aminoácidos, sea al menos 98% en un tramo continuo de al menos 20, 30, 40, 50, 60, 70 u 80 aminoácidos, o sea al menos 99% en un tramo continuo de al menos 20, 30, 40, 50, 60, 70 u 80 aminoácidos de la respectiva unidad de referencia NR3 o NR4.

Un fragmento (o variante de eliminación) de una unidad NR3 o NR4 tiene preferiblemente una eliminación de hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55 o 60 aminoácidos en su terminal N y/o en su terminal C. La eliminación también puede ser internamente.

Además, la variante o fragmento de la unidad NR3 o NR4 solo se considera como una variante o fragmento de la unidad NR3 o NR4 dentro del contexto de la presente invención, si las modificaciones con respecto a la secuencia de aminoácidos en la que se encuentra la variante o fragmento no cambian la propiedad de una proteína de seda de ser insoluble y retener su insolubilidad, de modo que dicha proteína de seda todavía pueda separarse de las partes insolubles de la célula huésped solubilizadas en la etapa c) del método de la presente invención. Preferiblemente, una proteína de seda que comprende la variante o fragmento de la unidad NR3 o NR4 es insoluble en una solución acuosa, por ejemplo, solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada y permanece insoluble en más del 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o incluso hasta 100% en una solución acuosa, por ejemplo, solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o agua desionizada que comprende una base (por ejemplo, NaOH, 0,005 a 1 M) durante un período de 10 a 90 minutos, por ejemplo, durante 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85 o 90 minutos y/o a una temperatura entre 4 y 60 °C, por ejemplo, a 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C. En realizaciones preferidas de la presente invención, una proteína de seda que comprende la variante o fragmento de la unidad NR3 o NR4, que permanece/queda insoluble, se separa/aísla de las partes insolubles de la célula huésped solubilizadas mediante filtración, preferiblemente usando una membrana filtrante que tiene un tamaño de poro de 10 nm a 200 nm, y/o por centrifugación, preferiblemente de 3000 a 12000 xg y más preferiblemente de 4000 a 8000 xg, durante 10 a 40 min, preferiblemente 20 a 40 min.

Preferiblemente, la proteína de seda aislada con el método de la presente invención se selecciona del grupo que consta de ADF-3 (SEQ ID NO: 1) o variantes de la misma, ADF-4 (SEQ ID NO: 2) o variantes de la misma, MaSp I (SEQ ID NO: 43) o variantes de la misma, MaSp II (SEQ ID NO: 44) o variantes de la misma, $(C)_m$, $(C)_mNR_z$, $NR_z(C)_m$, $NR_z(C)_mNR_z$, $(AQ)_n$, $(AQ)_nNR_z$, $NR_z(AQ)_n$, $NR_z(AQ)_nNR_z$, $(QAQ)_o$, $NR_z(QAQ)_o$, $(QAQ)_oNR_z$, Y_p , X_p y K_p , en las que m es un número entero de 2 a 64 (es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67 o 68), n es un número entero de 6 a 24 (es decir, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 o 24), o es un número entero de 8 a 16 (es decir, 8, 9, 10, 11, 12, 13, 14, 15 o 16), p es un número entero de 8 a 16 (es decir, 8, 9, 10, 11, 12, 13, 14, 15 o 16) y z es un entero de 1 a 3 (es decir 1, 2 o 3) y NR representa una unidad no repetitiva. Las fórmulas mencionadas anteriormente se definen por uno de los siguiente: en la fórmula

(i) $(C)_m$, un número "m" de módulos C, es decir, de 2 a 64 módulos C, representados por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 21, se combinan entre sí,

(ii) $(C)_mNR_z$, un número "m" de módulos C, es decir, de 2 a 64 módulos C, representados por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 21, se combinan entre sí, en los que dichos módulos C se combinan además con un número "z" de unidades no repetitivas (NR), es decir, 1 a 3 unidades no repetitivas (NR), por ejemplo, las unidades NR3 no repetitivas (NR) representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42,

(iii) $NR_z(C)_m$, un número "z" de unidades no repetitivas (NR), es decir, 1 a 3 unidades no repetitivas (NR), por ejemplo, las unidades NR3 no repetitivas (NR) representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representada por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42, están presentes ($z = 1$) o se combinan entre sí ($z = 2$ o 3), en el que dicha unidad o unidades no repetitivas (NR) se combinan adicionalmente con un número "m" de módulos C, es decir, de 2 a 64 módulos C, representados por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 21,

(iv) $NR_z(C)_mNR_z$, un número "z" de unidades no repetitivas (NR), es decir, 1 a 3 unidades no repetitivas (NR), por ejemplo, las unidades NR3 no repetitivas (NR) representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representada por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42, están presentes ($z = 1$) o se combinan entre sí ($z = 2$ o 3), en el que dicha unidad o unidades no repetitivas (NR) se combinan con un número "m" de módulos C, es decir, de 2 a 64 módulos C, representados por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 21, y en el que dichos módulos C se combinan además con un número "z" de unidades no repetitivas (NR), es decir, 1 a 3 unidades no repetitivas (NR), por ejemplo las unidades NR3 no repetitivas (NR)

representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42,

(v) $(AQ)_n$, un número "n" de combinaciones de módulos A y Q, es decir, 6 a 24 combinaciones de módulos A y Q, en las que el módulo A está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 20 y el módulo

5 Q está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 22, se combinan entre sí,

(vi) $(AQ)_nNR_z$, un número "n" de combinaciones de módulos A y Q, es decir, 6 a 24 combinaciones de módulos A y Q, en las que el módulo A está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 20 y el módulo Q está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 22, se combinan entre sí, y en el que dichas combinaciones de módulos A y Q se combinan además con un número "z" de unidades no repetitivas (NR), es decir, 1 a 3 unidades no repetitivas (NR), por ejemplo, las unidades NR3 no repetitivas (NR) representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42,

(vii) $NR_z(AQ)_n$, un número "z" de unidades no repetitivas (NR), es decir, 1 a 3 unidades no repetitivas (NR), por ejemplo, las unidades NR3 no repetitivas (NR) representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representada por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42, está presente ($z = 1$) o se combinan entre sí ($z = 2$ o 3), en el que dicha unidad o unidades no repetitivas (NR) se combinan además con un número "n" de combinaciones de módulos A y Q, es decir, 6 a 24 combinaciones de módulos A y Q, en el que el módulo A está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 20 y el módulo Q está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 22,

(viii) $NR_z(AQ)_nNR_z$, un número "z" de unidades no repetitivas (NR), es decir, 1 a 3 unidades no repetitivas (NR), por ejemplo, las unidades NR3 no repetitivas (NR) representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representada por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42, está presente ($z = 1$) o se combinan entre sí ($z = 2$ o 3), en el que dicha unidad o unidades no repetitiva (NR) se combinan además con un número "n" de combinaciones de módulos A y Q, es decir, 6 a 24 combinaciones de módulos A y Q, en el que el módulo A está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 20 y el módulo Q está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 22, y en el que dichas combinaciones de módulos A y Q se combinan adicionalmente con un número "z" de unidades no repetitivas (NR), es decir, de 1 a 3 unidades no repetitivas (NR), por ejemplo, las unidades NR3 no repetitivas (NR) representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representada por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42,

(ix) $(QAQ)_o$, un número "o" de combinaciones de módulos Q, A y Q, es decir, 8 a 16 combinaciones de módulos Q, A y Q, en las que el módulo Q está representado por una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 22 y el módulo A está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 20, se combinan entre sí,

(x) $(QAQ)_oNR_z$, un número "o" de combinaciones de módulos Q, A y Q, es decir, 8 a 16 combinaciones de módulos Q, A y Q, en las que el módulo Q está representado por una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 22 y el módulo A está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 20, se combinan entre sí, y en el que dichas combinaciones de módulos Q, A y Q se combinan además con un número "z" de unidades no repetitivas (NR), es decir 1 a 3 unidades no repetitivas (NR), por ejemplo, las unidades NR3 no repetitivas (NR) representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42,

(xi) $NR_z(QAQ)_o$, un número "z" de unidades no repetitivas (NR), es decir, 1 a 3 unidades no repetitivas (NR), por ejemplo, las unidades NR3 no repetitivas (NR) representadas por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 41 o NR4 representada por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 42, está presente ($z = 1$) o se combinan entre sí ($z = 2$ o 3), en las que dicha unidad o unidades no repetitivas (NR) se combinan además con un número "o" de combinaciones de módulos Q, A y Q, es decir, 8 a 16 combinaciones de módulos A y Q, en las que el módulo Q está representado por una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 22 y el módulo A está representado por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 20,

(xii) Y_p , un número "p" de módulos Y, es decir, 8 a 16 módulos Y, representados por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 28, se combinan entre sí,

(xiii) X_p , un número "p" de módulos X, es decir 8 a 16 módulos X, representados por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 27, se combinan entre sí, y

(xiv) K_p , un número "p" de módulos K, es decir, 8 a 16 módulos K, representados por la secuencia de aminoácidos de acuerdo con la SEQ ID NO: 23, se combinan entre sí.

55 Más preferiblemente,

(i) z en $(C)_mNR_z$ o $NR_z(C)_m$ es 1 y m es un número entero de 2 a 64 (es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 o 64), z en $(C)_mNR_z$ o $NR_z(C)_m$ es 2 y m es un número entero de 2 a 64 (es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 o 64), o z en $(C)_mNR_z$ o $NR_z(C)_m$ es 3 y m es un número entero de 2 a 64 (es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 o 64),

(ii) z en $(AQ)_nNR_z$ o $NR_z(AQ)_n$ es 1 y n es un número entero de 6 a 24 (es decir, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

17, 18, 19, 20, 21, 22, 23 o 24), z en $(AQ)_nNR_z$ o $NR_z(AQ)_n$ es 2 y n es un número entero de 6 a 24 (es decir, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 o 24), o z en $(AQ)_nNR_z$ o $NR_z(AQ)_n$ es 3 y n es un número entero de 6 a 24 (es decir, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 o 24),

(iii) z en $NR_z(QAQ)_o$ o $(QAQ)_oNR_z$ es 1 y o es un número entero de 8 a 16 (es decir, 8, 9, 10, 11, 12, 13, 14, 15 o 16), z en $NR_z(QAQ)_o$ o $(QAQ)_oNR_z$ es 2 y o es un número entero de 8 a 16 (es decir, 8, 9, 10, 11, 12, 13, 14, 15 o 16), o z en $NR_z(QAQ)_o$ o $(QAQ)_oNR_z$ es 3 y o es un número entero de 8 a 16 (es decir, 8, 9, 10, 11, 12, 13, 14, 15 o 16),

(iv) z en $NR_z(C)_mNR_z$ es 1 y m es un número entero de 2 a 64 (es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 o 64), z en $NR_z(C)_mNR_z$ es 2 y m es un número entero de 2 a 64 (es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 o 64), o z en $NR_z(C)_mNR_z$ es 3 y m es un número entero de 2 a 64 (es decir, 2, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 o 64), o

(v) z en $NR_z(AQ)_nNR_z$ es 1 y n es un número entero de 6 a 24 (es decir, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 o 24), z en $NR_z(AQ)_nNR_z$ es 2 y n es un número entero de 6 a 24 (es decir, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 o 24), o z en $NR_z(AQ)_nNR_z$ es 3 y n es un número entero de 6 a 24 (es decir, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 o 24),

en la que NR representa una unidad no repetitiva, preferiblemente NR3 o NR4.

Más preferiblemente, la proteína de seda aislada con el método de la presente invención es $C_{16}NR_4$, $C_{32}NR_4$, $(AQ)_{12}NR_3$, $(AQ)_{24}NR_3$, $(AQ)_{12}$, $(AQ)_{24}$, C_{16} , C_{32} , $NR_4C_{16}NR_4$, $NR_4C_{32}NR_4$, $NR_3C_{16}NR_3$, $NR_3C_{32}NR_3$, $NR_4(AQ)_{12}NR_4$, $NR_4(AQ)_{24}NR_4$, $NR_3(AQ)_{12}NR_3$, $NR_3(AQ)_{24}NR_3$, $(QAQ)_8$ o $(QAQ)_{16}$.

Una variante de ADF-3, ADF-4, MaSp I o MaSp II difiere del polipéptido de referencia ADF-3 (SEQ ID NO: 1), ADF-4 (SEQ ID NO: 2), MaSp I (SEQ ID NO: 43) o MaSp II (SEQ ID NO: 44) (tipo silvestre) del cual se deriva por hasta 150 (hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140 o 150) cambios de aminoácidos en la secuencia de aminoácidos (es decir, sustituciones, inserciones, eliminaciones, truncamientos del terminal N y/o truncamientos del terminal C). Dicha variante puede caracterizarse de manera alternativa o adicional por un cierto grado de identidad de secuencia con el polipéptido de referencia (tipo silvestre) del que se deriva. Por lo tanto, una variante de ADF-3, ADF-4, MaSp I o MaSp II tiene una identidad de secuencia de al menos 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o incluso 99,9% con los respectivos polipéptidos ADF-3, ADF-4, MaSp I o MaSp II de referencia (tipo silvestre). Preferiblemente, la identidad de secuencia está en un tramo continuo de al menos 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 120, 150, 180, 200, 250, 300, 350, 400, o más aminoácidos, preferiblemente a lo largo de toda la longitud del respectivo polipéptido ADF-3, ADF-4, MaSp I o MaSp II de referencia (tipo silvestre).

Se prefiere particularmente que la identidad de la secuencia sea al menos del 80% en toda la longitud, sea al menos del 85% en toda la longitud, sea al menos del 90% en toda la longitud, sea al menos del 95% en toda la longitud, sea al menos del 98% en toda la longitud, o sea al menos del 99% en toda la longitud del respectivo polipéptido ADF-3, ADF-4, MaSp I o MaSp II de referencia (tipo silvestre). Además, se prefiere particularmente que la identidad de secuencia sea al menos del 80% en un tramo continuo de al menos 20, 30, 50, 100, 150, 200, 250 o 300 aminoácidos, sea al menos del 85% en un tramo continuo de al menos 20, 30, 50, 100, 150, 200, 250 o 300 aminoácidos, sea al menos del 90% en un tramo continuo de al menos 20, 30, 50, 100, 150, 200, 250 o 300 aminoácidos, sea al menos del 95% en un tramo continuo de al menos 20, 30, 50, 100, 150, 200, 250 o 300 aminoácidos, sea al menos del 98% en un tramo continuo de al menos 20, 30, 50, 100, 150, 200, 250 o 300 aminoácidos, o sea al menos del 99% en un tramo continuo de al menos 20, 30, 50, 100, 150, 200, 250 o 300 aminoácidos del respectivo polipéptido ADF-3, ADF-4, MaSp I o MaSp II (tipo silvestre).

Un fragmento (o variante de eliminación) del polipéptido ADF-3 (SEQ ID NO: 1) tiene preferiblemente una eliminación de hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 120, 150, 170, 200, 220, 250, 270, 300, 320, 350, 370, 400, 420, 450, 470, 500, 520, 550, 570, 600, o 610 aminoácidos en su terminal N y/o en su terminal C. La eliminación también puede ser internamente.

Un fragmento (o variante de eliminación) del polipéptido ADF-4 (SEQ ID NO: 2) tiene preferiblemente una eliminación de hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 120, 150, 170, 200, 220, 250, 270, 300, 320, 330, 340, 350, 360, 370, 380 o 390 aminoácidos en su terminal N y/o en su terminal C. La eliminación también puede ser internamente.

Un fragmento (o variante de eliminación) del polipéptido MaSp I (SEQ ID NO: 43) tiene preferiblemente una eliminación de hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 620, 640, 660, 670, 680 o 690 aminoácidos en su terminal N y/o en su terminal C. La eliminación también puede ser internamente.

Un fragmento (o variante de eliminación) del polipéptido MaSp II (SEQ ID NO: 44) tiene preferiblemente una eliminación de hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 520, 540, 560 o 570 aminoácidos en su terminal N y/o en su terminal C. La eliminación también puede ser internamente.

- 5 Además, la variante o fragmento de ADF-3, ADF-4, MaSp I o MaSp II solo se considera como una variante o fragmento de ADF-3, ADF-4, MaSp I o MaSp II dentro del contexto de la presente invención, si las modificaciones con respecto a la secuencia de aminoácidos en la que se basa la variante o fragmento no cambian la propiedad de la proteína de seda de ser insoluble y de retener su insolubilidad, de modo que dicha proteína de seda todavía puede separarse de las partes insolubles de la célula huésped solubilizadas en la etapa c) del método de la presente invención.
- 10 Preferiblemente, la variante o fragmento de ADF-3, ADF-4, MaSp I o MaSp II todavía es insoluble en una solución acuosa, por ejemplo, solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o en agua desionizada y permanece insoluble en más del 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99,9% o incluso hasta 100% en una solución acuosa, por ejemplo, solución acuosa tamponada, en un medio de cultivo, en un medio de fermentación, en agua técnica o agua desionizada que comprende una base (por ejemplo, NaOH, 0,005 a 1 M) durante un período de 10 a 90 minutos, por ejemplo durante 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85 o 90 minutos y/o a una temperatura entre 4 y 60 °C, por ejemplo, a los 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55 o 60 °C. En realizaciones preferidas de la presente invención, la variante o fragmento de ADF-3, ADF-4, MaSp I o MaSp II, que permanece/queda insoluble, se separa/aísla de las partes insolubles de la célula huésped solubilizadas mediante
- 20 filtración, preferiblemente usando un filtro de membrana que tiene un tamaño de poro de 10 nm a 200 nm, y/o por centrifugación, preferiblemente de 3000 a 12000 x g y más preferiblemente de 4000 a 8000 x g, durante 10 a 40 min, preferiblemente 20 a 40 min.

En otra realización, la proteína de seda comprende además un TAG en el terminal amino y/o en el terminal carboxilo seleccionado del grupo que consta de

- 25 (i) TAG^{CYS1} que consta de la secuencia de aminoácidos GCGGGGGGSGGGG (SEQ ID NO: 35),
 (ii) TAG^{CYS2} que consta de la secuencia de aminoácidos GCGGGGGG (SEQ ID NO: 36),
 (iii) TAG^{CYS3} que consta de la secuencia de aminoácidos GCGGSGGGGSGGGG (SEQ ID NO: 37),
 (iv) TAG^{LYS1} que consta de la secuencia de aminoácidos GKGGGGGGSGGGG (SEQ ID NO: 38), y
 (v) TAG^{LYS2} que consta de la secuencia de aminoácidos GKGGGGGG (SEQ ID NO: 39).
- 30 Estos TAG contienen cisteína y/o lisina y pueden usarse para unir covalentemente sustancias a dicha proteína de seda después de su aislamiento. Preferiblemente, la sustancia unida/acoplada covalentemente se selecciona del grupo que consta de un polipéptido, un lípido, un tinte, un metal conjugado, carbón activado y un agente.

Lo más preferiblemente, el polipéptido de seda aislado con el método de la presente invención comprende o consta de TAG^{CYS1}C₁₆, C₁₆TAG^{CYS1}, TAG^{CYS1}C₁₆TAG^{CYS1}, TAG^{CYS2}C₁₆, C₁₆TAG^{CYS2}, TAG^{CYS2}C₁₆TAG^{CYS2}, TAG^{CYS3}C₁₆, C₁₆TAG^{CYS3}, TAG^{CYS3}C₁₆TAG^{CYS3}, TAG^{LYS1}C₁₆, C₁₆TAG^{LYS1}, TAG^{LYS1}C₁₆TAG^{LYS1}, TAG^{LYS2}C₁₆, C₁₆TAG^{LYS2} o TAG^{LYS2}C₁₆TAG^{LYS2}.

35

En un segundo aspecto, se describe una proteína objetivo insoluble obtenible por el método del primer aspecto.

En otro aspecto, la invención proporciona un método para aislar una proteína objetivo insoluble de una solución de proteína que comprende los etapas de:

- 40 a) proporcionar una solución de proteína que comprende la proteína objetivo insoluble,
 b) agregar una solución acuosa de al menos una base a dicha solución en una cantidad que sea suficiente para romper dichas células huésped y/o solubilizar dichas partes insolubles de la célula huésped, y
 c) separar la proteína objetivo insoluble de las partes insolubles de la célula huésped solubilizadas,

45 en el que en la etapa b) la proteína objetivo permanece insoluble y al menos el 80% de las partes insolubles de la célula huésped se solubilizan.

Preferiblemente, la solución de proteína es leche de un animal transgénico, por ejemplo, cabra, oveja o ganado, que expresa la proteína objetivo insoluble. En cuanto a las definiciones, la selección de la base añadida en la etapa b), los reactivos utilizados, las condiciones y parámetros del proceso y la proteína objetivo aislable, se mencionan en las explicaciones hechas anteriormente con respecto al primer aspecto de la invención.

- 50 Breve descripción de las tablas y las figuras

Fig. 1: Visualización de la separación de fases. Los residuos celulares están completamente disueltos, mientras que la proteína objetivo C₁₆ permanece insoluble en la fracción sedimentada. La adición de diferentes agentes afecta el resultado de la separación de fases del sedimento de proteína objetivo C₁₆. La Figura 1 muestra la purificación en una etapa como se describe en el Ejemplo 2. Los residuos celulares se incubaron durante 45 minutos con A: NaOH 0,05 M; B: NaOH 0,05 M y de urea 4,8 M, C NaOH : 0,05 M y Tween 20 al 0,1%, D: urea 4,8 M, E: clorhidrato de guanidina 2 M, F: Tween 20 al 0.1%. Después de la incubación y una etapa siguiente de centrifugación, los sedimentos de A, B

55

y C muestran un gránulo blanco que consiste casi por completo en proteína objetivo C₁₆ de alta pureza. En las muestras D, E y F, no se observó separación aguda de fases.

Figura 2: Análisis de inmunotransferencia (transferencia Western) de la proteína C₁₆ recombinante de seda de araña purificada de acuerdo con el Ejemplo 2. El carril 2, cargado con C₁₆, muestra una señal de proteína única (control positivo de C₁₆ en el carril 1). No se pudieron observar degradaciones de proteínas. La detección específica de la proteína objetivo está mediada por el péptido de etiqueta T7 conjugado de acuerdo con el protocolo adjunto de SERVAGel^{MR} TG10 (43210). En detalle: Carril 1: 0,5 µg de control positivo C₁₆ (purificación de acuerdo con Huemmerich et. al., Biochemistry 2004, 43, 13604-13612); Carril 2: 5 µg de C₁₆ (purificación de acuerdo con el ejemplo 2). Anticuerpo: conjugado de HRP de anticuerpo T7-Tag®, Catálogo No: 69048-3, Novagen; membrana de PVDF: Roti® Fluoro-PVDF, PVM020C-099, Roth; La inmunotransferencia (transferencia Western) se realizó de acuerdo con el protocolo adjunto de SERVAGel^{MR} TG10 (43210). Solución de bloqueo: leche en polvo al 5% en tampón de PBS 1x, incubación 30 min; la detección se realizó con solución de Lumi-light^{Plus}, 12015196001, Roche, de acuerdo con el protocolo adjunto.

Figura 3: El análisis de SDS-Page (teñido con plata) de una proteína recombinante C₁₆ de seda de araña purificada de acuerdo con el proceso descrito en el Ejemplo 2 (carril 2), cargada con 10 µg de proteína C₁₆, muestra una única señal de proteína (control positivo de proteína C₁₆ en el carril 1). De acuerdo con la Figura 1, no se pudieron observar contaminaciones. Carril 1: control positivo 10 µg de C₁₆ (purificación de acuerdo con Huemmerich et. al., Biochemistry 2004, 43, 13604-13612); Carril 2: 10 µg de C₁₆ (purificación como se describe en el Ejemplo 2). SDS-Page se realizó con SERVAGel^{MR} TG10 (43210) de acuerdo con el protocolo adjunto. La tinción se realizó con el kit de tinción de plata SERVA (35076) de acuerdo con el protocolo adjunto.

Figura 4: Análisis por HPLC de una proteína recombinante C₁₆ de seda de araña purificada de acuerdo con el Ejemplo 2. El cromatograma muestra un único pico agudo a 15,057 min (identificado como proteína objetivo C₁₆, detectada a una longitud de onda de 280 nm). No se pudieron observar contaminaciones. Parámetros de columna: columna de HPLC G3000 SWXL, TSK-Gel® Tosoh Bioscience, No. 08541, diámetro interior: 7,8 mm, longitud: 30,0 cm. El caudal fue constante a 0,5 mL/min.

Figura 5: El método de purificación de Huemmerich et al., Biochemistry 2004, 43, 13604-13612 y el método de purificación de acuerdo con el Ejemplo 2 se compararon con respecto a la cantidad de proteína C₁₆ purificada (proteína objetivo). Se podría demostrar que el rendimiento de la proteína objetivo C₁₆ es significativamente mayor (aproximadamente nueve veces) cuando se purifica de acuerdo con el método de la invención, véase el Ejemplo 2 (89%), en comparación con el método de Huemmerich et al., Biochemistry 2004, 43, 13604-13612 (10,4%). La Figura 5 muestra el rendimiento de la proteína C₁₆ objetivo purificada de acuerdo con el Ejemplo 2. Ambos métodos de purificación (Huemmerich et. al., Biochemistry 2004, 43, 13604-13612) y el método de purificación como se describe en el Ejemplo 2 se compararon con respecto a la cantidad de proteína C₁₆ purificada.

Figura 6: Ensayo de estabilidad de la proteína C₁₆ recombinante purificada de seda de araña tratada con hidróxido de sodio 0,05 M como se describe en el Ejemplo 2. No se pudo observar una degradación proteica significativa después de la incubación con 0,05 M durante un período de tiempo de entre 5 min a 24 horas. La identificación de la proteína objetivo está mediada a través del detector de fluorescencia en procesos de HPLC (longitud de onda de excitación: 275 nm - longitud de onda de emisión: 305 nm; gradiente: isocrático, eluyente: Tris/HCl 100 mM, pH 7,5). Se tomaron muestras en diferentes puntos de tiempo de incubación a los 5 min, 10 min, 20 min, 30 min, 1 h, 3 h, 7 h y 24 h. Los datos de cada cromatograma se estandarizaron al 100% para una mejor detección de diferencias menores.

Ejemplos

Para separar las proteínas objetivo, los inventores modificaron a título de ejemplo el polipéptido C₁₆ de seda sintética que se deriva de la proteína de seda de dragalina ADF-4 de la araña *Araneus diadematus* a través del jardín europeo. La proteína se eligió con base en las observaciones previas de que ADF-3 y/o ADF-4, así como sus variantes, muestran un comportamiento de ensamblaje eficiente.

Ejemplo I: Preparación de una suspensión de células microbianas

Las células microbianas se lisaron/rompieron dos veces usando un homogeneizador [APV Gaulin GmbH, LAB60-15 RBF] a una presión de 900 bares a temperatura ambiente. La solución resultante de residuos celulares se centrifugó a 9000 xg durante 30 min, 20 °C. Se descartó el sobrenadante y se utilizó el sedimento de residuos celulares para la purificación en una sola etapa.

Ejemplo II: Aislamiento de C₁₆

Se determinó el peso húmedo del sedimento de residuos celulares de la preparación de la solución de células microbianas (Ejemplo 1) para ajustar el volumen de dilución requerido. La proporción de sedimento de residuos celulares (peso húmedo) y NaOH 0,05 M fue 1:3 (p/p). En este ejemplo, se midió que el sedimento de residuos celulares era de 5 gramos. Por lo tanto, el peso de los aditivos de muestra se calculó en 15 gramos. Los aditivos en este experimento fueron A: NaOH 0,05 M; B: NaOH 0,05 M y urea 4,8 M, C: NaOH 0,05 M y Tween 20 al 0,1%, D: urea 4,8 M, E: clorhidrato de guanidina 2 M o F: Tween 20 al 0,1%. Todas las muestras se incubaron durante 45 minutos a 20 °C bajo buena agitación. Después de la incubación, las muestras se centrifugaron a 8000 xg durante 30 minutos, 20 °C (véase la Figura 1). Los sedimentos de A, B y C mostraron un gránulo blanco que consiste casi por completo en proteína objetivo C₁₆ de alta pureza. En las muestras D, E y F, no se observó separación aguda de fases.

Después de esta etapa, la proteína C₁₆ agregada (proteína objetivo) estaba presente en la fracción del sedimento, mientras que el sobrenadante se ha descartado. Se realizaron dos etapas de lavado adicionales con H₂O (centrifugación a 8000 x g, 30 min, 20 °C) para aumentar la pureza de la proteína objetivo C₁₆. La proteína objetivo C₁₆ así obtenida se analizó en un análisis de SDS-Page (tinción de plata) [La tinción se realizó con el kit de tinción de plata SERVA (35076) de acuerdo con el protocolo adjunto] (véase la Figura 3). El análisis de inmunotransferencia (transferencia Western) [el análisis de inmunotransferencia (transferencia Western) se realizó de acuerdo con el protocolo adjunto de SERVAGel^{MR} TG10 (43210). Solución de bloqueo: leche en polvo al 5% en tampón de PBS 1x, incubación 30 min; la detección se realizó con solución de Lumi-light^{Plus}, 12015196001, Roche, de acuerdo con el protocolo adjunto] (véase la Figura 2) y análisis de HPLC [Parámetros de columna: columna de HPLC TSK-Gel® G3000 SWXL, Tosoh Bioscience, No. 08541, diámetro interior: 7,8 mm, longitud: 30,0 cm. El caudal fue constante a 0,5 mL/min] (Véase la Figura 4).

Ejemplo III: Aislamiento de C₁₆ en comparación con el proceso estándar:

Se determinó el peso húmedo del sedimento de residuos celulares de la preparación de células microbianas para ajustar el volumen de dilución requerido. La proporción de sedimento de residuos celulares (peso húmedo) y NaOH 0,05 M fue de uno a tres (p/p). En este experimento, se midió que el sedimento de residuos celulares era de un kilogramo. Por lo tanto, el peso de NaOH 0,05 M se calculó en 3 kg (peso total: 4 kg). La muestra se incubó durante 45 minutos a 20 °C bajo agitación. Después de la incubación, la muestra se centrifugó a 8000 xg durante 30 min, 20 °C. Después de esta etapa, la proteína objetivo C₁₆ estaba presente en la fracción del sedimento, mientras que el sobrenadante se había descartado. Dos etapas de lavado adicionales con H₂O (centrifugación a 8000 xg, 30 min, 20 °C) aumentaron la pureza de la proteína objetivo C₁₆. En este ejemplo, la cantidad de proteína objetivo C₁₆ purificada se comparó con el método de purificación descrito por Huemmerich et al., Biochemistry 2004, 43, 13604-13612. En contraste con el método de la invención, la purificación de acuerdo con Huemmerich et al., se realizó con el sobrenadante después del rompimiento de las células de la preparación de células microbianas.

Tabla 1:

Método de purificación	Rendimiento [g de C ₁₆ /kg de solución de células microbianas]
Huemmerich et al., Biochemistry 2004, 43, 13604-13612	0,20
Experimento [-Ib-] 1,72	1,72

Con respecto a la cantidad de proteína objetivo C₁₆ purificada, se podría demostrar que el rendimiento de la proteína objetivo C₁₆ fue significativamente mayor (aproximadamente nueve veces) cuando se purificó como se describe en el Ejemplo 2 (véase la Tabla 1, Figura 5).

Ejemplo IV: Estabilidad de la proteína de C₁₆

Una solución de células microbianas que contenía la proteína objetivo C₁₆ se trató con NaOH 0,05 M. Como se describe en los Ejemplos 1 y 2, las células microbianas se lisaron dos veces usando un homogeneizador [APV Gaulin GmbH, LAB60-15 RBF] a una presión de 900 bares a temperatura ambiente. La solución de residuos celulares resultante se centrifugó a 9000 xg durante 30 min, 20 °C. Se descartó el sobrenadante y se utilizó el sedimento de residuos celulares para la purificación en una sola etapa. El peso húmedo del sedimento de residuos celulares de la preparación de la solución de células microbianas (Ejemplo 1) se determinó para ajustar el volumen de dilución requerido. La proporción de sedimento de residuos celulares (peso húmedo) y NaOH 0,05 M fue 1:3 (p/p).

Se tomaron muestras en diferentes momentos de incubación a los 5 min, 10 min, 20 min, 30 min, 1 h, 3 h, 7 h y 24 h. Los datos de cada cromatograma se estandarizaron al 100% para una mejor detección de diferencias menores. Como se muestra en la Figura 6, no tiene lugar una degradación proteica significativa después de la incubación con NaOH 0,05 M durante un período de tiempo entre 5 min y 24 horas. Parámetros del instrumento: columna de HPLC TSK-Gel® G3000 SWXL, diámetro interno: 7,8 mm, longitud: 30,0 cm. Caudal: 0,5 ml/min, Detección de fluorescencia: Longitud de onda de excitación: 275 nm - Longitud de onda de emisión: 305 nm; Gradiente: isocrático, eluyente: Tris/HCl 100 mM, pH 7,5.

Listado de secuencias

<110> AMSilk GmbH, et al.

<120> Separación de proteínas objetivo insolubles

<130> 558-36 PCT

<150> US 61/319.542

<151> 2010-03-31

<160> 44

<170> PatentIn versión 3.5
 <210> 1
 <211> 636
 <212> PRT
 5 <213> Araneus diadematus

<220>
 <221> PÉPTIDO
 <222> (1)..(636)
 10 <223> ADF-3

<400> 1

Ala Arg Ala Gly Ser Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly
 1 5 10 15

Gln Gln Gly Pro Gly Gln Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala
 20 25 30

Ala Ala Ala Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly Gln Gln Gly
 35 40 45

Pro Ser Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gly Gln Gly Pro
 50 55 60

Tyr Gly Pro Gly Ala Ser Ala Ala Ala Ala Ala Ala Gly Gly Tyr Gly
 65 70 75 80

Pro Gly Ser Gly Gln Gln Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro
 85 90 95

Gly Ser Ser Ala Ala Ala Ala Ala Ala Gly Gly Asn Gly Pro Gly Ser
 100 105 110

Gly Gln Gln Gly Ala Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly
 115 120 125

Ala Ser Ala Ala Ala Ala Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly
 130 135 140

15

ES 2 799 431 T3

Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gly Gln Gly Pro Tyr Gly
 145 150 155 160

Pro Gly Ala Ser Ala Ala Ala Ala Ala Ala Gly Gly Tyr Gly Pro Gly
 165 170 175

Ser Gly Gln Gly Pro Gly Gln Gln Gly Pro Gly Gly Gln Gly Pro Tyr
 180 185 190

Gly Pro Gly Ala Ser Ala Ala Ala Ala Ala Ala Gly Gly Tyr Gly Pro
 195 200 205

Gly Ser Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly
 210 215 220

Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ala Ala
 225 230 235 240

Ala Ala Gly Gly Tyr Gly Pro Gly Tyr Gly Gln Gln Gly Pro Gly Gln
 245 250 255

Gln Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala
 260 265 270

Ser Ala Ala Ser Gly Gly Tyr Gly Pro Gly Ser Gly Gln Gln Gly Pro
 275 280 285

Gly Gln Gln Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ala Ser
 290 295 300

Ala Ala Ala Ala Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly Gln Gln
 305 310 315 320

Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly
 325 330 335

Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ala Ala
 340 345 350

Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly Gln Gln Gly Pro Gly Gln
 355 360 365

Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln
 370 375 380

Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly
 385 390 395 400

ES 2 799 431 T3

Pro Gly Gln Gln Gly Pro Gly Gly Gln Gly Ala Tyr Gly Pro Gly Ala
 405 410 415

Ser Ala Ala Ala Gly Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly Gln
 420 425 430

Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln
 435 440 445

Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly Pro Gly Gln Gln Gly
 450 455 460

Pro Gly Gln Gln Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ala Ala
 465 470 475 480

Ala Ala Gly Gly Tyr Gly Pro Gly Ser Gly Gln Gln Gly Pro Gly Gln
 485 490 495

Gln Gly Pro Gly Gln Gln Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro
 500 505 510

Gly Ala Ala Ser Ala Ala Val Ser Val Gly Gly Tyr Gly Pro Gln Ser
 515 520 525

Ser Ser Val Pro Val Ala Ser Ala Val Ala Ser Arg Leu Ser Ser Pro
 530 535 540

Ala Ala Ser Ser Arg Val Ser Ser Ala Val Ser Ser Leu Val Ser Ser
 545 550 555 560

Gly Pro Thr Lys His Ala Ala Leu Ser Asn Thr Ile Ser Ser Val Val
 565 570 575

Ser Gln Val Ser Ala Ser Asn Pro Gly Leu Ser Gly Cys Asp Val Leu
 580 585 590

Val Gln Ala Leu Leu Glu Val Val Ser Ala Leu Val Ser Ile Leu Gly
 595 600 605

Ser Ser Ser Ile Gly Gln Ile Asn Tyr Gly Ala Ser Ala Gln Tyr Thr
 610 615 620

Gln Met Val Gly Gln Ser Val Ala Gln Ala Leu Ala
 625 630 635

<210> 2

5 <211> 410

ES 2 799 431 T3

<212> PRT
 <213> Araneus diadematus

<220>
 <221> PÉPTIDO
 <222> (1)..(410)
 <223> ADF-4

<400> 2

10

Ala Gly Ser Ser Ala Ala Ala Ala Ala Ala Ala Ser Gly Ser Gly Gly
 1 5 10 15

Tyr Gly Pro Glu Asn Gln Gly Pro Ser Gly Pro Val Ala Tyr Gly Pro
 20 25 30

Gly Gly Pro Val Ser Ser Ala Ala Ala Ala Ala Ala Ala Gly Ser Gly
 35 40 45

Pro Gly Gly Tyr Gly Pro Glu Asn Gln Gly Pro Ser Gly Pro Gly Gly
 50 55 60

Tyr Gly Pro Gly Gly Ser Gly Ser Ser Ala Ala Ala Ala Ala Ala Ala
 65 70 75 80

Ala Ser Gly Pro Gly Gly Tyr Gly Pro Gly Ser Gln Gly Pro Ser Gly
 85 90 95

Pro Gly Gly Ser Gly Gly Tyr Gly Pro Gly Ser Gln Gly Ala Ser Gly
 100 105 110

Pro Gly Gly Pro Gly Ala Ser Ala Ala Ala Ala Ala Ala Ala Ala Ala
 115 120 125

Ala Ser Gly Pro Gly Gly Tyr Gly Pro Gly Ser Gln Gly Pro Ser Gly
 130 135 140

Pro Gly Ala Tyr Gly Pro Gly Gly Pro Gly Ser Ser Ala Ala Ala Ala
 145 150 155 160

Ala Ala Ala Ala Ser Gly Pro Gly Gly Tyr Gly Pro Gly Ser Gln Gly
 165 170 175

Pro Ser Gly Pro Gly Val Tyr Gly Pro Gly Gly Pro Gly Ser Ser Ala
 180 185 190

Ala Ala Ala Ala Ala Ala Gly Ser Gly Pro Gly Gly Tyr Gly Pro Glu
 195 200 205

ES 2 799 431 T3

Asn Gln Gly Pro Ser Gly Pro Gly Gly Tyr Gly Pro Gly Gly Ser Gly
 210 215 220

Ser Ser Ala Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gly Tyr
 225 230 235 240

Gly Pro Gly Ser Gln Gly Pro Ser Gly Pro Gly Gly Ser Gly Gly Tyr
 245 250 255

Gly Pro Gly Ser Gln Gly Gly Ser Gly Pro Gly Ala Ser Ala Ala Ala
 260 265 270

Ala Ala Ala Ala Ala Ser Gly Pro Gly Gly Tyr Gly Pro Gly Ser Gln
 275 280 285

Gly Pro Ser Gly Pro Gly Tyr Gln Gly Pro Ser Gly Pro Gly Ala Tyr
 290 295 300

Gly Pro Ser Pro Ser Ala Ser Ala Ser Val Ala Ala Ser Val Tyr Leu
 305 310 315 320

Arg Leu Gln Pro Arg Leu Glu Val Ser Ser Ala Val Ser Ser Leu Val
 325 330 335

Ser Ser Gly Pro Thr Asn Gly Ala Ala Val Ser Gly Ala Leu Asn Ser
 340 345 350

Leu Val Ser Gln Ile Ser Ala Ser Asn Pro Gly Leu Ser Gly Cys Asp
 355 360 365

Ala Leu Val Gln Ala Leu Leu Glu Leu Val Ser Ala Leu Val Ala Ile
 370 375 380

Leu Ser Ser Ala Ser Ile Gly Gln Val Asn Val Ser Ser Val Ser Gln
 385 390 395 400

Ser Thr Gln Met Ile Ser Gln Ala Leu Ser
 405 410

5 <210> 3
 <211> 5
 <212> PRT
 <213> Secuencia Artificial

10 <220>
 <223> sintética

<220>
 <221> REPETICIÓN

ES 2 799 431 T3

<222> (1)..(5)
 <223> motivo de péptido de consenso

5 <400> 3

Gly Pro Gly Xaa Xaa
 1 5

<210> 4
 <211> 5
 10 <212> PRT
 <213> Araneus diadematus

<220>
 <221> VARIANTE
 15 <222> (4)..(4)
 <223> Q en la posición 4 también puede ser alanina, serina, glicina, tirosina, prolina, o glutamina

<220>
 <221> VARIANTE
 20 <222> (5)..(5)
 <223> Q en la posición 5 también puede ser alanina, serina, glicina, tirosina, prolina, or glutamina

<400> 4

Gly Pro Gly Gln Gln
 1 5

25 <210> 5
 <211> 5
 <212> PRT
 30 <213> Araneus diadematus

<220>
 <221> REPETICIÓN
 <222> (1)..(5)
 35 <223> motivo peptídico (ADF-3)

<400> 5

Gly Pro Gly Ala Ser
 1 5

40 <210> 6
 <211> 5
 <212> PRT
 <213> Araneus diadematus

45 <220>
 <221> REPETICIÓN
 <222> (1)..(5)
 <223> motivo peptídico (ADF-3)

50 <400> 6

Gly Pro Gly Ser Gly
 1 5

55 <210> 7
 <211> 5
 <212> PRT
 <213> Araneus diadematus

60 <220>
 <221> REPETICIÓN

<222> (1)..(5)
 <223> motivo peptídico (ADF-4)

5 <400> 7

Gly Pro Gly Gly Tyr
1 5

<210> 8
 <211> 5
 10 <212> PRT
 <213> Araneus diadematus

<220>
 <221> REPETICIÓN
 15 <222> (1)..(5)
 <223> motivo peptídico (ADF-4)

20 <400> 8

Gly Pro Gly Gly Pro
1 5

<210> 9
 <211> 5
 25 <212> PRT
 <213> Nephila clavipes

<220>
 <221> REPETICIÓN
 30 <222> (1)..(5)
 <223> motivo peptídico (proteína flageliforme)

35 <400> 9

Gly Pro Gly Gly Ala
1 5

<210> 10
 <211> 5
 40 <212> PRT
 <213> Arthropoda

<220>
 <221> REPETICIÓN
 45 <222> (1)..(5)
 <223> motivo peptídico (resilina)

50 <400> 10

Gly Pro Gly Gly Gly
1 5

<210> 11
 <211> 5
 50 <212> PRT
 <213> Nephila clavipes

<220>
 <221> REPETICIÓN
 55 <222> (1)..(5)
 <223> motivo peptídico (proteína flageliforme)

60 <400> 11

Gly Pro Gly Gly Ser
1 5

ES 2 799 431 T3

<210> 12
<211> 5
<212> PRT
<213> Secuencia Artificial

5
<220>
<223> sintética

<220>
10 <221> REPETICIÓN
<222> (1)..(5)
<223> motivo peptídico Ax

<400> 12

15
Ala Ala Ala Ala Ala
1 5

<210> 13
<211> 6
20 <212> PRT
<213> Araneus diadematus

<220>
<221> REPETICIÓN
25 <222> (1)..(6)
<223> motivo peptídico Ax (ADF 3)

<400> 13

30
Ala Ala Ala Ala Ala Ala
1 5

<210> 14
<211> 7
35 <212> PRT
<213> Araneus diadematus

<220>
<221> REPETICIÓN
40 <222> (1)..(7)
<223> motivo peptídico Ax (ADF-4)

<400> 14

45
Ala Ala Ala Ala Ala Ala Ala
1 5

<210> 15
<211> 8
50 <212> PRT
<213> Araneus diadematus

<220>
<221> REPETICIÓN
<222> (1)..(8)
<223> motivo peptídico Ax (ADF-4)

55
<400> 15

60
Ala Ala Ala Ala Ala Ala Ala Ala
1 5

<210> 16
<211> 9
<212> PRT
<213> Secuencia Artificial

- <220>
<223> sintética
- 5 <220>
<221> REPETICIÓN
<222> (1)..(9)
<223> motivo peptídico Ax
- 10 <400> 16
- Ala Ala Ala Ala Ala Ala Ala Ala Ala**
1 5
- 15 <210> 17
<211> 10
<212> PRT
<213> Araneus diadematus
- 20 <220>
<221> REPETICIÓN
<222> (1)..(10)
<223> motivo peptídico Ax (ADF-4)
- 25 <400> 17
- Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala**
1 5 10
- 30 <210> 18
<211> 9
<212> PRT
<213> Arthropoda
- 35 <220>
<221> REPETICIÓN
<222> (1)..(9)
<223> motivo peptídico (basado en resilina)
- <400> 18
- Gly Gly Arg Pro Ser Asp Thr Tyr Gly**
1 5
- 40 <210> 19
<211> 9
<212> PRT
<213> Arthropoda
- 45 <220>
<221> REPETICIÓN
<222> (1)..(9)
<223> motivo peptídico (basado en resilina)
- 50 <400> 19
- Gly Gly Arg Pro Ser Ser Ser Tyr Gly**
1 5
- 55 <210> 20
<211> 24
<212> PRT
<213> Secuencia Artificial
- 60 <220>
<223> sintética

ES 2 799 431 T3

<220>
 <223> sintética

5 <220>
 <221> DOMINIO
 <222> (1)..(27)
 <223> Módulo K (proteína flageliforme)

10 <400> 23

Gly Pro Gly Gly Ala Gly Gly Pro Tyr Gly Pro Gly Gly Ala Gly Gly
 1 5 10 15

Pro Tyr Gly Pro Gly Gly Ala Gly Gly Pro Tyr
 20 25

15 <210> 24
 <211> 28
 <212> PRT
 <213> Secuencia Artificial

20 <220>
 <223> sintética

25 <220>
 <221> DOMINIO
 <222> (1)..(28)
 <223> Módulo sp (proteína flageliforme)

<400> 24

Gly Gly Thr Thr Ile Ile Glu Asp Leu Asp Ile Thr Ile Asp Gly Ala
 1 5 10 15

Asp Gly Pro Ile Thr Ile Ser Glu Glu Leu Thr Ile
 20 25

30 <210> 25
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

35 <220>
 <223> sintética

40 <220>
 <221> DOMINIO
 <222> (1)..(34)
 <223> Módulo S (resilina)

<400> 25

Pro Gly Ser Ser Ala Ala Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly
 1 5 10 15

45 Gln Gly Gln Gly Gln Gly Gln Gly Gln Gly Gly Arg Pro Ser Asp Thr
 20 25 30

Tyr Gly

50 <210> 26
 <211> 39

ES 2 799 431 T3

Gly Pro Gly Gly Ala Gly Pro Gly Gly Tyr Gly Pro Gly Gly Ser Gly
 1 5 10 15

Pro Gly Gly Tyr Gly Pro Gly Gly Ser Gly Pro Gly Gly Tyr
 20 25 30

5 <210> 29
 <211> 24
 <212> PRT
 <213> Secuencia Artificial

10 <220>
 <223> sintética

15 <220>
 <221> DOMINIO
 <222> (1)..(24)
 <223> Módulo Ac

<400> 29

Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ala Ala Ala Ala Gly Gly
 1 5 10 15

Tyr Gly Pro Gly Cys Gly Gln Gln
 20

20 <210> 30
 <211> 24
 <212> PRT
 <213> Secuencia Artificial

25 <220>
 <223> sintética

30 <220>
 <221> DOMINIO
 <222> (1)..(24)
 <223> Módulo Ak

<400> 30

35 Gly Pro Tyr Gly Pro Gly Ala Ser Ala Ala Ala Ala Ala Ala Gly Gly
 1 5 10 15

Tyr Gly Pro Gly Lys Gly Gln Gln
 20

40 <210> 31
 <211> 35
 <212> PRT
 <213> Secuencia Artificial

45 <220>
 <223> sintética

50 <220>
 <221> DOMINIO
 <222> (1)..(35)
 <223> Módulo Cc

<400> 31

ES 2 799 431 T3

Gly Ser Ser Ala Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gly
 1 5 10 15

Tyr Gly Pro Glu Asn Gln Gly Pro Cys Gly Pro Gly Gly Tyr Gly Pro
 20 25 30

Gly Gly Pro
 35

5 <210> 32
 <211> 35
 <212> PRT
 <213> Secuencia Artificial

10 <220>
 <223> sintética

<220>
 <221> DOMINIO
 <222> (1)..(35)
 15 <223> Módulo Ck1

<400> 32

Gly Ser Ser Ala Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gly
 1 5 10 15

Tyr Gly Pro Glu Asn Gln Gly Pro Lys Gly Pro Gly Gly Tyr Gly Pro
 20 25 30

Gly Gly Pro
 35

20 <210> 33
 <211> 35
 <212> PRT
 <213> Secuencia Artificial

25 <220>
 <223> sintética

30 <220>
 <221> DOMINIO
 <222> (1)..(35)
 <223> Módulo Ck2

35 <400> 33

Gly Ser Ser Ala Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gly
 1 5 10 15

Tyr Gly Pro Lys Asn Gln Gly Pro Ser Gly Pro Gly Gly Tyr Gly Pro
 20 25 30

Gly Gly Pro
 35

ES 2 799 431 T3

<212> PRT
 <213> Secuencia Artificial

5 <220>
 <223> sintética

<220>
 <221> DOMINIO
 <222> (1)..(14)
 10 <223> TAG cys3

<400> 37

Gly Cys Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
 1 5 10

15 <210> 38
 <211> 13
 <212> PRT
 <213> Secuencia Artificial

20 <220>
 <223> sintética

<220>
 25 <221> DOMINIO
 <222> (1)..(13)
 <223> TAG lys1

<400> 38

30 <210> 39
 <211> 8
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> sintética

40 <220>
 <221> DOMINIO
 <222> (1)..(8)
 <223> TAG lys2

45 <400> 39

Gly Lys Gly Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly
 1 5 10

50 <210> 40
 <211> 5
 <212> PRT
 <213> Arthropoda

55 <220>
 <221> REPETICIÓN
 <222> (1)..(5)
 <223> motivo peptídico (resilina)

60 <400> 40

Gly Pro Gly Gln Gly
 1 5

ES 2 799 431 T3

<210> 41
 <211> 124
 <212> PRT
 <213> Secuencia Artificial
 5
 <220>
 <223> basada en ADF-3
 <220>
 10 <221> DOMINIO
 <222> (1)..(124)
 <223> NR3 (ADF-3)
 <400> 41
 15
 Gly Ala Ala Ser Ala Ala Val Ser Val Gly Gly Tyr Gly Pro Gln Ser
 1 5 10 15
 Ser Ser Ala Pro Val Ala Ser Ala Ala Ala Ser Arg Leu Ser Ser Pro
 20 25 30
 Ala Ala Ser Ser Arg Val Ser Ser Ala Val Ser Ser Leu Val Ser Ser
 35 40 45
 Gly Pro Thr Asn Gln Ala Ala Leu Ser Asn Thr Ile Ser Ser Val Val
 50 55 60
 Ser Gln Val Ser Ala Ser Asn Pro Gly Leu Ser Gly Cys Asp Val Leu
 65 70 75 80
 Val Gln Ala Leu Leu Glu Val Val Ser Ala Leu Val Ser Ile Leu Gly
 85 90 95
 Ser Ser Ser Ile Gly Gln Ile Asn Tyr Gly Ala Ser Ala Gln Tyr Thr
 100 105 110
 Gln Met Val Gly Gln Ser Val Ala Gln Ala Leu Ala
 115 120
 20 <210> 42
 <211> 109
 <212> PRT
 <213> Secuencia Artificial
 25 <220>
 <223> basada en ADF-4
 <220>
 30 <221> DOMINIO
 <222> (1)..(109)
 <223> NR4 (ADF-4)
 <400> 42

ES 2 799 431 T3

Gly Ala Tyr Gly Pro Ser Pro Ser Ala Ser Ala Ser Val Ala Ala Ser
 1 5 10 15

Arg Leu Ser Ser Pro Ala Ala Ser Ser Arg Val Ser Ser Ala Val Ser
 20 25 30

Ser Leu Val Ser Ser Gly Pro Thr Asn Gly Ala Ala Val Ser Gly Ala
 35 40 45

Leu Asn Ser Leu Val Ser Gln Ile Ser Ala Ser Asn Pro Gly Leu Ser
 50 55 60

Gly Cys Asp Ala Leu Val Gln Ala Leu Leu Glu Leu Val Ser Ala Leu
 65 70 75 80

Val Ala Ile Leu Ser Ser Ala Ser Ile Gly Gln Val Asn Val Ser Ser
 85 90 95

Val Ser Gln Ser Thr Gln Met Ile Ser Gln Ala Leu Ser
 100 105

- 5 <210> 43
- <211> 747
- <212> PRT
- 10 <213> Araneus diadematus
- <220>
- <221> PÉPTIDO
- <222> (1)..(747)
- 15 <223> MaSp I
- <400> 43

ES 2 799 431 T3

Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly
 1 5 10 15
 Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly
 20 25 30
 Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala Gly Ala Ala Ala Ala
 35 40 45
 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser
 50 55 60
 Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala
 65 70 75 80
 Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly
 85 90 95
 Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala
 100 105 110
 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Asn
 115 120 125
 Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Ala Ala Ala Ala Gly
 130 135 140
 Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly
 145 150 155 160

ES 2 799 431 T3

Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala
 165 170 175

Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala
 180 185 190

Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly
 195 200 205

Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly
 210 215 220

Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala
 225 230 235 240

Gly Ala Ser Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly
 245 250 255

Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Glu Gly Ala Gly Ala
 260 265 270

Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu
 275 280 285

Gly Gly Gln Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln
 290 295 300

Gly Ala Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala
 305 310 315 320

Ala Gly Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln
 325 330 335

Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly
 340 345 350

Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly
 355 360 365

Gln Gly Ala Gly Ala Val Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln
 370 375 380

Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln
 385 390 395 400

Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Arg Gly
 405 410 415

ES 2 799 431 T3

Tyr Gly Gly Leu Gly Asn Gln Gly Ala Gly Arg Gly Gly Leu Gly Gly
 420 425 430

Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln
 435 440 445

Gly Gly Tyr Gly Gly Leu Gly Asn Gln Gly Ala Gly Arg Gly Gly Gln
 450 455 460

Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly
 465 470 475 480

Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala
 485 490 495

Ala Ala Ala Ala Val Gly Ala Gly Gln Glu Gly Ile Arg Gly Gln Gly
 500 505 510

Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Ser Gly Arg
 515 520 525

Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly
 530 535 540

Gly Ala Gly Gln Gly Gly Leu Gly Gly Gln Gly Ala Gly Gln Gly Ala
 545 550 555 560

Gly Ala Ala Ala Ala Ala Ala Gly Gly Val Arg Gln Gly Gly Tyr Gly
 565 570 575

Gly Leu Gly Ser Gln Gly Ala Gly Arg Gly Gly Gln Gly Ala Gly Ala
 580 585 590

Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu
 595 600 605

Gly Gly Gln Gly Val Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly
 610 615 620

Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Val Gly
 625 630 635 640

Ser Gly Ala Ser Ala Ala Ser Ala Ala Ala Ser Arg Leu Ser Ser Pro
 645 650 655

Gln Ala Ser Ser Arg Val Ser Ser Ala Val Ser Asn Leu Val Ala Ser
 660 665 670

ES 2 799 431 T3

Gly Pro Thr Asn Ser Ala Ala Leu Ser Ser Thr Ile Ser Asn Val Val
 675 680 685

Ser Gln Ile Gly Ala Ser Asn Pro Gly Leu Ser Gly Cys Asp Val Leu
 690 695 700

Ile Gln Ala Leu Leu Glu Val Val Ser Ala Leu Ile Gln Ile Leu Gly
 705 710 715 720

Ser Ser Ser Ile Gly Gln Val Asn Tyr Gly Ser Ala Gly Gln Ala Thr
 725 730 735

Gln Ile Val Gly Gln Ser Val Tyr Gln Ala Leu
 740 745

5 <210> 44
 <211> 627
 <212> PRT
 <213> Araneus diadematus

10 <220>
 <221> PÉPTIDO
 <222> (1)..(627)
 <223> MaSp II

15 <400> 44

Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro
 1 5 10 15

Gly Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala Ala Ala Ala Ala Ala
 20 25 30

Ala Ala Ala Ala Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro
 35 40 45

Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Arg Tyr Gly Pro Gly
 50 55 60

Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala Ala Ala Ala Ala Ala Gly
 65 70 75 80

Ser Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Arg Gln Gln Gly Pro
 85 90 95

Gly Gly Tyr Gly Gln Gly Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala
 100 105 110

ES 2 799 431 T3

Ala Ala Ala Ser Ala Ala Ala Ser Ala Glu Ser Gly Gln Gln Gly Pro
 115 120 125

Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly
 130 135 140

Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Ser Gly
 145 150 155 160

Pro Gly Ser Ala Ala Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gln
 165 170 175

Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr
 180 185 190

Gly Pro Gly Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala Ala Ala Ala
 195 200 205

Ala Ala Ala Ala Ser Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly
 210 215 220

Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Leu
 225 230 235 240

Ser Gly Pro Gly Ser Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly Gln
 245 250 255

Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Ser Gly Pro
 260 265 270

Gly Ser Ala Gly Pro Gly Gly Tyr
 275 280 285

Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Gly
 290 295 300

Pro Ser Gly Ala Gly Ser Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly
 305 310 315 320

Gln Gln Gly Leu Gly Gly Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly
 325 330 335

Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Gly Pro Gly Ser Ala
 340 345 350

Ser Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly Gln Gln Gly Pro Gly Gly
 355 360 365

ES 2 799 431 T3

Tyr Gly Pro Gly Gln Gln Gly Pro Ser Gly Pro Gly Ser Ala Ser Ala
 370 375 380

Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly Gly Tyr Gly Pro Gly Gln
 385 390 395 400

Gln Gly Pro Gly Gly Tyr Ala Pro Gly Gln Gln Gly Pro Ser Gly Pro
 405 410 415

Gly Ser Ala Ser Ala Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly Gly
 420 425 430

Tyr Gly Pro Gly Gln Gln Gly Pro Gly Gly Tyr Ala Pro Gly Gln Gln
 435 440 445

Gly Pro Ser Gly Pro Gly Ser Ala Ala Ala Ala Ala Ala Ala Ala
 450 455 460

Gly Pro Gly Gly Tyr Gly Pro Ala Gln Gln Gly Pro Ser Gly Pro Gly
 465 470 475 480

Ile Ala Ala Ser Ala Ala Ser Ala Gly Pro Gly Gly Tyr Gly Pro Ala
 485 490 495

Gln Gln Gly Pro Ala Gly Tyr Gly Pro Gly Ser Ala Val Ala Ala Ser
 500 505 510

Ala Gly Ala Gly Ser Ala Gly Tyr Gly Pro Gly Ser Gln Ala Ser Ala
 515 520 525

Ala Ala Ser Arg Leu Ala Ser Pro Asp Ser Gly Ala Arg Val Ala Ser
 530 535 540

Ala Val Ser Asn Leu Val Ser Ser Gly Pro Thr Ser Ser Ala Ala Leu
 545 550 555 560

Ser Ser Val Ile Ser Asn Ala Val Ser Gln Ile Gly Ala Ser Asn Pro
 565 570 575

Gly Leu Ser Gly Cys Asp Val Leu Ile Gln Ala Leu Leu Glu Ile Val
 580 585 590

Ser Ala Cys Val Thr Ile Leu Ser Ser Ser Ser Ile Gly Gln Val Asn
 595 600 605

Tyr Gly Ala Ala Ser Gln Phe Ala Gln Val Val Gly Gln Ser Val Leu
 610 615 620

ES 2 799 431 T3

Ser Ala Phe
625

REIVINDICACIONES

1. Un método para aislar una proteína de seda objetivo insoluble de una suspensión de células huésped intactas o rotas que comprende las etapas de:
- 5 a) proporcionar una suspensión de células huésped intactas o rotas que comprende una proteína de seda objetivo insoluble y partes insolubles de la célula huésped,
 b) agregar una solución acuosa de al menos una base a dicha suspensión, en la que la base está en una cantidad que es suficiente para romper dichas células huésped y/o solubilizar dichas partes insolubles de la célula huésped, y
 c) separar la proteína de seda objetivo insoluble de las partes insolubles de la célula huésped solubilizadas,
- 10 en el que en la etapa b) la proteína de seda objetivo permanece insoluble, y en el que al menos el 80% de las partes insolubles de la célula huésped se solubilizan.
2. El método de la reivindicación 1, en el que las células huésped son células huésped bacterianas, de levadura, de plantas o de insectos.
- 15 3. El método de las reivindicaciones 1 o 2, en el que la proteína de seda objetivo insoluble aislada tiene una pureza de al menos 80%, preferiblemente de al menos 90%, y más preferiblemente de al menos 95%.
4. El método de una cualquiera de las reivindicaciones 1 a 3, en el que el método comprende además la adición de al menos un reactivo (i) antes de la etapa b), (ii) en la etapa b) y/o (iii) después de la etapa b).
- 20 5. El método de la reivindicación 4, en el que el reactivo se selecciona del grupo que consta de un agente desnaturizante, un agente cosmotrópico y un detergente, o es una combinación de los mismos.
6. El método de una cualquiera de las reivindicaciones 1 a 5, en el que la separación de la proteína de seda objetivo insoluble de las partes insolubles de la célula huésped solubilizadas en la etapa c) se logra mediante centrifugación, sedimentación y/o filtración.
- 25 7. El método de una cualquiera de las reivindicaciones 1 a 6, en el que el método comprende adicionalmente después a la etapa b) y antes de la etapa c) una etapa b') de homogeneización de dicha suspensión.
8. El método de una cualquiera de las reivindicaciones 1 a 7, en el que el método comprende adicionalmente después a la etapa c) una etapa d) de lavar la proteína de seda objetivo insoluble separada, preferiblemente el centrifugado, sedimento o retenido, con una solución acuosa, un solución orgánica y/o urea.
- 30 9. El método de una cualquiera de las reivindicaciones 1 a 8, en el que la proteína de seda objetivo permanece insoluble al menos 90%, preferiblemente al menos 95%, después de la adición de una solución acuosa que comprende al menos una base en la etapa b) sobre un período de tiempo de entre 10 y 40 min y a una temperatura de entre 15 °C y 25 °C.
10. El método de una cualquiera de las reivindicaciones 1 a 9, en el que la proteína de seda objetivo insoluble forma un agregado de proteína que comprende al menos 85% de dicha proteína de seda objetivo.
- 35 11. El método de una cualquiera de las reivindicaciones 1 a 10, en el que la proteína de seda objetivo insoluble es una proteína de seda que comprende al menos dos unidades repetitivas idénticas.
12. El método de la reivindicación 11, en el que la proteína de seda comprende al menos dos unidades repetitivas idénticas, cada una de las cuales comprende al menos una secuencia de consenso seleccionada del grupo que consta de:
- 40 i) GPGXX (SEQ ID NO: 3), en la que X es cualquier aminoácido, preferiblemente en cada caso independientemente seleccionado de A, S, G, Y, P y Q;
 ii) GGX, en la que X es cualquier aminoácido, preferiblemente en cada caso independientemente seleccionado de Y, P, R, S, A, T, N y Q; y
 iii) A_x, en la que x es un número entero de 5 a 10.
- 45 13. El método de las reivindicaciones 11 o 12, en el que las unidades repetitivas se seleccionan independientemente del módulo A (SEQ ID NO: 20), módulo C (SEQ ID NO: 21), módulo Q (SEQ ID NO: 22), módulo K (SEQ ID NO: 23), módulo sp (SEQ ID NO: 24), módulo S (SEQ ID NO: 25), módulo R (SEQ ID NO: 26), módulo X (SEQ ID NO: 27) o módulo Y (SEQ ID NO: 28).
14. El método de una cualquiera de las reivindicaciones 11 a 13, en el que la proteína de seda comprende además al menos una unidad no repetitiva (NR).
- 50

FIG. 1

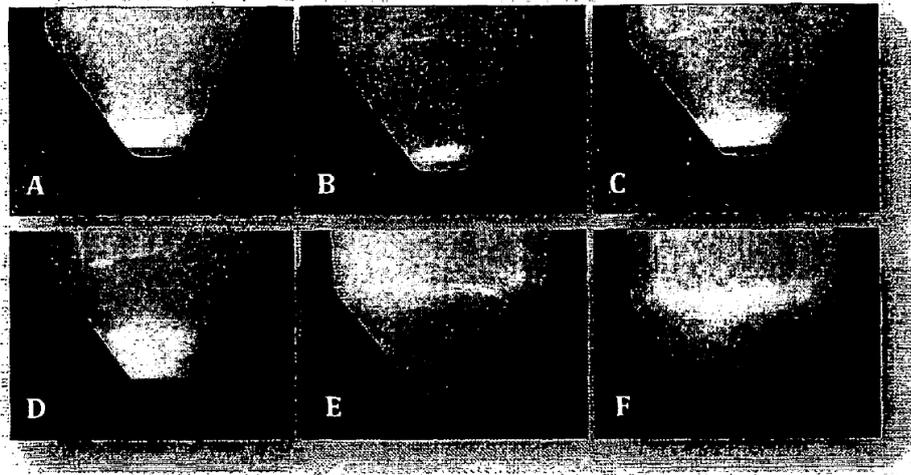


FIG. 2



FIG. 3

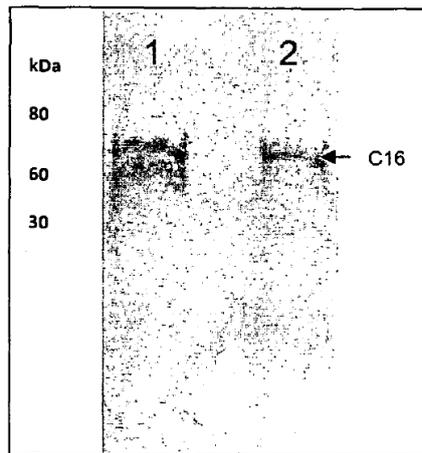


FIG. 4

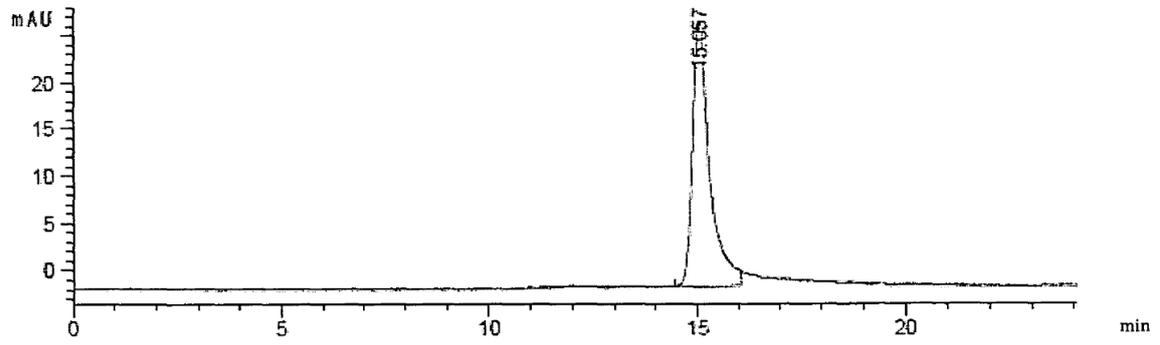


FIG. 5

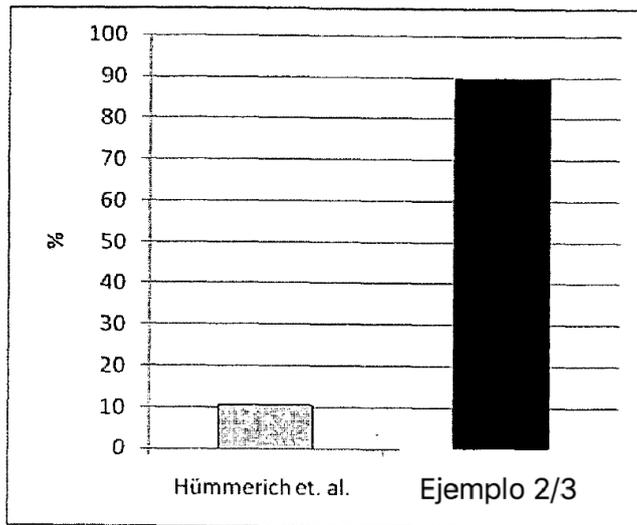


FIG. 6

