

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 986 499**

51 Int. Cl.:

C07K 7/08 (2006.01)
C07K 14/33 (2006.01)
C07K 14/315 (2006.01)
C07K 17/00 (2006.01)
C07K 1/16 (2006.01)
C07K 14/31 (2006.01)
C12N 15/62 (2006.01)
C07K 1/22 (2006.01)
C07K 16/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.05.2018** **PCT/EP2018/063435**

87 Fecha y número de publicación internacional: **29.11.2018** **WO18215503**

96 Fecha de presentación y número de la solicitud europea: **23.05.2018** **E 18726972 (5)**

97 Fecha y número de publicación de la concesión europea: **10.07.2024** **EP 3630793**

54 Título: **Proteína recombinante**

30 Prioridad:

24.05.2017 GB 201708277

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

11.11.2024

73 Titular/es:

CYTIVA BIOPROCESS R&D AB (100.0%)
Björkgatan 30
751 84 Uppsala, SE

72 Inventor/es:

ANDER, MATS;
BERGMAN, MAGNUS;
BJORKMAN, TOMAS;
GALLI, JOAKIM y
RODRIGO, GUSTAV

74 Agente/Representante:

ELZABURU, S.L.P

Observaciones:

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

ES 2 986 499 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

Proteína recombinante

Campo técnico de la invención

- 5 La presente invención se refiere a proteínas recombinantes, y más particularmente a proteínas recombinantes expresadas en bacterias gram negativas tales como *Escherichia coli* (*E. coli*). La invención también se refiere a ácidos nucleicos, vectores y bacterias gram negativas para la expresión de proteínas recombinantes, así como a matrices de separación con ligandos de proteínas recombinantes unidos covalentemente y a métodos de separación de inmunoglobulinas en dichas matrices.

Antecedentes de la invención

- 10 La expresión de proteínas heterólogas en *E. coli* (*E. coli*) se usa comúnmente para proteínas recombinantes a escala de laboratorio y comercial. La expresión en *E. coli* usando secreción, generalmente se refiere al transporte de la proteína producida sobre la membrana interna que separa el citoplasma y el periplasma. Mediante secreción al periplasma, la proteína muchas veces también se filtra hacia el medio extracelular (Mergulhão et al., Biotech Adv 23, 177-202, 2005). La secreción tiene muchas ventajas en comparación con la expresión citoplásmica, tales como facilitar el plegamiento correcto de la proteína, el procesamiento correcto del extremo N-terminal, la simplificación del procesamiento posterior y la prevención de la agregación en cuerpos de inclusión. Sin embargo, lejos de todas las proteínas se han expresado con éxito en forma soluble en el periplasma. Algunos de los problemas que podrían surgir
- 15 el plegamiento correcto de la proteína, el procesamiento correcto del extremo N-terminal, la simplificación del procesamiento posterior y la prevención de la agregación en cuerpos de inclusión. Sin embargo, lejos de todas las proteínas se han expresado con éxito en forma soluble en el periplasma. Algunos de los problemas que podrían surgir es una secreción pobre y ningún procesamiento, o incorrecto, del péptido señal. Un ejemplo particular de proteínas que tienen problemas con una secreción inadecuada es la expresión de aglutinantes de inmunoglobulina basados en los dominios B o C de unión a Fc nativos o mutados de la Proteína A de *Staphylococcus aureus* (L. Abrahmsén et al. EMBO J 4 (13B), 3901-3906, 1985). Dichos aglutinantes de inmunoglobulina se usan comúnmente como ligandos en la separación por cromatografía de afinidad de anticuerpos monoclonales, una categoría principal de productos farmacéuticos modernos.
- 20

- 25 La expresión de la proteína depende de una secuencia promotora que inicia la transcripción del ácido ribonucleico mensajero (mRNA) seguido de un sitio de unión a ribosoma (RBS) que atrae la maquinaria de traducción seguido de una secuencia de péptido señal que facilita el transporte de la proteína al periplasma. La proteína madura se clona a menudo después del péptido señal y la proteína madura se escinde del péptido señal mediante una peptidasa señal cuando pasa por la membrana. Sin embargo, un problema cuando se clonan constructos después de un péptido señal es que las enzimas de restricción a menudo necesitan una secuencia específica para escindir el DNA, y esto deja una cicatriz de clonación después de la secuencia del péptido señal.
- 30

Por consiguiente, existe la necesidad de mejoras en la expresión de proteínas heterólogas en *E. coli* y otras bacterias gram negativas, en particular para las proteínas de unión a inmunoglobulinas derivadas de los dominios B y C de la Proteína A de *Staphylococcus aureus*.

Compendio de la invención

- 35 Un aspecto de la invención es proporcionar una proteína funcional que se expresa y secreta fácilmente en bacterias gram negativas como *E. coli*. Esto se logra con una proteína recombinante que comprende un polipéptido de unión a Fc y, unido al extremo N-terminal de dicho polipéptido de unión a Fc, un espaciador N-terminal que tiene una longitud tal que la distancia entre un sitio de escisión del péptido señal y una unidad estructural proximal del extremo N-terminal del polipéptido de unión a Fc es de 14-24 restos de aminoácidos.

- 40 Una ventaja es que el nivel de expresión se mejora mediante la introducción del espaciador N-terminal. Una ventaja adicional es que se mejora la selectividad de la escisión del péptido señal.

Un segundo aspecto de la invención es proporcionar una molécula de ácido nucleico que codifica la proteína recombinante. Esto se logra con una molécula de ácido nucleico que comprende los siguientes elementos en la dirección 5' a 3', estando dichos elementos unidos operativamente:

- 45 a) una secuencia de DNA promotora inducible o constitutiva;
- b) una secuencia de DNA que codifica un péptido señal;
- c) una secuencia de DNA que codifica un espaciador N-terminal; y
- d) una secuencia de DNA que codifica un polipéptido de unión a inmunoglobulina.

- 50 Un tercer aspecto de la invención es proporcionar un vector de clonación que expresa y secreta la proteína recombinante de cualquier reivindicación anterior en el periplasma bacteriano de una célula gram negativa. Esto se logra con un vector de clonación que comprende la molécula de ácido nucleico anterior.

Un cuarto aspecto de la invención es proporcionar una bacteria gram negativa que comprende el vector de clonación.

Un quinto aspecto de la invención es proporcionar un método para expresar y secretar la proteína recombinante en una bacteria gram negativa. Esto se logra mediante un método que comprende las etapas de proporcionar la bacteria gram negativa y cultivar la bacteria gram negativa.

5 Un sexto aspecto de la invención es proporcionar una matriz de separación que comprende la proteína recombinante unida covalentemente a un soporte.

Un séptimo aspecto de la invención es proporcionar un método para separar una inmunoglobulina, que comprende las etapas de:

i) proporcionar la matriz de separación anterior en donde la proteína recombinante comprende un polipéptido de unión a inmunoglobulina;

10 ii) poner en contacto la matriz de separación con una muestra líquida que contiene una inmunoglobulina, para unir la inmunoglobulina;

iii) opcionalmente lavar la matriz de separación con un líquido de lavado;

iv) poner en contacto la matriz de separación con un líquido de elución, para eluir la inmunoglobulina.

Realizaciones adicionales adecuadas de la invención se describen en las reivindicaciones dependientes.

15 **Dibujos**

La Figura 1 muestra una secuencia de ejemplo de un sitio de clonación en donde se insertó la secuencia espaciadora. Los sitios de escisión de la enzima de restricción están marcados en la parte superior de la secuencia.

20 La Figura 2 muestra el área integrada del pico a 237 nm de los eluatos de la columna de Sepharose 6FF Tricorn 10 de IgG.

La Figura 3 muestra las mediciones de UV a 210 nm, área integrada del pico. a) OmpA-AQGT de pGE120 (referencia), 52% de péptido señal procesado correcto. b) OmpA-DsbA8AA de pGE144, 96% de escisión del péptido señal correcta, c) DsbA-DsbA8AA de pGE140, 97% de escisión de péptido señal correcta.

25 La Figura 4 muestra la expresión de proteína medida en caldo de fermentación tratado térmicamente para Zvar2₆, con y sin el espaciador N-terminal DsbA8AA, usando el análisis de concentración con una curva estándar. La flecha indica el tiempo de inducción.

La Figura 5 muestra mediciones de UV a 210 nm, área integrada del pico. a) OmpA-AQGT-Zvar2₆ de pGE0002 (referencia). b) OmpA-DsbA8AA-Zvar2₆ de pGE0180.

30 La Figura 6 muestra una imagen esquemática del constructo, con péptido señal, espaciador N-terminal y polipéptido de unión a Fc.

La Figura 7 muestra un resumen de los resultados de expresión de los cultivos en matraz de agitación. Las barras de error correspondientes a una desviación estándar se incluyen cuando sea aplicable. Se realizaron tres mediciones repetidas para el constructo pGE180. Se realizaron dos mediciones repetidas para las muestras DsbA7 y DsbA4. Para todos los demás constructos, se realizó una medición.

35 La Figura 8 muestra a) un ejemplo del pico relevante del cromatograma iónico total (TIC) de un constructo con escisión correcta del péptido señal (pGE0180 en este caso) y b) un constructo con escisión incorrecta del péptido señal (pGE0002 en este caso).

40 La Figura 9 muestra ejemplos de picos del TIC desconvolucionados después de 24 horas de incubación en NaOH 1 M. La imagen izquierda muestra el patrón típico visto en pGE0180 (el grupo de picos entre 500 y 800 m/z). La imagen derecha muestra la misma área en DsbA8_noGT donde no hay picos, excepto el fondo.

La Figura 10 muestra las áreas integradas de cromatograma iónico extraído (XIC) de picos peptídicos correspondientes a secuencias N-terminales escindidas de diferentes longitudes después de 0 horas (superior izquierda), 4 horas (superior derecha) y 24 horas (inferior izquierda). Las áreas se dan en unidades arbitrarias (AU).

45 La Figura 11 muestra una vista ampliada de las áreas peptídicas relevantes de los candidatos más prometedores. Las áreas se dan en unidades arbitrarias (AU).

La Figura 12 muestra una descomposición de los péptidos que aparecen más destacadamente en los constructos que muestran la estabilidad alcalina más baja en la región N-terminal. Las áreas se dan en unidades arbitrarias (AU).

La Figura 13 muestra las concentraciones de proteína medidas, las concentraciones de proteína estimadas a partir de los coeficientes de extinción a 280 nm, y la OD600 al final de los cultivos con alimentación por lotes. Se realizó una prueba por duplicado para pGE180, todos los demás constructos se cultivaron solo una vez. Las barras de error correspondientes a un intervalo de confianza del 95% se incluyen cuando los datos están disponibles.

La Figura 14 muestra una comparación entre las capacidades de unión dinámica de los constructos candidatos (DsbA8_noGT y DsbA8_DT), la muestra de referencia con Zvar2 y los datos previos de otras inmobilizaciones usando matrices base con pesos secos, porosidades y densidades de ligando similares.

Definiciones

Los términos "anticuerpo" e "inmunoglobulina" se usan indistintamente en la presente memoria, y se entiende que incluyen también fragmentos de anticuerpos, proteínas de fusión que comprenden anticuerpos o fragmentos de anticuerpos y conjugados que comprenden anticuerpos o fragmentos de anticuerpos.

Los términos "polipéptido de unión a Fc" y "proteína de unión a Fc" significan un polipéptido o proteína respectivamente, capaz de unirse a la parte cristalizable (Fc) de un anticuerpo e incluye, por ejemplo, la Proteína A de *Staphylococcus aureus* y la proteína G de *Streptococcus*, o cualquier fragmento o proteína de fusión de las mismas que haya mantenido dicha propiedad de unión.

Los términos "polipéptido de unión a Fab" y "proteína de unión a Fab" significan un polipéptido o proteína respectivamente, capaz de unirse a la parte de unión a antígeno (Fab) de un anticuerpo e incluye, por ejemplo, la proteína L de *Peptostreptococcus magnus*, la proteína G de *Streptococcus*, la proteína A nativa de *Staphylococcus aureus* o cualquier fragmento o proteína de fusión de las mismas que haya mantenido dicha propiedad de unión.

El término "enlazador" en la presente memoria significa un elemento que une dos unidades polipeptídicas, monómeros o dominios entre sí en un multímero.

El término "% de identidad" con respecto a comparaciones de secuencias de aminoácidos se determina mediante algoritmos de alineación estándar tales como, por ejemplo, la herramienta de alineación local básica (BLAST™) descrita en Altschul et al. (1990) J. Mol. Biol., 215: 403-410. Un software basado en la web para esto está disponible libremente en US National Library of Medicine en http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome. En la presente memoria, el algoritmo "blastp (BLAST proteína-proteína) se usa para la alineación de una secuencia problema con una secuencia objetivo y determinar, entre otros, el % de identidad.

La abreviatura "DsbA" en la presente memoria significa *E. coli*, Tiol: proteína de intercambio disulfuro, UniProt P0AEG4.

La abreviatura "OmpA" en la presente memoria significa *E. coli*, Proteína A de membrana externa, UniProt P0A910.

La abreviatura "PrA" en la presente memoria significa proteína A de *Staphylococcus aureus*, UniProt P38507.

La abreviatura "GIII" en la presente memoria significa el Gen 3 del Bacteriófago M13, UniProt P69168.

El término "péptido señal" en la presente memoria significa un péptido corto (normalmente de 16-30 aminoácidos de longitud) presente en el extremo N-terminal de la mayoría de las proteínas recién sintetizadas que están destinadas a la ruta secretora. También puede denominarse como secuencia señal, señal de direccionamiento, señal de localización, secuencia de localización, péptido de tránsito, secuencia líder o péptido líder. El péptido señal se escinde normalmente de la proteína mediante una enzima peptidasa señal.

La expresión "sitio de escisión de péptido señal" en la presente memoria significa un dipéptido entre el que la peptidasa señal escinde el péptido señal de la proteína madura. En la mayoría de los casos (pero no en todos) el dipéptido es Ala-Ala. El sitio de escisión del péptido señal puede calcularse con algoritmos tales como SignalP 4.1, disponible en línea en <http://www.cbs.dtu.dk/services/SignalP/>

(Center for Biological Sequence Analysis, Technical University of Denmark).

El término "expresión heteróloga" en la presente memoria significa la expresión de un gen o parte de un gen en un organismo huésped, que no tiene de forma natural este gen o fragmento del gen. "Secretado" se refiere a través de la membrana interna de bacterias gram negativas, tales como *E. coli*.

El citoplasma (citoplásmico) es el espacio dentro de la membrana celular interna en bacterias gram negativas que contienen el material genético. El término "expresión citoplásmica" en la presente memoria significa la expresión de proteínas dentro del citoplasma.

El periplasma es una matriz concentrada de tipo gel en el espacio entre la membrana citoplásmica interna y la membrana externa bacteriana denominado espacio periplásmico en bacterias gram negativas.

El término "promotor" en la presente memoria significa una región de DNA que inicia la transcripción (escribiendo en mRNA) de un gen particular. Los promotores se localizan normalmente cerca de los sitios de inicio de la transcripción de los genes, en la misma cadena y la cadena anterior en el DNA (hacia la región 5' de la cadena sentido). Un promotor puede ser inducible, lo que significa que la expresión de los genes unidos operativamente al promotor puede activarse mediante la presencia de una sustancia inductora. Alternativamente, el promotor puede ser constitutivo, es decir, que no está regulado por ninguna sustancia inductora.

La abreviatura "RBS" en la presente memoria significa un sitio de unión a ribosomas, o sitio de unión ribosomal. Esta es una secuencia de nucleótidos anteriores del codón de inicio de un transcrito del mRNA que es responsable del reclutamiento de un ribosoma durante el inicio de la traducción de proteínas.

La abreviatura "RhaBAD" en la presente memoria significa el promotor del operón de la ramnosa de *E. coli* (también denominado promotor de ramnosa) de los genes RhaB, RhaA y RhaD. Este es un promotor ampliamente usado en biología molecular.

La abreviatura "T5" en la presente memoria significa el promotor del bacteriófago T5 para RNA polimerasa de *E. coli*, con un operador lac incorporado. Un operador es un segmento de DNA al que se une un factor de transcripción para regular la expresión génica reprimiéndolo.

La abreviatura "pD861-SR" en la presente memoria se refiere a un plásmido para la expresión de proteínas de *E. coli* con el promotor de la ramnosa (RhaBAD) y un sitio de unión a ribosomas fuerte (SR).

La abreviatura "pJ401" en la presente memoria se refiere a un plásmido para la expresión de proteína de *E. coli*, con el promotor del bacteriófago T5 y el operador lac doblemente incorporado reflejados en cada lado del promotor.

La abreviatura "OptEc" en la presente memoria significa la expresión optimizada para *E. coli*, es decir, se escogen tripletes de codones para que se ajusten a la maquinaria de traducción de *E. coli*.

La abreviatura "FspI" en la presente memoria significa una enzima de restricción de DNA de la especie *Escherichia* (ATCC 29114). Se escinde de manera roma en la secuencia TGCGCA.

La abreviatura "KpnI" en la presente memoria significa una enzima de restricción de DNA de *Klebsiella pneumoniae* OK8 (ATCC 49790). Se escinde con el saliente en la secuencia GGTACC.

La abreviatura "SRP" en la presente memoria significa la ruta de partículas de reconocimiento de señal, una ruta conservada de manera universal para dirigirse a polipéptidos para la secreción a través de la ruta co-traduccional.

La abreviatura "Sec" en la presente memoria significa la secreción o la ruta secretora de tipo II, un sistema que es responsable de la secreción de proteínas a través de la membrana celular.

El término "K12-017 de *E. coli*" en la presente memoria significa una cepa de expresión de *E. coli*, tal y como se describe por Olsson M.O. e Isaksson L.A. en *Molec. Gen. Genet.* 169, 251-257 (1979).

El término "polipéptido funcional" en la presente memoria significa un polipéptido con una propiedad técnicamente útil. Ejemplos de dichas propiedades son: a) unión altamente específica a una especie diana (para su uso como aglutinantes de afinidad, en particular como ligandos para cromatografía de afinidad), b) propiedades terapéuticas (para su uso como un medicamento), c) propiedades enzimáticas (para su uso como biocatalizadores) y d) propiedades emisoras de señal (para su uso como proteínas indicadoras, tales como proteínas indicadoras fluorescentes).

Tal y como se usa en la presente memoria, los términos "comprende", "que comprende", "que contiene", "que tiene" y similares pueden tener el significado atribuido a ellos en la ley de patentes de Estados Unidos y pueden significar "incluye", "que incluye" y similares; "que consiste esencialmente en" o "consiste esencialmente" de manera similar tiene el significado atribuido en la ley de patentes de Estados Unidos y el término es abierto, permitiendo la presencia de más de lo que se indica siempre que las características básicas o novedosas de lo que se indica no se cambien por la presencia de más de lo que se indica, pero excluye las realizaciones de la técnica anterior.

Descripción detallada de realizaciones

En un aspecto, mostrado por las Figuras 1-3, la presente invención describe una proteína recombinante que comprende un polipéptido de unión a Fc y, unido al extremo N-terminal del polipéptido de unión a Fc, un espaciador N-terminal que tiene una longitud de manera que el número de restos de aminoácidos entre un sitio de escisión del péptido señal y una unidad estructural proximal del extremo N-terminal del polipéptido de unión a Fc es de 14-24.

El polipéptido de unión a Fc es un polipéptido de unión a inmunoglobulina. Dicho polipéptido puede comprender, por ejemplo, uno o más dominios de unión a inmunoglobulina derivados de una proteína bacteriana seleccionada del grupo que consiste en Proteína A de *Staphylococcus aureus*, Proteína L de *Peptostreptococcus magnus* y Proteína G de *Streptococcus*, tal como del grupo que consiste en Proteína A de *Staphylococcus aureus* y Proteína L de *Peptostreptococcus magnus*. El dominio(s) de unión a inmunoglobulina puede tener, por ejemplo, al menos un 80%,

tal como al menos un 90 o un 95%, de identidad de secuencia con el dominio E, D, A, B o C de la Proteína A de *Staphylococcus aureus*, con la proteína Z (una variante del dominio B de la Proteína A de *Staphylococcus aureus*), Zvar o Zvar2 (mutantes estabilizados con álcali de la Proteína Z) o con el dominio 1, 2, 3, 4 o 5 de la Proteína L de *Peptostreptococcus magnus*. En este contexto, el dominio(s) que contiene inmunoglobulina puede definirse por, o tener al menos un 80%, tal como al menos un 90 o 95% de identidad de secuencia con, una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 1-11. SEQ ID NO. 1-7 (dominios E, D, A, B y C de la Proteína A de *Staphylococcus aureus*, proteína Z y Zvar) se muestran en la Figura 15, y SEQ ID NO. 8-11 se especifican a continuación.

SEQ ID NO. 8 - Zvar2

VDAKFDKEAQ EAFYEILHLP NLTEEQRNAP IQSLKDDPSV SKAILAEAKK
LNDAPK

SEQ ID NO. 9 – Dominio C truncado de la Proteína A de *Staphylococcus aureus*

QQ NAFYEILHLP NLTEEQRNGF IQSLKDDPSV SKEILAEAKK LNDAP

SEQ ID NO.10 - versión truncada de Zvar

QQ NAFYEILHLP NLTEEQRNGF IQSLKDDPSV SKEILAEAKK LNDAP

SEQ ID NO. 11 - versión truncada de Zvar2

AQ EAFYEILHLP NLTEEQRNAP IQSLKDDPSV SKAILAEAKK LNDAP

SEQ ID NO. 1-11 pueden caracterizarse todos como dominios de unión a Fc derivados de la Proteína A de *Staphylococcus aureus*. Dichos dominios pueden estabilizarse adicionalmente con álcali mediante mutaciones de los dominios nativos, como se ha hecho en Zvar y Zvar2. Ejemplos adicionales de dichos dominios estabilizados con álcali pueden ser SEQ ID NO. 48-93 (mostrados en el Ejemplo 6) y otros ejemplos se dan, por ejemplo, en los documentos de Patente US 8,329,860, US 8,754,196, US 9,040,661, US 9,403,883, JP 2006304633A, US 8,674,073, US 2010/0221844, US 2012/0208234, US 9,051,375, US 2014/0031522, US 2014/0107315, US 2013/0096276, US 2013/0274451, US 2005/0143566, US 2016/0159855, US 2016/0168209, US 2016/0237124, WO 2014/146350, WO 2016/079033, WO 2016/079034, WO 2016/152946, PCT EP2017/061162, PCT EP2017/061164, PCT EP2017/061160, PCT EP2017/061158, PCT EP2017/061159, US 14/961164, US 15/348699 y US 15/2823 67.

Específicamente, los dominios de unión a Fc estabilizados con álcali pueden tener al menos un 80%, tal como al menos un 90% o al menos un 95%, de identidad de secuencia con una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO. 7-11, 48-64 y 74-93.

El polipéptido de unión a inmunoglobulina puede ser adecuadamente un multímero de dominios de unión a inmunoglobulina, por ejemplo dominios de unión a Fc, como se ha descrito anteriormente. El multímero puede ser, por ejemplo, un dímero, trímero, tetrámero, pentámero, hexámero o un heptámero, tal como un dímero, tetrámero o hexámero. Adecuadamente, el multímero puede comprender al menos cuatro dominios. Los dominios pueden estar directamente unidos entre sí (por ejemplo, en el caso de SEQ ID NO. 1-8, 48-64 y 74-93), pero también pueden unirse entre sí a través de enlazadores, que comprenden típicamente 1-25 (tal como 3-20) restos de aminoácidos (por ejemplo, en el caso de SEQ ID NO 9-11). Los ejemplos de enlazadores adecuados incluyen APKADNKFNKE, APKVDAKFDKE, APK, APKVDA, AKFDKE, APKVFDKE, APAKFDKE, VDAKFDKE, APKKFDKE, APKYEDGKQYTVDAKFDKE y APKYEDGVDAKFDKE. Algunos ejemplos específicos de multímeros incluyen la SEQ ID NO. 12 (tetrámero), SEQ ID NO. 13 (hexámero) y SEQ ID NO. 65-73 (dímeros).

SEQ ID NO 12 - tetrámero de Zvar con cisteína C-terminal

VDAKFDKEQQ NAFYEILHLP NLTEEQRNAP IQSLKDDPSQ SANLLAEAKK
LNDAPK
VDAKFDKEQQ NAFYEILHLP NLTEEQRNAP IQSLKDDPSQ SANLLAEAKK
LNDAPK
VDAKFDKEQQ NAFYEILHLP NLTEEQRNAP IQSLKDDPSQ SANLLAEAKK
LNDAPK
VDAKFDKEQQ NAFYEILHLP NLTEEQRNAP IQSLKDDPSQ SANLLAEAKK
LNDAPKC

SEQ ID NO 13 - Hexámero de Zvar2 con cisteína C-terminal

VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQSLKDDPSV SKAILAEAKK
 LNDQAQPK VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQSLKDDPSV
 SKAILAEAKK LNDQAQPK
 VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQSLKDDPSV SKAILAEAKK
 LNDQAQPK VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQSLKDDPSV
 SKAILAEAKK LNDQAQPK VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF
 IQSLKDDPSV SKAILAEAKK LNDQAQPK VDAKFDKEAQ EAFYEILHLP
 NLTEEQRNAF IQSLKDDPSV SKAILAEAKK LNDQAQPKC

La proteína recombinante puede comprender un resto de acoplamiento en o proximal al extremo C-terminal o al extremo N-terminal, tal y como en el extremo C-terminal. Este resto de acoplamiento puede usarse para el acoplamiento específico a un soporte como se describe a continuación y puede comprender una cisteína, lo que permite el acoplamiento mediante enlaces tioéter. Alternativamente, o adicionalmente, el resto de acoplamiento puede comprender una o más lisinas, tal como un agrupamiento de 2-8 lisinas.

El polipéptido de unión a Fc puede tener estructuras secundarias y terciarias y puede comprender adecuadamente una o más unidades estructurales, como se ejemplifica mediante hélices alfa, láminas beta y/o barriles beta. En particular, el polipéptido puede comprender una pluralidad de hélices alfa, tales como al menos tres hélices alfa. Los dominios de unión a Fc derivados de la Proteína A de *Staphylococcus aureus* (por ejemplo, SEQ ID NO. 1-11, 48-64 y 74-93) comprende cada uno tres hélices alfa, de modo que el número de hélices alfa en un multímero como se ha descrito anteriormente puede ser tres veces el número de dominios en el multímero. La primera hélice alfa en los dominios como se ejemplifica por las SEQ ID NO. 1-11, 48-64 y 74-93 comienza en la posición 9 (usando la nomenclatura de posiciones de la Figura 15), que en el caso de las SEQ ID NO 1-7 y 9-10 es una glutamina y en las SEQ ID NO. 8, 11, 48-64 y 74-93 es una alanina.

El espaciador N-terminal tiene adecuadamente una longitud tal que el número de restos de aminoácidos entre el sitio de escisión del péptido señal y una unidad estructural proximal N-terminal del polipéptido de unión a Fc es de 14-24. Como se ha descrito anteriormente, esta unidad estructural puede ser adecuadamente una hélice alfa (o alternativamente una lámina beta o barril beta). El número de restos de aminoácidos entre el extremo N-terminal del polipéptido de unión a Fc y la unidad estructural proximal del extremo N-terminal puede variar, por ejemplo, entre 0 (SEQ ID NO. 9-11) y 11 (SEQ ID NO. 2), y por consiguiente, la longitud del espaciador N-terminal puede variar, por ejemplo, entre 3-24 restos de aminoácidos, tal como entre 8-24 restos de aminoácidos o 14-24 restos de aminoácidos. Por ejemplo, si el espaciador N-terminal está unido a un dominio de unión a Fc que tiene al menos un 90% de identidad con las SEQ ID NO. 1-8, 48-64 o 74-93, el espaciador N-terminal puede tener, por ejemplo, una longitud de 8-12 restos de aminoácidos y si el espaciador N-terminal está unido a un dominio de unión a Fc que tiene al menos un 90% de identidad con las SEQ ID NO. 9-11, el espaciador N-terminal puede tener, por ejemplo, una longitud de 16-20 restos de aminoácidos. El espaciador N-terminal puede consistir, por ejemplo, en restos de aminoácidos seleccionados del grupo que consiste en alanina, asparagina, ácido aspártico, glutamina, ácido glutámico, glicina, histidina, lisina, fenilalanina, serina, treonina, triptófano, tirosina y valina. Para una estabilidad con álcali mejorada puede ser ventajoso excluir asparaginas. En este caso, el espaciador N-terminal puede consistir, por ejemplo, en restos de aminoácidos seleccionados del grupo que consiste en alanina, ácido aspártico, glutamina, ácido glutámico, glicina, histidina, lisina, fenilalanina, serina, treonina, triptófano, tirosina y valina. Puede ser ventajoso si el espaciador N-terminal no comprende ningún agrupamiento de argininas o lisinas, y/o si comprende como máximo dos restos de aminoácidos seleccionados del grupo que consiste en lisina y arginina. Los dos restos de aminoácidos N-terminales del espaciador N-terminal son AQ (alanina seguida de glutamina).

En particular, el espaciador N-terminal puede comprender, comprender esencialmente o tener, una secuencia de aminoácidos que tiene al menos un 80% de identidad de secuencia con, o estar definida por, una secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NO. 16-18, 29-30, 33-40, 43-45 y 47. El espaciador N-terminal puede comprender además, comprender esencialmente o tener una secuencia de aminoácidos que tiene al menos un 80% de identidad de secuencia con, o estar definida por, una secuencia de aminoácidos seleccionada del grupo que consiste en AQYEDGKQYTADNKFNGE, AQYEDGKQYTVDAKFDKE, AQKDQTWYTGVDKFDKE, AQHDEAQQEAVDAKFDKE, AQGGGSGGGSVDAKFDKE, AQYEDGKQYGTVDKFDKE, AQYEDGKQGTVDKFDKE, AQYEDGKQYTTLEKGTVDKFDKE, AQYEDGKQYTTLEKPVAGGTVDKFDKE, AQYEDGKQYTVDAKFDKE, AQYEDGKQYTETVDKFDKE, AQYEDGKQYTDTVDAKFDKE, AQYEDGKQYTATVDKFDKE, AQYEDGKQYEDTVDAKFDKE, AQHHHHHHHHHGTVDKFDKE, AQHHHHHHHGTVDKFDKE y AQHDEAQQEAGTVDAKFDKE. Las últimas secuencias son particularmente ventajosas en combinación con polipéptidos de unión a Fc derivados de las SEQ ID NO. 9-11.

En términos de estabilidad alcalina, puede ser ventajoso además usar un espaciador N-terminal que comprende, comprende esencialmente o tiene, una secuencia de aminoácidos que tiene al menos un 80% de identidad de secuencia con, o estar definido por, una secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NO. 35, 37, 38, 40, 43, 44 y 47. El espaciador N-terminal para su estabilidad alcalina también puede comprender, comprender esencialmente o tener una secuencia de aminoácidos que tiene al menos un 80% de identidad de secuencia con, o estar definido por, una secuencia de aminoácidos seleccionada del grupo que consiste en AQYEDGKQYTVDAKFDKE, AQYEDGKQYTETVDKFDKE, AQYEDGKQYTDTVDAKFDKE,

AQYEDGKQYEDTVDKFDKE, AQHHHHHHHGTVDKFDKE, AQHHHHHHGTVDKFDKE y AQHDEAQQEAGTVDKFDKE. Las últimas secuencias son particularmente ventajosas en combinación con polipéptidos de unión a inmunoglobulina derivados de las SEQ ID NO. 9-11.

La estabilidad alcalina de la proteína recombinante puede evaluarse acoplándola a un chip de SPR, por ejemplo a un chip sensor Biacore CM5 tal y como se describe en los ejemplos del documento de Patente WO2016079033, usando, por ejemplo, químicas de acoplamiento de NHS- o maleimida, y midiendo la capacidad de unión a inmunoglobulina del chip, usando típicamente IgG policlonal humana, antes y después de la incubación en soluciones alcalinas a una temperatura específica, por ejemplo, 22 +/- 2°C. La incubación puede realizarse, por ejemplo, en NaOH 0,5 M durante un número de ciclos de 10 min, tal como 100, 200 o 300 ciclos. La capacidad de la IgG de la matriz después de 100 ciclos de incubación de 10 min en NaOH 0,5 M a 22 +/- 2°C puede ser al menos 55, tal como al menos 60, al menos 80 o al menos 90% de la capacidad de la IgG antes de la incubación. Alternativamente, la capacidad de la IgG restante después de 100 ciclos para un mutante particular medido como anteriormente puede compararse con la capacidad de la IgG restante para una proteína recombinante parental. En este caso, la capacidad de la IgG restante para el mutante puede ser al menos 105%, tal como al menos 110%, al menos 125%, al menos 150% o al menos 200% de la proteína recombinante parental.

En un segundo aspecto, la presente invención describe una molécula de ácido nucleico que codifica la proteína recombinante como se ha descrito anteriormente. La molécula de ácido nucleico comprende los siguientes elementos unidos operativamente en la dirección 5' a 3':

a) una secuencia de DNA promotora inducible o constitutiva, por ejemplo, la secuencia promotora RhaBAD o T5. RhaBAD es inducible con ramnosa y T5 con isopropil β-D-1-tiogalactopiranosido (IPTG). Un ejemplo de un promotor constitutivo es el promotor spa (el promotor de la Proteína A natural de *Staphylococcus aureus*);

b) una secuencia de DNA que codifica un péptido señal, por ejemplo, los péptidos señal OmpA (SEQ ID NO. 14) o DsbA (SEQ ID NO. 15), o un péptido señal que tiene al menos un 80% de identidad de secuencia con uno cualquiera de estos;

c) una secuencia de DNA que codifica un espaciador N-terminal como se ha descrito anteriormente; y

d) una secuencia de DNA que codifica un polipéptido de unión a Fc como se ha descrito anteriormente.

La molécula de ácido nucleico puede comprender además un sitio de unión a ribosomas (RBS) y un origen de replicación. Puede comprender también adecuadamente un marcador de resistencia a antibióticos.

En un tercer aspecto, la presente invención describe un vector de clonación, por ejemplo, un plásmido, que expresa y secreta la proteína recombinante como se ha descrito anteriormente en el periplasma bacteriano de una célula gram negativa, por ejemplo, *E. coli*. El vector de clonación comprende la molécula de ácido nucleico como se ha descrito anteriormente.

En un cuarto aspecto, la presente invención describe una bacteria gram negativa que comprende el vector de clonación descrito anteriormente. La bacteria puede identificarse, por ejemplo, como *Escherichia coli*, en particular *E. coli* de la cepa K12, tal como K12-017 de *E. coli*. Otros ejemplos de bacterias gram negativas incluyen, por ejemplo, el género *Pseudomonas*, tal como, por ejemplo, *Pseudomonas fluorescens*. La transformación puede lograrse, por ejemplo, mediante el método de choque térmico, aunque también son posibles otros métodos tales como electroporación.

En un quinto aspecto, la presente invención describe un método para expresar y secretar la proteína recombinante descrita anteriormente en una bacteria gram negativa. Este método comprende las etapas de:

i) proporcionar la bacteria gram negativa descrita anteriormente; y

ii) cultivar la bacteria gram negativa.

Si la molécula de ácido nucleico como se ha descrito anteriormente comprende un promotor inducible, el método puede comprender además una etapa de inducción de la expresión de la proteína recombinante en la bacteria gram negativa.

En un sexto aspecto, la presente invención describe una matriz de separación que comprende la proteína recombinante descrita anteriormente, unida covalentemente a un soporte.

Como entenderá el experto, la proteína recombinante expresada debería purificarse en un grado apropiado antes de inmovilizarse en un soporte. Dichos métodos de purificación son bien conocidos en el campo, y la inmovilización de ligandos basados en proteínas a soportes se lleva a cabo fácilmente usando métodos estándar. Los métodos y soportes adecuados se describirán a continuación con más detalle.

La estabilidad alcalina de la matriz se puede evaluar midiendo la capacidad de unión a inmunoglobulina, típicamente usando una IgG policlonal humana, antes y después de la incubación en soluciones alcalinas a una temperatura específica, por ejemplo, 22 +/- 2°C. La incubación se puede realizar, por ejemplo, en NaOH 0,5 M durante un número

de ciclos de 15 min, tal como 100, 200 o 300 ciclos, correspondientes a un tiempo de incubación total de 25, 50 o 75 h. La capacidad de la IgG de la matriz después de 96-100 ciclos de incubación de 15 min o un tiempo de incubación total de 24 o 25 h en NaOH 0,5 M a 22 +/- 2°C puede ser al menos 80, tal como al menos 85, al menos 90 o al menos 95% de la capacidad de la IgG antes de la incubación.

El soporte sólido de la matriz según la invención puede ser de cualquier tipo adecuado bien conocido. Una matriz de separación por afinidad convencional es a menudo de naturaleza orgánica y se basa en polímeros que exponen una superficie hidrófila al medio acuoso usado, es decir, exponen hidroxilo (-OH), carboxi (-COOH), carboxamido (-CONH₂, posiblemente en formas N-sustituidas), amino (-NH₂, posiblemente en forma sustituida), grupos oligo- o polietileno en sus superficies externas y, si están presentes, también en superficies internas. El soporte sólido puede ser adecuadamente poroso. La porosidad puede expresarse como un valor de Kav o Kd (la fracción del volumen de poro disponible para una molécula sonda de un tamaño particular) medida mediante cromatografía de exclusión por tamaño inversa, por ejemplo, según los métodos descritos en *Gel Filtration Principles and Methods*, Pharmacia LKB Biotechnology 1991, pp. 6-13. Por definición, tanto los valores de Kd como de Kav siempre se encuentran dentro del intervalo 0-1. El valor de Kav puede ser ventajosamente de 0,6 - 0,95, por ejemplo, 0,7 - 0,90 o 0,6 - 0,8, medido con dextrano de Mw 110 kDa como una molécula sonda. Una ventaja de esto es que el soporte tiene una gran fracción de poros capaces de acomodar tanto la proteína recombinante de la invención como las inmunoglobulinas que se unen a la proteína recombinante y proporcionar el transporte de masa de las inmunoglobulinas hacia y desde los sitios de unión.

La proteína recombinante puede unirse al soporte mediante técnicas de acoplamiento convencionales utilizando, por ejemplo, grupos tiol, amino y/o carboxi presentes en el ligando. Los bisepóxidos, epiclorhidrina, CNBr, N-hidroxisuccinimida (NHS) etc, son reactivos de acoplamiento bien conocidos. Entre el soporte y la proteína recombinante, se puede introducir una molécula conocida como espaciador, que mejora la disponibilidad de la proteína recombinante y facilita el acoplamiento químico de la proteína recombinante al soporte. Dependiendo de la naturaleza de la proteína recombinante y de las condiciones de acoplamiento, el acoplamiento puede ser un acoplamiento multipunto (por ejemplo, a través de una pluralidad de lisinas) o un acoplamiento de un solo punto (por ejemplo, a través de una única cisteína). Alternativamente, la proteína recombinante puede unirse al soporte mediante enlace no covalente, tal como adsorción física o adsorción bioespecífica.

En algunas realizaciones, la matriz comprende 5 - 25, tal como 5 - 20 mg/ml, 5 - 15 mg/ml, 5 - 11 mg/ml o 6 - 11 mg/ml de la proteína recombinante acoplada al soporte. La cantidad de proteína acoplada puede controlarse mediante la concentración de proteína usada en el proceso de acoplamiento, mediante las condiciones de activación y acoplamiento usadas y/o mediante la estructura de poros del soporte usado. Como regla general, la capacidad de unión absoluta de la matriz aumenta con la cantidad de proteína acoplada, al menos hasta un punto en donde los poros se estrechan significativamente por la proteína acoplada. La capacidad de unión relativa por mg de proteína acoplada disminuirá a niveles de acoplamiento altos, dando como resultado un coste-beneficio óptimo dentro de los intervalos especificados anteriormente.

En determinadas realizaciones, la proteína se acopla al soporte a través de enlaces tioéter. Los métodos para realizar dicho acoplamiento son bien conocidos en este campo y se realizan fácilmente por el experto en este campo usando técnicas y equipos estándar. Los enlaces tioéter son flexibles y estables y generalmente adecuados para su uso en cromatografía de afinidad. En particular, cuando el enlace tioéter es a través de un resto de cisteína terminal o casi terminal en la proteína recombinante, la movilidad de la proteína recombinante acoplada se mejora, lo que proporciona una mejora de la capacidad de unión y la cinética de unión. En algunas realizaciones, la proteína recombinante se acopla a través de una cisteína C-terminal proporcionada en la proteína como se ha descrito anteriormente. Esto permite un acoplamiento eficaz del tiol de la cisteína a grupos electrófilos, por ejemplo, grupos epóxido, grupos halohidrina, etc., sobre un soporte, dando como resultado un acoplamiento de puente tioéter. Alternativamente, la proteína recombinante puede estar unida covalentemente al soporte a través de uno o más enlaces amida. Esto se puede lograr, por ejemplo, mediante la reacción entre una o más lisinas en la proteína y uno o más grupos carboxilo activados en el soporte. La activación puede realizarse, por ejemplo, mediante el reactivo de N-hidroxisuccinimida (NHS) conocido comúnmente. Otra alternativa más es que la proteína esté unida a través de uno o más enlaces de amina secundaria. Dichos enlaces pueden formarse a partir de lisinas en la proteína y o bien grupos hidroxilo en el soporte, que se han activado usando, por ejemplo, química de cloruro de tresilo o cloruro de tosilo, o mediante reacción de aminación reductora entre las lisinas y aldehídos en el soporte. Los aldehídos pueden formarse, por ejemplo, a partir de dioles vecinales en el soporte mediante oxidación con periodato.

En determinadas realizaciones, el soporte comprende un polímero de polihidroxiolos, tal como un polisacárido. Los ejemplos de polisacáridos incluyen, por ejemplo, dextrano, almidón, celulosa, pululano, agar, agarosa, etc. Los polisacáridos son inherentemente hidrófilos con bajos grados de interacciones no específicas, proporcionan un alto contenido de grupos hidroxilo reactivos (activables) y generalmente son estables frente a soluciones de limpieza alcalinas usadas en el bioprocesamiento.

En algunas realizaciones, el soporte comprende agar o agarosa. Los soportes usados en la presente invención pueden prepararse fácilmente según los métodos estándar, tales como gelificación en suspensión inversa (S Hjertén: *Biochim Biophys Acta* 79(2), 393-398 (1964). Alternativamente, las matrices base son productos disponibles comercialmente, tales como perlas de agarosa reticuladas vendidas bajo el nombre de SEPHAROSE™ FF (GE Healthcare). En una

realización, que es especialmente ventajosa para separaciones a gran escala, el soporte se ha adaptado para aumentar su rigidez usando los métodos descritos en los documentos de Patente US6602990 o US7396467 y por lo tanto hace la matriz más adecuada para caudales altos.

- 5 En determinadas realizaciones, el soporte, tal como un soporte de polisacárido o de agarosa, está reticulado, tal como con reticulaciones de hidroxialquil éter. Los reactivos reticulantes que producen dichas reticulaciones pueden ser, por ejemplo, epihalohidrinatos como epiclorhidrina, diepóxidos como butanodiol diglicidil éter, reactivos alilantes como haluros de alilo o alil glicidil éter. La reticulación es beneficiosa para la rigidez del soporte y mejora la estabilidad química. Las reticulaciones de hidroxialquil éter son estables frente a álcalis y no provocan una adsorción no específica significativa.
- 10 Alternativamente, el soporte sólido se basa en polímeros sintéticos, tales como alcohol polivinílico, acrilatos de polihidroxialquilo, metacrilatos de polihidroxialquilo, poliácridamidas, polimetacrilamidas, etc. En el caso de polímeros hidrófobos, tales como matrices basadas en bencenos sustituidos con divinilo y monovinilo, la superficie de la matriz a menudo se hidrofiliiza para exponer grupos hidrófilos como se ha definido anteriormente a un líquido acuoso circundante. Dichos polímeros se producen fácilmente según métodos estándar, véase, por ejemplo, "Styrene based polymer supports developed by suspension polymerization" (R Arshady: *Chimica e L'Industria* 70(9), 70-75 (1988)).
- 15 Alternativamente, se utiliza un producto disponible comercialmente, tal como SOURCE™ (GE Healthcare). En otra alternativa, el soporte sólido según la invención comprende un soporte de naturaleza inorgánica, por ejemplo, sílice, óxido de circonio, etc.

- 20 En determinadas realizaciones, el soporte sólido está en otra forma tal como una superficie, un chip, capilares o un filtro (por ejemplo, una membrana o una matriz de filtro de profundidad).

En un séptimo aspecto, la presente invención describe un método para separar una inmunoglobulina, que comprende las etapas de:

- 25 i) proporcionar la matriz de separación como se ha descrito anteriormente, en donde la proteína recombinante comprende un polipéptido de unión a inmunoglobulina, por ejemplo, que comprende uno o más dominios de unión a Fc derivados de la Proteína A de *Staphylococcus aureus*;
- ii) poner en contacto la matriz de separación con una muestra líquida que contiene una inmunoglobulina, para unir la inmunoglobulina;
- iii) opcionalmente lavar la matriz de separación con un líquido de lavado;
- iv) poner en contacto la matriz de separación con un líquido de elución, para eluir la inmunoglobulina;
- 30 v) opcionalmente limpiar la matriz de separación con un líquido de limpieza. Los líquidos de limpieza alcalinos se usan comúnmente en el bioprocesamiento y, siempre que la proteína recombinante sea estable frente a álcalis, el líquido de limpieza puede comprender al menos NaOH o KOH 0,1 M, tal como al menos NaOH o KOH 0,5 M, o NaOH o KOH 0,5-2,5 M.

EJEMPLOS

- 35 En primer lugar, se ensayaron dos péptidos señal diferentes junto con cuatro espaciadores N-terminales diferentes para encontrar la mejor escisión y expresión de proteínas. Este experimento (Ejemplo 1) se realizó con el monómero de Zvar2 usando el sistema promotor RhaBAD. En el Ejemplo 2, el péptido señal más prometedor y el inicio N-terminal se clonaron en el hexámero de Zvar2. Este constructo se ensayó en fermentaciones con alimentación por lotes con y sin la secuencia de inicio N-terminal. Este estudio se realizó usando el sistema de expresión del promotor T5. En el
- 40 Ejemplo 3, todos los dominios nativos de la Proteína A de *Staphylococcus aureus*, monómero de Zvar (Zvar₁), monómero de Zvar2 (Zvar₂) y tetrámero de Zvar (Zvar₄) se ensayaron con el péptido señal y el inicio N-terminal elegidos para ver si la secuencia de inicio también tenía un efecto sobre dominios estrechamente relacionados. También se realizó el Ejemplo 3 con el sistema de expresión de T5. El Ejemplo 4 se realizó con un conjunto de diferentes espaciadores N-terminales, el Ejemplo 5 fue un escalado de fermentaciones usando un par de espaciadores
- 45 seleccionados y el Ejemplo 6 fue una investigación usando el espaciador N-terminal AQYEDGKQYTGT en combinación con diferentes mutantes de una proteína de unión a inmunoglobulina.

Materiales y métodos

Constructos

- 50 Los genes como se describen en la Tabla 1 fueron sintetizados por un fabricante contratado de genes sintéticos (ATUM, CA, USA). Se sintetizó ácido didesoxiribonucleico (dsDNA) bicatenario basándose en la secuencia de aminoácidos (AA) y se optimizó para la expresión en *E. coli* mediante un algoritmo patentado por el fabricante. El dsDNA se insertó en el vector de expresión pD861-SR o pJ401 con un péptido señal de DsbA u OmpA (Tabla 2).

Tabla 1. Plásmidos pedidos a ATUM.

Número de plásmido	Inserto	Alias	Promotor	Péptido señal	Inicio N-terminal	Vector de clonación	Secuencia del inserto
pGE0002	Zvar2 ₆	Zvar2 ₆	T5	OmpA	AQGT	pJ401	(VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQSLKDDPSV SKAHLAEAKK LNDQAQPK) _{x6C}
pGE0068	Zvar2 ₁	Zvar2 ₁	RhaBAD	DsbA	AQGT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQSLKDDPSV SKAHLAEAKK LNDQAQPKC
pGE0091	Zvar4_Opt Ec	Zvar4	T5	OmpA	AQGT	pJ401	(VDAKFDKEQO NAFYEILHLP NLTEEQRNAF IQSLKDDPSQ SANLLAEAKK LNDQAQPK) _{x4C}
pGE0096	Zvar1_Opt Ec	Zvar1	T5	OmpA	AQGT	pJ401	VDAKFDKEQO NAFYEILHLP NLTEEQRNAF IQSLKDDPSQ SANLLAEAKK LNDQAQPKC
pGE0119	Zvar2 ₁	Zvar2 ₁	RhaBAD	OmpA	AQGT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQSLKDDPSV SKAHLAEAKK LNDQAQPKC
pGE0120	Zvar2 ₁	Zvar2 ₁	T5	OmpA	AQGT	pJ401	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQSLKDDPSV SKAHLAEAKK LNDQAQPKC
pGE0127	Dominio E	Dominio E	T5	OmpA	AQGT	pJ401	*YDEAQQNAFY QVLNMPNLNA DQRNGFIQSLKDDPSQSANV LGEAQKLNDQ APKC
pGE0128	Dominio D	Dominio D	T5	OmpA	AQGT	pJ401	*YDAQQNNFNK DQQAIFYELL NMPNLNEEQR NGFIQSLKDD PSQSTNVTIGE AKKLNESQA PKC
pGE0129	Dominio A	Dominio A	T5	OmpA	AQGT	pJ401	*YDNNFNKEQO NAFYEILNMP NLNEEQRNGF IQSLKDDPSQ SANLLAEAKK LNFQAQPKC
pGE0130	Dominio B	Dominio B	T5	OmpA	AQGT	pJ401	*YDNKFNKEQO NAFYEILHLP NLNEEQRNGF IQSLKDDPSQ SANLLAEAKK LNDQAQPKC
pGE0131	Dominio C	Dominio C	T5	OmpA	AQGT	pJ401	*YDNKFNKEQO NAFYEILHLP NLTEEQRNGF IQSLKDDPSV SKHLAEAKK LNDQAQPKC
* Los dominios E, D, A, B y C tienen una mutación A1V para añadir un sitio de restricción en el inicio de la secuencia (VD) y el dominio D tiene un D insertado, para tener sitios de clonación múltiple uniformes de los vectores.							

Tabla 2. Péptidos señal.

Nombre	Secuencia AA	SEQ ID NO
OmpA	MKKTAIAIAVALAGFATVAQA	14
DsbA	MKKIWLALAGLVLAFSASA	15

Clonación de los extremos N-terminales

Todos los plásmidos se transformaron en K12-017 de *E. coli* usando células químicamente competentes. Para modificar los vectores, se insertó una secuencia de espaciador corto de 8 AA entre los sitios de escisión FspI y KpnI. Las secuencias de 8 AA se tomaron a partir del primer AA en la proteína madura de OmpA (AQKDQTWYTGTT, SEQ ID NO 16), DsbA (AQYEDGKQYTGT, SEQ ID NO 17), SpA (AQHDEAQQEAGT, SEQ ID NO 18) o la estructura enlazadora flexible del fago GIII AA 236-243 de M13 (AQGGGSGGGSGT, SEQ ID NO 19), con modificaciones en OmpA para ajustar el sitio de restricción y en OmpA y SpA para mutar asparaginas a glutaminas (Figura 1). Además, se clonó una variedad de diferentes secuencias en el plásmido pGE0002 (Tabla 1).

Digestión de plásmidos

Los plásmidos se escindieron con las enzimas de restricción FspI y KpnI (New England Biolabs (NEB), MA, USA), se mezclaron 6 µl de tampón NEB 2.1 con 6 µg de plásmido y 2 µl de FspI en un total de 58 µl. La solución se incubó a 37°C durante 1 hora antes de añadir KpnI. La incubación continuó durante 2 h, seguido de la adición de 1 µl de fosfatasa de intestino de ternero (CIP) seguido de 30 min de incubación adicional. Las bases escindidas se retiraron de los plásmidos digeridos usando el kit de purificación por PCR QIAquick (Qiagen, Hilden, Alemania).

Hibridación

Los oligonucleótidos se pidieron a Integrated DNA technologies (IDT, IA, USA). Todos los oligonucleótidos fueron modificados por el fabricante con un grupo fosfato cinco prima. Se mezclaron dos pares de oligonucleótidos complementarios en tampón de ligación y se calentaron a 95°C, 4 min, seguido de enfriamiento a temperatura ambiente. El fragmento hibridado se unió en los plásmidos escindidos con FspI y KpnI mediante el uso de DNA ligasa de T4 (NEB, MA, USA). Una lista completa de los oligonucleótidos usados se encuentra en la Tabla 3.

Tabla 3. Lista de oligonucleótidos usados para la adición de la secuencia de inicio N-terminal.

Oligo	Nombre	Secuencia	SEQ ID NO
GEOLI_191	Nterm_Omp_F	GCAGAAAGATCAGACCTGGTACACCGGCGGTAC	20
GEOLI_192	Nterm_Omp_R	CGCCGGTGTACCAGGTCTGATCTTTCTGC	21
GEOLI_193	Nterm_SPA_F	GCAGCATGATGAAGCGCAGCAGGAAGCGGGTAC	22
GEOLI_194	Nterm_SPA_R	CCGCTTCCTGCTGCGCTTCATCATGCTGC	23
GEOLI_195	Nterm_dsbA_F	GCAGTATGAAGATGGCAAACAGTACACCGGTAC	24
GEOLI_196	Nterm_dsbA_R	CGGTGTACTGTTTGCCATCTTCATACTGC	25
GEOLI_197	Nterm_flex_F	GCAGGGTGGCGGTTCTGGCGGTGGCAGCGGTAC	26
GEOLI_198	Nterm_flex_R	CGCTGCCACCGCCAGAACCGCCACCCTGC	27
GEOLI_284	Nterm_dsbA7_F	GCAGTATGAAGATGGCAAACAGTACGGTAC	94
GEOLI_285	Nterm_dsbA7_R	CGTACTGTTTGCCATCTTCATACTGC	95
GEOLI_286	Nterm_dsbA6_F	GCAGTATGAAGATGGCAAACAGGGTAC	96
GEOLI_287	Nterm_dsbA6_R	CCTGTTTGCCATCTTCATACTGC	97
GEOLI_288	Nterm_dsbA5_F	GCAGTATGAAGATGGCAAAGGTAC	98
GEOLI_289	Nterm_dsbA5_R	CTTTGCCATCTTCATACTGC	99
GEOLI_290	Nterm_dsbA4_F	GCAGTATGAAGATGGCGGTAC	100

ES 2 986 499 T3

GEOLI_291	Nterm_dsbA4_R	CGCCATCTTCATACTGC	101
GEOLI_292	Nterm_dsbA8_noGT_F	GCAGTATGAAGATGGCAAACAGTACACCGT	102
GEOLI_293	Nterm_dsbA8_noGT_R	CTACGGTGTACTGTTTGCCATCTTCATACTGC	103
GEOLI_294	Nterm_dsbA6_noGT_F	GCAGTATGAAGATGGCAAACAGGT	104
GEOLI_295	Nterm_dsbA6_noGT_R	CTACCTGTTTGCCATCTTCATACTGC	105
GEOLI_296	Nterm_dsbA8_ET_F	GCAGTATGAAGATGGCAAACAGTACACCGAAACCGT	106
GEOLI_297	Nterm_dsbA8_ET_R	CTACGGTTTCGGTGTACTGTTTGCCATCTTCATAC TGC	107
GEOLI_298	Nterm_dsbA8_DT_F	GCAGTATGAAGATGGCAAACAGTACACCGATAACCGT	108
GEOLI_299	Nterm_dsbA8_DT_R	CTACGGTATCGGTGTACTGTTTGCCATCTTCATAC TGC	109
GEOLI_300	Nterm_dsbA8_AT_F	GCAGTATGAAGATGGCAAACAGTACACCGCGACCGT	110
GEOLI_301	Nterm_dsbA8_AT_R	CTACGGTTCGGGTGTACTGTTTGCCATCTTCATAC TGC	111
GEOLI_302	Nterm_dsbA7_EDT_F	GCAGTATGAAGATGGCAAACAGTACGAAGATACCGT	112
GEOLI_303	Nterm_dsbA7_EDT_R	CTACGGTATCTTCGTACTGTTTGCCATCTTCATAC TGC	113
GEOLI_304	Nterm_dsbA12_F	GCAGTATGAAGATGGCAAACAGTACACCACCCTG GAAAAAGGTAC	114
GEOLI_305	Nterm_dsbA12_R	CTTTTTCACGGGTGGTGTACTGTTTGCCATCTTCA TACTGC	115
GEOLI_306	Nterm_dsbA16_F	GCAGTATGAAGATGGCAAACAGTACACCACCCTG GAAAAACCGGTGGCGGGCGGTAC	116
GEOLI_307	Nterm_dsbA16_R	CGCCCCCACCAGGTTTTTCCAGGGTGGTGTACTGT TTGCCATCTTCATACTGC	117
GEOLI_308	Nterm_R8_F	GCAGCGCCGTCGTCGCCGTCGCCGTCGTGGTAC	118
GEOLI_309	Nterm_R8_R	CACGACGGCGACGGCGACGACGGCGCTGC	119
GEOLI_310	Nterm_K8_F	GCAGAAGAAGAAAAAGAAGAAGAAAAAGGGTAC	120
GEOLI_311	Nterm_K8_R	CCTTTTTCTTCTCTTTTTCTTCTTCTGC	121
GEOLI_312	Nterm_H8_F	GCAGCATCACCACCATCACCATCACCATGGTAC	122
GEOLI_313	Nterm_H8_R	CATGGTGATGGTGATGGTGGTGATGCTGC	123

GEOLI_314	Nterm_H6_F	GCAGCATCACCACCATCACCATGGTAC	124
GEOLI_315	Nterm_H6_R	CATGGTGATGGTGGTGGTGGTGGTGC	125
GEOLI_316	Nterm_H4_F	GCAGCATCACCACCATGGTAC	126
GEOLI_317	Nterm_H4_R	CATGGTGGTGGTGGTGGTGGTGGTGC	127
GEOLI_318	Nterm_Strep_F	GCAGTGGAGCCATCCGCAGTTTGAAAAAGGTAC	128
GEOLI_319	Nterm_Strep_R	/5Phos/CTTTTTCAAACTGCGGATGGCTCCACTGC	129

Constructos clonados

- 5 Los plásmidos ligados se transformaron en células K12-017 de *E. coli* químicamente competentes. Las células se descongelaron en hielo durante 30 min, se añadieron 100 µl de células competentes a 10 µl de reacción de ligación. Las células se incubaron en hielo durante 20 min seguido de choque térmico a 42°C durante 1 min, seguido de incubación en hielo 5 min. Seguido de la adición de 900 µl de medio SOC (NEB, MA, USA). La reacción de transformación se incubó a 37°C en una incubadora con agitación rotatoria durante 60 min, se extendieron 100 µl de cada reacción sobre placas de agar Luria que contenían antibióticos selectivos apropiados. Los clones positivos se seleccionaron con reacción en cadena de la polimerasa (PCR) y se seleccionaron los clones con inserto correcto. Los clones seleccionados se cultivaron en 10 ml de caldo Luria (LB) durante la noche (o/n) que contenía los antibióticos apropiados, seguido de preparaciones de plásmidos usando el kit Qiagen Plasmid Miniprep (Qiagen, Hilden, Alemania). Los plásmidos se enviaron a GATC Biotech para la verificación de secuencia (GATC Biotech, Colonia, Alemania).

Tabla 4. Plásmidos resultantes después de la adición de la secuencia de inicio N-terminal.

Número de plásmido	Inserto	Alias	Promotor	Péptido señal	Inicio N-terminal	Vector de clonación	Secuencia del inserto
pGE0138	Zvar2 ₁	OmpA8AA-Zvar2 ₁	RhaBAD	DsbA	AQKDQWTYGGT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQLKDDPSV SKAILAEAKK LNDQAAPKC
pGE0139	Zvar2 ₁	SpA8AA-Zvar2 ₁	RhaBAD	DsbA	AQHDEAQQEAGT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQLKDDPSV SKAILAEAKK LNDQAAPKC
pGE0140	Zvar2 ₁	DsbA8AA-Zvar2 ₁	RhaBAD	DsbA	AQYEDGKQYTGT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQLKDDPSV SKAILAEAKK LNDQAAPKC
pGE0141	Zvar2 ₁	Flex8AA-Zvar2 ₁	RhaBAD	DsbA	AQGGGSGGGSGT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQLKDDPSV
							SKAILAEAKK LNDQAAPKC

pGE0142	Zvar2 ₁	OmpA8AA-Zvar2 ₁	RhaBAD	OmpA	AQKDQTWYTGTT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAP IQSLKDDPSV SKAILAEAKK LNDQAAPKC
pGE0143	Zvar2 ₁	SpA8AA-Zvar2 ₁	RhaBAD	OmpA	AQHDEAQQEAGT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAP IQSLKDDPSV SKAILAEAKK LNDQAAPKC
pGE0144	Zvar2 ₁	DsbA8AA-Zvar2 ₁	RhaBAD	OmpA	AQYEDGKQYTGT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAP IQSLKDDPSV SKAILAEAKK LNDQAAPKC
pGE0145	Zvar2 ₁	Flex8AA-Zvar2 ₁	RhaBAD	OmpA	AQGGGSGGGSGT	pD861-SR	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAP IQSLKDDPSV SKAILAEAKK LNDQAAPKC
pGE0180	Zvar2 ₆	DsbA8AA-Zvar2 ₆	T5	OmpA	AQYEDGKQYTGT	pJ401	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAP IQSLKDDPSV SKAILAEAKK LNDQAAPK _{186C}
pGE0293	Dominio E	Dominio E de DsbA8AA	T5	OmpA	AQYEDGKQYTGT	pJ401	VDEAQQNAPFY QVLEMPNLNA DQRNGFIQSL KDDPSQSANY LGEAQKLNDSSQ APKC
pGE0294	Dominio D	Dominio D de DsbA8AA	T5	OmpA	AQYEDGKQYTGT	pJ401	VDAQQNNFNK DQQSIFYEIL NMPNLNEEQR NGFIQSLKDD PSQSTNVLGE AKKLNLSQA PKC
pGE0295	Dominio A	Dominio A de DsbA8AA	T5	OmpA	AQYEDGKQYTGT	pJ401	VDNENKEQQ NIFYEILNMP NLNEEQRNGF IQSLKDDPSQ SANLLAEAKK LNEQAAPKC
pGE0296	Dominio B	Dominio B de DsbA8AA	T5	OmpA	AQYEDGKQYTGT	pJ401	VDNKNKEQQ NIFYEILHLP NLNEEQRNGF IQSLKDDPSQ SANLLAEAKK LNDQAAPKC
pGE0297	Dominio C	Dominio C de DsbA8AA	T5	OmpA	AQYEDGKQYTGT	pJ401	VDNKNKEQQ NIFYEILHLP NLTEEQRNGF IQSLKDDPSV SKEILAEAKK LNDQAAPKC

pGE0298	Zvar ₁ _OptExp	DsbA8AA-Zvar ₁	T5	OmpA	AQYEDGKQYTGT	pJ401	VDAKFDKEQQ NAFYEILHLP NLTEEQRNAF IQSLKDDPSQ SANLLAEAKK LNDAAQAPKC
pGE0299	Zvar ₂	DsbA8AA-Zvar ₂	T5	OmpA	AQYEDGKQYTGT	pJ401	VDAKFDKEAQ EAFYEILHLP NLTEEQRNAF IQSLKDDPSV SKALLAEAKK LNDAAQAPKC
pGE0300	Zvar ₄ _OptEc	DsbA8AA-Zvar ₄	T5	OmpA	AQYEDGKQYTGT	pJ401	(VDAKFDKEQQ NAFYEILHLP NLTEEQRNAF IQSLKDDPSQ SANLLAEAKK LNDAAQAPK)x4C

Ejemplo 1

Expresión y purificación de proteínas en matraces de agitación

La cepa K12-017 de *Escherichia coli* transformada con los plásmidos recombinantes pGE0138-145 se cultivó en 30°C durante 4 h en 100 ml de medio de caldo Terrific (TB) (que contiene 12 g de triptona, 24 g de extracto de levadura, 5 g de glicerol (85%), 2,31 g de KH₂PO₄, 12,54 g de K₂HPO₄ y 50 mg de sulfato de kanamicina por litro). Los cultivos se indujeron cuando la OD_{600 nm} alcanzó 1-2, usando L-ramnosa 4 mM (Sigma Aldrich, MO, USA). Los cultivos se incubaron adicionalmente en 30°C 17-20 h. La solución de cultivo se sometió a centrifugación a baja velocidad (a 4000 rpm) durante 20 min en un rotor oscilante para recoger el sedimento celular húmedo. Los sedimentos celulares bacterianos se suspendieron en 20 ml de una solución tampón de fosfato 25 mM (pH 7,4), y las células se lisaron mediante tratamiento térmico a 85°C durante 10 min en un bloque térmico. Seguido de una centrifugación a alta velocidad (a 13000 rpm) durante 10 min para separar el sobrenadante. Después, los sobrenadantes se filtraron a través de un filtro de jeringa de 0,2 µm para eliminar cualquier partícula residual antes de aplicarlos a una columna de cromatografía de afinidad de IgG Sepharose 6FF (GE Healthcare Bio-Sciences, Uppsala, Suecia). La columna se equilibró con tampón de carga (fosfato 25 mM pH 7, NaCl 250 mM) seguido de la carga de la muestra. Después de la carga, la columna se lavó con 5 volúmenes de columna (CV) de tampón de carga y un CV de tampón de lavado bajo en sal (acetato 50 mM pH 6), seguido de elución con ácido acético 50 mM, pH 2,8. La absorbancia a 237 nm se midió en línea usando un sistema de cromatografía ÄKTA explorer 100 (GE Healthcare Bio-Sciences, Uppsala, Suecia) y la integración del pico del pico eluido se realizó en el software del sistema (Unicorn 5.1).

El matraz de agitación da como resultado 2 péptidos señal y 4 inicios N-terminales

En el primer trabajo de optimización del sistema de expresión, se probaron dos péptidos señal diferentes, utilizando diferentes rutas de secreción en la célula. DsbA utiliza la ruta de partículas de reconocimiento de señal (SRP) y OmpA la ruta Sec. Los resultados de la expresión de proteínas usando los péptidos señal DsbA y OmpA, mostraron que el péptido señal OmpA tenía una mayor expresión de proteínas; 2-5 veces mayor en comparación con el péptido señal DsbA. Además, el primer AA N-terminal después del sitio de escisión del péptido señal tuvo un impacto significativo en los niveles de expresión. En este estudio, el AA de inicio de DsbA tenía el nivel de expresión más alto junto con ambos péptidos señal. Se observó el nivel de expresión más bajo para la secuencia de inicio de Flex8AA usando ambos péptidos señal.

Resultados de la espectrometría de masas acoplada a cromatografía líquida (LC/MS)

Los eluatos de la IgG Sepharose 6FF se analizaron usando LC/MS (Waters, PA, USA). Los resultados mostraron que el eluato del constructo de referencia pGE0120, con AQGT como secuencia de inicio N-terminal tenía un intervalo de diferentes sitios de escisión del péptido señal (Figura 3; a). El área integrada del pico con masa correcta fue 52% del área total con péptido señal procesado correctamente. Otros picos correspondieron a sitios de escisión con 6 AA extra, 9 AA extra, péptido señal intacto (21 AA extra) y 7 AA perdidos del constructo principal. Después de la adición de 8-AA de DsbA, el péptido señal se escindió correctamente hasta el 96% (Figura 3; b). Se observó un resultado similar con el péptido señal DsbA y 8AA del extremo N-terminal DsbA, con el 97% del péptido señal correctamente escindiado.

Ejemplo 2**Expresión de proteínas en fermentador con alimentación por lotes**

Se usaron seis fermentadores de 1 l de volumen de trabajo, GRETA (Belach Bioteknik, SkogAs, Suecia). El volumen de partida se ajustó a 750 ml, y el volumen final fue de aproximadamente 1 l. La aireación se ajustó a 1 L/min, la temperatura, el pH y el antiespumante se controlaron de manera automática. El pH se mantuvo en el valor ajustado mediante la adición de amoníaco al 25% y ácido fosfórico 2 M. El control antiespumante, Breox FMT 30 (BASF, Ludwigshafen, Germany), se añadió de manera automática cuando se producía una alta formación de espuma. El pH y el oxígeno disuelto (DO) se controlaron con sondas de Broadley James (CA, USA). El DO se mantuvo constante al 30% aumentando la velocidad del agitador desde 300 rpm hasta 1500 rpm. Cuando el agitador alcanzó la velocidad máxima, el segundo punto de ajuste de DO fue del 20%, y se mantuvo constante mediante la adición de oxígeno puro mezclado en el flujo de aire. Se usó caldo Terrific (TB) suplementado con 50 mg/L de kanamicina o neomicina en los precultivos del matraz de agitación y se inició mediante la adición de 100 µl de suspensión celular a 100 mL de TB suplementado con 50 µg/mL de kanamicina o neomicina y se incubó a 37°C durante 17 h. Se inoculó el medio de fermentación principal con 10 ml de precultivo. La alimentación de glucosa con 60% (p/v) (VWR, PA, USA) se alimentó al fermentador según un perfil preestablecido después de que se consumió la glucosa del lote inicial. La tiempo total de la fermentación principal fue de 26 h.

Análisis de concentración usando una curva estándar

Las muestras de fermentación se trataron térmicamente a 85°C durante 5 min en tubos de 1,5 ml. Seguido de centrifugación a alta velocidad (a 13500 rpm, en centrífuga de mesa) durante 5 min para separar el sobrenadante. Los sobrenadantes se filtraron a través de un filtro de jeringa de 0,2 µm para eliminar cualquier partícula residual antes de aplicarlos a una columna de IgG Sepharose 6FF™ (GE Healthcare Bio-Sciences, Uppsala Suecia). La columna se equilibró con tampón de carga (solución salina tamponada con fosfato (Medicago, Uppsala, Suecia)) seguido de la carga de 50 µl de muestra. Después de la carga, la columna se lavó con 5 CV de tampón de carga, seguido de elución con tampón fosfato 200 mM, pH 2,9. Se midió la absorbancia a 237 nm en línea y se realizó la integración del pico del pico eluido en el software del sistema de cromatografía. La concentración de proteína de las muestras se determinó mediante una curva estándar usando una proteína purificada de concentración conocida.

Resultados de fermentación, expresión de proteínas

El plásmido pGE0002 que contiene el hexámero de Zvar2 (Zvar2₆) se escindió con las enzimas de restricción FspI y KpnI. El inicio N-terminal de DsbA se ligó en el vector escindido, y el plásmido resultante se denominó pGE0180. Estos dos vectores se transformaron en K12-017 de *E. coli* y los constructos resultantes se expresaron en fermentaciones de alimentación por lotes para ver diferencias en la expresión de proteínas y los patrones de escisión de péptidos señal. El análisis de concentración de proteína usando una curva estándar mostró que la expresión de proteína aumentó de 3,5 g/L a 16,8 g/L, es decir, un aumento de más de 4 veces en la expresión de proteína (Figura 4).

Resultados de LC/MS

Los eluatos de IgG Sepharose se analizaron usando LC/MS (Waters, PA, USA). Los resultados mostraron que el eluato del constructo de referencia pGE0002 mostró un intervalo de diferentes sitios de péptido señal incluyendo sitios de escisión con 6 AA y 9 AA extra, observados previamente en el experimento 1 (Figura 5, a). Sin embargo, cuando se añade 8-AA de DsbA en el extremo N-terminal, el péptido señal se escindió correctamente, dando como resultado un pico singular claro de masa correcta (Figura 5; b).

Ejemplo 3

Expresión de proteínas en fermentador con alimentación por lotes y análisis de concentración usando una curva estándar realizada como en el Ejemplo 2.

Preparación de proteínas

Se aplicaron muestras filtradas (véase anteriormente) a una columna IgG Sepharose 6FF Tricorn 10 con un CV de 10,5 ml (GE Healthcare Bio-Sciences, Uppsala, Suecia). La columna se equilibró con tampón de carga de alto contenido en sal (fosfato 50 mM, pH 7,0, NaCl 500 mM) seguido de la carga de 0,5-5 mL de muestra basándose en cálculos de concentración (véase anteriormente). Después de la carga, la columna se lavó con 5 CV de tampón de carga, seguido de elución con ácido acético 100 mM. Las muestras eluidas se recogieron y se concentraron a aproximadamente 1 mg/ml usando Vivaspin 5, corte de 3000 Da (GE Healthcare Bio-Sciences, Uppsala, Suecia). La cantidad de proteína se estimó mediante mediciones de absorbancia a 280 nm usando coeficientes de absorbancia específicos, calculados basándose en Pace et al. Protein Science 4, 2411-2423, (1995) y la ley de Beer Lamberts. La concentración de la muestra se calculó de nuevo usando la masa de proteína total en el conjunto de elución y el volumen de la muestra inyectado. La concentración se confirmó mediante análisis de aminoácidos. La homogeneidad y los pesos moleculares se analizaron usando espectrometría de masas en un Waters Q-ToF (PA, USA).

Resultados de la expresión de proteínas de constructos adicionales

Para determinar los efectos de la secuencia de inicio N-terminal en otros dominios de la Proteína A de *Staphylococcus aureus* y monómero de Zvar del dominio estabilizado con álcali (Zvar₁) y tetrámero (Zvar₄) se construyeron nuevos plásmidos con el extremo N-terminal adicional y se expresaron en fermentaciones con alimentación por lotes usando el plásmido pJ401. Los resultados mostraron que el dominio E no tenía expresión medible sin inserto. Sin embargo, con el inicio de DsbA de 8AA, la expresión de proteína se midió a 2,9 g/L. Tanto el dominio D como el dominio A mostraron una baja expresión que no mejoró añadiendo el inicio N-terminal. Tanto el dominio C como el dominio B (que tienen la misma secuencia de AA en el inicio) mostraron una expresión muy similar, con un claro aumento en la expresión con el inicio de DsbA de 8AA. Sin embargo, el mayor aumento de expresión se observó en el monómero de Zvar- y Zvar2 (Zvar₁ y Zvar₂) con más de 4 veces y 10 veces la expresión de proteína, respectivamente. También, el tetrámero de Zvar (Zvar₄) mostró un aumento significativo en la expresión de proteína con el inicio de DsbA de 8AA, aunque el aumento no fue tan drástico como para el monómero.

Tabla 5. Expresión de proteínas de ejemplo, todos los dominios de prA, Zvar y tetrámero de Zvar, con y sin secuencia de inicio N-terminal.

Plásmido	Dominio	N-terminal	Concentración [g/L]	Método de análisis*
pGE0127	Dominio E	AQGT	0,0	Preparación de proteínas
pGE0293	Dominio E	DsbA de 8AA	2,9	Curva estándar
pGE0128	Dominio D	AQGT	0,6	Preparación de proteínas
pGE0294	Dominio D	DsbA de 8AA	0,1	Preparación de proteínas
pGE0129	Dominio A	AQGT	1,9	Preparación de proteínas
pGE0295	Dominio A	DsbA de 8AA	0,2	Preparación de proteínas
pGE0130	Dominio B	AQGT	2,5	Curva estándar
pGE0296	Dominio B	DsbA de 8AA	6,3	Curva estándar
pGE0131	Dominio C	AQGT	2,5	Curva estándar
pGE0297	Dominio C	8AA DsbA	5,8	Curva estándar
pGE0132	Zvar ₁	AQGT	2,5	Curva estándar
pGE0298	Zvar ₁	8AA DsbA	10,2	Curva estándar
pGE0120	Zvar ₂	AQGT	0,7	Preparación de proteínas
pGE0299	Zvar ₂	DsbA de 8AA	7,3	Curva estándar
pGE0091	Zvar ₄ _OptEc	AQGT	9,6	Curva estándar
pGE0300	Zvar ₄ _OptEc	DsbA de 8AA	15,3	Curva estándar
*O la curva estándar o el método de preparación de proteínas, como se ha descrito anteriormente.				

En el primer experimento con el espaciador N-terminal, se probaron dos péptidos señal diferentes de dos rutas de secreción, y se demostró que el péptido señal OmpA, que utiliza la ruta dependiente de Sec, tenía la mayor expresión y escisión del péptido señal en este experimento. Otros péptidos señal que utilizan esta ruta son, por ejemplo, PhoA (fosfatasa alcalina de *E. coli*), MalE (proteína de unión a maltosa de *E. coli*) y PelB (Pectato liasa B de *Erwinia*

carotovora). Por el contrario, el constructo de referencia (con inicio AQGT) tenía una mayor expresión de proteínas y escisión de péptidos señal usando el péptido señal DsbA, lo que indica que es una combinación de péptido señal y extremo N-terminal. Además, se probaron cuatro extremos N-terminales diferentes y todos funcionaron mejor en comparación con el constructo de referencia usando el péptido señal OmpA. Sorprendentemente, cuando la escisión del péptido señal se realizó óptimamente, también aumentó la expresión de proteínas, lo que indica que la escisión de la peptidasa señal podría ser la enzima limitante de la velocidad en el transporte de proteínas sobre la membrana interna. Sin estar ligado por la teoría, una hipótesis es que la proteína se "atasca" en los poros de la membrana interna y detiene la expresión adicional cuando la secreción no funciona de manera óptima. En cultivos de alta expresión de proteínas, se encontró más del 50% de la proteína fuera de las células en caldo de fermentación lo que indica una fuga periplásmica al espacio extracelular. Cuando se probaron otros dominios de la Proteína A de *Staphylococcus aureus*, se observó que el dominio E no tenía expresión de proteína medible, lo que fue inesperado dado que es el primer dominio después del péptido señal en la Proteína A nativa de *Staphylococcus aureus*. Sin embargo, una diferencia importante es que el primer AA después del sitio de escisión se ha cambiado de AQHDEA a AQGTVDEA para fines de clonación, y podría ser esta sustitución/inserto la que afecta negativamente a la función del péptido señal. Otra diferencia es que el péptido señal se ha sustituido de SpA nativa a OmpA, esto en combinación con el inicio N-terminal podría no ser óptimo. Cuando se añade el extremo N-terminal de DsbA de 8AA al dominio E en combinación con el péptido señal OmpA, el transporte y la expresión de proteínas funcionan adecuadamente. Además, los dominios A y D que están muy estrechamente relacionados en la unidad estructural proximal N-terminal expresan poco y no mejoran con AA adicional, mientras que los dominios B y C, idénticos en esta unidad, expresan bien con el inicio de AQGT, y la expresión aumenta adicionalmente cuando se añaden DsbA de 8AA en el extremo N-terminal. Además, Zvar como monómero y tetrámero aumenta la expresión con AA adicional. Sin embargo, el aumento más drástico se observa en el monómero de Zvar2 donde la expresión de la proteína va de 0,68 g/L a 7,28 g/L. Sin estar limitado por la teoría, una hipótesis de la diferencia entre Zvar y Zvar2 es que Zvar2 tiene dos mutaciones en la primera hélice alfa en la posición Q9 y N11, que se estima que aumenta la estructura de la hélice alfa en el inicio de la proteína, como se calcula con algoritmos tales como GORIV, disponible en línea en https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html

(PRABI, Rhone-Alpes Bioinformatics Center), aumentando así el impedimento estérico de la peptidasa señal para escindir el péptido señal de la proteína naciente.

Ejemplo 4

Variantes del espaciador N-terminal

Los insertos N-terminales se diseñaron cada uno para investigar el efecto de la secuencia de inserto en uno o más de los siguientes parámetros: rendimiento de proteína, escisión de péptido señal y estabilidad alcalina. Los diseños pueden dividirse en cuatro categorías separadas como sigue:

1. Variación de la longitud del inserto existente sin modificaciones.
2. Eliminación del supuesto sitio de truncamiento (glicina-treonina) en condiciones alcalinas, mientras que varía la longitud del inserto restante.
3. Sustitución de la glicina en el sitio conocido de truncamiento para otros diversos restos.
4. Sustitución de la secuencia completa.

El objetivo de la categoría 1 fue principalmente investigar el efecto de la longitud del inserto en la escisión del péptido señal y el rendimiento de proteínas, y no abordaba explícitamente el sitio de truncamiento sospechoso en condiciones alcalinas. Las categorías 2 y 3 tenían principalmente como objetivo abordar el problema de estabilidad alcalina, aunque también proporcionan algunos datos más para los otros dos problemas. La categoría 4 fue una categoría más abierta, proporcionando un conjunto más amplio de secuencias para proporcionar datos en las tres áreas de interés.

Las secuencias de nucleótidos se optimizaron por codones para *E. coli*, pero con alguna degeneración añadida en las secuencias más repetitivas. Los insertos que pertenecen a las categorías 1 y 4 se diseñaron para ser complementarios a los sitios de digestión de FspI en el extremo 5' y KpnI en el extremo 3'. Puesto que el sitio de truncamiento de glicina-treonina era una parte del sitio de restricción KpnI, los insertos que pertenecen a las categorías 2 y 3 se diseñaron para ser complementarios al sitio de restricción AccI adyacente en el extremo 3', manteniendo la complementariedad en 5' con FspI. Las secuencias de aminoácidos de todos los insertos se muestran en la Tabla 6. En la Figura 6 se muestra una ilustración esquemática del constructo general.

Tabla 6. Todos los insertos usados en los experimentos. Los restos resultantes del sitio de restricción “cicatriz” están marcados en *cursiva*.

Categoría	Nombre del constructo	Secuencia de aminoácidos del inserto	SEQ NO	ID
Constructo original sin inserto de DsbA	pGE0002	<i>AQGT</i>	28	
Construcción original con inserto de DsbA de 8 aa	pGE0180	<i>AQYEDGKQYTGT</i>	17	
1	DsbA7	<i>AQYEDGKQYGT</i>	29	
	DsbA6	<i>AQYEDGKQGT</i>	30	
	DsbA5	<i>AQYEDGKGT</i>	31	
	DsbA4	<i>AQYEDGGT</i>	32	
	DsbA12	<i>AQYEDGKQYTTLEKGT</i>	33	
	DsbA16	<i>AQYEDGKQYTTLEKPVAGGT</i>	34	
2	DsbA8_noGT	<i>AQYEDGKQYT</i>	35	
	DsbA6_noGT	<i>AQYEDGKQ</i>	36	
3	DsbA8_ET	<i>AQYEDGKQYTET</i>	37	
	DsbA8_DT	<i>AQYEDGKQYTD</i>	38	
	DsbA8_AT	<i>AQYEDGKQYTAT</i>	39	
	DsbA7_EDT	<i>AQYEDGKQYEDT</i>	40	
4	R8	<i>AQRRRRRRRRGT</i>	41	
	K8	<i>AQKKKKKKKKGT</i>	42	
	H8	<i>AQHSHHHHHHGT</i>	43	
	H6	<i>AQHSHHHHGT</i>	44	
	H4	<i>AQHSHHHGT</i>	45	
	Strep	<i>AQWSHPQFEKGT</i>	46	
	SPA	<i>AQHDEAQQEAGT</i>	47	

5 Los constructos se ensamblaron uniendo diferentes oligonucleótidos hibridados en el plásmido de pGE0002 o pGE0180, digerido con las enzimas de restricción relevantes para eliminar el inserto N-terminal original. Después de la unión, los plásmidos se transformaron en células K12-017 de *E. coli*. Las colonias positivas se seleccionaron usando PCR de colonias y electroforesis.

Se eligieron dos o tres clones de cada constructo y se enviaron a GATC biotech (Colonia, Alemania) para la verificación de secuencia.

Se seleccionó un clon de secuencia verificado para cada constructo para la expresión de proteínas en cultivo en matraz de agitación. La Figura 7 muestra un resumen de los resultados de la cuantificación de los rendimientos de proteínas en los diferentes cultivos en matraz de agitación realizados con IgG Sepharose 6FF. Para cada constructo, la barra izquierda es la concentración de proteína después de la corrección lineal usando como referencia una solución estándar de 5 g/L y 10 g/L. La barra derecha es esta concentración dividida por la OD₆₀₀ nm al final del cultivo. Este valor pretende dar una idea de si una concentración de proteína baja se debe a una velocidad baja real de producción de proteína por célula, o simplemente debido a una densidad de cultivo baja. Obsérvese que las concentraciones se miden después de granular las células y re-suspenderlas en solución salina tamponada con fosfato (Medicago, Uppsala, Suecia), y como tales no son directamente comparables a las concentraciones de muestras tomadas de, por ejemplo, el multif fermentador.

Las soluciones de proteína para cada constructo producidas a partir del cultivo en matraz de agitación se purificaron usando IgG Sepharose 6FF (GE Healthcare, Uppsala, Suecia). Las muestras de proteína purificada se investigaron para la escisión del péptido señal usando análisis de LC/MS (Waters, PA, USA). Una proteína con un péptido señal correctamente escindido solo tiene una masa principal (ejemplo, recuento iónico total (TIC) mostrado en la Figura 8a), mientras que una proteína con un péptido señal incorrectamente escindido produce uno o más picos secundarios (tal y como se muestra en la Figura 8b). Los picos laterales más a la izquierda y a la derecha de los picos principales probablemente se deben a la fuga del citoplasma durante la etapa de tratamiento térmico cuando se extrae la proteína del cultivo celular. Los picos C1 a C5, N1 y N2 son probablemente diferentes versiones parcialmente digeridas de la proteína del citoplasma. El pico pequeño a la derecha del pico principal es proteína con péptido señal no escindido, que sería razonable ver si se había producido fuga del citoplasma.

En muchos casos, los picos eran difíciles de interpretar de manera concluyente debido a la pureza de las muestras y a la insuficiente separación durante la etapa de cromatografía líquida. Sin embargo, se pudo observar que los siguientes constructos tenían una escisión de péptido señal correcta: pGE0180, DsbA7, DsbA6, DsbA8_noGT, DsbA8_DT, DsbA8_ET, DsbA8_AT, DsbA7_EDT, DsbA12, DsbA16, H8, H6 y SPA.

Después de la investigación de escisión de péptido señal, las soluciones de proteína de constructos con escisión del péptido señal correcta se purificaron adicionalmente usando una columna de intercambio aniónico Capto Q ImpRes (GE Healthcare, Uppsala, Suecia). Las muestras purificadas se sometieron a tratamiento con hidróxido de sodio 1 M (NaOH) durante 0, 4 y 24 horas. Se usaron columnas Vivaspin (GE Healthcare, Uppsala, Suecia) para separar la proteína principal de cualquier péptido pequeño que pudiera haber sido truncado. Tanto las muestras peptídicas como las proteínas principales se investigaron usando LC/MS (Waters, PA, USA).

Los péptidos aparecieron en los espectros de masas con +2 de carga, es decir, a un valor de masa sobre carga (m/z) correspondiente a la mitad de su peso molecular. En los casos investigados en la presente memoria, todos los péptidos tenían un peso molecular entre aproximadamente 800 Da y 1600 Da. Como tal, el agrupamiento relevante de picos apareció entre 400 y 800 m/z, tal y como puede verse para pGE0180 a la izquierda en la Figura 9.

Si estaban presentes los péptidos, se usó la función de búsqueda MassLynx en la TIC para extraer los picos de diferentes masas peptídicas como cromatogramas iónicos extraídos (XIC). Los picos de XIC se integraron y las áreas integradas se usaron después para comparar las cantidades de los diferentes péptidos en las muestras.

La Figura 10 muestra un resumen de los resultados del estudio de estabilidad alcalina inicial realizado con columnas Vivaspin sin lavar. A pesar de un poco de ruido de fondo solapado, son visibles estándares claros después de 24 horas de incubación.

La Figura 11 muestra una vista diferente de los mismos datos para los constructos más prometedores (incluyendo pGE0180 como referencia), mientras que la Figura 12 muestra una descomposición de los constructos más destacadamente escindidos. Las secuencias alrededor del sitio de escisión se incluyen para dar una idea de qué restos podrían ser vulnerables. Notablemente, para las áreas de pico bajas (alrededor de 5000 unidades de área), el ruido de fondo tiene una influencia significativa.

El segundo análisis con columnas de centrifugación lavadas con isopropanol condujeron a resultados en su mayor parte similares a la ejecución inicial, aunque las áreas de los picos se redujeron. La notable excepción es DsbA6, que se reveló que tenía alguna escisión. Esta no fue detectable inmediatamente a partir del análisis de péptidos, sino que se hizo evidente cuando se investigó el retenido, en donde DsbA6 mostró un pico truncado claro de manera muy similar a pGE0180, mientras que DsbA8_noGT y DsbA_DT no mostraron dicho pico.

Ejemplo 5

Se seleccionaron constructos interesantes del estudio en matraz con agitación y análisis posterior de la proteína purificada para un cultivo a mayor escala en un multif fermentador (Belach Bioteknik, Skogås, Suecia). Al final de la fermentación, se midieron la OD₆₀₀ nm y concentración de proteína (véase la Figura 13).

Puesto que las diferentes proteínas tienen coeficientes de extinción algo variables, es lógico que también pueden variar en absorbancia a 237 nm, que es donde se estimó la concentración de proteína. Por lo tanto, la relación entre

el coeficiente de extinción de pGE0180 (0,294) y el coeficiente de extinción calculado para cada proteína distinta se usó para estimar un valor más "verdadero".

Los cultivos de alimentación por lotes de los constructos DsbA8_noGT y DsbA8_DT se trataron térmicamente para extraer la proteína periplásmica, y se centrifugaron y diafiltraron para eliminar los restos celulares. El filtrado se acondicionó para eliminar algunos contaminantes mediante precipitación y preparar la solución para purificación con Capto S ImpAct (GE Healthcare, Uppsala, Suecia) usando un gradiente de pH. El eluato se redujo para eliminar los dímeros, y se desaló para eliminar el DTT y obtener el tampón correcto para la purificación por intercambio aniónico, donde se usó un gradiente salino. La solución final purificada se concentró mediante ultrafiltración con membrana para alcanzar la concentración necesaria para la inmovilización a las densidades de ligando deseadas. El producto final se analizó con SEC para determinar la pureza y LC/MS para determinar el peso molecular, con resultados satisfactorios.

Las proteínas purificadas (>90% de pureza) se acoplaron a perlas de agarosa altamente reticuladas, teniendo como objetivo densidades de ligando similares a las de un experimento previo en donde Zvar2 se inmovilizó sobre una matriz base de propiedades similares y se analizó para determinar la capacidad de unión dinámica. Los geles acoplados se empaquetaron en columnas Tricorn 5/100 (GE Healthcare, Uppsala, Suecia). El relleno del lecho de gel se evaluó a través de la asimetría de un pico de acetona. Las columnas empaquetadas se usaron para medir la capacidad de unión dinámica de los geles usando IgG. Se realizó una medición por constructo. Los resultados se resumen en la Tabla 7. La Figura 14 muestra una comparación entre las mediciones realizadas en este caso, y las realizadas en el experimento previo mencionado anteriormente.

Tabla 7. Asimetría y valores de la capacidad de unión dinámica para los diferentes geles acoplados a proteínas.

	Zvar2	DsbA8_noGT	DsbA8_DT
Qb10 (mg IgG/mL gel)	73,7	74,2	71,6
Qb80 (mg IgG/mL gel)	93,6	92,2	90,7

Ejemplo 6

El espaciador N-terminal AQYEDGKQYTGT y una cisteína C-terminal se introdujeron en varios mutantes adicionales monoméricos y diméricos de Zvar2 (como se muestra a continuación), usando el péptido señal OmpA. La estructura de la proteína madura fue, por lo tanto: AQYEDGKQYTGT-IgG-polipéptido de unión a C.

Los plásmidos se transformaron en células K12-017 de *E. coli* competentes. Las células se descongelaron en hielo durante 30 min, se añadieron 50 µl de células competentes a 20 ng de plásmidos. Las células se incubaron en hielo durante 20 min seguido de choque térmico a 42°C durante 1 min, seguido de incubación en hielo 5 min y adición de 400 µl de medio SOC (NEB, MA, USA). La reacción de transformación se incubó a 37°C en una incubadora con agitación rotatoria durante 60 min, se extendieron 200 µl de cada reacción sobre placas de agar Luria que contenían antibióticos selectivos apropiados.

La cepa K12-017 de *E. coli* transformada con los plásmidos recombinantes (Tabla 8) se cultivó a 37°C en 100 ml de medio de caldo Terrific (TB) (que contenía 12 g de triptona, 24 g de extracto de levadura, 5 g de glicerol (85%), 2,31 g de KH₂PO₄, 12,54 g de K₂HPO₄ y 50 mg de sulfato de kanamicina por litro). Los cultivos se indujeron cuando la OD_{600 nm} alcanzó 1, usando IPTG 1 mM (Sigma Aldrich, MO, USA). Los cultivos se incubaron adicionalmente en 30°C 17-20 h. La solución de cultivo se sometió a centrifugación a baja velocidad (a 4000 rpm) durante 20 min en un rotor oscilante para recoger el sedimento celular húmedo. Los sedimentos celulares bacterianos se suspendieron en 10 ml de una solución tampón de fosfato 25 mM (pH 7,4), y las células se lisaron mediante tratamiento térmico a 85°C durante 15 min en un bloque térmico. Seguido de una centrifugación a alta velocidad (a 10000 x g) durante 10 min para separar el sobrenadante. Después, los sobrenadantes se filtraron a través de un filtro de jeringa de 0,2 µm para eliminar cualquier partícula residual antes de aplicarla a una IgG Sepharose 6FF de XK 16-6 (GE Healthcare Bio-Sciences, Uppsala, Suecia). La columna se equilibró con tampón de carga (fosfato 25 mM pH 7, NaCl 250 mM) seguido de la carga de la muestra. Después de cargar la columna se lavó con 5 volúmenes de columna (CV) de tampón de carga y un CV de tampón de lavado bajo en sal (acetato 50 mM pH 6), seguido de elución con ácido acético 50 mM, pH 2,8. La absorbancia a 237 nm se midió en línea usando un explorador ÁKTA 100 (GE Healthcare Bio-Sciences, Uppsala, Suecia) y la integración del pico del pico eluido se realizó en el software del sistema (Unicorn 5.1).

La fracción de unión a IgG se eluyó y se recogió en un conjunto y se observaron los volúmenes y concentraciones de proteína de los conjuntos - véase la Tabla 8. La cantidad de proteína de unión a IgG expresada puede ser en varios casos mayor que la cantidad recuperada, debido a la pérdida de proteína secretada en el medio de cultivo y/o a la sobrecarga de la columna de IgG.

Tabla 8. Expresión de proteínas de unión a IgG con el espaciador N-terminal AQYEDGKQYTGT y con una cisteína C-terminal.

Proteína de unión a IgG	SEQ ID NO	Conc. de proteína mg/mL	Volumen total, mL	Proteína recuperada, mg
Zvar2(A29G) ₁	48	1,4	3	4,2
Zvar2 (A29S) ₁	49	0,8	3	2,4
Zvar2 (A29Y) ₁	50	2,0	4	8,0
Zvar2 (A29Q) ₁	51	2,3	3	6,9
Zvar2 (A29T) ₁	52	2,1	3	6,3
Zvar2 (A29N) ₁	53	2,0	4	8,0
Zvar2 (A29F) ₁	54	2,0	4	8,0
Zvar2 (A29L) ₁	55	2,3	3	6,9
Zvar2 (A29W) ₁	56	2,1	3	6,3
Zvar2 (A29I) ₁	57	1,6	3	4,8
Zvar2 (A29M) ₁	58	2,0	4	8,0
Zvar2 (A29V) ₁	59	2,1	3	6,3
Zvar2 (A29D) ₁	60	2,3	3	6,9
Zvar2 (A29E) ₁	61	2,2	4	8,8
Zvar2 (A29H) ₁	62	1,2	4	4,8
Zvar2 (A29R) ₁	63	1,9	4	7,6
Zvar2 (A29K) ₁	64	2,0	4	8,0
Zvar2 (Δ 235,236,237) ₂	65	2,2	4	8,8
Zvar2 (Δ 233,234,235) ₂	66	2,9	4	11,6
Zvar2 (Δ 5-1) ₂	67	2,1	5	10,5
Zvar2 (Δ 5-2) ₂	68	2,8	6	16,8
Zvar2 (D3C-term) ₂	69	3,3	5	16,5
Zvar2 (D3N-term) ₂	70	3,5	2,7	9,4
Zvar2 (D8N-term) ₂	71	2,6	2,1	5,46
Zvar2 (Enlazador + 8) ₂	72	3,0	3,4	10,2

ES 2 986 499 T3

Zvar2 (Enlazador + 4) ₂	73	3,0	2,7	8,1
Zvar2 (Δ Q9) ₁	74	1,4	1,7	2,4
Zvar2 (Δ Q40) ₁	75	2,2	3,7	8,1
Zvar2 (Δ A42) ₁	76	1,7	1,7	2,9
Zvar2 (Δ N43) ₁	77	1,6	2	3,2
Zvar2 (Δ L44) ₁	78	1,3	1,2	1,6
Zvar2 (E11N,A12F) ₁	79	1,3	1,1	1,4
Zvar2 (E11N,A12Y) ₁	80	1,2	0,75	0,9
Zvar2 (E11N,A12K) ₁	81	1,4	0,72	1,0
Zvar2 (E11N,A12R) ₁	82	1,8	1,49	2,7
Zvar2 (L22F) ₁	83	2,9	2,77	8,0
Zvar2 (A43N,I44F) ₁	84	1,9	4	7,6
Zvar2 (A43N,I44Y) ₁	85	2,4	3	7,2
Zvar2 (A43N,I44W) ₁	86	1,1	3	3,3
Zvar2 (A43N,I44R) ₁	87	2,4	3,9	9,4
Zvar2 (A43N,I44K) ₁	88	2,4	4	9,6
Zvar2 (D53F) ₁	89	2,7	3	8,1
Zvar2 (D53Y) ₁	90	2,8	3	8,4
Zvar2 (D53W) ₁	91	1,3	2,9	3,8
Zvar2 (D53K) ₁	92	2,2	4	8,8
Zvar2 (D53R) ₁	93	2,2	3	6,6

Monómero de Zvar2(A29G) (SEQ ID NO 48)

VDAKFDKEAQEAFYEILHLPNLTEEQRNQFIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29S) (SEQ ID NO 49)

VDAKFDKEAQEAFYEILHLPNLTEEQRNSFIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29Y) (SEQ ID NO 50)

VDAKFDKEAQEAFYEILHLPNLTEEQRNYFIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29Q) (SEQ ID NO 51)

VDAKFDKEAQEAFYEILHLPNLTEEQRNQFIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29T) (SEQ ID NO 52)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29N) (SEQ ID NO 53)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

5 Monómero de Zvar2 (A29F) (SEQ ID NO 54)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29L) (SEQ ID NO 55)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29W) (SEQ ID NO 56)

10 VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29I) (SEQ ID NO 57)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29M) (SEQ ID NO 58)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

15 Monómero de Zvar2 (A29V) (SEQ ID NO 59)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29D) (SEQ ID NO 60)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29E) (SEQ ID NO 61)

20 VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29H) (SEQ ID NO 62)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Monómero de Zvar2 (A29R) (SEQ ID NO 63)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

25 Monómero de Zvar2 (A29K) (SEQ ID NO 64)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFTIQSLKDDPSVSKAILAEAKKLNDQAQPK

Dímero de Zvar2 (Δ 235,236,237) (SEQ ID NO 65)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQPKV
FDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQPK

Dímero de Zvar2 (Δ 233, 234, 235) (SEQ ID No 66)

30 VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPAK
FDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQPK

Dímero de Zvar2 (Δ 5-1) (SEQ ID No 67)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAKFD
KEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQPK

Dímero de Zvar2 (Δ 5-2) (SEQ ID No 68)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQPKV
DAAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQPK

ES 2 986 499 T3

Dímero de Zvar2 (D3C-term) (SEQ ID NO 69)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQVDAK
FDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

Dímero de Zvar2 (D3N-term) (SEQ ID NO 70)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPKK
FDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

5 Dímero de Zvar2 (D8N-term) (SEQ ID NO 71)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPKA
QEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

Dímero de Zvar2 (Enlazador + 8) (SEQ ID NO 72)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPKY
EDGKQYTVDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

Dímero de Zvar2 (Enlazador + 4) (SEQ ID NO 73)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPKY
EDGVDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

Monómero de Zvar2 (Δ Q9) (SEQ ID NO 74)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

Monómero de Zvar2 (Δ Q40) (SEQ ID NO 75)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSSKAILAEAKKLNDQAQAPK

15 Monómero de Zvar2 (Δ A42) (SEQ ID NO 76)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSAILAEAKKLNDQAQAPK

Monómero de Zvar2 (Δ N43) (SEQ ID NO 77)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKILAEAKKLNDQAQAPK

Monómero de Zvar2 (Δ L44) (SEQ ID NO 78)

20 VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKALAEAKKLNDQAQAPK

Monómero de Zvar2 (E11N,A12F) (SEQ ID NO 79)

VDAKFDKEAQNFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

Monómero de Zvar2 (E11N,A12Y) (SEQ ID NO 80)

VDAKFDKEAQNYFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

25 Monómero de Zvar2 (E11N,A12K) (SEQ ID NO 81)

VDAKFDKEAQNKFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

Monómero de Zvar2 (E11N,A12R) (SEQ ID NO 82)

VDAKFDKEAQNRFYEILHLPNLTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

Monómero de Zvar2 (L22F) (SEQ ID NO 83)

30 VDAKFDKEAQEAFYEILHLPNFTTEEQRNAFIQSLKDDPSVSKAILAEAKKLNDQAQAPK

Monómero de Zvar2 (A43N,I44F) (SEQ ID NO 84)

VDAKFDKEAQEAFYEILHLPNLTEEQRNAFIQSLKDDPSVSKNFLAEAKKLNDQAQAPK

Monómero de Zvar2 (A43N,I44Y) (SEQ ID NO 85)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFIQSLKDDPSVSKNYLAEAKKLNDQAQPK

Monómero de Zvar2 (A43N,I44W) (SEQ ID NO 86)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFIQSLKDDPSVSKNWLAEAKKLNDQAQPK

5 Monómero de Zvar2 (A43N,I44R) (SEQ ID NO 87)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFIQSLKDDPSVSKNRLAEAKKLNDQAQPK

Monómero de Zvar2 (A43N,I44K) (SEQ ID NO 88)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFIQSLKDDPSVSKNKLAEAKKLNDQAQPK

Monómero de Zvar2 (D53F) (SEQ ID NO 89)

10 VDAKFDKEAQEAFYEILHLPNLTEEQRNFIQSLKDDPSVSKAILAEAKKLNFAQAPK

Monómero de Zvar2 (D53Y) (SEQ ID NO 90)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFIQSLKDDPSVSKAILAEAKKLNYAQAPK

Monómero de Zvar2 (D53W) (SEQ ID NO 91)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFIQSLKDDPSVSKAILAEAKKLNWAQAPK

15 Monómero de Zvar2 (D53K) (SEQ ID NO 92)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFIQSLKDDPSVSKAILAEAKKLNKAQAPK

Monómero de Zvar2 (D53R) (SEQ ID NO 93)

VDAKFDKEAQEAFYEILHLPNLTEEQRNFIQSLKDDPSVSKAILAEAKKLNRAQAPK

REIVINDICACIONES

- 5 1. Una proteína recombinante que comprende un polipéptido de unión a Fc y, unido al extremo N-terminal de dicho polipéptido de unión a Fc, un espaciador N-terminal que tiene una longitud tal que el número de restos de aminoácidos entre un sitio de escisión del péptido señal y una hélice alfa proximal del extremo N-terminal de dicho polipéptido de unión a Fc es de 14-24, y en donde los dos aminoácidos N-terminales de dicho espaciador N-terminal son alanina seguido de glutamina.
2. La proteína recombinante de la reivindicación 1, en donde dicho polipéptido de unión a Fc comprende al menos tres hélices alfa.
- 10 3. La proteína recombinante de una cualquiera de las reivindicaciones anteriores, en donde como máximo dos restos de aminoácidos en dicho espaciador N-terminal se seleccionan del grupo que consiste en lisina y arginina, y/o en donde dicho espaciador N-terminal consiste en 8-24 restos de aminoácidos.
4. La proteína recombinante de una cualquiera de las reivindicaciones anteriores, en donde dicho espaciador N-terminal tiene al menos un 80% de identidad de secuencia con, o está definido por, una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO. 16-18, 29-30, 33-40, 43-45 y 47.
- 15 5. La proteína recombinante de una cualquiera de las reivindicaciones anteriores, en donde dicho polipéptido de unión a Fc comprende uno o más dominios de unión a Fc derivados de la Proteína A de *Staphylococcus aureus*.
6. La proteína recombinante de la reivindicación 5, en donde dichos dominios de unión a Fc son dominios de unión a Fc estabilizados con álcali, que tienen opcionalmente al menos un 90% de identidad de secuencia con una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO. 1-11 y 48-64.
- 20 7. La proteína recombinante de una cualquiera de las reivindicaciones 5-6, en donde dicho polipéptido de unión a Fc comprende un multímero de al menos cuatro dominios de unión a Fc.
8. La proteína recombinante de una cualquiera de las reivindicaciones anteriores, que comprende además en, o adyacente a, el extremo C-terminal un resto de acoplamiento, que comprende opcionalmente un resto de cisteína y/o una pluralidad de restos de lisina.
- 25 9. La proteína recombinante de una cualquiera de las reivindicaciones anteriores, en donde dicho espaciador N-terminal es estable frente a los álcalis debido a la ausencia de asparaginas, y/o consiste en restos de aminoácidos seleccionados del grupo que consiste en alanina, ácido aspártico, glutamina, ácido glutámico, glicina, histidina, lisina, fenilalanina, serina, treonina, triptófano, tirosina y valina.
- 30 10. Una molécula de ácido nucleico que codifica la proteína recombinante de una cualquiera de las reivindicaciones anteriores, comprendiendo dicha molécula de ácido nucleico los siguientes elementos en la dirección 5' a 3', estando dichos elementos unidos operativamente:
 - a) una secuencia de DNA promotora inducible o constitutiva;
 - b) una secuencia de DNA que codifica un péptido señal;
 - c) una secuencia de DNA que codifica un espaciador N-terminal; y
 - 35 d) una secuencia de DNA que codifica un polipéptido de unión a Fc.
11. Un vector de clonación que expresa y secreta la proteína recombinante de una cualquiera de las reivindicaciones anteriores en el periplasma bacteriano de una célula gram negativa, comprendiendo dicho vector de clonación la molécula de ácido nucleico de la reivindicación 10.
- 40 12. Una bacteria gram negativa que comprende el vector de clonación de la reivindicación 11, bacteria que se identifica opcionalmente como *Escherichia coli*, tal como K12 de *Escherichia coli* o K12-017 de *Escherichia coli*.
13. Un método para expresar y secretar la proteína recombinante de una cualquiera de las reivindicaciones anteriores en una bacteria gram-negativa, comprendiendo dicho método las etapas de:
 - i) proporcionar la bacteria gram negativa de la reivindicación 12; y
 - ii) cultivar dicha bacteria gram negativa.
- 45 14. Una matriz de separación que comprende la proteína recombinante de una cualquiera de las reivindicaciones 1-9, unida covalentemente a un soporte, que comprende opcionalmente partículas porosas, tales como partículas de polisacáridos reticuladas.
15. La matriz de separación de la reivindicación 14, en donde dicha proteína recombinante está unida covalentemente a dicho soporte a través de un enlace tioéter o uno o más enlaces amida.

16. La matriz de separación de una cualquiera de las reivindicaciones 14-15, en donde dicha matriz de separación es estable frente a álcalis, tal como en donde la capacidad de la IgG de la matriz después de 24 h de tiempo de incubación en NaOH 0,5 M a 22 +/- 2°C es al menos el 80 % de la capacidad de la IgG antes de la incubación.

17. Un método para separar una inmunoglobulina, que comprende las etapas de:

- 5 i) proporcionar la matriz de separación de una cualquiera de las reivindicaciones 14-16;
- ii) poner en contacto dicha matriz de separación con una muestra líquida que contiene una inmunoglobulina, para unir dicha inmunoglobulina;
- iii) opcionalmente lavar dicha matriz de separación con un líquido de lavado;
- iv) poner en contacto dicha matriz de separación con un líquido de elución, para eluir dicha inmunoglobulina.

- 10 18. El método de la reivindicación 17, en donde en la etapa a) dicha proteína recombinante comprende uno o más dominios de unión a Fc derivados de la Proteína A de *Staphylococcus aureus* y en donde dicho método comprende además, después de la etapa iv), una etapa v) de limpieza de dicha matriz de separación con un líquido de limpieza, en donde dicho líquido de limpieza comprende opcionalmente al menos NaOH o KOH 0,1 M, tal como al menos NaOH o KOH 0,5 M, o NaOH o KOH 0,5-2,5 M.

ES 2 986 499 T3

EspI
KpnI
AccI

~~~~~
~~~~~-----

5' GCCCAGGCTGCGCAGTATGAAGATGGCAAACAGTACACCGGTACCGTAGACGCGAAG 3'

A Q A A Q Y E D G K Q Y T G T V D A K

3' CGGGTCCGACGCGTCATACTTCTACCGTTTGTCATGTGGCCATGGCATCTGCGCTTC 5'

Fig. 1.

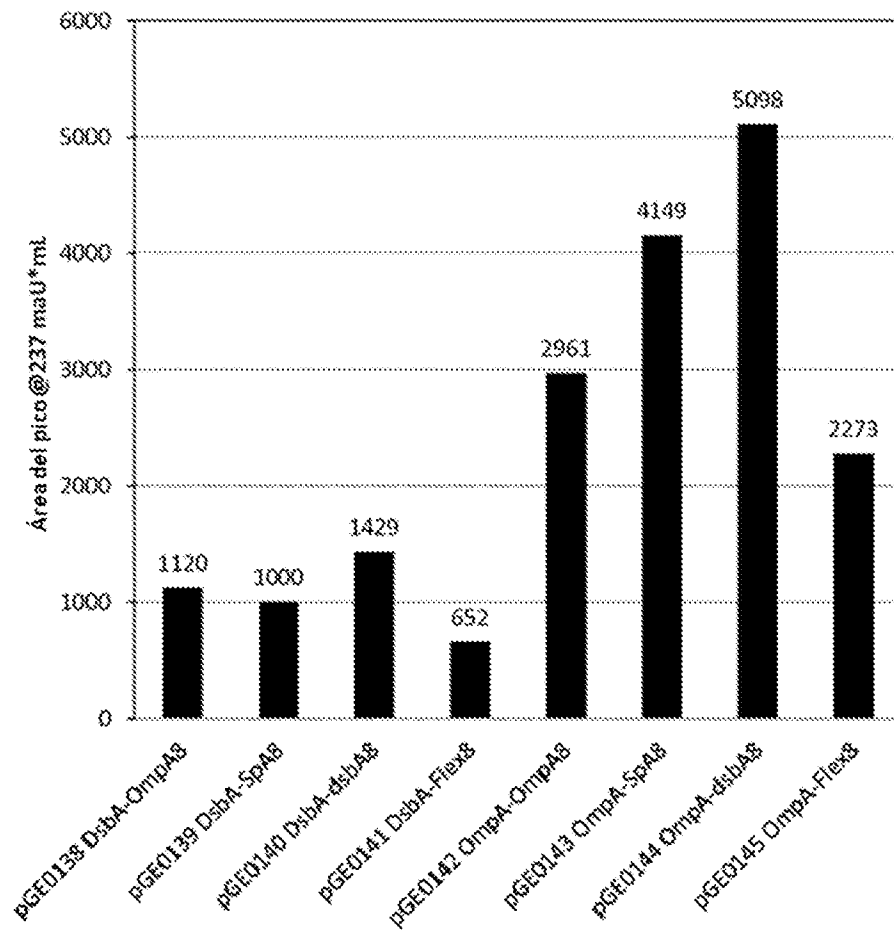


Fig. 2.

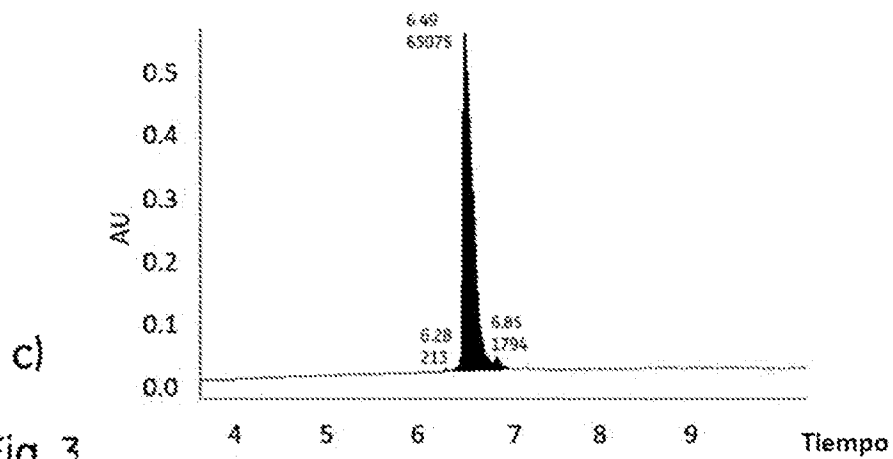
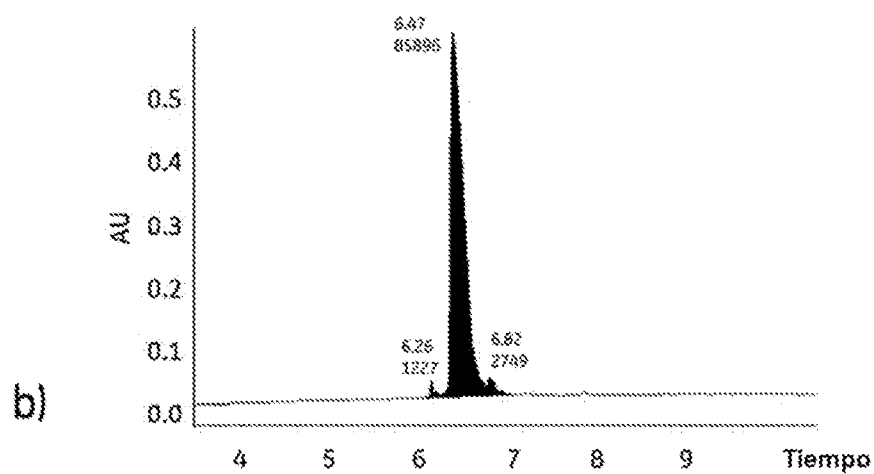
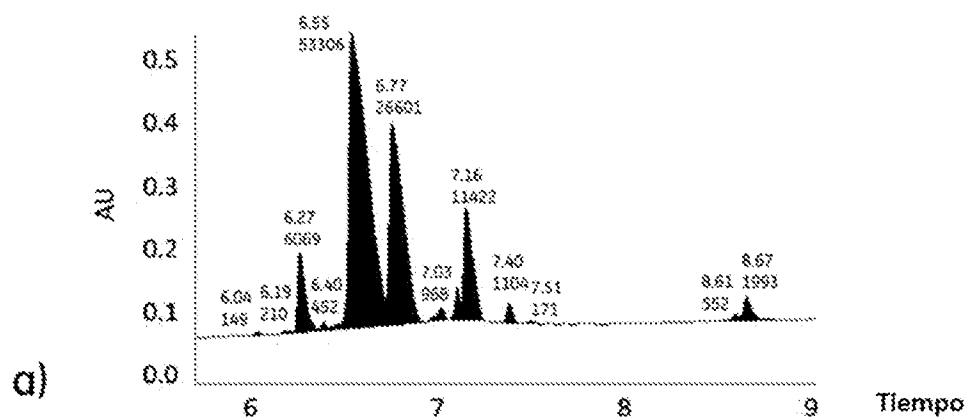


Fig. 3

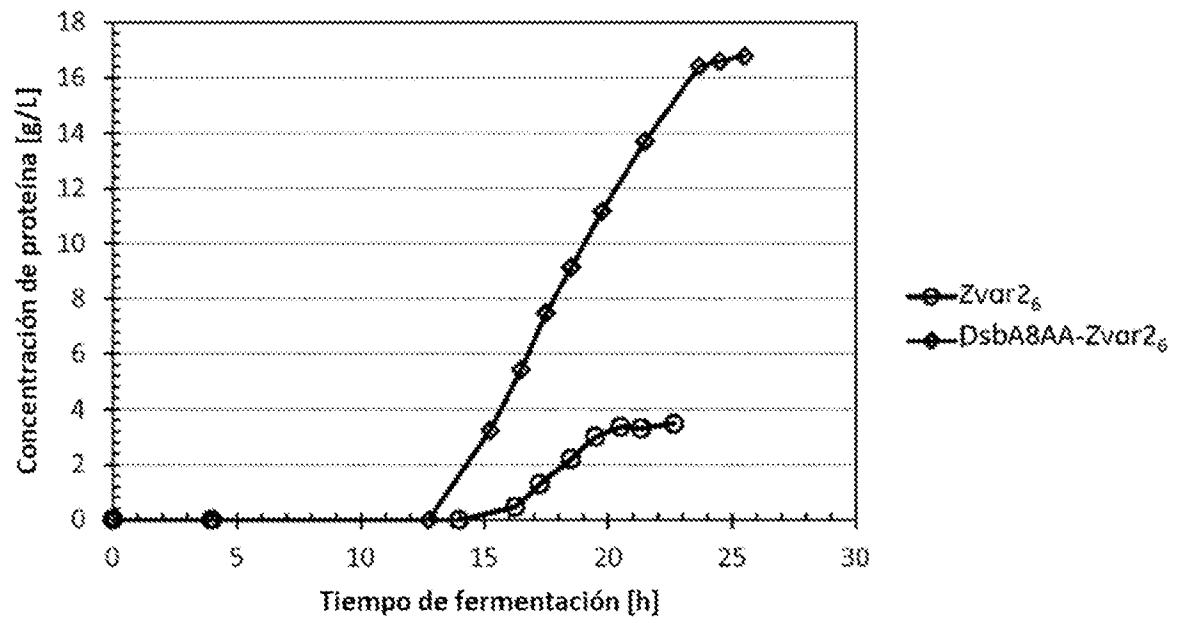
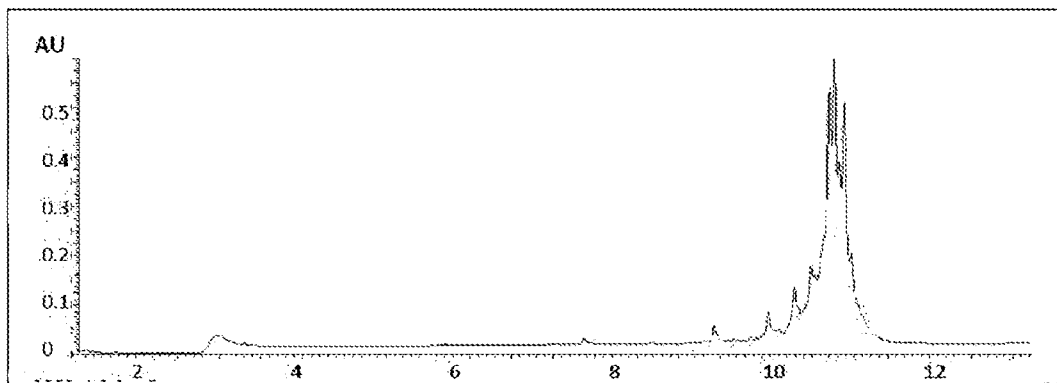
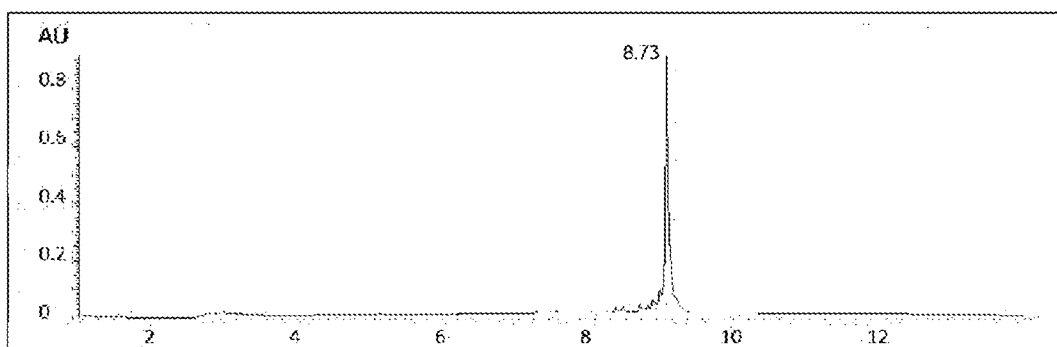


Fig. 4.



a)



b)

Fig. 5.

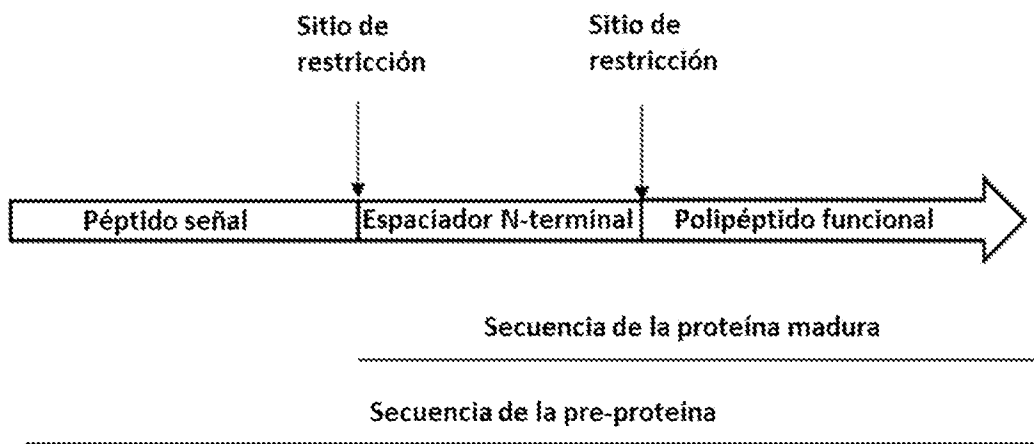


Fig. 6.

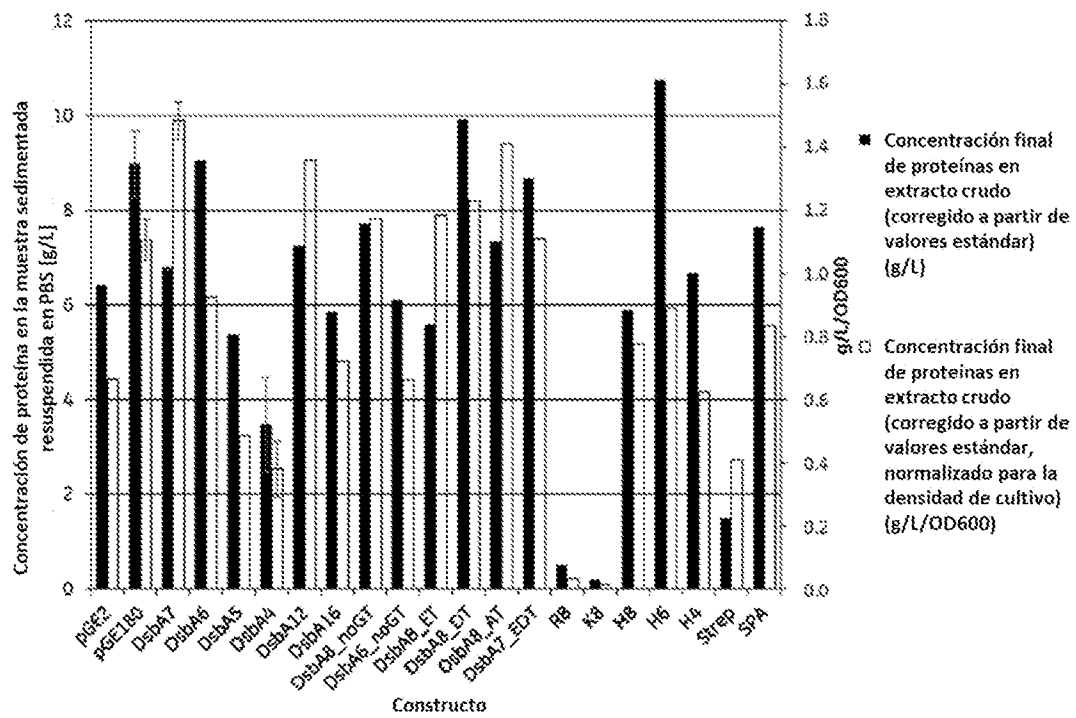


Fig. 7.

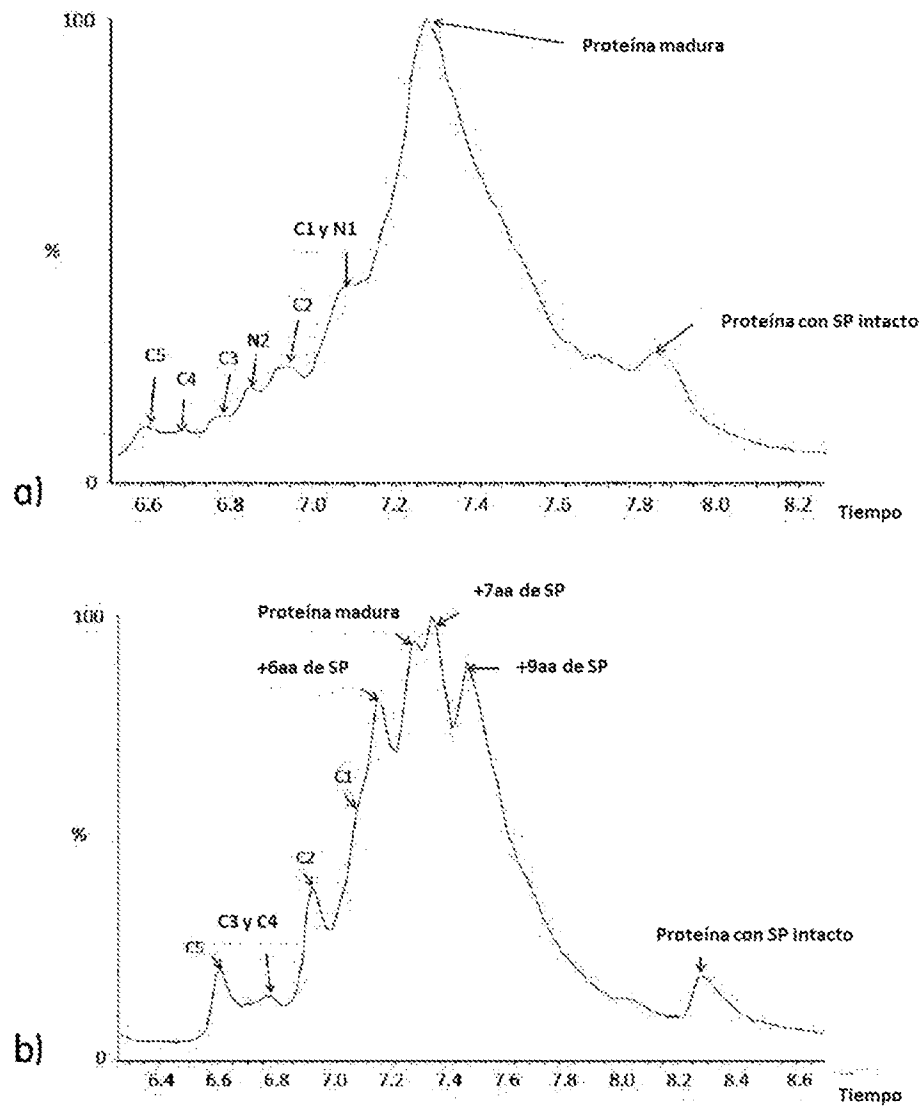


Fig. 8.

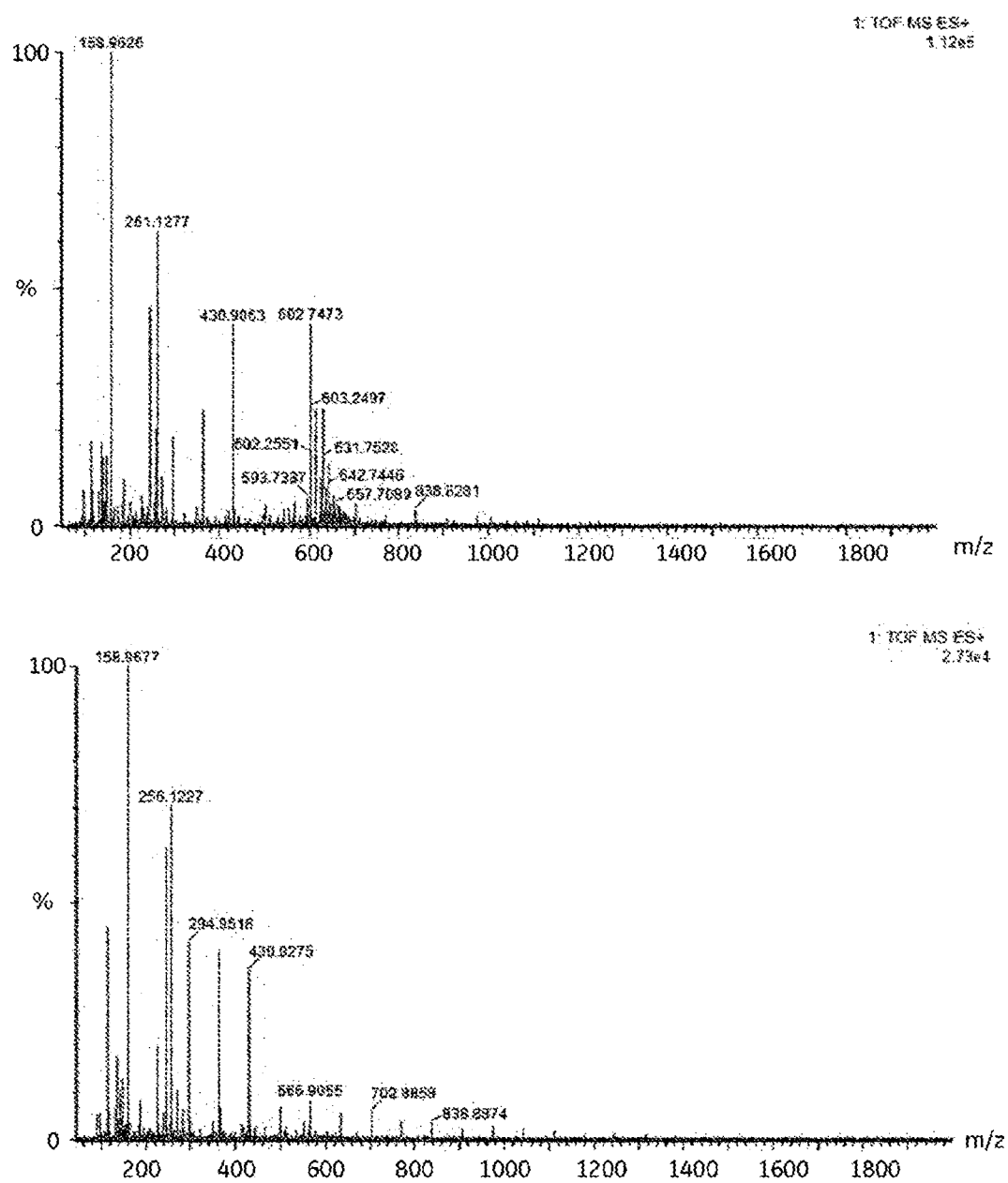


Fig. 9.

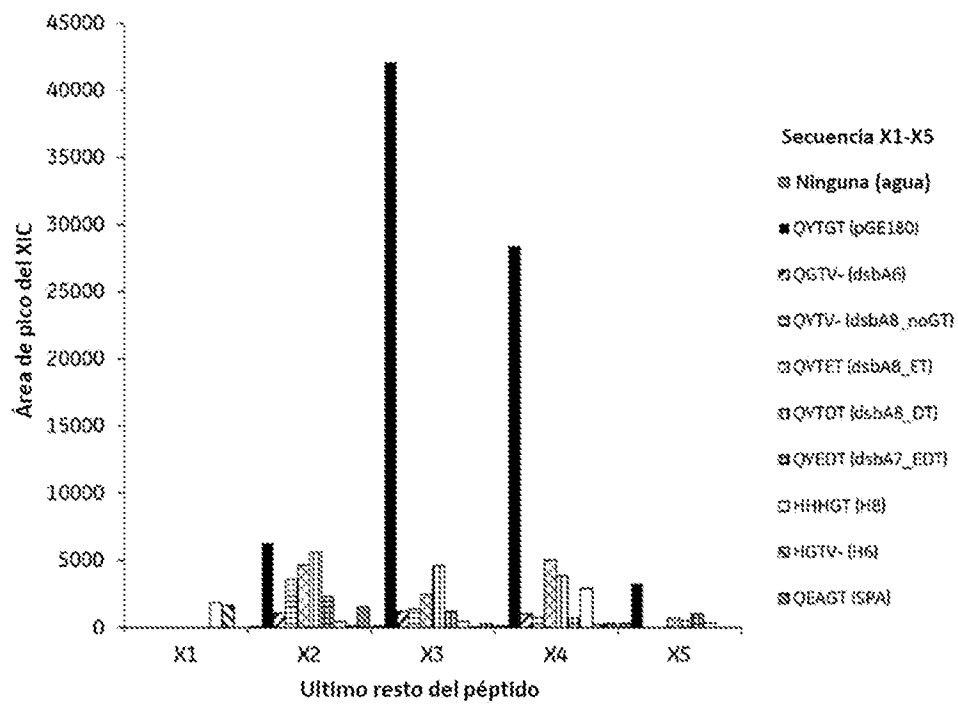


Fig. 11.

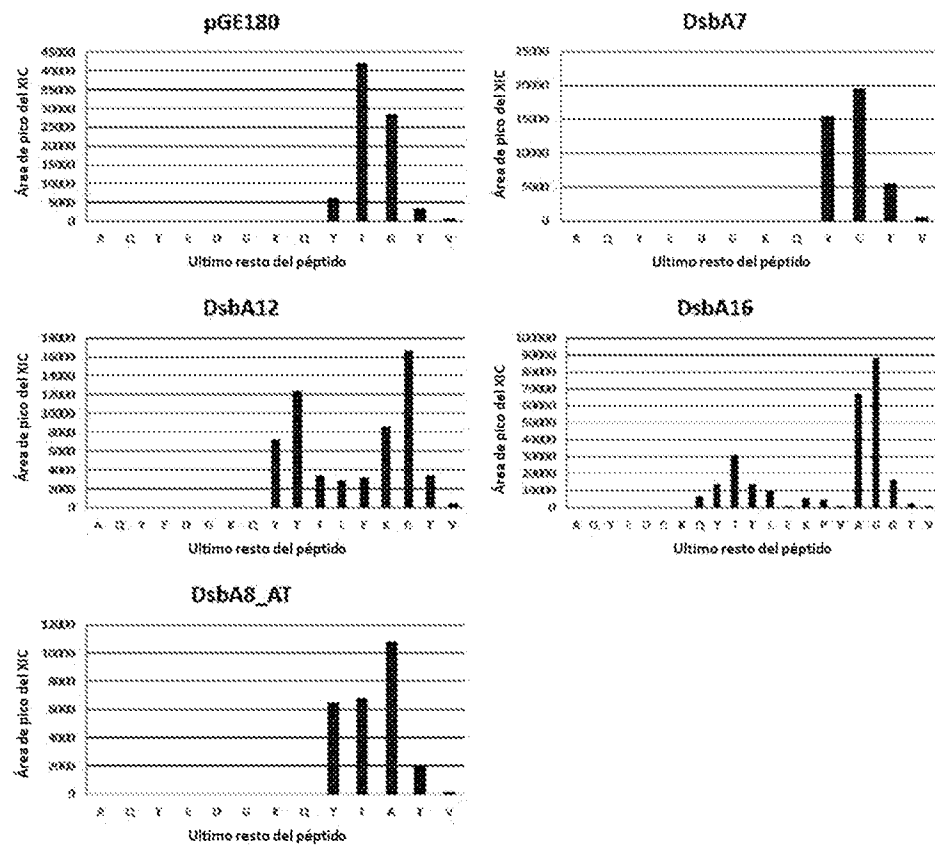


Fig. 12.

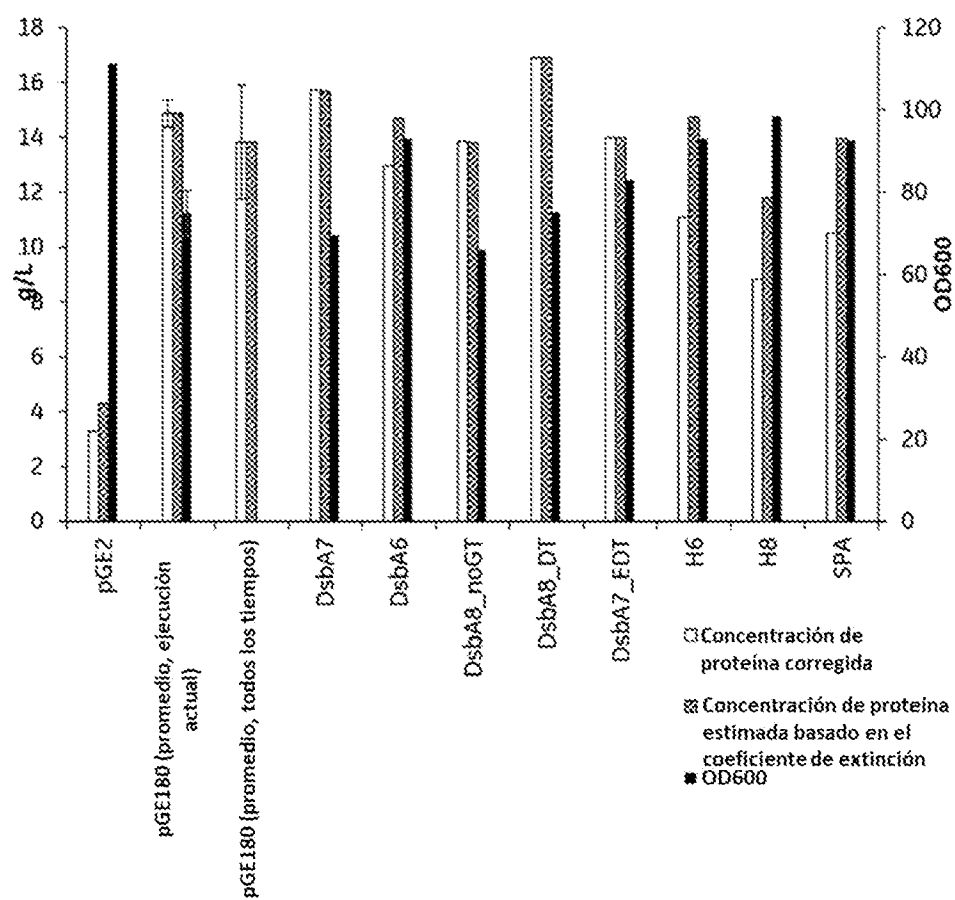


Fig. 13.

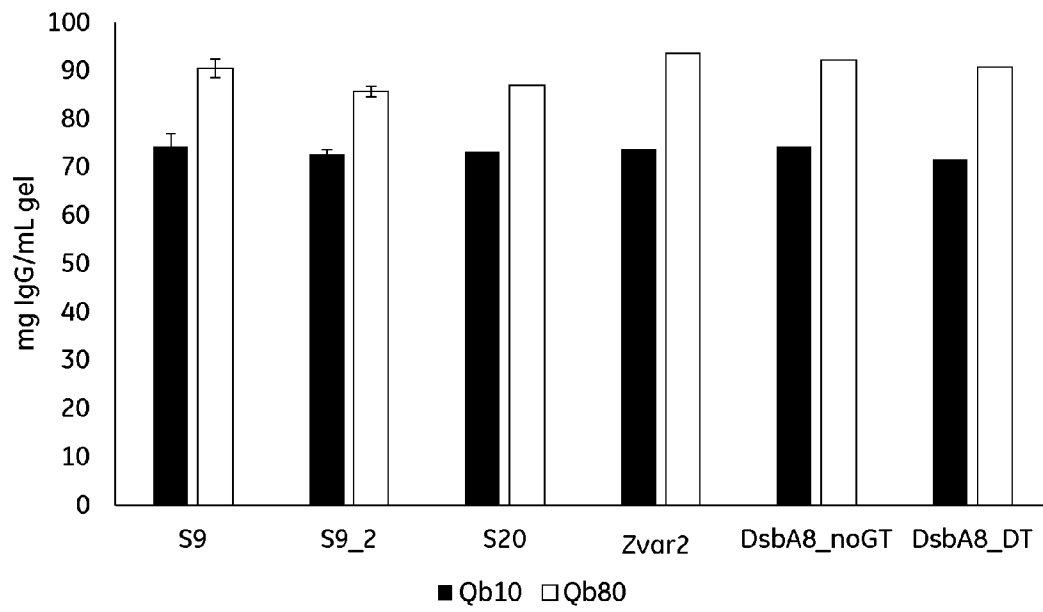


Fig. 14.

Alineación de los dominios de unión a Fc

| | | | | | | | | | |
|------|-------|-----------|------------|-----------|------------|------------|------------|----|-----------------|
| E | |AQ | NATYCVLMNF | NLNDQRMGF | IOSLEDDPSQ | SANVLGEAQK | INDGQAPK | 51 | (SEQ ID NO: 1) |
| D | ADA | QNTFNDDQ | SATYELIMNF | NLNEQRMGF | IOSLEDDPSQ | STNVLSQAPK | INDGQAPK | 61 | (SEQ ID NO: 2) |
| A | --A | DNN-FNNEQ | NATYELIMNF | NLNEQRMGF | IOSLEDDPSQ | SANLLAEAKK | INDGQAPK | 58 | (SEQ ID NO: 3) |
| B | --- | ADNKNNEEQ | NATYELILNP | NLNEQRMGF | IOSLEDDPSQ | SANLLAEAKK | INDGQAPK | 58 | (SEQ ID NO: 4) |
| C | --- | ADNKNNEEQ | NATYELILNP | NLNEQRMGF | IOSLEDDPSV | SEKLLAEAKK | INDGQAPK | 58 | (SEQ ID NO: 5) |
| E | --- | YDNKNNEEQ | NATYELILNP | NLNEQRMGF | IOSLEDDPSQ | SANLLAEAKK | INDGQAPK | 58 | (SEQ ID NO: 6) |
| Zoar | --- | YDNKNNEEQ | NATYELILNP | NLNEQRMGF | IOSLEDDPSQ | SANLLAEAKK | INDGQAPK | 58 | (SEQ ID NO: 7) |
| | --- |QQ | NATYELILNP | NLNEQRMGF | IOSLEDDPSV | SEKLLAEAKK | INDGQ--- | 47 | (SEQ ID NO: 8) |
| | |QQ | NATYELILNP | NLNEQRMGF | IOSLEDDPSQ | SANLLAEAKK | INDGQ----- | 47 | (SEQ ID NO: 10) |
| Pos | 1 | 10 | 20 | 30 | 40 | 50 | 58 | | |

Fig. 15.