

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 609 066**

51 Int. Cl.:

C07K 14/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.10.2010 PCT/EP2010/065047**

87 Fecha y número de publicación internacional: **14.04.2011 WO11042516**

96 Fecha de presentación y número de la solicitud europea: **07.10.2010 E 10762693 (9)**

97 Fecha y número de publicación de la concesión europea: **05.10.2016 EP 2486047**

54 Título: **Sistema de expresión**

30 Prioridad:

08.10.2009 GB 0917647

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

18.04.2017

73 Titular/es:

**GLAXOSMITHKLINE BIOLOGICALS S.A. (100.0%)
Rue de l'Institut, 89
1330 Rixensart, BE**

72 Inventor/es:

**BLAIS, NORMAND;
DEHOTTAY, PHILIPPE MARC HELENE;
DEWERCHIN, MARIANNE;
GOFFIN, PHILIPPE y
MARTIN, DENIS**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 609 066 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

Sistema de expresión

5 La presente invención se refiere al campo de la expresión de toxinas bacterianas, en particular de toxinas diftéricas (incluyendo formas mutantes de toxina diftérica, tales como CRM197) e incluye procedimientos adecuados para la expresión y elaboración de cultivos en grandes cantidades de tales toxinas. Polinucleótidos y polipéptidos que se pueden usar o producir durante el procedimiento de la invención también se divulgan.

La toxina diftérica es una exotoxina proteica producida por la bacteria *Corynebacterium diphtheria*. Se produce como un polipéptido individual que contiene una secuencia señal (la secuencia señal tox) que se retira por la bacteria en secreción de la proteína.

10 La toxina de la difteria se somete a ajuste fácilmente para formar dos subunidades ligadas por un puente disulfuro, Fragmento A y Fragmento B, como un resultado de escisión en residuo 190, 192 o 193 (Moskaug y col. Biol. Chem. 264: 15709-15713, 1989). El Fragmento A es la parte catalíticamente activa y es una ADP-ribosiltransferasa dependiente de NAD que específicamente tiene como objetivo un factor de síntesis de proteínas EF-2, inactivando de este modo EF-2 y cancelando la síntesis de proteínas en una célula.

15 La inmunidad a una toxina bacteriana tal como toxina de difteria se puede adquirir de forma natural durante el curso de la infección, o artificialmente por inyección de una forma detoxificada de la toxina (toxóide) (Germanier, Bacterial Vaccines, Academic Press, Orlando, FL., 1984). Los toxóides se han fabricado tradicionalmente por modificación química de toxinas nativas (Lingood y col. Brit. J. Exp. Path. 44; 177, 1963), volviéndolas no tóxicas reteniendo mientras una antigenicidad que protege al animal vacunado contra el reto subsiguiente por la toxina natural.
20 Alternativamente, se han descrito varias toxinas mutadas de difteria que tienen toxicidad reducida (documentos US4709017, US4950740).

CRM197 es una forma no tóxica de la toxina diftérica pero es inmunológicamente indistinguible de la toxina diftérica. CRM197 se produce por *C. diphtheriae* infectada por el β 197tox- en fase no toxigénica creado por mutagénesis de nitrosoguanidina del carinéfago toxigénico b (Uchida y col. Nature New Biology (1971) 233; 8-11). La proteína
25 CRM197 tiene el mismo peso molecular que la toxina diftérica pero difiere de ella por un cambio de base única en el gen estructural. Esto conduce a un cambio de glicina por glutamina de aminoácidos en posición 52 que hace al fragmento A incapaz de unir NAD y por lo tanto no tóxico (Pappenheimer 1977, Ann Rev, Biochem. 46; 69-94, Rappuoli Applied and Environmental Microbiology sept. 1983 p. 560-564).

30 Difteria toxóide y una forma mutante con toxicidad reducida, CRM197, son componentes en muchas vacunas que proporcionan inmunidad contra *Corynebacterium diphtheriae*. Se conocen varias vacunas de combinación que pueden evitar *Bordetella pertussis*, *Clostridium tetani*, *Corynebacterium diphtheriae* y opcionalmente virus de la Hepatitis B y/o *Haemophilus influenzae* tipo b (véanse, por ejemplo, documentos WO 93/24148 y WO 97/00697, WO 02/055105).

35 La toxina diftérica y formas mutantes que incluyen CRM197 se han usado también en vacunas según sea seguro y en vehículos dependientes de células T efectivos para sacáridos. CRM197 se usa actualmente en la vacuna de conjugados de CRM197 con oligosacáridos de *Haemophilus influenzae* tipo b (HibTitre®; Lederle Praxis Biologicals, Rochester, N.Y.).

40 Se conocen bien en la técnica procedimientos de preparación de toxóide de difteria (DT). Por ejemplo, DT se puede producir por purificación de la toxina a partir de un cultivo de *Corynebacterium diphtheriae* seguida por detoxificación química, o puede hacerse por purificación de un recombinante, o se puede detoxificar genéticamente de forma análoga a la toxina (por ejemplo, CRM197, u otros mutantes según se describen en los documentos US 4,709,017, US 5,843,711, US 5,601,827 y US 5,917,017).

45 La producción de cantidades significativas de toxinas diftéricas tales como CRM197 para usar en vacunas se ha dificultado debido a la baja abundancia de proteínas. Este problema se ha tratado previamente expresando CRM197 en *E. coli* (Bishai y col. J Bacteriol. 169:5140-5151), Bishai y col. describe la expresión de una proteína de condensación recombinante que contiene toxina diftérica (que incluye la secuencia señal tox) esto condujo a la producción de proteína degradada.

50 La clonación de los fragmentos de difteria que contienen la secuencia señal tox y la expresión de estas secuencias en *Escherichia coli* implica ciertas dificultades. La proteína expresada se secreta en el espacio periplásmico y esta secreción está asociada con la viabilidad reducida de las células huésped (O'Keefe y col. Proc. Natl. Acad. Sci. 86: 343-346) y con la proteólisis incrementada de la proteína recombinante (Bishai y col. J Bacteriol. 169: 5140-5151). Por estas razones se ha sugerido la retirada de la secuencia señal tox de tal forma que la expresión ya no sea más periplásmica, esto puede incrementar expresión de toxóides de difteria (Bishai y col.).

55 De acuerdo con ello, la presente solicitud proporciona un procedimiento mejorado para fabricar una toxina bacteriana por expresión periplásmica comprendiendo las etapas de a) cultivar un cultivo de la célula huésped bacteriana que contiene un vector de expresión en el que las secuencias señal particulares están unidas a la

secuencia de una toxina bacteriana y b) inducir expresión del polipéptido que contiene secuencias señal particulares unidas a una toxina bacteriana de tal forma que una toxina bacteriana se expresa periplásmicamente. La presente solicitud también proporciona polinucleótidos que se usan en el procedimiento de la invención.

5 La producción de toxinas bacterianas en el periplasma puede tener una o más ventajas por encima de la producción citoplásmica.

(1) La proteína se produce en su forma madura después de la escisión del péptido señal, y/o;

(2) El periplasma de *E. coli* es un ambiente oxidante que permite la formación de enlaces disulfuro, esto puede ayudar a producir proteínas solubles, plegadas correctamente, y/o;

10 (3) El periplasma de *E. coli* contiene menos proteasas que el citoplasma, esto puede ayudar a evitar escisión proteolítica de la proteína expresada, y/o;

(4) El periplasma también contiene menos proteínas, esto permite que se obtenga proteína recombinante más pura.

Sumario de la invención

15 En un primer aspecto de la invención se proporciona un procedimiento de producción industrial de una toxina bacteriana expresada periplásmicamente por:

A. en una fase de alimentación discontinua, cultivar un cultivo de una célula huésped gram-negativa que comprende un polinucleótido que comprende una parte de toxina 3' que codifica una toxina bacteriana y una parte de señal 5' que codifica un polipéptido capaz de dirigir el transporte de dicha toxina bacteriana al periplasma de la célula huésped, en el que el pH es de 6,0-7,0; y

20 B. inducir expresión de dicha toxina bacteriana a un pH que es y se mantiene a, un pH de 6,5-8,5, en el que dicho pH es al menos 0,5 unidades de pH superior que el pH en la etapa A y la velocidad de alimentación del sustrato se mantiene a entre el 20 % y el 80 % de la velocidad de alimentación del sustrato de la etapa A;

en el que la toxina bacteriana tiene al menos el 90 % de identidad de secuencia con la SEC ID N.º: 32 y el procedimiento se lleva a cabo en un fermentador que contiene desde 10-2000 litros de cultivo.

25 Se divulga también un polinucleótido que comprende una parte de secuencia señal 5' y una parte de toxina 3' en el que;

(a) La parte de secuencia señal 5' codifica un polipéptido heterólogo que tiene una secuencia de aminoácidos capaz de dirigir transporte de una proteína heteróloga al periplasma bacteriano; y

30 (b) la parte de toxina 3' codifica un polipéptido que tiene una secuencia de aminoácidos al menos el 90 % idéntica a la SEC ID N.º: 32 o fragmentos de la misma que codifican al menos 15 aminoácidos y/o al menos un epitopo de células B o T.

También se divulga un polinucleótido que comprende una parte de secuencia señal 5' y una parte de toxina 3' en la que;

35 (a) la parte de secuencia señal 5' codifica un polipéptido que tiene una secuencia de aminoácidos capaz de dirigir transporte de una proteína heteróloga al periplasma bacteriano y en el que la secuencia señal 5' no se deriva de *C. diphtheriae*; y

(b) la parte de toxina 3' codifica un polipéptido que tiene una secuencia de aminoácidos al menos el 90 % idéntica a la SEC ID N.º: 32 o fragmentos de la misma que codifican al menos 15 aminoácidos y/o al menos un epitopo de células B o T.

40 Adicionalmente, se divulga un polinucleótido que comprende una parte de secuencia señal 5' y una parte de toxina 3' en el que la secuencia de la parte señal 5' codifica un péptido señal que tiene una secuencia de aminoácidos de

(a) SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 o 26;

45 (b) variantes de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26, que varían a partir de las secuencias correspondientes por 1, 2 o 3 mutaciones puntuales, inserciones de aminoácidos o deleciones de aminoácidos, que son capaces de dirigir una proteína expresada al periplasma; o

(c) fragmentos de al menos 10 aminoácidos de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 que son capaces de dirigir una proteína expresada al periplasma;

y la parte de toxina 3' codifica una toxina bacteriana o un fragmento o variante de la misma.

Un vector que comprende el polinucleótido de la invención unido a un promotor inducible y una célula huésped que comprende el polinucleótido o el vector se divulga también así como un polipéptido codificado por el polinucleótido.

5 En un aspecto de la invención se proporciona un procedimiento de fabricación de un conjugado que comprende las etapas de a) fabricar una toxina bacteriana usando el procedimiento de la invención y b) conjugar la toxina bacteriana de la etapa a) a un antígeno.

En un aspecto adicional de la invención se proporciona un procedimiento de elaboración de una vacuna que comprende las etapas de a) marcar una toxina bacteriana o conjugado usando el procedimiento de la invención y b) mezclar la toxina bacteriana o conjugado de la misma con un excipiente farmacéuticamente aceptable.

Descripción de las figuras

10 Figura 1 – Visión general del procedimiento de clonación para producir las construcciones de CRM197-secuencia señal. Una sección de DNA que codifica una secuencia señal condensada con el extremo N-terminal de CRM197 se corta del plásmido A por digestión con NdeI y SacI. La secuencia C-terminal de CRM197 SEC ID N.º: 31 se corta de plásmido B por digestión con AatII y XhoI. Estas dos secuencias se ajustan en un vector pET26b por digestión del vector pET26b con NdeI, SacI, AatII y XhoI y ligación usando DNA ligasa. El vector resultante (pRIT16668) se usó para producir las otras construcciones de secuencia señal-CRM197 digiriendo plásmidos que contienen las secuencias señal y el vector pRIT16668 con AgeI y NdeI y ajustando estas conjuntamente usando DNA ligasa.

Figura 2 – Visión general del procedimiento de clonación para producir construcciones que contienen la secuencia CRM197 sin una secuencia señal N-terminal (pRIT16669). Esto se llevó a cabo insertando un fragmento de PCR que codifica CRM197 dentro de pRIT16668 usando digestión con NdeI y AatII seguida por ligación.

20 Figura 3 – Visión general del procedimiento de clonación para producir las construcciones que contienen la secuencia señal FlgI. Esto se llevó a cabo insertando un fragmento de PCR que codifica la CRM197 y la secuencia señal FlgI (SEC ID N.º: 23) dentro de pRIT16669 usando digestión con NdeI y AatII seguida por ligación. Esto produjo el plásmido pRIT16681.

25 Figura 4 – Geles que demuestran la inducción de DsbA-CRM197 3 horas después de tratamiento con IPTG a 30 °C en tres cepas diferentes. Se usó gel Novex TG al 10-20 %. El gel Panel A es una transferencia Western teñida usando anti-DTPa y NBT-BCIP antirrátón. El gel Panel B se tiñe usando azul de Coomassie. El carril 1 contiene el marcador de peso molecular, el carril 2 contiene una cepa BLR(DE3) que no se ha inducido, el carril 3 contiene una cepa BLR(DE3) que se ha inducido, el carril 4 contiene una cepa B834 (DE3) que no se ha inducido, el carril 5 contiene una cepa B384 (DE3) que se ha inducido y el carril 6 contiene una cepa HMS174(DE3) que se ha inducido.

30 Figura 5 – Geles que demuestran expresión de CRM-197 en la célula bacteriana y en el medio de cultivo. El gel Panel B muestra una transferencia Western teñida usando anti-DTPa y NBT-BCIP antirrátón. El Panel A se tiñe usando azul de Coomassie. Los contenidos de cada carril se describen en la tabla a continuación:

Tabla 1

Carril Número	Secuencia señal	Temperatura de cultivo	SEC ID N.º de nucleótidos	SEC ID N.º de aminoácidos
1	DsbA	Durante toda una noche 23 °C	25	26
2	OmpA	Durante toda una noche 23 °C	5	6
3	NspA	Durante toda una noche 23 °C	7	8
4	FlgI	Durante toda una noche 23 °C	23	24
5	TolB	Durante toda una noche 23 °C	19	20
6	SfmC	Durante toda una noche 23 °C	11	12
7	TorT	Durante toda una noche 23 °C	9	10

Carril Número	Secuencia señal	Temperatura de cultivo	SEC ID N.º de nucleótidos	SEC ID N.º de aminoácidos
8	FocC	Durante toda una noche 23 °C	13	14
9	Ccmh	Durante toda una noche 23 °C	15	16
10	Yra1	Durante toda una noche 23 °C	17	18
11	NikA	Durante toda una noche 23 °C	21	22
12	PhtE	Durante toda una noche 23 °C	1	2
13	SipA	Durante toda una noche 23 °C	3	4
14	Marcador	250, 150, 100 75, 50, 37, 25, 20		
15	DsbA	No inducido	25	26
16	DsbA	Durante toda una noche 30 °C	25	26
17	FocC	Durante toda una noche 30 °C	13	14
18	CcmH	Durante toda una noche 30 °C	15	16

Figura 6 - Geles que demuestran inducción de DsbA-CRM197 durante 3 horas a 30 °C en tres cepas diferentes. Se usó gel Novex TG al 10-20 %. El Panel A es una transferencia Western teñida usando anti-DTPa y NBT-BCIP antirrátón. El gel Panel B se tiñe usando azul de Coomassie. El carril 1 contiene el marcador de peso molecular, el carril 2 contiene la cepa BLR(DE3) que no se ha inducido, el carril 3 contiene la cepa BLR(DE3) que se ha inducido, el carril 4 contiene la fracción soluble de una cepa BLR(DE3) que se ha inducido, el carril 5 contiene la fracción insoluble de una cepa BLR(DE3) que se ha inducido, el carril 6 contiene la cepa B834(DE3) que no se ha inducido, el carril 7 contiene la fracción soluble inducida de B834(DE3) y el carril 8 contiene la fracción insoluble inducida por B834(DE3).

Figura 7 – Geles que comparan la expresión de la construcción Flgl optimizada con o sin la adición de tampón de fosfato de potasio 1 M a una concentración de 100 mM en la fase de inducción. Se usó gel Novex TG al 10-20 %. El carril 1 contiene el marcador de peso molecular, el carril 2 contiene el extracto celular a 0 minutos a partir de inducción (inducción a 30 °C), el carril 3 contiene extracto celular a 2 horas a partir de la inducción (inducción a 30 °C), el carril 4 contiene extracto celular a 4 horas a partir de la inducción (inducción a 30 °C), el carril 5 contiene extracto celular después de inducción durante toda una noche a 30 °C, el carril 6 contiene el medio después de 4 horas a 30 °C, el carril 7 contiene el medio otra vez después de inducción a 30 °C durante toda una noche, el carril 8 contiene extracto celular a dos horas a partir de la inducción (inducción a 23 °C), el carril 9 contiene extracto celular a 4 horas a partir de inducción (inducción a 23 °C), el carril 10 contiene extracto celular después de inducción durante toda una noche, el carril 11 contiene medio 4 horas a partir de la inducción (inducción a 23 °C) y el carril 12 contiene medio después de inducción durante toda una noche (inducción a 23 °C).

Figura 8 – Representación de un perfil de fermentación con los parámetros del procedimiento sometidos a seguimiento durante fermentación de alimentación discontinua a escala de 20 litros. La línea 1 describe la cantidad de sustrato añadido (gramos), la línea 2 describe el pH, la línea 3 describe la velocidad de agitación (rpm), la línea 4 describe el pO₂ (%), la línea 5 describe la temperatura (°C) y la línea 6 describe la cantidad de base añadida (gramos).

Figura 9 – Listados de secuencias de polipéptidos y polinucleótidos de la invención.

Figura 10 – Representación de la producción de CRM197 en las fracciones periplásmicas y las fracciones asociadas

a células como una función de la velocidad de alimentación y pH durante la inducción, para crecimiento llevado a cabo a pH 6,8. El panel izquierdo muestra producción de CRM197 periplásmico. El panel derecho describe la producción de CRM197 asociada a células.

Descripción detallada de la invención

5 Definiciones

El término "Polinucleótido(s)" hace referencia generalmente a cualquier polirribonucleótido o polidesoxirribonucleótido, que puede ser RNA o DNA no modificado o RNA o DNA modificado que incluye regiones/formas de cadena simple y de cadena doble.

El término polipéptido hace referencia a cualquier péptido que comprenda al menos 10 aminoácidos.

10 La expresión "polinucleótido que codifica un péptido" como se usa en la presente memoria comprende polinucleótidos que incluyen una secuencia que codifica un péptido o polipéptido de la invención. El término también comprende polinucleótidos que incluyen una región continua individual o que incluyen regiones discontinuas que codifican el péptido o el polipéptido (por ejemplo, polinucleótidos interrumpidos por fago integrado, una secuencia de inserción integrada, una secuencia de vector integrada, una secuencia de trasposón integrada, o debido a edición de
15 RNA o reorganización de DNA genómico) conjuntamente con regiones adicionales, que también pueden contener secuencias codificantes y/o secuencias no codificantes.

El polinucleótido "variantes de la parte de secuencia señal 5'" son secuencias de polinucleótidos que codifican secuencias señal que contienen, sustitución de 1, 2, 3, 4 o 5 aminoácidos, o mutaciones de adición de aminoácidos o mutaciones de delección de aminoácidos comparados con la correspondiente señal de tipo silvestre.

20 Una "delección de aminoácido" es la eliminación de un aminoácido a partir de la secuencia de aminoácidos de una proteína.

Una "adición de aminoácido" es la adición de un aminoácido a partir de la secuencia de aminoácidos de una proteína.

25 Una "sustitución de aminoácido" es el reemplazo de un aminoácido con otro aminoácido, en la secuencia de una proteína.

Las "variantes de la parte de toxina 3'" de polinucleótidos son secuencias de polinucleótidos que codifican secuencias polipeptídicas que tienen identidad del 80 %, el 85 %, el 90 %, el 95 %, el 98 % o el 100 % de un polipéptido de toxina. Se da más adelante una definición para identidad.

30 En general "variantes" son opcionalmente polipéptidos que varían a partir de los referentes por sustituciones de aminoácidos conservadoras, a través de las que un residuo está sustituido por otro con características similares. Típicamente tales sustituciones son entre Ala, Val, Leu e Ile; entre Ser y Thr; entre los residuos ácidos Asp y Glu; entre Asn y Gln; y entre los residuos básicos Lys y Arg; o los residuos aromáticos Phe y Tyr.

"Fragmentos de la parte de secuencia señal 5'" son secuencias que codifican al menos 10, 15 o 20 aminoácidos de un péptido señal.

35 "Fragmentos de la parte de toxina 3'" son secuencias que codifican una parte contigua de al menos 5, 10, 15, 20, 30, 40, 50, 100 o 200 aminoácidos de un polipéptido de toxina, que tiene opcionalmente actividad inmunógena. Un péptido con "actividad inmunógena" o que es "inmunógeno" es capaz (si es necesario cuando se acopla a un vehículo) de provocar una respuesta inmune que reconoce la toxina respectiva. Los fragmentos preferidos son aquellos polinucleótidos que codifican un epítipo de células B o células T y los polinucleótidos que comprenden
40 dichos fragmentos de polinucleótido, recombinantes. Opcionalmente fragmentos de estas toxinas son fragmentos que son inmunogénicos.

"Identidad", como se conoce en la técnica, es una relación entre dos o más secuencias polipeptídicas o entre dos o más secuencias polinucleotídicas, como puede ser el caso, como se determina comparando las secuencias. En la
45 técnica, "identidad" también significa el grado de parentesco entre secuencias de polipéptidos o polinucleótidos, como puede ser el caso, según se determina por la coincidencia entre sucesiones de tales secuencias. La "identidad" se puede calcular fácilmente por procedimientos conocidos incluyendo pero no limitados a aquellos descritos en (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, Nueva York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, Nueva York, 1993; Computer
50 Analysis of Sequence Data, Parte I, Griffin, A.M. y Griffin, H.G., eds., Humana Press, Nueva Jersey, 1994; Sequence Analysis in Molecular Biology, von Heine, G., Academic Press, 1987; y Sequence Analysis Primer, Gribskov, M. y Devereux, J., eds., M Stockton Press, Nueva York, 1991; y Carillo, H. y Lipman, D., SIAM J. Applied Math., 48: 1073 (1988)). Procedimientos de determinación de identidad se diseñan dando la coincidencia más grande entre las secuencias puestas a prueba. Además, los procedimientos de determinación de identidad están codificados en programas de ordenador disponibles públicamente. Los procedimientos de determinación de identidad entre dos

5 secuencias incluyen, pero no se limitan a, el programa GAP en el paquete de programas GCG (Devereux, J. y col., Nucleic Acids Research 12 (1): 387 (1984)), BLASTP, BLASTN (Altschul, S.F. y col., J. Molec. Biol. 215: 403-410 (1990)) y FASTA (Pearson y Lipman Proc. Natl. Acad. Sci. EE.UU. 85: 2444-2448 (1988)). La familia de programas BLAST está públicamente disponible a partir de NCBI y de otras fuentes (BLAST Manual, Altschul, S. y col., NCBI NLM NIH Bethesda, MD 20894; Altschul, S. y col., J. Mol. Biol. 215: 403-410 (1990)). El bien conocido algoritmo de Smith Waterman se puede usar también para determinar identidad.

Los parámetros para comparación de secuencia polipeptídica incluyen lo siguiente:

Algorithm: Needleman y Wunsch, J. Mol Biol. 48: 443-453 (1970)

Matriz de comparación: BLOSSUM62 de Henikoff y Henikoff,

10 Proc. Natl. Acad. Sci. EE.UU. 89: 10915-10919 (1992)

Penalización de Hueco: 8

Penalización de Longitud de Hueco: 2

15 Un programa útil con estos parámetros está disponible públicamente como el programa de "huecos" a partir del Genetics Computer Group, Madison WI. Los parámetros ya mencionados son los parámetros por defecto para comparaciones peptídicas (junto con ninguna penalización para los huecos finales).

Los parámetros para comparación de nucleótidos incluyen los siguientes:

Algoritmo: Needleman y Wunsch, J. Mol Biol. 48: 443-453 (1970)

Matriz de comparación: coincidencias = +10, desajustes = 0

Penalización de Hueco: 50

20 Penalización de Longitud de Hueco: 3

Disponible como: el programa de "huecos" de Genetics Computer Group, Madison WI. Estos son los parámetros por defecto para comparaciones de ácidos nucleicos.

La expresión 'péptido señal' o 'polipéptido señal' se refiere a un péptido que es capaz de dirigir una proteína expresada al periplasma.

25 La expresión "toxina bacteriana" comprende tanto toxinas como toxoides.

El término "toxoides" describe toxinas que están parcialmente o completamente inactivadas por, por ejemplo la introducción de mutaciones, deleciones o inserciones puntuales.

30 Los términos "comprendiendo", "comprender" y "comprende" en la presente memoria se desean por los autores de la invención para ser opcionalmente sustituibles con las expresiones "constituido por", "consisten en" y "consiste en", respectivamente, en cada caso.

La expresión 'proteína heteróloga' se refiere a una proteína que no es nativa para el tipo celular en el que se expresa. De forma similar, la expresión 'polipéptido heterólogo' se refiere a un polipéptido que no es nativo para el tipo celular en el que se expresa.

35 La expresión 'polipéptido no derivado de *C. diphtheriae*' hace referencia a un polipéptido que es diferente en secuencia a un polipéptido hallado en *C. diphtheriae* nativa (no recombinante).

Polinucleótidos de uso en la invención

Un aspecto de la divulgación se refiere a polinucleótidos que codifican toxinas que se expresan periplásmicamente.

40 En general la presencia de una secuencia señal en la proteína facilita el transporte de la proteína dentro del periplasma (huéspedes procariontes) o la secreción de la proteína (huéspedes eucariotas). En los huéspedes procariontes la secuencia señal dirige a la proteína naciente a través de la membrana interna dentro del espacio periplásmico; la secuencia señal se escinde después. Una secuencia señal es capaz de dirigir una proteína expresada al periplasma si, cuando se une a un polipéptido de interés, durante la traducción del polipéptido en una bacteria gram negativa, más de dicho polipéptido se encuentra en el periplasma de una bacteria gram negativa que en la ausencia de una secuencia señal. En una realización al menos el 50, 60, 70, 80, 90 o 100 % del polipéptido de interés se dirige al periplasma cuando se expresa en una bacteria gram negativa tal como *E. coli*.

45 Un ensayo para probar si una secuencia señal es capaz de dirigir la expresión periplásmica se puede llevar a cabo usando elementos comunicadores. Por ejemplo una secuencia señal periplásmica se puede ajustar corriente arriba

de un gen para una proteína fluorescente verde, esta proteína se puede expresar en una célula huésped de la invención. Se puede usar un microscopio para juzgar los niveles comparativos de la proteína fluorescente verde en el citoplasma y el periplasma.

5 Una proteína o péptido se transporta co-traduccionalmente si el transporte tiene lugar antes de la síntesis de una cantidad sustancial de la cadena polipeptídica. La SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26 codifica péptidos que son capaces de dirigir una proteína expresada al periplasma a través de transporte co-traducciona, mientras que las SEC ID N.ºs: 6 y 8 codifican péptidos que son capaces de dirigir una proteína expresada al periplasma por transporte postraducciona.

10 Se divulga un polinucleótido (por ejemplo para expresión de una toxina en una célula bacteriana) que comprende una parte de secuencia señal 5' y una parte de toxina 3', en el que

(a) La parte de secuencia señal 5' codifica un polipéptido heterólogo (señal) que tiene una secuencia de aminoácidos capaz de dirigir transporte de una proteína heteróloga (toxina) al periplasma bacteriano; y

15 (b) la parte de toxina 3' codifica un polipéptido que tiene una secuencia de aminoácidos del al menos el 75 %, 80 %, 85 %, 90 %, 95 % o 99 % idéntico a la SEC ID N.º: 32 o a fragmentos de la misma (que pueden ser inmunógenos) que codifican al menos 15 aminoácidos (contiguos) y/o al menos un epitopo de células B o T.

Se divulga adicionalmente un polinucleótido (para expresión de un polipéptido en una célula bacteriana) que comprende una parte de secuencia señal 5' y una parte de toxina 3', en el que

20 (a) la parte de secuencia señal 5' codifica un polipéptido que tiene una secuencia de aminoácidos capaz de dirigir transporte de una proteína heteróloga al periplasma bacteriano y en el que la secuencia señal 5' no se deriva de *C. diphtheriae*; y

(b) la parte de toxina 3' codifica un polipéptido que tiene una secuencia de aminoácidos al menos el 90 % idéntica a la SEC ID N.º: 32 o fragmentos de la misma que codifican al menos 15 aminoácidos y/o al menos un epitopo de células B o T.

En un polinucleótido la secuencia señal 5' no es la secuencia señal tox de *Corynebacterium diphtheriae*.

25 Está dentro de las capacidades de la persona experta identificar epitopos de células B y T. Se pueden identificar epitopos de células B por predicción de estructura bidimensional, por ejemplo usando el programa PSIPRED (de David Jones, Brunel Bioinformatics Group, Dept. Biological Sciences, Brunel University, Uxbridge UB8 3PH, Reino Unido) (Fig. 4). El índice antigénico se calcula en base al procedimiento descrito por Jameson y Wolf (CABIOS 4:181-186 [1988]). Los parámetros usados en este programa son el índice antigénico y la longitud mínima para un péptido antigénico. Se usó un índice antigénico de 0,9 para un mínimo de 5 aminoácidos consecutivos como umbral en el programa. Los epitopos de células T se puede identificar, por ejemplo, por el procedimiento de epitopo descrito por Sturniolo y col. (Nature Biotech. 17: 555-561 [1999]).

30 Por claridad la frase 'que tiene una secuencia de aminoácidos capaz de dirigir transporte de una proteína heteróloga al periplasma bacteriano' quiere decir lo mismo que 'que tiene una secuencia de aminoácidos capaz de dirigir transporte al periplasma bacteriano de una proteína heteróloga'.

Este polinucleótido puede codificar un polipéptido que tiene una secuencia de aminoácidos capaz de dirigir transporte co-traducciona de una proteína heteróloga al periplasma bacteriano.

Opcionalmente la parte de la toxina 3' codifica SEC ID N.º 32. Opcionalmente la parte de la toxina 3' codifica DT. Opcionalmente la toxina 3' comprende SEC ID N.º 31.

40 Opcionalmente la parte de la secuencia señal 5' codifica una cualquiera de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26, o una cualquiera de SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26, o una cualquiera de SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, o 24, o SEC ID N.º: 24 o una cualquiera de SEC ID N.º: 2, 4, o 24, o una cualquiera de SEC ID N.º: 2, 10, o 24, o una cualquiera de SEC ID N.º: 2, 12, o 24, o una cualquiera de SEC ID N.º: 2, 14, o 24, o una cualquiera de SEC ID N.º: 4, 10 o 24, o una cualquiera de SEC ID N.º: 4, 12, o 24, o una cualquiera de SEC ID N.º: 4, 16, o 24, o una cualquiera de SEC ID N.º: 4, 18 o 24, o una cualquiera de SEC ID N.º: 4, 20 o 24, o una cualquiera de SEC ID N.º: 4, 22, o 24, o una cualquiera de SEC ID N.º: 10, 12, o 24, o una cualquiera de SEC ID N.º: 10, 14, o 24, o una cualquiera de SEC ID N.º: 10, 16, o 24, o una cualquiera de SEC ID N.º: 10, 18, o 24, o una cualquiera de SEC ID N.º: 10, 22 o 24, o una cualquiera de SEC ID N.º: 12, 14, o 24, o una cualquiera de SEC ID N.º: 12, 16, o 24, o una cualquiera de SEC ID N.º: 12, 18, o 24, o una cualquiera de SEC ID N.º: 12, 20, o 24, o una cualquiera de SEC ID N.º: 12, 22, o 24, o una cualquiera de SEC ID N.º: 14, 16, o 24, o una cualquiera de SEC ID N.º: 14, 18, o 24, o una cualquiera de SEC ID N.º: 14, 20 o 24, o una cualquiera de SEC ID N.º: 14, 22, o 24, o una cualquiera de SEC ID N.º: 16, 18, o 24, o una cualquiera de SEC ID N.º: 16, 20 o 24, o una cualquiera de SEC ID N.º: 16, 22, o 24, o una cualquiera de SEC ID N.º: 18, 20 o 24, o una cualquiera de SEC ID N.º: 18, 22, o 24.

Opcionalmente la parte de la secuencia señal 5' divulgada codifica (conteniendo variantes) 1, 2 o 3 mutaciones puntuales, inserciones o deleciones, de una cualquiera de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 o una cualquiera de SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26, o una cualquiera de SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, o 24, o SEC ID N.º: 24 o una cualquiera de SEC ID N.º: 2, 4, o 24, o una cualquiera de SEC ID N.º: 2, 10, o 24, o una cualquiera de SEC ID N.º: 2, 12, o 24, o una cualquiera de SEC ID N.º: 2, 14, o 24, o una cualquiera de SEC ID N.º: 4, 10 o 24, o una cualquiera de SEC ID N.º: 4, 12, o 24, o una cualquiera de SEC ID N.º: 4, 16, o 24, o una cualquiera de SEC ID N.º: 4, 18 o 24, o una cualquiera de SEC ID N.º: 4, 20 o 24, o una cualquiera de SEC ID N.º: 4, 22, o 24, o una cualquiera de SEC ID N.º: 10, 12, o 24, o una cualquiera de SEC ID N.º: 10, 14, o 24, o una cualquiera de SEC ID N.º: 10, 16, o 24, o una cualquiera de SEC ID N.º: 10, 18, o 24, o una cualquiera de SEC ID N.º: 10, 22 o 24, o una cualquiera de SEC ID N.º: 12, 14, o 24, o una cualquiera de SEC ID N.º: 12, 16, o 24, o una cualquiera de SEC ID N.º: 12, 18, o 24, o una cualquiera de SEC ID N.º: 12, 20, o 24, o una cualquiera de SEC ID N.º: 12, 22, o 24, o una cualquiera de SEC ID N.º: 14, 16, o 24, o una cualquiera de SEC ID N.º: 14, 18, o 24, o una cualquiera de SEC ID N.º: 14, 20 o 24, o una cualquiera de SEC ID N.º: 14, 22, o 24, o una cualquiera de SEC ID N.º: 16, 18, o 24, o una cualquiera de SEC ID N.º: 16, 20 o 24, o una cualquiera de SEC ID N.º: 16, 22, o 24, o una cualquiera de SEC ID N.º: 18, 20 o 24, o una cualquiera de SEC ID N.º 18, 22, o 24.

Opcionalmente, la parte de la secuencia señal 5' divulgada codifica fragmentos de al menos 10, 12, 15, 18 o 20 aminoácidos de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 o una cualquiera de SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26, o una cualquiera de SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, o 24, o SEC ID N.º: 24, o una cualquiera de SEC ID N.º: 2, 4, o 24, o una cualquiera de SEC ID N.º: 2, 10, o 24, o una cualquiera de SEC ID N.º: 2, 12, o 24, o una cualquiera de SEC ID N.º: 2, 14, o 24, o una cualquiera de SEC ID N.º: 4, 10 o 24, o una cualquiera de SEC ID N.º: 4, 12, o 24, o una cualquiera de SEC ID N.º: 4, 16, o 24, o una cualquiera de SEC ID N.º: 4, 18 o 24, o una cualquiera de SEC ID N.º: 4, 20 o 24, o una cualquiera de SEC ID N.º: 4, 22, o 24, o una cualquiera de SEC ID N.º: 10, 12, o 24, o una cualquiera de SEC ID N.º: 10, 14, o 24, o una cualquiera de SEC ID N.º: 10, 16, o 24, o una cualquiera de SEC ID N.º: 10, 18, o 24, o una cualquiera de SEC ID N.º: 10, 22 o 24, o una cualquiera de SEC ID N.º: 12, 14, o 24, o una cualquiera de SEC ID N.º: 12, 16, o 24, o una cualquiera de SEC ID N.º: 12, 18, o 24, o una cualquiera de SEC ID N.º: 12, 20, o 24, o una cualquiera de SEC ID N.º: 12, 22, o 24, o una cualquiera de SEC ID N.º: 14, 16, o 24, o una cualquiera de SEC ID N.º: 14, 18, o 24, o una cualquiera de SEC ID N.º: 14, 20 o 24, o una cualquiera de SEC ID N.º: 14, 22, o 24, o una cualquiera de SEC ID N.º: 16, 18, o 24, o una cualquiera de SEC ID N.º: 16, 20 o 24, o una cualquiera de SEC ID N.º: 16, 22, o 24, o una cualquiera de SEC ID N.º 18, 20 o 24, o una cualquiera de SEC ID N.º 18, 22, o 24 que son capaces de dirigir transporte de una proteína al periplasma bacteriano.

Además, se divulga un polinucleótido que comprende una secuencia señal 5' y una parte de toxina 3', en el que la secuencia de parte señal 5' codifica un péptido señal que tiene una secuencia de aminoácidos de;

(a) SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, o 24; o SEC ID N.º: 24, o SEC ID N.º: 2, 4, o 24; o SEC ID N.º: 2, 10, o 24; o SEC ID N.º: 2, 12, o 24; o SEC ID N.º: 2, 14, o 24; o SEC ID N.º: 4, 10 o 24; o SEC ID N.º: 4, 12, o 24; o SEC ID N.º: 4, 16, o 24; o SEC ID N.º: 4, 18 o 24; o SEC ID N.º: 4, 20 o 24; o SEC ID N.º: 4, 22, o 24; o SEC ID N.º: 10, 12, o 24; o SEC ID N.º: 10, 14, o 24; o SEC ID N.º: 10, 16, o 24; o SEC ID N.º: 10, 18, o 24; o SEC ID N.º: 10, 22 o 24; o SEC ID N.º: 12, 14, o 24; o SEC ID N.º: 12, 16, o 24; o SEC ID N.º: 12, 18, o 24; o SEC ID N.º: 12, 20, o 24; o SEC ID N.º: 12, 22, o 24; o SEC ID N.º: 14, 16, o 24; o SEC ID N.º: 14, 18, o 24; o SEC ID N.º: 14, 20 o 24; o SEC ID N.º: 14, 22, o 24; o SEC ID N.º: 16, 18, o 24; o SEC ID N.º: 16, 20 o 24; o SEC ID N.º: 16, 22, o 24; o SEC ID N.º: 18, 20 o 24; o una cualquiera de SEC ID N.º 18, 22, o 24.

(b) (variantes de) SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, o 24; o SEC ID N.º: 24, o SEC ID N.º: 2, 4, o 24; o SEC ID N.º: 2, 10, o 24; o SEC ID N.º: 2, 12, o 24; o SEC ID N.º: 2, 14, o 24; o SEC ID N.º: 4, 10 o 24; o SEC ID N.º: 4, 12, o 24; o SEC ID N.º: 4, 16, o 24; o SEC ID N.º: 4, 18 o 24; o SEC ID N.º: 4, 20 o 24; o SEC ID N.º: 4, 22, o 24; o SEC ID N.º: 10, 12, o 24; o SEC ID N.º: 10, 14, o 24; o SEC ID N.º: 10, 16, o 24; o SEC ID N.º: 10, 18, o 24; o SEC ID N.º: 10, 22 o 24; o SEC ID N.º: 12, 14, o 24; o SEC ID N.º: 12, 16, o 24; o SEC ID N.º: 12, 18, o 24; o SEC ID N.º: 12, 20, o 24; o SEC ID N.º: 12, 22, o 24; o SEC ID N.º: 14, 16, o 24; o SEC ID N.º: 14, 18, o 24; o SEC ID N.º: 14, 20 o 24; o SEC ID N.º: 14, 22, o 24; o SEC ID N.º: 16, 18, o 24; o SEC ID N.º: 16, 20 o 24; o SEC ID N.º: 16, 22, o 24; o SEC ID N.º: 18, 20 o 24; o una cualquiera de SEC ID N.º 18, 22, o 24 que varían a partir de las secuencias correspondientes en 1, 2 o 3 mutaciones puntuales, inserciones de aminoácidos o deleciones de aminoácidos que son capaces de dirigir una proteína expresada al periplasma; o

(c) fragmentos de al menos 10 aminoácidos de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, o 24; o SEC ID N.º: 24, o SEC ID N.º: 2, 4, o 24; o SEC ID N.º: 2, 10, o 24; o SEC ID N.º: 2, 12, o 24; o SEC ID N.º: 2, 14, o 24; o SEC ID N.º: 4, 10 o 24; o SEC ID N.º: 4, 12, o 24; o SEC ID N.º: 4, 16, o 24; o SEC ID N.º: 4, 18 o 24; o SEC ID N.º: 4, 20 o 24; o SEC ID N.º: 4, 22, o 24; o SEC ID N.º: 10, 12, o 24; o SEC ID N.º: 10, 14, o 24; o SEC ID N.º: 10, 16, o 24; o SEC ID N.º: 10, 18, o 24; o SEC ID N.º: 10, 22 o 24; o SEC ID N.º: 12, 14, o 24; o SEC ID N.º: 12, 16, o 24; o SEC ID N.º: 12, 18, o 24; o SEC ID N.º: 12, 20, o 24; o SEC ID N.º: 12, 22, o 24; o SEC ID N.º: 14, 16, o 24; o SEC ID N.º: 14, 18, o 24; o SEC ID N.º: 14, 20 o 24; o SEC ID N.º: 14, 22, o 24; o SEC ID N.º: 16, 18, o 24; o SEC ID N.º: 16, 20 o 24; o SEC ID N.º: 16, 22, o 24; o SEC ID N.º: 18, 20 o 24; o una cualquiera de SEC ID N.º 18, 22, o 24 que son capaces

de dirigir una proteína expresada al periplasma.

y la parte de toxina 3' codifica una toxina bacteriana o un fragmento o variante de la misma.

Se divulga también un polinucleótido que comprende una secuencia señal 5' y una toxina 3', en el que la secuencia señal 5' codifica un péptido señal que tiene una secuencia de aminoácidos de

5 (d) SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26;

(e) variantes de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26, que varían a partir de las secuencias correspondientes por 1, 2 o 3 mutaciones puntuales, inserciones de aminoácidos o deleciones de aminoácidos, que son capaces de dirigir una proteína expresada al periplasma; o

10 (f) fragmentos de al menos 10 aminoácidos de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 que son capaces de dirigir una proteína expresada al periplasma;

y la parte de toxina 3' codifica una toxina bacteriana o un fragmento o variante de la misma.

Opcionalmente la parte de la secuencia señal 5' del polinucleótido codifica al menos una de las SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26. Por ejemplo la parte de la secuencia señal 5' del polinucleótido de la invención codifica una de las SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26. Opcionalmente la parte de la secuencia señal 5' del polinucleótido de la invención codifica al menos una de las SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEC ID N.º: 2, 4, 10, 12, 14, 16, 18, 20, 22, o 24; o SEC ID N.º: 24, o SEC ID N.º: 2, 4, o 24; o SEC ID N.º: 2, 10, o 24; o SEC ID N.º: 2, 12, o 24; o SEC ID N.º: 2, 14, o 24; o SEC ID N.º: 4, 10 o 24; o SEC ID N.º: 4, 12, o 24; o SEC ID N.º: 4, 16, o 24; o SEC ID N.º: 4, 18 o 24; o SEC ID N.º: 4, 20 o 24; o SEC ID N.º: 4, 22, o 24; o SEC ID N.º: 10, 12, o 24; o SEC ID N.º: 10, 14, o 24; o SEC ID N.º: 10, 16, o 24; o SEC ID N.º: 10, 18, o 24; o SEC ID N.º: 10, 22 o 24; o SEC ID N.º: 12, 14, o 24; o SEC ID N.º: 12, 16, o 24; o SEC ID N.º: 12, 18, o 24; o SEC ID N.º: 12, 20, o 24; o SEC ID N.º: 12, 22, o 24; o SEC ID N.º: 14, 16, o 24; o SEC ID N.º: 14, 18, o 24; o SEC ID N.º: 14, 20 o 24; o SEC ID N.º: 14, 22, o 24; o SEC ID N.º: 16, 18, o 24; o SEC ID N.º: 16, 20 o 24; o SEC ID N.º: 16, 22, o 24; o SEC ID N.º: 18, 20 o 24; o una cualquiera de SEC ID N.º 18, 22, o 24. Las secuencias de nucleótidos codificadas por SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 pueden ser idénticas a las secuencias polinucleótídicas correspondientes de SEC ID N.º: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, o 25. Alternativamente ella puede ser cualquier secuencia, que como un resultado de la redundancia (degeneración) del código genético, también codifica polipéptidos de la SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26. Alternativamente el polinucleótido puede comprender una parte que codifica SEC ID N.º: 33-45.

Los polinucleótidos también pueden comprender una parte de secuencia señal 5' que codifica una variante de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 que es capaz de dirigir una proteína al periplasma.

También se divulga una secuencia señal 5' que comprende un fragmento de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26. Los fragmentos de la invención consisten en partes contiguas de al menos, 10, 15, o 20 aminoácidos de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26. Opcionalmente, los fragmentos conducen a que se transporte al periplasma al menos el 50, 60, 70, 90, 90 o 100 % de la toxina codificada polipeptídica. Opcionalmente los fragmentos tienen las mismas o sustancialmente las mismas propiedades de transporte periplásmico que el polipéptido que comprende la secuencia de aminoácidos correspondiente de SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26. Opcionalmente, los fragmentos son capaces de dirigir transporte cotraduccional de un polipéptido al periplasma.

La parte de la toxina 3' codifica cualquier toxina bacteriana, tal como la toxina de difteria, de tétanos, pertúsica o neumolisina.

Como se divulga en la presente memoria, la parte de toxina 3' codifica un toxoide bacteriano por ejemplo toxoide de *Diphtheria* o CRM 197. Opcionalmente la parte de toxina 3' codifica la secuencia de aminoácidos descrita en la SEC ID N.º:32. Opcionalmente la parte de toxina 3' comprende la secuencia de nucleótidos de SEC ID N.º: 31. En una realización el polipéptido codifica SEC ID N.º:33-45.

45 Alternativamente la parte 3' puede codificar cualquier fragmento de al menos 10, 15, 25, 35, 50, 100, 150 aminoácidos o variante con homología de secuencia del 80 %, 85 %, 90 %, 95 %, 98 % o 99 % de la toxina o el toxoide anteriormente descrito.

Opcionalmente la parte de secuencia señal 5' está directamente 5' de la parte de toxina 3'. Alternativamente están separados por al menos 3, 6, 9, 12, 21, 51, 99, 300 o 1000 nucleótidos adicionales. Por ejemplo estos nucleótidos pueden codificar una o más secuencias peptídicas de al menos 10, 20, 30, 50, 100, 200 o 500 aminoácidos.

Además para todos y cada uno de los polinucleótidos de la invención se proporciona un polinucleótido complementario a él. Se prefiere que estos polinucleótidos complementarios sean totalmente complementarios para cada polinucleótido con los que son complementarios.

Además la divulgación contempla la expresión de cualesquiera de los polinucleótidos dentro de una secuencia que

codifica para una proteína más grande tal como una proteína precursora o una proteína de fusión. A menudo es ventajoso incluir una secuencia de aminoácidos adicional que contiene secuencias secretoras o líderes, pro-secuencias, secuencias que ayudan en purificación tales como residuos de histidina múltiples, o una secuencia adicional para estabilidad durante producción recombinante. Además, se considera también la adición de polipéptido exógeno o cola lipídica o secuencias polinucleotídicas para incrementar el potencial inmunogénico de la molécula final.

Opcionalmente la secuencia señal 5' no es la secuencia señal tox de *C. diphtheriae*.

Polipéptido de uso en la invención

La presente divulgación también proporciona polipéptidos codificados por los polinucleótidos divulgados. Además se divulgan fragmentos y variantes de estos polipéptidos.

Fragmentos de estos polipéptidos se divulgan también. Los fragmentos contienen segmentos tanto de la parte de la secuencia señal 5' como de la parte de la toxina 3'. Opcionalmente estos fragmentos comprenden al menos 10, 15 o 20 aminoácidos de un péptido señal. Opcionalmente estos fragmentos comprenden al menos 10, 15 o 20 de los aminoácidos a partir de la SEC ID N.º: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 o de la SEC ID N.º: 33-45.

Opcionalmente estos fragmentos están al menos dirigidos el 50 %, 60 %, 70 %, 80 %, 90 % o el 100 % al periplasma cuando se expresan en una bacteria gram negativa tal como *E. coli*.

Se divulgan también proteínas de fusión incluyendo el polipéptido o los fragmentos polipeptídicos.

Los polipéptidos, o los fragmentos inmunogénicos de la invención, pueden ser una parte de una proteína más grande tal como un precursor o una proteína de fusión. A menudo es ventajoso incluir una secuencia de aminoácidos adicional que contiene secuencias que ayudan en purificación tales como residuos de histidina múltiples, o una secuencia adicional para estabilidad durante producción recombinante. Además, se considera también la adición de polipéptido exógeno o cola lipídica o secuencias polinucleotídicas para incrementar el potencial inmunogénico de la molécula final.

Vectores y células huésped

La divulgación se refiere también a vectores que comprenden un polinucleótido o polinucleótidos tal como se divulgan, células huésped que están modificadas mediante ingeniería genética con los vectores y la producción de polipéptidos por técnicas recombinantes. Sistemas de traducción libres de células se pueden emplear también para producir tales proteínas usando RNA derivados de las construcciones de DNA de la invención.

Los polipéptidos recombinantes se pueden preparar por procedimientos bien conocidos por aquellos expertos en la técnica a partir de células hospedadoras modificadas mediante ingeniería genética que comprenden sistemas de expresión. De acuerdo con ello, la presente divulgación también se refiere a sistemas de expresión que comprenden un polinucleótido o polinucleótidos de la presente invención, a células huésped que están modificadas mediante ingeniería genética con tales sistemas de expresión y a la producción de polipéptidos de la invención por técnicas recombinantes. Ello se refiere adicionalmente a un polinucleótido ligado a un promotor inducible tal que cuando el promotor se induce se expresa un polipéptido codificado por el polinucleótido. Opcionalmente el promotor inducible está activado por adición de una cantidad suficiente de IPTG. Opcionalmente este está a una concentración de entre 0,1 y 10 mM, 0,1 y 5 mM, 0,1 y 2,5 mM, 0,2 y 10 mM, 0,2 y 5 mM, 0,2 y 2,5 mM, 0,4 y 10 mM, 1 y 10 mM, 1 y 5 mM, 2,5 y 10 mM, 2,5 y 5 mM, 5 y 10 mM.

Para producción recombinante de los polipéptidos de la invención, las células huésped se pueden modificar mediante ingeniería genética para incorporar sistemas de expresión o partes de los mismos o polinucleótidos de la invención. La introducción de un polinucleótido dentro de la célula huésped se puede llevar a cabo por procedimientos descritos en muchos manuales de laboratorios estándar, tales como Davis y col., BASIC METHODS IN MOLECULAR BIOLOGY, (1986) y Sambrook y col., MOLECULAR CLONING: A LABORATORY MANUAL, 2ª Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), tales como, transfección con fosfato de calcio, transfección mediada por DEAE-dextrano, transvección, microinyección, transfección mediada por lípidos catiónicos, electroporación, conjugación, transducción, introducción mediante raspado de la membrana celular, introducción balística e infección.

Los ejemplos representativos de los huéspedes apropiados incluyen células bacterianas gram negativas, tales como células de *E. coli*, *Acinetobacter*, *Actinobacillus*, *Bordetella*, *Brucella*, *Campylobacter*, *Cyanobacteria*, *Enterobacter*, *Erwinia*, *Franciscella*, *Helicobacter*, *Hemophilus*, *Klebsiella*, *Legionella*, *Moraxella*, *Neisseria*, *Pasteurella*, *Proteus*, *Pseudomonas*, *Salmonella*, *Serratia*, *Shigella*, *Treponema*, *Vibrio*, *Yersinia*, particularmente en *Escherichia coli*.

Se puede usar una gran variedad de sistemas de expresión para producir los polipéptidos. En una realización el vector se deriva de plásmidos bacterianos. Generalmente, cualquier sistema de expresión o vector adecuado para mantener, propagar o expresar polinucleótidos y/o para expresar un polipéptido en un huésped se puede usar para expresión a este respecto. La secuencia de DNA apropiada se puede insertar en el sistema de expresión por cualquiera de una diversidad de técnicas bien conocidas y de rutina, tales como, por ejemplo, aquellas expuestas en

Sambrook y col., MOLECULAR CLONING, A LABORATORY MANUAL, (citado anteriormente).

Fermentación

En un aspecto de la invención se proporciona un procedimiento de producción industrial de una toxina bacteriana expresada periplásmicamente por:

- 5 A. en una fase de alimentación discontinua, cultivar un cultivo de una célula huésped gram-negativa que comprende un polinucleótido que comprende una parte de toxina 3' que codifica una toxina bacteriana y una parte de señal 5' que codifica un polipéptido capaz de dirigir el transporte de dicha toxina bacteriana al periplasma de la célula huésped, en el que el pH es de 6,0-7,0; y
- 10 B. inducir expresión de dicha toxina bacteriana a un pH que es y se mantiene a, un pH de 6,5-8,5, en el que dicho pH es al menos 0,5 unidades de pH superior que el pH en la etapa A; y la velocidad de alimentación del sustrato se mantiene a entre el 20 % y el 80 % de la velocidad de alimentación del sustrato de la etapa A;
- en el que la toxina bacteriana tiene al menos el 90 % de identidad de secuencia con la SEC ID N.º: 32 y el procedimiento se lleva a cabo en un fermentador que contiene desde 10-2000 litros de cultivo.
- 15 La expresión de polipéptidos se induce cuando un agente inductor tal como IPTG se añade a un cultivo de células huésped, causando expresión de polipéptido a una velocidad incrementada.
- La persona experta en la técnica reconocerá como efectuar un cambio entre la etapa a) y la etapa b). Un cambio en una afección (por ejemplo pH, temperatura o velocidad de alimentación de sustrato) entre la etapa a) y la etapa b) quiere decir que la condición promedio general durante la etapa a) o la etapa b) es como se reporta y se puede evaluar por ejemplo, si no hay otra intervención, justamente antes (por ejemplo 5 segundos, 15 segundos, 30 segundos, 1 minuto, 15 minutos, 30 minutos o una hora) de la inducción (etapa a) o justamente después (por ejemplo 5 segundos, 15 segundos, 30 segundos, 1 minuto, 15 minutos, 30 minutos o una hora) de la inducción (etapa b). Claramente los autores de la invención también conciben que la intervención para cambiar las condiciones de fermentación puede ocurrir ligeramente antes o ligeramente después de la inducción para lograr el mismo resultado técnico pero en un escenario tal de nuevo la condición general o promedio durante la etapa a) o etapa b) se cambiará según se revela en la presente memoria.
- 20 En una realización adicional el pH de la etapa a) varía desde 5,0-7,0, 5,0-6,0, 6,0-7,0 o desde 6,5-7,0.
- En una realización el pH de la etapa b) se mantiene a pH mayor de 6,5, 7,0, 7,5, 8,0, o entre 6,5 y 8,5, 6,5 y 7,5, 6,5 y 7,0, 7,0 y 8,5, 7,0 y 7,5, 7,5 y 8,5, 7,5 y 8,0, u 8,0 y 8,5. En una realización adicional el pH se mantiene usando un tampón del grupo constituido por tampón fosfato, tampón Tris y tampón de histidina. Opcionalmente el tampón está a una concentración de 10-200 mM, 50-100 mM, 100-200 mM, 10-50 mM o 50-150 mM. Opcionalmente el tampón es tampón fosfato a 80-120 mM, 80-100 mM o 100 mM.
- 30 En una realización el pH en la etapa b) es al menos, exacta o aproximadamente 2,0, 1,5, 1,0, o 0,5, unidades de pH más altas que el pH en una etapa a).
- En una realización el procedimiento de la invención se lleva a cabo en un fermentador. El pH está incrementado de tal forma que el pH en la etapa b) es mayor que en la etapa a). Opcionalmente este incremento en pH se logra por adición de base por ejemplo hidróxido sódico o amoniaco.
- 35 En una realización adicional el pH de la etapa (b) se mantiene entre 6,5 y 8,5, 6,5 y 7,5, 6,5 y 7,0, 7,0 y 8,5, 7,0 y 8,5, 7,0 y 8,0, 7,0 y 7,5, 7,5 y 8,5, 7,5 y 8,0, 8,0 y 8,5, por la adición de base por ejemplo hidróxido de sodio o amoniaco durante la etapa b).
- 40 El procedimiento de la invención comprende una primera velocidad de alimentación de sustrato en la etapa a) y una segunda velocidad de alimentación de sustrato en la etapa b) en el que la segunda velocidad de alimentación de sustrato es menor que la primera velocidad de alimentación de sustrato. La segunda velocidad de alimentación de sustrato se mantiene entre el 20 % y el 80 % o el 20 % y el 30 % de la velocidad de alimentación de sustrato mantenida durante la etapa a).
- 45 La velocidad de alimentación de sustrato (o velocidad de provisión de sustrato) es la velocidad de adición de sustrato durante las fases discontinua y de inducción (ml min^{-10}) es decir no incluye periodo de fase de alimentación no discontinua inicial alguno de fermentación.
- En una realización adicional la toxina es CRM 197, opcionalmente está codificada por SEC ID N.º: 32.
- 50 La célula huésped bacteriana de este procedimiento se definió en la sección en vectores y células bacterianas. En una realización la célula huésped bacteriana se selecciona del grupo constituido por *E. coli*, *Acinetobacter*, *Actinobacillus*, *Bordetella*, *Brucella*, *Campylobacter*, *Cyanobacteria*, *Enterobacter*, *Erwinia*, *Franciscella*, *Helicobacter*, *Hemophilus*, *Klebsiella*, *Legionella*, *Moraxella*, *Neisseria*, *Pasteurella*, *Proteus*, *Pseudomonas*, *Salmonella*, *Serratia*, *Shigella*, *Treponema*, *Vibrio* y *Yersinia*.

En una realización la célula huésped bacteriana es una cepa de *E. coli*. En una realización adicional la célula huésped bacteriana es una cepa K o una cepa B de *E. coli*. En una realización adicional tal célula huésped bacteriana es una cepa K12 o B834 de *E. coli*.

5 En una realización la temperatura de la etapa a) es más alta que la temperatura de la etapa b). En una realización la etapa a) del procedimiento de la invención se lleva a cabo a una temperatura de 20-40 °C. Opcionalmente la etapa b) del procedimiento de la invención se lleva a cabo a una temperatura de 20-28 °C, 21-27 °C, 22-26 °C, 23-24 °C, 21-24 °C o 22-23 °C

10 En una realización la expresión adicional se induce en la etapa b) por la adición de una cantidad suficiente de IPTG (isopropil-beta-D-tiogalactopiranosido). Opcionalmente este está a una concentración de entre 0,1 y 10 mM, 0,1 y 5 mM, 0,1 y 2,5 mM, 0,2 y 10 mM, 0,2 y 5 mM, 0,2 y 2,5 mM, 0,4 y 10 mM, 1 y 10 mM, 1 y 5 mM, 2,5 y 10 mM, 2,5 y 5 mM, 5 y 10 mM.

En una realización la densidad óptica DO₆₅₀ de la bacteria esta entre 0-2,5 o entre 0,4-1,5 en inducción. La expresión 'en inducción' hace referencia al punto en el procedimiento en el que se añade un agente inductor tal como IPTG, esto ocurrirá en el mismo principio de la etapa b).

15 Un fermentador es cualquier aparato adecuado para la producción industrial de cultivos bacterianos. Sin embargo este término no incluye matraces de cultivo que se usan típicamente para crecimiento de bacterias en una escala menor. Opcionalmente el fermentador contiene 20-500, 50-500, 100-500, 250-500, 400-500, 20-400, 50-400, 100-400, 250-400, 20-250, 100-200, 100-250, 250-300, 300-500, 10-2000, 500-2000, 1000-2000, 1500-2000 o 1000-1500 o alrededor de 150 litros de cultivo.

20 Opcionalmente el cultivo se agita en el fermentador. La agitación es opcionalmente removiendo el cultivo en el fermentador pero puede ser por otros medios adecuados, por ejemplo por agitación, vibromezclador y/o burbujeo de gas.

25 En una realización el nivel de oxígeno disuelto (OD) en el fermentador es (manteniéndose) mayor del 5 %, 10 %, 15 %, 16 %, 17 %, 18 %, 19 %, 20 %, 21 %, 22 %, 23 %, 24 %, 25 % o 30 %. En una realización adicional el OD en el fermentador es 5 %-50 %, 10 %-40 %, 15 %-25 % o 17 %-22 %. Una OD del 100 % es la cantidad de oxígeno presente cuando el medio (en ausencia de un cultivo) está saturado con oxígeno tras hacer burbujear aire comprimido a través del medio a 28 °C y a presión de 0,5 bar.

30 La etapa de fermentación puede estar sometida a una gran cantidad de producción de espuma. Con el fin de controlar la formación de espuma se añade opcionalmente un agente antiespumante al fermentador. Opcionalmente se usa una sonda de espuma o un rompedor de espuma mecánico en el fermentador, esto se puede usar así como el agente antiespumante.

35 En una realización el procedimiento de fabricación de la toxina bacteriana implica eliminación de un péptido señal de la toxina bacteriana dentro de la célula huésped bacteriana para obtener una toxina bacteriana madura. Opcionalmente esta eliminación se lleva a cabo por maquinaria de la célula huésped. En una realización la escisión se lleva a cabo por la peptidasa señal de *E. coli*. Opcionalmente la secuencia señal eliminada se escinde adicionalmente por peptidasa del péptido señal.

En una realización el procedimiento de la invención comprende una etapa adicional c) de cosechar la pasta celular a partir del cultivo.

40 En una realización la centrifugación se usa para cosechar las células. Opcionalmente esto tiene lugar a 5.000-8.000 g, 5.500-7.500 g, 6.000-7.000 g. Opcionalmente esto tiene lugar entre 4 °C-10 °C 5 °C-9 °C, 6 °C-8 °C o 7 °C-8 °C.

En una realización adicional el procedimiento de la invención comprende una etapa adicional de purificar la toxina bacteriana, por ejemplo como una toxina bacteriana madura. En una realización esta etapa implica la disrupción celular y la purificación adicional usando cromatografía y técnicas de filtración.

45 En una realización las células se alteran usando choque osmótico, procedimientos mecánicos o procedimientos enzimáticos. Opcionalmente el procedimiento mecánico comprende usar un homogeneizador mecánico, agitador, sonicación, usando una prensa francesa o un molino de perlas. Opcionalmente el procedimiento enzimático comprende usar lisozima, zimolasa o digestión con lisostafina.

50 En una realización la técnica de cromatografía es cromatografía de afinidad, filtración de gel, cromatografía líquida de alta presión (HPLC) o cromatografía de intercambio iónico. Opcionalmente la cromatografía de afinidad usa una columna de purificación de marca de afinidad, una columna de purificación de anticuerpo, una columna de afinidad de lectina, una columna de purificación de prostaglandinas o una columna de estreptavidina. Opcionalmente el HPLC usa una columna de intercambio iónico, una columna de fase reversa o una columna de exclusión por tamaño. Opcionalmente la columna de intercambio iónico es una columna de intercambio aniónico o una columna de intercambio catiónico.

Conjugación

La invención también proporciona un procedimiento de fabricación de un conjugado que comprende las etapas de i) fabricar una toxina bacteriana usando el procedimiento de la invención y b) conjugar la toxina bacteriana de la etapa ii) a un antígeno.

- 5 En una realización el antígeno es un sacárido bacteriano. Por ejemplo el sacárido bacteriano es un sacárido capsular a partir de una bacteria seleccionada del grupo constituido por *S. pneumoniae*, *H. influenzae*, *N. meningitidis*, *Streptococcus* del grupo B, *Streptococcus* del grupo A, *Salmonella* Vi, enterococos y *S. aureus*. Según se define en la presente memoria un "sacárido" puede ser bien un oligosacárido o bien un polisacárido.

- 10 Los conjugados de sacáridos presentes en las composiciones inmunogénicas se pueden preparar por cualquier técnica de acoplamiento conocida. El procedimiento de la conjugación puede contar con activación del sacárido con tetrafluoroborato de 1-ciano-4-dimetilaminopiridinio (CDAP) para formar un éster de cianato. El sacárido activado puede acoplarse así directamente o por medio de un grupo espaciador (engarce) a un grupo amino en la proteína vehículo. Por ejemplo, el espaciador podría ser cistamina o cisteamina dando un polisacárido tiolado que podría acoplarse al vehículo por medio de un enlace de tioéter obtenido después de reacción con una proteína vehículo
- 15 activada por maleimida (por ejemplo usando GMBS) o una proteína vehículo haloacetilada (por ejemplo usando yodoacetimida [por ejemplo yodoacetimida de etilo HCl] o bromoacetato de N-succinimido o SIAB, o SIA, o SBAP). Preferentemente, el éster de cianato (opcionalmente fabricado por química de CDAP) está acoplado con diamina de hexano o con ADH y el sacárido derivado de amina está conjugado a la proteína vehículo usando química de carbodiimidas (por ejemplo EDAC o EDC) por medio de un grupo carboxilo en el vehículo de la proteína. Tales
- 20 conjugados se describen en la solicitud publicada PCT WO 93/15760 Uniformed Services University y en los documentos WO 95/08348 y WO 96/29094.

- Otras técnicas adecuadas usan carbodiimidas, hidrazidas, ésteres activos, norborano, ácido p-nitrobenzoico, N-hidroxisuccinimida, S-NHS, EDC, TSTU. Muchas se describen en el documento WO 98/42721. La conjugación puede implicar un engarce carbonilo que puede formarse por reacción de un grupo hidroxilo libre del sacárido con
- 25 CDI (Bethell y col. J. Biol. Chem. 1979, 254; 2572-4, Hearn y col. J. Chromatogr. 1981. 218; 509-18) seguido por reacción con una proteína para formar un enlace carbamato. Esto puede implicar reducción del término anomérico a un grupo hidroxilo primario, protección/desprotección opcional del grupo hidroxilo primario, reacción del grupo hidroxilo primario con CDI para formar un intermedio de carbamato de CDI y acoplamiento del intermedio de carbamato de CDI con un grupo amino en una proteína.

- 30 Los conjugados se pueden preparar también por procedimientos de aminación reductora directos como se describe en los documentos US 4365170 (Jennings) y US 4673574 (Anderson). Se describen otros procedimientos en los documentos EP-0-161-188, EP-208375 y EP-0-477508.

- Un procedimiento adicional implica el acoplamiento de un sacárido activado de bromuro de cianógeno (o CDAP) derivatizado con dihidrazida de ácido adípico (ADH) al vehículo de proteínas por condensación de carbodiimida (Chu
- 35 C. y col. Infect. Immunity, 1983 245 256), por ejemplo usando EDAC.

- En una realización, un grupo hidroxilo (preferentemente un grupo hidroxilo activado por ejemplo un grupo hidroxilo activado para fabricar un éster de cianato [por ejemplo con CDAP]) en un sacárido está unido a un grupo amino o carboxílico en una proteína bien directamente o bien indirectamente (a través de un engarce). Donde un engarce está presente, un grupo hidroxilo en un sacárido está preferentemente unido a un grupo amino en un engarce, por
- 40 ejemplo usando conjugación de CDAP. Un grupo amino adicional en el engarce (por ejemplo ADH) puede conjugarse a un grupo de ácido carboxílico en una proteína, por ejemplo usando química de carbodiimidas, por ejemplo usando EDAC. En una realización, el/los sacárido(s) capsular(es) pneumocócico(s) está(n) conjugado(s) al engarce primero antes de que el engarce esté conjugado a la proteína vehículo. Alternativamente el engarce puede estar conjugado al vehículo antes de conjugación al sacárido.

- 45 En general los tipos siguientes de grupos químicos en un vehículo de proteínas se pueden usar para acoplamiento/conjugación:

A) Carboxilo (por ejemplo por medio de ácido aspártico o de ácido glutámico). En una realización este grupo se une a grupos amino en sacáridos directamente o a un grupo amino en un engarce con química de carbodiimidas por ejemplo con EDAC.

- 50 B) Grupo amino (por ejemplo por medio de lisina). En una realización este grupo se une a grupos carboxilo en sacáridos directamente o a un grupo carboxilo en un engarce con química de carbodiimidas por ejemplo con EDAC. En otra realización este grupo está unido a grupos hidroxilo activados con CDAP o CNBr en sacáridos directamente o a tales grupos en un engarce; a sacáridos o engarces que tienen un grupo aldehído; a sacáridos o engarces que tienen un grupo éster de succinimida.

- 55 C) Sulfhidrilo (por ejemplo por medio de cisteína). En una realización este grupo está unido a un sacárido acetilado de bromo o de cloro o a un engarce con química de maleimidas. En una realización este grupo está activado/modificado con bisdiazobencidina.

D) Grupo hidroxilo (por ejemplo por medio de tirosina). En una realización este grupo está activado/modificado con bisdiazobencidina.

E) Grupo imidazolilo (por ejemplo por medio de histidina). En una realización este grupo está activado/modificado con bisdiazobencidina.

5 F) Grupo guanidilo (por ejemplo por medio de arginina).

G) Grupo indolilo (por ejemplo por medio de triptófano).

En un sacárido, en general se pueden usar los siguientes grupos para un acoplamiento: OH, COOH o NH₂. Los grupos aldehído se pueden generar después de diferentes tratamientos conocidos en la técnica tales como: peryodato, hidrólisis ácida, peróxido de hidrógeno, etc.

10 Enfoques de acoplamiento directo:

Sacárido-OH + CNBr o CDAP ----> éster de cianato + NH₂-Prot ----> conjugado

Sacárido-aldehído + NH₂-Prot ----> base de Schiff + NaCNBH₃ ----> conjugado

Sacárido-COOH + NH₂-Prot + EDAC ----> conjugado

Sacárido-NH₂ + COOH-Prot + EDAC ----> conjugado

15 Enfoques de acoplamiento indirecto por medio de espaciador (engarce):

Sacárido-OH + CNBr o CDAP ---> éster de cianato + NH₂----NH₂ ----> sacárido----NH₂ + COOH-Prot + EDAC ----> conjugado

20 Sacárido-OH + CNBr o CDAP ----> éster de cianato + NH₂----SH ----> sacárido----SH + SH-Prot (proteína nativa con una cisteína expuesta u obtenida después de modificación de grupos amino de la proteína por SPDP por ejemplo) ----> sacárido-S-S-Prot

Sacárido-OH + CNBr o CDAP ---> éster de cianato + NH₂----SH ----> sacárido----SH + maleimida-Prot (modificación de grupos amino) ----> conjugado

Sacárido-OH + CNBr o CDAP ---> éster de cianato + NH₂----SH ---> Sacárido-SH + Prot haloacetilada ----> Conjugado

25 Sacárido-COOH + EDAC + NH₂----NH₂ ---> sacárido----NH₂ + EDAC + COOH-Prot ----> conjugado

Sacárido-COOH + EDAC+ NH₂----SH ----> sacárido----SH + SH-Prot (proteína nativa con una cisteína expuesta u obtenida después de la modificación de grupos amino de la proteína por SPDP por ejemplo) ----> sacárido-S-S-Prot

Sacárido-COOH + EDAC+ NH₂----SH ----> sacárido----SH + maleimida-Prot (modificación de grupos amino) ----> conjugado

30 Sacárido-COOH + EDAC + NH₂----SH ---> Sacárido-SH + Prot haloacetilada ----> Conjugado

Sacárido-Aldehído + NH₂----NH₂ ---> sacárido----NH₂ + EDAC + COOH-Prot ----> conjugado

Nota: en vez de EDAC anterior, se puede usar cualquier carbodiimida adecuada.

35 En resumen, los tipos de grupo químico vehículo de proteínas que se pueden usar generalmente para acoplamiento con un sacárido son grupos amino (por ejemplo en residuos de lisina), grupos COOH (por ejemplo en residuos de ácido aspártico y glutámico) y grupos SH (si son accesibles) (por ejemplo en residuos de cisteína).

Vacuna o composiciones inmunogénicas

La presente invención proporciona adicionalmente un procedimiento de elaboración de una vacuna de composición inmunogénica que comprende las etapas de

- A. fabricar una toxina bacteriana o conjugado bacteriano usando el procedimiento de la invención y;
- 40 B. mezclar la toxina bacteriana o el conjugado de la misma con un excipiente farmacéuticamente aceptable.

La vacuna o composición inmunogénica producida por este procedimiento puede adicionalmente comprender antígenos a partir de especies bacterianas adicionales. En una realización la vacuna o la composición inmunogénica comprende antígenos seleccionados de *S. pneumoniae*, *H. influenzae*, *N. meningitidis*, *E. coli*, *M. cattarhalis*, tétanos, difteria, tosferina, *S. epidermidis*, enterococos, o *S. aureus*.

En una etapa adicional los polipéptidos se pueden mezclar con un coadyuvante. La elección de un coadyuvante adecuado para mezclarse con toxinas bacterianas o conjugados fabricados usando los procedimientos de la invención está dentro del conocimiento de la persona experta en la técnica. Los coadyuvantes adecuados incluyen una sal de aluminio tal como gel de hidróxido de aluminio o fosfato de aluminio o alumbre, pero pueden ser también otras sales metálicas tales como aquellas de calcio, magnesio, hierro o cinc, o pueden ser una suspensión insoluble de tirosina acilada, o azúcares acilados, sacáridos derivatizados catiónicamente o aniónicamente, o polifosfatos.

Las preparaciones de vacunas que contienen composiciones inmunogénicas de la presente invención se pueden usar para proteger o tratar a un mamífero susceptible de infección, por medio de administrar dicha vacuna por medio de vía sistémica o mucosal. Estas administraciones pueden incluir inyección *por medio de* las vías intramuscular, intraperitoneal, intradérmica o subcutánea; o *por medio de* administración mucosal a los tractos oral/alimentario, respiratorio, genitourinario. Se prefiere la administración intranasal de vacunas para el tratamiento de neumonía u otitis media (ya que el transporte nasofaríngeo de pneumococos puede prevenirse más eficazmente, atenuando así la infección en su fase más temprana). Aunque la vacuna de la invención puede administrarse como una dosis única, los componentes de la misma se pueden coadministrar conjuntamente al mismo tiempo o a tiempos diferentes (por ejemplo los conjugados de sacáridos pneumocócicos pueden administrarse por separado, al mismo tiempo o 1-2 semanas después de la administración del componente proteico bacteriano cualquiera de la vacuna para coordinación óptima de las respuestas inmunes unas con respecto a otras). Además de una vía individual de administración, se pueden usar 2 vías diferentes de administración. Por ejemplo, sacáridos o conjugados de sacárido se pueden administrar intramuscularmente (o intradérmicamente) y las proteínas bacterianas se pueden administrar intranasalmente (o intradérmicamente). Además, las vacunas de la invención se pueden administrar intramuscularmente para dosis de iniciación e intranasalmente para dosis de refuerzo.

El contenido de toxinas en la vacuna estará típicamente en el intervalo 1-100 µg, preferentemente 5-50 µg, lo más típicamente en el intervalo 5 – 25 µg. Tras una vacunación inicial, los sujetos pueden recibir una o varias inmunizaciones de refuerzo adecuadamente espaciadas.

La preparación de vacunas se describe generalmente en Vaccine Design ("The subunit and adjuvant approach" (eds. Powell M.F. & Newman M.J.) (1995) Plenum Press Nueva York). La encapsulación en liposomas se describe por Fullerton, Patente de los Estados Unidos 4.235.877.

Listado de secuencias

<110> DEWERCHIN Marianne

MARTIN Denis

DEHOTTAY Philippe

GOFFIN Philippe

RIOUX Stephane

<120> Sistema de expression

<130> VB63702

<160> 45

<170> FastSEQ para Windows Version 4.0

<210> 1

<211> 72

<212> DNA

<213> Secuencia artificial

<220>

<221> misc_señal

<222> (0)...(0)

5 <400> 1

```
atgaaattta gtaaaaaata tatagcagct ggatcagctg ttatcgtatc cttgagtcta 60
tgtgcctatg ca                                     72
```

<210> 2

10 <211> 24

<212> PRT

<213> Secuencia artificial

<220>

15 <221> señal

<222> (0)...(0)

<400> 2

```
Met Lys Phe Ser Lys Lys Tyr Ile Ala Ala Gly Ser Ala Val Ile Val
  1      5      10      15
Ser Leu Ser Leu Cys Ala Tyr Ala
20
```

<210> 3

<211> 75

<212> DNA

25 <213> Secuencia artificial

<220>

<221> misc_señal

<222> (0)...(0)

30

<400> 3

ES 2 609 066 T3

atgaaaatga ataaaaaggt actattgaca tcgacaatgg cagcttcgct attatcagtc 60
gcaagtgttc aagca 75

<210> 4

<211> 25

5 <212> PRT

<213> Secuencia artificial

<220>

<221> SEÑAL

10 <222> (0)...(0)

<400> 4

Met	Lys	Met	Asn	Lys	Lys	Val	Leu	Leu	Thr	Ser	Thr	Met	Ala	Ala	Ser
1				5					10					15	
Leu	Leu	Ser	Val	Ala	Ser	Val	Gln	Ala							
			20					25							

15

<210> 5

<211> 63

<212> DNA

<213> Secuencia artificial

20

<220>

<221> misc_señal

<222> (0)...(0)

25 <400> 5

atgaaaaaga cagctatcgc gattgcagtg gcactggctg gtttcgctac cgtagcgcag 60
gcc 63

<210> 6

30 <211> 21

<212> PRT

<213> Secuencia artificial

<220>

<221> SEÑAL

<222> (0)...(0)

5

<400> 6

Met	Lys	Lys	Thr	Ala	Ile	Ala	Ile	Ala	Val	Ala	Leu	Ala	Gly	Phe	Ala
1				5					10					15	
Thr	Val	Ala	Gln	Ala											
			20												

10 <210> 7

<211> 57

<212> DNA

<213> Secuencia artificial

15 <220>

<221> misc_señal

<222> (0)...(0)

<400> 7

20

atgaaaaaag cactgcccac actgattgcc ctgctctcc cggccgccgc actggcg 57

<210> 8

<211> 19

25 <212> PRT

<213> Secuencia artificial

<220>

<221> SEÑAL

30 <222> (0)...(0)

<400> 8

ES 2 609 066 T3

Met Lys Lys Ala Leu Ala Thr Leu Ile Ala Leu Ala Leu Pro Ala Ala
1 5 10 15
Ala Leu Ala

<210> 9

<211> 54

5 <212> DNA

<213> Secuencia artificial

<220>

<221> misc_ señal

10 <222> (0)...(0)

<400> 9

atgcggtac tgctatfff acttcttcc ctttcatgt tgccggcatt ttcg 54

15

<210> 10

<211> 18

<212> PRT

<213> Secuencia artificial

20

<220>

<221> SEÑAL

<222> (0)...(0)

25 <400> 10

Met Arg Val Leu Leu Phe Leu Leu Leu Ser Leu Phe Met Leu Pro Ala
1 5 10 15
Phe Ser

<210> 11

30 <211> 69

<212> DNA

<213> Secuencia artificial

<220>

<221> misc_señal

<222> (0)...(0)

5 <400> 11

```
atgatgacta aaataaagtt attgatgctc atttatatttt atttaatcat ttcggccagc 60
gcccatgct                                         69
```

<210> 12

10 <211> 23

<212> PRT

<213> Secuencia artificial

<220>

15 <221> SEÑAL

<222> (0)...(0)

<400> 12

```
Met Met Thr Lys Ile Lys Leu Leu Met Leu Ile Ile Phe Tyr Leu Ile
  1      5      10      15
Ile Ser Ala Ser Ala His Ala
20
```

<210> 13

<211> 75

<212> DNA

25 <213> Secuencia artificial

<220>

<221> misc_señal

<222> (0)...(0)

30

<400> 13

ES 2 609 066 T3

atgatgaagc acatgcgtat atggggcgtt ctggcatcat ttttagtctt tttttatatt 60
ccgcagagct atgcc 75

<210> 14

<211> 24

5 <212> PRT

<213> Secuencia artificial

<220>

<221> SEÑAL

10 <222> (0)...(0)

<400> 14

Met	Met	Lys	His	Met	Arg	Ile	Trp	Ala	Val	Leu	Ala	Ser	Phe	Leu	Val	Phe
1				5					10					15		
Phe	Tyr	Ile	Pro	Gln	Ser	Tyr	Ala									
			20													

15

<210> 15

<211> 54

<212> DNA

<213> Secuencia artificial

20

<220>

<221> misc_ señal

<222> (0)...(0)

25

<400> 15

atgaggtttt tattgggcgt gctgatgctg atgatctccg gctcagcgt ggcg 54

<210> 16

30 <211> 18

<212> PRT

<213> Secuencia artificial

ES 2 609 066 T3

<220>

<221> SEÑAL

<222> (0)...(0)

5 <400> 16

Met	Arg	Phe	Leu	Leu	Gly	Val	Leu	Met	Leu	Met	Ile	Ser	Gly	Ser	Ala
1				5					10					15	
Leu	Ala														

<210> 17

10 <211> 69

<212> DNA

<213> Secuencia artificial

<220>

15 <221> misc_señal

<222> (0)...(0)

<400> 17

20	atgtcaaaac	gaacattcgc	ggtgatatta	accttgttgt	gtagcttctg	tattggccag	60
	gcgcttgca						69

<210> 18

<211> 23

<212> PRT

25 <213> Secuencia artificial

<220>

<221> SEÑAL

<222> (0)...(0)

30

<400> 18

ES 2 609 066 T3

Met Ser Lys Arg Thr Phe Ala Val Ile Leu Thr Leu Leu Cys Ser Phe
1 5 10 15
Cys Ile Gly Gln Ala Leu Ala
20

<210> 19

<211> 66

5 <212> DNA

<213> Secuencia artificial

<220>

<221> misc_señal

10 <222> (0)...(0)

<400> 19

atgatgaagc aggcattacg agtagcattt ggttttctca tactgtgggc atcagttctg 60
catgct 66

15

<210> 20

<211> 22

<212> PRT

<213> Secuencia artificial

20

<220>

<221> SEÑAL

<222> (0)...(0)

25 <400> 20

Met Met Lys Gln Ala Leu Arg Val Ala Phe Gly Phe Leu Ile Leu Trp
1 5 10 15
Ala Ser Val Leu His Ala
20

<210> 21

30 <211> 66

<212> DNA

ES 2 609 066 T3

<213> Secuencia artificial

<220>

<221> misc_señal

5 <222> (0)...(0)

<400> 21

```
atgctctcca cactccgccg cactctattt gcgctgctgg ctgtgctgc tttatcgtc 60
catgcc                                           66
```

10

<210> 22

<211> 22

<212> PRT

<213> Secuencia artificial

15

<220>

<221> SEÑAL

<222> (0)...(0)

20 <400> 22

```
Met Leu Ser Thr Leu Arg Arg Thr Leu Phe Ala Leu Leu Ala Cys Ala
 1           5           10           15
Ser Phe Ile Val His Ala
                20
```

<210> 23

25 <211> 57

<212> DNA

<213> Secuencia artificial

<220>

30 <221> misc_señal

<222> (0)...(0)

<400> 23

ES 2 609 066 T3

atgattaaat ttctctctgc attaattctt ctactgtca cgacggcggc tcaggct 57

<210> 24

5 <211> 19

<212> PRT

<213> Secuencia artificial

<220>

10 <221> SEÑAL

<222> (0)...(0)

<400> 24

Met Ile Lys Phe Leu Ser Ala Leu Ile Leu Leu Leu Val Thr Thr Ala
1 5 10 15
Ala Gln Ala

15

<210> 25

<211> 57

<212> DNA

20 <213> Secuencia artificial

<220>

<221> misc_señal

<222> (0)...(0)

25

<400> 25

atgaaaaaga ttggctggc gctggctggt ttagtttag cgtttagcgc atcggcg 57

30 <210> 26

<211> 19

<212> PRT

<213> Secuencia artificial

ES 2 609 066 T3

<220>

<221> SEÑAL

<222> (0)...(0)

5 <400> 26

Met	Lys	Lys	Ile	Trp	Leu	Ala	Leu	Ala	Gly	Leu	Val	Leu	Ala	Phe	Ser
1				5					10					15	
Ala	Ser	Ala													

<210> 27

10 <211> 39

<212> DNA

<213> Secuencia artificial

<220>

15 <221> cebador_unión

<222> (0)...(0)

<400> 27

20 ggcggcata tgggtcgga tgatgtggtg gatagcagc 39

<210> 28

<211> 33

<212> DNA

25 <213> Secuencia artificial

<220>

<221> cebador_unión

<222> (0)...(0)

30

<400> 28

gcfgagctcg agttattagc tttgatttc gaa 33

ES 2 609 066 T3

<210> 29

<211> 90

<212> DNA

<213> Secuencia artificial

5

<220>

<221> cebador_unión

<222> (0)...(0)

10 <400> 29

```
ggagcgcata tgattaaatt tctctctgca ttaattcttc tactggtcac gacggcggct 60
caggctgggt cggatgatgt ggtggatagc                                     90
```

<210> 30

15 <211> 24

<212> DNA

<213> Secuencia artificial

<220>

20 <221> cebador_unión

<222> (0)...(0)

<400> 30

25 cagccgcat agttgcacc cgca 24

<210> 31

<211> 1605

<212> DNA

30 <213> C. diphtheriae

<400> 31

ES 2 609 066 T3

ggtgCGgatg atgtggtgga tagcagcaaa tcttttGTga tggaaaactt tagcagctat 60
 catggcacca aaccgggcta tgtggatagc attcagaaag gcatccagaa accgaaaagc 120
 ggcacccagG gcaactatga tgatgattgg aaagaatttt atagcaccga taacaaatat 180
 gatgCGgcgG gttatagcgt ggataacgaa aatccgctgt ctggcaaagc gggcgggtgG 240
 gtgaaagtga cctatccggg cctgacccaa gtgctggccc tgaaagtgga taacgcggaa 300
 accatcaaaa aagaactggg cctgagcctg accgaaccgc tgatggaaca ggtgggcacc 360
 gaagaattta ttaaacgctt tggcgatggc gcgagccgtg tggttctgag cctgccgttt 420
 gcggaaggca gcagcagcgt ggaatatatt aacaactggg aacaggcgaa agccctgagc 480
 gtggaactgg aaattaactt tgaaaccctg ggcaaactgG gccaggatgc gatgtatgaa 540
 tacatggcgc aggcgtgcgc gggcaatcgt gtgcgtcgta gcgtgggcag cagcctgagc 600

tgcattaacc tggattggga cgtcattcgt gataaaacca aaaccaaact cgaaagcctg 660
 aaagaacatg gcccgatcaa aaacaaaatg agcgaaagcc cgaacaaaac cgtgagcgaa 720
 gaaaaagcga aacagtatct ggaagaattt catcagaccg cgctggaaca tccggaactg 780
 agcgaactga aaaccgtgac cggcaccaat ccggtgtttg cgggtgCGaa ctatgCGgcg 840
 tgggCGgtga atgtggcgca ggtgattgat agcgaaaccg cggataacct ggaaaaaac 900
 accgCGgccc tgagcattct gccgggcatt ggcagcgtga tgggcattgc ggatggcgcg 960
 gtgcatcata acaccgaaga aattgtggcg cagagcattg ccctgagcag cctgatggtg 1020
 gcgcaggcga ttccgctggt tggcgaactg gtggatattg gctttgcggc gtacaacttt 1080
 gtggaagca tcatcaacct gtttcaggtg gtgcataaca gctataaccg tccggcgat 1140
 tctccgggtc ataaaaccca gccgtttctg catgatggct atgCGgtgag ctggaacacc 1200
 gtggaagata gcattattcg taccggcttt cagggcgaaa gcggccatga tattaaaatt 1260
 accgCGgaaa acaccccgct gccgattgog ggtgttctgc tgccgacat tccgggcaaa 1320
 ctggatgtga acaaaagcaa aaccatatt agcgtgaacg gtcgtaaaat tcgtatgCGt 1380
 tgccgtgcga ttgatggcga tgtgacctt tgcCGtccga aaagcccggt gtatgtgggc 1440
 aacggcgtgc acgcgaacct gcatgtggcg tttcatcgta gcagcagcga aaaaatccat 1500
 agcaacgaaa ttagcagcga tagcattggc gtgctgggct atcagaaaac cgtggacat 1560
 accaaagtga actctaaact gagcctgttc ttcgaaatca aaagc 1605

<210> 32

<211> 535

5 <212> PRT

<213> C. diphtheriae

<400> 32

ES 2 609 066 T3

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

ES 2 609 066 T3

```

                325                330                335
Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp
                340                345                350
Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe
                355                360                365
Gln Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly His
                370                375                380
Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr
385                390                395                400
Val Glu Asp Ser Ile Ile Arg Thr Gly Phe Gln Gly Glu Ser Gly His
                405                410                415
Asp Ile Lys Ile Thr Ala Glu Asn Thr Pro Leu Pro Ile Ala Gly Val
                420                425                430
Leu Leu Pro Thr Ile Pro Gly Lys Leu Asp Val Asn Lys Ser Lys Thr
                435                440                445
His Ile Ser Val Asn Gly Arg Lys Ile Arg Met Arg Cys Arg Ala Ile
                450                455                460
Asp Gly Asp Val Thr Phe Cys Arg Pro Lys Ser Pro Val Tyr Val Gly
465                470                475                480
Asn Gly Val His Ala Asn Leu His Val Ala Phe His Arg Ser Ser Ser
                485                490                495
Glu Lys Ile His Ser Asn Glu Ile Ser Ser Asp Ser Ile Gly Val Leu
                500                505                510
Gly Tyr Gln Lys Thr Val Asp His Thr Lys Val Asn Ser Lys Leu Ser
                515                520                525
Leu Phe Phe Glu Ile Lys Ser
                530                535

```

<210> 33

<211> 54

5 <212> PRT

<213> Secuencia artificial

<400> 33

```

Met Lys Phe Ser Lys Lys Tyr Ile Ala Ala Gly Ser Ala Val Ile Val
 1                5                10                15
Ser Leu Ser Leu Cys Ala Tyr Ala Gly Ala Asp Asp Val Val Asp Ser
                20                25                30
Ser Lys Ser Phe Val Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys
                35                40                45
Pro Gly Tyr Val Asp Ser
                50

```

10

<210> 34

<211> 55

<212> PRT

15 <213> Secuencia artificial

ES 2 609 066 T3

<400> 34

```

Met Lys Met Asn Lys Lys Val Leu Leu Thr Ser Thr Met Ala Ala Ser
 1      5      10      15
Leu Leu Ser Val Ala Ser Val Gln Ala Gly Ala Asp Asp Val Val Asp
      20      25      30
Ser Ser Lys Ser Phe Val Met Glu Asn Phe Ser Ser Tyr His Gly Thr
      35      40      45
Lys Pro Gly Tyr Val Asp Ser
      50      55
    
```

5 <210> 35

<211> 51

<212> PRT

<213> Secuencia artificial

10 <400> 35

```

Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
 1      5      10      15
Thr Val Ala Gln Ala Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser
      20      25      30
Phe Val Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr
      35      40      45
Val Asp Ser
      50
    
```

<210> 36

15 <211> 49

<212> PRT

<213> Secuencia artificial

<400> 36

20

```

Met Lys Lys Ala Leu Ala Thr Leu Ile Ala Leu Ala Leu Pro Ala Ala
 1      5      10      15
Ala Leu Ala Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val
      20      25      30
Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp
      35      40      45
Ser
    
```

<210> 37

ES 2 609 066 T3

<211> 48

<212> PRT

<213> Secuencia artificial

5 <400> 37

```

Met Arg Val Leu Leu Phe Leu Leu Leu Ser Leu Phe Met Leu Pro Ala
 1           5           10           15
Phe Ser Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met
      20           25           30
Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp Ser
      35           40           45
    
```

<210> 38

10 <211> 53

<212> PRT

<213> Secuencia artificial

<400> 38

15

```

Met Met Thr Lys Ile Lys Leu Leu Met Leu Ile Ile Phe Tyr Leu Ile
 1           5           10           15
Ile Ser Ala Ser Ala His Ala Gly Ala Asp Asp Val Val Asp Ser Ser
      20           25           30
Lys Ser Phe Val Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro
      35           40           45
      Gly Tyr Val Asp Ser
      50
    
```

<210> 39

<211> 55

20 <212> PRT

<213> Secuencia artificial

<400> 39

ES 2 609 066 T3

Met Met Lys His Met Arg Ile Trp Ala Val Leu Ala Ser Phe Leu Val
 1 5 10 15
 Phe Phe Tyr Ile Pro Gln Ser Tyr Ala Gly Ala Asp Asp Val Val Asp
 20 25 30
 Ser Ser Lys Ser Phe Val Met Glu Asn Phe Ser Ser Tyr His Gly Thr
 35 40 45
 Lys Pro Gly Tyr Val Asp Ser
 50 55

<210> 40

<211> 48

5 <212> PRT

<213> Secuencia artificial

<400> 40

Met Arg Phe Leu Leu Gly Val Leu Met Leu Met Ile Ser Gly Ser Ala
 1 5 10 15
 Leu Ala Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met
 20 25 30
 Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp Ser
 35 40 45

10

<210> 41

<211> 53

<212> PRT

15 <213> Secuencia artificial

<400> 41

Met Ser Lys Arg Thr Phe Ala Val Ile Leu Thr Leu Leu Cys Ser Phe
 1 5 10 15
 Cys Ile Gly Gln Ala Leu Ala Gly Ala Asp Asp Val Val Asp Ser Ser
 20 25 30
 Lys Ser Phe Val Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro
 35 40 45
 Gly Tyr Val Asp Ser
 50

20

<210> 42

<211> 52

<212> PRT

<213> Secuencia artificial

ES 2 609 066 T3

<400> 42

```

Met Met Lys Gln Ala Leu Arg Val Ala Phe Gly Phe Leu Ile Leu Trp
 1      5      10
Ala Ser Val Leu His Ala Gly Ala Asp Asp Val Val Asp Ser Ser Lys
      20      25      30
Ser Phe Val Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly
      35      40      45
Tyr Val Asp Ser
    50
    
```

5

<210> 43

<211> 52

<212> PRT

<213> Secuencia artificial

10

<400> 43

```

Met Leu Ser Thr Leu Arg Arg Thr Leu Phe Ala Leu Leu Ala Cys Ala
 1      5      10      15
Ser Phe Ile Val His Ala Gly Ala Asp Asp Val Val Asp Ser Ser Lys
      20      25      30
Ser Phe Val Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly
      35      40      45
Tyr Val Asp Ser
    50
    
```

15 <210> 44

<211> 49

<212> PRT

<213> Secuencia artificial

20 <400> 44

```

Met Ile Lys Phe Leu Ser Ala Leu Ile Leu Leu Leu Val Thr Thr Ala
 1      5      10      15
Ala Gln Ala Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val
      20      25      30
Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp
      35      40      45
Ser
    
```

ES 2 609 066 T3

<210> 45

<211> 49

<212> PRT

<213> Secuencia artificial

5

<400> 45

```
Met Lys Lys Ile Trp Leu Ala Leu Ala Gly Leu Val Leu Ala Phe Ser
 1          5          10          15
Ala Ser Ala Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val
 20          25          30
Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp
 35          40          45
Ser
```

REIVINDICACIONES

1. Un procedimiento de producción industrial de una toxina bacteriana expresada periplásmicamente por:
 - A. en una fase de alimentación discontinua, cultivar un cultivo de una célula huésped gram-negativa que comprende un polinucleótido que comprende una parte de toxina 3' que codifica una toxina bacteriana y una parte de señal 5' que codifica un polipéptido capaz de dirigir el transporte de dicha toxina bacteriana al periplasma de la célula huésped, en el que el pH es de 6,0-7,0; y
 - B. inducir expresión de dicha toxina bacteriana a un pH que es y se mantiene a, un pH de 6,5-8,5, en el que dicho pH es al menos 0,5 unidades de pH superior que el pH en la etapa A; y la velocidad de alimentación del sustrato se mantiene a entre el 20 % y el 80 % de la velocidad de alimentación del sustrato de la etapa A;
- 5 en el que la toxina bacteriana tiene al menos el 90 % de identidad de secuencia con la SEC ID N.º: 32 y el procedimiento se lleva a cabo en un fermentador que contiene desde 10-2000 litros de cultivo.
- 10 2. El procedimiento de la reivindicación 1, en el que la etapa A se lleva a cabo a una temperatura de 20-40 °C y la etapa B se lleva a cabo a una temperatura de 20-28 °C, en el que la temperatura de la etapa A es más alta que la de la etapa B.
- 15 3. El procedimiento de la reivindicación 1, en el que el pH se mantiene usando un tampón del grupo que consiste en tampón fosfato, tampón Tris y tampón de histidina opcionalmente en el que el tampón está a una concentración de 10-200 mM, 50-100 mM, 10-50 mM, 100-200 mM o 50-150 nM.
4. El procedimiento de una cualquiera de las reivindicaciones 1-3, en el que un incremento en pH en la etapa B se logra por adición de base por ejemplo hidróxido de sodio o amoníaco.
- 20 5. El procedimiento de una cualquiera de las reivindicaciones 1-4, en el que la expresión se induce en la etapa B por la adición de una cantidad suficiente de IPTG.
6. El procedimiento de una cualquiera de las reivindicaciones 1-5, en el que el nivel de oxígeno disuelto es mayor que el 5 % y/o el nivel de oxígeno disuelto es del 15-25 %.
- 25 7. El procedimiento de una cualquiera de las reivindicaciones 1-6, que comprende una etapa adicional c) de cosechar la pasta celular del cultivo y/o una etapa adicional de purificar la toxina bacteriana, por ejemplo como una toxina bacteriana madura.
8. Un procedimiento de fabricación de un conjugado que comprende las etapas de i) fabricar una toxina bacteriana usando el procedimiento de una cualquiera de las reivindicaciones 1-7 e ii) conjugar la toxina bacteriana de la etapa i) a un antígeno.
- 30 9. El procedimiento de la reivindicación 8, en el que el antígeno es un sacárido bacteriano opcionalmente un sacárido capsular de *S. pneumoniae*, *H. influenzae*, *N. meningitidis* o *S. aureus*.
10. Un procedimiento de elaboración de una vacuna que comprende las etapas de:
 - A. fabricar una toxina bacteriana o un conjugado usando el procedimiento de una cualquiera de las reivindicaciones 1-9 y;
 - B. mezclar la toxina bacteriana o el conjugado de la misma con un excipiente farmacéuticamente aceptable.
- 35 11. El procedimiento de la reivindicación 10 en el que la toxina bacteriana se mezcla con antígenos adicionales en la etapa B y/o en el que la toxina bacteriana se mezcla con un coadyuvante en la etapa B.

Figura 1

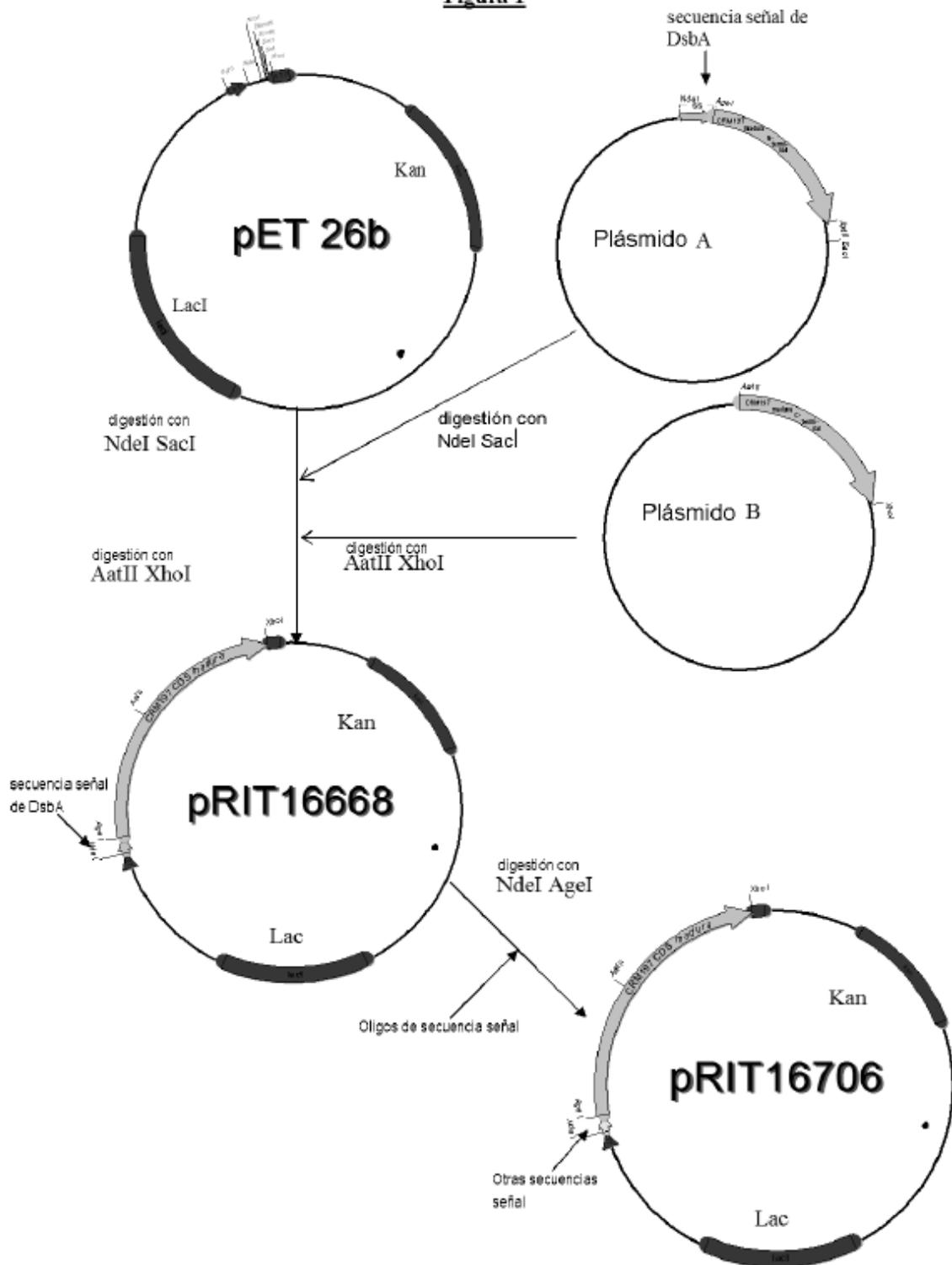


Figura 2

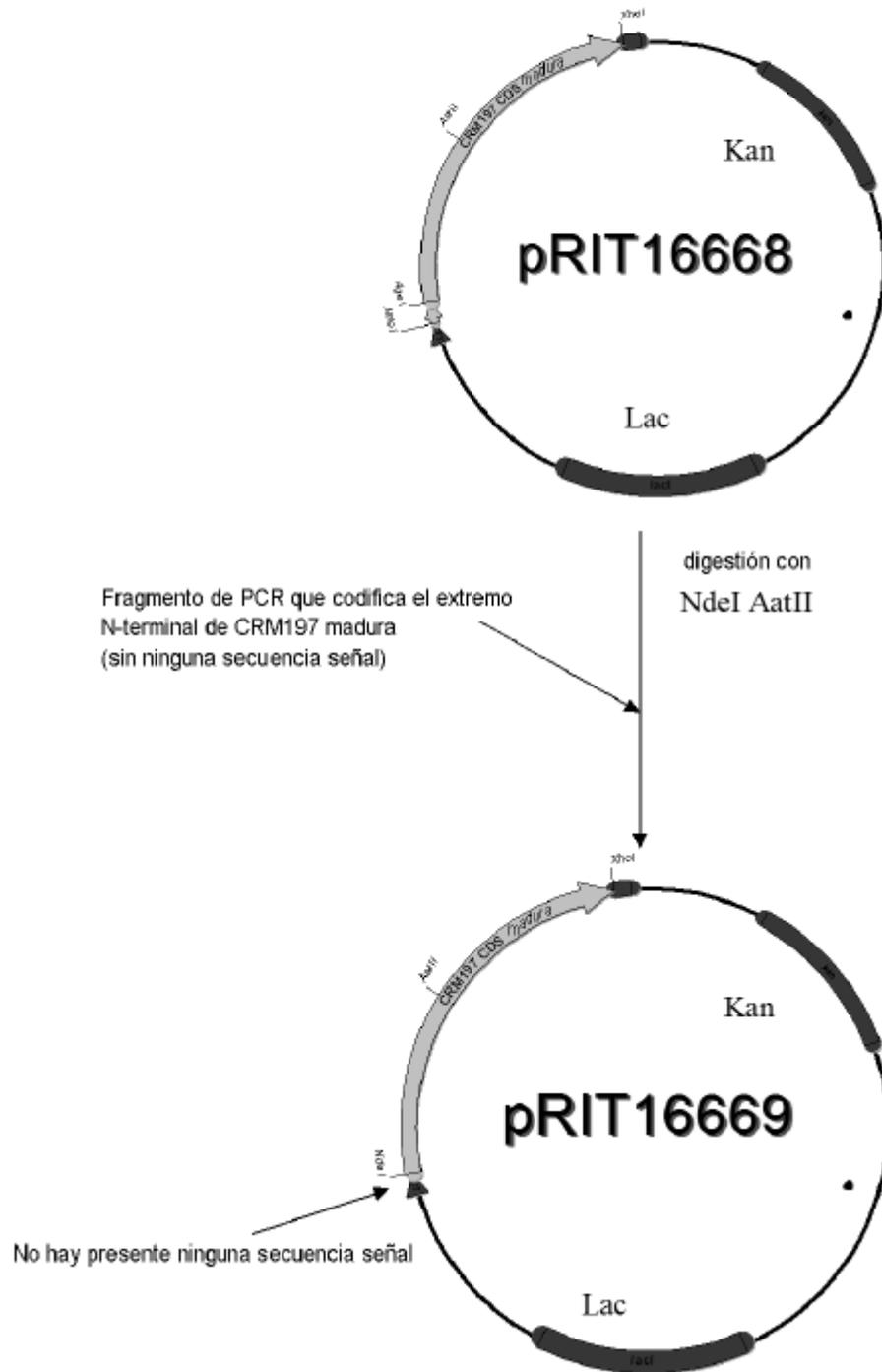


Figura 3

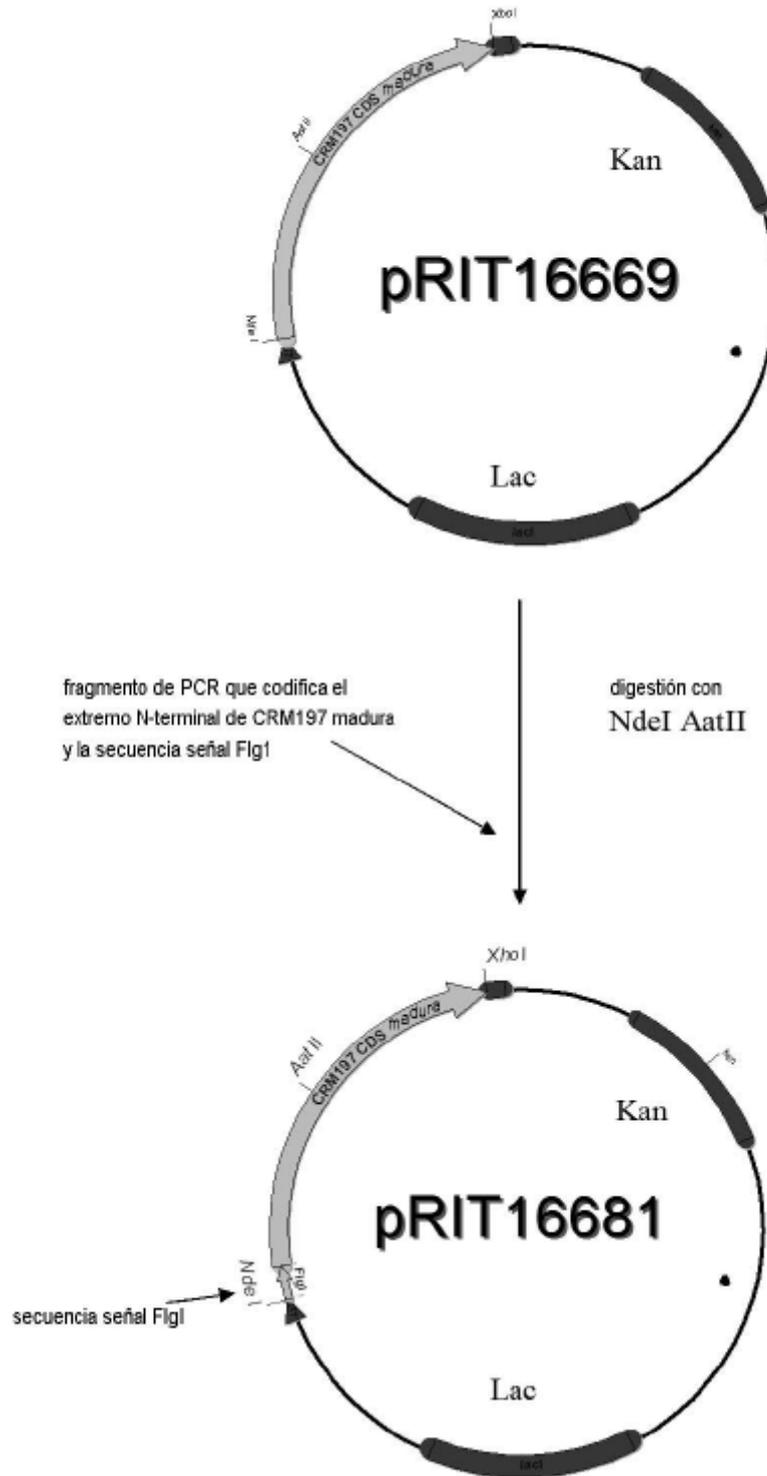


Figura 4

Figura 2: geles que muestran la inducción de CRM197 3 horas a 30 °C en cepa diferente, medio sin tamponar

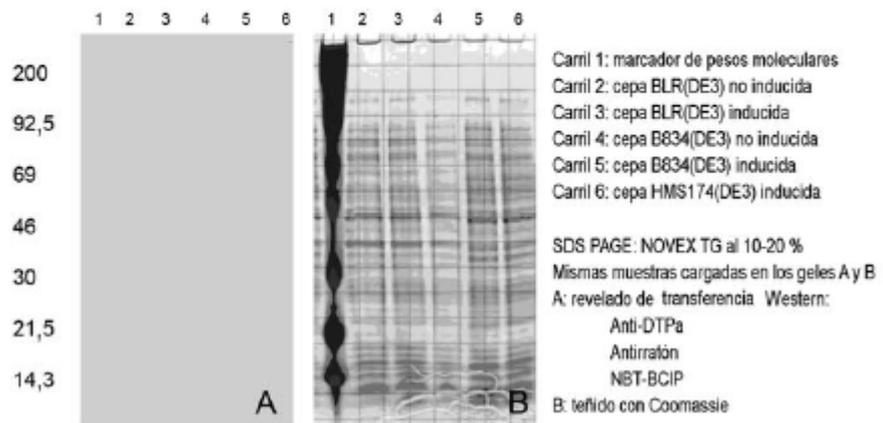
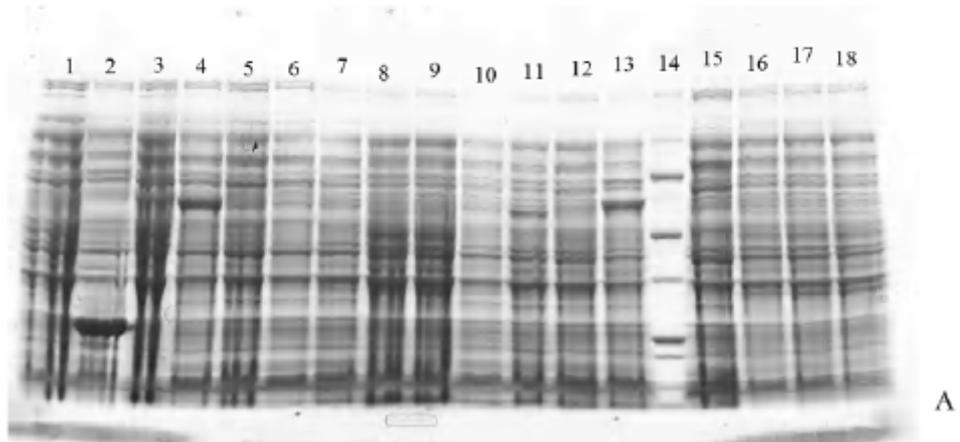
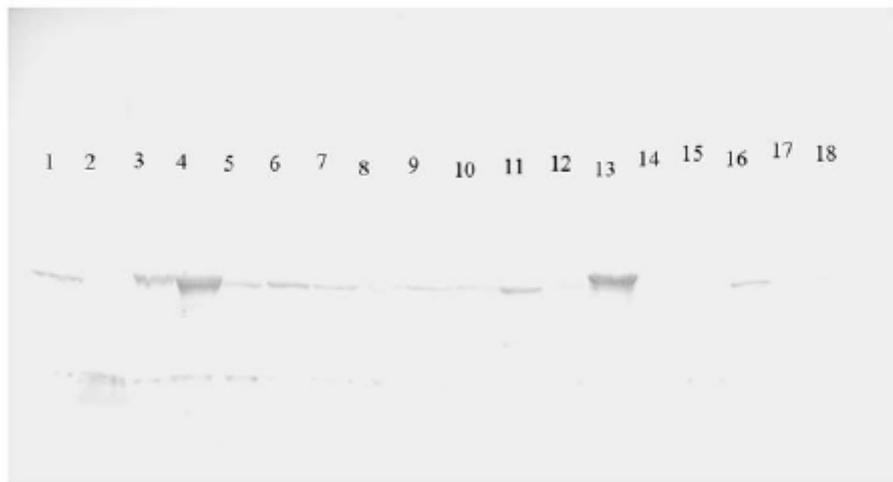


Figura 5
Expresión de CRM197 en la célula bacteriana (A) y en el medio de cultivo (B)



A



B

Figura 6

Geles que muestran inducción de CRM197 3 horas a 30°C en diferentes cepas y solubilidad de la proteína expresada

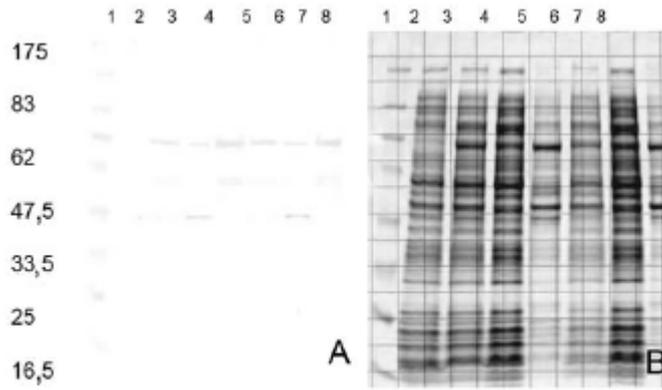


Figura 7

Figura 7: nivel de expresión de CRM197 en función de pH y temperatura

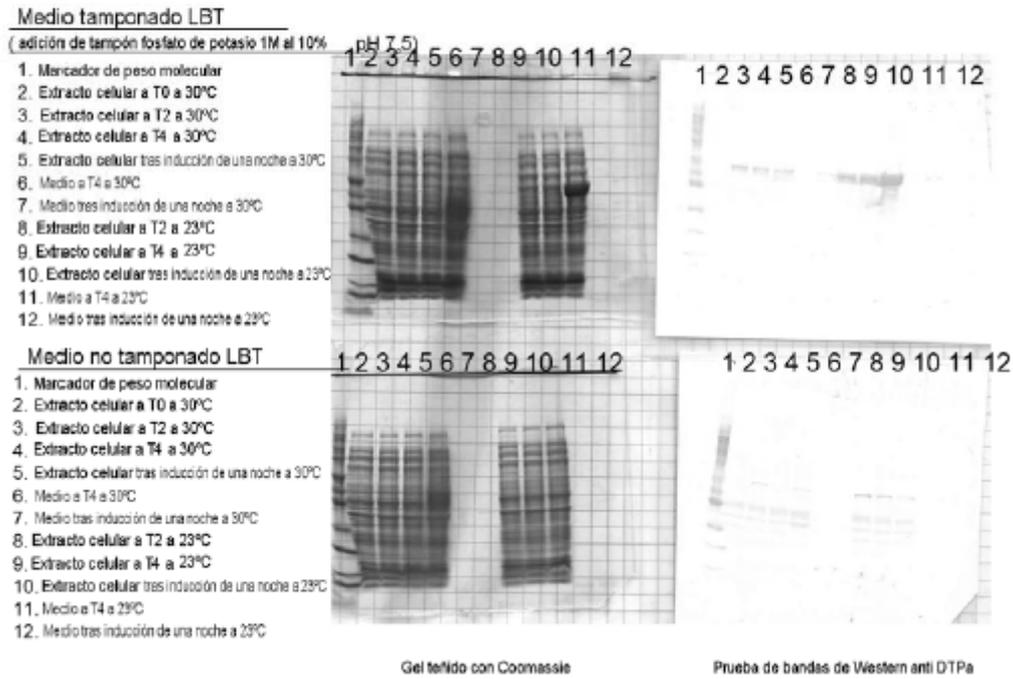


Figura 8

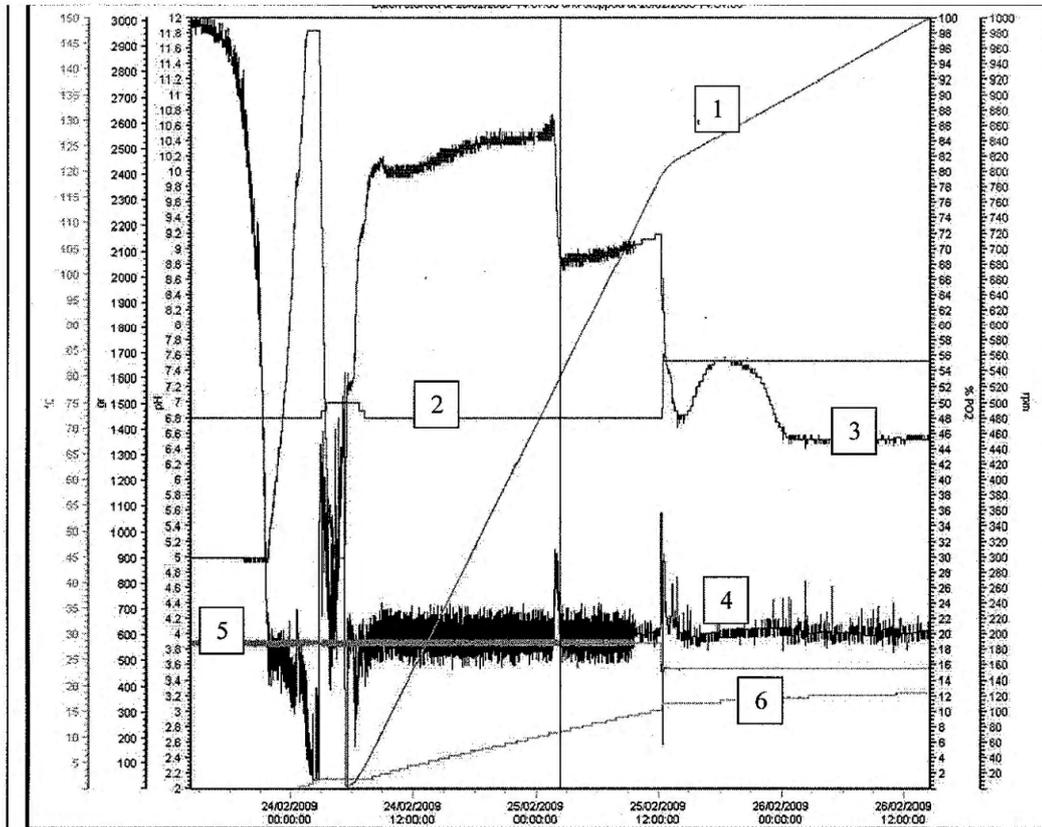


Figura 9

SEC ID N.º 1: secuencia de nucleótidos de la secuencia señal PhtE

ATGAAATTTAGTAAAAAATATATAGCAGCTGGATCAGCTGTTATCGTATCCTTGAGTCTATGTGCCTATGCA

5

SEC ID N.º 2: secuencia de aminoácidos de la secuencia señal PhtE

MKFSKKYIAAGSAVIVSLSLCAYA

SEC ID N.º 3: secuencia de nucleótidos de la secuencia señal SipA

10 ATGAAAATGAATAAAAAGGTACTATTGACATCGACAATGGCAGCTTCGCTATTATCAGTCGCAAGTGTTCAAGCA

SEC ID N.º 4: secuencia de aminoácidos de la secuencia señal SipA

MKMNKVLLTSTMAASLLSVASVQA

15 **SEC ID N.º 5: secuencia de nucleótidos de la secuencia señal OmpA**

ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTTCGCTACCGTAGCGCAGGCC

SEC ID N.º 6: secuencia de aminoácidos de la secuencia señal OmpA

MKKTAIAlAVALAGFATVAQA

20

SEC ID N.º 7: secuencia de nucleótidos de la secuencia señal NspA

ATGAAAAAAGCACTTGCCCACTGATTGCCCTCGCTCTCCCGGCCGCCCACTGGCG

SEC ID N.º 8: secuencia de aminoácidos de la secuencia señal NspA

25 MKKALATLIALALPAAALA

SEC ID N.º 9: secuencia de nucleótidos de la secuencia señal TorT

ATGCGCGTACTGCTATTTTTACTTCTTTCCCTTTTCATGTTGCCGGCATTTCG

30 **SEC ID N.º 10: secuencia de aminoácidos de la secuencia señal TorT**

MRVLLFLLLSLFLPAFS

SEC ID N.º 11: secuencia de nucleótidos de la secuencia señal SfmC

ATGATGACTAAAATAAAGTTATTGATGCTCATTATATTTTATTTAATCATTTCGGCCAGCGCCCATGCT

35

SEC ID N.º 12: secuencia de aminoácidos de la secuencia señal SfmC

MMTKIKLLMLIIFYLIISASAHA

SEC ID N.º 13: secuencia de nucleótidos de la secuencia señal FocC

ATGATGAAGCACATGCGTATATGGGCCGTTCTGGCATCATTTTTAGTCTTTTTTATATTCCGCAGAGCTATGCC

5 **SEC ID N.º 14: secuencia de aminoácidos de la secuencia señal FocC**

MMKHMRIWAVLASFLVFFYIPQSYA

SEC ID N.º 15: secuencia de nucleótidos de la secuencia señal CcmH

ATGAGGTTTTTATTGGGCGTGCTGATGCTGATGATCTCCGGCTCAGCGCTGGCG

10

SEC ID N.º 16: secuencia de aminoácidos de la secuencia señal CcmH

MRFLLGVLMLMISGSALA

SEC ID N.º 17: secuencia de nucleótidos de la secuencia señal Yral

15 ATGTCAAACGAACATTTCGCGGTGATATTAACCTTGTTGTGTAGCTTCTGTATTGGCCAGGCGCTTGCA

SEC ID N.º 18: secuencia de aminoácidos de la secuencia señal Yral

MSKRTEFAVILTLLCSFCIGQALA

20 **SEC ID N.º 19: secuencia de nucleótidos de la secuencia señal ToIB**

ATGATGAAGCAGGCATTACGAGTAGCATTTGGTTTTCTCATACTGTGGGCATCAGTTCTGCATGCT

SEC ID N.º 20: secuencia de aminoácidos de la secuencia señal ToIB

MMKQALRVAFGFLILWASVLHA

25

SEC ID N.º 21: secuencia de nucleótidos de la secuencia señal Nika

ATGCTCTCCCACTCCGCCGCACTCTATTTGCGCTGCTGGCTTGTGCGTCTTTTATCGTCCATGCC

SEC ID N.º 22: secuencia de aminoácidos de la secuencia señal Nika

30 MLSTLRRTLFALLACASFIVHA

SEC ID N.º 23: secuencia de nucleótidos de la secuencia señal Flgl

ATGATTAATTTCTCTCTGCATTAATTCTTCTACTGGTCACGACGGCGGCTCAGGCT

35 **SEC ID N.º 24: secuencia de aminoácidos de la secuencia señal Flgl**

MIKFLSALILLLVTTAAQA

SEC ID N.º 25: secuencia de nucleótidos de la secuencia señal DsbA

ATGAAAAAGATTTGGCTGGCGCTGGCTGGTTTAGTTTTAGCGTTTAGCGCATCGGCG

SEC ID N.º 26: secuencia de aminoácidos de la secuencia señal DsbA

5 MKKIWLALAGLVLAFSASA

SEC ID N.º 27: secuencia de DNA del cebador MDSCRM1

5' GCGCGGCATATGGGTGCGGATGATGTGGTGGATAGCAGC 3'

10 **SEC ID N.º:28: secuencia de DNA del cebador MDSCRM2**

5' GCGGAGCTCGAGTTATTAGCTTTTGATTTCGAA 3'

SEC ID N.º:29: secuencia de DNA del cebador

5' gga gcg cat atg att aaa ttt ctc tct gca tta att ctt cta ctg gtc acg acg gcg gct cag gct ggt gcg gat gat gtg gtg gat agc 3'

15

SEC ID N.º:30: secuencia de DNA del cebador MDSCRM906NC

cebador MDSCRM906NC

5' CAC GCC GCA TAG TTC GCA CCC GCA 3'

20 **SEC ID N.º 31: secuencia de nucleótidos de la región codificante de CRM197 optimizada de codones madura**

GGTGCGGATGATGTGGTGGATAGCAGCAAATCTTTTGTGATGGAAAACCTTTAGCAGCTATCATGGCACCAAACCGGGCTA
TGTGGATAGCATTTCAGAAAGGCATCCAGAAACCGAAAAGCGGCACCCAGGGCAACTATGATGATGATTGGAAAAGAATTTT
ATAGCACCGATAACAAATATGATGCGGCGGGTTATAGCGTGGATAACGAAAATCCGCTGTCTGGCAAAGCGGGCGGTGTG
25 GTGAAAGTGACCTATCCGGGCTGACCAAAGTGCTGGCCCTGAAAGTGGATAACGCGGAAACCATCAAAAAAGAACTGGG
CCTGAGCCTGACCGAACCGCTGATGGAAACAGGTGGGCACCGAAGAATTTATTAAACGCTTTGGCGATGGCGGAGCCGTG
TGGTTCTGAGCCTGCCGTTTGCAGGAGCAGCAGCGTGGAAATATATTAACAACCTGGGAACAGGCGAAAAGCCCTGAGC
GTGGAACCTGAAATTAACCTTTGAAACCCGTGGCAAACGTGGCCAGGATGCGATGTATGAATACATGGCGCAGGCGTGCGC
GGGCAATCGTGTGCGTCGTAGCGTGGGCAGCAGCCTGAGCTGCATTAACCTGGATTGGGACGTCATTCGTGATAAAACCA
30 AAACCAAAATCGAAAGCCTGAAAGAACATGGCCCGATCAAAAACAAAATGAGCGAAAGCCCGAACAAAACCGTGAGCGAA
GAAAAAGCGAAACAGTATCTGGAAGAATTTTCATCAGACCGCGCTGGAACATCCGGAACCTGAGCGAACCTGAAACCGTGAC
CGGCACCAATCCGGTGTGTTGCGGGTGGAACTATGCGGCGTGGGCGGTGAATGTGGCGCAGGTGATTGATAGCGAAACCG
CGGATAACCTGGAAAAAACCCGCGCCCTGAGCATTTCTGCCGGGCATTTGGCAGCGTATGGGCATTTGCGGATGGCGCG
GTGCATCATAACACCGAAGAAATTTGTGGCGCAGAGCATTTGCCCTGAGCAGCCTGATGGTGGCGCAGGCGATTCCGCTGGT
35 TGGCGAACTGGTGGATATTGGCTTTGCGGCGTACAACCTTTGTGGAAAGCATCATCAACCTGTTTCAGGTGGTGCATAACA
GCTATAACCGTCCGGCGTATTCTCCGGGTGATAAAACCCAGCCGTTTCTGCATGATGGCTATGCGGTGAGCTGGAACACC
GTGGAAGATAGCATTATTCGTACCGGCTTTTCAGGGCGAAAAGCGGCCATGATATTAATAATTACCGCGAAAAACCCCGCT

GCCGATTGCGGGTGTCTGCTGCCGACCATTCCGGGCAAACCTGGATGTGAACAAAAGCAAAAACCCATATTAGCGTGAACG
GTCGTAAAAATTCGTATGCGTTGCCGTGCGATTGATGGCGATGTGACCTTTTGGCGTCCGAAAAGCCCGGTGTATGTGGGC
AACGGCGTGACGCGAACCTGCATGTGGCGTTTCATCGTAGCAGCAGCGAAAAAATCCATAGCAACGAAATTAGCAGCGA
TAGCATTGGCGTGTGGGCTATCAGAAAACCGTGGACCATAACAAAGTGAACCTCTAAACTGAGCCTGTTCTTCGAAATCA
5 AAAGC

SEC ID N.º 32: secuencia de aminoácidos de CRM197 madura

GADDVVDSSKSFVMENFSSYHGTPGYVDS
IQKGIQPKSGTQGNYYDDWKEFYSTDNKY
10 DAAGYSVDNENPLSGKAGGVVKVTPGLTK
VLALKVDNAETIKKELGLSLTEPLMEQVGT
EEFIKRFDGASRVVLSLPAEGSSSVEYI
NNWEQAKALSVELEINFETRGRGQDAMYE
YMAQACAGNRVRRSVGSSLSCINLDWDVIR
15 DKTKTKIESLKEHGPIKNKMSSEPNKTVSE
EKAKQYLEEFHQTALEHPELSELKTVTGTN
PVFAGANYAAWAVNVAQVIDSETADNLEKT
TAALSILPGIGSVMGIADGAVHHNTEEIVA
QSIALSSLMVAQAIPLVGELVDIGFAAYNF
20 VESIINLFQVVHNSYNRPAYSPGHKTQPFL
HDGYAVSWNTVEDSIIRTFQGESGHDIKI
TAENTPLPIAGVLLPTIPGKLDVNKSKTHI
SVNGRKIRMRCRAIDGDVTFCRPKSPVYVG
NGVHANLHVAFHRSSSEKIHSNEISSDSIG
25 VLG YQKTVDHTKVNKLSLFFFEIKS.

SECUENCIA N.º 33: secuencia de aminoácidos de la secuencia señal PhtE y los primeros 30 aminoácidos de CRM197

MKFSKKYIAAGSAVIVLSLCAAYAGADDVVDSSKSFVMENFSSYHGTPGYVDS
30

SEC ID N.º 34: secuencia de aminoácidos de la secuencia señal Sip y los primeros 30 aminoácidos de CRM197

MKMNKKVLLTSTMAASLLSVASVQAGADDVVDSSKSFVMENFSSYHGTPGYVDS

SEC ID N.º 35: secuencia de aminoácidos de la secuencia señal OmpA y los primeros 30 aminoácidos de CRM197

MKKTAIAIAVALAGFATVAQAGADDVVDSSKSFVMENFSSYHGTPGYVDS

SEC ID N.º 36: secuencia de aminoácidos de la secuencia señal NspA y los primeros 30 aminoácidos de CRM197

MKKALATLIALALPAAALAGADDVVDSSKSFVMENFSSYHGTKPGYVDS

5 **SEC ID N.º 37: secuencia de aminoácidos de la secuencia señal TorT y los primeros 30 aminoácidos de CRM197**

MRVLLFLLLSLFLMLPAFSGADDVVDSSKSFVMENFSSYHGTKPGYVDS

10 **SEC ID N.º 38: secuencia de aminoácidos de la secuencia señal SfmC y los primeros 30 aminoácidos de CRM197**

MMTKIKLLMLIIFYLIISASAHAGADDVVDSSKSFVMENFSSYHGTKPGYVDS

SEC ID N.º 39: secuencia de aminoácidos de la secuencia señal FocC y los primeros 30 aminoácidos de CRM197

15 MMKHMRIWAVLASFLVFFYIPQSYAGADDVVDSSKSFVMENFSSYHGTKPGYVDS

SEC ID N.º 40: secuencia de aminoácidos de la secuencia señal CcmH y los primeros 30 aminoácidos de CRM197

MRFLGLVLMISGSALAGADDVVDSSKSFVMENFSSYHGTKPGYVDS

20 **SEC ID N.º 41: secuencia de aminoácidos de la secuencia señal Yral y los 30 primeros aminoácidos de CRM197**

MSKRTFAVILTLLCSFCIGQALAGADDVVDSSKSFVMENFSSYHGTKPGYVDS

25 **SEC ID N.º 42: secuencia de aminoácidos de la secuencia señal TolB y los 30 primeros aminoácidos de CRM197**

MMKQALRVAFGFLILWASVLHAGADDVVDSSKSFVMENFSSYHGTKPGYVDS

30 **SEC ID N.º 43: secuencia de aminoácidos de la secuencia señal NikA y los 30 primeros aminoácidos de CRM197**

MLSTLRRTLFALLACASFIVHAGADDVVDSSKSFVMENFSSYHGTKPGYVDS

SEC ID N.º 44: secuencia de aminoácidos de la secuencia señal FlgI y los 30 primeros aminoácidos de CRM197

35 MIKFLSALILLVTTAAQAGADDVVDSSKSFVMENFSSYHGTKPGYVDS

SEC ID N.º 45: secuencia de aminoácidos de la secuencia señal DsbA y los 30 primeros aminoácidos de CRM197

MKKIWLALAGLVLAFSASAGADDVVDSSKSFVMENFSSYHGTKPGYVDS

40

Figura 10

