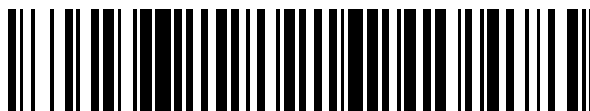


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 750 023**

51 Int. Cl.:

C07K 14/34 (2006.01)

C12N 15/62 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.10.2010** **E 16192258 (8)**

97 Fecha y número de publicación de la concesión europea: **21.08.2019** **EP 3170837**

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:

24.03.2020

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 2 750 023 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

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. La invención también proporciona nuevos polinucleótidos y polipéptidos que se pueden usar o producir durante el procedimiento de la invención.

La toxina diftérica es una exotoxina proteica producida por la bacteria *Corynebacterium diphtheriae*. 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.

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.

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. Alternativamente, se han descrito varias toxinas mutadas de difteria que tienen toxicidad reducida (documentos US4709017, US4950740).

La CRM197 es una forma no tóxica de la toxina diftérica pero es inmunológicamente indistinguible de la toxina diftérica. La 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 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).

La 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).

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.).

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).

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.

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 periplasmático y esta secreción está asociada con la viabilidad reducida de las células hospedadoras (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 periplasmática, esto puede incrementar expresión de toxóides de difteria (Bishai y col.).

De acuerdo con ello, la presente solicitud proporciona un procedimiento mejorado para fabricar una toxina bacteriana por expresión periplasmática comprendiendo las etapas de a) cultivar un cultivo de la célula hospedadora 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 de manera periplasmática. La presente solicitud también proporciona polinucleótidos que se usan en el procedimiento de la invención.

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;
- (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

En un primer aspecto de la invención se proporciona un polinucleótido que comprende una parte de secuencia señal 5' y una parte de toxina 3', en el que;

- (a) La parte de señal 5' codifica un polipéptido capaz de dirigir el transporte de una proteína heteróloga al periplasma bacteriano; teniendo el polipéptido una secuencia de aminoácidos de (i) SEQ ID NO: 24, (ii) variantes de SEQ ID NO: 24, en el que dichas variantes contienen 1, 2 o 3 mutaciones de sustitución, adición o delección de aminoácidos, en el que una adición de aminoácidos es la adición de un aminoácido, una sustitución de aminoácidos es el reemplazo de un aminoácido por otro aminoácido y una delección de aminoácidos es la eliminación de un aminoácido, o (iii) fragmentos de al menos 10 aminoácidos de SEQ ID NO: 24; y
- (b) La parte de toxina 3' codifica un polipéptido que tiene una secuencia de aminoácidos al menos el 90 % idéntica a la SEQ ID NO: 32.

En un segundo aspecto de la invención, se proporciona un vector que comprende el polinucleótido de la invención unido a un promotor inducible.

En un tercer aspecto de la invención, se proporciona una célula hospedadora que comprende el polinucleótido o el vector de la invención.

En un cuarto aspecto de la invención, se proporciona un polipéptido codificado por el polinucleótido de la invención.

En un quinto aspecto de la invención, se proporciona un procedimiento de fabricación de una toxina bacteriana que comprende las etapas de a) cultivar un cultivo de la célula hospedadora bacteriana de la invención y b) inducir la expresión del polipéptido de manera que un toxoide bacteriano se exprese de manera periplasmática.

En un sexto 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) conjuguar la toxina bacteriana de la etapa a) a un antígeno.

En un séptimo aspecto de la invención se proporciona un procedimiento de elaboración de una vacuna que comprende las etapas de a) fabricar 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

Figura 1 – Visión general del procedimiento de clonación para producir las construcciones de CRM197-secuencia señal. Una sección de ADN 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 SEQ ID NO: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 ADN 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 ADN 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.

Figura 3 – Visión general del procedimiento de clonación para producir las construcciones que contienen la secuencia señal Flgl. Esto se llevó a cabo insertando un fragmento de PCR que codifica la CRM197 y la secuencia señal Flgl (SEQ ID NO:23) dentro de pRIT16669 usando digestión con NdeI y AatII seguida por ligación. Esto

produjo el plásmido pRIT16681.

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

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	SEQ ID NO: de nucleótidos	SEQ ID NO: 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
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 FlgI 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).

5 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).

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

Figura 10 – Representación de la producción de CRM197 en las fracciones periplasmáticas y las fracciones asociadas a células como una función de la velocidad de alimentación y pH durante la inducción, para el cultivo llevado a cabo a pH 6,8. El panel izquierdo muestra la producción de CRM197 periplasmática. El panel derecho describe la producción de CRM197 asociada a células.

15 **Descripción detallada de la invención**

Definiciones

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

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

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 transposición integrada, o debido a edición de ARN o reorganización de ADN 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.

30 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.

35 Una "sustitución de aminoácido" es el reemplazo de un aminoácido por 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.

40 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.

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

"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 "immunógeno" 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 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

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 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 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 la comparación de secuencias de polipéptidos incluyen los siguientes:

Algoritmo: Needleman y Wunsch, J. Mol Biol. 48: 443-453 (1970)
Matriz de comparación: BLOSSUM62 de Henikoff y Henikoff, Proc. Natl. Acad. Sci. EE.UU. 89: 10915-10919 (1992)
Penalización por Hueco: 8
Penalización por Longitud de Hueco: 2

Un programa útil con estos parámetros está disponible públicamente como el programa de "huecos" de 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 la 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 por Hueco: 50
Penalización por Longitud de Hueco: 3
Disponible como: el programa de "huecos" de Genetics Computer Group, Madison WI. Estos son los parámetros por defecto para las 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.

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

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

Los autores de la presente memoria pretenden que los términos "comprendiendo", "comprender" y "comprende" sean ser opcionalmente sustituibles por las expresiones "constituido por", "consiste en" y "consta de", 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.

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 la invención

Un aspecto de la invención se refiere a polinucleótidos que codifican toxinas que se expresan de manera periplasmática.

En general la presencia de una secuencia señal en la proteína facilita el transporte de la proteína dentro del periplasma (hospedadores procariotas) o la secreción de la proteína (hospedadores eucariotas). En los hospedadores procariotas la secuencia señal dirige a la proteína nascente a través de la membrana interna dentro del espacio periplasmático; 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*.

Un ensayo para probar si una secuencia señal es capaz de dirigir la expresión periplasmática se puede llevar a cabo usando elementos comunicadores. Por ejemplo una secuencia señal periplasmática se puede ajustar corriente arriba de un gen para una proteína fluorescente verde, esta proteína se puede expresar en una célula hospedadora de la invención. Se puede usar un microscopio para juzgar los niveles comparativos de la proteína fluorescente verde en el citoplasma y el periplasma.

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 SEQ ID NO: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-traduccion, mientras que las SEC ID NO:6 y 8 codifican péptidos que son capaces de dirigir una proteína expresada al periplasma por transporte postraduccion.

En una realización se proporciona 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 capaz de dirigir el transporte de una proteína heteróloga al periplasma bacteriano; teniendo el polipéptido una secuencia de aminoácidos de (i) SEQ ID NO: 24, (ii) variantes de SEQ ID NO: 24, en el que dichas variantes contienen 1, 2 o 3 mutaciones de sustitución, adición o delección de aminoácidos, en el que una adición de aminoácidos es la adición de un aminoácido, una sustitución de aminoácidos es el reemplazo de un aminoácido por otro aminoácido y una delección de aminoácidos es la eliminación de un aminoácido, o (iii) fragmentos de al menos 10 aminoácidos de SEQ ID NO: 24; y

(b) La parte de toxina 3' codifica un polipéptido que tiene una secuencia de aminoácidos al menos el 90 %, idéntica a la SEQ ID NO: 32.

En una realización, la secuencia señal 5' no es la secuencia señal tox de *Corynebacterium diphtheriae*.

Está dentro de las capacidades de la persona experta identificar epítomos de células B y T. Se pueden identificar epítomos 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 epítomos de células T se puede identificar, por ejemplo, por el procedimiento de epítomo descrito por Sturniolo y col. (Nature Biotech. 17: 555-561 [1999]).

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'.

En una realización adicional, este polinucleótido puede codificar un polipéptido que tiene una secuencia de aminoácidos capaz de dirigir el transporte co-traduccion de una proteína heteróloga al periplasma bacteriano.

Opcionalmente la parte de la toxina 3' codifica la SEQ ID NO:32. Opcionalmente la parte de la toxina 3' codifica DT. Opcionalmente la parte de toxina 3' comprende la SEQ ID NO:31.

En un aspecto adicional de la divulgación, la parte de la secuencia señal 5' codifica (variantes que contienen) 1, 2 o 3 mutaciones, inserciones o delecciones puntuales de una cualquiera de SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26, o una cualquiera de SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26, o una cualquiera de SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, o 24, o SEQ ID NO:24 o una cualquiera de SEQ ID NO:2, 4, o 24, o una cualquiera de SEQ ID NO:2, 10, o 24, o una cualquiera de SEQ ID NO:2, 12, o 24, o una cualquiera de SEQ ID NO:2, 14, o 24, o una cualquiera de SEQ ID NO:4, 10 o 24, o una cualquiera de SEQ ID NO:4, 12, o 24, o una cualquiera de SEQ ID NO:4, 16, o 24, o una cualquiera de SEQ ID NO:4, 18 o 24, o una cualquiera de SEQ ID NO:4, 20 o 24, o una cualquiera de SEQ ID NO:4, 22, o 24, o una cualquiera de SEQ ID NO:10, 12, o 24, o una cualquiera de SEQ ID NO:10, 14, o 24, o una cualquiera de SEQ ID NO:10, 16, o 24, o una cualquiera de SEQ ID NO:10, 18, o 24, o una cualquiera de SEQ ID NO:10, 22 o 24, o una cualquiera de SEQ ID NO:12, 14, o 24, o una cualquiera de SEQ ID NO:12, 16, o 24, o una cualquiera de SEQ ID NO:12, 18, o 24, o una cualquiera de SEQ ID NO:12, 20, o 24, o una cualquiera de SEQ ID NO:12, 22, o 24, o una cualquiera de SEQ ID NO:14, 16, o 24, o una cualquiera de SEQ ID NO:14, 18, o 24, o una cualquiera de SEQ ID NO:14, 20 o 24, o una cualquiera de SEQ ID NO:14, 22, o 24, o una cualquiera de SEQ ID NO:16, 18, o 24, o una cualquiera de SEQ ID NO:16, 20 o 24, o una cualquiera de SEQ ID NO:16, 22, o 24, o una cualquiera de SEQ ID NO:18, 20 o 24, o una cualquiera de SEQ ID NO:18, 22 o 24.

En un aspecto adicional de la divulgación, la parte de la secuencia señal 5' codifica fragmentos de al menos 10, 12, 15, 18 o 20 aminoácidos de SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 o una cualquiera de SEQ ID

NO:2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26, o una cualquiera de SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, o 24, o SEQ ID NO:24, o una cualquiera de SEQ ID NO:2, 4, o 24, o una cualquiera de SEQ ID NO:2, 10, o 24, o una cualquiera de SEQ ID NO:2, 12, o 24, o una cualquiera de SEQ ID NO:2, 14, o 24, o una cualquiera de SEQ ID NO:4, 10 o 24, o una cualquiera de SEQ ID NO:4, 12, o 24, o una cualquiera de SEQ ID NO:4, 16, o 24, o una cualquiera de SEQ ID NO:4, 18 o 24, o una cualquiera de SEQ ID NO:4, 20 o 24, o una cualquiera de SEQ ID NO:4, 22, o 24, o una cualquiera de SEQ ID NO:10, 12, o 24, o una cualquiera de SEQ ID NO:10, 14, o 24, o una cualquiera de SEQ ID NO:10, 16, o 24, o una cualquiera de SEQ ID NO:10, 18, o 24, o una cualquiera de SEQ ID NO:10, 22 o 24, o una cualquiera de SEQ ID NO:12, 14, o 24, o una cualquiera de SEQ ID NO:12, 16, o 24, o una cualquiera de SEQ ID NO:12, 18, o 24, o una cualquiera de SEQ ID NO:12, 20, o 24, o una cualquiera de SEQ ID NO:12, 22, o 24, o una cualquiera de SEQ ID NO:14, 16, o 24, o una cualquiera de SEQ ID NO:14, 18, o 24, o una cualquiera de SEQ ID NO:14, 20 o 24, o una cualquiera de SEQ ID NO:14, 22, o 24, o una cualquiera de SEQ ID NO:16, 18, o 24, o una cualquiera de SEQ ID NO:16, 20 o 24, o una cualquiera de SEQ ID NO:16, 22, o 24, o una cualquiera de SEQ ID NO:18, 20 o 24, o una cualquiera de SEQ ID NO:18, 22, o 24 que son capaces de dirigir el transporte de una proteína al periplasma bacteriano.

En un aspecto adicional de la divulgación se proporciona un polinucleótido que comprende una parte de secuencia señal 5' y una parte de toxina 3', en el que la parte de secuencia señal 5' codifica un péptido señal que tiene una secuencia de aminoácidos de;

(a) SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, o 24; o SEQ ID NO:24, o SEQ ID NO:2, 4, o 24; o SEQ ID NO:2, 10, o 24; o SEQ ID NO:2, 12, o 24; o SEQ ID NO:2, 14, o 24; o SEQ ID NO:4, 10 o 24; o SEQ ID NO:4, 12, o 24; o SEQ ID NO:4, 16, o 24; o SEQ ID NO:4, 18 o 24; o SEQ ID NO:4, 20 o 24; o SEQ ID NO:4, 22, o 24; o SEQ ID NO:10, 12, o 24; o SEQ ID NO:10, 14, o 24; o SEQ ID NO:10, 16, o 24; o SEQ ID NO:10, 18, o 24; o SEQ ID NO:10, 22 o 24; o SEQ ID NO:12, 14, o 24; o SEQ ID NO:12, 16, o 24; o SEQ ID NO:12, 18, o 24; o SEQ ID NO:12, 20, o 24; o SEQ ID NO:12, 22, o 24; o SEQ ID NO:14, 16, o 24; o SEQ ID NO:14, 18, o 24; o SEQ ID NO:14, 20 o 24; o SEQ ID NO:14, 22, o 24; o SEQ ID NO:16, 18, o 24; o SEQ ID NO:16, 20 o 24; o SEQ ID NO:16, 22, o 24; o SEQ ID NO:18, 20 o 24; o una cualquiera de SEQ ID NO:18, 22, o 24.

(b) (variantes de) SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, o 24; o SEQ ID NO:24, o SEQ ID NO:2, 4, o 24; o SEQ ID NO:2, 10, o 24; o SEQ ID NO:2, 12, o 24; o SEQ ID NO:2, 14, o 24; o SEQ ID NO:4, 10 o 24; o SEQ ID NO:4, 12, o 24; o SEQ ID NO:4, 16, o 24; o SEQ ID NO:4, 18 o 24; o SEQ ID NO:4, 20 o 24; o SEQ ID NO:4, 22, o 24; o SEQ ID NO:10, 12, o 24; o SEQ ID NO:10, 14, o 24; o SEQ ID NO:10, 16, o 24; o SEQ ID NO:10, 18, o 24; o SEQ ID NO:10, 22 o 24; o SEQ ID NO:12, 14, o 24; o SEQ ID NO:12, 16, o 24; o SEQ ID NO:12, 18, o 24; o SEQ ID NO:12, 20, o 24; o SEQ ID NO:12, 22, o 24; o SEQ ID NO:14, 16, o 24; o SEQ ID NO:14, 18, o 24; o SEQ ID NO:14, 20 o 24; o SEQ ID NO:14, 22, o 24; o SEQ ID NO:16, 18, o 24; o SEQ ID NO:16, 20 o 24; o SEQ ID NO:16, 22, o 24; o SEQ ID NO:18, 20 o 24; o una cualquiera de SEQ ID NO:18, 22, o 24 que varían 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 SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, o 24; o SEQ ID NO:24, o SEQ ID NO:2, 4, o 24; o SEQ ID NO:2, 10, o 24; o SEQ ID NO:2, 12, o 24; o SEQ ID NO:2, 14, o 24; o SEQ ID NO:4, 10 o 24; o SEQ ID NO:4, 12, o 24; o SEQ ID NO:4, 16, o 24; o SEQ ID NO:4, 18 o 24; o SEQ ID NO:4, 20 o 24; o SEQ ID NO:4, 22, o 24; o SEQ ID NO:10, 12, o 24; o SEQ ID NO:10, 14, o 24; o SEQ ID NO:10, 16, o 24; o SEQ ID NO:10, 18, o 24; o SEQ ID NO:10, 22 o 24; o SEQ ID NO:12, 14, o 24; o SEQ ID NO:12, 16, o 24; o SEQ ID NO:12, 18, o 24; o SEQ ID NO:12, 20, o 24; o SEQ ID NO:12, 22, o 24; o SEQ ID NO:14, 16, o 24; o SEQ ID NO:14, 18, o 24; o SEQ ID NO:14, 20 o 24; o SEQ ID NO:14, 22, o 24; o SEQ ID NO:16, 18, o 24; o SEQ ID NO:16, 20 o 24; o SEQ ID NO:16, 22, o 24; o SEQ ID NO:18, 20 o 24; o una cualquiera de SEQ ID NO: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.

La divulgación también proporciona un polinucleótido que comprende una parte de secuencia señal 5' y una parte de toxina 3', en el que la parte de secuencia señal 5' codifica un péptido señal que tiene una secuencia de aminoácidos de

(d) SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26;

(e) variantes de SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26, que varían 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

(f) fragmentos de al menos 10 aminoácidos de SEQ ID NO: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.

- En un aspecto de la divulgación, la parte de la secuencia señal 5' del polinucleótido de la divulgación codifica al menos una de las SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26. En un aspecto adicional de la divulgación la parte de la secuencia señal 5' del polinucleótido de la invención codifica al menos una de las SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26. En un aspecto adicional de la divulgación la parte de la secuencia señal 5' del polinucleótido de la invención codifica al menos una de las SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, 24, o 26; o SEQ ID NO:2, 4, 10, 12, 14, 16, 18, 20, 22, o 24; o SEQ ID NO:24, o SEQ ID NO:2, 4, o 24; o SEQ ID NO:2, 10, o 24; o SEQ ID NO:2, 12, o 24; o SEQ ID NO:2, 14, o 24; o SEQ ID NO:4, 10 o 24; o SEQ ID NO:4, 12, o 24; o SEQ ID NO:4, 16, o 24; o SEQ ID NO:4, 18 o 24; o SEQ ID NO:4, 20 o 24; o SEQ ID NO:4, 22, o 24; o SEQ ID NO:10, 12, o 24; o SEQ ID NO:10, 14, o 24; o SEQ ID NO:10, 16, o 24; o SEQ ID NO:10, 18, o 24; o SEQ ID NO:10, 22 o 24; o SEQ ID NO:12, 14, o 24; o SEQ ID NO:12, 16, o 24; o SEQ ID NO:12, 18, o 24; o SEQ ID NO:12, 20, o 24; o SEQ ID NO:12, 22, o 24; o SEQ ID NO:14, 16, o 24; o SEQ ID NO:14, 18, o 24; o SEQ ID NO:14, 20 o 24; o SEQ ID NO:14, 22, o 24; o SEQ ID NO:16, 18, o 24; o SEQ ID NO:16, 20 o 24; o SEQ ID NO:16, 22, o 24; o SEQ ID NO:18, 20 o 24; o una cualquiera de SEQ ID NO:18, 22, o 24.
- Las secuencias de nucleótidos codificadas por las SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 pueden ser idénticas a las secuencias polinucleotídicas correspondientes de SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, o 25. Alternativamente esta puede ser cualquier secuencia, que como resultado de la redundancia (degeneración) del código genético, que también codifica polipéptidos de las SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26. Alternativamente el polinucleótido puede comprender una parte que codifica las SEQ ID NO:33-45.
- Los polinucleótidos de la divulgación también pueden comprender una parte de secuencia señal 5' que codifica una variante de SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 que es capaz de dirigir una proteína al periplasma.
- La presente invención también proporciona una parte de secuencia señal 5' que comprende un fragmento de SEQ ID NO:24. Los fragmentos de la invención consisten en partes contiguas de al menos, 10 aminoácidos de SEQ ID NO:24. En una realización, los fragmentos conducen a que se transporte al periplasma al menos el 50, 60, 70, 90, 90 o 100 % de la toxina codificada polipeptídica. En una realización adicional, los fragmentos tienen las mismas o sustancialmente las mismas propiedades de transporte periplasmático que el polipéptido que comprende la secuencia de aminoácidos correspondiente de SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26. En una realización adicional, los fragmentos son capaces de dirigir el transporte cotraduccional de un polipéptido al periplasma.
- La parte de la toxina 3' codifica cualquier toxina bacteriana. En una realización, la toxina bacteriana codificada es la toxina de la difteria, del tétanos, pertúsica o neumolisina.
- En una realización, la parte de toxina 3' codifica un toxoide bacteriano, por ejemplo, el toxoide de *Diphtheria* o CRM 197. Opcionalmente la parte de toxina 3' codifica la secuencia de aminoácidos descrita en la SEC ID NO:32. En una realización adicional, la parte de toxina 3' comprende la secuencia de nucleótidos de SEQ ID NO:31. En una realización el polipéptido codifica las SEC ID NO:33-45.
- 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 en 5' de la parte de toxina 3'. Alternativamente están separadas 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 invención contempla la expresión de cualesquiera de los polinucleótidos dentro de una secuencia que codifica 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 la purificación tales como residuos de histidina múltiples, o una secuencia adicional para la estabilidad durante la producción recombinante. Además, se considera también la adición de secuencias exógenas de polipéptidos o de cola lipídica o de polinucleótidos para incrementar el potencial inmunogénico de la molécula final.
- En una realización, la parte de secuencia señal 5' no es la secuencia señal tox de *C. diphtheriae*.

Polipéptidos de la invención

- La presente invención también proporciona polipéptidos codificados por los polinucleótidos de la invención. Además se proporcionan fragmentos y variantes de estos polipéptidos.

La invención abarca polipéptidos que codifican polinucleótidos de la invención.

La invención también abarca fragmentos de estos polipéptidos. Los fragmentos contienen segmentos tanto de la parte de la secuencia señal 5' como de la parte de la toxina 3'. En una realización adicional, estos fragmentos comprenden al menos 10, 15 o 20 aminoácidos de un péptido señal. En una realización, estos fragmentos comprenden al menos 10, 15 o 20 de los aminoácidos a partir de las SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, o 26 o de las SEQ ID NO:33-45.

En una realización, 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*.

La invención también abarca proteínas de fusión que incluyen el polipéptido o los fragmentos polipeptídicos de la invención.

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

Vectores y células hospedadoras

La divulgación se refiere también a vectores que comprenden un polinucleótido o polinucleótidos de la invención, células hospedadoras que están modificadas mediante ingeniería genética con los vectores de la invención y la producción de polipéptidos por técnicas recombinantes. También pueden emplearse sistemas de traducción sin células para producir tales proteínas usando ARN derivados de las construcciones de ADN de la invención.

Los polipéptidos recombinantes de la presente invención 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 invención también se refiere a sistemas de expresión que comprenden un polinucleótido o polinucleótidos de la presente invención, a células hospedadoras que están modificadas mediante ingeniería genética con tales sistemas de expresión y a la producción de polipéptidos de la invención por técnicas recombinantes. En un aspecto adicional de la invención, la presente invención se refiere a un polinucleótido de la invención ligado a un promotor inducible de tal manera que cuando el promotor se induce se expresa un polipéptido codificado por el polinucleótido. Un aspecto adicional de la invención comprende dicho vector en el que el promotor inducible está activado por la 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 la producción recombinante de los polipéptidos de la invención, las células hospedadoras 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 hospedadora 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 hospedadores apropiados incluyen células bacterianas gram negativas, tales como células de *E. coli*, *Acinetobacter*, *Actinobacillus*, *Bordetella*, *Brucella*, *Campylobacter*, *Cyanobacteria*, *Enterobacter*, *Erwinia*, *Francisella*, *Helicobacter*, *Hemophilus*, *Klebsiella*, *Legionella*, *Moraxella*, *Neisseria*, *Pasteurella*, *Proteus*, *Pseudomonas*, *Salmonella*, *Serratia*, *Shigella*, *Treponema*, *Vibrio*, *Yersinia*. En un aspecto adicional de la invención la célula hospedadora es una célula de *Escherichia coli*.

Se puede usar una gran variedad de sistemas de expresión para producir los polipéptidos de la invención. En una realización el vector 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 hospedador se puede usar para la expresión a este respecto. La secuencia de ADN 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 adicional de la invención se proporciona un procedimiento para fabricar una toxina bacteriana que comprende las etapas de a) cultivar un cultivo de la célula hospedadora bacteriana de la invención y b) inducir la expresión del polipéptido de tal forma que una toxina bacteriana se expresa de manera periplasmática.

En una realización adicional de la invención, se proporciona un procedimiento para la expresión periplasmática de un recombinante

- A. Cultivando un cultivo de una célula hospedadora gram negativa; y
- B. Induciendo la expresión de un polipéptido de tal forma que dicha proteína se exprese de manera periplasmática;

en el que durante la expresión se accionan las siguientes etapas:

- i. El pH de la etapa a) es inferior al pH de la etapa b);
- ii. La temperatura de la etapa a) es más alta que la temperatura de la etapa b); o
- iii. La velocidad de alimentación del sustrato de la etapa a) es mayor que la velocidad de alimentación del sustrato de la etapa b).

En una etapa de la realización, se acciona la etapa i, la etapa ii y la etapa iii, la etapa i y la etapa ii, la etapa i y la etapa iii, la etapa ii y la etapa iii o la etapa i, la etapa ii y la etapa iii. En una realización adicional, se expresa un polinucleótido de la invención.

La expresión de polipéptidos se induce cuando un agente inductor tal como IPTG se añade a un cultivo de células hospedadoras, causando la expresión del polipéptido a una velocidad incrementada.

En una realización, el pH de la etapa a) es el mismo que el pH de la etapa b), en una segunda realización, el pH de la etapa a) es más bajo que el pH de la etapa b), es decir, el pH de la etapa b) aumenta en comparación con el de la etapa a).

La persona experta en la técnica reconocerá como efectuar un cambio entre la etapa a) y la etapa b). Un cambio en una condición (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 indica 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.

En una realización adicional el pH de la etapa a) varía de 5,0-7,0, 5,0-6,0, 6,0-7,0 o de 6,5-7,0.

En una realización el pH de la etapa b) se mantiene. En una realización el pH se mantiene a un pH mayor de 6,5, 7,0, 7,5, 8,0, 8,5, 9,0, 9,5, 10,0 o entre 6,5 y 10,0, 6,5 y 9,5, 6,5 y 9,0, 6,5 y 8,5, 6,5 y 7,5, 6,5 y 7,0 y 10,0, 7,0 y 9,5, 7,0 y 9,0, 7,0 y 8,5, 7,0 y 8,0, 7,0 y 7,5, 7,5 y 10,0, 7,5 y 9,5, 7,5 y 9,0, 7, y 8,5, 7,5 y 8,0, 8,0 y 10,0, 8,0 y 9,5, 8,0 y 9,0, 8,0 y 8,5, 8,5 y 10,0, 8,5 y 9,5, 8,0 y 9,0, 8,0 y 8,5, 8,5 y 10,0, 8,5 y 9,5, 8,5 y 9,0, 9,0 y 10,0, 9,0 y 9,5 o 9,5 y 10,0. 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.

En una realización el pH en la etapa b) es al menos, exacta o aproximadamente 2,0, 1,5, 1,0, 0,5, 0,3, 0,2, o 0,1 unidades de pH más altas que el pH en la etapa a).

En una realización el procedimiento de la invención se lleva a cabo en un fermentador. En una realización el pH está incrementado de tal forma que el pH en la etapa b) es mayor que el pH en la etapa a). Opcionalmente este incremento en pH se logra mediante la adición de una base, por ejemplo, hidróxido de sodio o amoníaco.

En una realización adicional el pH de la etapa (b) se mantiene entre 6,5 y 10,0, 6,5 y 9,5, 6,5 y 9,0, 6,5 y 8,5, 6,5 y 7,5, 6,5 y 7,0, 7,0 y 10,0, 7,0 y 9,5, 7,0 y 9,0, 7,0 y 8,5, 7,0 y 8,0, 7,0 y 7,5, 7,5 y 10,0, 7,5 y 9,5, 7,5 y 9,0, 7,5 y 8,5, 7,5 y 8,0, 8,0 y 10,0, 8,0 y 9,5, 8,0 y 9,0, 8,0 y 8,5, 8,5 y 10,0, 8,5 y 9,5, 8,0 y 9,0, 8,0 y 8,5, 8,5 y 10,0, 8,5 y 9,5, 8,5 y 9,0, 9,0 y 10,0, 9,0 y 9,5 o 9,5 y 10,0 mediante la adición de una base, por ejemplo, hidróxido de sodio o amoníaco, durante la etapa b).

En una realización, 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. En una realización adicional, la segunda velocidad de alimentación de sustrato se mantiene entre el 5 % y el 90 %, el 20 % y el 80 % o el 20 % y el 30 % de la velocidad de alimentación de sustrato mantenida durante la etapa a).

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, la toxina bacteriana es toxina de neumolisina, toxina diftérica, una toxina tetánica o una toxina tosferínica o una variante mutada puntual de la misma. En una realización, la toxina es toxoide de difteria. En una realización adicional la toxina es CRM 197, que opcionalmente está codificada por la SEQ ID NO:32. En una realización adicional, la toxina bacteriana es un fragmento o una variante mutada puntual de la toxina de neumolisina, toxina diftérica, toxina tetánica o toxina tosferínica.

La célula hospedadora bacteriana de este procedimiento se definió en la sección sobre vectores y células hospedadoras. En una realización, la célula hospedadora bacteriana se selecciona del grupo constituido por *E. coli*, *Acinetobacter*, *Actinobacillus*, *Bordetella*, *Brucella*, *Campylobacter*, *Cyanobacteria*, *Enterobacter*, *Erwinia*, *Francisella*, *Helicobacter*, *Hemophilus*, *Klebsiella*, *Legionella*, *Moraxella*, *Neisseria*, *Pasteurella*, *Proteus*, *Pseudomonas*, *Salmonella*, *Serratia*, *Shigella*, *Treponema*, *Vibrio* y *Yersinia*.

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

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.

En una realización adicional, la expresión 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).

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 el cultivo de bacterias a menor escala. 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.

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

En una realización el nivel de oxígeno disuelto (OD) en el fermentador es (se mantiene) 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 de 5 %-50 %, 10 %-40 %, 15 %-25 % o 17 %-22 %. Un 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.

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.

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 hospedadora bacteriana para obtener una toxina bacteriana madura. Opcionalmente esta eliminación se lleva a cabo por maquinaria de la célula hospedadora. 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 recoger la pasta celular del cultivo.

En una realización la centrifugación se usa para recoger 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 alteración celular y la purificación adicional usando cromatografía y técnicas de filtración.

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

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 la 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.

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.

Los conjugados de sacáridos presentes en las composiciones inmunogénicas de la invención 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 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-succinimidilo 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 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 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.

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 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 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.

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.
- 5 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.
- 10 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.
- 15 E) Grupo imidazolilo (por ejemplo por medio de histidina). En una realización este grupo está activado/modificado con bisdiazobencidina.
- 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.
- 20 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
- 25 Sacárido-NH₂ + COOH-Prot + EDAC ----> conjugado
- 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
- 30 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
- 35 Sacárido-OH + CNBr o CDAP ----> éster de cianato + NH₂----SH ----> Sacárido-SH + Prot haloacetilada ----> Conjugado
- 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
- 40 Sacárido-COOH + EDAC+ NH₂----SH ----> sacárido----SH + maleimida-Prot (modificación de grupos amino) ----> conjugado
- 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.
- 45 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 de la invención

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;

5 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*.

10 En una etapa adicional, los polipéptidos de la invención 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 polifosfazenos.

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.

35 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.

Ejemplos**Ejemplo 1 diseño de construcciones de secuencia señal - CRM197**

40 Utilizando técnicas estándar de biología molecular, se crearon plásmidos que contenían una secuencia codificante de una secuencia señal N-terminal fusionada al toxoide CRM197. Se utilizaron trece secuencias de señal diferentes (tabla 2) que se seleccionaron de la base de datos GenEMBL.

Tabla 2

Secuencia señal	SEQ ID NO: de nucleótidos	SEQ ID NO: de aminoácidos	Origen
PhtE	1	2	<i>S. pneumoniae</i>
SipA	3	4	Strep del grupo B
OmpA	5	6	<i>E. coli</i>
NspA	7	8	<i>N. meningitidis</i>

(continuación)

Secuencia señal	SEQ ID NO: de nucleótidos	SEQ ID NO: de aminoácidos	Origen
TorT	9	10	<i>E. coli</i>
SfmC	11	12	<i>E. coli</i>
FocC	13	14	<i>E. coli</i>
Ccmh	15	16	<i>E. coli</i>
Yra1	17	18	<i>E. coli</i>
TolB	19	20	<i>E. coli</i>
NikA	21	22	<i>E. coli</i>
FlgI	23	24	<i>E. coli</i>
DsbA	25	26	<i>E. coli</i>

La primera construcción se creó a partir de tres plásmidos (véase la figura 1). El plásmido A, creado por Geneart, contenía la secuencia señal DsbA (SEQ ID NO: 25) en 3' con respecto a una secuencia que codifica un fragmento N-terminal de CRM197 (aminoácidos 1-206 de la SEQ ID NO 32). Este plásmido se clonó en el plásmido estructural pET26b (Novagen) por digestión de ambos plásmidos utilizando enzimas de restricción NdeI y SacI seguido de ligamiento. El plásmido B, creado por Geneart, contenía un fragmento C terminal de CRM197 (aminoácidos 206-533 de SEQ ID NO 32). Este se clonó en el plásmido estructural pET26b usando restricción con enzimas de restricción AatII y XhoI, el plásmido resultante se denominó pRIT16668. Esta construcción contiene la secuencia señal de DsbA hasta el sitio de unión de la signalasa, seguido de dos aminoácidos de DsbA encontrados después de la secuencia de unión a signalasa, otros dos aminoácidos del sitio de unión de la enzima de restricción, y finalmente CRM197 pierde los dos primeros aminoácidos. Esto significa que cuando estas construcciones se expresan, la proteína producida después de la expresión periplasmática y de la eliminación de la secuencia señal por signalasa contiene CRM197 que pierde los dos primeros amino ácidos pero con cuatro aminoácidos adicionales de DsbA y el sitio de restricción.

El plásmido pRIT16668 contiene sitios de restricción NdeI y AgeI que flanquean la secuencia señal DsbA.

Se crearon plásmidos que contenían las otras 12 secuencias de señal digiriendo el plásmido pRIT16668 con NdeI y AgeI para eliminar la secuencia señal DsbA. Los oligonucleótidos hibridados que codifican cada una de las otras 12 secuencias de señal estaban, a la vez, ligados entre estos dos sitios.

Los plásmidos resultantes se transformaron en células Novablue químicamente competentes (Novagen cat 70181-3, utilizado según lo recomendado por el fabricante). Se confirmó la presencia del inserto correcto para cada plásmido mediante secuenciación. Después, para su expresión, los plásmidos se transformaron individualmente en células B834(DE3) químicamente competentes (Novagen cat. 69041-3HMS174 (DE3) células químicamente competentes o, alternativamente, en células BLR(DE3) químicamente competentes (Novagen cat. 69053-3, utilizado según lo recomendado por el fabricante).

Ejemplo 2 diseño de una construcción CRM197 para la expresión citoplasmática (que carece de la secuencia señal periplasmática)

En este experimento, se creó un plásmido similar a los creados en el Ejemplo 1 pero que carecía de la secuencia señal periplasmática. Este procedimiento de diseño se resume en la Figura 2. En primer lugar, la secuencia CRM197 (SEQ ID NO: 31) se amplificó mediante técnicas de PCR estándar utilizando dos cebadores que se denominaron MDSCRM1 (SEQ ID NO: 27) y MDSCRM2 (SEQ ID NO: 28) y pRIT16668 como molde (los ciclos de temperatura utilizados fueron (94 °C 2'- 55 °C 2'- 72 °C 2'30) X 25 – 72 °C 10'- final 4 °C).)

Este fragmento se insertó en el plásmido pRIT16668 en lugar del inserto de secuencia señal CRM197. Esto se realizó utilizando técnicas estándar de biología molecular a través de la digestión del producto de PCR y pRIT16668 con las enzimas de restricción NdeI y XhoI. El plásmido resultante contiene la secuencia CRM197 madura (SEQ ID NO: 31) pero no contiene secuencia señal y se denominó pRIT16669.

Los plásmidos resultantes se transformaron en células Novablue químicamente competentes (Novagen cat 70181-3, utilizado según lo recomendado por el fabricante). La presencia del inserto correcto para cada plásmido se confirmó mediante secuenciación. Después, para la expresión, los plásmidos se transformaron en células B834(DE3)

químicamente competentes (Novagen cat 69041-3, utilizado según lo recomendado por el fabricante).

Ejemplo 3 diseño de una construcción óptima de secuencia señal - CRM197

La secuencia señal Flgl se seleccionó para su uso para crear una construcción mejorada. El procedimiento de clonación seguido se resume en la figura 3.

- 5 Una región de ADN que contenía la secuencia señal Flgl fusionada con la parte N terminal de la secuencia CRM197, se amplificó utilizando técnicas estándar de PCR. Se utilizaron dos cebadores, denominados Flgl c-CRMopt3e (SEQ ID NO: 29) y MDSCRM906NC (SEQ ID NO: 30) utilizando, como molde, pRIT16669, con ciclos (94 °C 2' (94 °C 1'- 60 °C 1'- 72 °C 2'30) X3 (94 °C 1'- 58 °C 1'- 72 °C 2'30) X3 (94 °C 1'-56 °C 1'- 72 °C 2'30) X3 (94 °C 1'-54 °C 1'- 72 °C 2'30) X20 72 °C 10'- final 4 °C).
- 10 Este fragmento se insertó en el plásmido pRIT16669 utilizando técnicas estándar de biología molecular a través de digestión del producto de PCR y el plásmido pRIT16669 con las enzimas de restricción NdeI y AatII. El plásmido resultante contiene el extremo N maduro completo de CRM197 (SEQ ID NO: 31) y la secuencia señal Flgl que termina en el sitio de unión de signalasa (SEQ ID NO: 23) y se denominó pRIT16681
- 15 Los plásmidos resultantes se transformaron en células Novablue químicamente competentes (Novagen cat 70181-3, utilizado según lo recomendado por el fabricante). La presencia del inserto correcto para cada plásmido se confirmó mediante secuenciación. Después, para la expresión, los plásmidos se transformaron en células B834(DE3) químicamente competentes (Novagen cat 69041-3).

Ejemplo 4 expresión de la construcción secuencia señal DsbA - CRM197

- 20 En este experimento, las células transformadas con las construcciones de secuencia señal DsbA, creadas en el Ejemplo 1, se cultivaron y se indujo su expresión. Precultivos (3 ml) de la cepa recombinante B834(DE3), HMS174(DE3) y BLR(DE3) se cultivaron en medio LBT complementado con glucosa al 1 % en presencia de kanamicina (50 µg/ml) durante una noche a 37 °C con agitación. Al día siguiente, muestras de este precultivo se añadieron a 20 ml de medio LBT con 50 µg/ml de kanamicina después de que la densidad óptica DO₆₂₀ del medio alcanzara un valor de 0,1. Se dejó que estas crecieran a 30 °C con agitación. En este plásmido, la expresión de
- 25 CRM197 está bajo el control del operador *lac*, por lo tanto, la expresión de CRM197 codificado puede inducirse con la adición de isopropil-beta-D tiogalactopiranosido (IPTG). Cuando la densidad óptica alcanzó un valor de 0,6 se realizó la inducción con IPTG (añadido hasta una concentración final de 1 mM) y se permitió que el cultivo creciese durante 3 horas más a 30 °C.
- 30 El nivel de expresión de CRM197 se evaluó procesando el producto celular bacteriano total en un gel de SDS-PAGE (Novex Tris-glicina 10-20 %) y tiñendo con azul brillante de Coomassie. Se realizó una transferencia Western usando anticuerpos policlonales anti-DTPa de ratón con NBT-BCIP.

- 35 La Figura 4 describe los resultados de este experimento. El gel de Coomassie no muestra diferencias en el perfil de expresión entre los cultivos inducidos y no inducidos. La transferencia Western demuestra una ausencia de CRM197 en los cultivos tanto inducidos como no inducidos. En estas condiciones, no hay, o hay poca, CRM197 periplasmática expresada.

Ejemplo 5 optimización de las condiciones de pH y temperatura y análisis de expresión de las construcciones

- 40 Las etapas del ejemplo 4 se repitieron, sin embargo, en este caso, antes de la inducción, al cultivo se añadieron 100 mM de tampón K₂HPO₄/KH₂PO₄ de tal manera que el medio se tamponó a un pH de 7,5. La expresión se indujo por la adición de IPTG 1 mM y se dejó que el cultivo creciera durante la noche a 23 °C o a 30 °C. El nivel de proteína expresado en el sobrenadante y en el sedimento se evaluó utilizando técnicas de tinción de Coomassie y transferencia Western como se describe en el Ejemplo 4.

- 45 Con respecto a las secuencias señal postraduccionales, el nivel de expresión de CRM197 fue muy alto para la secuencia señal OmpA, sin embargo, la proteína se escinde para formar una proteína de aproximadamente 27 kDa. Esta forma truncada está asociada a células (no se encuentra ni en la fracción de liberación periplasmática ni en el medio de cultivo). Con respecto a la secuencia señal de NspA, la expresión es positiva pero se obtiene un nivel de CRM197 maduro bajo.

- 50 Con respecto a las secuencias co-traduccionales, el nivel de expresión es alto con las secuencias señal de Flgl y SipA, sin embargo, la CRM197 producida no está completamente madura. La expresión con la secuencia señal NikA conduce a la expresión de la proteína madura. El nivel de expresión global es más bajo que con la secuencia señal Flgl y SipA, pero sigue siendo detectable en el extracto total por tinción de Coomassie.

Ejemplo 6 análisis de expresión de las construcciones CRM197 de expresión citoplasmática

Células DE3 se transformaron con los plásmidos producidos en el ejemplo 2, se cultivaron y se indujo la expresión de CRM197 como se describe en el ejemplo 4. El nivel de expresión de CRM197 se evaluó ejecutando el producto

de la célula bacteriana total en un gel de SDS-PAGE y usando técnicas de tinción con azul Coomassie y transferencia de Western como se describe en el ejemplo 4. Los extractos bacterianos se realizaron mediante una técnica de disparo (sistema constante) en tampón PBS estándar. Después de la centrifugación, las fracciones solubles e insolubles obtenidas se cargaron en un gel.

- 5 En la figura 9 se muestran los resultados de este experimento. Aunque el gen está bien expresado en estas cepas, una parte sustancial de la proteína es insoluble y permanece en el sedimento. Además, la expresión citoplasmática que aparece en el gel, conduce a proteínas con un peso molecular aparente más bajo.

Ejemplo 7 análisis de expresión de la construcción Flgl optimizada

- 10 Células BLR(DE3) de *E. coli* se transformaron con la construcción producida en el ejemplo 3, se cultivaron y se indujo la expresión como se describe en el protocolo optimizado del ejemplo 5. Además, algunas células de muestra se cultivaron sin la adición de tampón K_2HPO_4/KH_2PO_4 100 mM. De manera similar, el nivel de proteína expresada se evaluó usando técnicas de tinción de Coomassie y de transferencia Western como se describe en el Ejemplo 4. La temperatura de inducción se mantuvo a 23 °C o a 30 °C y se tomaron muestras en el momento de la inducción, 2 horas después de la inducción, 4 horas después de la inducción y después de una inducción durante toda la noche.
- 15 En la Figura 7 se muestran los resultados de este experimento. La transferencia Western demuestra un nivel de expresión de CRM197 fácilmente detectable con tampón a 23 °C. Se observó que los niveles de expresión de proteínas en las células que no estaban tamponados durante la fase de inducción eran más bajos.

Ejemplo 8 Precultivo de B2355 en *Escherichia coli*

- 20 Se preparó un precultivo usando un cultivo de semillas congelado de *Escherichia coli* cepa B2355. Esta cepa es una cepa B834(DE3) transformada con un derivado de pET26b que contiene una secuencia que codifica una proteína de fusión entre el péptido señal de Flgl de *E. coli* y la parte madura de CRM197 (este es el plásmido pRIT16681 descrito en la figura 3 y en el ejemplo 3). La capacidad de cultivo de semillas se determinó que era de aproximadamente 1×10^{10} unidades formadoras de colonias por ml.

- 25 El cultivo de semillas se descongeló a temperatura ambiente y se usaron 400 µl para inocular un matraz Erlenmeyer de 2 litros que contenía 400 ml de medio de precultivo (adaptado de Zabriskie et al. (J. Ind. Microbiol. 2: 87-95 (1987))).

Después, el matraz inoculado se incubó a 37 °C (± 1 °C) y a 200 rpm. El precultivo se detuvo cuando la densidad óptica a 650 nm (DO_{650} nm) alcanzó 2,50, (alrededor de 6 h de incubación). El precultivo se utilizó para inocular medio en un fermentador tan pronto como se detuvo el cultivo (ejemplo 9).

Ejemplo 9: fermentación semicontinua a escala de 20 l

Procedimiento

- 35 Se usó un fermentador de 20 litros (Biolafitte). Se transfirieron asépticamente nueve litros de medio de fase discontinua al fermentador (adaptado de Zabriskie et al. (J. Ind. Microbiol. 2: 87-95 (1987))). El pH del medio se reajustó a 6,8 con adición de base. También se añadieron 3 ml de antiespumante irradiado sin diluir (SAG 471) al fermentador. Después, antes de la inoculación, se ajustó la temperatura (28 °C), la presión de descarga (0,5 bar), la velocidad de oxigenación (20 litros de aire rociado por minuto) y la velocidad de agitación inicial (300 rpm). El nivel de oxígeno disuelto en estas condiciones fue del 100 %. La presión de descarga y la velocidad de oxigenación se mantuvieron a un nivel constante durante la fermentación.

La inoculación se realizó añadiendo 18 ml de precultivo (preparado como se describe en el Ejemplo 8).

- 40 Durante la fase discontinua (0-15 h), la temperatura se mantuvo a 28 °C. El nivel de oxígeno disuelto se estableció al 20 %. El nivel de oxígeno disuelto (OD) se reguló aumentando la agitación cuando el OD disminuyó por debajo del 20 %. El agotamiento de glucosa produjo un aumento del OD y la disminución simultánea de la agitación.

Después de 15 h de fermentación, se añadió más sustrato según el siguiente perfil de adición de alimentación.

Tabla 3

Tiempo de fermentación	Tasa de alimentación de sustrato adicional (ml/min)	Alimentación acumulativa en peso (g)
0	0,000	0
1	0,000	0

ES 2 750 023 T3

(continuación)		
Tiempo de fermentación	Tasa de alimentación de sustrato adicional (ml/min)	Alimentación acumulativa en peso (g)
2	0,000	0
3	0,000	0
4	0,000	0
5	0,000	0
6	0,000	0
7	0,000	0
8	0,000	0
9	0,000	0
10	0,000	0
11	0,000	0
12	0,000	0
13	0,000	0
14	0,000	0
15	0,000	0
16	0,600	21
17	1,150	81
18	1,150	161
19	1,150	241
20	1,150	321
21	1,150	400
22	1,150	480
23	1,150	560
24	1,150	639
25	1,150	719
26	1,150	799
27	1,150	878
28	1,150	958
29	1,150	1038
30	1,150	1117
31	1,150	1197
32	1,150	1277

ES 2 750 023 T3

(continuación)

Tiempo de fermentación	Tasa de alimentación de sustrato adicional (ml/min)	Alimentación acumulativa en peso (g)
33	1,150	1357
34	1,150	1436
35	1,150	1516
36	1,150	1596
37	1,150	1675
38	1,150	1755
39	1,150	1835
40	1,150	1914
41	1,150	1994
42	1,150	2074
43	1,150	2153
44	1,150	2233
45	1,150	2313
46	1,150	2393
47	0,325	2444
48	0,325	2466
49	0,325	2489
50	0,325	2511
51	0,325	2534
52	0,325	2556
53	0,325	2579
54	0,325	2601
55	0,325	2624
56	0,325	2646
57	0,325	2669
58	0,325	2691
59	0,325	2714
60	0,325	2736
61	0,325	2759
62	0,325	2782
63	0,325	2804

(continuación)

Tiempo de fermentación	Tasa de alimentación de sustrato adicional (ml/min)	Alimentación acumulativa en peso (g)
64	0,325	2827
65	0,325	2849
66	0,325	2872
67	0,325	2894
68	0,325	2917
69	0,325	2939
70	0,325	2962
71	0,325	2984
72	0,325	3007

Durante la fase semicontinua (15-46 h), el pH se mantuvo a 6,8 mediante la adición de base, la temperatura se reguló a 28 °C, y el nivel de OD se mantuvo al 20 % controlando la velocidad de agitación.

- 5 A las 46 horas se añadió IPTG a una concentración final de 1 mM para inducir las bacterias. Además el pH se aumentó gradualmente después de 46 horas mediante la adición de base, y la temperatura disminuyó a 23 °C (estos cambios son necesarios para altos niveles de expresión periplasmática). El pH y la temperatura se mantuvieron durante toda la fase de inducción (46-72 h). El nivel de OD se mantuvo al 20 % controlando la velocidad de agitación.
- 10 Al final de la fase de inducción (72 h), la pasta celular se recogió por centrifugación (6 500xg, 4 °C durante 1 h), y se conservó a -20 °C.

La extracción periplasmática se realizó mediante choque osmótico usando un procedimiento adaptado de Chen et al. (Biochem. Ing. J. 19: 211-215 (2004)). Mediante un ensayo Elisa se analizó el contenido de CRM197 en las fracciones periplasmática y citoplasmática.

- 15 La Figura 8 muestra un perfil de fermentación típico con los parámetros del proceso monitorizados durante la fermentación semicontinua a escala de 20 l.

Al final de la fermentación, mediante un ensayo Elisa se analizó la productividad periplasmática de CRM197:

Periplasmática	Citoplasmática	Eficacia de secreción
3180 mg/l	394 mg/l	87 %

- 20 Esta técnica demostró niveles de expresión y eficacia de secreción sin precedentes.

Ejemplo 10: determinación de las velocidades y temperaturas de alimentación óptimas a utilizar durante la fase de inducción

- 25 En este experimento, se utilizó la metodología de superficie de respuesta (J.ind. Microbiol. Biotechnol. 37: 195-204 (2010)) para determinar los valores óptimos de tres parámetros, con el fin de maximizar la producción periplasmática de una proteína recombinante. Los tres parámetros de fermentación investigados fueron el pH durante la fase de cultivo, el pH durante la inducción y la velocidad de alimentación durante la inducción. Los valores de estos tres parámetros se eligieron de acuerdo con un diseño de coraza uniforme de Doehlert (Doehlert (Applied Statistics 19: 231-239 (1970))). Utilizando los valores descritos en la tabla 5 se realizaron quince fermentaciones.

- 30 Las fermentaciones se realizaron utilizando la cepa B2284, esta es una cepa de células BLR(DE3) transformadas con un derivado de pET26b que contiene una secuencia que codifica una proteína de fusión entre el péptido señal de FlgI de *E. coli* y la parte madura de CRM197 (este es el plásmido pRIT16681 descrito en la figura 3 y en el ejemplo 3).

Para cada fermentación, el cultivo de semillas se descongeló a temperatura ambiente y se utilizaron 500 ml para inocular un matraz Erlenmeyer de 2 litros que contenía 400 ml de medio de precultivo (adaptado de Zabriskie et al. (J. Ind. Microbiol. 2: 87-95)(1987))).

Después, el matraz inoculado se incubó a 37 °C ($\pm$ 1 °C) y a 200 rpm. El precultivo se detuvo cuando la densidad óptica a 650 nm (DO_{650} nm) alcanzó aproximadamente 2,50, (alrededor de 6 h de incubación). El precultivo se utilizó para inocular medio en un fermentador tan pronto como se detuvo el cultivo (adaptado de Zabriskie et al. (J. Ind. Microbiol.2: 87-95 (1987))).

Se utilizó un fermentador de 20 litros (BiolaFitte). Al fermentador se transfirieron asépticamente nueve litros de medio de fase discontinua. El pH del medio se reajustó a 6,8 con adición de base. También se añadieron 3 ml de antiespumante irradiado sin diluir (SAG 471) al fermentador. Después, antes de la inoculación, se ajustó la temperatura (28 °C), la presión de descarga (0,5 bar), la velocidad de oxigenación (20 litros de aire rociado por minuto) y la velocidad de agitación inicial (300 rpm). El nivel de oxígeno disuelto (OD) en estas condiciones fue del 100 %. La presión de descarga y la velocidad de oxigenación se mantuvieron a un nivel constante durante la fermentación.

La inoculación se realizó añadiendo 15-20 ml de precultivo.

Durante la fase discontinua (0-15 h), la temperatura se mantuvo a 28 °C. El nivel de oxígeno disuelto se estableció al 20 % y se reguló aumentando la agitación cuando el OD disminuyó por debajo del 20 %

Durante la fase discontinua (15-46 h), el pH se mantuvo según una de las condiciones descritas en la tabla 5 mediante la adición de base. La temperatura se reguló a 28 °C. La velocidad de agitación se mantuvo a un valor de referencia constante (máximo 800 rpm), y el nivel de OD se mantuvo al 20% mediante la adición automática de solución de alimentación concentrada (adaptada de Zabriskie et al. (J. Ind. Microbiol. 2: 87-95) (1987)) cuando el OD aumentó por encima del 20%.

Cuando el cultivo alcanzó una DO_{650} nm de aproximadamente 90, el valor de referencia del pH se modificó según una de las condiciones descritas en la tabla 5 por adición de base o ácido y la temperatura se redujo a 23 °C. Una vez alcanzadas estas condiciones, se añadió IPTG a una concentración final de 1 mM. El pH y la temperatura se mantuvieron durante toda la fase de inducción (24 h). Se usó una velocidad de alimentación de sustrato constante durante toda la fase de inducción, según una de las condiciones descritas en la tabla 5. El nivel de OD se mantuvo al 20% controlando la velocidad de agitación.

Al final de la fase de inducción, la pasta celular se recogió por centrifugación (típicamente 6 500xg, 4 °C durante 1 h), y se conservó a -20 °C.

La extracción periplasmática se realizó mediante choque osmótico usando un procedimiento adaptado de Chen et al. (Biochem.Eng. J. 19: 211-215 (2004)). Mediante un ensayo Elisa se analizó el contenido de CRM197 en las fracciones periplasmática y citoplasmática (tabla 6).

Tabla 5

Cultivo nº	pH		Velocidad de alimentación durante la inducción (ml/min)	DO_{650nm}	
	Antes de la inducción	Durante la inducción		En el momento de la inducción	Al final de la fermentación
CDT337	7,0	7,8	1,10	93,0	104,0
CDT338	7,0	7,8	0,28	93,0	102,4
CDT341	7,0	8,7	0,89	94,4	40,0
CDT342	7,0	6,9	0,48	89,2	98,0
CDT344	7,0	6,9	0,89	90,0	89,0
CDT345	7,0	8,7	0,48	92,8	42,0
CDT348	7,0	7,8	0,69	89,2	97,6
CDT349	7,0	7,8	0,69	96,0	109,0

(continuación)

Cultivo nº	pH		Velocidad de alimentación durante la inducción (ml/min)	DO650nm	
	Antes de la inducción	Durante la inducción		En el momento de la inducción	Al final de la fermentación
CDT360	7,4	8,1	0,89	88,4	98,8
CDT351	6,6	7,5	0,48	89,2	99,0
CDT354	6,6	7,5	0,89	89,0	93,6
CDT355	6,6	8,4	0,69	87,0	40,0
CDT357	7,4	7,1	0,48	84,8	88,8
CDT358	7,4	7,2	0,69	84,0	86,4
CDT358	7,0	7,8	0,69	94,8	108,0

Tabla 6

Cultivo nº	pH antes de la inducción	pH durante la inducción	Velocidad de alimentación durante la inducción (ml/min)	CRM197 (mg/l por Elisa)	
				Periplasmática	Citoplasmática
CDT337	7,0	7,8	1,10	1500	422
CDT338	7,0	7,8	0,28	921	357
CDT341	7,0	8,7	0,89	13	11
CDT342	7,0	6,9	0,48	1058	341
CDT344	7,0	6,9	0,89	822	275
CDT345	7,0	8,7	0,48	10	20
CDT348	7,0	7,8	0,69	1166	558
CDT349	7,0	7,8	0,69	889	652
CDT360	7,4	8,1	0,89	77	50
CDT351	6,6	7,5	0,48	1533	427
CDT354	6,6	7,5	0,89	803	595
CDT355	6,6	8,4	0,69	20	32
CDT357	7,4	7,1	0,48	54	29
CDT358	7,4	7,2	0,69	681	310
CDT359	7,0	7,8	0,69	1523	685

- 5 En función de los resultados de las 15 fermentaciones, para modelar la producción de CRM197 en las fracciones periplasmática y citoplasmática, se utilizó el programa informático NEMROD-W (LPRAI, Marsella, Francia).

10 Como se muestra en la Fig. 10, la producción de CRM197 periplasmática es máxima a velocidades de alimentación bajas durante la inducción (Fig. 10a), mientras que la acumulación de CRM197 dentro de la célula es máxima a velocidades de alimentación más altas (Fig. 10b). La diferencia en la velocidad de alimentación óptima para la producción de CRM197 periplasmática o asociada a células, permite definir condiciones que mejoren selectivamente la producción de CRM197 periplasmática. También se requiere un aumento del pH en la inducción para una

producción eficaz de CRM197 periplasmática (Fig.10a). A continuación, en la Tabla 7, se describe el pH óptimo durante el cultivo, el pH óptimo durante la inducción y la velocidad de alimentación óptima, obtenido con el programa informático NEMROD-W.

Tabla 7

Parámetro	Valor
pH durante el cultivo	6,8
pH durante la inducción	7,5
Velocidad de alimentación durante la inducción	0,330 ml min ⁻¹

5

La optimización de las condiciones de pH y velocidad de alimentación, condujo a una secreción óptima de CRM97 en el periplasma.

LISTADO DE SECUENCIAS

10 <110> BLAIS Nomand
 DEHOTTAY Philippe Marc Helene
 DEWERCHIN Marianne
 GOFFIN Philippe
 MARTIN Denis

15 <120> Sistema de expresión
 <130> VB63702
 <150> GB 0917647.0
 <151> 08-10-2009

20 <150> PCT/EP2010/065047
 <151> 07-10-2010
 <160> 45
 <170> FastSEQ for Windows Versión 4.0

25 <210> 1
 <211> 72
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <221> misc_signal
 <222> (0)...(0)

atgaaatttta gtaaaaaata tatagcagct ggatcagctg ttatcgtatc cttgagtcta 60
 tgtgcctatg ca 72

35 <210> 2
 <211> 24
 <212> PRT
 <213> Secuencia artificial

40 <220>
 <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> ADN

<213> Secuencia artificial

<220>

<221> misc_signal

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

<400> 3

atgaaaatga ataaaaaggt actattgaca togacaatgg cagcttcgct attatcagtc 60
gcaagtgttc aagca 75

<210> 4

<211> 25

<212> PRT

<213> Secuencia artificial

<220>

<221> SEÑAL

<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

<210> 5

<211> 63

<212> ADN

<213> Secuencia artificial

<220>

<221> misc_signal

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

<400> 5

atgaaaaaga cagctatcgc gattgcagtg gcaactggctg gtttcgctac cgtagcgcag 60
gcc 63

<210> 6

<211> 21

<212> PRT

<213> Secuencia artificial

<220>

<221> SEÑAL

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

<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

```

<210> 7
<211> 57
<212> ADN
<213> Secuencia artificial

5  <220>
    <221> misc_signal
    <222> (0)...(0)

10 <400> 7
    atgaaaaaag cacttgccac actgattgcc ctcgctctcc cggccgccgc actggcg 57

    <210> 8
    <211> 19
    <212> PRT
    <213> Secuencia artificial

15 <220>
    <221> SEÑAL
    <222> (0)...(0)

    <400> 8
        Met Lys Lys Ala Leu Ala Thr Leu Ile Ala Leu Ala Leu Pro Ala Ala
        1 5 10 15
20 Ala Leu Ala

    <210> 9
    <211> 54
    <212> ADN
    <213> Secuencia artificial

25 <220>
    <221> misc_signal
    <222> (0)...(0)

    <400> 9
30 atgcgcgtac tgctattttt actcttttcc cttttcatgt tgccggcatt ttcg 54

    <210> 10
    <211> 18
    <212> PRT
    <213> Secuencia artificial

35 <220>
    <221> SEÑAL
    <222> (0)...(0)

    <400> 10
        Met Arg Val Leu Leu Phe Leu Leu Leu Ser Leu Phe Met Leu Pro Ala
        1 5 10 15
40 Phe Ser

    <210> 11
    <211> 69
    <212> ADN
    <213> Secuencia artificial

45 <220>
    <221> misc_signal
    <222> (0)...(0)

    <400> 11
        atgatgacta aaataaagtt attgatgctc atttatatttt atttaaatcat ttcggccagc 60
        gcccatgct 69
50

```

<210> 12
 <211> 23
 <212> PRT
 <213> Secuencia artificial
 5 <220>
 <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
 10
 <210> 13
 <211> 75
 <212> ADN
 <213> Secuencia artificial
 15 <220>
 <221> misc_signal
 <222> (0)...(0)
 <400> 13
 atgatgaagc acatgcgtat atgggccggt ctggcatcat ttttagtctt tttttatatt 60
 ccgcagagct atgcc 75
 20
 <210> 14
 <211> 24
 <212> PRT
 <213> Secuencia artificial
 25 <220>
 <221> SEÑAL
 <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
 30
 <210> 15
 <211> 54
 <212> ADN
 <213> Secuencia artificial
 35 <220>
 <221> misc_signal
 <222> (0)...(0)
 <400> 15
 atgaggtttt tattgggcgt gctgatgctg atgatctccg gctcagcgct ggcg 54
 40
 <210> 16
 <211> 18
 <212> PRT
 <213> Secuencia artificial
 45 <220>
 <221> SEÑAL
 <222> (0)...(0)

<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
 <211> 69
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> misc_signal
 <222> (0)...(0)

<400> 17
atgtcaaaac gaacattcgc ggtgatatta accttgttgt gtagcttctg tattggccag 60
gcgcttgca 69

<210> 18
 <211> 23
 <212> PRT
 <213> Secuencia artificial

<220>
 <221> SEÑAL
 <222> (0)...(0)

<400> 18
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
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> misc_signal
 <222> (0)...(0)

<400> 19
atgatgaagc aggcattacg agtagcattt ggttttctca tactgtgggc atcagttctg 60
catgct 66

<210> 20
 <211> 22
 <212> PRT
 <213> Secuencia artificial

<220>
 <221> SEÑAL
 <222> (0)...(0)

<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
 <211> 66

<212> ADN
 <213> Secuencia artificial
 <220>
 <221> misc_signal
 <222> (0)...(0)

5

<400> 21

atgctctcca cactccgcgc cactctattt gcgctgctgg cttgtgcgctc ttttatogtc 60
catgcc 66

10

<210> 22
 <211> 22
 <212> PRT
 <213> Secuencia artificial

15

<220>
 <221> SEÑAL
 <222> (0)...(0)

<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

20

<210> 23
 <211> 57
 <212> ADN
 <213> Secuencia artificial

25

<220>
 <221> misc_signal
 <222> (0)...(0)

<400> 23

atgattaaat ttctctctgc attaatctt ctactggcca cgacggcggc tcaggct 57

30

<210> 24
 <211> 19
 <212> PRT
 <213> Secuencia artificial

35

<220>
 <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

40

<210> 25
 <211> 57
 <212> ADN
 <213> Secuencia artificial

45

<220>
 <221> misc_signal
 <222> (0)...(0)

<400> 25

atgaaaaaga ttggctggc gctggctggt ttagtttag cgtttagcgc atcggcg 57

<210> 26
 <211> 19
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <221> SEÑAL
 <222> (0)...(0)

10

<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
 <211> 39
 <212> ADN
 <213> Secuencia artificial

15

<220>
 <221> primer_bind
 <222> (0)...(0)

20

<400> 27
 gcgcggcata tgggtgcgga tgatgtggtg gatagcagc 39

<210> 28
 <211> 33
 <212> ADN
 <213> Secuencia artificial

25

<220>
 <221> primer_bind
 <222> (0)...(0)

30

<400> 28
 gcggagctcg agttattagc ttttgatttc gaa 33

<210> 29
 <211> 90
 <212> ADN
 <213> Secuencia artificial

35

<220>
 <221> primer_bind
 <222> (0)...(0)

40

<400> 29
ggagcgcata tgattaaatt tctctctgca ttaattcttc tactgggtcac gacggcggct 60
caggctggtg cggatgatgt ggtggatagc 90

<210> 30
 <211> 24
 <212> ADN
 <213> Secuencia artificial

45

<220>
 <221> primer_bind
 <222> (0)...(0)

50

<400> 30
 caccgccgat agttcgcacc cgca 24

<210> 31
 <211> 1605

<212> ADN
<213> *C. diphtheriae*

<400> 31

```

ggtgcggtatg atgtggtgga tagcagcaaa tcttttgtga tggaaaactt tagcagctat 60
catggcacca aaccgggcta tgtggatagc attcagaaag gcatccagaa accgaaaagc 120
ggcaccagg gcaactatga tgatgattgg aaagaatttt atagcaccga taacaaatat 180
gatgcggcgg gttatagcgt ggataacgaa aatccgctgt ctggcaaagc gggcgggtgtg 240
gtgaaagtga cctatccggg cctgaccaa gtgctggccc tgaaagtgga taacgcggaa 300
accatcaaaa aagaactggg cctgagcctg accgaaccgc tgatggaaca ggtgggcacc 360
gaagaattta ttaaacgctt tggcgtatgg gcgagccgtg tggttctgag cctgccgttt 420
gcggaaggca gcagcagcgt ggaatatatt acaactggg aacaggcgaa agccctgagc 480
gtggaactgg aaattaactt tgaaaccogt ggcaaacgtg gccaggatgc gatgtatgaa 540
tacatggcgc aggcgtgcgc gggcaatogt gtgcgtcgta gcgtgggcag cagcctgagc 600
tgcattaacc tggattggga cgtcattogt gataaaacca aaaccaaact cgaaagcctg 660
aaagaacatg gccgatcaa aaacaaaatg agcgaaaagc cgaacaaaac cgtgagcgaa 720
gaaaaagcga aacagtatct ggaagaattt catcagaccg cgctggaaca tccggaaactg 780
agcgaactga aaaccgtgac cggcaccaat ccggtgtttg cgggtgcgaa ctatgcggcg 840
tgggcgggtga atgtggcgca ggtgattgat agcgaaaacc cgataacct ggaaaaaacc 900
accgcggccc tgagcattct gccgggcatt ggcagcgtga tgggcattgc ggatggcgcg 960
gtgcatcata acaccgaaga aattgtggcg cagagcattg ccctgagcag cctgatggtg 1020
gcgcaggcga ttccgctggt tggcgaactg gtggatattg gctttgcggc gtacaacttt 1080
gtggaagca tcatcaacct gtttcagggtg gtgcataaca gctataaccg tccggcgtat 1140
tctccgggtc ataaaaccca gccgtttctg catgatggct atgcggtgag ctggaacacc 1200
gtggaagata gcattattog tacoggcttt cagggcgaaa gcggccatga tattaaaatt 1260
accgcgga aaaccccgct gccgattgcg ggtgttctgc tgccgaccat tccgggcaaa 1320
ctggatgtga acaaaagcaa aaccatatt agcgtgaacg gtcgtaaaat tcgtatgctg 1380
tgccgtgcga ttgatggcga tgtgaacctt tgccgtccga aaagcccggt gtatgtgggc 1440
aacggcgtgc acgcgaacct gcatgtggcg tttcatcgta gcagcagcga aaaaatccat 1500
agcaacgaaa ttagcagcga tagcattggc gtgctgggct atcagaaaac cgtggaccat 1560
accaaagtga actctaaact gagcctgttc ttcgaaatca aaagc 1605

```

5

<210> 32
<211> 535
<212> PRT
<213> *C. diphtheriae*

10

<400> 32

```

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
				165					170					175
Ala	Met	Tyr	Glu	Tyr	Met	Ala	Gln	Ala	Cys	Ala	Gly	Asn	Arg	Val
			180					185					190	
Arg	Ser	Val	Gly	Ser	Ser	Leu	Ser	Cys	Ile	Asn	Leu	Asp	Trp	Asp
		195					200					205		Val
Ile	Arg	Asp	Lys	Thr	Lys	Thr	Lys	Ile	Glu	Ser	Leu	Lys	Glu	His
	210					215					220			Gly
Pro	Ile	Lys	Asn	Lys	Met	Ser	Glu	Ser	Pro	Asn	Lys	Thr	Val	Ser
225					230					235				240
Glu	Lys	Ala	Lys	Gln	Tyr	Leu	Glu	Glu	Phe	His	Gln	Thr	Ala	Leu
				245					250					255
His	Pro	Glu	Leu	Ser	Glu	Leu	Lys	Thr	Val	Thr	Gly	Thr	Asn	Pro
			260					265					270	Val
Phe	Ala	Gly	Ala	Asn	Tyr	Ala	Ala	Trp	Ala	Val	Asn	Val	Ala	Gln
	275					280						285		Val
Ile	Asp	Ser	Glu	Thr	Ala	Asp	Asn	Leu	Glu	Lys	Thr	Thr	Ala	Ala
	290					295					300			Leu
Ser	Ile	Leu	Pro	Gly	Ile	Gly	Ser	Val	Met	Gly	Ile	Ala	Asp	Gly
305					310					315				320
Val	His	His	Asn	Thr	Glu	Glu	Ile	Val	Ala	Gln	Ser	Ile	Ala	Leu
				325					330					335
Ser	Leu	Met	Val	Ala	Gln	Ala	Ile	Pro	Leu	Val	Gly	Glu	Leu	Val
			340					345					350	Asp
Ile	Gly	Phe	Ala	Ala	Tyr	Asn	Phe	Val	Glu	Ser	Ile	Ile	Asn	Leu
		355					360					365		Phe
Gln	Val	Val	His	Asn	Ser	Tyr	Asn	Arg	Pro	Ala	Tyr	Ser	Pro	Gly
	370					375					380			His
Lys	Thr	Gln	Pro	Phe	Leu	His	Asp	Gly	Tyr	Ala	Val	Ser	Trp	Asn
385					390					395				400
Val	Glu	Asp	Ser	Ile	Ile	Arg	Thr	Gly	Phe	Gln	Gly	Glu	Ser	Gly
				405					410					415
Asp	Ile	Lys	Ile	Thr	Ala	Glu	Asn	Thr	Pro	Leu	Pro	Ile	Ala	Gly
			420					425					430	Val
Leu	Leu	Pro	Thr	Ile	Pro	Gly	Lys	Leu	Asp	Val	Asn	Lys	Ser	Lys
		435					440					445		Thr
His	Ile	Ser	Val	Asn	Gly	Arg	Lys	Ile	Arg	Met	Arg	Cys	Arg	Ala
	450					455					460			Ile
Asp	Gly	Asp	Val	Thr	Phe	Cys	Arg	Pro	Lys	Ser	Pro	Val	Tyr	Val
465					470					475				480
Asn	Gly	Val	His	Ala	Asn	Leu	His	Val	Ala	Phe	His	Arg	Ser	Ser
				485					490					495
Glu	Lys	Ile	His	Ser	Asn	Glu	Ile	Ser	Ser	Asp	Ser	Ile	Gly	Val
			500					505					510	Leu
Gly	Tyr	Gln	Lys	Thr	Val	Asp	His	Thr	Lys	Val	Asn	Ser	Lys	Leu
		515					520					525		Ser
Leu	Phe	Phe	Glu	Ile	Lys	Ser								
	530					535								

<210> 33
 <211> 54
 <212> PRT
 <213> Secuencia artificial

5

<220>

<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

<210> 34
 <211> 55
 <212> PRT
 <213> Secuencia artificial

<220>

<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

<210> 35
 <211> 51
 <212> PRT
 <213> Secuencia artificial

<220>

<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
 <211> 49
 <212> PRT
 <213> Secuencia artificial

<220>

<400> 36

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
 <211> 48

<212> PRT

<213> Secuencia artificial

<220>

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

<211> 53

<212> PRT

10 <213> Secuencia artificial

<220>

<400> 38

```

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

```

15 <210> 39

<211> 55

<212> PRT

<213> Secuencia artificial

<220>

20 <400> 39

```

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

25 <212> PRT

<213> Secuencia artificial

<220>

<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

```

30

<210> 41
 <211> 53
 <212> PRT
 <213> Secuencia artificial

5 <220>

<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
    
```

10 <210> 42
 <211> 52
 <212> PRT
 <213> Secuencia artificial

<220>

15 <400> 42

```

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 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
    
```

20 <210> 43
 <211> 52
 <212> PRT
 <213> Secuencia artificial

<220>

<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
    
```

25 <210> 44
 <211> 49
 <212> PRT
 <213> Secuencia artificial

<220>

30 <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

```

<210> 45

<211> 49

<212> PRT

5 <213> Secuencia artificial

<220>

<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

```

10

REIVINDICACIONES

1. Un polinucleótido que comprende una parte de secuencia señal 5' y una parte de toxina 3', en el que
5 (a) la parte de secuencia señal 5' codifica un polipéptido capaz de dirigir el transporte de una proteína heteróloga al periplasma bacteriano, teniendo el polipéptido una secuencia de aminoácidos de (i) SEQ ID NO: 24, (ii) variantes de SEQ ID NO: 24, en el que dichas variantes contienen 1, 2 o 3 mutaciones de sustitución, adición o
10 delección de aminoácidos, en el que una adición de aminoácidos es la adición de un aminoácido, una sustitución de aminoácidos es el reemplazo de un aminoácido por otro aminoácido y una delección de aminoácidos es la eliminación de un aminoácido, o (iii) fragmentos de al menos 10 aminoácidos de SEQ ID NO: 24; y
15 (b) la parte de toxina 3' codifica un polipéptido que tiene una secuencia de aminoácidos al menos 90 % idéntica a la SEQ ID NO: 32.
2. El polinucleótido de la reivindicación 1, en el que la parte de toxina 3' codifica la SEQ ID NO: 32.
3. El polinucleótido de las reivindicaciones 1-2, en el que la parte de toxina 3' comprende la SEQ ID NO: 31.
4. El polinucleótido de las reivindicaciones 1-3, que codifica un polipéptido que comprende la SEQ ID NO: 44.
5. El polinucleótido de las reivindicaciones 1-4, en el que la parte de secuencia señal 5' está directamente en 5' de la
15 parte de toxina 3'.
6. Un vector que comprende el polinucleótido de las reivindicaciones 1-5 unido a un promotor inducible.
7. El vector de la reivindicación 6, en el que el promotor inducible se activa mediante la adición de una cantidad suficiente de IPTG.
8. Una célula hospedadora que comprende el vector de expresión de la reivindicación 6 o 7 o el polinucleótido de
20 una cualquiera de las reivindicaciones 1-5.
9. La célula hospedadora de la reivindicación 8, en la que la célula hospedadora es una bacteria gram negativa.
10. La célula hospedadora de las reivindicaciones 8 o 9, en la que la célula hospedadora se selecciona del grupo que consiste en *E. coli*, *Acinetobacter*, *Actinobacillus*, *Bordetella*, *Brucella*, *Campylobacter*, *Cyanobacteria*,
25 *Enterobacter*, *Erwinia*, *Franciscella*, *Helicobacter*, *Hemophilus*, *Klebsiella*, *Legionella*, *Moraxella*, *Neisseria*, *Pasteurella*, *Proteus*, *Salmonella*, *Serratia*, *Shigella*, *Treponema*, *Vibrio* y *Yersinia*.
11. La célula hospedadora de las reivindicaciones 8-10, en la que la célula hospedadora es una célula de *Escherichia coli*.
12. Un procedimiento para la fabricación de una toxina bacteriana que comprende las etapas de a) cultivar un cultivo de la célula hospedadora bacteriana de las reivindicaciones 8-11 y b) inducir la expresión del polipéptido de manera
30 que un toxoide bacteriano se exprese de manera periplasmática.

Figura 1

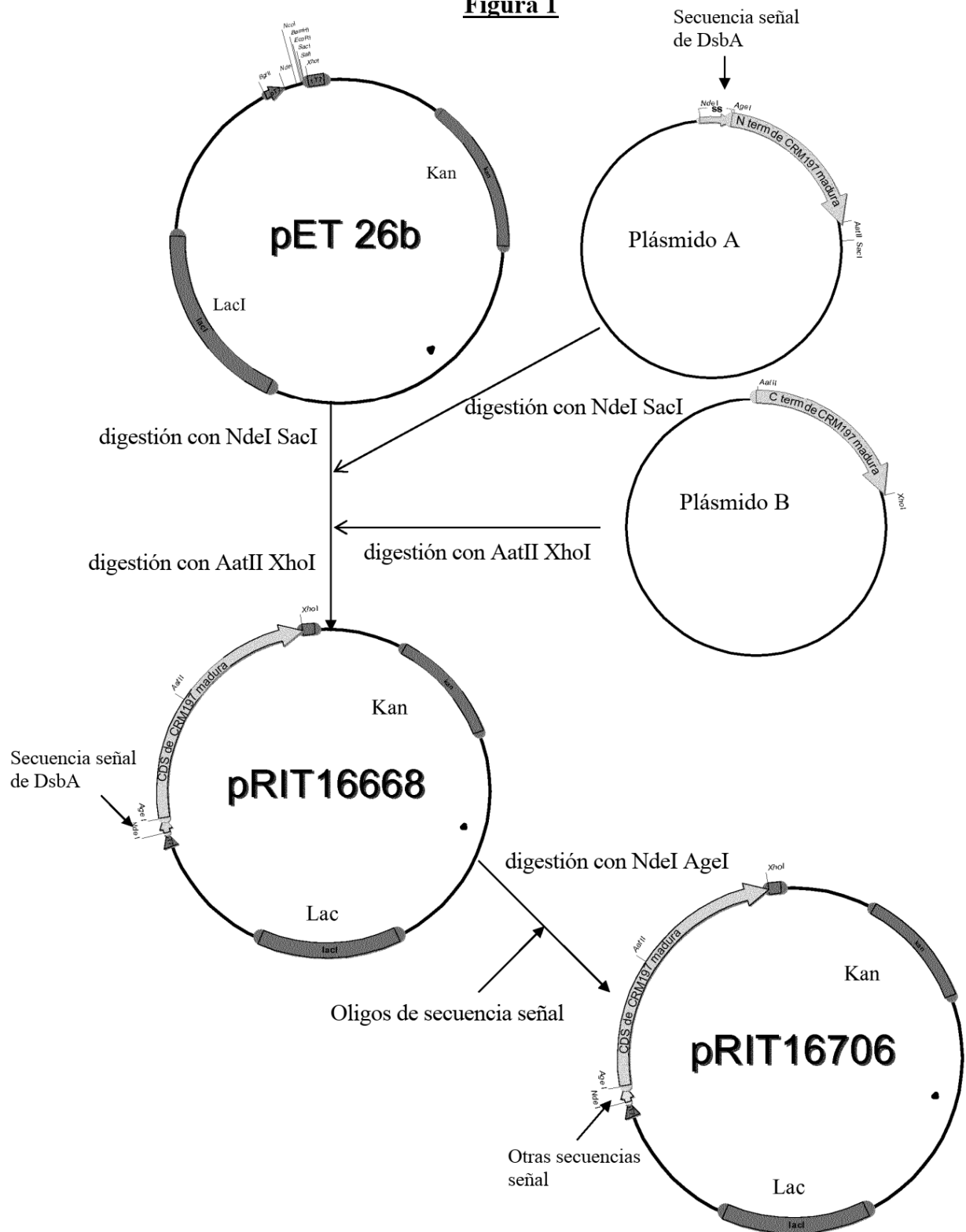


Figura 2

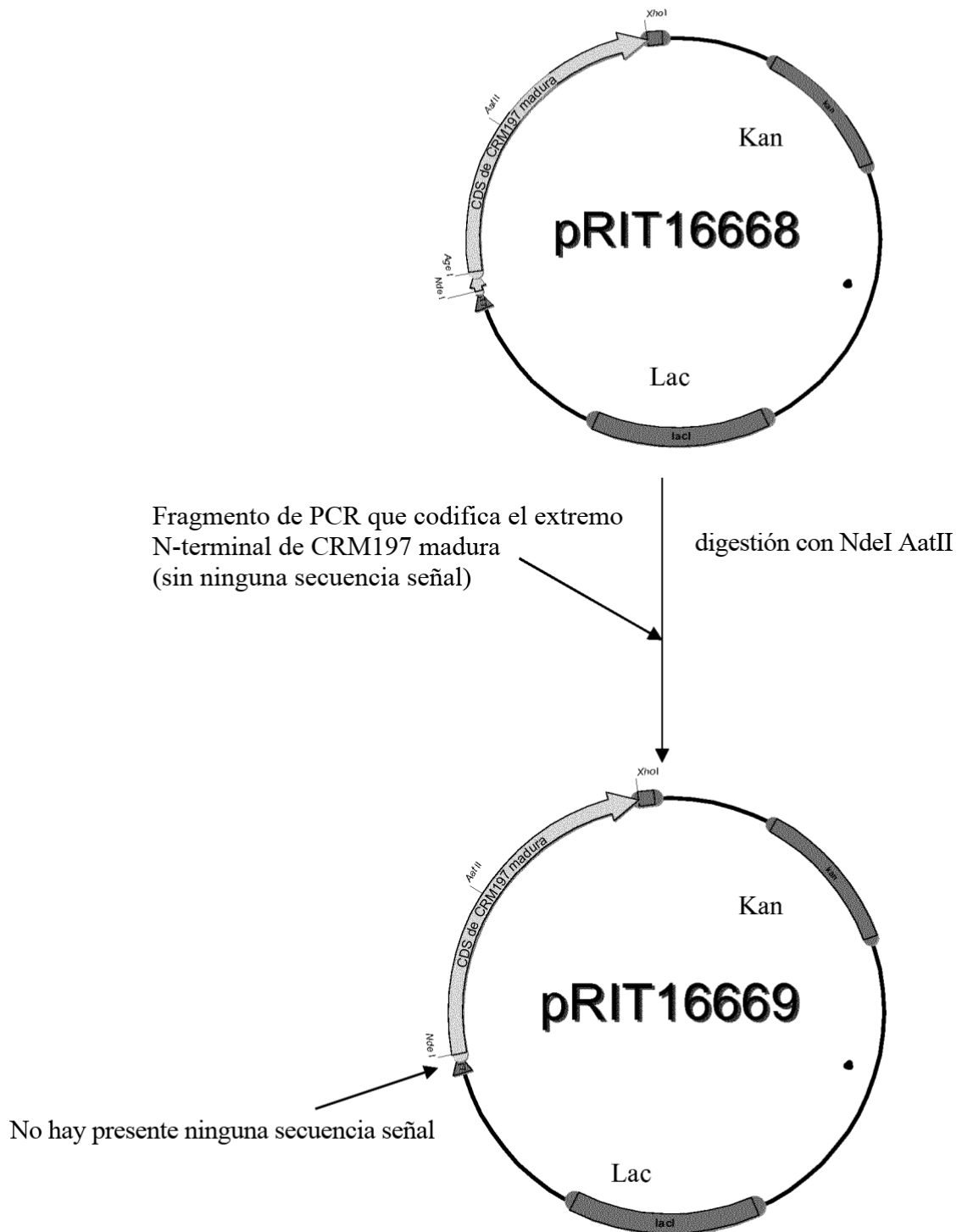


Figura 3

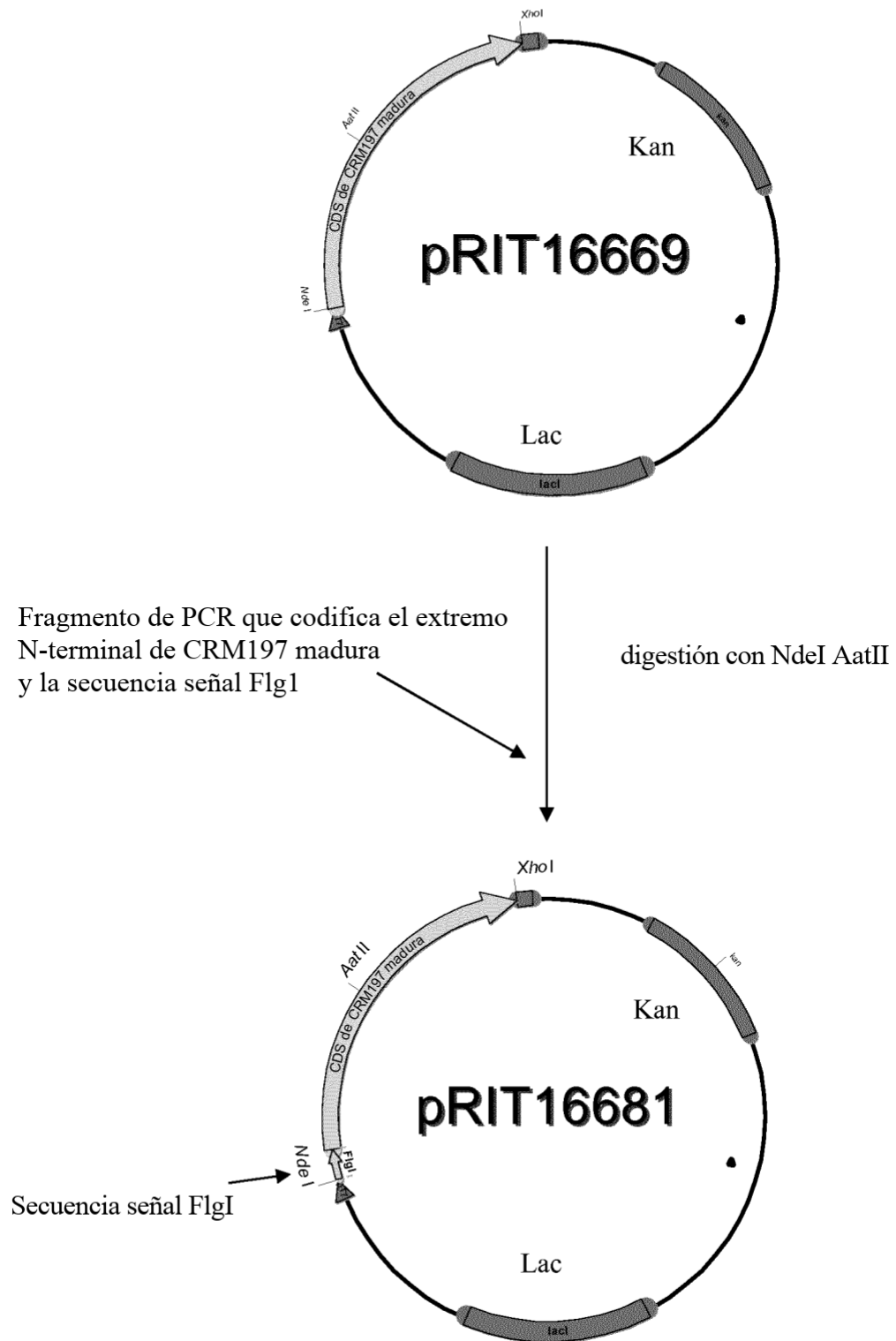


Figura 4

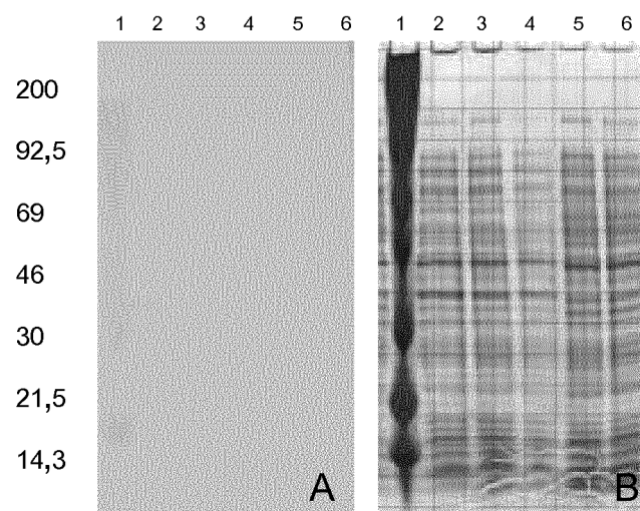
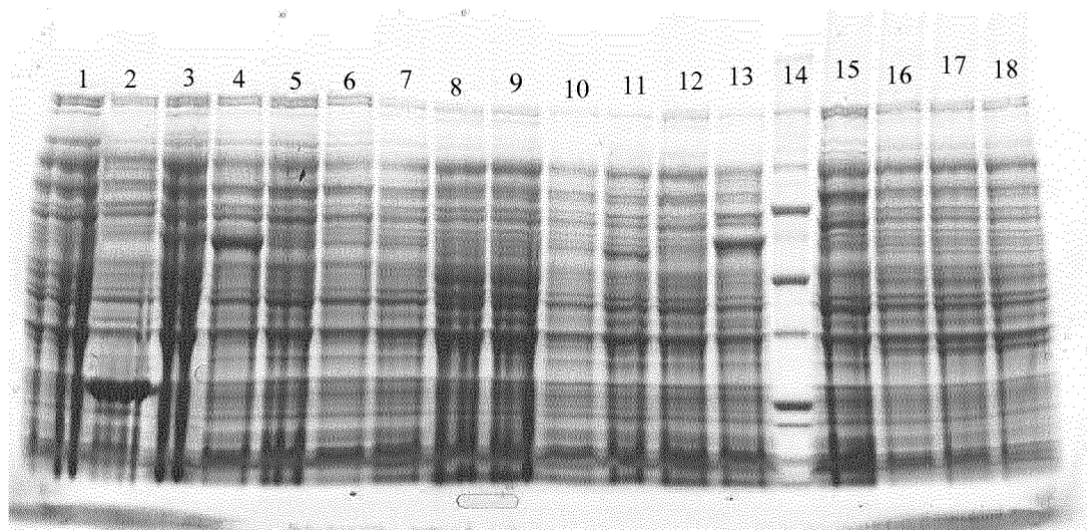
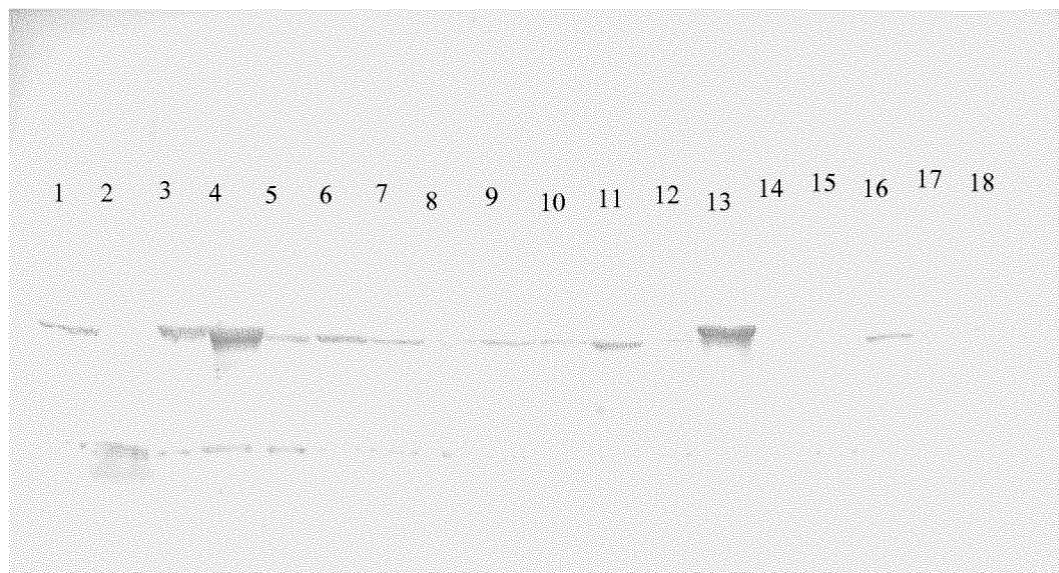


Figura 5

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



A



B

Figura 6

Figura 5: Geles que muestran inducción de CRM197 3 horas a 30° C en diferentes cepas y solubilidad de la proteína expresada

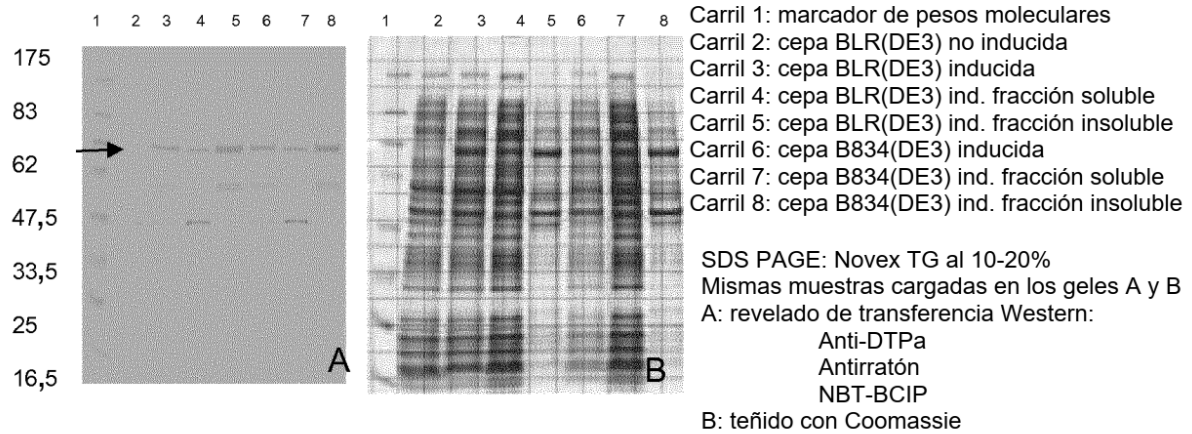


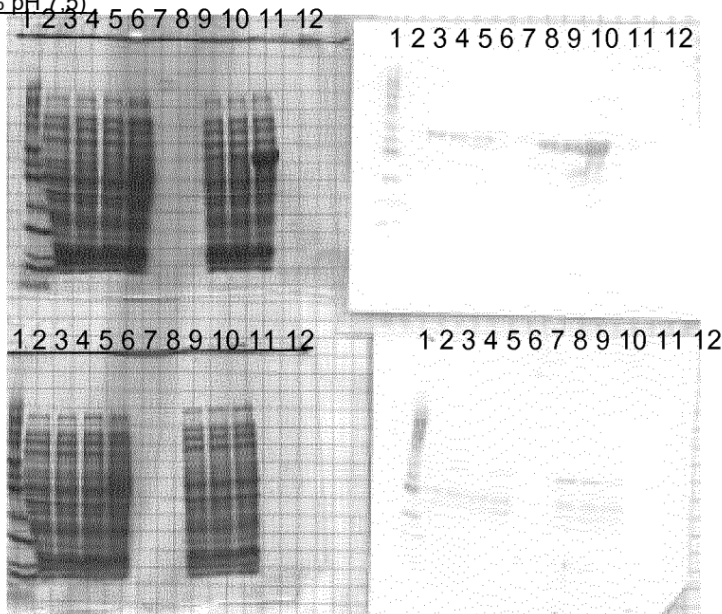
Figura 7

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

Medio tamponado con LBT

(adición de tampón fosfato de potasio 1M al 10% pH 7.5)

1. Marcador de peso molecular
2. Extracto celular a T0 a 30°C
3. Extracto celular a T2 a 30°C
4. Extracto celular a T4 a 30°C
5. Extracto celular tras inducción de una noche a 30°C
6. Medio a T4 a 30°C
7. Medio tras inducción de una noche a 30°C
8. Extracto celular a T2 a 23°C
9. Extracto celular a T4 a 23°C
10. Extracto celular tras inducción de una noche a 23°C
11. Medio a T4 a 23°C
12. Medio tras inducción de una noche a 23°C



Medio no tamponado con LBT

1. Marcador de peso molecular
2. Extracto celular a T0 a 30°C
3. Extracto celular a T2 a 30°C
4. Extracto celular a T4 a 30°C
5. Extracto celular tras inducción de una noche a 30°C
6. Medio a T4 a 30°C
7. Medio tras inducción de una noche a 30°C
8. Extracto celular a T2 a 23°C
9. Extracto celular a T4 a 23°C
10. Extracto celular tras inducción de una noche a 23°C
11. Medio a T4 a 23°C
12. Medio tras inducción de una noche a 23°C

Gel teñido con Coomassie

Transferencia Western anti DTPa

Figura 8

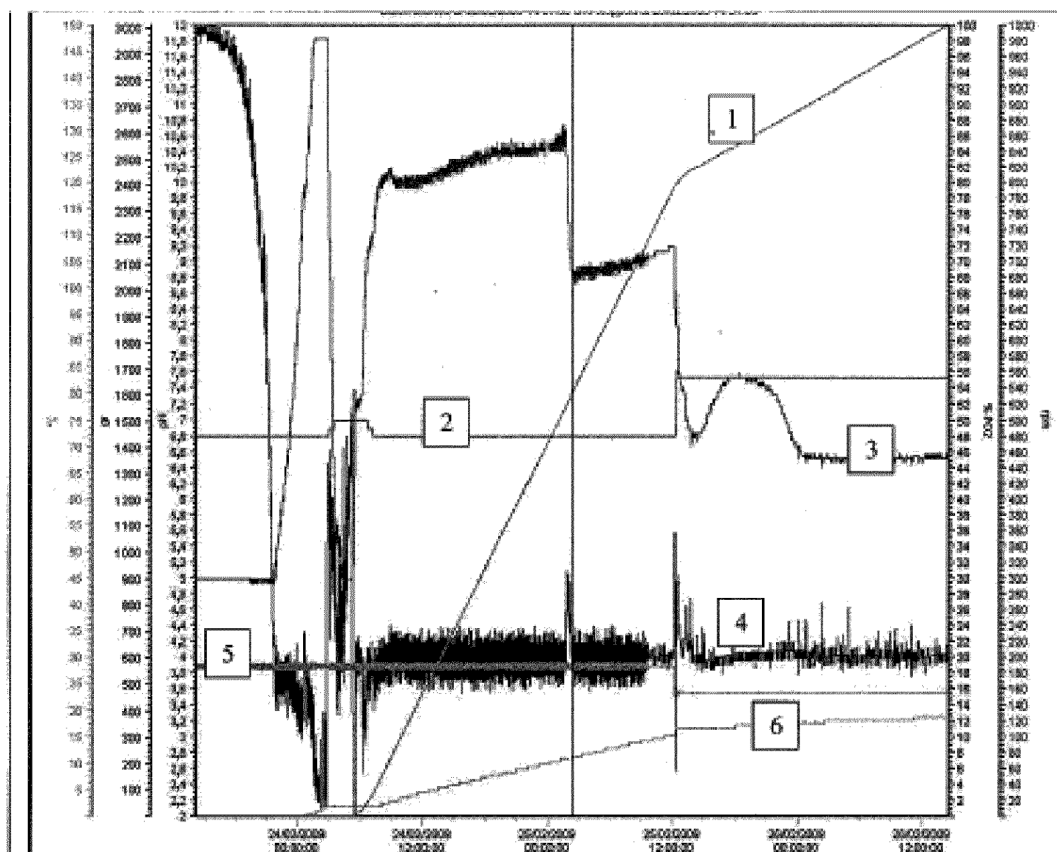


Figura 9

SEQ ID NO 1: secuencia de nucleótidos de la secuencia señal PhtE

ATGAAATTTAGTAAAAAATATATAGCAGCTGGATCAGCTGTTATCGTATC
CTTGAGTCTATGTGCCTATGCA

SEQ ID NO 2: secuencia de aminoácidos de la secuencia señal PhtE

MKFSKKYIAAGSAVIVSLSLCAYA

SEQ ID NO 3: secuencia de nucleótidos de la secuencia señal SipA

ATGAAAATGAATAAAAAGGTACTATTGACATCGACAATGGCAGCTTCGCT
ATTATCAGTCGCAAGTGTTCAAGCA

SEQ ID NO 4: secuencia de aminoácidos de la secuencia señal SipA

MKMNKKVLLTSTMAASLLSVASVQA

SEQ ID NO 5: secuencia de nucleótidos de la secuencia señal OmpA

ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTAC
CGTAGCGCAGGCC

SEQ ID NO 6: secuencia de aminoácidos de la secuencia señal OmpA

MKKTAIAIAVALAGFATVAQA

SEQ ID NO 7: secuencia de nucleótidos de la secuencia señal NspA

ATGAAAAAAGCACTTGCCCACTGATTGCCCTCGCTCTCCCGGCCGCCGC
ACTGGCG

SEQ ID NO 8: secuencia de aminoácidos de la secuencia señal NspA

MKKALATLIALALPAAALA

SEQ ID NO 9: secuencia de nucleótidos de la secuencia señal TorT

ATGCGCGTACTGCTATTTTTACTTCTTTCCCTTTTCATGTTGCCGGCATTTC
CG

SEQ ID NO 10: secuencia de aminoácidos de la secuencia señal TorT

MRVLLFLLLSLFMLPAFS

SEQ ID NO 11: secuencia de nucleótidos de la secuencia señal SfmC

ATGATGACTAAAATAAAGTTATTGATGCTCATTATATTTTATTTAATCATT
TCGGCCAGCGCCCATGCT

SEQ ID NO 12: secuencia de aminoácidos de la secuencia señal SfmC

MMTKIKLLMLIIFYLIISASAH

SEQ ID NO 13: secuencia de nucleótidos de la secuencia señal FocC

ATGATGAAGCACATGCGTATATGGGCGTTCTGGCATCATTTTATAGTCTTT
TTTATATTCCGCAGAGCTATGCC

SEQ ID NO 14: secuencia de aminoácidos de la secuencia señal FocC

MMKHMRIWAVLASFLVFFYIPQSYA

SEQ ID NO 15: secuencia de nucleótidos de la secuencia señal CcmH

ATGAGGTTTTTATTGGGCGTGCTGATGCTGATGATCTCCGGCTCAGCGCTG
GCG

SEQ ID NO 16: secuencia de aminoácidos de la secuencia señal CcmH

MRFLLGVLMLMISGSALA

SEQ ID NO 17: secuencia de nucleótidos de la secuencia señal YraI

ATGTCAAACGAACATTCGCGGTGATATTAACCTTGTTGTGTAGCTTCTGT
ATTGGCCAGGCGCTTGCA

SEQ ID NO 18: secuencia de aminoácidos de la secuencia señal YraI

MSKRTFAVILTLLCSFCIGQALA

SEQ ID NO 19: secuencia de nucleótidos de la secuencia señal TolB

ATGATGAAGCAGGCATTACGAGTAGCATTTGGTTTTCTCATACTGTGGGCA
TCAGTTCTGCATGCT

SEQ ID NO 20: secuencia de aminoácidos de la secuencia señal TolB
MMKQALRVAFGFLILWASVLHA

SEQ ID NO 21: secuencia de nucleótidos de la secuencia señal NikA
ATGCTCTCCACACTCCGCCGCACTCTATTTGCGCTGCTGGCTTGTGCGTCT
TTTATCGTCCATGCC

SEQ ID NO 22: secuencia de aminoácidos de la secuencia señal NikA
MLSTLRRTLFALLACASFIVHA

SEQ ID NO 23: secuencia de nucleótidos de la secuencia señal FlgI
ATGATTAAATTTCTCTCTGCATTAATTCTTCTACTGGTCACGACGGCGGCT
CAGGCT

SEQ ID NO 24: secuencia de aminoácidos de la secuencia señal FlgI
MIKFLSALILLVTTAAQA

SEQ ID NO 25: secuencia de nucleótidos de la secuencia señal DsbA
ATGAAAAAGATTTGGCTGGCGCTGGCTGGTTTAGTTTTAGCGTTTAGCGCA
TCGGCG

SEQ ID NO 26: secuencia de aminoácidos de la secuencia señal DsbA
MKKIWLALAGLVLAFSASA

SEQ ID NO:27: secuencia de ADN del cebador MDSCRM1
5' GCGCGGCATATGGGTGCGGATGATGTGGTGGATAGCAGC 3'

SEQ ID NO:28: secuencia de ADN del cebador MDSCRM2
5' GCGGAGCTCGAGTTATTAGCTTTTGATTTCGAA 3'

SEQ ID NO:29: secuencia de ADN 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
ggg gcg gat gat gtg gtg gat agc 3'

SEQ ID NO:30: secuencia de ADN del cebador MDSCRM906NC

MDSCRM906NC primer

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

SEQ ID NO 31: secuencia de nucleótidos madura de la región codificante de CRM197 optimizada con codones

GGTGCGGATGATGTGGTGGATAGCAGCAAATCTTTTGTGATGGAAACTTTAGCAGCTATCATGGCACCAAACCGGGCTA
TGTGGATAGCATTGAGAAAGGCATCCAGAAACCGAAAAGCGGCACCCAGGGCAACTATGATGATGATTGGAAAGAATTTT
ATAGCACCGGATAACAAATATGATGCGGCGGGTTATAGCGTGGATAACGAAAATCCGCTGTCTGGCAAAGCGGGCGGTGTG
GTGAAAGTGACCTATCCGGGCGCTGACCAAAGTGCTGGCCCTGAAAGTGATAACGCGGAAACCATCAAAAAAGAACTGGG
CCTGAGCCTGACCGAACCGCTGATGGAACAGGTGGGCACCGAAGAATTTATTAAACGCTTTGGCGATGGCGCGAGCCGTG
TGGTCTGAGCCTGCCGTTTTCGGAAGGCAGCAGCAGCGTGGAAATATATTAACAACCTGGGAACAGGCGAAGCCCTGAGC
GTGGAACCTGGAATTAACCTTTGAAACCCGTGGCAAACGTGCCAGGATGCGATGTATGAATACATGGCGCAGGCGTGCGC
GGGCAATCGTGTGCGTCGTAGCGTGGGCAGCAGCCTGAGCTGCATTAACTGGATTGGGACGTCATTCGTGATAAAACCA
AAACCAAAATCGAAAGCCTGAAAGAACATGGCCCGATCAAAAACAAAATGAGCGAAAGCCCGAACAACAAACCGTGAGCGAA
GAAAAAGCGAAACAGTATCTGGAAGAATTTTCATCAGACCGCGCTGGAACATCCGGAACCTGAGCGAACTGAAAACCGTGAC
CGGCACCAATCCGGTGTTCGCGGTGCGAACTATGCGCGCTGGGCGGTGAATGTGGCGCAGGTGATTGATAGCGAAACCG
CGGATAACCTGGAAAAAACACCGCGGCCCTGAGCATTTCGCCGGGCATTGGCAGCGTGATGGGCATTGCGGATGGCGCG
GTGCATCATAAACACCGAAGAAATTTGTGGCGCAGAGCATTGCCCTGAGCAGCCTGATGGTGGCGCAGGCGATTCCGCTGGT
TGGCGAACTGGTGGATATTGGCTTTCGCGCGTACAACTTTGTGGAAGCATCATCAACCTGTTTCAGGTGGTGCATAACA
GCTATAACCGTCCGCGTATTCTCCGGTGCATAAAACCCAGCCGTTTCTGCATGATGGCTATGCGGTGAGCTGGAACACC
GTGGAAGATAGCATTATTCGTACCGGCTTTCAGGGCGAAAGCGCCATGATATTAAATTAACCGCGAAACACCCCGCT
GCCGATTGCGGGTGTTCGTGCTGCCGACCATTCGGGCAAACTGGATGTGAACAAAGCAAAACCCATATTAGCGTGAACG
GTCGTAATAATTCGTATGCGTTGCCGTGCGATTGATGGCGATGTGACCTTTTGCCGTCCGAAAAGCCCGGTGTATGTGGC
AACGGCGTGACCGCAACCTGCATGTGGCGTTTCATCGTAGCAGCAGCGAAAAATCCATAGCAACGAAATAGCAGCGA
TAGCATTTGGCGTGCTGGGCTATCAGAAAACCGTGGACCATAACCAAGTGAACCTCTAACTGAGCCTGTCTTCGAAATCA
AAAGC

SEQ ID NO 32: secuencia de aminoácidos de CRM197 madura

GADDVVDSSKSFVMENFSSYHGTPGYVDS

IQKGIQPKPGSGTQGNYYDDWKEFYSTDNKY

DAAGYSVDNENPLSGKAGGVVKVTYPGLTK

VLALKVDNAETIKKELGLSLTEPLMEQVGT

EEFIKRFGDGASRVVLSLPFAEGSSSVEYI

NNWEQAKALSVELEINFETRGRGQDAMYE

YMAQACAGNRVRRSVGSSLSCINLDWDVIR

DKTKTKIESLKEHGPIKNKMSESPNKTVSE

EKAKQYLEEFHQTALEHPELSELKTVTGTN
PVFAGANYAAWAVNVAQVIDSETADNLEKT
TAALSILPGIGSVMGIADGAVHHNTEEIVA
QSIALLSLMVAQAIPLVGELVDIGFAAYNF
VESIINLFQVVHNSYNRPAYSPGHKTQPFL
HDGYAVSWNTVEDSIIRTGFQGESGHDIKI
TAENTPLPIAGVLLPTIPGKLDVNKSKTHI
SVNGRKRIRMRCRAIDGDVTFCRPKSPVYVG
NGVHANLHVAFHRSSSEKIHHSNEISSDSIG
VLGYQKTVDHTKVNSKLSLFFEIKS.

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

MKFSKKYIAAGSAVIVSLSLCAYAGADDVVDSSKSFVMENFSSYHGTPKPGYV
DS

SEQ ID NO 34: secuencia de aminoácidos de la secuencia de señal Sip y los primeros 30 aminoácidos de CRM197

MKMNKKVLLTSTMAASLLSVASVQAGADDVVDSSKSFVMENFSSYHGTPK
GYVDS

SEQ ID NO 35: secuencia de aminoácidos de la secuencia de señal OmpA y los primeros 30 aminoácidos de CRM197

MKKTAIAIAVALAGFATVAQAGADDVVDSSKSFVMENFSSYHGTPKPGYVDS

SEQ ID NO 36: secuencia de aminoácidos de la secuencia de señal NspA y los primeros 30 aminoácidos de CRM197

MKKALATLIALALPAAALAGADDVVDSSKSFVMENFSSYHGTPKPGYVDS

SEQ ID NO 37: Amino acid sequence of the TorT signal sequence and the first 30 amino acids of CRM197

MRVLLFLLLSLFMLPAFSGADDVVDSSKSFVMENFSSYHGTPKPGYVDS

SEQ ID NO 38: secuencia de aminoácidos de la secuencia de señal SfmC y los primeros 30 aminoácidos de CRM197

MMTKIKLLMLIIFYLIISASAHAGADDVVDSSKSFVMENFSSYHGTPKPGYVDS

SEQ ID NO 39: secuencia de aminoácidos de la secuencia de señal FocC y los primeros 30 aminoácidos de CRM197

MMKHMRIWAVLASFLVFFYIPQSYAGADDVVDSSKSFVMENFSSYHGTPKPGYVDS

SEQ ID NO 40: secuencia de aminoácidos de la secuencia de señal CcmH y los primeros 30 aminoácidos de CRM197

MRFLLGVLMLMISGSALAGADDVVDSSKSFVMENFSSYHGTPKPGYVDS

SEQ ID NO 41: secuencia de aminoácidos de la secuencia de señal YraI y los primeros 30 aminoácidos de CRM197

MSKRTFAVILTLLCSFCIGQALAGADDVVDSSKSFVMENFSSYHGTPKPGYVDS

SEQ ID NO 42: secuencia de aminoácidos de la secuencia de señal TolB y los primeros 30 aminoácidos de CRM197

MMKQALRVAFGFLILWASVLHAGADDVVDSSKSFVMENFSSYHGTPKPGYVDS

SEQ ID NO 43: secuencia de aminoácidos de la secuencia de señal NikA y los primeros 30 aminoácidos de CRM197

MLSTLRRTLFALLACASFIVHAGADDVVDSSKSFVMENFSSYHGTPGYVDS

SEQ ID NO 44: secuencia de aminoácidos de la secuencia de señal FlgI y los primeros 30 aminoácidos de CRM197

MIKFLSALILLVTTAAQAGADDVVDSSKSFVMENFSSYHGTPGYVDS

SEQ ID NO 45: secuencia de aminoácidos de la secuencia de señal DsbA y los primeros 30 aminoácidos de CRM197

MKKIWLALAGLVLAFSASAGADDVVDSSKSFVMENFSSYHGTPGYVDS

Figura 10

