

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 627 826**

51 Int. Cl.:

**C07K 14/245** (2006.01)

**C12N 15/70** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **13.01.2011 PCT/EP2011/050416**

87 Fecha y número de publicación internacional: **21.07.2011 WO11086139**

96 Fecha de presentación y número de la solicitud europea: **13.01.2011 E 11701045 (4)**

97 Fecha y número de publicación de la concesión europea: **22.03.2017 EP 2523969**

54 Título: **Cepa bacteriana hospedadora que expresa DsbC recombinante y que tiene una actividad Tsp reducida**

30 Prioridad:

**14.01.2010 GB 201000591**

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

**31.07.2017**

73 Titular/es:

**UCB BIOPHARMA SPRL (100.0%)  
Allée de La Recherche 60  
1070 Brussels, BE**

72 Inventor/es:

**ELLIS, MARK y  
HUMPHREYS, DAVID, PAUL**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 627 826 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

Cepa bacteriana hospedadora que expresa DsbC recombinante y que tiene una actividad Tsp reducida

La invención se refiere a una cepa bacteriana hospedadora recombinante, en particular *E. coli*. La invención también se refiere a un método para producir una proteína de interés en una célula de este tipo.

5 **Antecedentes de la invención**

Las células bacterianas tales como *E. coli*, se utilizan comúnmente para la producción de proteínas recombinantes. Existen muchas ventajas del uso de células bacterianas, tales como *E. coli*, para la producción de proteínas recombinantes, en particular debido a la naturaleza versátil de las células bacterianas como células hospedadoras que permiten la inserción de genes a través de plásmidos. *E. coli* se ha utilizado para producir muchas proteínas recombinantes incluyendo la insulina humana.

10

A pesar de las muchas ventajas de utilizar células bacterianas para producir proteínas recombinantes, todavía existen limitaciones significativas, incluyendo la dificultad de producir proteínas sensibles a las proteasas. Las proteasas desempeñan un papel importante en el recambio de proteínas viejas, dañadas o mal plegadas en el periplasma y el citoplasma de *E. coli*. Las proteasas bacterianas actúan para degradar la proteína recombinante de interés, lo que con frecuencia reduce significativamente el rendimiento de proteína activa.

15

Se ha identificado una serie de proteasas bacterianas. En *E. coli* las proteasas que incluyen Proteasa III (ptr), DegP, OmpT, Tsp, prlC, ptrA, ptrB, pepA-T, tsh, espc, eatA, clpP e lon han sido identificadas.

Tsp (conocida también como Prc) es una proteasa periplásmica de 60 kDa. El primer sustrato conocido de Tsp fue la proteína-3 que se une a la penicilina (PBP3) (Determination of the cleavage site involved in C-terminal processing of penicillin-binding protein 3 of *Escherichia coli*; Nagasawa H., Sakagami Y., Suzuki A., Suzuki H., Hara H., Hirota Y. *J. Bacteriol.* Nov. de 1989; 171(11):5890-3 y Cloning, mapping and characterization of the *Escherichia coli* Tsp gene which is involved in C-terminal processing of penicillin-binding protein 3; Hara H, Yamamoto Y, Higashitani A, Suzuki H, Nishimura Y. *J Bacteriol.* 1991 Ago; 173 (15):4799-813) pero más tarde se descubrió que Tsp también era capaz de escindir las proteínas de la cola del fago y, por lo tanto, se renombró como Proteasa Específica de la Cola (Tsp, del inglés "Tail Specific Protease") (Silber *et al.*, *Proc. Natl. Acad. Sci. USA*, 89: 295-299 (1992)). Silber *et al.* (Deletion of the prc(tsp) gene provides evidence for additional tail-specific proteolytic activity in *Escherichia coli* K-12; Silber, K.R., Sauer, R.T.; *Mol Gen Genet* 1994 242:237-240) describen una cepa con delección de prc (KS1000) en la que la mutación se creó mediante la sustitución de un segmento del gen prc por un fragmento que comprende un marcador de Kan<sup>r</sup>.

20

25

30

La reducción de la actividad de Tsp (prc) es deseable para reducir la proteólisis de proteínas de interés. Sin embargo, se descubrió que las células que carecen de proteasa prc presentan un crecimiento termosensible con osmolaridad reducida. Hara *et al.* aislaron inversores termorresistentes que contienen mutaciones supresoras (spr) extragénicas (Hara *et al.*, *Microbial Drug Resistance*, 2: 63-72 (1996)). Spr es una proteasa periplásmica unida a la membrana de 18 kDa y los sustratos de spr son Tsp y peptidoglucanos en la membrana externa involucrados en la hidrólisis de la pared celular durante la división celular. El gen spr se designa UniProtKB/Swiss-Prot POAFV4 (SPR\_ECOLI).

35

Se han descrito cepas carentes de proteasa mejoradas que comprenden el gen spr mutante. Chen *et al.* (Chen C., Snedecor B., Nishihara J.C., Joly J.C., McFarland N., Andersen D.C., Battersby J.E., Champion K.M. *Biotechnol Bioeng.* 2004 Mar 5;85(5):463-74) describen la construcción de cepas de *E. coli* que son portadoras de diferentes combinaciones de mutaciones en prc (Tsp) y otra proteasa, DegP, creada mediante la amplificación de las regiones aguas arriba y aguas abajo del gen y ligándolas entre sí en un vector que comprende marcadores de selección y una mutación sprW174R (la acumulación de alto nivel de un fragmento de anticuerpo recombinante en el periplasma de *Escherichia coli* requiere una cepa hospedadora mutante triple ( $\Delta$ DegP  $\Delta$ prc sprW174R)). Se encontró que la combinación de las mutaciones  $\Delta$ DegP,  $\Delta$ prc y sprW174R proporcionaba los niveles más altos de cadena ligera de anticuerpo, cadena pesada de anticuerpo y F(ab')<sub>2</sub>-LZ. El documento EP1341899 da a conocer una cepa de *E. coli* que carece de las proteasas que codifican DegP y prc cromosómicas, respectivamente, y alberga un gen spr mutante que codifica una proteína que suprime los fenotipos de crecimiento mostrados por las cepas que albergan mutantes de prc.

40

45

La proteína disulfuro isomerasa es una enzima que cataliza la formación y la rotura de enlaces disulfuro entre residuos de cisteína dentro de las proteínas cuando se pliegan. Es conocida la coexpresión de proteínas que catalizan la formación de enlaces disulfuro para mejorar la expresión de proteínas en una célula hospedadora. El documento WO98/56930 describe un método para producir polipéptidos que contienen enlaces disulfuro heterólogos en células bacterianas, en donde una disulfuro isomerasa procariota, tal como DsbC o DsbG es coexpresada con un polipéptido eucariota. El documento US6673569 describe un operón artificial que comprende polinucleótidos que codifican cada una de DsbA, DsbB, DsbC y DsbD para uso en la producción de una proteína extraña. El documento EP0786009 describe un procedimiento para producir un polipéptido heterólogo en bacterias, en el que la expresión de ácido nucleico que codifica DsbA o DsbC se induce antes de la inducción de la expresión de ácido nucleico que codifica el polipéptido heterólogo.

50

55

DsbC es una proteína procariota que se encuentra en el periplasma de *E. coli*, la cual cataliza la formación de enlaces disulfuro en *E. coli*. DsbC tiene una longitud de la secuencia de aminoácidos de 236 (incluido el péptido señal) y un peso molecular de 25,6 KDa (UniProt nº P0AEG6). DsbC se identificó en primer lugar en 1994 (Missiakas *et al.* The *Escherichia coli* dsbC (xprA) gene encodes a periplasmic protein involved in disulfide bond formation, The EMBO Journal vol. 13, nº 8, págs. 2013-2020, 1994 y Shevchik *et al.* Characterization of DsbC, a periplasmic protein of *Erwinia chrysanthemi* and *Escherichia coli* with disulfide isomerase activity, The EMBO Journal vol. 13, nº 8, págs. 2007-2012, 1994).

Se ha encontrado sorprendentemente que la hiperexpresión de DsbC a partir de DsbC recombinante en una célula bacteriana gram-negativa, mejora el fenotipo de lisis celular de células que carecen de la proteasa Tsp. Por consiguiente, los presentes inventores han proporcionado una nueva cepa que tiene propiedades ventajosas para la producción de una proteína de interés.

### Compendio de la invención

La presente invención proporciona una célula bacteriana gram-negativa recombinante, caracterizada porque la célula:

- a) comprende un vector de expresión que comprende un polinucleótido recombinante que codifica DsbC;
- b) tiene un gen Tsp mutado que codifica una proteína Tsp que tiene una reducción de la actividad proteasa en comparación con una célula de tipo silvestre,
- c) comprende un gen spr mutado que codifica una proteína spr mutante capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado, y
- d) tiene un genoma que es isogénico al de la cepa de *E. coli* de tipo silvestre W3110, excepto por el gen Tsp mutado y el gen spr mutado.

La célula bacteriana gram-negativa que tiene la anterior combinación específica de modificaciones genéticas muestra fenotipos de crecimiento ventajoso y de producción de proteínas.

La presente invención también proporciona un método para producir una proteína de interés que comprende expresar la proteína de interés en una célula bacteriana gram-negativa recombinante tal como se ha definido anteriormente. El método comprende cultivar una célula bacteriana gram-negativa recombinante como se ha definido anteriormente en un medio de cultivo en condiciones eficaces para expresar la proteína de interés y el polinucleótido recombinante que codifica DsbC; y recuperar la proteína de interés a partir del periplasma de la célula bacteriana gram-negativa recombinante y/o del medio de cultivo.

### Breve descripción de los dibujos

La Figura 1 muestra el perfil de crecimiento de cepas MXE001 y MXE008 que expresan Fab' anti-TNF y de cepas MXE001 y MXE008 que expresan Fab' anti-TNF $\alpha$  y DsbC recombinante.

La Figura 2 muestra el rendimiento de Fab' anti-TNF a partir del periplasma (símbolos sombreadas) y el material sobrenadante (símbolos blancos sin sombrear) de cepas MXE001, MXE008 y MXE009 de *E. coli* que expresan Fab' anti-TNF y DsbC recombinante.

La Figura 3 muestra el rendimiento de Fab' anti-TNF a partir del periplasma (símbolos sombreados) y el material sobrenadante (símbolos blancos sin sombrear) de cepas MXE001 y MXE008 de *E. coli* que expresan Fab' anti-TNF y cepas MXE001 y MXE008 de *E. coli* que expresan Fab' anti-TNF y DsbC recombinante.

La Figura 4 muestra el rendimiento de Fab' anti-TNF a partir del periplasma de la cepa W3110 de *E. coli* que expresa Fab A y Fab B y de la cepa MXE008 de *E. coli* que expresa Fab A y DsbC recombinante o Fab B y DsbC recombinante.

La Figura 5a muestra el extremo 5' de la proteína ptr de tipo silvestre (proteasa III) y la proteína ptr mutada con el gen inactivado (proteasa III) y las secuencias génicas.

La Figura 5b muestra el extremo 5' de la proteína Tsp de tipo silvestre y Tsp mutada con el gen inactivado mutado y las secuencias génicas.

La Figura 5c muestra una región de la proteína DegP de tipo silvestre y DegP mutada y las secuencias génicas.

La Figura 6 muestra la construcción de un vector para uso en la producción de una célula de acuerdo con una realización de la presente invención.

La Figura 7 muestra los perfiles de crecimiento comparativo de fermentaciones de 5 L y 200 L de la cepa MXE008 de *E. coli* que expresa Fab' anti-TNF y DsbC recombinante.

La Figura 8 muestra títulos de Fab' comparativo de fermentaciones de 5 L y 200 L de la cepa MXE008 de *E. coli* que expresa Fab' anti-TNF y DsbC recombinante.

La Figura 9 muestra los perfiles de crecimiento comparativo de fermentaciones de cepas MXE008 y MXE009 de *E. coli* que expresan Fab' anti-TNF y DsbC recombinante.

- 5 La Figura 10 muestra títulos de Fab' comparativo de fermentaciones de cepas MXE008 y MXE009 de *E. coli* que expresan Fab' anti-TNF y DsbC recombinante.

**Breve descripción de las secuencias**

SEQ ID NO: 1 es la secuencia de ADN del gen Tsp de tipo silvestre que incluye los 6 nucleótidos ATGAAC aguas arriba del codón de iniciación.

- 10 SEQ ID NO: 2 es la secuencia de aminoácidos de la proteína Tsp de tipo silvestre.

SEQ ID NO: 3 es la secuencia de ADN de un gen Tsp mutado inactivado que incluye los 6 nucleótidos ATGAAT aguas arriba del codón de iniciación.

SEQ ID NO: 4 es la secuencia de ADN del gen Proteasa III de tipo silvestre.

SEQ ID NO: 5 es la secuencia de aminoácidos de la proteína Proteasa III de tipo silvestre.

- 15 SEQ ID NO: 6 es la secuencia de ADN de un gen Proteasa III mutado inactivado.

SEQ ID NO: 7 es la secuencia de ADN del gen DegP de tipo silvestre.

SEQ ID NO: 8 es la secuencia de aminoácidos de la proteína DegP de tipo silvestre.

SEQ ID NO: 9 es la secuencia de ADN de un gen DegP mutado.

SEQ ID NO: 10 es la secuencia de aminoácidos de una proteína DegP mutada.

- 20 SEQ ID NO: 11 es la secuencia de aminoácidos de la región variable de la cadena ligera de un anticuerpo anti-TNF.

SEQ ID NO:12 es la secuencia de aminoácidos de la región variable de la cadena pesada de un anticuerpo anti-TNF.

SEQ ID NO: 13 es la secuencia de aminoácidos de la cadena ligera de un anticuerpo anti-TNF.

SEQ ID NO: 14 es la secuencia de aminoácidos de la cadena pesada de un anticuerpo anti-TNF.

- 25 SEQ ID NO: 15 es la secuencia del cebador oligonucleótido 3' para la región del gen Tsp mutado que comprende el sitio de restricción *Ase I*.

SEQ ID NO: 16 es la secuencia del cebador oligonucleótido 5' para la región del gen Tsp mutado que comprende el sitio de restricción *Ase I*.

- 30 SEQ ID NO: 17 es la secuencia del cebador oligonucleótido 3' para la región del gen Proteasa III mutado que comprende el sitio de restricción *Ase I*.

SEQ ID NO: 18 es la secuencia del cebador oligonucleótido 5' para la región del gen Proteasa III mutado que comprende el sitio de restricción *Ase I*.

SEQ ID NO: 19 es la secuencia del cebador de oligonucleótido 5' para la región del gen DegP mutado que comprende el sitio de restricción *Ase I*.

- 35 SEQ ID NO: 20 es la secuencia del cebador de oligonucleótido 3' para la región del gen DegP mutado que comprende el sitio de restricción *Ase I*.

SEQ ID NO: 21 es la secuencia del gen spr de tipo silvestre que incluye la secuencia señal que son los primeros 26 residuos de aminoácidos.

SEQ ID NO: 22 es la secuencia del gen spr de tipo silvestre sin la secuencia señal.

- 40 SEQ ID NO: 23 es la secuencia de nucleótidos de una secuencia de OmpT mutada que comprende las mutaciones D210A y H212A.

SEQ ID NO: 24 es la secuencia de aminoácidos de una secuencia de OmpT mutada que comprende las mutaciones D210A y H212A.

SEQ ID NO: 25 es la secuencia de nucleótidos de una secuencia de OmpT mutada inactivada.

SEQ ID NO: 26 es la secuencia de nucleótidos de DsbC marcada con his.

SEQ ID NO: 27 es la secuencia de aminoácidos de DsbC marcada con his.

SEQ ID NO: 28 muestra la secuencia de aminoácidos de CDRH1 de hTNF40.

- 5 SEQ ID NO: 29 muestra la secuencia de aminoácidos de CDRH2 de hTNF40 que es una CDR híbrida en la que los seis aminoácidos C-terminales proceden de la secuencia H2 CDR de un anticuerpo de la línea germinal del subgrupo 3 humano y los cambios de aminoácidos en la secuencia resultantes de esta hibridación están señalados de la siguiente manera: WINTYIGEPI YADSVKG.

SEQ ID NO: 30 muestra la secuencia de aminoácidos de CDRH3 de hTNF40.

- 10 SEQ ID NO: 31 muestra la secuencia de aminoácidos de CDRL1 de hTNF40.

SEQ ID NO: 32 muestra la secuencia de aminoácidos de CDRL2 de hTNF40.

SEQ ID NO: 33 muestra la secuencia de aminoácidos de CDRL3 de hTNF40.

SEQ ID NO: 34 muestra la secuencia de aminoácidos de CDRH2 de hTNF40.

SEQ ID NO: 35 muestra la secuencia del adaptador de oligonucleótidos OmpA.

- 15 SEQ ID NO: 36 muestra el casete de oligonucleótidos que codifica la secuencia intergénica 1 (IGS1) para la expresión de Fab en *E. coli*.

SEQ ID NO: 37 muestra el casete de oligonucleótidos que codifica la secuencia intergénica 2 (IGS2) para la expresión de Fab en *E. coli*.

- 20 SEQ ID NO: 38 muestra el casete de oligonucleótidos que codifica la secuencia intergénica 3 (IGS3) para la expresión de Fab en *E. coli*.

SEQ ID NO: 39 muestra el casete de oligonucleótidos que codifica la secuencia intergénica 4 (IGS4) para la expresión de Fab en *E. coli*.

### Descripción detallada de las realizaciones preferidas de la invención

A continuación se describirá la presente invención con más detalle.

- 25 Los términos "proteína" y "polipéptido" se utilizan indistintamente en la presente memoria, a menos que el contexto indique lo contrario. Se entiende que "péptido" hace referencia a 10 o menos aminoácidos.

El término "polinucleótido" incluye un gen, ADN, ADNc, ARN, ARNm, etc., a menos que el contexto indique lo contrario.

- 30 Según se utiliza en la presente memoria, la expresión "que comprende" en el contexto de la presente memoria descriptiva debe ser interpretada como "que incluye".

La célula no mutada o célula de control en el contexto de la presente invención significa una célula del mismo tipo que la célula gram-negativa recombinante de la invención, en donde la célula no se ha modificado para reducir la actividad de la proteína Tsp y es portadora de la secuencia de DsbC recombinante y, opcionalmente, el gen spr mutante. Por ejemplo, una célula no mutada puede ser una célula de tipo silvestre y se puede obtener a partir de la misma población de células hospedadoras que las células de la invención antes de la modificación para introducir cualquier mutación.

- 35 Las expresiones "célula", "línea celular", "cultivo celular" y "cepa" se utilizan indistintamente.

La expresión "fenotipo de una célula que comprende un gen Tsp mutado" en el contexto de la presente invención significa el fenotipo mostrado por una célula que alberga un gen Tsp mutante. Normalmente las células que comprenden un gen Tsp mutante se pueden lisar, especialmente a altas densidades celulares. La lisis de estas células provoca la fuga de cualquier proteína recombinante al material sobrenadante. Células portadoras del gen Tsp mutado también pueden mostrar un crecimiento termosensible con baja osmolaridad. Por ejemplo, las células no presentan ninguna tasa de crecimiento o una reducida, o las células mueren en medios hipotónicos a una temperatura elevada, tal como a 40°C o más.

- 45 El término "isogénico" en el contexto de la presente invención significa que el genoma de la célula comprende la misma o sustancialmente la misma secuencia o secuencias de ácido nucleico, en comparación con la célula de tipo silvestre a partir de la cual se obtiene la célula, excepto por la modificación requerida para reducir la actividad de la

proteína Tsp en comparación con una célula de tipo silvestre y, opcionalmente, el gen spr mutado. En esta realización, el genoma de la célula no comprende más mutaciones adicionales de origen no natural o mediante ingeniería genética. En una realización, el genoma de la célula de la presente invención tiene sustancialmente la misma o la misma secuencia genómica en comparación con células de tipo silvestre, excepto por la modificación requerida para reducir la actividad de la proteína Tsp en comparación con una célula de tipo silvestre y, opcionalmente, el gen spr mutado. En una realización, la célula de acuerdo con la presente invención puede tener sustancialmente la misma secuencia genómica en comparación con la célula de tipo silvestre, excepto por la modificación requerida para reducir la actividad de la proteína Tsp en comparación con una célula de tipo silvestre y, opcionalmente, el gen spr mutado teniendo en cuenta cualquier mutación de origen natural que pueda tener lugar. En una realización, la célula de acuerdo con la presente invención puede tener exactamente la misma secuencia genómica, en comparación con la célula de tipo silvestre, excepto por la modificación requerida para reducir la actividad de la proteína Tsp en comparación con una célula de tipo silvestre y, opcionalmente, el gen spr mutado.

El polinucleótido recombinante que codifica DsbC puede estar presente en un vector de expresión adecuado, transformado en la célula y/o integrado en el genoma de la célula hospedadora. En la realización en la que el polinucleótido que codifica DsbC se inserta en el genoma del hospedador, el genoma de la célula también será diferente de una célula de tipo silvestre debido a la secuencia de polinucleótido insertada que codifica DsbC. Preferiblemente, el polinucleótido que codifica DsbC está en un vector de expresión en la célula, causando con ello una alteración mínima del genoma de la célula hospedadora.

En una realización, la célula de la presente invención comprende un polinucleótido que codifica la proteína de interés. En esta realización, el polinucleótido que codifica la proteína de interés puede estar contenido dentro de un vector de expresión adecuado, transformado en la célula y/o integrado en el genoma de la célula hospedadora. En la realización en la que el polinucleótido que codifica la proteína de interés se inserta en el genoma del hospedador, el genoma también será diferente de una célula de tipo silvestre debido a la secuencia de polinucleótido insertada que codifica la proteína de interés. Preferiblemente, el polinucleótido que codifica la proteína de interés está en un vector de expresión en la célula, causando con ello una alteración mínima del genoma de la célula hospedadora.

La expresión "de tipo silvestre" en el contexto de la presente invención, significa una cepa de una célula bacteriana gram-negativa, como puede presentarse en la naturaleza o se puede aislar del entorno, que no es portadora de ninguna mutación producida por ingeniería genética. Un ejemplo de una cepa de tipo silvestre de *E. coli* es W3110, tal como la cepa W3110 K-12.

Los presentes inventores han proporcionado una célula bacteriana gram-negativa recombinante adecuada para expresar una proteína de interés que tiene una actividad de proteína Tsp reducida, en comparación con una célula de tipo silvestre y comprende un polinucleótido recombinante que codifica DsbC.

Las células de acuerdo con la presente invención comprenden un polinucleótido recombinante que codifica DsbC. Tal como se utiliza en la presente memoria, un "polipéptido recombinante" se refiere a una proteína que se construye o produce empleando tecnología de ADN recombinante. La secuencia de polinucleótido que codifica DsbC puede ser idéntica a la secuencia endógena que codifica DsbC encontrada en células bacterianas. Alternativamente, la secuencia de polinucleótido recombinante que codifica DsbC es una versión mutada de la secuencia de DsbC de tipo silvestre, por ejemplo, que tiene un sitio de restricción eliminado, tal como un sitio EcoRI y/o una secuencia que codifica un marcador His. Un ejemplo de secuencia nucleotídica de DsbC modificada para su utilización en la presente invención se muestra en SEQ ID NO: 26, la cual codifica la secuencia de aminoácidos de DsbC marcada con his mostrada en SEQ ID NO: 27.

Se ha encontrado que la combinación específica de la expresión de un polinucleótido recombinante que codifica DsbC en una célula bacteriana que tiene una la actividad reducida de la proteína Tsp en comparación con una célula de tipo silvestre y en una realización preferida, además, un gen spr mutado proporciona un hospedador mejorado para la expresión de proteínas de interés. La combinación específica de las mutaciones genéticas anteriores proporciona nuevas cepas que han mejorado la salud celular y el fenotipo de crecimiento en comparación con las células que son portadoras de un gen Tsp mutado inactivado. Las células que son portadoras de un gen Tsp mutado pueden tener una buena tasa de crecimiento celular, pero una limitación de estas células es su tendencia a la lisis, especialmente a altas densidades celulares. En consecuencia, el fenotipo de una célula que comprende un gen Tsp mutado es una tendencia a la lisis, especialmente a altas densidades celulares. Sin embargo, la expresión de DsbC en las células de la presente invención, suprime el fenotipo Tsp reducido y, por lo tanto, las células muestran una lisis reducida.

Las células de acuerdo con la presente invención muestran un rendimiento mejorado de la producción de proteínas en comparación con las células que son portadoras de un gen Tsp mutado inactivado. El rendimiento mejorado de proteínas puede ser la tasa de producción de proteínas y/o la duración de la producción de proteínas desde la célula. El rendimiento mejorado de proteínas puede ser el rendimiento de proteínas periplásmicas y/o el rendimiento de proteínas en el material sobrenadante. En una realización, las células de la presente invención muestran un rendimiento mejorado de proteínas periplásmicas en comparación con una célula que es portadora de un gen Tsp mutado, debido a una reducción de la fuga desde la célula. Las células bacterianas recombinantes pueden ser capaces de una tasa de producción de una proteína de interés más rápida y, por lo tanto, la misma cantidad de una proteína

de interés se puede producir en menos tiempo en comparación con una célula que comprende un gen Tsp mutado. La tasa de producción de una proteína de interés más rápida puede ser especialmente significativa durante el período de crecimiento inicial de la célula, por ejemplo, durante las primeras 5, 10, 20 o 30 horas después de la inducción de la expresión de la proteína.

- 5 Las células de acuerdo con la presente invención expresan preferiblemente un rendimiento máximo en el periplasma y/o el medio de aproximadamente 1,0 g/L, 1,5 g/L, 1,8 g/L, 2,0 g/L, 2,4 g/L, 2,5 g/L, 3,0 g/L, 3,5 g/L o 4,0 g/L de una proteína de interés.

- 10 Las células proporcionadas por la presente invención han reducido la actividad proteasa en comparación con una célula de tipo silvestre, lo que puede reducir la proteólisis de una proteína recombinante de interés, en particular las proteínas de interés que son proteolíticamente sensibles a la proteasa Tsp. Por lo tanto, las células proporcionadas por la presente invención proporcionan un mayor rendimiento de proteínas intactas, preferiblemente de la proteína de interés y un rendimiento inferior, o preferiblemente una falta de fragmentos proteolíticos de proteínas, preferentemente de la proteína de interés, en comparación con una célula bacteriana de tipo silvestre.

- 15 El experto en la materia será capaz con facilidad de someter a ensayo un clon celular candidato para observar si tiene el rendimiento deseado de una proteína de interés, usando métodos bien conocidos en la técnica que incluyen un método de fermentación, ELISA y HPLC de la proteína G. Los métodos de fermentación adecuados se describen en Humphreys D P, et al. (1997). Formation of dimeric Fabs in *E. coli*: effect of hinge size and isotype, presence of interchain disulphide bond, Fab' expression levels, tail piece sequences and growth conditions. J. IMMUNOL. METH. 209: 193-202; Backlund E. Reeks D. Markland K. Weir N. Bowering L. Larsson G. Fedbatch design for periplasmic product retention in *Escherichia coli*, Journal Article. Research Support, Non-U.S. Gov't Journal of Biotechnology. 20 135(4):358-65, 2008 Jul 31; Champion KM. Nishihara JC. Joly JC. Arnott D. Similarity of the *Escherichia coli* proteome upon completion of different biopharmaceutical fermentation processes. [Artículo de Revista] Proteomics. 1(9):1133-48, 2001 Sep; y Horn U. Strittmatter W. Krebber A. Knupfer U. Kujau M. Wenderoth R. Muller K. Matzku S. Pluckthun A. Riesenberg D. High volumetric yields of functional dimeric miniantibodies in *Escherichia coli*, using an optimized expression vector and high-cell-density fermentation under non-limited growth conditions, Artículo de Re- 25 vista. Research Support, Non-U.S. Gov't Applied Microbiology & Biotechnology. 46(5-6):524-32, 1996 Dic. La persona experta también será capaz fácilmente de someter a ensayo la proteína secretada para observar si la proteína está plegada correctamente usando métodos bien conocidos en la técnica, tales como HPLC de la proteína G, di-  
croísmo circular, RMN, cristalografía de rayos X y métodos de medición de la afinidad de epítopos.

- 30 En una realización, la célula de acuerdo con la presente invención también expresa una o varias proteínas adicionales del modo siguiente:

- una o varias proteínas capaces de facilitar el plegamiento de proteínas, tales como FkpA, Skp, SurA, PPIA y PPIID; y/o
- una o varias proteínas capaces de facilitar la secreción o translocación de proteínas, tales como SecY, 35 SecE, SecG, SecYEG, SecA, SecB, FtsY y Lep; y/o
- una o varias proteínas capaces de facilitar la formación de enlaces disulfuro, tales como DsbA, DsbB, DsbD, DsbG.

Una o varias de las proteínas anteriores puede(n) estar integrada(s) en el genoma de la célula y/o insertada(s) en un vector de expresión.

- 40 En una realización, la célula de acuerdo con la presente invención no comprende un polinucleótido recombinante que codifica una o varias de las siguientes proteínas adicionales:

- una o varias proteínas capaces de facilitar el plegamiento de proteínas, tales como FkpA, Skp, SurA, PPIA y PPIID;
- una o varias proteínas capaces de facilitar la secreción o translocación de proteínas, tales como SecY, 45 SecE, SecG, SecYEG, SecA, SecB, FtsY y Lep; y
- una o varias proteínas capaces de facilitar la formación de enlaces disulfuro, tales como DsbA, DsbB, DsbD, DsbG.

En realizaciones de la presente invención, la célula comprende además un gen spr mutado. La proteína spr es una proteasa periplásmica de *E. coli* unida a la membrana.

- 50 La secuencia de aminoácidos de tipo silvestre de la proteína spr se muestra en SEQ ID NO: 21 con la secuencia señal en el extremo N-terminal y en SEQ ID NO: 22 sin la secuencia señal de 26 aminoácidos (según el número de registro P0AFV4 de UniProt). La numeración de los aminoácidos de la secuencia de la proteína spr en la presente

invención incluye la secuencia señal. Por consiguiente, el aminoácido 1 de la proteína spr es el primer aminoácido (Met) mostrado en SEQ ID NO: 21.

5 En las realizaciones en las que la célula de acuerdo con la presente invención comprende un gen spr mutado, el gen spr mutado es preferiblemente el gen spr cromosómico de la célula. El gen spr mutado codifica una proteína spr capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado. Las células que son portadoras de un gen Tsp mutado pueden tener una buena tasa de crecimiento celular, pero una limitación de esas células es su tendencia a la lisis, especialmente a altas densidades celulares. En consecuencia, el fenotipo de una célula que comprende un gen Tsp mutado es una tendencia a la lisis, especialmente a altas densidades celulares. Las células que son portadoras de un gen Tsp mutado también muestran un crecimiento termosensible con baja osmolaridad. Sin embargo, las mutaciones spr realizadas mediante las células de la presente invención, cuando se introducen en una célula que tiene una reducción de la actividad Tsp, suprimen el fenotipo Tsp reducido y, por lo tanto, la célula muestra una lisis reducida, particularmente con una alta densidad celular. El fenotipo de crecimiento de una célula puede ser medido fácilmente por una persona experta en la técnica durante una técnica de agitación en matraz o de fermentación de alta densidad celular. La supresión del fenotipo de lisis celular se puede observar a partir de la tasa de crecimiento mejorada y/o la producción de proteína recombinante, en particular en el periplasma, mostrada por una célula que es portadora de spr mutante y en la que se ha reducido la actividad Tsp en comparación con una célula portadora del mutante Tsp y una spr de tipo silvestre.

20 Las células de acuerdo con la presente invención comprenden preferiblemente un gen spr mutante que codifica una proteína spr que tiene una mutación en uno o varios aminoácidos seleccionados a partir de N31, R62, I70, Q73, C94, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144, H145, G147, H157 y W174, más preferentemente en uno o varios aminoácidos seleccionados a partir de C94, S95, V98, Y115, D133, V135, H145, G147, H157 y W174. Preferiblemente, el gen spr mutante codifica una proteína spr que tiene una mutación en uno o varios aminoácidos seleccionados a partir de N31, R62, I70, Q73, C94, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144, H145, G147 y H157, más preferentemente en uno o varios aminoácidos seleccionados a partir de C94, S95, V98, Y115, D133, V135, H145, G147 y H157. En esta realización, la proteína spr preferiblemente no tiene ninguna mutación adicional. Preferiblemente, el gen spr mutante codifica una proteína spr que tiene una mutación en uno o varios aminoácidos seleccionados a partir de N31, R62, I70, Q73, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144 y G147, más preferentemente en uno o varios aminoácidos seleccionados a partir de S95, V98, Y115, D133, V135 y G147. En esta realización, la proteína spr preferiblemente no tiene ninguna mutación adicional.

Los presentes inventores han identificado mutaciones spr que son capaces de suprimir el fenotipo de crecimiento de una célula que comprende un gen Tsp mutado.

35 Los inventores también han encontrado sorprendentemente que las células que son portadoras de un gen recombinante DsbC, un nuevo gen spr mutado y que tienen una actividad reducida de la proteína Tsp en comparación con una célula de tipo silvestre, muestran una tasa de crecimiento celular incrementada y un aumento de la duración de la supervivencia celular, en comparación con una célula que comprende un gen Tsp mutado. Específicamente, las células que son portadoras de un gen DsbC recombinante, una mutación spr y que tienen una actividad reducida de la proteína Tsp, muestran un fenotipo reducido de lisis celular en comparación con las células que son portadoras de un gen Tsp mutado.

40 La mutación de uno o varios de los aminoácidos de spr anteriores puede ser cualquier mutación de sentido erróneo adecuada para uno, dos o tres de los nucleótidos que codifican el aminoácido. La mutación cambia el residuo del aminoácido a cualquier aminoácido adecuado, lo que se traduce en una proteína spr mutada, capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado. La mutación de sentido erróneo puede cambiar el aminoácido a uno que tiene un tamaño diferente y/o tiene propiedades químicas diferentes, en comparación con el aminoácido de tipo silvestre.

En una realización, la mutación es en una, dos o tres de la tríada catalítica de residuos de aminoácidos de C94, H145 y H157 (Solution NMR Structure of the NlpC/P60 Domain of Lipoprotein Spr from Escherichia coli Structural Evidence for a Novel Cysteine Peptidase Catalytic Triad, Biochemistry, 2008, 47, 9715-9717).

Por consiguiente, el gen spr mutado puede comprender:

- 50
- una mutación en C94; o
  - una mutación en H145; o
  - una mutación en H157; o
  - una mutación en C94 y H145; o
  - una mutación en C94 y H157; o

- una mutación en H145 y H157; o
- una mutación en C94, H145 y H157.

En esta realización, la proteína spr preferiblemente no tiene ninguna mutación adicional.

5 Uno, dos o tres de C94, H145 y H157 se pueden mutar a cualquier aminoácido adecuado, lo que se traduce en una proteína spr capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado. Por ejemplo, uno, dos o tres de C94, H145 y H157 se pueden mutar a un aminoácido pequeño tal como Gly o Ala. En consecuencia, la proteína spr puede tener una, dos o tres de las mutaciones C94A, H145A y H157A. Preferiblemente, el gen spr comprende la mutación de sentido erróneo H145A, que se ha encontrado que produce una proteína spr capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado.

10 La designación para un mutante por sustitución en la presente memoria consiste en una letra seguida de un número seguido de una letra. La primera letra designa el aminoácido en la proteína de tipo silvestre. El número hace referencia a la posición del aminoácido en la que se está realizando la sustitución de aminoácido, y la segunda letra designa el aminoácido que se utiliza para remplazar al aminoácido de tipo silvestre.

15 En una realización preferida, la proteína spr mutante comprende una mutación en uno o varios aminoácidos seleccionados a partir de N31, R62, I70, Q73, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144 y G147, preferiblemente una mutación en uno o varios aminoácidos seleccionados a partir de S95, V98, Y115, D133, V135 y G147. En esta realización, la proteína spr preferiblemente no tiene ninguna mutación adicional. Por consiguiente, el gen spr mutado puede comprender:

- una mutación en N31; o
- 20 • una mutación en R62; o
- una mutación en I70; o
- una mutación en Q73; o
- una mutación en S95; o
- una mutación en V98; o
- 25 • una mutación en Q99; o
- una mutación en R100; o
- una mutación en L108; o
- una mutación en Y115; o
- una mutación en D133; o
- 30 • una mutación en V135; o
- una mutación en L136; o
- una mutación en G140; o
- una mutación en R144; o
- una mutación en G147.

35 En una realización, la proteína spr mutante comprende múltiples mutaciones en los aminoácidos:

- S95 y Y115; o
- N31, Q73, R100 y G140; o
- Q73, R100 y G140; o
- R100 y G140; o
- 40 • Q73 y G140; o

- Q73 y R100; o
- R62, Q99 y R144; o
- Q99 y R144.

5 Uno o varios de los aminoácidos N31, R62, I70, Q73, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144 y G147 puede estar mutado a cualquier aminoácido adecuado, lo que se traduce en una proteína spr capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado. Por ejemplo, uno o varios de N31, R62, I70, Q73, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140 y R144 se pueden mutar a un aminoácido pequeño tal como Gly o Ala.

10 En una realización preferida, la proteína spr comprende una o varias de las siguientes mutaciones: N31Y, R62C, I70T, Q73R, S95F, V98E, Q99P, R100G, L108S, Y115F, D133A, V135D o V135G, L136P, G140C, R144C y G147C. Preferiblemente, la proteína spr comprende una o varias de las siguientes mutaciones: S95F, V98E, Y115F, D133A, V135D o V135G y G147C. En esta realización, la proteína spr preferiblemente no tiene ninguna mutación adicional.

15 En una realización, la proteína spr tiene una mutación seleccionada a partir de N31Y, R62C, I70T, Q73R, S95F, V98E, Q99P, R100G, L108S, Y115F, D133A, V135D o V135G, L136P, G140C, R144C y G147C. En esta realización, la proteína spr preferiblemente no tiene ninguna mutación adicional.

En una realización adicional, la proteína spr tiene múltiples mutaciones seleccionadas a partir de:

- S95F e Y115F;
- N31Y, Q73R, R100G y G140C;
- Q73R, R100G y G140C;
- 20 • R100G y G140C;
- Q73R y G140C;
- Q73R y R100G;
- R62C, Q99P y R144C; o
- Q99P y R144C.

25 Preferiblemente, el gen spr mutante codifica una proteína spr que tiene una mutación seleccionada a partir de C94A, D133A, H145A y H157A.

En otra realización, el gen spr mutado codifica una proteína spr que tiene la mutación W174R. En una realización alternativa, la proteína spr no tiene la mutación W174R.

30 La célula de acuerdo con la presente invención tiene una actividad reducida de la proteína Tsp, en comparación con una célula de tipo silvestre. La expresión "actividad reducida de la proteína Tsp en comparación con una célula de tipo silvestre" significa que la actividad de Tsp de la célula está reducida en comparación con la actividad de Tsp de una célula de tipo silvestre. La célula puede estar modificada por cualquier medio adecuado para reducir la actividad de Tsp.

35 En una realización, la actividad reducida de Tsp procede de la modificación del polinucleótido endógeno que codifica Tsp y/o secuencias reguladoras de la expresión asociadas. La modificación puede reducir o detener la transcripción y traducción del gen Tsp o puede proporcionar la proteína Tsp expresada con actividad proteasa reducida en comparación con la proteína Tsp de tipo silvestre.

En una realización, se modifica la secuencia reguladora de la expresión asociada para reducir la expresión de Tsp. Por ejemplo, el promotor del gen Tsp puede mutarse para evitar la expresión del gen.

40 En una realización preferida, la célula de acuerdo con la presente invención es portadora de un gen Tsp mutado que codifica una proteína Tsp con actividad proteasa reducida o un gen Tsp mutado inactivado. Preferiblemente el gen Tsp cromosómico está mutado.

45 Tal y como se usa en el presente documento, "gen Tsp" significa un gen que codifica la proteasa Tsp (también conocida como Prc) que es una proteasa periplásmica capaz de actuar sobre la proteína 3 que se une a penicilina (PBP3) y proteínas de la cola del fago. La secuencia del gen Tsp de tipo silvestre se muestra en SEQ ID NO: 1 y la secuencia de la proteína Tsp de tipo silvestre se muestra en SEQ ID NO: 2.

La referencia al gen Tsp mutado o gen Tsp mutado que codifica Tsp, se refiere a cualquier gen Tsp mutado que codifica una proteína Tsp que tiene una actividad proteasa reducida o un gen Tsp mutado inactivado, a menos que se indique lo contrario.

5 La expresión "gen Tsp mutado que codifica una proteína Tsp que tiene una actividad proteasa reducida" en el contexto de la presente invención significa que el gen Tsp mutado no tiene la actividad proteasa completa, en comparación con el gen Tsp no mutado de tipo silvestre.

10 Preferentemente, el gen Tsp mutado codifica una proteína Tsp que tiene 50% o menos, 40% o menos, 30% o menos, 20% o menos, 10% o menos o 5% o menos de la actividad proteasa de una proteína Tsp de tipo silvestre sin mutar. Más preferiblemente, el gen Tsp mutado codifica una proteína Tsp que no tiene actividad proteasa. En esta realización, la célula no carece de Tsp cromosómico, es decir, la secuencia del gen Tsp no se ha eliminado o mutado para evitar la expresión de cualquier forma de proteína Tsp.

15 Se puede introducir cualquier mutación adecuada en el gen Tsp con el fin de producir una proteína que tenga actividad proteasa reducida. La actividad proteasa de una proteína Tsp expresada a partir de una bacteria gram-negativa, puede ser comprobada fácilmente por una persona experta en la técnica a través de cualquier método adecuado en la técnica, tal como el método descrito en Keiler et al. (Identification of Active Site Residues of the Tsp Protease\* THE JOURNAL OF BIOLOGICAL CHEMISTRY vol. 270, n° 48, del 1 de Diciembre 1, págs. 28864-28868, 1995 Kenneth C. Keiler y Robert T. Sauer) en donde se somete a ensayo la actividad de Tsp.

20 Se ha descrito en Keiler et al. (supra) que Tsp tiene un sitio activo que comprende los residuos S430, D441 y K455, y los residuos G375, G376, E433 y T452 son importantes para mantener la estructura de Tsp. Keiler et al. (supra) informan sobre hallazgos de que los genes Tsp mutados S430A, D441A, K455A, K455H, K455R, G375A, G376A, E433A y T452A no tienen actividad proteasa detectable. Se ha informado además de que el gen Tsp mutado S430C muestra una actividad de tipo silvestre de aproximadamente 5-10%. De acuerdo con ello, la mutación de Tsp para producir una proteína que tiene actividad proteasa reducida, puede comprender una mutación, tal como una mutación de sentido erróneo en uno o varios de los residuos S430, D441, K455, G375, G376, E433 y T452. Preferiblemente, la mutación de Tsp para producir una proteína que tiene actividad proteasa reducida puede comprender una mutación, tal como una mutación de sentido erróneo en uno, dos o los tres residuos del sitio activo S430, D441 y K455.

Por consiguiente, el gen Tsp mutado puede comprender:

- una mutación en S430; o
- 30 • una mutación en D441; o
- una mutación en K455; o
- una mutación en S430 y D441; o
- una mutación en S430 y K455; o
- una mutación en D441 y K455; o
- 35 • una mutación en S430, D441 y K455.

Se puede mutar uno o varios de S430, D441, K455, G375, G376, E433 y T452 a cualquier aminoácido adecuado, lo que se traduce en una proteína que tiene actividad proteasa reducida. Ejemplos de mutaciones adecuadas son S430A, S430C, D441A, K455A, K455H, K455R, G375A, G376A, E433A y T452A. El gen Tsp mutado puede comprender una, dos o tres mutaciones en los residuos del sitio activo, por ejemplo el gen puede comprender:

- 40 • S430A o S430C; y/o
- D441A; y/o
- K455A o K455H o K455R.

Preferiblemente, el gen Tsp tiene la mutación puntual S430A o S430C.

45 La expresión "gen Tsp mutado inactivado" en el contexto de la presente invención, significa que el gen Tsp comprende una o varias mutaciones que impiden la expresión de la proteína Tsp codificada por el gen de tipo silvestre, para proporcionar una célula que carece de proteína Tsp. El gen inactivado (del inglés, "knockout") se puede transcribir parcial o completamente, pero no se puede traducir en la proteína codificada. El gen Tsp mutado inactivado se puede mutar de cualquier manera adecuada, por ejemplo, mediante una o varias deleciones, inserciones, mutaciones puntuales, de sentido erróneo, sin sentido y de desplazamiento del marco, para causar una falta de expresión de

la proteína. Por ejemplo, el gen se puede inactivar mediante la inserción de una secuencia de ADN extraño, tal como un marcador de resistencia a antibióticos, en la secuencia que codifica el gen.

5 En una realización preferida, el gen Tsp no está mutado mediante la inserción de una secuencia de ADN extraño, tal como un marcador de resistencia a antibióticos, en la secuencia que codifica el gen. En una realización, el gen Tsp comprende una mutación en el codón de inicio del gen y/o en uno o varios codones de parada situados aguas abajo del codón de inicio del gen y aguas arriba del gen de codón de parada, evitando de este modo la expresión de la proteína Tsp. La mutación del codón de inicio puede ser una mutación de sentido erróneo de uno, dos o los tres nucleótidos del codón de inicio. Alternativa o adicionalmente, el codón de inicio se puede mutar por una mutación en el marco de lectura por inserción o delección. El gen Tsp comprende dos codones ATG en el extremo 5' de la secuencia codificante, uno o ambos de los codones ATG puede estar mutado con una mutación de sentido erróneo. El gen Tsp se puede mutar en el segundo codón ATG (codón 3) a TCG, como se muestra en la Figura 5b. El gen Tsp puede comprender alternativa o adicionalmente uno o varios codones de parada situados aguas abajo del codón de inicio del gen y aguas arriba del codón de parada del gen. Preferiblemente, el gen Tsp mutado inactivado comprende tanto una mutación de sentido erróneo en el codón de inicio como uno o varios codones de parada insertados. En una realización preferida, el gen TSP se muta para eliminar "T" del quinto codón, provocando de este modo un desplazamiento del marco lo que da como resultado codones de parada en los codones 11 y 16, como se muestra en la Figura 5b. En una realización preferida, el gen Tsp se muta para insertar un sitio de restricción *Ase I* para crear un tercer codón de parada en marco en el codón 21, como se muestra en la Figura 5b.

20 En una realización preferida, el gen Tsp mutado inactivado tiene la secuencia de ADN de SEQ ID NO: 3, que incluye los 6 nucleótidos ATGAAT aguas arriba del codón de inicio. Las mutaciones que se han realizado en la secuencia de Tsp mutada inactivada de SEQ ID NO: 3 se muestran en la Figura 5b. En una realización, el gen Tsp mutado tiene la secuencia de ADN de los nucleótidos 7 a 2048 de SEQ ID NO: 3.

25 En una realización preferida de la presente invención, la célula bacteriana gram-negativa recombinante comprende además un gen DegP mutado que codifica una proteína DegP que tiene actividad de tipo chaperona y actividad proteasa reducida y/o un gen ptr mutado, en donde el gen ptr mutado codifica una proteína Proteasa III que tiene actividad proteasa reducida o es un gen ptr mutado inactivado y/o un gen OmpT mutado, en donde el gen OmpT mutado codifica una proteína OmpT que tiene actividad proteasa reducida o es un gen OmpT mutado inactivado.

En una realización, la presente invención proporciona una célula bacteriana gram-negativa recombinante que comprende

- 30 a. un polinucleótido recombinante que codifica DsbC;
- b. un gen Tsp mutado que codifica una proteína Tsp que tiene actividad proteasa reducida o un gen Tsp mutado inactivado;
- 35 c. un gen DegP mutado que codifica una proteína DegP que tiene actividad de tipo chaperona y actividad proteasa reducida y/o una OmpT mutada en donde el gen OmpT mutado codifica una proteína OmpT que tiene actividad proteasa reducida o es un gen OmpT mutado inactivado; y
- d. opcionalmente un gen spr mutado.

Preferiblemente, en esta realización el genoma celular es isogénico a una célula bacteriana de tipo silvestre, excepto para las mutaciones anteriores b, c y d.

40 En una realización, la presente invención proporciona una célula bacteriana gram-negativa recombinante que comprende:

- a. un polinucleótido recombinante que codifica DsbC;
- b. un gen Tsp mutado que codifica una proteína Tsp que tiene actividad proteasa reducida o un gen Tsp mutado inactivado;
- 45 c. un gen ptr mutado, en donde el gen ptr mutado codifica una proteína Proteasa III que tiene actividad proteasa reducida o es un gen ptr mutado inactivado y/o una OmpT mutada en donde el gen OmpT mutado codifica una proteína OmpT que tiene actividad proteasa reducida o es un gen OmpT mutado inactivado; y
- d. opcionalmente un gen spr mutado.

Preferiblemente, en esta realización el genoma celular es isogénico a una célula bacteriana de tipo silvestre, excepto para las mutaciones anteriores b, c y d.

50 En una realización, la presente invención proporciona una célula que comprende

- a. un polinucleótido recombinante que codifica DsbC;

b. un gen Tsp mutado que codifica una proteína Tsp que tiene actividad proteasa reducida o un gen Tsp mutado inactivado;

c. un gen DegP mutado que codifica una proteína DegP que tiene actividad de tipo chaperona y actividad proteasa reducida;

5 d. un gen ptr mutado, en donde el gen ptr mutado codifica una proteína Proteasa III que tiene actividad proteasa reducida o es un gen ptr mutado inactivado;

e. opcionalmente una OmpT mutada, en donde el gen OmpT mutado codifica una proteína OmpT que tiene actividad proteasa reducida o es un gen OmpT mutado inactivado; y

f. opcionalmente un gen spr mutado.

10 Preferiblemente, en esta realización el genoma de la célula es isogénico a una célula bacteriana de tipo silvestre, excepto para las mutaciones anteriores b, c, d, e y f.

En realizaciones de la presente invención, la célula comprende un gen DegP mutado. Tal como se usa en este documento, "DegP" significa un gen que codifica una proteína DegP (también conocida como HtrA), que tiene doble función como chaperona y proteasa (familias de serina peptidasas; Rawlings ND, Barrett AJ. *Methods Enzymol.* 1994;244:19-61). La secuencia del gen DegP sin mutar se muestra en SEQ ID NO: 7 y la secuencia de la proteína DegP sin mutar se muestra en SEQ ID NO: 8.

A bajas temperaturas, DegP actúa como una chaperona y a altas temperaturas DegP tiene una preferencia para actuar como una proteasa (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. *Cell*, volumen 97, publicación 3, páginas 339 - 347. Spiess C, Beil A, Ehrmann M) y The proteolytic activity of the HtrA (DegP) protein from *Escherichia coli* at low temperatures, Skorko-Glonek J et al. *Microbiology* 2008, 154, 3649-3658).

En las realizaciones en las que la célula comprende la mutación DegP, la mutación DegP en la célula proporciona un gen DegP mutado que codifica una proteína DegP que tiene actividad de tipo chaperona pero no una actividad proteasa completa.

25 La expresión "que tiene actividad de tipo chaperona" en el contexto de la presente invención, significa que la proteína DegP mutada tiene la misma actividad de tipo chaperona o sustancialmente la misma, en comparación con la proteína DegP de tipo silvestre sin mutar. Preferiblemente, el gen DegP mutado codifica una proteína DegP que tiene 50% o más, 60% o más, 70% o más, 80% o más, 90% o más o 95% o más de la actividad de tipo chaperona de una proteína DegP de tipo silvestre sin mutar. Más preferiblemente, el gen DegP mutado codifica una proteína DegP que tienen la misma actividad de tipo chaperona, en comparación con la de DegP de tipo silvestre.

30 La expresión "que tiene actividad proteasa reducida" en el contexto de la presente invención, significa que la proteína DegP mutada no tiene la actividad proteasa completa en comparación con la proteína DegP de tipo silvestre sin mutar. Preferiblemente, el gen DegP mutado codifica una proteína DegP que tiene 50% o menos, 40% o menos, 30% o menos, 20% o menos, 10% o menos o 5% o menos de la actividad proteasa de una proteína DegP de tipo silvestre sin mutar. Más preferiblemente, el gen DegP mutado codifica una proteína DegP que no tiene actividad proteasa. La célula no carece de DegP cromosómica, es decir, las secuencias génicas de DegP no se han eliminado o mutado para evitar la expresión de cualquier forma de proteína DegP.

Cualquier mutación adecuada se puede introducir en el gen DegP con el fin de producir una proteína que tiene actividad de tipo chaperona y actividad proteasa reducida. La actividad proteasa y chaperona de una proteína DegP expresada a partir de una bacteria gram-negativa, se puede comprobar fácilmente por una persona experta en la técnica a través de cualquier método adecuado, tal como el método descrito en Spiess et al., en el que se sometieron a ensayo las actividades proteasa y chaperona de DegP en MalS, un sustrato natural de DegP (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. *Cell*, volumen 97, publicación 3, páginas 339 - 347. Spiess C, Beil A, Ehrmann M) y también el método descrito en The proteolytic activity of the HtrA (DegP) protein from *Escherichia coli* at low temperatures, Skorko-Glonek J et al. *Microbiology* 2008, 154, 3649-3658.

40 DegP es una serina proteasa y tiene un centro activo que consiste en una triada catalítica de residuos de aminoácidos de His105, Asp135 y Ser210 (Families of serine peptidases, *Methods Enzymol.*, 1994, 244:19-61 Rawlings N y Barrett A). La mutación DegP para producir una proteína que tiene actividad de tipo chaperona y actividad proteasa reducida puede comprender una mutación, tal como una mutación de sentido erróneo en uno, dos o tres de His105, Asp135 y Ser210.

Por consiguiente, el gen DegP mutado puede comprender:

- una mutación en His105; o

- una mutación en Asp135; o
- una mutación en Ser210; o
- una mutación en His105 y Asp135; o
- una mutación en His105 y Ser210; o
- 5 • una mutación en Asp135 y Ser210; o
- una mutación en His105, Asp135 y Ser210.

Uno, dos o tres de His105, Asp135 y Ser210 se pueden mutar a cualquier aminoácido adecuado, lo que se traduce en una proteína que tiene actividad de tipo chaperona y actividad proteasa reducida. Por ejemplo, uno, dos o tres de His105, Asp135 y Ser210 se pueden mutar a un aminoácido pequeño tal como Gly o Ala. Una mutación adecuada adicional es cambiar uno, dos o tres de His105, Asp135 y Ser210 a un aminoácido que tiene propiedades opuestas, tal como mutar Asp135 a Lys o Arg, mutar His105 polar a un aminoácido no polar tal como Gly, Ala, Val o Leu y mutar Ser210 pequeña hidrófila a un residuo grande o hidrófobo tal como Val, Leu, Phe o Tyr. Preferiblemente, el gen DegP comprende la mutación puntual S210A, como se muestra en la Figura 5c, que se ha encontrado que produce una actividad de tipo chaperona pero no tiene actividad proteasa (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. Cell, volumen 97, publicación 3, páginas 339 - 347. Spiess C, Beil A, Ehrmann M).

DegP tiene dos dominios PDZ, PDZ1 (residuos 260-358) y PDZ2 (residuos 359-448), que median en la interacción proteína-proteína (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. Cell, volumen 97, publicación 3, páginas 339 - 347. Spiess C, Beil A, Ehrmann M). En una realización de la presente invención, el gen DegP está mutado para deleccionar el dominio pDZ1 y/o el dominio PDZ2. La delección de pDZ1 y PDZ2 da como resultado la pérdida completa de la actividad proteasa de la proteína DegP y una actividad reducida de tipo chaperona, en comparación con la proteína DegP de tipo silvestre, mientras que la delección de cualquiera entre pDZ1 o PDZ2, da como resultado una actividad proteasa del 5% y una actividad de tipo chaperona similar en comparación con la proteína DegP de tipo silvestre (A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. Cell, volumen 97, publicación 3, páginas 339 - 347. Spiess C, Beil A, Ehrmann M).

El gen DegP mutado también puede comprender un sitio de restricción silencioso que no es de origen natural, tal como *Ase I*, con el fin de ayudar en los métodos de identificación y escrutinio, por ejemplo, como se muestra en la Figura 5c.

30 La secuencia preferida del gen DegP mutado que comprende la mutación puntual S210A y un sitio de restricción marcador *Ase I*, se proporciona en SEQ ID NO: 9 y la secuencia de proteínas codificadas se muestra en SEQ ID NO: 10. Las mutaciones que se han realizado en la secuencia DegP mutada de SEQ ID NO: 9, se muestran en la Figura 5c.

35 En las realizaciones de la presente invención, en donde la célula comprende un gen DegP mutado que codifica una proteína DegP que tiene actividad de tipo chaperona y actividad proteasa reducida, una o varias de las células proporcionadas por la presente invención pueden proporcionar un rendimiento mejorado de proteínas correctamente plegadas a partir de la célula, respecto a las células mutadas en las que el gen DegP se ha mutado a DegP inactivado que evita la expresión de DegP, tal como DegP cromosómica defectuosa. En una célula que comprende un gen DegP mutado inactivado que evita la expresión DegP, la actividad de tipo chaperona de DegP se pierde completamente, mientras que en la célula de acuerdo con la presente invención, la actividad de tipo chaperona de DegP se conserva mientras que se pierde la actividad proteasa completa. En estas realizaciones, una o varias células según la presente invención tienen una actividad proteasa inferior para evitar la proteólisis de la proteína, mientras que mantienen la actividad de tipo chaperona para permitir un plegamiento correcto y un transporte de la proteína en la célula hospedadora.

45 El experto en la materia será capaz fácilmente de analizar la proteína secretada para ver si la proteína está plegada correctamente usando métodos bien conocidos en la técnica, tales como la HPLC de proteína G, difracción circular, RMN, métodos de cristalografía de rayos X y de medición de la afinidad del epítipo.

En estas realizaciones, una o varias células según la presente invención pueden haber mejorado el crecimiento celular en comparación con células que son portadoras de un gen DegP mutado inactivado que evita la expresión de DegP. Sin desear estar ligado a ninguna teoría, un crecimiento celular mejorado tal vez se muestre debido a la proteasa DegP que conserva la actividad de tipo chaperona, lo que puede aumentar la capacidad de la célula para procesar todas las proteínas que requieren una actividad de tipo chaperona. En consecuencia, la producción de proteínas correctamente plegadas necesarias para el crecimiento y la reproducción de la célula, se puede aumentar en una o varias de las células de la presente invención, en comparación con las células que son portadoras de una mutación inactivadora de DegP, mejorando de este modo las rutas celulares que regulan el crecimiento. Además,

- 5 cepas conocidas que carecen de una proteasa DegP son generalmente sensibles a la temperatura y no crecen normalmente a temperaturas superiores a aproximadamente 28°C. Sin embargo, las células de acuerdo con la presente invención no son sensibles a la temperatura y pueden crecer a temperaturas de 28°C o superiores, incluyendo temperaturas de aproximadamente 30°C a aproximadamente 37°C, que se utilizan normalmente para la producción a escala industrial de proteínas procedentes de bacterias.
- En realizaciones de la presente invención, la célula comprende un gen ptr mutado. Tal como se usa en el presente documento, "gen ptr" significa un gen que codifica la Proteasa III, una proteasa que degrada proteínas de peso molecular elevado. La secuencia del gen ptr no mutado se muestra en SEQ ID NO: 4 y la secuencia de la proteína Proteasa III no mutada se muestra en SEQ ID NO: 5.
- 10 La referencia al gen ptr mutado o gen ptr mutado que codifica la Proteasa III, se refiere a cualquier gen ptr mutado que codifica una proteína de la Proteasa III que tiene actividad proteasa reducida o un gen ptr mutado inactivado, a menos que se indique lo contrario.
- La expresión "gen ptr mutado que codifica una proteína Proteasa III que tiene actividad proteasa reducida" en el contexto de la presente invención, significa que el gen ptr mutado no tiene la actividad proteasa completa en comparación con el gen ptr de tipo silvestre sin mutar.
- 15 Preferiblemente, el gen ptr mutado codifica una Proteasa III que tiene 50% o menos, 40% o menos, 30% o menos, 20% o menos, 10% o menos o 5% o menos de la actividad proteasa de una proteína Proteasa III de tipo silvestre sin mutar. Más preferiblemente, el gen ptr mutado codifica una proteína Proteasa III que no tiene actividad proteasa. En esta realización, la célula no carece de un ptr cromosómico, es decir, la secuencia del gen ptr no se ha delecionado o mutado para evitar la expresión de cualquier forma de proteína Proteasa III.
- 20 Cualquier mutación adecuada se puede introducir en el gen ptr con el fin de producir una proteína Proteasa III que tiene actividad proteasa reducida. La actividad proteasa de una proteína Proteasa III expresada a partir de una bacteria gram-negativa se puede someter a ensayo fácilmente por una persona experta en la técnica, a través de cualquier método adecuado en la técnica.
- 25 La expresión "gen ptr mutado inactivado" en el contexto de la presente invención, significa que el gen comprende una o varias mutaciones con lo que de este modo no se provoca ninguna expresión de la proteína codificada por el gen para proporcionar una célula que carece de la proteína codificada por el gen mutado inactivado. El gen inactivado se puede transcribir parcial o completamente, pero no se puede traducir en la proteína codificada. El gen ptr mutado inactivado puede estar mutado de cualquier manera adecuada, por ejemplo, mediante una o varias delecciones, inserciones, mutaciones puntuales, de sentido erróneo, sin sentido y de desplazamiento del marco, para causar una falta de expresión de la proteína. Por ejemplo, el gen se puede inactivar mediante la inserción de una secuencia de ADN extraño, tal como un marcador de resistencia a antibióticos, en la secuencia que codifica el gen.
- 30 En una realización preferida, el gen no está mutado mediante la inserción de una secuencia de ADN extraño, tal como un marcador de resistencia a antibióticos, en la secuencia que codifica el gen. Preferiblemente, el gen Proteasa III comprende una mutación en el codón de inicio del gen y/o uno o varios codones de parada situados aguas abajo del codón de inicio del gen y aguas arriba del gen de codón de parada, evitando de este modo la expresión de la proteína Proteasa III.
- 35 Una mutación en el codón de inicio del gen diana inactivado causa la pérdida de la función del codón de inicio y de ese modo asegura que el gen diana no comprende un codón de inicio adecuado al inicio de la secuencia codificante. La mutación del codón de inicio puede ser una mutación de sentido erróneo de uno, dos o los tres nucleótidos del codón de inicio. Alternativa o adicionalmente, el codón de inicio se puede mutar con una mutación del marco de lectura por inserción o delección.
- En una realización preferida, el gen ptr está mutado para cambiar el codón de inicio ATG a ATT, como se muestra en la Figura 5a.
- 45 El gen ptr mutado inactivado puede comprender alternativa o adicionalmente uno o varios codones de parada situados aguas abajo del codón de inicio del gen y aguas arriba del codón de parada del gen. Preferiblemente, el gen ptr mutado inactivado comprende tanto una mutación de sentido erróneo en el codón de inicio como uno o varios codones de parada insertados.
- 50 Uno o varios de los codones de parada insertados son preferiblemente codones de parada en marco. Sin embargo, uno o varios de los codones de parada insertados pueden ser alternativa o adicionalmente codones de parada fuera de marco. Uno o varios de los codones de parada fuera de marco pueden ser necesarios para detener la traducción, en donde se cambia un codón de inicio fuera de marco a un codón de inicio en marco mediante una mutación de desplazamiento de marco de lectura por inserción o delección. Uno o varios de los codones de parada se pueden introducir mediante cualquier mutación adecuada incluyendo una mutación puntual sin sentido y una mutación de desplazamiento de marco de lectura. Uno o varios de los codones de parada se introducen preferiblemente mediante una mutación de desplazamiento de marco de lectura y/o una mutación por inserción, preferiblemente mediante la
- 55

sustitución de un segmento de la secuencia del gen por una secuencia que comprende un codón de parada. Por ejemplo, se puede insertar un sitio de restricción *Ase I*, el cual comprende el codón de parada TAA.

En una realización preferida, el gen *ptr* está mutado para insertar un codón de parada en marco mediante la inserción de un sitio de restricción *Ase I*, como se muestra en la Figura 5a. En una realización preferida, el gen *ptr* mutado inactivado tiene la secuencia de ADN de SEQ ID NO: 6. Las mutaciones que se han realizado en la secuencia del gen *ptr* mutado inactivado de SEQ ID NO: 6 se muestran en la Figura 5a.

Las mutaciones por inactivación del gen descritas anteriormente son ventajosas debido a que causan una alteración mínima o ninguna en el ADN cromosómico, aguas arriba o aguas abajo del sitio del gen *diana* inactivado y no requieren la inserción y la retención de ADN extraño, tal como marcadores de la resistencia a antibióticos, lo que puede afectar a la idoneidad de la célula para expresar una proteína de interés, en particular proteínas terapéuticas. En consecuencia, una o varias de las células de acuerdo con la presente invención pueden presentar características de crecimiento y/o expresión de la proteína mejoradas, en comparación con células en las que el gen de la proteasa se ha inactivado mediante la inserción de ADN extraño en la secuencia que codifica el gen.

En realizaciones de la presente invención, la célula es portadora de un gen *OmpT* mutado. Tal y como se usa en el presente documento, "gen *OmpT*" significa un gen que codifica la proteasa *OmpT* (proteasa T de la membrana externa) que es una proteasa de la membrana externa. La secuencia del gen *OmpT* de tipo silvestre sin mutar es SWISS-PROT P09169.

La referencia a un gen mutado *OmpT* o un gen *OmpT* mutado que codifica *OmpT*, se refiere a cualquier gen *OmpT* mutado que codifica una proteína *OmpT* que tiene actividad proteasa reducida o un gen *OmpT* mutado inactivado, a menos que se indique lo contrario.

La expresión "gen *OmpT* mutado que codifica una proteína *OmpT* que tiene actividad proteasa reducida" en el contexto de la presente invención, significa que el gen *OmpT* mutado no tiene la actividad proteasa completa en comparación con el gen *OmpT* de tipo silvestre sin mutar. El gen *OmpT* mutado puede codificar una proteína *OmpT* que tiene 50% o menos, 40% o menos, 30% o menos, 20% o menos, 10% o menos o 5% o menos de la actividad proteasa de una proteína *OmpT* de tipo silvestre sin mutar. El gen *OmpT* mutado puede codificar una proteína *OmpT* que no tiene actividad proteasa. En esta realización, la célula no carece de *OmpT* cromosómica, es decir, la secuencia del gen *OmpT* no se ha deletado o mutado para evitar la expresión de cualquier forma de proteína *OmpT*.

Cualquier mutación adecuada se puede introducir en el gen *OmpT* con el fin de producir una proteína que tenga actividad proteasa reducida. La actividad proteasa de una proteína *OmpT* expresada a partir de una bacteria gramnegativa puede ser comprobada fácilmente por una persona experta en la técnica, a través de cualquier método adecuado en la técnica, tal como el método descrito en Kramer et al. (Identification of essential acidic residues of outer membrane protease *OmpT* supports a novel active site, FEBS Letters 505 (2001) 426-430) y Dekker et al. (Substrate Specificity of the Integral Membrane Protease *OmpT* Determined by Spatially Addressed Peptide Libraries, Biochemistry 2001, 40, 1694-1701).

*OmpT* se ha descrito en Kramer et al. (Identification of active site serine and histidine residues in *Escherichia coli* outer membrane protease *OmpT* FEBS Letters 2000 468, 220-224) los cuales describen que la sustitución de serinas, histidinas y residuos ácidos por alaninas da como resultado una actividad reducida ~10 veces para Glu27, Asp97, Asp208 o His101, una actividad reducida ~500 veces para Ser99 y una actividad reducida ~10000 veces para Asp83, Asp85, Asp210 o His212. Vandeputte-Rutten et al. (Crystal Structure of the Outer Membrane Protease *OmpT* from *Escherichia coli* suggests a novel catalytic site, The EMBO Journal 2001, vol 20 n° 18 5033-5039) describen que tiene un sitio activo que comprende una pareja Asp83-Asp85 y una pareja His212-Asp210. Además, Kramer et al. (Lipopolysaccharide regions involved in the activation of *Escherichia coli* outer membrane protease *OmpT*, Eur. J. Biochem. FEBS 2002, 269, 1746-1752) describen que las mutaciones D208A, D210A, H212A, H212N, H212Q, G216K/K217G, K217T y R218L en el bucle L4, dieron lugar todas a una pérdida parcial o prácticamente completa de la actividad enzimática.

Por consiguiente, la mutación *OmpT* para producir una proteína que tiene actividad proteasa reducida puede comprender una mutación, tal como una mutación de sentido erróneo en uno o varios de los residuos E27, D43, D83, D85, D97, S99, H101 E111, E136, E193, D206, D208, D210, H212 G216, K217, R218 & E250.

Uno o varios de E27, D43, D83, D85, D97, S99, H101 E111, E136, E193, D206, D208, D210, H212 G216, K217, R218 & E250 se pueden mutar a cualquier aminoácido adecuado, el cual da como resultado una proteína que tiene actividad proteasa reducida. Por ejemplo, uno o varios de E27, D43, D83, D85, D97, S99, H101 E111, E136, E193, D206, D208, D210, H212 G216, K217, R218 & E250 se pueden mutar a alanina. Ejemplos de mutaciones adecuadas son E27A, D43A, D83A, D85A, D97A, S99A, H101A E111A, E136A, E193A, D206A, D208A, D210A, H212A, H212N, H212Q, G216K, K217G, K217T, R218L & E250A. En una realización, el gen *OmpT* mutado comprende las mutaciones D210A y H212A. Una secuencia de *OmpT* mutada adecuada que comprende las mutaciones D210A y H212A se muestra en SEQ ID NO: 23.

La expresión "gen *OmpT* mutado inactivado" en el contexto de la presente invención, significa que el gen comprende una o varias mutaciones que provocan de este modo la falta de expresión de la proteína codificada por el gen para

- proporcionar una célula que carece de la proteína codificada por el gen mutado inactivado. El gen inactivado se puede transcribir parcial o completamente, pero no se puede traducir en la proteína codificada. El gen OmpT mutado inactivado se puede mutar de cualquier manera adecuada, por ejemplo, mediante una o varias deleciones, inserciones, mutaciones puntuales, de sentido erróneo, sin sentido y de desplazamiento del marco, para causar la falta de expresión de la proteína. Por ejemplo, el gen se puede inactivar mediante la inserción de una secuencia de ADN extraño, tal como un marcador de resistencia a antibióticos, en la secuencia que codifica el gen.
- En una realización, el gen OmpT comprende una mutación del codón de inicio del gen y/o uno o varios codones de parada situados aguas abajo del codón de inicio del gen y aguas arriba del codón de parada del gen, evitando de este modo la expresión de la proteína OmpT. La mutación del codón de inicio puede ser una mutación de sentido erróneo de uno, dos o los tres nucleótidos del codón de inicio. Alternativa o adicionalmente, el codón de inicio se puede mutar por una mutación del marco de lectura por inserción o deleción.
- Una secuencia de OmpT mutada inactivada, adecuada se muestra en SEQ ID NO: 24.
- En una realización, la célula bacteriana gram-negativa según la presente invención no es portadora de un gen OmpT mutado inactivado, ya que carece de OmpT cromosómica.
- En una realización, la célula bacteriana gram-negativa según la presente invención no es portadora de un gen DegP mutado inactivado, ya que carece de DegP cromosómica. En una realización, la célula bacteriana gram-negativa según la presente invención no es portadora de un gen DegP mutado.
- En una realización, la célula bacteriana gram-negativa según la presente invención no es portadora de un gen ptr mutado inactivado, ya que carece de ptr cromosómica.
- Muchas mutaciones por ingeniería genética, incluyendo las mutaciones por inactivación de un gen, implican el uso de marcadores de resistencia a antibióticos que permiten la selección e identificación de células mutadas con éxito. Sin embargo, como se ha mencionado anteriormente, hay una serie de inconvenientes para el uso de marcadores de resistencia a antibióticos.
- Una realización adicional de la presente invención supera las desventajas anteriores de la utilización de marcadores de resistencia a antibióticos, en donde el gen Tsp mutado, opcionalmente la spr mutada, opcionalmente el gen DegP mutado y/o un gen ptr mutado y/o un gen OmpT mutado, están mutados para comprender uno o varios sitios de marcadores de restricción. Los sitios de restricción se modifican por ingeniería genética en el gen y son de origen no natural. Los sitios de marcadores de restricción son ventajosos porque permiten la detección e identificación de células modificadas correctamente que comprenden las mutaciones cromosómicas requeridas. Las células que han sido modificadas para ser portadoras de uno o varios de los genes de proteasas mutadas, se pueden analizar mediante PCR del ADN genómico a partir de lisados celulares, usando parejas de oligonucleótidos diseñados para amplificar una región del ADN genómico que comprende un sitio de marcador de restricción de origen no natural. El ADN amplificado se puede analizar entonces mediante electroforesis en gel de agarosa antes y después de la incubación con una enzima de restricción adecuada capaz de digerir el ADN en el sitio del marcador de restricción de origen no natural. La presencia de fragmentos de ADN después de la incubación con la enzima de restricción, confirma que las células se han modificado con éxito para ser portadoras de uno o varios de los genes mutados.
- En la realización en la que la célula es portadora de un gen ptr mutado inactivado que tiene la secuencia de ADN de SEQ ID NO: 6, las secuencias de cebadores oligonucleótidos que se muestran en SEQ ID NO: 17 y SEQ ID NO:18 se pueden utilizar para amplificar la región del ADN que comprende el sitio de restricción de *Ase I* de origen no natural, a partir del ADN genómico de células transformadas. El ADN genómico amplificado se puede incubar entonces con la enzima de restricción *Ase I* y se analiza por electroforesis en gel para confirmar la presencia del gen ptr mutado en el ADN genómico.
- En la realización en la que la célula comprende un gen Tsp mutado inactivado que tiene la secuencia de ADN de SEQ ID NO: 3 o los nucleótidos 7 a 2048 de SEQ ID NO: 3, las secuencias de cebadores oligonucleótidos mostradas en SEQ ID NO: 15 y SEQ ID NO:16 se pueden utilizar para amplificar la región del ADN que comprende el sitio de restricción *Ase I* de origen no natural, a partir del ADN genómico de células transformadas. El ADN genómico amplificado se puede incubar entonces con la enzima de restricción de *Ase I* y analizar por electroforesis en gel para confirmar la presencia del gen Tsp mutado en el ADN genómico.
- En la realización en la que la célula comprende un gen DegP mutado que tiene la secuencia de ADN de SEQ ID NO: 9, las secuencias de cebadores oligonucleótidos que se muestran en SEQ ID NO: 19 y SEQ ID NO: 20 se pueden utilizar para amplificar la región del ADN que comprende el sitio de restricción *Ase I* de origen no natural, a partir del ADN genómico de células transformadas. El ADN genómico amplificado se puede incubar entonces con la enzima de restricción *Ase I* y analizar por electroforesis en gel para confirmar la presencia del gen DegP mutado en el ADN genómico.
- Uno o varios de los sitios de restricción se pueden introducir mediante cualquier mutación adecuada, incluyendo deleciones, inserciones, mutaciones puntuales, de sentido erróneo, sin sentido y de desplazamiento del marco. Un sitio de restricción se puede introducir mediante la mutación del codón de inicio y/o una mutación para introducir uno

o varios de los codones de parada, como se ha descrito anteriormente. Esta realización es ventajosa debido a que el sitio del marcador de restricción es un marcador directo y único de las mutaciones inactivadoras de genes introducidas.

5 Un sitio de marcador de restricción se puede insertar, el cual comprende un codón de parada en marco, tal como un sitio de restricción *Ase I*. Esto es particularmente ventajoso debido a que el sitio de restricción insertado sirve como un sitio de marcador de restricción y un codón de parada para evitar la transcripción completa de la secuencia que codifica el gen. Por ejemplo, en la realización en la que un codón de parada se introduce en el gen *ptr* mediante la introducción de un sitio *Ase I*, esto también crea un sitio de restricción, como se muestra en la Figura 5a. Por ejemplo, en la realización en la que un codón de parada se introduce en el gen *Tsp* en el codón 21 mediante la introducción de un sitio *Ase I*, esto también crea un sitio de restricción, como se muestra en la Figura 5b.

10 Un sitio de marcador de restricción se puede insertar mediante la mutación del codón de inicio y, opcionalmente, una o varias mutaciones puntuales adicionales. En esta realización, el sitio del marcador de restricción es preferiblemente un sitio de restricción *EcoR I*. Esto es particularmente ventajoso debido a que la mutación del codón de inicio también crea un sitio de marcador de restricción. Por ejemplo, en la realización en la que el codón de inicio del gen *ptr* se cambia a ATT, esto crea un sitio marcador *EcoR I*, como se muestra en la Figura 5a. Por ejemplo, en la realización en la que el codón de inicio (codón 3) del gen *Tsp* se cambia de ATG a TCG, como se muestra en la Figura 1b, una mutación puntual adicional en el codón 2 de AAC de AAT y una mutación en el codón 3 de ATG a TCG, crea un sitio de restricción marcador *EcoR I*, como se muestra en la Figura 5b.

15 En la realización de la presente invención, en donde la célula es portadora de un gen *OmpT* mutado, uno o varios de los sitios de restricción se pueden introducir mediante cualquier mutación adecuada, incluyendo mediante una o varias deleciones, inserciones, mutaciones puntuales, de sentido erróneo, sin sentido y de desplazamiento del marco. Por ejemplo, en la realización en la que el gen *OmpT* comprende las mutaciones D210A y H212A, estas mutaciones introducen un sitio de restricción *HindIII* silencioso que se puede utilizar como marcador de selección.

20 En el gen *DegP* o el gen *spr*, un sitio de restricción marcador se puede introducir usando cambios a codones silenciosos. Por ejemplo, un sitio *Ase I* se puede utilizar como un sitio de restricción silencioso marcador, en el que el codón de parada TAA está fuera de marco, como se muestra en la Figura 5c para el gen *DegP* mutado.

25 En las realizaciones de la presente invención, en donde el gen *ptr* y/o el gen *Tsp* están mutados para codificar una Proteasa III o *Tsp* que tiene actividad proteasa reducida, se pueden introducir uno o varios sitios de restricción marcadores usando cambios a codones silenciosos.

30 Cualquier bacteria gram-negativa adecuada se puede utilizar como la célula parental para la producción de la célula recombinante de la presente invención. Una bacteria gram-negativa adecuada incluye *Salmonella typhimurium*, *Pseudomonas fluorescens*, *Erwinia carotovora*, *Shigella*, *Klebsiella pneumoniae*, *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Acinetobacter baumannii* y *E. coli*. Preferiblemente la célula parental es *E. coli*. Cualquier cepa adecuada de *E. coli* se puede usar en la presente invención, pero se emplea preferiblemente una cepa W3110 de tipo silvestre, tal como W3110 K-12.

35 Un inconveniente asociado con las cepas bacterianas que carecen de proteasas, creadas previamente y utilizadas para expresar proteínas recombinantes, es que comprenden mutaciones adicionales en genes implicados en el metabolismo celular y la replicación del ADN, tales como por ejemplo *phoA*, *fhuA*, *lac*, *rec*, *gal*, *ara*, *arg*, *thi* y *pro* en cepas de *E. coli*. Estas mutaciones pueden tener muchos efectos deletéreos en la célula hospedadora, incluidos los efectos sobre el crecimiento celular, la estabilidad, el rendimiento de la expresión de proteínas recombinantes y la toxicidad. Las cepas que tienen una o varias de estas mutaciones genómicas, en particular las cepas que tienen una cantidad elevada de estas mutaciones, pueden mostrar una pérdida de capacidad, lo que reduce la tasa de crecimiento bacteriano a un nivel que no es adecuado para la producción de proteínas a escala industrial. Además, cualquiera de las mutaciones genómicas anteriores puede afectar a otros genes en *cis* y/o en *trans* de forma perjudicial impredecible, alterando de este modo el fenotipo de la cepa, la capacidad y el perfil de proteínas. Además, el uso de células fuertemente mutadas no es generalmente adecuado para la producción de proteínas recombinantes para uso comercial, en particular de agentes terapéuticos, debido a que tales cepas tienen generalmente vías metabólicas defectuosas y por lo tanto pueden crecer mal o no crecer nada en medios mínimos o químicamente definidos.

40 En una realización preferida, las células solo son portadoras de las mutaciones mínimas para introducir el polinucleótido recombinante que codifica *DsbC*; la modificación requerida para reducir la actividad de la proteína *Tsp* y opcionalmente el gen *spr* mutado. Solo se realizan mutaciones mínimas en el genoma de la célula para introducir un polinucleótido recombinante y mutaciones. Las células no son portadoras de ninguna otra mutación, lo que puede tener efectos nocivos sobre el crecimiento de la célula y/o la capacidad para expresar una proteína de interés. En consecuencia, una o varias de las células hospedadoras recombinantes de la presente invención pueden mostrar una expresión proteica mejorada y/o características de crecimiento mejoradas en comparación con células que comprenden mutaciones por ingeniería genética adicionales en la secuencia genómica. Las células proporcionadas por la presente invención son también más adecuadas para el uso en la producción de proteínas terapéuticas, en comparación con células que comprenden alteraciones adicionales en el genoma celular.

Por consiguiente, la presente invención también proporciona una célula bacteriana gram-negativa que tiene un genoma isogénico a una célula bacteriana de tipo silvestre, excepto por la modificación recombinante necesaria para reducir la actividad de la proteína Tsp y opcionalmente el gen spr mutado. La célula es portadora adicionalmente de un polinucleótido recombinante que codifica DsbC y, opcionalmente, una secuencia de polinucleótido que codifica una proteína de interés.

En una realización preferida, la célula es isogénica a una célula de *E. coli* de tipo silvestre, tal como la cepa W3110, a excepción de la modificación necesaria para reducir la actividad de la proteína Tsp, en comparación con una célula de tipo silvestre y, opcionalmente, el gen spr mutado. La célula es portadora adicionalmente de un polinucleótido recombinante que codifica DsbC y, opcionalmente, una secuencia de polinucleótido que codifica una proteína de interés.

La célula de acuerdo con la presente invención puede comprender además una secuencia de polinucleótido que codifica una proteína de interés. La secuencia de polinucleótido que codifica la proteína de interés puede ser exógena o endógena. La secuencia de polinucleótido que codifica la proteína de interés puede estar integrada en el cromosoma del hospedador o puede no estar integrada en un vector, normalmente un plásmido. Preferiblemente, el polinucleótido está en un vector de expresión en la célula, causando con ello una alteración mínima en el genoma de la célula hospedadora.

"Proteína de interés" en el contexto de la presente memoria descriptiva se entiende que se refiere a un polipéptido para la expresión, por lo general un polipéptido recombinante. Sin embargo, la proteína de interés puede ser una proteína endógena expresada a partir de un gen endógeno en la célula hospedadora.

Tal como se utiliza en la presente memoria, un "polipéptido recombinante" se refiere a una proteína que se construye o produce empleando tecnología de ADN recombinante. La proteína de interés puede ser una secuencia exógena idéntica a una proteína endógena o una versión mutada de la misma, por ejemplo, con una actividad biológica atenuada, o un fragmento de la misma expresado a partir de un vector exógeno. Alternativamente, la proteína de interés puede ser una proteína heteróloga, que normalmente no se expresa en la célula hospedadora.

La proteína de interés puede ser cualquier proteína adecuada que incluye una proteína terapéutica, profiláctica o de diagnóstico.

En una realización, la proteína de interés es útil en el tratamiento de enfermedades o trastornos que incluyen enfermedades y trastornos inflamatorios, enfermedades y trastornos inmunes, trastornos fibróticos y cánceres.

Las expresiones "enfermedad o trastorno inflamatorio" y "enfermedad o trastorno inmunológico" incluyen artritis reumatoide, artritis psoriática, enfermedad de Still, enfermedad de Muckle Wells, psoriasis, enfermedad de Crohn, colitis ulcerosa, SLE (lupus eritematoso sistémico), asma, rinitis alérgica, dermatitis atópica, esclerosis múltiple, vasculitis, diabetes mellitus de tipo I, trasplantes y enfermedad de injerto frente al receptor.

La expresión "trastorno fibrótico" incluye fibrosis pulmonar idiopática (IPF), esclerosis sistémica (o esclerodermia), fibrosis renal, nefropatía diabética, nefropatía por IgA, hipertensión, enfermedad renal en fase terminal, fibrosis peritoneal (diálisis peritoneal ambulatoria continua), cirrosis hepática, degeneración macular asociada con la edad (ARMD), retinopatía, fibrosis reactiva cardíaca, cicatrización, queloides, quemaduras, úlceras de la piel, angioplastia, cirugía de derivación coronaria, artroplastia y cirugía de cataratas.

El término "cáncer" incluye un nuevo crecimiento maligno que surge del epitelio, se encuentra en la piel, o de modo más habitual, en el revestimiento de los órganos corporales, por ejemplo: mama, ovario, próstata, pulmón, riñón, páncreas, estómago, vejiga o intestino. Los cánceres tienden a infiltrarse en el tejido adyacente y a propagarse (metastatizar) hacia órganos distantes, por ejemplo: a hueso, hígado, pulmón o el cerebro.

La proteína puede ser un polipéptido proteolíticamente sensible, es decir, proteínas que son propensas a ser escindidas, susceptibles de escisión o escindidas por una o varias bacterias gram-negativas, tales como *E. coli*, proteasas, ya sea en el estado natural o durante la secreción. En una realización, la proteína de interés es proteolíticamente sensible a una proteasa seleccionada entre DegP, Proteasa III y Tsp. En una realización, la proteína de interés es proteolíticamente sensible a la proteasa Tsp. En una realización, la proteína de interés es proteolíticamente sensible a las proteasas DegP y Proteasa III. En una realización, la proteína de interés es proteolíticamente sensible a las proteasas DegP y Tsp. En una realización, la proteína de interés es proteolíticamente sensible a las proteasas Tsp y Proteasa III. En una realización, la proteína de interés es proteolíticamente sensible a las proteasas DegP, Proteasa III y Tsp.

Preferiblemente, la proteína es un polipéptido eucariota.

La proteína de interés expresada por las células de acuerdo con la invención puede ser, por ejemplo, un inmunógeno, una proteína de fusión que comprende dos proteínas heterólogas o un anticuerpo. Los anticuerpos para uso como la proteína de interés incluyen anticuerpos monoclonales, multivalentes, multiespecíficos, humanizados, completamente humanos o quiméricos. El anticuerpo puede ser de cualquier especie, pero se obtiene preferiblemente a partir de un anticuerpo monoclonal, un anticuerpo humano o un fragmento humanizado. El anticuerpo se puede

obtener a partir de cualquier clase (por ejemplo, IgG, IgE, IgM, IgD o IgA) o subclase de molécula de inmunoglobulina y se puede obtener a partir de cualquier especie, incluyendo, por ejemplo, ratón, rata, tiburón, conejo, cerdo, hámster, camello, llama, cabra o ser humano. Las porciones del fragmento de anticuerpo se pueden obtener de más de una especie, por ejemplo los fragmentos de anticuerpo pueden ser quiméricos. En un ejemplo las regiones constantes son de una especie y las regiones variables de otra.

5 El anticuerpo puede ser una molécula de anticuerpo completa con cadenas pesadas y ligeras de longitud completa o un fragmento del mismo, por ejemplo, VH, VL, VHH, Fab, Fab modificado, Fab', F(ab')<sub>2</sub>, Fv, fragmento scFv, Fab-Fv, o un anticuerpo con especificidad dual, tal como un Fab-dAb, como se describe en el documento PCT/GB2008/003331.

10 El anticuerpo puede ser específico de cualquier antígeno diana. El antígeno puede ser una proteína asociada a una célula, por ejemplo una proteína de la superficie celular sobre células tales como células bacterianas, células de levadura, linfocitos T, células endoteliales o células tumorales, o puede ser una proteína soluble. Los antígenos de interés también pueden ser cualquier proteína importante desde el punto de vista médico, tales como las proteínas hiperreguladas durante una enfermedad o infección, por ejemplo receptores y/o sus correspondientes ligandos.

15 Ejemplos particulares de proteínas de la superficie celular incluyen moléculas de adhesión, por ejemplo integrinas tales como integrinas  $\beta$ 1, por ejemplo, VLA-4, E-selectina, P-selectina o L-selectina, CD2, CD3, CD4, CD5, CD7, CD8, CD11a, CD11b, CD18, CD19, CD20, CD23, CD25, CD33, CD38, CD40, CD40L, CD45, CDW52, CD69, CD134 (OX40), ICOS, BCMP7, CD137, CD27L, CDCP1, CSF1 o CSF1-Receptor, DPCR1, DPCR1, dudulin2, FLJ20584, FLJ40787, HEK2, KIAA0634, KIAA0659, KIAA1246, KIAA1455, LTBP2, LTK, MAL2, MRP2, nectina de tipo 2,

20 NKCC1, PTK7, RAIG1, TCAM1, SC6, BCMP101, BCMP84, BCMP11, DTD, antígeno carcinoembrionario (CEA), globulina de grasa de la leche humana (HMFG1 y 2), antígenos del MHC de clase I y de clase II, KDR y VEGF, y cuando sea apropiado, receptores de los mismos.

Los antígenos solubles incluyen interleucinas tales como IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-8, IL-12, IL-13, IL-14, IL-16 o IL-17, tales como IL17A y/o IL17F, antígenos virales, por ejemplo antígenos del virus respiratorio sincitial o citomegalovirus, inmunoglobulinas, tales como IgE, interferones tales como interferón  $\alpha$ , interferón  $\beta$  o interferón  $\gamma$ , factor de necrosis tumoral TNF (anteriormente conocido como factor de necrosis tumoral- $\alpha$ ), factor de necrosis tumoral- $\beta$ , factores estimulantes de colonias tales como G-CSF o GM-CSF, y factores de crecimiento derivados de plaquetas tales como PDGF- $\alpha$ , y PDGF- $\beta$  y cuando sea apropiado, receptores de los mismos. Otros antígenos incluyen antígenos de la superficie celular bacteriana, toxinas bacterianas, virus, tales como de la gripe, EBV, HepA, B y C,

25 agentes de terrorismo biológico, radionúclidos y metales pesados, y venenos y toxinas de serpientes y arañas.

30 En una realización, se puede utilizar el anticuerpo para alterar funcionalmente la actividad del antígeno de interés. Por ejemplo, el anticuerpo puede neutralizar, suscitar antagonismo o suscitar agonismo de la actividad de dicho antígeno, directa o indirectamente.

35 La presente invención también proporciona una célula bacteriana gram-negativa recombinante que comprende un polinucleótido recombinante que codifica DsbC, un gen Tsp mutado, en donde el gen Tsp mutado codifica una proteína Tsp que tiene actividad proteasa reducida o es un gen Tsp mutado inactivado, y una secuencia de polinucleótido que codifica un anticuerpo o un fragmento que se une a antígeno del mismo, específico de TNF. En una realización preferida, la célula comprende además un gen spr mutante que codifica una proteína spr mutante.

40 En una realización preferida, la proteína de interés expresada por las células de acuerdo con la presente invención es un anticuerpo anti-TNF, más preferiblemente un Fab' anti-TNF, como se describe en el documento WO01/094585.

45 En una realización, el anticuerpo que tiene especificidad para TNF $\alpha$  humano, comprende una cadena pesada en la que el dominio variable comprende una CDR que tiene la secuencia mostrada en SEQ ID NO: 28 para CDRH1, la secuencia mostrada en SEQ ID NO: 29 o SEQ ID NO: 34 para CDRH2 o la secuencia mostrada en SEQ ID NO: 30 para CDRH3.

En una realización el anticuerpo comprende una cadena ligera en donde el dominio variable comprende una CDR que tiene la secuencia mostrada en SEQ ID NO: 31 para CDRL1, la secuencia mostrada en SEQ ID NO: 32 para CDRL2 o la secuencia mostrada en SEQ ID NO: 33 para CDRL3.

50 Las CDRs proporcionadas en las SEQ IDS NOs: 28 y 30 a 34 mencionadas anteriormente se obtienen a partir de un anticuerpo monoclonal de ratón hTNF40. Sin embargo, SEQ ID NO: 29 consiste en una CDR híbrida. La CDR híbrida comprende una parte de CDR2 de la cadena pesada de un anticuerpo monoclonal de ratón hTNF40 (SEQ ID NO: 34) y una parte de CDR2 de la cadena pesada procedente de una secuencia humana de la región V de la línea germinal del grupo 3.

55 En una realización, el anticuerpo comprende una cadena pesada en la que el dominio variable comprende una CDR que tiene la secuencia mostrada en SEQ ID NO:28 para CDRH1, la secuencia mostrada en SEQ ID NO:29 o SEQ ID NO:34 para CDRH2 o la secuencia mostrada en SEQ ID NO: 30 para CDRH3 y una cadena ligera en la que el dominio variable comprende una CDR que tiene la secuencia mostrada en SEQ ID NO: 31 para CDRL1, la secuencia mostrada en SEQ ID NO: 32 para CDRL2 o la secuencia mostrada en la SEQ ID NO: 33 para CDRL3.

En una realización, el anticuerpo comprende SEQ ID NO: 28 para CDRH1, SEQ ID NO: 29 o SEQ ID NO: 34 para CDRH2, SEQ ID NO:30 para CDRH3, SEQ ID NO:31 para CDRL1, SEQ ID NO:32 para CDRL2 y SEQ ID NO:33 para CDRL3. Preferiblemente, el anticuerpo comprende SEQ ID NO: 29 para CDRH2.

5 El anticuerpo anti-TNF es preferiblemente una molécula de anticuerpo injertada con CDR. En una realización preferida, el dominio variable comprende regiones marco aceptoras humanas y CDRs donantes no humanas.

Preferiblemente, la molécula de anticuerpo tiene especificidad para TNF humano (anteriormente conocido como TNF $\alpha$ ), en donde la cadena ligera comprende la región variable de la cadena ligera de SEQ ID NO: 11 y la cadena pesada comprende la región variable de la cadena pesada de SEQ ID NO: 12.

El anticuerpo anti-TNF es preferiblemente un fragmento Fab o Fab'.

10 Preferiblemente, la molécula de anticuerpo que tiene especificidad hacia TNF humano es un Fab' y tiene una secuencia de la cadena ligera que comprende o que consiste en SEQ ID NO: 13 y una secuencia de la cadena pesada que comprende o consiste en SEQ ID NO: 14.

Después de la expresión, los fragmentos de anticuerpo se pueden procesar adicionalmente, por ejemplo, mediante conjugación con otra entidad, tal como una molécula efectora.

15 La expresión 'molécula efectora', tal y como se utiliza en la presente memoria incluye, por ejemplo, agentes antineoplásicos, fármacos, toxinas (tales como toxinas enzimáticamente activas de origen bacteriano o vegetal y fragmentos de las mismas, p. ej., ricina y fragmentos de la misma), proteínas biológicamente activas, por ejemplo, enzimas, otros anticuerpos o fragmentos de anticuerpo, polímeros sintéticos o de origen natural, ácidos nucleicos y fragmentos de los mismos, p. ej., ADN, ARN y fragmentos de los mismos, radionúclidos, particularmente radioyoduro, radioisótopos, metales quelados, nanopartículas y grupos indicadores tales como compuestos fluorescentes o compuestos que pueden detectarse por espectroscopia RMN o RES. El efector molecular puede estar fijado al anticuerpo o a uno de sus fragmentos por cualquier método adecuado, por ejemplo el fragmento de anticuerpo puede modificarse para fijar al menos una molécula efectora como se describe en los documentos WO05/003171 o WO05/003170. Los documentos WO05/003171 o WO05/003170 describen además moléculas efectoras adecuadas.

25 En una realización, el anticuerpo o un fragmento del mismo, tal como un Fab, está PEGilado para generar un producto con las propiedades requeridas, por ejemplo, similar a los anticuerpos completos, si es necesario. Por ejemplo, el anticuerpo puede ser un Fab' anti-TNF- $\alpha$  PEGilado, como se describe en el documento WO01/094585, preferiblemente que tiene fijado a uno de los residuos de cisteína en el extremo C-terminal de la cadena pesada un grupo obtenido a partir de lisil-maleimida, en donde cada uno de los dos grupos amino del residuo de lisilo tiene unido covalentemente al mismo un residuo de metoxipoli(etilenglicol) que tiene un peso molecular de aproximadamente 20.000 Da, de modo que el peso molecular promedio total de los residuos de metoxipoli(etilenglicol) es de aproximadamente 40.000 Da, más preferiblemente el grupo obtenido a partir de lisil-maleimida es [1-[[[2-[[3-(2,5-dioxo-1-pirrolidinil)-1-oxopropil]amino]etil]amino]carbonil]-1,5-pentanodil]bis(iminocarbonilo).

30 La célula también puede comprender otras secuencias de polinucleótidos que codifican una o varias proteínas de interés.

35 En una realización una o varias de las proteínas hospedadoras de *E. coli* que en el tipo silvestre se sabe que se copurifican con la proteína recombinante de interés durante la purificación, se seleccionan para una modificación genética, como se describe en Humphreys et al. "Engineering of Escherichia coli to improve the purification of periplasmic Fab' fragments: changing the pI of the chromosomally encoded PhoS/PstS protein", Protein Expression and Purification 37 (2004) 109-118 y el documento WO04/035792). El uso de tales proteínas hospedadoras modificadas mejora el proceso de purificación de las proteínas de interés, especialmente anticuerpos, producidos en *E. coli* mediante una alteración de las propiedades físicas de las proteínas de *E. coli* seleccionadas, de modo que ya no se copurifican con el anticuerpo recombinante. Preferiblemente, la proteína de *E. coli* alterada se selecciona a partir de una o varias entre la proteína que se une a fosfato (PhoS/PstS), proteína que se une a dipéptido (DppA), proteína que se une a maltosa (MBP) y tiorredoxina.

40 En una realización, una propiedad física de una proteína hospedadora contaminante se altera por la adición de un marcador de aminoácidos en el extremo C-terminal o N-terminal. En una realización preferida, la propiedad física que se altera es el punto isoeléctrico y el marcador de aminoácido es un marcador de poli(ácido aspártico) fijado al extremo C-terminal. En una realización, las proteínas de *E. coli* alteradas por la adición de dicho marcador son la proteína que se une a dipéptido (DppA), proteína que se une a maltosa (MBP), proteína que se une a tiorredoxina y fosfato (PhoS/PstS). En una realización específica, el pI de la proteína que se une a fosfato de *E. coli* (PhoS/PstS) se reduce desde 7,2 hasta 5,1 mediante la adición de un marcador de poli(ácido aspártico) (poliD), que contiene 6 residuos de ácido aspártico en el extremo C-terminal.

45 También se prefiere la modificación de residuos específicos de la proteína contaminante de *E. coli* para alterar sus propiedades físicas, ya sea solos o en combinación con la adición de marcadores del extremo N o C-terminal. Tales cambios pueden incluir inserciones o deleciones para alterar el tamaño de las sustituciones de proteínas o aminoácidos para alterar el pI o la hidrofobicidad. En una realización, estos residuos están situados en la superficie de la

proteína. En una realización preferida, los residuos de la superficie de la proteína PhoS se alteran con el fin de reducir el pI de la proteína. Preferiblemente se evitan los residuos que han estado implicados por ser importantes en el enlace fosfato (Bass, documento de patente de EE.UU. nº 5.304.472) para mantener una proteína PhoS funcional. Preferiblemente se dirigen a los residuos de lisina que se proyectan lejos de la superficie de la proteína o que están en grandes grupos de residuos básicos o cerca de ellos. En una realización, la proteína PhoS tiene un marcador de hexa poli(ácido aspártico) fijado al extremo C-terminal mientras que se dirigen a los residuos de la superficie en el extremo opuesto de la molécula para su sustitución. Residuos de lisina seleccionados preferiblemente se sustituyen por ácido glutámico o ácido aspártico para conferir un mayor potencial de cambio de pI que cuando se cambian los residuos neutros a los ácidos. La designación para un mutante por sustitución en la presente memoria consiste en una letra seguida de un número seguido de una letra. La primera letra designa el aminoácido en la proteína de tipo silvestre. El número hace referencia a la posición del aminoácido en la que se está realizando la sustitución del aminoácido, y la segunda letra designa el aminoácido que se utiliza para remplazar al aminoácido de tipo silvestre. En mutaciones preferidas de PhoS, los residuos de lisina de la presente invención (K) 275, 107, 109, 110, 262, 265, 266, 309, 313 están sustituidos por ácido glutámico (E) o glutamina (Q), como mutaciones simples o combinadas, además, la lisina (K)318 puede estar sustituida por ácido aspártico (D) como una mutación simple o combinada. Preferiblemente las mutaciones individuales son K262E, K265E y K266E. Preferiblemente las mutaciones combinadas son K265/266E y K110/265/266E. Más preferiblemente, todas las mutaciones se combinan con el marcador de poli(ácido aspártico) (poliD) fijado al extremo C-terminal y opcionalmente también con la sustitución K318D. En una realización preferida las mutaciones dan lugar a una reducción del pI de al menos 2 unidades. Preferiblemente, las mutaciones de la presente invención reducen el pI de PhoS desde 7,2 a entre aproximadamente 4 y aproximadamente 5,5. En una realización de la presente invención, el pI de la proteína PhoS de *E. coli* se reduce desde 7,2 hasta aproximadamente 4,9, aproximadamente 4,8 y aproximadamente 4,5 utilizando las mutaciones poliD K318D, poliD K265/266E y poliD K110/265/266E respectivamente.

La célula bacteriana gram-negativa recombinante según la presente invención se puede producir por cualquier medio adecuado.

En las realizaciones de la presente invención en las que el gen Tsp cromosómico está mutado y opcionalmente uno o varios de los genes cromosómicos spr, DegP, ptr y OmpT están mutados, el experto conoce las técnicas adecuadas que se pueden usar para sustituir una secuencia de gen cromosómico por una secuencia de gen mutado. Se pueden emplear vectores adecuados, los cuales permiten la integración en el cromosoma hospedador mediante recombinación homóloga. En la realización en la que la célula comprende dos o tres de los genes mutados anteriores, los genes mutados se pueden introducir en la bacteria gram-negativa en el mismo vector o en diferentes vectores.

Se describen métodos de sustitución de genes adecuados, por ejemplo, en Hamilton *et al.* (New Method for Generating Deletions and Gene Replacements in Escherichia coli, Hamilton C. M. et al., Journal of Bacteriology Sept. 1989, Vol. 171, nº 9 p 4617-4622), Skorupski *et al.* (Positive selection vectors for allelic exchange, Skorupski K y Taylor R. K., Gene, 1996, 169, 47-52), Kiel *et al.* (A general method for the construction of Escherichia coli mutants by homologous recombination and plasmid segregation, Kiel J.A.K.W. et al., Mol Gen Genet 1987, 207:294-301), Blomfield *et al.* (Allelic exchange in Escherichia coli using the Bacillus subtilis sacB gene and a temperature sensitive pSC101 replicon, Blomfield I. C. et al., Molecular Microbiology 1991, 5(6), 1447-1457) y Ried *et al.* (An nptI-sacB-sacR cartridge for constructing directed, unmarked mutations in Gram-negative bacteria by marker exchange- eviction mutagenesis, Ried J. L. y Collmer A., Gene 57 (1987) 239-246). Un plásmido adecuado que permite la recombinación/sustitución homóloga es el plásmido pKO3 (Link et al., 1997, Journal of Bacteriology, 179, 6228-6237).

El experto conoce las técnicas adecuadas que se pueden emplear para insertar el polinucleótido recombinante que codifica DsbC. El polinucleótido recombinante que codifica DsbC puede estar integrado en el genoma de la célula usando un vector adecuado, tal como el plásmido pKO3.

Alternativa o adicionalmente, el polinucleótido recombinante que codifica DsbC puede no estar integrado en un casete de expresión recombinante. En una realización, un casete de expresión se emplea en la presente invención para ser portador del polinucleótido que codifica DsbC que normalmente comprende una secuencia que codifica una proteína que codifica DsbC y una o varias secuencias reguladoras de la expresión. Una o varias de las secuencias reguladoras de la expresión pueden incluir un promotor. Una o varias de las secuencias reguladoras de la expresión también pueden incluir una región no traducida 3' tal como una secuencia de terminación. Los promotores adecuados se describen con más detalle a continuación.

En una realización, se emplea un casete de expresión en la presente invención que es portador del polinucleótido que codifica la proteína de interés y/o el polinucleótido recombinante que codifica DsbC que comprende normalmente una o varias secuencias reguladoras de la expresión, una o varias secuencias codificantes que codifican una o varias proteínas de interés y/o una secuencia codificante que codifica DsbC. Una o varias de las secuencias reguladoras de la expresión pueden incluir un promotor. Una o varias de las secuencias reguladoras de la expresión también pueden incluir una región no traducida 3' tal como una secuencia de terminación. Los promotores adecuados se describen con más detalle a continuación.

5 En una realización, la célula de acuerdo con la presente invención comprende uno o varios vectores, tales como un plásmido. El vector comprende preferiblemente uno o varios de los casetes de expresión tal y como se han definido anteriormente. En una realización, la secuencia de polinucleótido que codifica una proteína de interés y el polinucleótido que codifica DsbC se insertan en un vector. Alternativamente, la secuencia de polinucleótido que codifica una proteína de interés y el polinucleótido que codifica DsbC se insertan en vectores distintos.

10 En la realización en la que la proteína de interés es un anticuerpo que comprende las cadenas tanto pesadas como ligeras, la línea celular se puede transfectar con dos vectores, un primer vector que codifica un polipéptido de la cadena ligera y un segundo vector que codifica un polipéptido de la cadena pesada. Alternativamente, se puede utilizar un único vector, incluyendo el vector secuencias que codifican los polipéptidos de la cadena ligera y la cadena pesada. Alternativamente, la secuencia de polinucleótido que codifica el anticuerpo y el polinucleótido que codifica DsbC, se insertan en un vector. Preferiblemente, el vector comprende las secuencias que codifican los polipéptidos de la cadena ligera y pesada del anticuerpo.

En la realización en la que la célula también expresa una o varias proteínas adicionales del modo siguiente:

- 15
- una o varias proteínas capaces de facilitar el plegamiento de proteínas, tales como FkpA, Skp, SurA, PPIA y PPIID; y/o
  - una o varias proteínas capaces de facilitar la secreción o translocación de proteínas, tales como SecY, SecE, SecG, SecYEG, SecA, SecB, FtsY y Lep; y/o
  - una o varias proteínas capaces de facilitar la formación de un enlace disulfuro, tales como DsbA, DsbB, DsbD, DsbG;

20 la proteína adicional o varias proteínas adicionales se pueden expresar a partir de uno o varios polinucleótidos insertados en el mismo vector que el polinucleótido que codifica DsbC y/o la secuencia de polinucleótido que codifica una proteína de interés. Alternativamente, uno o varios de los polinucleótidos se pueden insertar en vectores distintos.

25 El vector para uso en la presente invención se puede producir mediante la inserción de uno o varios casetes de expresión tal y como se han definido anteriormente, en un vector adecuado. Alternativamente, las secuencias reguladoras de la expresión para dirigir la expresión de la secuencia de polinucleótido, pueden estar contenidas en el vector y por lo tanto solo la región codificante del polinucleótido puede ser necesaria para completar el vector.

30 El polinucleótido que codifica DsbC y/o el polinucleótido que codifica la proteína de interés se insertan adecuadamente en un vector replicable, normalmente un vector de replicación autónoma, para la expresión en la célula bajo el control de un promotor adecuado para la célula. Muchos vectores son conocidos en la técnica para este fin y la selección del vector apropiado puede depender del tamaño del ácido nucleico y del tipo de célula particular.

Ejemplos de vectores que se pueden emplear para transformar la célula hospedadora con un polinucleótido de acuerdo con la invención incluyen:

- 35
- un plásmido, tal como pBR322 o pACYC184, y/o
  - un vector vírico tal como un fago bacteriano
  - un elemento genético transponible tal como un transposón.

Tales vectores comprenden normalmente un origen plasmídico de replicación del ADN, un marcador seleccionable de antibióticos, un promotor y un terminador transcripcional separado por un sitio de clonación múltiple (casete de expresión) y una secuencia de ADN que codifica un sitio de unión al ribosoma.

40 Los promotores empleados en la presente invención se pueden enlazar con el polinucleótido relevante directamente o de forma alternativa, estar situados en una posición apropiada, por ejemplo, en un vector de modo que cuando se inserta el polipéptido relevante, el promotor relevante puede actuar sobre el mismo. En una realización, el promotor se localiza antes de la porción que codifica el polinucleótido sobre el que actúa, por ejemplo, un promotor relevante antes de cada porción que codifica el polinucleótido. "Antes", tal y como se usa en el presente documento, se entiende que significa que el promotor se encuentra en el extremo 5' en relación con la porción que codifica el polinucleótido.

45

Los promotores pueden ser endógenos o exógenos a las células hospedadoras. Los promotores adecuados incluyen Lac, tac, trp, PhoA, lpp, Arab, Tet y T7.

Uno o varios de los promotores empleados pueden ser promotores inducibles.

50 En la realización en la que el polinucleótido que codifica DsbC y el polinucleótido que codifica la proteína de interés se insertan en un vector, las secuencias de nucleótidos que codifican DsbC y la proteína de interés pueden estar bajo el control de un único promotor o de promotores distintos. En la realización en la que las secuencias de nucleó-

tidos que codifican DsbC y la proteína de interés están bajo el control de promotores distintos, el promotor pueden ser promotores inducibles independientemente.

Los promotores para uso en sistemas bacterianos también contienen generalmente una secuencia de Shine-Dalgarno (S.D.) ligada funcionalmente al ADN que codifica el polipéptido de interés. El promotor se puede eliminar del ADN de origen bacteriano mediante digestión con enzimas de restricción e insertar en el vector que contiene el ADN deseado.

En las realizaciones de la presente invención en las que una secuencia de polinucleótido comprende dos o más secuencias que codifican dos o más proteínas de interés, por ejemplo una cadena ligera de anticuerpo y una cadena pesada de anticuerpo, la secuencia de polinucleótido puede comprender una o varias secuencias internas de sitios de entrada al ribosoma (IRES) que permiten la iniciación de la traducción en el centro de un ARNm. Una secuencia IRES puede estar situada entre las secuencias de polinucleótidos codificantes para potenciar una traducción independiente del ARNm para producir las secuencias de polipéptidos codificados.

El vector de expresión también comprende preferiblemente un mensaje dicistrónico para producir el anticuerpo o un fragmento del mismo que se une a antígeno, tal y como se describe en el documento WO 03/048208 o WO2007/039714. Preferiblemente, el cistrón aguas arriba contiene ADN que codifica la cadena ligera del anticuerpo y el cistrón aguas abajo contiene ADN que codifica la cadena pesada correspondiente, y la secuencia intergénica dicistrónica (IGS) comprende preferiblemente una secuencia seleccionada a partir de IGS1 (SEQ ID NO: 36), IGS2 (SEQ ID NO: 37), IGS3 (SEQ ID NO: 38) y IGS4 (SEQ ID NO: 39).

Los terminadores pueden ser endógenos o exógenos con relación a las células hospedadoras. Un terminador adecuado es rrnB.

Unos reguladores transcripcionales adecuados adicionales incluyen promotores y terminadores y los métodos de direccionamiento de proteínas se pueden encontrar en "Strategies for Achieving High-Level Expression of Genes in "Escherichia coli" Sawas C. Makrides, Microbiological Reviews, Sept 1996, p 512-538.

El polinucleótido DsbC insertado en el vector de expresión comprende preferiblemente el ácido nucleico que codifica la secuencia señal de DsbC y la secuencia que codifica DsbC. El vector contiene preferiblemente una secuencia de ácido nucleico que permite que el vector se replique en una o varias células hospedadoras seleccionadas, preferiblemente que se replique independientemente del cromosoma hospedador. Tales secuencias son bien conocidas para una variedad de bacterias.

En una realización, la DsbC y/o la proteína de interés comprende un marcador de histidina en el extremo N-terminal y/o C-terminal.

La molécula de anticuerpo puede ser secretada desde la célula o estar dirigida al periplasma mediante secuencias señal adecuadas. Alternativamente, las moléculas de anticuerpo se pueden acumular en el citoplasma de la célula. Preferiblemente la molécula de anticuerpo está dirigida al periplasma.

El polinucleótido que codifica la proteína de interés se puede expresar como una fusión con otro polipéptido, preferiblemente una secuencia señal u otro polipéptido que tiene un sitio de escisión específico en el extremo N-terminal del polipéptido maduro. La secuencia señal heteróloga seleccionada debe ser una que sea reconocida y procesada por la célula hospedadora. Para las células hospedadoras procariontas que no reconocen y procesan la secuencia señal del polipéptido natural o eucariota, la secuencia señal se sustituye por una secuencia señal procarionta. Las secuencias señal adecuadas incluyen *OmpA*, *PhoA*, *LamB*, *PeIB*, *DsbA* y *DsbC*.

La construcción de vectores adecuados que contienen uno o varios de los componentes anteriormente enumerados, emplea técnicas de ligación convencionales. Los plásmidos aislados o fragmentos de ADN se escinden, se adaptan y se ligan de nuevo en la forma deseada para generar los plásmidos requeridos.

En una realización preferida de la presente invención, la presente invención proporciona un vector multicistrónico que comprende la secuencia de polinucleótido que codifica DsbC y la secuencia de polinucleótido que codifica una proteína de interés. El vector multicistrónico se puede producir por un método de clonación ventajoso que permite una clonación secuencial repetida de secuencias de polinucleótidos en un vector. El método utiliza extremos cohesivos compatibles de una pareja de sitios de restricción, tales como los extremos "AT" de los sitios de restricción *Ase I* y *Nde I*. Una secuencia de polinucleótido que comprende una secuencia codificante y que tiene extremos cohesivos compatibles, tales como un fragmento *AseI-NdeI*, se puede clonar en un sitio de restricción en el vector, tal como *Nde I*. La inserción de la secuencia de polinucleótido destruye el sitio de restricción 5' pero crea un nuevo sitio de restricción 3', tal como *Nde I*, que después se puede utilizar para insertar una secuencia de polinucleótido adicional que comprende extremos cohesivos compatibles. El proceso se puede repetir después para insertar secuencias adicionales. Cada secuencia de polinucleótido insertada en el vector comprende una secuencia no codificadora en 3' del codón de parada que puede comprender un sitio *Ssp I* para el escrutinio, una secuencia de unión al ribosoma de Shine Dalgarno, un espaciador rico en A y un sitio *NdeI* que codifica un codón de inicio.

Una representación esquemática de la creación de un vector que comprende una secuencia de polinucleótido que codifica una cadena ligera de un anticuerpo (LC), una cadena pesada de un anticuerpo (HC), una secuencia de polinucleótido DsbC y una secuencia de polinucleótido adicional, se muestra en la Figura 6.

5 Las cepas mutadas con éxito se pueden identificar usando procedimientos bien conocidos en la técnica, que incluyen la secuenciación de ADN con PCR de colonias y el cartografiado con enzimas de restricción y PCR de colonias.

Las realizaciones de la invención descritas en este documento con referencia al polinucleótido, se aplican igualmente a realizaciones alternativas de la invención, por ejemplo, vectores, casetes de expresión y/o células hospedadoras que comprenden los componentes empleados en la misma, por lo que el aspecto relevante se puede aplicar a las mismas.

10 De acuerdo con un tercer aspecto de la presente invención, se proporciona un método para producir una proteína recombinante de interés que comprende expresar la proteína recombinante de interés en una célula bacteriana gram-negativa recombinante como se ha descrito anteriormente en el primer o segundo aspecto de la presente invención. El método para producir una proteína recombinante de interés comprende:

15 a. cultivar una célula bacteriana gram-negativa recombinante como se ha definido anteriormente en un medio de cultivo en condiciones eficaces para expresar la proteína recombinante de interés y el polinucleótido recombinante que codifica DsbC; y

b. recuperar la proteína recombinante de interés a partir del periplasma de la célula bacteriana gram-negativa recombinante y/o el medio de cultivo.

20 La célula bacteriana gram-negativa y la proteína de interés empleadas preferiblemente en el método de la presente invención, se han descrito con detalle anteriormente.

25 Cuando el polinucleótido que codifica la proteína de interés es exógeno, el polinucleótido se puede incorporar en la célula hospedadora usando cualquier medio adecuado conocido en la técnica. Como se ha descrito anteriormente, normalmente el polinucleótido se incorpora como parte de un vector de expresión que se transforma en la célula. Por consiguiente, en un aspecto, la célula de acuerdo con la presente invención comprende un casete de expresión que comprende el polinucleótido que codifica la proteína de interés y un casete de expresión que comprende el polinucleótido que codifica DsbC.

La secuencia del polinucleótido que codifica la proteína de interés y la secuencia del polinucleótido que codifica DsbC se pueden transformar en una célula utilizando técnicas convencionales, por ejemplo, empleando cloruro de rubidio, PEG o electroporación.

30 El método de acuerdo con la presente invención también puede emplear un sistema de selección para facilitar la selección de las células estables que han sido transformadas con éxito con el polinucleótido que codifica la proteína de interés. El sistema de selección emplea normalmente una cotransformación de una secuencia de polinucleótido que codifica un marcador de selección. En una realización, cada polinucleótido transformado en la célula comprende además una secuencia de polinucleótido que codifica uno o varios marcadores de selección. Por consiguiente, la transformación de los polinucleótidos que codifican DsbC y la proteína de interés y uno o varios de los polinucleótidos que codifican el marcador, se produce conjuntamente y el sistema de selección se puede emplear para seleccionar las células que producen las proteínas deseadas.

35 40 Las células capaces de expresar uno o varios de los marcadores, son capaces de sobrevivir/crecer/multiplicarse bajo ciertas condiciones impuestas artificialmente, por ejemplo, la adición de una toxina o antibiótico, debido a las propiedades con las que están dotadas gracias al polipéptido/gen o el componente de polipéptido del sistema de selección incorporado en las mismas (por ejemplo, resistencia a antibióticos). Las células que no son capaces de expresar uno o varios de los marcadores no son capaces de sobrevivir/crecer/multiplicarse en las condiciones impuestas de modo artificial. También puede elegirse que las condiciones impuestas de modo artificial sean más o menos vigorosas, según se requiera.

45 En la presente invención puede emplearse cualquier sistema de selección adecuado. Normalmente, el sistema de selección puede estar basado en la inclusión en el vector de uno o varios genes que proporcionan resistencia a un antibiótico conocido, por ejemplo, un gen de resistencia a tetraciclina, cloranfenicol, kanamicina o ampicilina. Las células que crecen en presencia de un antibiótico relevante se pueden seleccionar ya que expresan tanto el gen que proporciona resistencia al antibiótico como la proteína deseada.

50 En una realización, el método según la presente invención comprende además la etapa de cultivar la célula transformada en un medio para expresar de este modo la proteína de interés.

Un sistema de expresión inducible o un promotor constitutivo se pueden usar en la presente invención para expresar la proteína de interés y/o la DsbC. Los sistemas de expresión inducible y los promotores constitutivos adecuados son bien conocidos en la técnica.

En una realización en la que el polinucleótido que codifica DsbC y el polinucleótido que codifica la proteína de interés están bajo el control de los mismos o de distintos promotores inducibles, la expresión de la secuencia de polinucleótido que codifica una proteína de interés y el polinucleótido recombinante que codifica DsbC, se induce mediante la adición de un inductor al medio de cultivo.

- 5 Cualquier medio adecuado se puede emplear para cultivar la célula transformada. El medio puede estar adaptado a un sistema de selección específica, por ejemplo, el medio puede comprender un antibiótico, para permitir que solo aquellas células que han sido transformadas con éxito, puedan crecer en el medio.

Las células obtenidas a partir del medio pueden someterse a más escrutinios y/o purificaciones según sea necesario. El método puede comprender además una o varias etapas para extraer y purificar la proteína de interés, según sea necesario.

10 El polipéptido se puede recuperar a partir de la cepa, incluyendo desde el citoplasma, el periplasma o el material sobrenadante.

El o los métodos específicos que se utilizan para purificar una proteína, dependen del tipo de proteína. Los métodos adecuados incluyen el fraccionamiento sobre columnas de inmunoafinidad o de intercambio iónico; precipitación con etanol; HPLC de fase inversa; cromatografía de interacción hidrófoba; cromatografía sobre sílice; cromatografía en una resina de intercambio iónico tal como S-Sepharosa y DEAE; cromatografía; precipitación en sulfato de amonio; y filtración en gel.

En una realización, el método comprende además separar la proteína recombinante de interés de DsbC.

20 Los anticuerpos se pueden separar adecuadamente a partir del medio de cultivo y/o de un extracto de citoplasma y/o de un extracto de periplasma mediante procedimientos de purificación de anticuerpos convencionales, tales como, por ejemplo, proteína A-Sepharosa, cromatografía de la proteína G, cromatografía de la proteína L, resinas en modo mezclado, interacción tiofílica, marcador His, marcador Flag, cromatografía de hidroxapatita, electroforesis en gel, diálisis, cromatografía de afinidad, sulfato de amonio, etanol o fraccionamiento/precipitación con PEG, membranas de intercambio iónico, cromatografía de adsorción en lecho expandido (EBA) o cromatografía en lecho móvil simulado.

El método también puede incluir una etapa adicional de medición de la cantidad de expresión de la proteína de interés y la selección de células que tienen altos niveles de expresión de la proteína de interés.

El método también puede incluir una o varias etapas de procesamiento adicional aguas abajo, tales como la PEGilación de la proteína de interés, tal como un anticuerpo o un fragmento de anticuerpo.

30 Una o varias etapas del método descrito en la presente pueden realizarse en combinación en un recipiente adecuado, tal como un biorreactor.

## Ejemplos

### Líneas celulares

Para todos los experimentos se utilizó la celular W3110 de *E. coli* como línea celular natural original.

35 Se crearon líneas celulares portadoras de las mutaciones siguientes:

- a. un gen Tsp mutado;
- b. un gen Tsp mutado y portador de DsbC recombinante;
- c. un gen Tsp mutado y un gen spr mutado;
- d. un gen Tsp mutado y un gen spr mutado y portador de DsbC recombinante.

40 **Ejemplo 1 Generación de la línea celular MXE001 que es portadora del gen Tsp mutado ( $\Delta$ Tsp)**

La cepa MXE001 se generó de la manera siguiente:

45 El casete Tsp se trasladó como fragmentos de restricción Sal I, Not I a plásmidos pKO3 similarmente restringidos. El plásmido pKO3 utiliza el mutante sensible a la temperatura del origen de replicación (*RepA*) de pSC101 junto con un marcador de cloranfenicol para forzar y seleccionar eventos de integración cromosómica. El gen *sacB* que codifica la levansacarasa es letal para el cultivo de *E. coli* en sacarosa y por consiguiente (junto con el marcador de cloranfenicol y el origen de pSC101) se utiliza para forzar y seleccionar eventos de desintegración y curado de plásmidos. Esta metodología se ha descrito anteriormente (Hamilton *et al.*, 1989, *Journal of Bacteriology*, 171, 4617-4622 y Blomfield *et al.*, 1991, *Molecular Microbiology*, 5, 1447-1457). El sistema pKO3 elimina todos los marcadores selectivos del genoma del hospedador excepto el del gen insertado.

Se construyeron los plásmidos siguientes.

pMXE191 que comprende el gen Tsp mutado inactivado como se muestra en SEQ ID NO: 3 que comprende los marcadores de restricción *EcoR I* y *Ase I*.

5 El plásmido se transformó a continuación en células W3110 electrocompetentes de *E. coli* competente, preparadas utilizando el método encontrado en Miller, E.M. y Nickoloff, J.A., "*Escherichia coli* electrotransformation" en *Methods in Molecular Biology*, vol. 47, Nickoloff, J.A. (compilador), Humana Press, Totowa, NJ, 105 (1995).

10 Día 1 Se mezclaron 40 µl de células de *E. coli* con (10 pg) 1 µl de ADN de pKO3 en una cubeta de electroporación de 0,2 cm de BioRad refrigerada antes de la electroporación a 2.500 V, 25 µF y 200 Ω. Se añadieron inmediatamente 1000 µl de 2xPY, las células se recuperaron mediante agitación a 250 rpm en una incubadora a 30°C durante 1 hora. Las células se diluyeron 1/10 en serie en 2xPY antes de que se extendieran 100 µl de partes alícuotas sobre placas con 2xPY agar-agar que contenían cloranfenicol a 20 µg/ml precalentadas a 30°C y 43°C. Las placas se incubaron durante la noche a 30°C y 43°C.

15 Día 2 El número de colonias cultivadas a 30°C dio una estimación de la eficiencia de la electroporación mientras que las colonias que sobreviven creciendo a 43°C representan potenciales eventos de integración. Se seleccionaron colonias aisladas a partir de la placa a 43°C y se volvieron a poner en suspensión en 10 ml de 2xPY. Se sembraron 100 µl de esta suspensión sobre placas 2xPY con agar-agar que contenía sacarosa al 5% (p/v), precalentadas a 30°C para generar colonias aisladas. Las placas se incubaron durante la noche a 30°C.

20 Día 3 Las colonias aquí representan potenciales eventos simultáneos de desintegración y curado de plásmidos. Si los eventos de desintegración y curado ocurrían al principio del cultivo, entonces la mayor parte de la masa de la colonia era clonal. Se seleccionaron colonias aisladas y se sembraron en placas por duplicado sobre 2xPY agar-agar que contenía ya sea 20 µg/ml de cloranfenicol o sacarosa al 5% (p/v). Las placas se incubaron durante la noche a 30°C.

25 Día 4 Las colonias que crecen sobre sacarosa y mueren sobre cloranfenicol representan potenciales eventos de sustitución cromosómica y curado de plásmidos. Éstas se seleccionaron y escrutaron mediante PCR con un oligonucleótido específico para mutaciones. Las colonias que generaban una banda positiva de PCR con el tamaño correcto se separaban para producir colonias aisladas sobre 2xPY agar-agar que contenía sacarosa al 5% (p/v) y las placas se incubaron durante la noche a 30°C.

30 **Día 5** Colonias individuales de *E. coli* positivas para PCR, sensibles al cloranfenicol y resistentes a la sacarosa se utilizaron para preparar reservas en glicerol de células químicamente competentes y que actuaban como plantillas de PCR para una reacción de PCR con oligos flanqueantes en 5' y 3', para generar un producto de PCR para la secuenciación directa de ADN usando Taq polimerasa.

35 La cepa celular MXE001 se sometió a ensayo para confirmar la modificación con éxito del ADN genómico que era portador del gen Tsp mutado, mediante amplificación con PCR de la región del gen Tsp que comprendía un sitio de restricción *Ase I* de origen no natural, como se muestra en las Figuras 1a, 1b y 1c, utilizando cebadores oligonucleótidos. Las regiones amplificadas del ADN se analizaron a continuación por electroforesis en gel antes y después de la incubación con la enzima de restricción *Ase I* para confirmar la presencia del sitio de restricción artificial de *Ase I* en los genes mutados. Este método se llevó a cabo del modo siguiente:

Los siguientes oligos se utilizaron para amplificar, mediante PCR, ADN genómico procedente de lisados celulares de *E. coli* preparados a partir de MXE001 y W3110:

6284 Tsp 3'	5'-GCATCATAATTTTCTTTTACCTC-3' (SEQ ID NO: 15)
6283 Tsp 5'	5'-GGGAAATGAACCTGAGCAAACGC-3' (SEQ ID NO: 16)

40 Se prepararon lisados calentando una colonia aislada de células durante 10 minutos a 95°C en 20 µl de 1x tampón de PCR. Se dejó enfriar la mezcla a temperatura ambiente, a continuación centrifugación a 13.200 rpm durante 10 minutos. El material sobrenadante se retiró y se marcó como "lisado celular".

El ADN se amplificó utilizando un procedimiento de PCR convencional.

5 µl	Tampón x10 (Roche)
1 µl	Mezcla de dNTP (Roche, mezcla 10 mM)

1,5 µl	oligo 5' (5 pmol)
1,52 µl	oligo 3' (5 pmol)
2 µl	Lisado celular
0,5 µl	Taq ADN polimerasa (Roche 5U/µl )
38,5 µl	H <sub>2</sub> O

Ciclo de PCR.

94°C	1 minuto
94°C	1 minuto)
55°C	1 minuto) repetido durante 30 ciclos
72°C	1 minuto)
72°C	10 minutos

5 Una vez completadas las reacciones, se extrajeron 25 µl a un nuevo tubo de microcentrífuga para la digestión con *Ase I*. A los 25 µl de reacción de PCR se añadieron 19 µl de H<sub>2</sub>O, 5 µl de tampón 3 (NEB), 1 µl de *Ase I* (NEB), se mezclaron e incubaron a 37°C durante 2 horas.

10 A la reacción PCR restante se añadieron 5 µl de tampón de carga (x6) y se cargaron 20 µl sobre 200 ml de gel de agarosa TAE al 0,8% (Invitrogen) más bromuro de etidio (5 µl de 10 mg/ml de solución madre) y se ejecutaron a 100 voltios durante 1 hora. Se cargaron 10 µl de marcador de tamaño (Perfect DNA marker 0,1-12 Kb, Novagen) en el carril final.

Una vez terminadas las digestiones con *Ase I*, se añadieron 10 µl de tampón de carga (x6) y se cargaron 20 µl sobre un gel de agarosa TAE al 0,8% (Invitrogen) más bromuro de etidio (5 µl de 10 mg/ml de solución madre) y se ejecutaron a 100 voltios durante 1 hora. Se cargaron 10 µl de marcador de tamaño (Perfect DNA marker 0,1-12 Kb, Novagen) en el carril final. Se visualizaron ambos geles empleando un transiluminador de UV.

15 El fragmento genómico amplificado presentaba la banda de tamaño correcto de 2,8 Kb para *Tsp*. Después de una digestión con *Ase I*, ésta confirmó la presencia de los sitios de *Ase I* introducidos en la cepa MXE001 carente de *Tsp* pero no en el control W3110.

MXE001: El ADN genómico se amplificó utilizando el conjunto de cebador *Tsp* y el ADN resultante se digirió con *Ase I* para producir bandas de 2,2 y 0,6 Kbp.

20 El ADN amplificado por PCR de W3110 no fue digerido por la enzima de restricción *Ase I*.

**Ejemplo 2 Generación de líneas celulares que son portadoras del gen spr mutado y líneas celulares que son portadoras del gen Tsp mutado y el gen spr mutado**

Se generaron y seleccionaron mutaciones de *spr* para utilizar un ensayo de complementación.

25 Se mutó el gen *spr* utilizando el kit de PCR de diversidad de mutagénesis aleatoria de Clontech<sup>®</sup> que introducía 1 a 2 mutaciones por 1.000 pb. El ADN de la PCR con *spr* mutada se clonó en un vector de expresión inducible [pTTO CDP870] que expresaba Fab' CDP870 junto con el mutante *spr*. Esta ligación se electrotransformó a continuación en la cepa MXE001 ( $\Delta Tsp$ ) de *E. coli* preparada utilizando el método encontrado en Miller, E.M. y Nickoloff, J.A., "*Escherichia coli* electrotransformation" en *Methods in Molecular Biology*, vol. 47, Nickoloff, J.A. (compilador), Humana Press, Totowa, N.J., 105 (1995). Se utilizó el protocolo siguiente, se añadieron 40 µl de MXE001 electrocompetente, 30 2,5 µl de la ligación (100 pg de ADN) a una cubeta de electroporación de 0,2 cm, se llevó a cabo la electrotransformación utilizando BioRad Genepulser Xcell con las siguientes condiciones, 2.500 V, 25 µF y 200 Ω. Después de la electrotransformación, se añadió 1 ml de SOC (Invitrogen) (precalentado a 37°C) y las células se dejaron recuperar a 37°C durante 1 hora con agitación suave.

Las células se sembraron en placas en agar-agar hipotónico (5 g/l de extracto de levadura, 2,5 g/l de triptona, 15 g/l de agar-agar (todo de Difco)) y se incubaron a 40°C. Las células que formaban colonias se volvieron a sembrar en placas sobre HLB a 43°C para confirmar el restablecimiento de la capacidad de crecimiento en condiciones osmóticas bajas a temperatura elevada para la cepa MXE001. Se preparó ADN de plásmido a partir de los clones seleccionados y se secuenció para identificar mutaciones de spr.

5 Utilizando este método se aislaron ocho mutaciones individuales, una doble y dos múltiples en la proteína spr que complementaban el fenotipo  $\Delta$ Tsp del modo siguiente:

1. V98E
2. D133A
3. V135D
4. V135G
5. G147C
6. S95F e Y115F
7. I70T

- 10 3. V135D
4. V135G
5. G147C
6. S95F e Y115F
7. I70T
- 15 8. N31T, Q73R, R100G, G140C
9. R62C, Q99P, R144C
10. L108S
11. L136P

20 Las mutaciones individuales 1 a 5 identificadas anteriormente y tres mutaciones de la tríada catalítica de spr (C94A, H145A, H157A) y W174R se usaron para generar nuevas cepas usando la cepa de *E. coli* de tipo silvestre W3110 (genotipo: F- LAM- IN (rrnD-rrnE)1 rph1 (n° de ATCC 27325)) para crear cepas mutadas en spr o la cepa MXE001 del Ejemplo 1 para preparar cepas  $\Delta$ Tsp/spr mutante combinadas.

25 Las siguientes cepas de células mutantes de *E. coli* se generaron utilizando un sistema de vector de reemplazo de genes utilizando el plásmido pKO3 de recombinación/reemplazo homólogo (Link *et al.*, 1997, Journal of Bacteriology, 179, 6228-6237), como se describe en el Ejemplo 1 para la generación de MXE001.

Tabla 1

Cepa celular mutante de <i>E. coli</i>	Genotipo	Vectores de spr
MXE001	$\Delta$ Tsp	-
MXE008	$\Delta$ Tsp, spr D133A	pMXE339, pK03 spr D133A (-Sall)
MXE009	$\Delta$ Tsp, spr H157A	pMXE345, pK03 spr H157A (-Sall)
MXE010	spr G147C	pMXE338, pK03 spr G147C (-Sall)
MXE011	spr C94A	pMXE343, pK03 spr C94A (-Sall)
MXE012	spr H145A	pMXE344, pK03 spr H145A (-Sall)
MXE013	spr W174R	pMXE346, pK03 spr W174R (-Sall)
MXE014	$\Delta$ Tsp, spr V135D	pMXE340, pK03 spr V135D (-Sall)
MXE015	$\Delta$ Tsp, spr V98E	pMXE342, pK03 spr V98E (-Sall)
MXE016	$\Delta$ Tsp, spr C94A	pMXE343, pK03 spr C94A (-Sall)

Cepa celular mutante de <i>E. coli</i>	Genotipo	Vectores de spr
MXE017	$\Delta$ Tsp, spr H145A	pMXE344, pK03 spr H145A (-Sall)
MXE018	$\Delta$ Tsp, spr V135G	pMXE341, pK03 spr V135G (-Sall)

Los casetes de integración de spr mutante se trasladan como fragmentos de restricción Sal I, Not I a los plásmidos pK03 restringidos de manera similar.

5 El plásmido utiliza el mutante sensible a la temperatura del origen de replicación de pSC101 (*RepA*) junto con un marcador de cloranfenicol para forzar y seleccionar eventos de integración cromosómica. El gen *sacB* que codifica levansacarasa es letal para el cultivo de *E. coli* sobre sacarosa y, por consiguiente, (junto con el marcador de cloranfenicol y el origen de pSC101) se utiliza para forzar y seleccionar eventos de desintegración y curado de plásmidos. Esta metodología se ha descrito anteriormente (Hamilton *et al.*, 1989, *Journal of Bacteriology*, 171, 4617-4622 y Blomfield *et al.*, 1991, *Molecular Microbiology*, 5, 1447-1457). El sistema pK03 elimina todos los marcadores selectivos del genoma del hospedador, excepto para el gen insertado.

Los vectores de pK03 que se indican a continuación fueron construidos comprendiendo los genes spr mutados que incluían una mutación silenciosa dentro de la secuencia de spr que elimina un sitio de restricción Sall para la identificación del clon.

pMXE336, S95F en spr de pK03 (-Sall)

15 pMXE337, Y115F en spr de pK03 (-Sall)

pMXE338, G147C en spr de pK03 (-Sall)

pMXE339, D133A en spr de pK03 (-Sall)

pMXE340, V135D en spr de pK03 (-Sall)

pMXE341, V135G en spr de pK03 (-Sall)

20 pMXE342, V98E en spr de pK03 (-Sall)

pMXE343, C94A en spr de pK03 (-Sall)

pMXE344, H145A en spr de pK03 (-Sall)

pMXE345, H157A en spr de pK03 (-Sall)

pMXE346, W174R en spr de pK03 (-Sall)

25 Estos plásmidos se transformaron a continuación en células W3110 de *E. coli* químicamente competentes, preparadas empleando el método encontrado en Miller, E.M. y Nickoloff, J.A., "Escherichia coli electrotransformation" en *Methods in Molecular Biology*, vol. 47, Nickoloff, J.A. (compilador), Humana Press, Totowa, NJ, 105 (1995) o en la cepa MXE001 del Ejemplo 1, para preparar cepas  $\Delta$ Tsp/spr mutante combinadas, como se muestra en la Tabla 1.

30 **Día 1** Se mezclaron 40  $\mu$ l de células *E. coli* electrocompetentes o células MXE001 con (10 pg) 1  $\mu$ l de ADN de pK03 en una cubeta de electroporación de 0,2 cm de BioRad, refrigerada antes de la electroporación a 2.500 V, 25  $\mu$ F y 200  $\Omega$ . Se añadieron inmediatamente 1000  $\mu$ l de 2xPY, las células se recuperaron por agitación a 250 rpm en una incubadora a 30°C durante 1 hora. Las células se diluyeron 1/10 en serie en 2xPY antes de que se extendieran en placas partes alícuotas de 100  $\mu$ l sobre placas con 2xPY agar-agar que contenían cloranfenicol a 20  $\mu$ g/ml, precalentadas a 30°C y 43°C. Las placas se incubaron durante la noche a 30°C y 43°C.

35 **Día 2** El número de colonias cultivadas a 30°C proporcionaba una estimación de la eficiencia de la electroporación mientras que las colonias que sobrevivían el cultivo a 43°C representaban eventos potenciales de integración. Se seleccionaron colonias aisladas de la placa a 43°C y se volvieron a poner en suspensión en 10 ml de 2xPY. Se sembraron 100  $\mu$ l de esta suspensión sobre placas de 2xPY agar-agar que contenían sacarosa al 5% (p/v), precalentadas a 30°C para generar colonias aisladas. Las placas se incubaron durante la noche a 30°C.

40 **Día 3** Las colonias aquí representan eventos potenciales simultáneos de desintegración y curado de plásmidos. Si los eventos de desintegración y curado ocurrían al principio del cultivo, entonces la mayor parte de la masa de la colonia será clonal. Se seleccionaron colonias aisladas y se sembraron en placas por duplicado sobre 2xPY agar-agar que contenía ya sea 20  $\mu$ g/ml de cloranfenicol o sacarosa al 5% (p/v). Las placas se incubaron durante la noche a 30°C.

Día 4 Las colonias que crecen sobre sacarosa y mueren sobre cloranfenicol representan potenciales eventos de sustitución cromosómica y curado de plásmidos. Éstas se seleccionaron y detectaron de forma sistemática por PCR más digestión por restricción en busca de la pérdida de un sitio Sall. Las colonias que generaban una banda positiva en la PCR del tamaño correcto y resistencia a la digestión con Sall, se separaban para producir colonias aisladas en 2xPY agar-agar que contenía 5% (p/v) de sacarosa y las placas se incubaron durante la noche a 30°C.

Día 5 Colonias individuales de *E. coli* positivas para PCR, sensibles al cloranfenicol y resistentes a la sacarosa se utilizaron para preparar reservas en glicerol de células químicamente competentes y que actuaban como plantillas de PCR para una reacción de PCR con oligos flanqueantes en 5' y 3' para generar un producto de PCR para la secuenciación directa de ADN usando **Taq** polimerasa para confirmar la mutación correcta.

### 10 **Ejemplo 3 - Generación de plásmido para expresión de Fab' y DsbC**

Se construyó un plásmido que contenía secuencias tanto de la cadena pesada como de la ligera de un Fab' anti-TNF y la secuencia codificadora de DsbC.

Se creó un mensaje bicistrónico del fragmento Fab' anti-TNF $\alpha$  (denominado CDP870) descrito en el documento WO01/94585. El cistrón aguas arriba codificaba la cadena ligera del anticuerpo (SEQ ID NO: 13) mientras que el cistrón aguas abajo codificaba la cadena pesada del anticuerpo (SEQ ID NO: 14). Una secuencia de ADN que codificaba el péptido señal OmpA, se fusionó al extremo 5' del ADN que codificaba cada cadena ligera y cadena pesada para permitir una secreción eficaz en el periplasma. La secuencia intergénica (IGS) 2 se utilizó como se muestra en SEQ ID NO: 37.

El plásmido pDPH358 (pTTO 40.4 CDP870 IGS2), un vector de expresión para el Fab' CDP870 (un Fab' anti-TNF) y DsbC (un polipéptido periplásmico), se construyó empleando metodologías convencionales de clonación por restricción que pueden encontrarse en Sambrook *et al.* 1989, Molecular cloning: a laboratory manual. CSHL press, N.Y. El plásmido pDPH358 contenía los elementos siguientes; una secuencia de promotor *tac* potente y de operador *lac*. Como se muestra en la Figura 6, el plásmido contenía un único sitio de restricción EcoRI después de la región codificante de la cadena pesada de los Fab', seguido por una secuencia no codificante y luego un único sitio de restricción NdeI. El gen DsbC se clonó mediante PCR utilizando ADN cromosómico de W3110 en bruto como plantilla, de tal manera que el producto de la PCR codificaba un sitio EcoRI en 5' seguido de una unión fuerte al ribosoma, seguida del codón de inicio natural, secuencia señal y secuencia madura de DsbC, terminando con un marcador His en el extremo C-terminal y por último una secuencia NdeI no codificadora. El fragmento EcoRI-NdeI de PCR se restringió y se ligó en el vector de expresión, de tal manera que los tres polipéptidos: cadena ligera de Fab', cadena pesada de Fab' y DsbC se codificaban en un solo ARNm policistrónico.

Los genes de la cadena ligera y la cadena pesada del Fab y el gen DsbC se transcribieron como un solo mensaje policistrónico. El ADN que codificaba el péptido señal de la proteína OmpA de *E. coli* se fusionó con el extremo 5' de las secuencias génicas tanto de la cadena ligera como de la pesada, que dirigían la translocación de los polipéptidos al periplasma de *E. coli*. La transcripción se terminó utilizando un terminador doble de la transcripción *rnnB t1t2*. El gen *lacIq* codificaba la proteína represora Lac I expresada de forma constitutiva. Esta transcripción reprimida procedente del promotor *tac* hasta la desrepresión, fue provocada por la presencia de la alolactosa o IPTG. El origen de replicación utilizado era p15A, el cual conservaba un número bajo de copias. El plásmido contenía un gen de resistencia a la tetraciclina para la selección del antibiótico.

### **Ejemplo 4 - Expresión de Fab' anti-TNF y DsbC en las líneas celulares mutantes**

#### 40 **Expresión de Fab' anti-TNF y DsbC**

La cepa MXE001 proporcionada en el Ejemplo 1 y las cepas mutantes MXE008 y MXE009 proporcionadas en el Ejemplo 2, se transformaron con el plásmido generado en el Ejemplo 3.

La transformación de las cepas se llevó a cabo utilizando el método encontrado en Chung C.T. *et al.* Transformation and storage of bacterial cells in the same solution. PNAS 86:2172-2175 (1989).

#### 45 **Expresión de Fab' anti-TNF**

Las cepas MXE001, MXE008 y MXE009 se transformaron con el plásmido pMXE117 (pTTO CDP870 o 40.4 IGS2), un vector de expresión para el Fab' CDP870 (un Fab' anti-TNF que tiene una secuencia de la cadena ligera mostrada en SEQ ID NO: 13 y una secuencia de la cadena pesada mostrada en SEQ ID NO: 14), como se ha descrito en el Ejemplo 3, el cual fue construido usando metodologías de clonación por restricción convencionales que se pueden encontrar en Sambrook *et al.* 1989, Molecular cloning: a laboratory manual. CSHL press, N.Y. El plásmido pMXE117 (pTTO CDP870 o 40.4 IGS2) contenía los elementos siguientes; una secuencia de promotor *tac* potente y de operador *lac*. Los genes de la cadena ligera y pesada de Fab se transcribieron como un solo mensaje bicistrónico. El ADN que codificaba el péptido señal de la proteína OmpA de *E. coli* se fusionó con el extremo 5' de las secuencias génicas tanto de la cadena ligera como de la pesada, que dirigían la translocación de los polipéptidos al periplasma de *E. coli*. La transcripción se terminó utilizando un terminador doble de transcripción *rnnB t1t2*. El gen *lacIq* codificaba la proteína represora Lac I expresada de forma constitutiva. Esta transcripción reprimida procedente del promotor *tac*

hasta la desrepresión estaba provocada por la presencia de alolactosa o IPTG. El origen de replicación utilizado era p15A, el cual conservaba un número bajo de copias. El plásmido contenía un gen de resistencia a la tetraciclina para la selección con antibiótico.

5 La transformación de las cepas se llevó a cabo utilizando el método encontrado en Chung C.T. *et al.* Transformation and storage of bacterial cells in the same solution. PNAS 86:2172-2175 (1989).

#### **Ejemplo 5 - Crecimiento de cepas de *E. coli* mutadas y expresión de Fab' anti-TNF en cepas de *E. coli* mutadas usando fermentaciones de alta densidad**

Las siguientes cepas producidas en el Ejemplo 4 se sometieron a ensayo en experimentos de fermentación que comparaban el crecimiento y la expresión de un Fab' anti-TNF $\alpha$ :

- 10           MXE001 que expresa Fab' anti-TNF  
               MXE008 que expresa Fab' anti-TNF  
               MXE009 que expresa Fab' anti-TNF  
               MXE001 que expresa Fab' anti-TNF y DsbC  
               MXE008 que expresa Fab' anti-TNF y DsbC  
 15           MXE009 que expresa Fab' anti-TNF y DsbC

#### **Medio de crecimiento**

El medio de crecimiento de la fermentación se basaba en el medio SM6E (descrito en Humphreys et al., 2002, Protein Expression and Purification, 26, 309-320) con 3,86 g/l de NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O y 112 g/l de glicerol.

20 **Inóculo** Los cultivos de inóculo se cultivaron en el mismo medio complementado con 10  $\mu$ g/ml de tetraciclina. Los cultivos se incubaron a 30°C con agitación durante aproximadamente 22 horas.

**Fermentación.** Los fermentadores (2,5 litros de volumen total) se sembraron con un cultivo del inóculo a una DO<sub>600</sub> de 0,3-0,5. La temperatura se mantuvo a 30°C durante la fase de crecimiento y se redujo a 25°C antes de la inducción. La concentración de oxígeno disuelto se mantuvo por encima de la saturación de aire del 30% mediante agitación variable y flujo de aire. El pH del cultivo se controló en 7,0 mediante una valoración automática con 15% (v/v) de NH<sub>4</sub>OH y 10% (v/v) de H<sub>2</sub>SO<sub>4</sub> conc. La formación de espuma se controló mediante la adición de 10% (v/v) de solución Struktol J673 (Schill y Seilacher). Se realizó una serie de adiciones en diferentes etapas de la fermentación. Cuando la concentración de la biomasa alcanzó aproximadamente una DO<sub>600</sub> de 40, se añadieron sales de magnesio y NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O. Se realizaron más adiciones de NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O antes y durante la fase de inducción para asegurar que el fosfato se mantenía en exceso. Cuando el glicerol presente al principio de la fermentación se había agotado (aproximadamente una DO<sub>600</sub> de 75) se aplicó una alimentación continua de 80% (p/p) de glicerol. En el mismo punto en la fermentación, se aplicó una alimentación de IPTG a 170  $\mu$ M. El inicio de la alimentación de IPTG se tomó como inicio de la inducción. Las fermentaciones tenían normalmente una duración de 64-120 horas con tasas de alimentación de glicerol (que oscilaban entre 0,5 y 2,5 ml/h).

35 **Medición de la concentración de biomasa y tasa de crecimiento** La concentración de biomasa fue determinada midiendo la densidad óptica de cultivos a 600 nm.

40 **Extracción periplásmica** Las células se recogieron a partir de muestras de cultivo mediante centrifugación. La fracción de material sobrenadante fue conservada (a -20°C) para un análisis adicional. La fracción de sedimento celular se resuspendió hasta el volumen de cultivo original en tampón de extracción (Tris-HCl 100 mM, EDTA 10 mM; pH 7,4). Después de la incubación a 60°C durante aproximadamente 16 horas, el extracto se clarificó por centrifugación y la fracción de material sobrenadante se conservó (a -20°C) para el análisis.

45 **Cuantificación de Fab'** Las concentraciones de Fab' en extractos periplásmicos y material sobrenadante del cultivo se determinaron mediante ELISA acoplado a Fab' como se describe en Humphreys et al., 2002, Protein Expression and Purification, 26, 309-320 y empleando HPLC de Proteína G. Una columna de 1 ml de proteína-G HiTrap HP (GE-Healthcare o equivalente) se cargó con el analito (pH aproximadamente neutro, 30°C, filtro de 0,2  $\mu$ m) a 2 ml/min, la columna se lavó con fosfato 20 mM, NaCl 50 mM pH 7,4 y a continuación, el Fab' se eluyó usando una inyección de glicina/HCl 50 mM pH 2,7. El Fab' eluido se midió mediante A280 en un sistema de HPLC Agilent 1100 o 1200 y se cuantificó por referencia a una curva estándar de una proteína Fab' purificada de concentración conocida.

50 La Figura 1 muestra el perfil de crecimiento de cepas MXE001 y MXE008 que expresan Fab' anti-TNF y el perfil de crecimiento de cepas MXE001 y MXE008 que expresan Fab' anti-TNF $\alpha$  y DSBC recombinante. Se puede observar que las cepas que expresan DsbC muestran un crecimiento mejorado en comparación con las cepas de células

correspondientes que no expresan DsbC recombinante. También se puede observar que la presencia de la mutación spr en las cepas, mejora el crecimiento celular.

5 La Figura 2 muestra el rendimiento de Fab a partir del periplasma (símbolos sombreadas) y el material sobrenadante (símbolos blancos sin sombrear) de cepas MXE001, MXE008 y MXE009 de *E. coli* que expresan Fab' anti-TNF y DsbC recombinante. Se puede observar a partir de este gráfico que las tres cepas producían un alto rendimiento en Fab' anti-TNF $\alpha$ , en donde las cepas MXE008 y MXE009 producían más de 2,2 g/L de Fab' anti-TNF $\alpha$  periplásmico en 40 horas. También se puede observar que las cepas MXE008 y MXE009 que son portadoras de un gen spr mutante mostraban una lisis reducida en comparación con MXE001 en la que se puede observar menos Fab' anti-TNF $\alpha$  en el material sobrenadante (símbolos blancos).

10 La Figura 3 muestra el rendimiento de Fab a partir del periplasma (símbolos sombreados) y el material sobrenadante (símbolos blancos sin sombrear) a partir de cepas MXE001 y MXE008 de *E. coli* que expresaban Fab' anti-TNF $\alpha$  y a partir de cepas MXE001 y MXE008 de *E. coli* que expresaban Fab' anti-TNF $\alpha$  y DsbC recombinante. Se puede observar a partir de este gráfico que las cepas que expresaban DsbC recombinante producían un alto rendimiento de Fab' anti-TNF $\alpha$ , en donde la cepa MXE008 producía 2,4 g/L de Fab' anti-TNF $\alpha$  periplásmico en 40 horas. También se puede observar que las cepas MXE008 que son portadoras de un gen spr mutante, mostraban una lisis reducida en comparación con las cepas MXE001 en las que se puede observar menos Fab' anti-TNF $\alpha$  en el material sobrenadante (símbolos blancos).

#### **Ejemplo 6 - Crecimiento de cepas de *E. coli* mutadas y expresión de Fab A y Fab B en cepas de *E. coli* mutadas usando fermentaciones de alta densidad**

20 El efecto sobre el rendimiento de la producción de proteínas a partir de una célula de la presente invención también se llevó a cabo para dos proteínas Fab adicionales, Fab A que tiene especificidad para el antígeno A y Fab B que tiene especificidad para el antígeno B, utilizando el mismo método descrito anteriormente en el Ejemplo 5 y en comparación con las cepas W3110 que expresaban Fab A y Fab B.

25 La Figura 4 muestra el rendimiento de Fab total a partir del periplasma procedente de la cepa W3110 de *E. coli* que expresa Fab A y Fab B y de la cepa de *E. coli* MXE008 que expresa Fab A y DsbC recombinante o Fab B y DsbC recombinante. Se puede observar a partir de este gráfico que las cepas que expresan DsbC recombinante producían un rendimiento significativamente mayor de Fab A y Fab B en MXE008 que en comparación con W3110.

#### **Ejemplo 7 - Crecimiento de cepas de *E. coli* que expresan Fab' anti-TNF y DsbC usando fermentaciones a gran escala**

30 La siguiente cepa, como la producida en el ejemplo 4, se sometió a ensayo en experimentos de fermentación que comparaban el crecimiento de la cepa y la expresión de un Fab' anti-TNF $\alpha$ :

MXE008 que expresa Fab' anti-TNF y DsbC

Las fermentaciones se llevaron a cabo del modo siguiente:

35 Las células MXE008 que expresaban Fab' anti-TNF $\alpha$  y DsbC se cultivaron inicialmente utilizando un medio complejo de extracto de levadura y peptona en un cultivo en matraz de agitación. Las células se transfirieron después a un fermentador en la etapa inicial, usando un medio químicamente definido. Las células se cultivaron en condiciones de no limitación de nutrientes hasta un punto de transferencia definido. Las células se transfirieron entonces a un fermentador de producción de 5 L o 250 L, usando un medio definido químicamente similar, con un volumen final de aproximadamente 3,3 L o 230 L, respectivamente. El cultivo se desarrolló inicialmente en modo por lotes hasta el agotamiento de la fuente de carbono. Después del agotamiento de la fuente de carbono, se nutrió con una solución que limitaba la fuente de carbono con una tasa de aumento exponencial. Después de la adición de una cantidad definida de fuente de carbono, la tasa de adición de solución nutritiva se redujo y se añadió IPTG para inducir la expresión del Fab'. A continuación, la fermentación continuó y el Fab' se acumuló en el periplasma. En un periodo definido después de la inducción, el cultivo se recogió mediante centrifugación y el Fab' se extrajo de las células mediante resuspensión de las células recolectadas en un tampón Tris y EDTA y se calentó a 59°C.

Los perfiles de crecimiento de las fermentaciones se determinaron midiendo la densidad óptica del cultivo a 600 nm.

50 Los títulos de Fab' se determinaron mediante HPLC de Proteína G como se ha descrito en el Ejemplo 5 anterior, excepto que durante la extracción periplásmica se utilizaron células de nuevo aporte y se añadió 1 ml de tampón de extracción para el cultivo celular. El Fab' del material sobrenadante y periplásmico se midió como se ha descrito en el Ejemplo 5. La Figura 8 muestra el título de Fab' periplásmico.

La Figura 7 muestra los perfiles de crecimiento comparativos de fermentaciones de 5 L y 200 L de la cepa MXE008 de *E. coli* que expresa Fab' anti-TNF y DsbC recombinante.

La Figura 8 muestra los títulos de Fab' comparativos de fermentaciones de 5 L y 200 L de la cepa MXE008 de *E. coli* que expresa Fab' anti-TNF y DsbC recombinante.

**Ejemplo 8 - Crecimiento de cepas de *E. coli* que expresan Fab' anti-TNF y DsbC usando fermentaciones a gran escala**

Las siguientes cepas, como la producida por el Ejemplo 4, se sometieron a ensayo en experimentos de fermentación que comparaban el crecimiento de la cepa y la expresión de un Fab' anti-TNF $\alpha$ :

5    MXE008 que expresa Fab' anti-TNF y DsbC

      MXE009 que expresa Fab' anti-TNF y DsbC

Las fermentaciones se llevaron a cabo como se ha descrito anteriormente en el Ejemplo 7 usando un fermentador de producción de 20 L.

Los perfiles de crecimiento de las fermentaciones se determinaron midiendo la densidad óptica del cultivo a 600 nm.

10   Los títulos de Fab' se determinaron mediante HPLC de Proteína G como se ha descrito en el Ejemplo 7 anterior.

La Figura 9 muestra los perfiles de crecimiento comparativos de las cepas MXE008 y MXE009 de *E. coli* que expresan Fab' anti-TNF y DsbC recombinante.

La Figura 10 muestra los títulos de Fab' periplásmicos comparativos de las cepas MXE008 y MXE009 de *E. coli* que expresan Fab' anti-TNF y DsbC recombinante.

15   **LISTADO DE SECUENCIAS**

      <110> UCB Pharma S.A.

      <120> CEPA BACTERIANA HOSPEDADORA

20

      <130> G0109-WO

      <150> GB1000591.6

      <151> 14-01-2010

25

      <160> 39

      <170> PatentIn versión 3.5

30

      <210> 1

      <211> 2049

      <212> ADN

      <213> *E. coli*

35

      <400> 1

ES 2 627 826 T3

atgaacatgt tttttaggct taccgcgtta gctggcctgc ttgcaatagc aggccagacc 60  
 ttcgctgtag aagatatcac gcgtgctgat caaattccgg tattaagga agagacgcag 120  
 catgcgacgg taagtggagc cgtaacgtcg cgcttcaccc gttctcatta tcgccagttc 180  
 gacctcgatc aggcattttc ggccaaaatc tttgacogct acctgaatct gctcgattac 240  
 agccacaacg tgctgctggc aagcgatggt gaacagttcg cgaaaaagaa aaccgagtta 300  
 ggcgatgaac tgcgttcagg caaactcgac gttttctacg atctctacaa tctggcgcaa 360  
 aagcgccggt ttgagcggtta ccagtacgct ttgtcggtac tggaaaagcc gatggatttc 420  
 accggcaacg acacttataa ccttgaccgc agcaaaagcg cctggccgaa aaacgaggtc 480  
 gagttgaacg cgctgtggga cagtaaagtc aaattcgacg agttaagcct gaagctgaca 540  
 ggaaaaacgg ataagaaat tcgtgaaacc ctgactogcc gctacaaatt tgccattcgt 600  
 cgtctggcgc aaaccaacag cgaagatggt ttctcgctgg caatgacggc gtttgccgct 660  
 gaaatcgacc cgcataccaa ctatctttcc ccgcgtaata ccgaacagtt caaactgaa 720  
 atgagtttgt cgctggaagg tattggcgca gtgctgcaaa tggatgatga ctacaccggt 780  
 atcaattcga tggggcagc tgggtccggca gcgaagagta aagctatcag cgttgggtgac 840  
 aaaattgtcg gtgttggtca aacaggcaag ccgatggttg acgtgattgg ctggcgtctt 900  
 gatgatgtgg ttgccttaat taaagggccg aagggcagta aagttcgtct ggaaatttta 960  
 cctgctggta aagggaccaa gaccctact gtaacggtga cccgtgaacg tattcgtctc 1020  
 gaagaccgcg cggttaaaat gtcgggtaag accgtcggta aagagaaagt cggcgtgctg 1080  
 gatattccgg gcttctatgt gggtttgaca gacgatgtca aagtgcaact gcagaaactg 1140  
 gaaaaacaga atgtcagcag cgtcatcatc gacctcgta gcaatggcgg tggggcggtta 1200  
 actgaagccg tategctctc cggctctggtt attcctgctg gtccattgt tcaggtccgc 1260  
 gataacaacg gcaaggttcg tgaagatagc gataccgacg gacaggtttt ctataaaggc 1320  
 ccgctggtgg tgctggttga ccgcttcagt gcttcggctt cagaaatctt tgcccgcgca 1380  
 atgcaggatt acggtcgtgc gctggttggt ggtgaaccga cgtttggtaa aggcaccggt 1440  
 cagcaatacc gttcattgaa ccgtatttac gatcagatgt tacgtcctga atggccagcg 1500  
 ctgggttctg tgcagtacac gatccagaaa ttctatcgcg ttaacggcgg cagtacgcaa 1560  
 cgtaaaggcg taacgccaga catcatcatg ccgacgggta atgaagaaac ggaaacgggt 1620  
 gagaaattcg aagataacgc gctgccgtgg gatagcattg atgccgcgac ttatgtgaaa 1680  
 tcaggagatt taacggcctt tgaaccggag ctgctgaagg aacataatgc gcgtatcgcg 1740  
 aaagatcctg agttocagaa catcatgaag gatatcgcg gcttcaacgc tatgaaggac 1800  
 aagcgcaata tcgtttctct gaattacgct gtgcgtgaga aagagaataa tgaagatgat 1860  
 gcgacgcgctc tggcgcgttt gaacgaacgc tttaaacgcg aaggtaaacc ggagttgaag 1920  
 aaactggatg atctaccgaa agattaccag gagccggatc cttatctgga tgagacggtg 1980  
 aatatcgcac tcgatctggc gaagctttaa aaagccagac ccgcggaaca acccgctccc 2040  
 gtcaagtaa 2049

ES 2 627 826 T3

<211> 682  
 <212> PRT  
 <213> *E. coli*

5 <400> 2

Met Asn Met Phe Phe Arg Leu Thr Ala Leu Ala Gly Leu Leu Ala Ile  
 1 5 10 15

Ala Gly Gln Thr Phe Ala Val Glu Asp Ile Thr Arg Ala Asp Gln Ile  
 20 25 30

Pro Val Leu Lys Glu Glu Thr Gln His Ala Thr Val Ser Glu Arg Val  
 35 40 45

Thr Ser Arg Phe Thr Arg Ser His Tyr Arg Gln Phe Asp Leu Asp Gln  
 50 55 60

Ala Phe Ser Ala Lys Ile Phe Asp Arg Tyr Leu Asn Leu Leu Asp Tyr  
 65 70 75 80

Ser His Asn Val Leu Leu Ala Ser Asp Val Glu Gln Phe Ala Lys Lys  
 85 90 95

Lys Thr Glu Leu Gly Asp Glu Leu Arg Ser Gly Lys Leu Asp Val Phe  
 100 105 110

Tyr Asp Leu Tyr Asn Leu Ala Gln Lys Arg Arg Phe Glu Arg Tyr Gln  
 115 120 125

ES 2 627 826 T3

Tyr Ala Leu Ser Val Leu Glu Lys Pro Met Asp Phe Thr Gly Asn Asp  
 130 135 140

Thr Tyr Asn Leu Asp Arg Ser Lys Ala Pro Trp Pro Lys Asn Glu Ala  
 145 150 155 160

Glu Leu Asn Ala Leu Trp Asp Ser Lys Val Lys Phe Asp Glu Leu Ser  
 165 170 175

Leu Lys Leu Thr Gly Lys Thr Asp Lys Glu Ile Arg Glu Thr Leu Thr  
 180 185 190

Arg Arg Tyr Lys Phe Ala Ile Arg Arg Leu Ala Gln Thr Asn Ser Glu  
 195 200 205

Asp Val Phe Ser Leu Ala Met Thr Ala Phe Ala Arg Glu Ile Asp Pro  
 210 215 220

His Thr Asn Tyr Leu Ser Pro Arg Asn Thr Glu Gln Phe Asn Thr Glu  
 225 230 235 240

Met Ser Leu Ser Leu Glu Gly Ile Gly Ala Val Leu Gln Met Asp Asp  
 245 250 255

Asp Tyr Thr Val Ile Asn Ser Met Val Ala Gly Gly Pro Ala Ala Lys  
 260 265 270

Ser Lys Ala Ile Ser Val Gly Asp Lys Ile Val Gly Val Gly Gln Thr  
 275 280 285

Gly Lys Pro Met Val Asp Val Ile Gly Trp Arg Leu Asp Asp Val Val  
 290 295 300

Ala Leu Ile Lys Gly Pro Lys Gly Ser Lys Val Arg Leu Glu Ile Leu  
 305 310 315 320

Pro Ala Gly Lys Gly Thr Lys Thr Arg Thr Val Thr Leu Thr Arg Glu  
 325 330 335

Arg Ile Arg Leu Glu Asp Arg Ala Val Lys Met Ser Val Lys Thr Val  
 340 345 350

Gly Lys Glu Lys Val Gly Val Leu Asp Ile Pro Gly Phe Tyr Val Gly  
 355 360 365

Leu Thr Asp Asp Val Lys Val Gln Leu Gln Lys Leu Glu Lys Gln Asn  
 370 375 380

ES 2 627 826 T3

Val Ser Ser Val Ile Ile Asp Leu Arg Ser Asn Gly Gly Gly Ala Leu  
385 390 395 400

Thr Glu Ala Val Ser Leu Ser Gly Leu Phe Ile Pro Ala Gly Pro Ile  
405 410 415

Val Gln Val Arg Asp Asn Asn Gly Lys Val Arg Glu Asp Ser Asp Thr  
420 425 430

Asp Gly Gln Val Phe Tyr Lys Gly Pro Leu Val Val Leu Val Asp Arg  
435 440 445

Phe Ser Ala Ser Ala Ser Glu Ile Phe Ala Ala Ala Met Gln Asp Tyr  
450 455 460

Gly Arg Ala Leu Val Val Gly Glu Pro Thr Phe Gly Lys Gly Thr Val  
465 470 475 480

Gln Gln Tyr Arg Ser Leu Asn Arg Ile Tyr Asp Gln Met Leu Arg Pro  
485 490 495

Glu Trp Pro Ala Leu Gly Ser Val Gln Tyr Thr Ile Gln Lys Phe Tyr  
500 505 510

Arg Val Asn Gly Gly Ser Thr Gln Arg Lys Gly Val Thr Pro Asp Ile  
515 520 525

Ile Met Pro Thr Gly Asn Glu Glu Thr Glu Thr Gly Glu Lys Phe Glu  
530 535 540

Asp Asn Ala Leu Pro Trp Asp Ser Ile Asp Ala Ala Thr Tyr Val Lys  
545 550 555 560

Ser Gly Asp Leu Thr Ala Phe Glu Pro Glu Leu Leu Lys Glu His Asn  
565 570 575

Ala Arg Ile Ala Lys Asp Pro Glu Phe Gln Asn Ile Met Lys Asp Ile  
580 585 590

Ala Arg Phe Asn Ala Met Lys Asp Lys Arg Asn Ile Val Ser Leu Asn  
595 600 605

Tyr Ala Val Arg Glu Lys Glu Asn Asn Glu Asp Asp Ala Thr Arg Leu  
610 615 620

Ala Arg Leu Asn Glu Arg Phe Lys Arg Glu Gly Lys Pro Glu Leu Lys  
625 630 635 640

Lys Leu Asp Asp Leu Pro Lys Asp Tyr Gln Glu Pro Asp Pro Tyr Leu  
645 650 655

Asp Glu Thr Val Asn Ile Ala Leu Asp Leu Ala Lys Leu Glu Lys Ala  
660 665 670

Arg Pro Ala Glu Gln Pro Ala Pro Val Lys  
675 680

<210> 3

5 <211> 2048

ES 2 627 826 T3

<212> ADN  
 <213> *E. coli*

<400> 3

5

```

atgaattcgt ttttaggctt accgcgcttag ctggcctgct tgcaatagca ggccagacat      60
taattgtaga agatatcacg cgtgctgatc aaattccggg attaaaggaa gagacgcagc      120
atgcgacggg aagtgagcgc gtaacgctgc gcttcacccg ttctcattat cgccagttcg      180
acctcgatca ggcattttcg gccaaaatct ttgaccgcta cctgaatctg ctcgattaca      240
gccacaacgt gctgctggca agcgatgttg aacagttcgc gaaaaagaaa accgagttag      300
gcgatgaact gcgttcaggc aaactcgacg ttttctacga tctctacaat ctggcgcaaa      360
agcgcctggt tgagcgctac cagtacgctt tgtcgtgact ggaaaagccg atggatttca      420
ccggcaacga cacttataac cttgaccgca gcaaagcgcc ctggccgaaa aacgaggctg      480
agttgaacgc gctgtgggac agtaaagtca aattcgacga gttaagcctg aagctgacag      540
gaaaaacgga taaagaaatt cgtgaaaccc tgactcgccg ctacaaattt gccattcgtc      600
gtctggcgca aaccaacagc gaagatgttt tctcgtggc aatgacggcg tttgcgcggtg      660
aaatcgaccc gcataccaac tatctttccc cgcgtaatac cgaacagttc aacctgaaa      720
tgagtttgtc gctggaaggt attggcgcag tgctgcaaat ggatgatgac tacaccgtta      780
tcaattcgat ggtggcaggt ggtccggcag cgaagagtaa agctatcagc gttggtgaca      840
aaattgtcgg tgttgggtcaa acaggcaagc cgatggttga cgtgattggc tggcgtcttg      900
atgatgtggt tgcttaatt aaagggccga agggcagtaa agttcgtctg gaaattttac      960
ctgctggtaa agggaccaag acccgactg taacgttgac ccgtgaacgt attcgtctcg     1020
aagaccgcgc ggttaaaatg tcggtgaaga ccgtcggtaa agagaaagtc ggctgctgg      1080
atattccggg cttctatgtg ggtttgacag acgatgtcaa agtgcaactg cagaaactgg     1140
aaaaacagaa tgcagcagc gtcacatcg acctgcgtag caatggcggg ggggcgttaa     1200
ctgaagccgt atcgtctcc ggtctgttta ttctcgggg tcccattgtt caggtccgcg     1260
ataacaacgg caaggttcgt gaagatagcg ataccgacgg acaggttttc tataaaggcc     1320
cgctggtggt gctggttgac cgcttcagtg cttcggcttc agaaatcttt gccgcggcaa     1380
tgcaggatta cggctgtgcg ctgggtgtgg gtgaaccgac gtttggtaaa ggcaccgttc     1440
  
```

ES 2 627 826 T3

agcaataaccg ttcattgaac cgtatttacg atcagatggt acgtcctgaa tggccagcgc 1500  
 tgggttctgt gcagtacacg atccagaaat tctatcgcgt taacggcggc agtacgcaac 1560  
 gtaaaggcgt aacgccagac atcatcatgc cgacgggtaa tgaagaaacg gaaacgggtg 1620  
 agaaattcga agataacgcg ctgccgtggg atagcattga tgccgcgact tatgtgaaat 1680  
 caggagattt aacggccttt gaaccggagc tgctgaagga acataatgcg cgtatcgcga 1740  
 aagatcctga gttccagaac atcatgaagg atatcgcgcg cttcaacgct atgaaggaca 1800  
 agcgcaatat cgtttctctg aattacgctg tgctgagaa agagaataat gaagatgatg 1860  
 cgacgcgtct gccgcgtttg aacgaacgct ttaaaccgga aggtaaaccg gagttgaaga 1920  
 aactggatga tctaccgaaa gattaccagg agccgatcc ttatctggat gagacggtga 1980  
 atatcgcact cgatctggcg aagcttgaaa aagccagacc cgcggaacaa cccgctcccg 2040  
 tcaagtaa 2048

<210> 4  
 <211> 2889  
 <212> ADN  
 <213> *E. coli*

5

<400> 4

atgccccgca gcacctggtt caaagcatta ttgttgttag ttgccctttg ggcaccotta 60  
 agtcaggcag aaacgggatg gcagccgatt caggaaacca tccgtaaaag tgataaagat 120  
 aaccgccagt atcaggctat acgtctggat aacggtatgg tggctctgct ggtttctgat 180  
 ccgcaggcag ttaaatacgt ctcggcgctg gtggtgcccg ttgggtcgct ggaagatccc 240  
 gaggcgtacc aggggctggc acattacctt gaacatatga gtctgatggg gtcgaaaaag 300  
 taccgcagag ctgacagtct ggccgaatat ctcaaatgc acggcggtag tcacaatgcc 360  
 agcactgcgc cgtatcgcac ggctttctat ctggaagttg agaacgacgc cttgcctggt 420  
 gcggtagacc gcctggccga tgctattgct gaacctttgc tcgacaagaa atatgccgaa 480  
 cgtgagcgta atgcggtgaa cgctgaatta accatggcgc gtacgcgtga cgggatgcgc 540  
 atggcacagg tcagcgcaga aaccattaac ccggcacacc ccggttcaaa gttttctggt 600  
 ggtaacctcg aaactttaag cgacaaacct ggtaatccgg tgcagcaggc gctgaaagat 660  
 ttccacgaga agtactattc cgccaatttg atgaaggcgg ttatttacag taataaaccg 720  
 ctgccggagt tggcaaaaat gggggcggac acctttggtc gcgtgccgaa caaagagagc 780  
 aaaaaaccgg aatcaccgt gccggtagtc accgacgcgc aaaaggcat tatcattcat 840  
 tacgtccctg cgctgccgcg taaagtgtg cgcgttgagt ttogcatcga taacaactca 900  
 gcgaagttcc gtagtaaaac cgatgaattg attacctatc tgattggcaa tcgcagccca 960  
 ggtacacttt ctgactggct gcaaaagcag ggattagttg agggcattag cgccaactcc 1020  
 gatcctatcg tcaacggcaa cagcggcgta ttagcgatct ctgcgtcttt aaccgataaa 1080

10

ES 2 627 826 T3

ggccctggcta atcgcgatca ggttgtggcg gcaattttta gctatctcaa tctgttacgt 1140  
 gaaaaaggca ttgataaaca atacttcgat gaactggcga atgtgctgga tatcgacttc 1200  
 cgttatccgt cgatcaccog tgatatggat tacgtcgaat ggctggcaga taccatgatt 1260  
 cgcgttcoctg ttgagcatac gctggatgca gtcaatattg ccgatcggta cgatgctaaa 1320  
 gcagtaaagg aacgtctggc gatgatgacg ccgcagaatg cgcgtatctg gtatatcagc 1380  
 ccgaaagagc cgcacaacaa aacggcttac tttgtcgatg cgcctatca ggctgataaa 1440  
 atcagcgcac aaactttcgc cgactggcag aaaaaagccg ccgacattgc gctctctttg 1500  
 ccagagotta acccttatat tctgatgat ttctcgtga ttaagtcaga gaagaaatac 1560  
 gaccatccag agctgattgt tgatgagtcg aatctgcgog tgggtgatgc gccaaagccgt 1620  
 tattttgcc a gcgagcccaa agctgatgtc agcctgattt tgcgtaatcc gaaagccatg 1680  
 gacagcgcgc gcaatcaggt gatgtttgcg ctcaatgatt atctcgcagc gctggcgcct 1740  
 gatcagttaa gcaaccaggc gtcggttggt ggcataagtt ttccaccaaa cgctaacaac 1800  
 ggccttatgg ttaatgctaa tggttacacc cagcgtctgc cgcagctgtt ccaggcattg 1860  
 ctcgaggggt actttagcta taccgctacg gaagatcagc ttgagcaggc gaagtccctgg 1920  
 tataaccaga tgatggattc cgcagaaaag ggtaaagcgt ttgagcaggc gattatgcc 1980  
 gcgcagatgc tctcgcaagt gccgtacttc tcgcgagatg aacggcgtaa aattttgcc 2040  
 tccattacgt tgaaagaggt gctggcctat cgcgacgcct taaaatcagc ggctcgacca 2100  
 gagtttatgg ttatcggcaa catgaccgag gccagggcaa caacgctggc acgcgatgtg 2160  
 caaaaacagt tgggcgctga tggttcagag tgggtgcgaa acaaagatgt agtggctgat 2220  
 aaaaaacaat ccgtcatctt tgaaaaagcc ggtaacagca ccgactccgc actggcagcg 2280  
 gtattttgac cgactggcga cgatgaatac accagctcag cctatagctc tctgttgggg 2340  
 cagatcgtac agccgtggtt ctacaatcag ttgcgtaccg aagaacaatt gggctatgcc 2400  
 gtgtttgcgt ttccaatgag cgtggggcgt cagtggggca tgggcttcct tttgcaaagc 2460  
 aatgataaac agccttcatt cttgtgggag cgttacaagg cgtttttccc aaccgcagag 2520  
 gcaaaattgc gagcgatgaa gccagatgag tttgcgcaaa tccagcaggc ggtaattacc 2580  
 cagatcgtgc aggcaccgca aacgctcggc gaagaagcat cgaagttaag taaagatttc 2640  
 gatcgcggca atatgcgctt cgattcgcgt gataaaatcg tggcccagat aaaactgctg 2700  
 acgccgcaaa aacttgctga tttcttccat cagggcgtgg tcgagccgca aggcattggct 2760  
 attctgtcgc agatttccgg cagccagaac gggaaaagccg aatatgtaca ccctgaaggc 2820  
 tggaaagtgt gggagaacgt cagcgcggtt cagcaaaacaa tgcccctgat gagtgaaaag 2880  
 aatgagtga 2889

<210> 5  
 <211> 962  
 <212> PRT  
 <213> *E. coli*

5

<400> 5

ES 2 627 826 T3

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

Trp Ala Pro Leu Ser Gln Ala Glu Thr Gly Trp Gln Pro Ile Gln Glu  
20 25 30

Thr Ile Arg Lys Ser Asp Lys Asp Asn Arg Gln Tyr Gln Ala Ile Arg  
35 40 45

Leu Asp Asn Gly Met Val Val Leu Leu Val Ser Asp Pro Gln Ala Val  
50 55 60

Lys Ser Leu Ser Ala Leu Val Val Pro Val Gly Ser Leu Glu Asp Pro  
65 70 75 80

Glu Ala Tyr Gln Gly Leu Ala His Tyr Leu Glu His Met Ser Leu Met  
85 90 95

Gly Ser Lys Lys Tyr Pro Gln Ala Asp Ser Leu Ala Glu Tyr Leu Lys  
100 105 110

Met His Gly Gly Ser His Asn Ala Ser Thr Ala Pro Tyr Arg Thr Ala  
115 120 125

Phe Tyr Leu Glu Val Glu Asn Asp Ala Leu Pro Gly Ala Val Asp Arg  
130 135 140

Leu Ala Asp Ala Ile Ala Glu Pro Leu Leu Asp Lys Lys Tyr Ala Glu  
145 150 155 160

Arg Glu Arg Asn Ala Val Asn Ala Glu Leu Thr Met Ala Arg Thr Arg  
165 170 175

Asp Gly Met Arg Met Ala Gln Val Ser Ala Glu Thr Ile Asn Pro Ala  
180 185 190

His Pro Gly Ser Lys Phe Ser Gly Gly Asn Leu Glu Thr Leu Ser Asp  
195 200 205

Lys Pro Gly Asn Pro Val Gln Gln Ala Leu Lys Asp Phe His Glu Lys  
210 215 220

Tyr Tyr Ser Ala Asn Leu Met Lys Ala Val Ile Tyr Ser Asn Lys Pro  
225 230 235 240

ES 2 627 826 T3

Leu Pro Glu Leu Ala Lys Met Ala Ala Asp Thr Phe Gly Arg Val Pro  
 245 250 255  
 Asn Lys Glu Ser Lys Lys Pro Glu Ile Thr Val Pro Val Val Thr Asp  
 260 265 270  
 Ala Gln Lys Gly Ile Ile Ile His Tyr Val Pro Ala Leu Pro Arg Lys  
 275 280 285  
 Val Leu Arg Val Glu Phe Arg Ile Asp Asn Asn Ser Ala Lys Phe Arg  
 290 295 300  
 Ser Lys Thr Asp Glu Leu Ile Thr Tyr Leu Ile Gly Asn Arg Ser Pro  
 305 310 315 320  
 Gly Thr Leu Ser Asp Trp Leu Gln Lys Gln Gly Leu Val Glu Gly Ile  
 325 330 335  
 Ser Ala Asn Ser Asp Pro Ile Val Asn Gly Asn Ser Gly Val Leu Ala  
 340 345 350  
 Ile Ser Ala Ser Leu Thr Asp Lys Gly Leu Ala Asn Arg Asp Gln Val  
 355 360 365  
 Val Ala Ala Ile Phe Ser Tyr Leu Asn Leu Leu Arg Glu Lys Gly Ile  
 370 375 380  
 Asp Lys Gln Tyr Phe Asp Glu Leu Ala Asn Val Leu Asp Ile Asp Phe  
 385 390 395 400  
 Arg Tyr Pro Ser Ile Thr Arg Asp Met Asp Tyr Val Glu Trp Leu Ala  
 405 410 415  
 Asp Thr Met Ile Arg Val Pro Val Glu His Thr Leu Asp Ala Val Asn  
 420 425 430  
 Ile Ala Asp Arg Tyr Asp Ala Lys Ala Val Lys Glu Arg Leu Ala Met  
 435 440 445  
 Met Thr Pro Gln Asn Ala Arg Ile Trp Tyr Ile Ser Pro Lys Glu Pro  
 450 455 460  
 His Asn Lys Thr Ala Tyr Phe Val Asp Ala Pro Tyr Gln Val Asp Lys  
 465 470 475 480  
 Ile Ser Ala Gln Thr Phe Ala Asp Trp Gln Lys Lys Ala Ala Asp Ile  
 485 490 495  
 Ala Leu Ser Leu Pro Glu Leu Asn Pro Tyr Ile Pro Asp Asp Phe Ser



ES 2 627 826 T3

Glu Tyr Thr Ser Ser Ala Tyr Ser Ser Leu Leu Gly Gln Ile Val Gln  
770 775 780

Pro Trp Phe Tyr Asn Gln Leu Arg Thr Glu Glu Gln Leu Gly Tyr Ala  
785 790 795 800

Val Phe Ala Phe Pro Met Ser Val Gly Arg Gln Trp Gly Met Gly Phe  
805 810 815

Leu Leu Gln Ser Asn Asp Lys Gln Pro Ser Phe Leu Trp Glu Arg Tyr  
820 825 830

Lys Ala Phe Phe Pro Thr Ala Glu Ala Lys Leu Arg Ala Met Lys Pro  
835 840 845

Asp Glu Phe Ala Gln Ile Gln Gln Ala Val Ile Thr Gln Met Leu Gln  
850 855 860

Ala Pro Gln Thr Leu Gly Glu Glu Ala Ser Lys Leu Ser Lys Asp Phe  
865 870 875 880

Asp Arg Gly Asn Met Arg Phe Asp Ser Arg Asp Lys Ile Val Ala Gln  
885 890 895

Ile Lys Leu Leu Thr Pro Gln Lys Leu Ala Asp Phe Phe His Gln Ala  
900 905 910

Val Val Glu Pro Gln Gly Met Ala Ile Leu Ser Gln Ile Ser Gly Ser  
915 920 925

Gln Asn Gly Lys Ala Glu Tyr Val His Pro Glu Gly Trp Lys Val Trp  
930 935 940

Glu Asn Val Ser Ala Leu Gln Gln Thr Met Pro Leu Met Ser Glu Lys  
945 950 955 960

Asn Glu

<210> 6

<211> 2915

5 <212> ADN

<213> *E. coli*

<400> 6

attccccgca gcacctgggtt caaagcatta ttgttgtag ttgcccttg ggcacattaa 60

tgtcaggcag aaacgggatg gcagccgatt caggaacca tccgtaaaag tgataaagat 120

10 aaccgccagt atcaggctat acgtctggat aacggtatgg tggctctgct gggttctgat 180

ES 2 627 826 T3

ccgcaggcag ttaaatacgt ctccgcgctg gtggtgcccg ttgggtcgct ggaagatccc 240  
 gaggcgtacc aggggctggc acattacctt gaacatatga gtctgatggg gtcgaaaaag 300  
 taccgcgagg ctgacagtct ggccgaatat ctcaaaatgc acggcggtag tcacaatgcc 360  
 agcaactgcgc cgtatcgcac ggctttctat ctggaagtgg agaacgacgc ettgcctggt 420  
 gcggtagacc gcctggccga tgctattgct gaacctttgc tcgacaagaa atatgccgaa 480  
 cgtgagcgta atgcggtgaa cgctgaatta accatggcgc gtacgcgtga cgggatgcgc 540  
 atggcacagg tcagcgcaga aaccattaac ccggcacacc ccggttcaaa gttttctggt 600  
 ggtaacctcg aaactttaag cgacaaacct ggtaatccgg tgcagcaggc gctgaaagat 660  
 ttccaagaga agtactattc cgccaatttg atgaaggcgg ttatttacag taataaacgg 720  
 ctcccgagtg tggcaaaaat ggccggcgac acctttggtc gcgtgccgaa caaagagagc 780  
 aaaaaaccgg aaatcacctg gccggtagtc accgacgcgc aaaagggcat tatcattcat 840  
 tacgtccctg cgctgccgcg taaagtgttg cgcggtgagt ttcgcatcga taacaactca 900  
 gcgaagtcc gtagtaaaac cgatgaattg attacctatc tgattggcaa tcgcagccca 960  
 ggtacacttt ctgactggct gcaaaagcag ggattagttg agggcattag cgccaactcc 1020  
 gatcctatcg tcaacggcaa cagcggcgta tttagcatct ctgcgtcttt aaccgataaa 1080  
 ggccctggcta atcgcgatca ggttggtggcg gcaattttta gctatctcaa tctgttacgt 1140  
 gaaaaaggca ttgataaaca atacttcgat gaactggcga atgtgctgga tatcgacttc 1200  
 cgttatccgt cgateacccg tgatatggat tacgtcgaat ggctggcaga taccatgatt 1260  
 cgcgttcctg ttgagcatac gctggatgca gtcaatattg ccgatcggta cgatgctaaa 1320  
 gcagtaaagg aacgtctggc gatgatgacg ccgcagaatg cgcgtatctg gtatatcagc 1380  
 ccgaaagagc cgcacaacaa aacggcttac ttgttcgatg cgcctatca ggtcgataaa 1440  
 atcagcgcac aaactttcgc cgactggcag aaaaaagccg ccgacattgc gctctcttg 1500  
 ccagagctta acccttatat tcctgatgat ttctcgtga ttaagtcaaa gaagaaatac 1560  
 gaccatccag agctgattgt tgatgagtcg aatctgcgcg tgggtgatgc gccaaagcgt 1620  
 tattttgcca gcgagcccaa agctgatgtc agcctgattt tgcgtaatcc gaaagccatg 1680  
 gacagcggcc gcaatcaggt gatgtttgcg ctcaatgatt atctcgcagg gctggcgctt 1740  
 gatcagttaa gcaaccaggc gtcggttggg ggcataagtt tttccaccaa cgtaacaac 1800  
 ggccttatgg ttaatgctaa tggttacacc cagcgtctgc cgcagctggt ccaggcattg 1860  
 ctcgaggggt actttagcta taccgctacg gaagatcagc ttgagcaggc gaagtccctgg 1920  
 tataaccaga tgatggattc cgcagaaaag ggtaaagcgt ttgagcaggc gattatgcc 1980  
 gcgcagatgc tctcgaagt gccgtacttc tcgcgagatg aacggcgtaa aattttgccc 2040  
 tccattacgt tgaaagaggt gctggcctat cgcgacgcct taaaatcagg ggctcgacca 2100

ES 2 627 826 T3

gagtttatgg ttatcggcaa catgaccgag gcccaggcaa caacgctggc acgcgatgtg 2160  
caaaaacagt tgggcgctga tggttcagag tgggtogaa acaaagatgt agtggtcgat 2220  
aaaaaacaat ccgtcatcct tgaaaaagcc ggtaacagca ccgactccgc actggcagcg 2280  
gtatttgtac cgactggcta cgatgaatac accagctcag cctatagctc tctgttgggg 2340  
cagatcgtac agccgtgggt ctacaatcag ttgctaccg aagaacaatt gggctatgcc 2400  
gtgtttgogt ttccaatgag cgtggggcgt cagtggggca tgggcttctt tttgcaaagc 2460  
aatgataaac agccttcatt cttgtgggag cgttacaagg cgtttttccc aaccgcagag 2520  
gcaaaattgc gagcgatgaa gccagatgag tttgcgcaa tccagcaggc ggtaattacc 2580  
cagatgctgc aggcaccgca aacgctcggc gaagaagcat cgaagttaag taaagatttc 2640  
gatcgcggca atatgcgctt cgattcgcgt gataaaatcg tggcccagat aaaactgctg 2700  
acgccgcaa aacttgctga tttcttccat caggcgggtg tcgagccgca aggcattggt 2760  
attctgtgc agatttccgg cagccagaac gggaaagccg aatatgtaca ccctgaaggc 2820  
tggaaagtgt gggagaacgt cagcgcgtt cagcaaaaca tgcccctgat gagtgaaaag 2880  
aatgagtgat gtcgccgaga cactagatcc tttgc 2915

<210> 7  
<211> 1425  
5 <212> ADN  
<213> *E. coli*

<400> 7

atgaaaaaaaa ccacattagc actgagtgca ctggctctga gtttaggttt ggcgttatct 60  
ccgctctctg caacggcggc tgagacttct tcagcaacga cagcccagca gatgccaagc 120  
cttgcacoga tgctcgaaaa ggtgatgcct tcagtggca gcattaacgt agaaggtagc 180  
acaaccgtta atacgccgct tatgccgctt aatttccagc agttcttcgg tgatgattct 240  
ccgttctgcc aggaaggttc tccgttccag agctctccgt tctgccaggg tggccagggc 300  
ggtaatggtg gcggccagca acagaaattc atggcgcgtg gttccggcgt catcattgat 360  
gccgataaag gctatgtcgt caccaacaac cacgttgttg ataacgcgac ggtcattaaa 420  
gttcaactga gcgatggcgg taagttcgac gcgaagatgg ttggcaaaga tccgcgctct 480  
gatatcgcgc tgatccaaat ccagaacctg aaaaacctga ccgcaattaa gatggcggat 540  
tctgatgcac tgcgcgtggg tgattacacc gtagcgattg gtaaccctgt tggctctggg 600  
gagacggtaa cttccgggat tgtctctcgc ctggggcgta gcggcctgaa tgccgaaaac 660  
tacgaaaact tcatccagac cgatgcagcg atcaacctg gtaactccgg tggcgcgctg 720  
gttaacctga acggcgaact gatcggatc aacaccgcca tcctcgacc ggacggcggc 780  
aacatcggta tcggttttgc tatcccgagt aacatggtga aaaacctgac ctgcgagatg 840  
10 gtggaatacg gccagggtgaa acgcgggtgag ctgggtatta tggggactga gctgaactcc 900

ES 2 627 826 T3

gaactggcga aagcgatgaa agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg 960  
 cctaattcct ccgctgcaaa agcgggcatt aaagcgggtg atgtgatcac ctactgaac 1020  
 ggtaagccga tcagcagcct tgccgcactg cgtgctcagg tgggtactat gccggtaggc 1080  
 agcaaaactga ccctgggctt actgcgcgac ggtaagcagg ttaacgtgaa cctggaactg 1140  
 cagcagagca gccagaatca ggttgattcc agctccatct tcaacggcat tgaaggcgtt 1200  
 gagatgagca acaaaaggca agatcagggc gtggtagtga acaacgtgaa aacgggcact 1260  
 ccggctgctc agatcggcct gaagaaaggt gatgtgatta ttggcgcgaa ccagcaggca 1320  
 gtgaaaaaca tcgctgaact gcgtaaagtt ctcgacagca aaccgtctgt gctggcactc 1380  
 aacattcagc gcggcgacag caccatctac ctgttaatgc agtaa 1425

<210> 8  
 <211> 474  
 <212> PRT  
 <213> *E. coli*

5

<400> 8

Met Lys Lys Thr Thr Leu Ala Leu Ser Ala Leu Ala Leu Ser Leu Gly  
 1 5 10 15  
 Leu Ala Leu Ser Pro Leu Ser Ala Thr Ala Ala Glu Thr Ser Ser Ala  
 20 25 30  
 Thr Thr Ala Gln Gln Met Pro Ser Leu Ala Pro Met Leu Glu Lys Val  
 35 40 45  
 Met Pro Ser Val Val Ser Ile Asn Val Glu Gly Ser Thr Thr Val Asn  
 50 55 60  
 Thr Pro Arg Met Pro Arg Asn Phe Gln Gln Phe Phe Gly Asp Asp Ser  
 65 70 75 80  
 Pro Phe Cys Gln Glu Gly Ser Pro Phe Gln Ser Ser Pro Phe Cys Gln  
 85 90 95  
 Gly Gly Gln Gly Gly Asn Gly Gly Gly Gln Gln Gln Lys Phe Met Ala  
 100 105 110  
 Leu Gly Ser Gly Val Ile Ile Asp Ala Asp Lys Gly Tyr Val Val Thr  
 115 120 125  
 Asn Asn His Val Val Asp Asn Ala Thr Val Ile Lys Val Gln Leu Ser  
 130 135 140  
 Asp Gly Arg Lys Phe Asp Ala Lys Met Val Gly Lys Asp Pro Arg Ser  
 145 150 155 160

10

ES 2 627 826 T3

Asp Ile Ala Leu Ile Gln Ile Gln Asn Pro Lys Asn Leu Thr Ala Ile  
 165 170 175  
 Lys Met Ala Asp Ser Asp Ala Leu Arg Val Gly Asp Tyr Thr Val Ala  
 180 185 190  
 Ile Gly Asn Pro Phe Gly Leu Gly Glu Thr Val Thr Ser Gly Ile Val  
 195 200 205  
 Ser Ala Leu Gly Arg Ser Gly Leu Asn Ala Glu Asn Tyr Glu Asn Phe  
 210 215 220  
 Ile Gln Thr Asp Ala Ala Ile Asn Arg Gly Asn Ser Gly Gly Ala Leu  
 225 230 235 240  
 Val Asn Leu Asn Gly Glu Leu Ile Gly Ile Asn Thr Ala Ile Leu Ala  
 245 250 255  
 Pro Asp Gly Gly Asn Ile Gly Ile Gly Phe Ala Ile Pro Ser Asn Met  
 260 265 270  
 Val Lys Asn Leu Thr Ser Gln Met Val Glu Tyr Gly Gln Val Lys Arg  
 275 280 285  
 Gly Glu Leu Gly Ile Met Gly Thr Glu Leu Asn Ser Glu Leu Ala Lys  
 290 295 300  
 Ala Met Lys Val Asp Ala Gln Arg Gly Ala Phe Val Ser Gln Val Leu  
 305 310 315 320  
 Pro Asn Ser Ser Ala Ala Lys Ala Gly Ile Lys Ala Gly Asp Val Ile  
 325 330 335  
 Thr Ser Leu Asn Gly Lys Pro Ile Ser Ser Phe Ala Ala Leu Arg Ala  
 340 345 350  
 Gln Val Gly Thr Met Pro Val Gly Ser Lys Leu Thr Leu Gly Leu Leu  
 355 360 365  
 Arg Asp Gly Lys Gln Val Asn Val Asn Leu Glu Leu Gln Gln Ser Ser  
 370 375 380  
 Gln Asn Gln Val Asp Ser Ser Ser Ile Phe Asn Gly Ile Glu Gly Ala  
 385 390 395 400  
 Glu Met Ser Asn Lys Gly Lys Asp Gln Gly Val Val Val Asn Asn Val  
 405 410 415

ES 2 627 826 T3

Lys Thr Gly Thr Pro Ala Ala Gln Ile Gly Leu Lys Lys Gly Asp Val  
 420 425 430

Ile Ile Gly Ala Asn Gln Gln Ala Val Lys Asn Ile Ala Glu Leu Arg  
 435 440 445

Lys Val Leu Asp Ser Lys Pro Ser Val Leu Ala Leu Asn Ile Gln Arg  
 450 455 460

Gly Asp Ser Thr Ile Tyr Leu Leu Met Gln  
 465 470

<210> 9  
 <211> 1425  
 <212> ADN  
 <213> *E. coli*

<400> 9

atgaaaaaaaa ccacattagc actgagtgca ctggctctga gtttaggttt ggcgttatct 60  
 ccgctctctg caacggcggc tgagacttct tcagcaacga cagcccagca gatgccaaagc 120  
 cttgcaccga tgctcgaaaa ggtgatgcct tcagtggca gcattaacgt agaaggtagc 180  
 acaaccgtta atacgccgcg tatgccgcgt aattccagc agttcttcgg tgatgattct 240  
 ccgttctgcc aggaaggttc tccgttcag agctctcctg tctgccaggg tggccagggc 300  
 ggtaatggg gcggccagca acagaaattc atggcgctgg gttccggcgt catcattgat 360  
 gccgataaag gctatgtcgt caccaacaac cacgttggtg ataacgcgac ggtcattaaa 420  
 gttcaactga gcgatggccc taagttcgac gcgaagatgg ttggcaaaga tccgcgctct 480  
 gatatcggc tgatccaaat ccagaaccgg aaaaacctga ccgcaattaa gatggcggat 540  
 tctgatgcac tgcgcgtggg tgattacacc gttagcgattg gtaaccctgt tggctctgggc 600  
 gagacggtaa cttccgggat tgtctctcgc ctggggcgta gcggcctgaa tgccgaaaaac 660  
 tacgaaaact tcatccagac cgatgcagcg attaatcgtg gtaacgcccg tggcgcgctg 720  
 gttaacctga acggcgaact gatcggatc aacaccgca tcctcgacc ggacggcggc 780  
 aacatcggt tgcggtttgc tatcccagat aacatggtga aaaacctgac ctgcagatg 840  
 gtggaatacg gccagggtgaa acgcggtgag ctgggtatta tggggactga gctgaactcc 900  
 gaactggcga aagcgatgaa agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg 960  
 cctaattcct ccgctgcaaa agcgggcatt aaagcgggtg atgtgatcac ctactgaac 1020  
 ggtaagccga tcagcagctt tgccgcactg cgtgctcagg tgggtactat gccggtaggc 1080  
 agcaaaactga ccctgggctt actgscgac ggtaagcagg ttaacgtgaa cctggaactg 1140  
 cagcagagca gccagaatca ggttgattcc agctccatct tcaacggcat tgaaggcgt 1200  
 gagatgagca acaaaggcaa agatcagggc gtggtagtga acaacgtgaa aacgggcaact 1260  
 ccggtgcgc agatcggcct gaagaaaggat gatgtgatta ttggcgcgaa ccagcaggca 1320  
 gtgaaaaaca tcgctgaact gcgtaaagtt ctcgacagca aaccgtctgt gctggcactc 1380  
 aacattcagc gcggcgacag caccatctac ctgttaatgc agtaa 1425

<210> 10

ES 2 627 826 T3

<211> 474  
 <212> PRT  
 <213> *E. coli*

5 <400> 10

Met Lys Lys Thr Thr Leu Ala Leu Ser Ala Leu Ala Leu Ser Leu Gly  
 1 5 10 15

Leu Ala Leu Ser Pro Leu Ser Ala Thr Ala Ala Glu Thr Ser Ser Ala  
 20 25 30

Thr Thr Ala Gln Gln Met Pro Ser Leu Ala Pro Met Leu Glu Lys Val  
 35 40 45

Met Pro Ser Val Val Ser Ile Asn Val Glu Gly Ser Thr Thr Val Asn  
 50 55 60

Thr Pro Arg Met Pro Arg Asn Phe Gln Gln Phe Phe Gly Asp Asp Ser  
 65 70 75 80

Pro Phe Cys Gln Glu Gly Ser Pro Phe Gln Ser Ser Pro Phe Cys Gln  
 85 90 95

Gly Gly Gln Gly Gly Asn Gly Gly Gly Gln Gln Gln Lys Phe Met Ala  
 100 105 110

Leu Gly Ser Gly Val Ile Ile Asp Ala Asp Lys Gly Tyr Val Val Thr  
 115 120 125

Asn Asn His Val Val Asp Asn Ala Thr Val Ile Lys Val Gln Leu Ser  
 130 135 140

Asp Gly Arg Lys Phe Asp Ala Lys Met Val Gly Lys Asp Pro Arg Ser  
 145 150 155 160

Asp Ile Ala Leu Ile Gln Ile Gln Asn Pro Lys Asn Leu Thr Ala Ile  
 165 170 175

Lys Met Ala Asp Ser Asp Ala Leu Arg Val Gly Asp Tyr Thr Val Ala  
 180 185 190

Ile Gly Asn Pro Phe Gly Leu Gly Glu Thr Val Thr Ser Gly Ile Val  
 195 200 205

ES 2 627 826 T3

Ser Ala Leu Gly Arg Ser Gly Leu Asn Ala Glu Asn Tyr Glu Asn Phe  
 210 215 220

Ile Gln Thr Asp Ala Ala Ile Asn Arg Gly Asn Ala Gly Gly Ala Leu  
 225 230 235 240

Val Asn Leu Asn Gly Glu Leu Ile Gly Ile Asn Thr Ala Ile Leu Ala  
 245 250 255

Pro Asp Gly Gly Asn Ile Gly Ile Gly Phe Ala Ile Pro Ser Asn Met  
 260 265 270

Val Lys Asn Leu Thr Ser Gln Met Val Glu Tyr Gly Gln Val Lys Arg  
 275 280 285

Gly Glu Leu Gly Ile Met Gly Thr Glu Leu Asn Ser Glu Leu Ala Lys  
 290 295 300

Ala Met Lys Val Asp Ala Gln Arg Gly Ala Phe Val Ser Gln Val Leu  
 305 310 315 320

Pro Asn Ser Ser Ala Ala Lys Ala Gly Ile Lys Ala Gly Asp Val Ile  
 325 330 335

Thr Ser Leu Asn Gly Lys Pro Ile Ser Ser Phe Ala Ala Leu Arg Ala  
 340 345 350

Gln Val Gly Thr Met Pro Val Gly Ser Lys Leu Thr Leu Gly Leu Leu  
 355 360 365

Arg Asp Gly Lys Gln Val Asn Val Asn Leu Glu Leu Gln Gln Ser Ser  
 370 375 380

Gln Asn Gln Val Asp Ser Ser Ser Ile Phe Asn Gly Ile Glu Gly Ala  
 385 390 395 400

Glu Met Ser Asn Lys Gly Lys Asp Gln Gly Val Val Val Asn Asn Val  
 405 410 415

Lys Thr Gly Thr Pro Ala Ala Gln Ile Gly Leu Lys Lys Gly Asp Val  
 420 425 430

Ile Ile Gly Ala Asn Gln Gln Ala Val Lys Asn Ile Ala Glu Leu Arg  
 435 440 445

Lys Val Leu Asp Ser Lys Pro Ser Val Leu Ala Leu Asn Ile Gln Arg  
 450 455 460

Gly Asp Ser Thr Ile Tyr Leu Leu Met Gln

465 470

- 5 <210> 11
- <211> 107
- <212> PRT
- <213> Secuencia artificial

- 10 <220>
- <223> hTNF40-gL1

ES 2 627 826 T3

<400> 11

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

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

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

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

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

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Asn Ile Tyr Pro Leu  
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys  
100 105

- 5 <210> 12
- <211> 118
- <212> PRT
- <213> Secuencia artificial

- 10 <220>
- <223> gh3h TNF40.4

<400> 12

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly  
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Val Phe Thr Asp Tyr  
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met  
35 40 45

- 15 Gly Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Ala Asp Ser Val  
50 55 60

Lys Gly Arg Phe Thr Phe Ser Leu Asp Thr Ser Lys Ser Thr Ala Tyr  
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys  
85 90 95

Ala Arg Gly Tyr Arg Ser Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr  
100 105 110

Leu Val Thr Val Ser Ser  
115

<210> 13

ES 2 627 826 T3

<211> 214  
 <212> PRT  
 <213> Secuencia artificial

5 <220>  
 <223> Cadena Ligera Injertada

<400> 13

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

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

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

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

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

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Asn Ile Tyr Pro Leu  
 85 90 95

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

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

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

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

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

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

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

Phe Asn Arg Gly Glu Cys  
 210

<210> 14  
 <211> 229  
 <212> PRT  
 <213> Secuencia artificial

15 <220>

ES 2 627 826 T3

<223> Cadena Pesada Injertada

<400> 14

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly  
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Val Phe Thr Asp Tyr  
20 25 30

Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Met  
35 40 45

Gly Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Ala Asp Ser Val  
50 55 60

Lys Gly Arg Phe Thr Phe Ser Leu Asp Thr Ser Lys Ser Thr Ala Tyr  
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys  
85 90 95

Ala Arg Gly Tyr Arg Ser Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr  
100 105 110

Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro  
115 120 125

5 Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly  
130 135 140

Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn  
145 150 155 160

Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln  
165 170 175

Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser  
180 185 190

Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser  
195 200 205

Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr  
210 215 220

His Thr Cys Ala Ala  
225

10 <210> 15

<211> 24

<212> ADN

<213> Secuencia artificial

<220>

15 <223> Cebador oligonucleótido

<400> 15

gcatcataat tttctttta cctc 24

5 <210> 16  
 <211> 24  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <223> Cebador oligonucleótido

<400> 16

15 gggaaatgaa cctgagcaaa acgc 24

<210> 17  
 <211> 24  
 <212> ADN  
 <213> Secuencia artificial

20 <220>  
 <223> Cebador oligonucleótido

<400> 17

25 gtgccaggag atgcagcagc ttgc 24

<210> 18  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial

30 <220>  
 <223> Cebador oligonucleótido

35 <400> 18

tttcagcca gtcagaaagt g 21

40 <210> 19  
 <211> 24  
 <212> ADN  
 <213> Secuencia artificial

45 <220>  
 <223> Cebador oligonucleótido

<400> 19

50 ctgcctgcga tttcgccgg aacg 24

<210> 20  
 <211> 24  
 <212> ADN

55 <213> Secuencia artificial

<220>  
 <223> Cebador oligonucleótido

60 <400> 20

cgcatggtac gtgccacgat atcc 24

65 <210> 21  
 <211> 188  
 <212> PRT

ES 2 627 826 T3

<213> *Escherichia coli*

<400> 21

Met Val Lys Ser Gln Pro Ile Leu Arg Tyr Ile Leu Arg Gly Ile Pro  
1 5 10 15

Ala Ile Ala Val Ala Val Leu Leu Ser Ala Cys Ser Ala Asn Asn Thr  
20 25 30

Ala Lys Asn Met His Pro Glu Thr Arg Ala Val Gly Ser Glu Thr Ser  
35 40 45

Ser Leu Gln Ala Ser Gln Asp Glu Phe Glu Asn Leu Val Arg Asn Val  
50 55 60

Asp Val Lys Ser Arg Ile Met Asp Gln Tyr Ala Asp Trp Lys Gly Val  
65 70 75 80

Arg Tyr Arg Leu Gly Gly Ser Thr Lys Lys Gly Ile Asp Cys Ser Gly  
85 90 95

5 Phe Val Gln Arg Thr Phe Arg Glu Gln Phe Gly Leu Glu Leu Pro Arg  
100 105 110

Ser Thr Tyr Glu Gln Gln Glu Met Gly Lys Ser Val Ser Arg Ser Asn  
115 120 125

Leu Arg Thr Gly Asp Leu Val Leu Phe Arg Ala Gly Ser Thr Gly Arg  
130 135 140

His Val Gly Ile Tyr Ile Gly Asn Asn Gln Phe Val His Ala Ser Thr  
145 150 155 160

Ser Ser Gly Val Ile Ile Ser Ser Met Asn Glu Pro Tyr Trp Lys Lys  
165 170 175

Arg Tyr Asn Glu Ala Arg Arg Val Leu Ser Arg Ser  
180 185

<210> 22

10 <211> 162

<212> PRT

<213> *Escherichia coli*

<400> 22

15

ES 2 627 826 T3

Cys Ser Ala Asn Asn Thr Ala Lys Asn Met His Pro Glu Thr Arg Ala  
 1 5 10 15

Val Gly Ser Glu Thr Ser Ser Leu Gln Ala Ser Gln Asp Glu Phe Glu  
 20 25 30

Asn Leu Val Arg Asn Val Asp Val Lys Ser Arg Ile Met Asp Gln Tyr  
 35 40 45

Ala Asp Trp Lys Gly Val Arg Tyr Arg Leu Gly Gly Ser Thr Lys Lys  
 50 55 60

Gly Ile Asp Cys Ser Gly Phe Val Gln Arg Thr Phe Arg Glu Gln Phe  
 65 70 75 80

Gly Leu Glu Leu Pro Arg Ser Thr Tyr Glu Gln Gln Glu Met Gly Lys  
 85 90 95

Ser Val Ser Arg Ser Asn Leu Arg Thr Gly Asp Leu Val Leu Phe Arg  
 100 105 110

Ala Gly Ser Thr Gly Arg His Val Gly Ile Tyr Ile Gly Asn Asn Gln  
 115 120 125

Phe Val His Ala Ser Thr Ser Ser Gly Val Ile Ile Ser Ser Met Asn  
 130 135 140

Glu Pro Tyr Trp Lys Lys Arg Tyr Asn Glu Ala Arg Arg Val Leu Ser  
 145 150 155 160

Arg Ser

- <210> 23
- 5 <211> 951
- <212> ADN
- <213> Secuencia artificial

- <220>
- 10 <223> Secuencia de OmpT mutada

<400> 23

ES 2 627 826 T3

atgCGggGga aacttctggg aatagtcctg acaacccta ttgCgatcag ctcttttGct 60  
 tctaccgaga ctttatcggt tactcctgac aacataaatg cggacattag tcttggaact 120  
 ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggaggccg aaaagtcagt 180  
 caactcgact ggaattcaa taacgctgca attattaaag gtgcaattaa ttgggatttg 240  
 atgccccaga tatctatcgg ggctgctggc tggacaactc tcggcagccg aggtggcaat 300  
 atggtcgatc aggactggat ggattccagt aaccocggaa cctggacgga tgaaagtaga 360  
 caccctgata cacaactcaa ttatgccaac gaatttgatc tgaatatcaa aggctggctc 420  
 ctcaacgaac ccaattaccg cctgggactc atggccggat atcaggaaag ccgttatagc 480  
 tttacagcca gaggtgggtc ctatatctac agttctgagg agggattcag agatgatatc 540  
 ggctccttcc cgaatggaga aagagcaatc ggctacaaac aacgttttaa aatgcctac 600  
 attggcttga ctggaagtta tcgttatgaa gattttgaac tcggtggcac atttaaatac 660  
 agcggctggg tggaatcatc tgataacgct gaagcttatg acccgggaaa aagaatcact 720  
 tatcgcagta aggtcaaaga ccaaaattac tattctgttg cagtcaatgc aggttattac 780  
 gtcacaccta acgcaaaagt ttatgttgaa ggcgcatgga atcgggttac gaataaaaaa 840  
 ggtaatactt cactttatga tcacaataat aacacttcag actacagcaa aatggagca 900  
 ggtatagaaa actataactt catcactact gctggtotta agtacacatt t 951

<210> 24

<211> 317

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de OmpT mutada

<400> 24

Met Arg Ala Lys Leu Leu Gly Ile Val Leu Thr Thr Pro Ile Ala Ile  
 1 5 10 15

Ser Ser Phe Ala Ser Thr Glu Thr Leu Ser Phe Thr Pro Asp Asn Ile  
 20 25 30

ES 2 627 826 T3

Asn Ala Asp Ile Ser Leu Gly Thr Leu Ser Gly Lys Thr Lys Glu Arg  
 35 40 45

Val Tyr Leu Ala Glu Glu Gly Gly Arg Lys Val Ser Gln Leu Asp Trp  
 50 55 60

Lys Phe Asn Asn Ala Ala Ile Ile Lys Gly Ala Ile Asn Trp Asp Leu  
 65 70 75 80

Met Pro Gln Ile Ser Ile Gly Ala Ala Gly Trp Thr Thr Leu Gly Ser  
 85 90 95

Arg Gly Gly Asn Met Val Asp Gln Asp Trp Met Asp Ser Ser Asn Pro  
 100 105 110

Gly Thr Trp Thr Asp Glu Ser Arg His Pro Asp Thr Gln Leu Asn Tyr  
 115 120 125

Ala Asn Glu Phe Asp Leu Asn Ile Lys Gly Trp Leu Leu Asn Glu Pro  
 130 135 140

Asn Tyr Arg Leu Gly Leu Met Ala Gly Tyr Gln Glu Ser Arg Tyr Ser  
 145 150 155 160

Phe Thr Ala Arg Gly Gly Ser Tyr Ile Tyr Ser Ser Glu Glu Gly Phe  
 165 170 175

Arg Asp Asp Ile Gly Ser Phe Pro Asn Gly Glu Arg Ala Ile Gly Tyr  
 180 185 190

Lys Gln Arg Phe Lys Met Pro Tyr Ile Gly Leu Thr Gly Ser Tyr Arg  
 195 200 205

Tyr Glu Asp Phe Glu Leu Gly Gly Thr Phe Lys Tyr Ser Gly Trp Val  
 210 215 220

Glu Ser Ser Asp Asn Ala Glu Ala Tyr Asp Pro Gly Lys Arg Ile Thr  
 225 230 235 240

Tyr Arg Ser Lys Val Lys Asp Gln Asn Tyr Tyr Ser Val Ala Val Asn  
 245 250 255

Ala Gly Tyr Tyr Val Thr Pro Asn Ala Lys Val Tyr Val Glu Gly Ala  
 260 265 270

Trp Asn Arg Val Thr Asn Lys Lys Gly Asn Thr Ser Leu Tyr Asp His  
 275 280 285

Asn Asn Asn Thr Ser Asp Tyr Ser Lys Asn Gly Ala Gly Ile Glu Asn  
 290 295 300

Tyr Asn Phe Ile Thr Thr Ala Gly Leu Lys Tyr Thr Phe  
 305 310 315

5 <210> 25  
 <211> 954  
 <212> ADN

ES 2 627 826 T3

<213> Secuencia artificial

<220>

<223> Secuencia de OmpT mutada

5

<400> 25

```

attcgggaga aacttctggg aatagtcctg acaacccta ttgcgatcag ctcttttgc 60
tctaccgaga ctttatcggt tactcctgac aacataaatg cggacattag tcttggaact 120
ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggaggccg aaaagtcagt 180
caactcgact ggaattcaa taacgctgca attattaaag gtgcaattaa ttgggatttg 240
atgccccaga tatctatcgg ggctgctggc tggacaactc tcggcagccg aggtggcaat 300
atggtcgatc aggactggat ggattccagt aaccocggaa cctggacgga tgaaagtaga 360
caccctgata cacaactcaa ttatgccaac gaatttgatc tgaatatcaa aggctggctc 420
ctcaacgaac ccaattaccg cctgggactc atggccggat atcaggaaag ccgttatagc 480
tttacagcca gaggtgggtc ctatatctac agttctgagg agggattcag agatgatatc 540
ggctccttcc cgaatggaga aagagcaatc ggctacaaac aacgttttaa aatgccctac 600
attggcttga ctggaagtta tcgttatgaa gattttgaac tcggtggcac atttaaatac 660
agcggctggg tggaatcatc tgataacgat gaacactatg acccgggaaa aagaatcact 720
tatcgcagta aggtcaaaga ccaaaattac tattctgttg cagtcaatgc aggttattac 780
gtcacaccta acgcaaaagt ttatggtgaa ggcgcatgga atcgggttac gaataaaaaa 840
ggtaatactt cactttatga tcacaataat aacacttcag actacagcaa aatggagca 900
ggtatagaaa actataactt catcactact gctggtotta agtacacatt ttaa 954

```

10

<210> 26

<211> 729

<212> ADN

<213> Secuencia artificial

15

<220>

<223> Secuencia de DsbC mutada

<400> 26

```

atgaagaaag gttttatggt gtttactttg tttagcggcgt tttcaggctt tgctcaggct 60
gatgacgogg caattcaaca aacgtagcc aaaatgggca tcaaagcag cgatattcag 120
cccgcgctg tagctggcat gaagacagtt ctgactaaca gcggcgtggt gtacatcacc 180

```

20

ES 2 627 826 T3

gatgatggta aacatatcat tcaggggcca atgtatgacg ttagtggcac ggctccggtc 240  
aatgtcacca ataagatgct gttaaagcag ttgaatgcgc ttgaaaaaga gatgatcggt 300  
tataaagcgc cgcaggaaaa acacgtcatc accgtgttta ctgatattac ctgtggttac 360  
tgccacaaac tgcattgagca aatggcagac tacaacgcgc tggggatcac cgtgcgttat 420  
cttgctttcc cgcgccaggg gctggacagc gatgcagaga aagaaatgaa agctatctgg 480  
tgtgcgaaag ataaaaacaa agcgtttgat gatgtgatgg caggtaaaag cgtcgcacca 540  
gccagttgcy acgtggatat tgccgacatc tacgcacttg gcgtccagct tggcgttagc 600  
ggctactccgg cagttgtgct gagcaatggc acacttgctc cgggttacca gccgccgaaa 660  
gagatgaaag aattttctga cgaacaccaa aaaatgacca gcggtaaaca ccatcaccat 720  
caccactaa 729

<210> 27  
<211> 242  
5 <212> PRT  
<213> Secuencia artificial

<220>  
10 <223> Secuencia de DsbC mutada

<400> 27

Met Lys Lys Gly Phe Met Leu Phe Thr Leu Leu Ala Ala Phe Ser Gly  
1 5 10 15  
Phe Ala Gln Ala Asp Asp Ala Ala Ile Gln Gln Thr Leu Ala Lys Met  
20 25 30  
Gly Ile Lys Ser Ser Asp Ile Gln Pro Ala Pro Val Ala Gly Met Lys  
35 40 45  
Thr Val Leu Thr Asn Ser Gly Val Leu Tyr Ile Thr Asp Asp Gly Lys  
50 55 60  
His Ile Ile Gln Gly Pro Met Tyr Asp Val Ser Gly Thr Ala Pro Val  
65 70 75 80  
Asn Val Thr Asn Lys Met Leu Leu Lys Gln Leu Asn Ala Leu Glu Lys  
85 90 95  
Glu Met Ile Val Tyr Lys Ala Pro Gln Glu Lys His Val Ile Thr Val  
100 105 110  
Phe Thr Asp Ile Thr Cys Gly Tyr Cys His Lys Leu His Glu Gln Met  
115 120 125  
Ala Asp Tyr Asn Ala Leu Gly Ile Thr Val Arg Tyr Leu Ala Phe Pro



ES 2 627 826 T3

Gly Tyr Arg Ser Tyr Ala Met Asp Tyr  
1 5

5 <210> 31  
<211> 11  
<212> PRT  
<213> Secuencia artificial

<220>  
10 <223> Descripción de la Secuencia Artificial: CDRL1 de hTNF40

<400> 31

Lys Ala Ser Gln Asn Val Gly Thr Asn Val Ala  
1 5 10

15 <210> 32  
<211> 7  
<212> PRT  
<213> Secuencia artificial

20 <220>  
<223> Descripción de la Secuencia Artificial: CDRL2 de hTNF40

<400> 32

25 Ser Ala Ser Phe Leu Tyr Ser  
1 5

30 <210> 33  
<211> 9  
<212> PRT  
<213> Secuencia artificial

<220>  
35 <223> Descripción de la Secuencia Artificial: CDRL3 de hTNF40

<400> 33

Gln Gln Tyr Asn Ile Tyr Pro Leu Thr  
1 5

40 <210> 34  
<211> 17  
<212> PRT  
<213> Secuencia artificial

45 <220>  
<223> Descripción de la Secuencia Artificial: CDRH2 de hTNF40

<400> 34

50 Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Val Asp Asp Phe Lys  
1 5 10 15

Gly

<210> 35  
<211> 84

# ES 2 627 826 T3

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Adaptador de oligonucleótidos OmpA	
	<400> 35	
	tcgagttcta gataacgagg cgtaaaaaat gaaaaagaca gctatcgcaa ttgcagtggc	60
10	cttggctctg acgtacgagt cagg	84
	<210> 36	
	<211> 67	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Casete de IGS 1	
	<400> 36	
20	gagctcacca gtaacaaaaa gttttaatag aggagagtgt taatgaagaa gactgctata	60
	gcaattg	67
	<210> 37	
	<211> 69	
25	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Casete de IGS 2	
30	<400> 37	
	gagctcacca gtaacaaaaa gttttaatag aggggagtgt taaaatgaag aagactgcta	60
	tagcaattg	69
35	<210> 38	
	<211> 81	
	<212> ADN	
	<213> Secuencia artificial	
40	<220>	
	<223> Casete de IGS 3	
	<400> 38	
	gagctcacca gtaacaaaaa gctttaatag aggagagtgt tgaggaggaa aaaaaaatga	60
45	agaaaactgc tatagcaatt g	81
	<210> 39	
	<211> 81	
	<212> ADN	
50	<213> Secuencia artificial	
	<220>	
	<223> Casete de IGS 4	
55	<400> 39	

ES 2 627 826 T3

gagtcacca gtaacaaaaa gttttaatag aggagagtgt tgacgaggat tatataatga 60  
agaaaactgc tatagcaatt g 81

**REIVINDICACIONES**

- 1.- Una célula bacteriana gram-negativa recombinante, caracterizada porque la célula:
- a) comprende un vector de expresión que comprende un polinucleótido recombinante que codifica DsbC;
  - 5 b) tiene un gen *Tsp* mutado que codifica una proteína *Tsp* que tiene una reducción de la actividad proteasa en comparación con una célula de tipo silvestre,
  - c) comprende un gen *spr* mutado que codifica una proteína *spr* mutante capaz de suprimir el fenotipo de una célula que comprende un gen *Tsp* mutado, y
  - 10 d) tiene un genoma que es isogénico al de la cepa de *E. coli* de tipo silvestre W3110, excepto por el gen *Tsp* mutado y el gen *spr* mutado.
- 2.- Una célula según la reivindicación 1, en la que el gen *spr* mutado codifica una proteína *spr* como se define en SEQ ID NO: 21 que tiene una mutación en uno o varios aminoácidos seleccionados a partir de N31, R62, I70, Q73, C94, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144, H145, G147 y H157.
- 3.- Una célula según la reivindicación 2, en la que el gen *spr* mutante codifica una proteína *spr* que tiene una o varias mutaciones seleccionadas a partir de N31Y, R62C, I70T, Q73R, C94A, S95F, V98E, Q99P, R100G, L108S, Y115F, D133A, V135D, V135G, L136P, G140C, R144C, H145A, G147C y H157A.
- 15 4.- Una célula según la reivindicación 3, en la que el gen *spr* mutante codifica una proteína *spr* que tiene la mutación C94A.
- 5.- Una célula según la reivindicación 3, en la que el gen *spr* mutante codifica una proteína *spr* que tiene la mutación H145A.
- 20 6.- Una célula según la reivindicación 3, en la que el gen *spr* mutante codifica una proteína *spr* que tiene las mutaciones S95F e Y115F.
- 7.- Una célula según una cualquiera de las reivindicaciones 1 a 6, en donde la célula tiene un gen *Tsp* mutado inactivado que comprende una mutación del codón de inicio del gen y/o comprende uno o varios codones de parada situados aguas abajo del codón de inicio del gen y aguas arriba del codón de parada del gen.
- 25 8.- Una célula según la reivindicación 7, en la que el gen *Tsp* mutado inactivado comprende un sitio de restricción marcador creado mediante una mutación de sentido erróneo en el codón de inicio del gen y opcionalmente una o varias mutaciones puntuales adicionales.
- 9.- Una célula según la reivindicación 8, en la que el gen *Tsp* mutado inactivado comprende SEQ ID NO: 3.
- 30 10.- Una célula según una cualquiera de las reivindicaciones 1 a 9, en donde la célula comprende una secuencia de polinucleótido que codifica una proteína de interés.
- 11.- Una célula según la reivindicación 10, en donde la célula comprende un vector que comprende el polinucleótido recombinante que codifica DsbC y la secuencia de polinucleótido que codifica una proteína de interés.
- 35 12.- Una célula según la reivindicación 11, en la que el vector comprende un promotor que controla la expresión del polinucleótido recombinante que codifica DsbC y la secuencia de polinucleótido que codifica una proteína de interés.
- 13.- Una célula según una cualquiera de las reivindicaciones 10 a 12, en donde la proteína de interés es un anticuerpo o un fragmento del mismo que se une a antígeno.
- 14.- Una célula según la reivindicación 13, en la que el anticuerpo o un fragmento del mismo que se une a antígeno es un Fab, Fab modificado o Fab'.
- 40 15.- Una célula según la reivindicación 13 o 14, en la que el anticuerpo o un fragmento del mismo que se une a antígeno es específico de TNF.
- 16.- Una célula según la reivindicación 15, en la que el anticuerpo o un fragmento que se une a antígeno comprende
- 45 a) una cadena pesada en la que el dominio variable comprende la secuencia mostrada en SEQ ID NO: 28 para la CDRH1, la secuencia mostrada en SEQ ID NO: 29 o SEQ ID NO: 34 para la CDRH2 y la secuencia mostrada en SEQ ID NO: 30 para la CDRH3 y
  - b) una cadena ligera en la que el dominio variable comprende la secuencia mostrada en SEQ ID NO: 31 para la CDRL1, la secuencia mostrada en SEQ ID NO: 32 para la CDRL2 y la secuencia mostrada en SEQ ID NO: 33 para la CDRL3.

17.- Una célula según la reivindicación 15, en donde el anticuerpo o un fragmento que se une a antígeno es un Fab' y tiene una secuencia de la cadena ligera que comprende o consiste en SEQ ID NO: 13 y una secuencia de la cadena pesada que comprende o consiste en SEQ ID NO: 14.

18.- Un método para producir una proteína de interés que comprende:

5 a) cultivar una célula bacteriana gram-negativa recombinante según se define en una cualquiera de las reivindicaciones 1 a 17 en un medio de cultivo en condiciones eficaces para expresar la proteína recombinante de interés y el polinucleótido recombinante que codifica DsbC; y

b) recuperar la proteína de interés a partir del periplasma de la célula bacteriana gram-negativa recombinante y/o el medio de cultivo.

10 19.- Un método según la reivindicación 18, en el que la expresión de la secuencia de polinucleótido que codifica una proteína de interés y el polinucleótido recombinante que codifica DsbC se induce mediante la adición de un inductor al medio de cultivo.

20.- Un método según la reivindicación 18 o 19, en el que el método comprende adicionalmente la separación de la proteína de interés a partir de DsbC.

15 21.- El método según una cualquiera de las reivindicaciones 18 a 20, que comprende además la etapa de procesamiento aguas abajo mediante PEGilación de la proteína de interés.

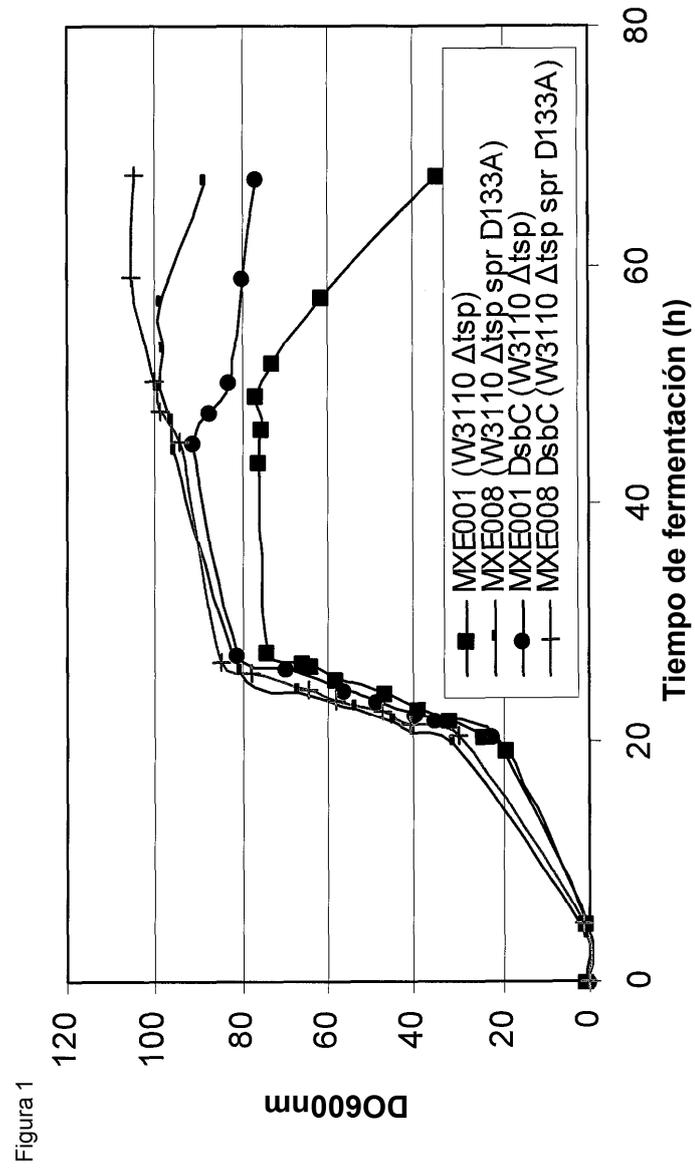
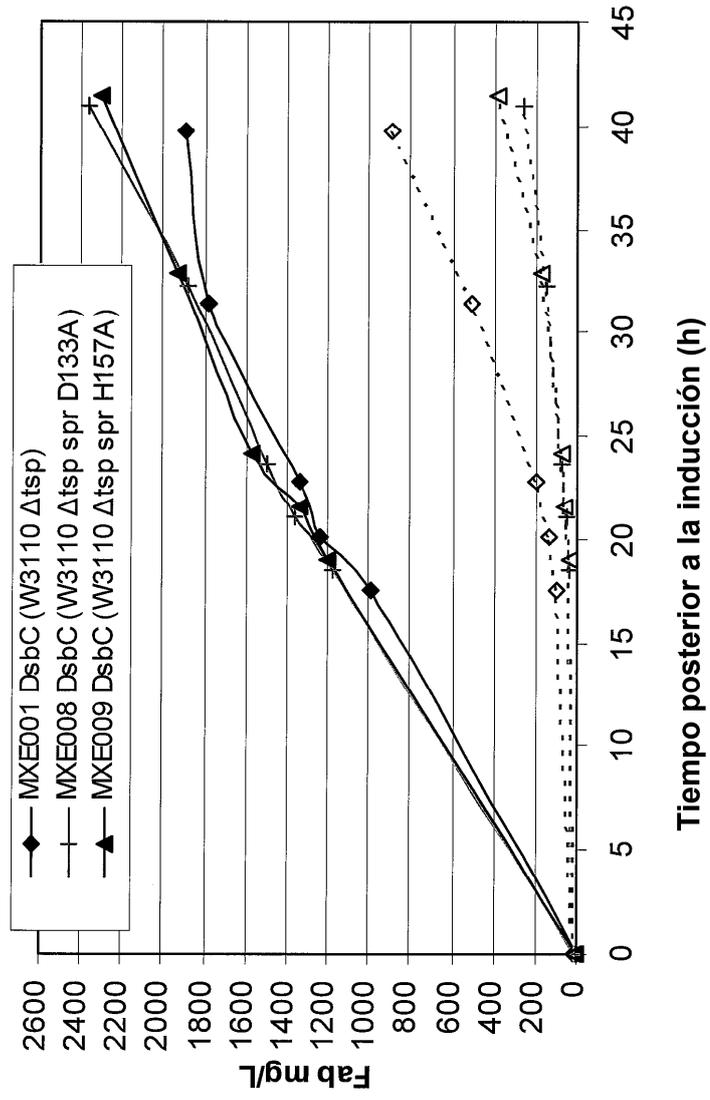


Figura 2



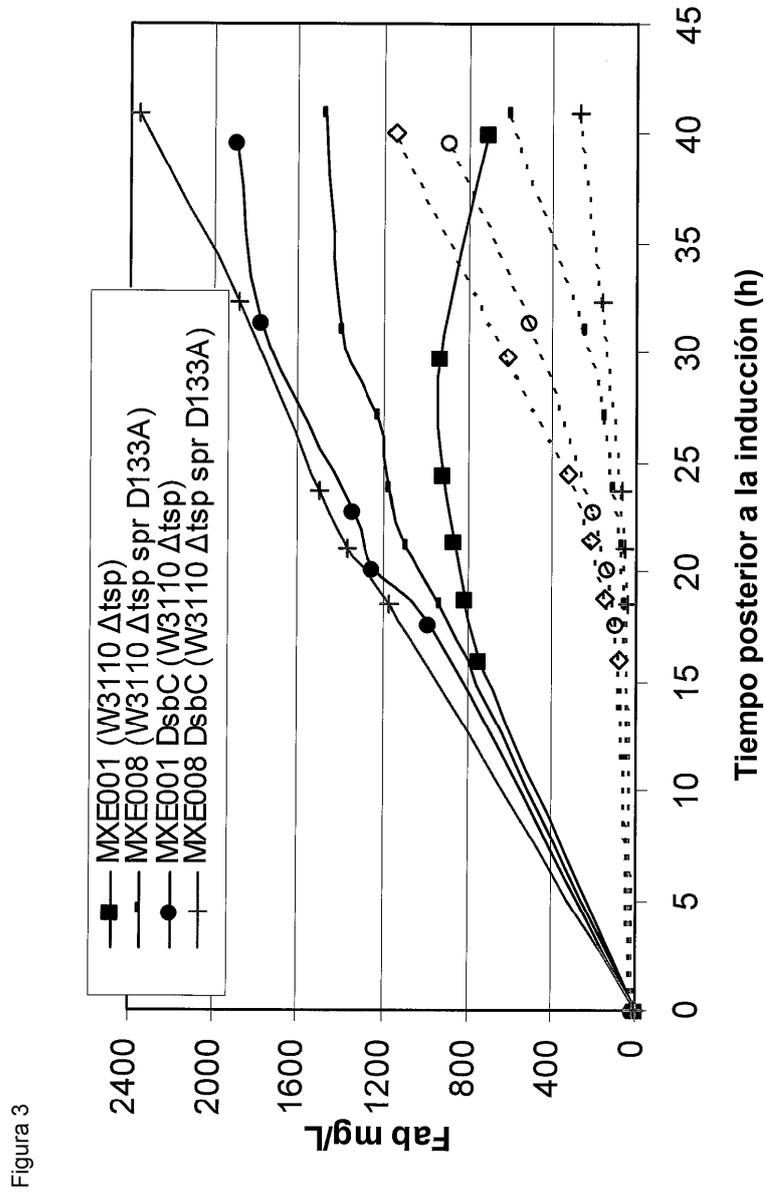


Figura 4

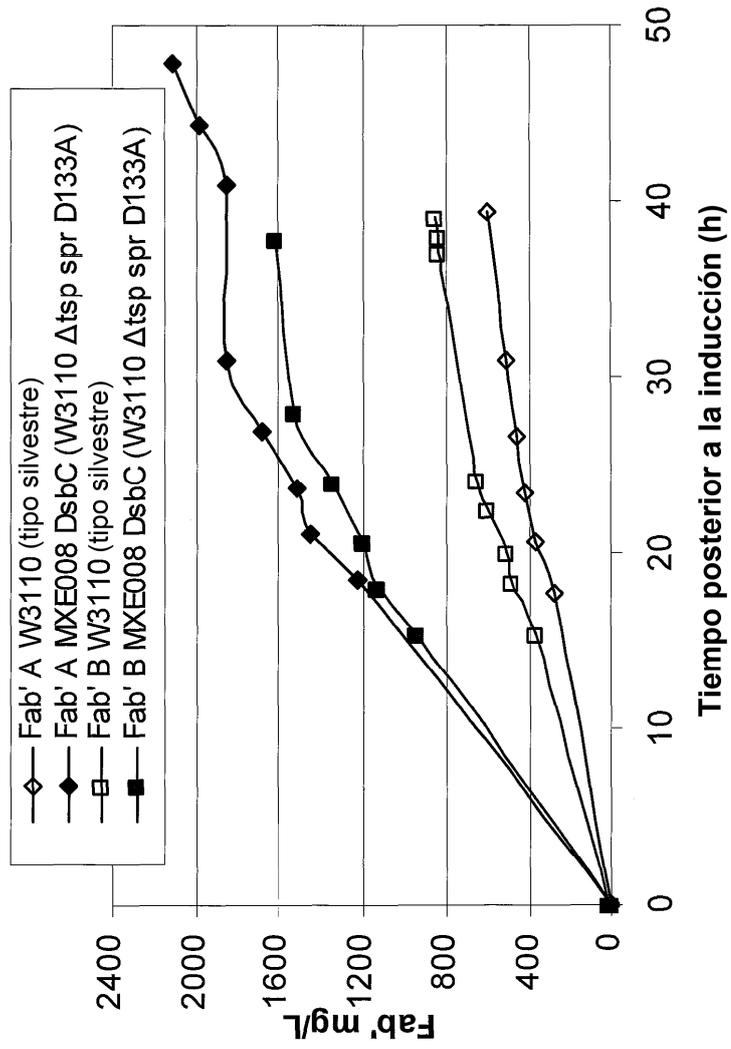


Figura 5a

**ptr de tipo silvestre (proteasa III) 5'.**

```

* M P R S T W F K A L L L L V
TGA ATG CCC CGC AGC ACC TGG TTC AAA GCA TTA TTG TTG TTA GTT

A L W A P L S
GCC CTT TGG GCA CCC TTA AGT
    
```

**Δ ptr mutada (proteasa III) 5'.**

```

EcoR I
~~~~~
* I P R S T W F K A L L L L V
TGA ATT CCC CGC AGC ACC TGG TTC AAA GCA TTA TTG TTG TTA GTT

Ase I
~~~~~
A L W A H * C
GCC CTT TGG GCA CAT TAA TGT
    
```

Figura 5b

**Tsp de tipo silvestre 5'.**

```

M N M F F R L T A L A G L L A
ATG AAC ATG TTT TTT AGG CTT ACC GCG TTA GCT GGC CTG CTT GCA

I A G Q T F A
ATA GCA GGC CAG ACC TTC GCT
    
```

**Δ Tsp mutada 5'.**

```

EcoR I
~~~~~
M N S F L G L P R * L A C L Q
ATG AAT TCG TTT TTA GGC TTA CCG CGT TAG CTG GCC TGC TTG CAA

Ase I
~~~~~
* Q A R H * L
TAG CAG GCC AGA CAT TAA TTG
    
```

Figura 5c

**DegP de tipo silvestre**

202 D A A I N R G N **S** G G  
949 GAT GCA GCG ATC AAC CGT GGT AAC **TCC** GGT GGT

**DegP mutada S210A**

*Ase I*  
~~~~~  
202 D A A I N R G N **A** G G  
949 GAT GCA GCG **ATT AAT** CGT GGT AAC **GCC** GGT GGT

Figura 6

