

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 539 931**

51 Int. Cl.:

C12N 15/70 (2006.01)

C07K 14/245 (2006.01)

C12N 9/52 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.01.2011 E 11701216 (1)**

97 Fecha y número de publicación de la concesión europea: **18.03.2015 EP 2524043**

54 Título: **Cepa hospedante bacteriana que comprende un gen spr mutante y que tiene actividad reducida de Tsp**

30 Prioridad:

14.01.2010 GB 201000587

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

07.07.2015

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:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 539 931 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 hospedante bacteriana que comprende un gen spr mutante y que tiene actividad reducida de Tsp

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

5 **Antecedentes de la invención**

Las células bacterianas, tales como *E. coli*, se usan habitualmente para producir proteínas recombinantes. Hay muchas ventajas en el uso de células bacterianas, tales como *E. coli*, para producir proteínas recombinantes, en particular debido a la naturaleza versátil de las células bacterianas como células hospedantes que permiten la inserción de genes mediante plásmidos. *E. coli* se ha usado para producir muchas proteínas recombinantes, incluyendo insulina humana.

A pesar de las muchas ventajas del uso de células bacterianas para producir proteínas recombinantes, todavía hay limitaciones significativas, que incluyen la dificultad de producir proteínas sensibles a proteasas. Las proteasas tienen una función importante en la renovación de proteínas viejas, dañadas o mal plegadas en el periplasma y citoplasma de *E. coli*. Las proteasas bacterianas actúan para degradar la proteína recombinante de interés, reduciendo así a menudo significativamente el rendimiento de la proteína activa.

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

La Tsp (también conocida como Prc) es una proteasa periplásmica de 60 kDa. El primer sustrato conocido de la Tsp fue la proteína de unión a penicilina 3 (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.* 1989 Nov;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 Aug;173 (15):4799-813), pero más tarde se descubrió que la Tsp también podía escindir las proteínas de la cola del fago, y por lo tanto, se renombró como proteasa específica de cola (Tsp, por sus siglas en 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) describe una cepa de eliminación de prc (KS1000) en donde la mutación se creó sustituyendo un segmento del gen prc por un fragmento que comprendía un marcador Kan^r.

La reducción de la actividad de la Tsp (prc) es deseable para reducir la proteólisis de proteínas de interés. Sin embargo, se ha encontrado que las células que carecen de la proteasa prc presentan crecimiento termosensible a osmolaridad baja. Hara et al., aislaron revertientes termorresistentes que contenían mutaciones supresoras (spr) extragénicas (Hara et al., *Microbial Drug Resistance*, 2: 63-72 (1996)). Spr es una proteasa periplásmica unida a membrana de 18 kDa y los sustratos de la spr son Tsp y peptidoglicanos en la membrana externa implicados en la hidrólisis de la pared celular durante la división celular. El gen de spr se denomina UniProtKB/SwissProt P0AFV4 (SPR_ECOLI).

Se han descrito cepas deficientes en proteasa mejoradas que comprenden el gen spr mutante. Chen et al. describen la construcción de cepas de *E. coli* que llevan diferentes combinaciones de mutaciones en prc (Tsp) y otra proteasa, DegP, creadas por amplificación de regiones en la dirección 5' y la dirección 3' del gen y ligándolas entre sí en un vector que comprende marcadores de selección y una mutación sprW174R ("High-level accumulation of a recombinant antibody fragment in the periplasm of *Escherichia coli* requires a triple-mutant (Δ DegP Δ prc sprW174R) host strain" (Chen C, Snedecor B, Nishihara JC, Joly JC, McFarland N, Andersen DC, Battersby JE, Champion KM. *Biotechnol Bioeng.* 2004 Mar 5;85(5):463-74). Se encontró que la combinación de las mutaciones Δ DegP, Δ prc y sprW174R proporcionaba los niveles mayores de cadena ligera de anticuerpo, cadena pesada de anticuerpo y F(ab')₂-LZ. El documento EP1341899 describe una cepa de *E. coli* que es deficiente en DegP y prc cromosómicos que codifican proteasas DegP y Prc, respectivamente, y albergan un gen spr mutante que codifica una proteína que suprime fenotipos de crecimiento presentados por cepas que albergan mutantes de prc.

La presente invención proporciona nuevas cepas bacterianas que llevan mutantes de spr alternativos que proporcionan medios ventajosos para producir proteínas recombinantes.

50 **Resumen de la invención**

En un primer aspecto de la presente invención, se proporciona una célula bacteriana Gram negativa recombinante que comprende un gen spr mutante que codifica una proteína spr que tiene una mutación en uno o más aminoácidos seleccionados de N31, R62, I70, Q73, C94, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144, H145, G147 y H157, y en donde la célula tiene actividad reducida de la proteína Tsp comparada con una célula de tipo natural.

En una realización, el genoma de la célula es isogénico con una célula bacteriana de tipo natural excepto por el gen spr mutado y la modificación requerida para reducir la actividad de la proteína Tsp comparada con una célula de tipo natural.

5 En un segundo aspecto, la presente invención proporciona una célula bacteriana Gram negativa recombinante que tiene actividad reducida de la proteína Tsp comparada con una célula de tipo natural y que comprende un gen spr mutante que codifica la proteína spr, en donde el genoma de la célula es isogénico con una célula bacteriana de tipo natural excepto por la modificación requerida para reducir la actividad de la proteína Tsp comparada con una célula de tipo natural y el gen spr mutado.

10 Las células proporcionadas por el primer y segundo aspectos de la presente invención muestra fenotipos de producción de proteínas y crecimiento ventajosos.

En un tercer aspecto, la presente invención proporciona un método para producir una proteína de interés que comprende expresar la proteína de interés en una célula bacteriana Gram negativa recombinante como se ha definido antes.

Breve descripción de los dibujos

15 La figura 1 muestra el perfil de crecimiento de cepas MXE008 y MXE009 que expresan Fab' anti-TNF α comparado con cepas W3110 de tipo natural y MXE001 que expresan Fab' anti-TNF α .

La figura 2 muestra la expresión de Fab' anti-TNF α de MXE008 y MXE009 comparado con W3110 de tipo natural y MXE001.

20 La figura 3 muestra el perfil de crecimiento de cepas MXE008 y MXE009 que expresan Fab' anti-TNF α comparado con cepas W3110 y MXE001de control que expresan Fab' anti-TNF α .

La figura 4 muestra el rendimiento de Fab' anti-TNF α del periplasma (símbolos sombreados) y el líquido sobrenadante (símbolos en blanco) de cepas de *E. coli* MXE008 y MXE009 comparado con W3110 y MXE001de control.

25 La figura 5 muestra el perfil de crecimiento de cepas MXE0012 y MXE017 que expresan Fab' anti-TNF α comparado con cepas W3110 y MXE001de tipo natural que expresan Fab' anti-TNF α .

La figura 6 muestra la expresión de Fab' anti-TNF α en MXE012 y MXE017 comparado con W3110 de tipo natural y MXE001.

La figura 7 muestra el perfil de crecimiento de W3110, MXE001 y MXE008 durante una fermentación de producción de Fab' anti-TNF α .

30 La figura 8 muestra la acumulación periplásmica de Fab' anti-TNF α (líneas y símbolos negros) y acumulación en el medio de Fab' anti-TNF α (líneas de trazos y símbolos blancos) para W3110, MXE001 (Δ tsp) y MXE008 (Δ tsp spr D112A) durante una fermentación de producción de Fab' anti-TNF α .

La figura 9 muestra los resultados de un ensayo de ADNbc de las cepas W3110, MXE001, MXE008 y MXE012.

La figura 10 muestra los resultados de un ensayo de proteínas de las cepas W3110, MXE001, MXE008 y MXE012.

35 La figura 11a muestra el extremo 5' de la proteína ptr (proteasa III) de tipo natural y ptr (proteasa III) mutada inactivada y secuencias de genes.

La figura 11b muestra el extremo 5' de la proteína Tsp de tipo natural y Tsp mutada inactivada y secuencias de genes.

La figura 11c muestra una región de la proteína DegP de tipo natural y DegP mutada y secuencias de genes.

40 Breve descripción de las secuencias

SEQ ID NO: 1 es la secuencia de ADN del gen Tsp de tipo natural que incluye 6 nucleótidos ATGAAC en la dirección 5' del codón de inicio.

SEQ ID NO: 2 es la secuencia de aminoácidos de la proteína Tsp de tipo natural.

45 SEQ ID NO: 3 es la secuencia de ADN del gen Tsp mutado inactivado que incluye los 6 nucleótidos ATGAAT en la dirección 5' del codón de inicio.

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

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

- SEQ ID NO: 6 es la secuencia de ADN de un gen de Proteasa III mutado inactivado.
- SEQ ID NO: 7 es la secuencia de ADN del gen DegP de tipo natural.
- SEQ ID NO: 8 es la secuencia de aminoácidos de la proteína DegP de tipo natural.
- SEQ ID NO: 9 es la secuencia de ADN de un gen DegP mutado.
- 5 SEQ ID NO: 10 es la secuencia de aminoácidos de la proteína DegP mutada.
- 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.
- 10 SEQ ID NO: 14 es la secuencia de aminoácidos de la cadena pesada de un anticuerpo anti-TNF.
- 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.
- 15 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.
- 20 SEQ ID NO: 19 es la secuencia del cebador oligonucleótido 5' para la región del gen DegP mutado que comprende el sitio de restricción Ase I.
- SEQ ID NO: 20 es la secuencia del cebador 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 natural que incluye la secuencia señal que son los primeros 26 restos de aminoácidos.
- 25 SEQ ID NO: 22 es la secuencia del gen spr no mutado sin la secuencia señal.
- 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.
- 30 SEQ ID NO: 25 es la secuencia de nucleótidos de una secuencia de OmpT mutada inactivada.
- SEQ ID NO: 26 muestra la secuencia de aminoácidos de CDRH1 de hTNF40.
- SEQ ID NO: 27 muestra la secuencia de aminoácidos de CDRH2 de hTNF40 que es una CDR híbrida en donde los 6 aminoácidos C-terminales son de la secuencia de CDR H2 de un anticuerpo de la línea germinal del subgrupo 3 humano y los cambios de aminoácidos de la secuencia que resulta de esta hibridación están subrayados como sigue: WINTYIGEPY YADSVKG.
- 35 SEQ ID NO: 28 muestra la secuencia de aminoácidos de CDRH3 de hTNF40.
- SEQ ID NO: 29 muestra la secuencia de aminoácidos de CDRL1 de hTNF40.
- SEQ ID NO: 30 muestra la secuencia de aminoácidos de CDRL2 de hTNF40.
- SEQ ID NO: 31 muestra la secuencia de aminoácidos de CDRL3 de hTNF40.
- 40 SEQ ID NO: 32 muestra la secuencia de aminoácidos de CDRH2 de hTNF40.
- SEQ ID NO: 33 muestra la secuencia del adaptador oligonucleótido OmpA.
- SEQ ID NO: 34 muestra el casete de oligonucleótido que codifica la secuencia intergénica 1 (IGS1) para la expresión de Fab en *E. coli*.

SEQ ID NO: 35 muestra el casete de oligonucleótido que codifica la secuencia intergénica 2 (IGS2) para la expresión de Fab en *E. coli*.

SEQ ID NO: 36 muestra el casete de oligonucleótido que codifica la secuencia intergénica 3 (IGS3) para la expresión de Fab en *E. coli*.

5 SEQ ID NO: 37 muestra el casete de oligonucleótido que codifica la secuencia intergénica 4 (IGS4) para la expresión de Fab en *E. coli*.

Descripción detallada de las realizaciones preferidas de la invención

10 Los autores de la presente invención han proporcionado una célula bacteriana Gram negativa recombinante adecuada para expresar una proteína de interés que comprende un gen spr mutado y la célula tiene actividad reducida de la proteína Tsp comparada con una célula de tipo natural.

En el primer aspecto de la invención, las nuevas mutaciones en el spr proporcionan nuevas cepas que tienen fenotipo de crecimiento celular mejorado comparado con células bacterianas de tipo natural y células que llevan un gen Tsp mutado.

15 Los autores de la presente invención han identificado nuevas mutaciones de spr que son capaces de suprimir el fenotipo de crecimiento de una célula que comprende un gen Tsp mutado. Los autores de la invención han encontrado sorprendentemente que las células que llevan el nuevo spr mutado y que tienen actividad reducida de Tsp, presentan mayor velocidad de crecimiento y mayor duración de la supervivencia celular comparado con una célula de tipo natural o una célula que comprende un gen Tsp mutado. Específicamente, las células que llevan las nuevas mutaciones de spr y que tienen actividad reducida de Tsp presentan fenotipo de lisis celular reducido comparado con células que llevan un gen Tsp mutado. Por consiguiente, las nuevas cepas reducen la pérdida de proteína de las células y permiten la acumulación periplásmica prolongada comparado con células que llevan un gen Tsp mutado.

20 Además, las células que llevan el nuevo spr mutante y que tienen actividad reducida de Tsp presentan mayor rendimiento de una proteína de interés comparado con una célula bacteriana de tipo natural o una célula que comprende un gen Tsp mutado. El rendimiento de proteínas mejorado puede ser rendimiento de proteínas periplásmico y/o rendimiento de proteínas en el líquido sobrenadante. En una realización, las células de la presente invención presentan rendimiento de proteínas periplásmico mejorado comparado con una célula que lleva un gen Tsp mutado, debido a la pérdida reducida de la célula. Las células bacterianas recombinantes pueden ser capaces de una mayor velocidad de producción de una proteína de interés, y por lo tanto, la misma cantidad de una proteína de interés se puede producir en un tiempo más corto comparado con una célula bacteriana de tipo natural o una célula que comprende un gen Tsp mutado. La mayor velocidad de producción de una proteína de interés puede ser especialmente significativa a lo largo del periodo de crecimiento inicial de la célula, por ejemplo, a lo largo de las primeras, 5, 10, 20 o 30 horas después de inducción de la expresión de proteína.

25 Las células de acuerdo con la presente invención, preferiblemente expresan un rendimiento máximo en el periplasma y/o en 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.

30 Las células proporcionadas por el primer y segundo aspectos de la presente invención tienen actividad reducida de proteasa Tsp comparado con una célula de tipo natural, lo cual puede reducir la proteólisis de una proteína de interés, en particular proteínas de interés que son proteolíticamente sensibles a la Tsp. Por lo tanto, las células proporcionadas por el primer y segundo aspectos de la presente invención pueden proporcionar rendimientos mayores de proteínas intactas, preferiblemente de la proteína de interés, y un rendimiento menor, o preferiblemente nada, de fragmentos proteolíticos de proteína, preferiblemente de la proteína de interés, comparado con una célula bacteriana de tipo natural.

35 En el segundo aspecto de la invención y una realización preferida del primer aspecto de la invención, las células llevan solo mutaciones mínimas en el genoma necesarias para introducir una o más mutaciones de spr y la modificación necesaria para reducir la actividad de la proteína Tsp comparado con una célula de tipo natural. La célula bacteriana solo difiere de una célula bacteriana de tipo natural por la una o más mutaciones del gen spr y la modificación necesaria para reducir la actividad de la proteína Tsp comparado con una célula de tipo natural. Las células no llevan ninguna otra mutación que pueda tener efectos perjudiciales en el crecimiento celular y/o la capacidad para expresar una proteína de interés. Por consiguiente, una o más de las células hospedantes recombinantes de la presente invención pueden presentar características de expresión de proteínas mejorada y/o de crecimiento mejorado comparado con células que comprenden mutaciones adicionales genéticamente diseñadas en la secuencia genómica. Las células proporcionadas por la presente invención también son más adecuadas para usar para producir proteína terapéuticas comparadas con células que comprenden además alteraciones del genoma celular.

55 El experto en la técnica podrá ensayar fácilmente un clon de célula candidata para ver si tiene el rendimiento deseado de una proteína de interés, usando métodos bien conocidos en la técnica, incluyendo un método de

5 fermentación, ELISA, HPLC de 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". Ayuda no gubernamental a la investigación, EE.UU., *Journal of Biotechnology*. 135(4):358-65, 31 de julio de 2008; Champion KM. Nishihara JC. Joly JC. Amott 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. Riesenberger 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 revista. Ayuda no gubernamental a la investigación, EE.UU., *Applied Microbiology & Biotechnology*. 46(5-6):524-32, diciembre1996. El experto en la técnica también podrá ensayar fácilmente la proteína secretada para ver si la proteína está plegada correctamente usando métodos bien conocidos en la técnica, tales como HPLC de proteína G, difracción circular, RMN, cristalografía de rayos X y métodos de medición de afinidad de epítopos.

La presente invención ahora se describirá con más detalle. Todas las realizaciones descritas en la presente memoria se refieren al primer, segundo y tercer aspectos de la presente invención salvo que se exponga específicamente de otra forma.

20 Los términos "proteína" y "polipéptido" se usan de forma intercambiable en la presente memoria, salvo que el contexto indique otra cosa. "Péptido" se refiere a 10 o menos aminoácidos.

El término "polinucleótido" incluye un gen, ADN, ADNc, ARN, ARNm etc., salvo que el contexto indique otra cosa.

Como se usa en la presente memoria, el término "comprende" en el contexto de la presente memoria descriptiva debe interpretarse como "incluye".

25 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 llevar la actividad reducida de proteína Tsp y para llevar el gen spr mutante. Por ejemplo, una célula no mutada será una célula de tipo natural y se puede obtener de la misma población de células hospedantes que las células de la invención antes de modificación para introducir cualquier mutación.

La expresión "célula", "línea celular", "célula de cultivo" y "cepa" se usan de forma intercambiable.

30 La expresión "fenotipo de una célula que comprende un gen Tsp mutado" en el contexto de la presente invención, significa el fenotipo presentado por una célula que alberga un gen Tsp mutante. Típicamente, las células que comprenden un gen Tsp mutante pueden producir lisis, en especial con altas densidades celulares. La lisis de estas células hace que cualquier proteína recombinante se pierda al líquido sobrenadante. Las células que llevan el gen Tsp mutado también pueden mostrar crecimiento termosensible a baja osmolalidad. Por ejemplo, las células no presentan o presentan velocidad reducida de crecimiento o las células mueren en medio hipotónico a una temperatura alta, tal como 40°C o más.

40 El término "isogénico" en el contexto de la presente invención significa que el genoma de la célula de la presente invención tiene sustancialmente la misma o la misma secuencia genómica comparada con una célula de tipo natural de la cual deriva la célula, excepto por el gen spr mutado y la modificación necesaria para reducir la actividad de proteína Tsp comparado con una célula de tipo natural. En esta realización, el genoma de la célula no comprende mutaciones adicionales que no se encuentran de forma natural o de ingeniería genética. En una realización, la célula de acuerdo con la presente invención puede tener sustancialmente la misma secuencia genómica compara con la célula de tipo natural excepto por el gen spr mutado y la modificación necesaria para reducir la actividad de proteína Tsp comparado con una célula de tipo natural teniendo en cuenta cualquier mutación natural que pueda haber. En una realización, la célula de acuerdo con la presente invención puede tener exactamente la misma secuencia genómica comparada con la célula de tipo natural excepto por el gen spr mutado y la modificación necesaria para reducir la actividad de proteína Tsp comparado con una célula de tipo natural.

50 La expresión "de tipo natural" en el contexto de la presente invención, significa una cepa de una célula bacteriana Gram negativa como puede encontrarse en la naturaleza o se puede aislar del entorno, que no lleva ninguna mutación genéticamente diseñada. Un ejemplo de una cepa de tipo natural de *E. coli* es W3110, tal como la cepa K-12 de W3110.

55 Se puede usar cualquier bacteria Gram negativa adecuada como célula parental para producir la célula recombinante de la presente invención, Las bacterias Gram negativas adecuadas incluyen *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*. Se puede usar cualquier cepa de *E. coli* en la presente invención, pero preferiblemente se usa una cepa W3110 de tipo natural, tal como K-12 W3110.

Un inconveniente asociado con las cepas bacterianas deficientes en proteasa previamente creadas y usadas para expresar proteínas recombinantes, es que comprenden mutaciones adicionales de genes implicados en el metabolismo celular y la replicación de ADN, tales como, por ejemplo, *phoA*, *thiA*, *lac*, *rec*, *gal*, *ara*, *arg*, *thi* y *pro* en cepas de *E. coli*. Estas mutaciones pueden tener muchos efectos perjudiciales en la célula hospedante, incluyendo efectos en el crecimiento celular, estabilidad, rendimiento de la expresión de la proteína recombinante y toxicidad. Las cepas que tienen una o más de estas mutaciones genómicas, en particular las cepas que tienen un número alto de estas mutaciones pueden presentar una pérdida de competencia que reduce la velocidad de crecimiento bacteriano a un nivel que no es adecuado para la producción industrial de proteínas. Además, cualquiera de las mutaciones genómicas anteriores puede afectar a otros genes en *cis* y/o en *trans* de formas dañinas impredecibles, alternado de esta forma el fenotipo de la cepa, capacidad y perfil de proteínas. Además, el uso de células muy mutadas en general no es adecuado para producir proteínas recombinantes para uso comercial, en particular para productos terapéuticos, porque dichas cepas en general tienen rutas metabólicas defectuosas y por lo tanto pueden crecer poco o nada en medios mínimos o químicamente definidos.

La célula de acuerdo con el segundo aspecto de la presente invención es isogénica con una célula bacteriana de tipo natural excepto por el gen *spr* mutado y la modificación necesaria para reducir la actividad de proteína Tsp comparado con una célula de tipo natural. La célula de acuerdo con el primer aspecto de la presente invención preferiblemente también es isogénica con una célula bacteriana de tipo natural excepto por el gen *spr* mutado y la modificación necesaria para reducir la actividad de proteína Tsp comparado con una célula de tipo natural. Solo se hacen mutaciones mínimas en el genoma de la célula para introducir las mutaciones. Las células no llevan ninguna otra mutación que pueda tener efectos perjudiciales en el crecimiento celular y/o en la capacidad para expresar una proteína de interés. Por consiguiente, una o más de las células hospedantes recombinantes de la presente invención pueden presentar expresión de proteína mejorada y/o características de crecimiento mejoradas comparadas con células que comprenden además mutaciones genéticamente diseñadas de la secuencia genómica. Las células proporcionadas por la presente invención también son más adecuadas para usar en la producción de proteínas terapéuticas comparadas con células que comprenden alteraciones adicionales del genoma celular.

En una realización preferida, la célula es isogénica con una célula de *E. coli* de tipo natural, tal como la cepa W3110, excepto por el gen *spr* mutado y la modificación necesaria para reducir la actividad de proteína Tsp comparado con una célula de tipo natural.

La célula de la presente invención puede diferir además de una célula de tipo natural comprendiendo un polinucleótido que codifica una proteína de interés. La secuencia del polinucleótido que codifica la proteína de interés puede ser exógena o endógena. El polinucleótido que codifica la proteína de interés puede estar contenido en un vector de expresión adecuado transformado en la célula y/o integrado en el genoma de la célula hospedante. En la realización donde el polinucleótido que codifica la proteína de interés se inserta en el genoma hospedante, la célula de la presente invención también diferirá de una célula de tipo natural por la secuencia de polinucleótido que codifica la proteína de interés insertada. Preferiblemente, el polinucleótido es un vector de expresión en la célula, produciendo así alteración mínima en el genoma de la célula hospedante.

La proteína *spr* es una proteína periplásmica unida a membrana de *E. coli*.

La secuencia de aminoácidos de tipo natural de la proteína *spr* se muestra en la SEQ ID NO: 21 con la secuencia señal en el extremo N y en la SEQ ID NO: 22 sin la secuencia señal de 26 aminoácidos (de acuerdo con el número de acceso en UniProt P0AFV4). La numeración de 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 la SEQ ID NO: 21.

El gen *spr* mutado preferiblemente es 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 llevan un gen Tsp mutado pueden tener una buena velocidad de crecimiento celular pero una limitación de estas células es su tendencia a la lisis, en especial con densidades celulares altas. Por consiguiente, el fenotipo de una célula que comprende un gen Tsp mutado es una tendencia a la lisis, en especial con densidades celulares altas. Las células que llevan un gen Tsp mutado también muestran crecimiento termosensible con baja osmolalidad. Sin embargo, las mutaciones *spr* que llevan las células de la presente invención, cuando se introducen en una célula que tiene actividad reducida de Tsp suprimen el fenotipo de Tsp reducido y, por lo tanto, la célula presenta lisis reducida, en particular con una densidad celular alta. El fenotipo de crecimiento de una célula lo puede medir fácilmente un experto en la técnica durante la técnica de fermentación en matraz agitado o de densidad celular alta. La supresión del fenotipo de lisis celular también puede verse a partir la velocidad de crecimiento y/o producción de proteína recombinante mejorados, en particular en el periplasma, presentados por una célula que lleva el mutante de *spr* y que tiene actividad reducida de Tsp comparado con una célula que lleva el mutante de Tsp y un *spr* de tipo natural.

Las células de acuerdo con el primer aspecto de la presente invención y una realización preferida del segundo aspecto, comprenden un gen *spr* mutante que codifica una proteína *spr* que tiene una mutación en uno o más aminoácidos seleccionados de N31, R62, I70, Q73, C94, S95, V98, Q99, R100, L108, Y115, D133, V135, L136,

G140, R144, H145, G147 y H157, preferiblemente una mutación en uno o más aminoácidos seleccionados de C94, S95, V98, Y115, D133, V135, H145, G147 y H157.

- 5 La mutación de uno o más de los aminoácidos anteriores puede ser cualquier mutación de aminoácido adecuada en 1, 2 o 3 de los nucleótidos que codifican el aminoácido. La mutación cambia el resto de aminoácido en cualquier aminoácido adecuado que de como resultado una proteína spr mutada capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado. La mutación de aminoácido puede cambiar el aminoácido por uno que tenga un tamaño diferente y/o tenga propiedades químicas diferentes comparadas con el aminoácido de tipo natural.

En una realización, el gen spr mutante codifica una proteína spr que tiene una o más mutaciones seleccionadas de C94A, S95F, V98E, Y115F, D133A, V135D o G, H145A, G147C y H157A.

- 10 En una realización, la mutación es en 1, 2 o 3 de la triada catalítica de los restos de aminoácidos 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:

- una mutación de C94; o
- 15 • una mutación de H145; o
- una mutación de H157; o
- una mutación de C94 y H145; o
- una mutación de C94 y H157; o
- una mutación de H145 y H157; o
- 20 • una mutación de C94, H145 y H157.

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

- 25 Uno, dos o tres de C94, H145 y H157 pueden estar mutados en cualquier aminoácido adecuado que de como resultado una proteína spr capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado. Por ejemplo, 1, 2 o 3 de C94, H145 y H157 pueden estar mutados en un aminoácido pequeño tal como Gly o Ala. Por consiguiente, la proteína spr puede tener 1, 2 o 3 de las mutaciones C94, H145 y H157. Preferiblemente, el gen spr comprende la mutación de aminoácido H145A, que se ha encontrado que produce una proteína spr capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado.

- 30 La designación de 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 natural. El número se refiere a la posición del aminoácido donde se está haciendo la sustitución de aminoácido, y la segunda letra designa el aminoácido que se usa para sustituir el aminoácido de tipo natural.

- 35 En una realización preferida, la proteína spr mutante comprende una mutación en uno o más aminoácidos seleccionados de N31, R62, I70, Q73, S95, V98, Q99, R100, L108, Y115, D133, V135, L136, G140, R144 y G147, preferiblemente una mutación en uno o más aminoácidos seleccionados 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 de N31; o
- una mutación de R62; o
- una mutación de I70; o
- 40 • una mutación de Q73; o
- una mutación de S95; o
- una mutación de V98; o
- una mutación de Q99; o
- una mutación de R100; o
- 45 • una mutación de L108; o

- una mutación de Y115; o
- una mutación de D133; o
- una mutación de V 135; o
- una mutación de L136; o
- 5 • una mutación de G140; o
- una mutación de R144; o
- una mutación de G147.

En una realización la proteína spr mutante comprende múltiples mutaciones de los aminoácidos:

- S95 y Y115; o
- 10 • N31, Q73, R100 y G140 ; o
- Q73, R100 y G140; o
- R100 y G140; o
- Q73 y G140; o
- Q73 y R100; o
- 15 • R62, Q99 y R144; o
- Q99 y R144.

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

En una realización preferida la proteína spr comprende una o más 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 más de las siguientes mutaciones: S95F, V98E, Y115F, D133A, V135D o V135G y G147C. En esta realización, preferiblemente la proteína spr no tiene ninguna mutación adicional.

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

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

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

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

40 En el segundo aspecto de la presente invención, se puede hacer cualquier mutación o mutaciones en el gen spr que den como resultado una proteína spr capaz de suprimir el fenotipo de una célula que comprende un gen Tsp mutado. Preferiblemente, la proteína spr puede tener una o más de las siguientes mutaciones N31Y, R62C, I70T,

Q73R, C94A, S95F, V98E, Q99P, R100G, L108S, Y115F, D133A, V135D, V135G, L136P, G140C, R144C, H145A, G147C, H157A y W174R. En una realización, la proteína spr no comprende la mutación W174R. Preferiblemente, el gen spr comprende una o más mutaciones descritas antes con respecto al primer aspecto de la presente invención.

5 Las células de acuerdo con la presente invención tienen actividad reducida de proteína Tsp comparada con una célula de tipo natural. La expresión "actividad reducida de proteína Tsp comparada con una célula de tipo natural" significa que la actividad de Tsp en la célula es menor comparada con la actividad de Tsp en una célula de tipo natural. La célula se puede modificar por cualquier medio adecuado para reducir la actividad de Tsp.

10 En una realización, la actividad reducida de Tsp es a partir de la modificación del polinucleótido endógeno que codifica la Tsp y/o secuencias reguladoras de expresión asociadas. La modificación puede reducir o detener la transcripción y traducción del gen Tsp o puede proporcionar una proteína Tsp expresada que tiene actividad reducida de proteasa compara con la proteína Tsp de tipo natural.

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

15 En una realización preferida, las células de acuerdo con la presente invención llevan un gen Tsp mutado que codifica una proteína Tsp que tiene actividad reducida de proteasa o un gen Tsp mutado inactivado.

Preferiblemente, el gen Tsp cromosómico está mutado.

20 Como se usa en la presente memoria, el "gen Tsp" significa un gen que codifica la proteasa Tsp (también conocida como Prc) que es una proteasa periplásmica capaz de actuar en la proteína 3 de unión a penicilina (PBP3) y proteínas de la cola de fago. La secuencia de un gen Tsp de tipo natural se muestra en la SEQ ID NO: 1 y la secuencia de la proteína Tsp de tipo natural se muestra en la SEQ ID NO: 2.

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

25 La expresión "gen Tsp mutado que codifica una proteína Tsp que tiene actividad reducida de proteasa" en el contexto de la presente invención significa que el gen Tsp mutado no tiene la actividad de proteasa completa comparado con el gen de Tsp no mutado de tipo natural.

30 Preferiblemente, 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 de proteasa de una proteína Tsp no mutada de tipo natural. Más preferiblemente, el gen Tsp mutado codifica una proteína Tsp que no tiene actividad de proteasa. En esta realización, la célula no es deficiente en Tsp cromosómico, es decir, la secuencia del gen Tsp no se ha eliminado o mutado para prevenir la expresión de cualquier forma de la proteína Tsp.

35 Se puede introducir cualquier mutación adecuada en el gen Tsp con el fin de producir una proteína que tenga actividad reducida de proteasa. La actividad de proteasa de una proteína Tsp expresada a partir de una bacteria Gram-negativa la puede ensayar fácilmente el experto en la técnica por 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, No. 48, presentado el 1 de diciembre, pág. 28864-28868, 1995 Kenneth C. Keiler y Robert T. Sauer) en donde se ensayó la actividad de proteasa de Tsp.

40 Keiler et al. (véase antes) han descrito que la Tsp tiene un sitio activo que comprende los restos S430, D441 y K455 y los restos G375, G376, E433 y T452 son importantes para mantener la estructura de la Tsp. Keiler et al (véase antes) describen que han encontrado que los genes de Tsp mutados S430A, D441A, K455A, K455H, K455R, G375A, G376A, E433A y T452A no tienen actividad de proteasa detectable. Se describe además que el gen Tsp mutado S430C presentaba aproximadamente 5-10% de la actividad natural. Por consiguiente, la mutación de Tsp para producir una proteína que tenga actividad de proteasa reducida puede comprender una mutación, tal como una mutación de aminoácido en uno o más de los restos S430, D441, K455, G375, G376, E433 y T452. Preferiblemente, la mutación de Tsp para producir una proteína que tenga actividad reducida de proteasa puede comprender una mutación, tal como una mutación de aminoácido en 1, 2 o 3 de los restos del sitio activo S430, D441 y K455.

Por consiguiente, el gen Tsp mutado puede comprender:

- una mutación de S430; o
- una mutación de D441; o
- 50 • una mutación de K455; o
- una mutación de S430 y D441; o
- una mutación de S430 y K455; o

- una mutación de D441 y K455; o
- una mutación de S430, D441 y K455.

Uno o más de S430, D441, K455, G375, G376, E433 y T452 se pueden mutar en cualquier aminoácido adecuado que de como resultado una proteína que tenga actividad de proteasa reducida. Los ejemplos de mutaciones adecuadas son S430A, S430C, D441A, K455A, K455H, K455R, G375A, G376A, E433A y T452A. El gen Tsp mutado puede comprender 1, 2 o 3 mutaciones en los restos del sitio activo, por ejemplo, el gen puede comprender:

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

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

La expresión "gen Tsp mutado inactivado" en el contexto de la presente invención significa que el gen comprende una o más mutaciones que previenen la expresión de la proteína Tsp codificada por el gen de tipo natural, para proporcionar una célula deficiente en proteína Tsp. El gen inactivado puede ser parcial o completamente transcrito pero no traducido en la proteína codificada. El gen Tsp mutado inactivado se puede mutar de cualquier forma adecuada, por ejemplo por una o más mutaciones por eliminación, inserción, puntual, de aminoácido, terminadora y de cambio del marco de lectura, para que no se produzca expresión de la proteína. Por ejemplo, el gen se puede inactivar por inserción de una secuencia de ADN extraña, tal como un marcador de resistencia a antibiótico, en la secuencia que codifica el gen.

En una realización preferida, el gen Tsp no se muta por inserción de una secuencia de ADN extraña, tal como un marcador de resistencia a antibiótico, en la secuencia que codifica el gen. En esta realización, el gen Tsp puede comprender una mutación en el codón de inicio del gen y/o uno o más codones de parada situados en la dirección 3' del codón de inicio del gen y en la dirección 5' del codón de parada del gen, de modo que se previene la expresión de la proteína Tsp. La mutación en el codón de inicio puede ser una mutación de aminoácido de 1, 2 o 3 de los nucleótidos del codón de inicio. Alternativa o adicionalmente, el codón de inicio se puede mutar por una mutación de cambio del marco de lectura por inserción o eliminación. El gen Tsp comprende dos codones ATG en el extremo 5' de la secuencia codificante, uno o ambos codones ATG se pueden mutar por una mutación de aminoácido. El gen Tsp se puede mutar en el segundo codón ATG (codón 3) en TCG, como se muestra en la figura 11b. El gen Tsp puede comprender alternativa o adicionalmente uno o más codones de parada situados en la dirección 3' del codón de inicio del gen y en la dirección 5' del codón de parada del gen. Preferiblemente el gen Tsp mutado inactivado comprende tanto una mutación de aminoácido en el codón de inicio como uno o más codones de parada insertados. En una realización preferida, el gen Tsp se muta para eliminar "T" del quinto codón produciendo de esta forma un cambio del marco de lectura que da como resultado codones de parada en los codones 11 y 16, como se muestra en la figura 11b. 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 el marco de lectura en el codón 21, como se muestra en la figura 11b.

En una realización preferida, el gen Tsp mutado inactivado tiene la secuencia de ADN de la SEQ ID NO: 3, que incluye 6 nucleótidos ATGAAT en la dirección 5' del codón de inicio. Las mutaciones que se han hecho en la secuencia de Tsp mutado inactivado de la SEQ ID NO: 3 se muestran en la figura 11b. En una realización, el gen Tsp mutado tiene la secuencia de ADN de los nucleótidos 7 a 2048 de la SEQ ID NO: 3.

En una realización, la célula bacteriana Gram negativa recombinante comprende además un polinucleótido recombinante que codifica DsbC. El polinucleótido 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 hospedante. En la realización donde el polinucleótido que codifica DsbC se inserta en el genoma del hospedante, la célula de la presente invención también diferirá de la célula de tipo natural 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 produciendo así una alteración mínima en el genoma de la célula hospedante.

Como se usa en la presente memoria, un "polipéptido recombinante" se refiere a una proteína que se construye o produce usando tecnología de ADN recombinante. La secuencia del polinucleótido que codifica DsbC puede ser idéntica a la secuencia endógena que codifica DsbC encontrada en células bacterianas. Alternativamente, la secuencia del polinucleótido recombinante que codifica DsbC es una versión mutada de la secuencia de DsbC de tipo natural, por ejemplo, que tiene un sitio de restricción eliminado, tal como un sitio de EcoRI, y/o una secuencia que codifica un marcador de his.

La DsbC es una proteína procariota encontrada en el periplasma de *E. coli* que cataliza la formación de enlaces disulfuro en *E. coli*. DsbC tiene un secuencia de aminoácidos de 236 de longitud (incluyendo el péptido señal) y un peso molecular de 25,6 kDa (UniProt No. POAEG6). DsbC se identificó por primera vez 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, pág. 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, pág. 2007-2012, 1994).

5 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 chaperona y actividad reducida de proteasa y/o un gen ptr mutado, en donde el gen ptr mutado codifica una proteína proteasa III que tiene actividad reducida de proteasa 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 reducida de proteasa o es un gen OmpT mutado inactivado.

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

10 a. un gen spr mutado

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

15 c. un gen DegP mutado que codifica una proteína DegP que tiene actividad de chaperona y actividad reducida de proteasa y/o un OmpT mutado en donde el gen OmpT mutado codifica una proteína OmpT que tiene actividad reducida de proteasa o es un gen OmpT mutado inactivado.

Preferiblemente, en esta realización el genoma de la célula es isogénico con una célula bacteriana de tipo natural excepto por las mutaciones anteriores.

En una realización, la presente invención proporciona una célula bacteriana Gram negativa recombinante que comprende:

20 a. un gen spr mutado

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

25 c. un gen ptr mutado, en donde el gen ptr mutado codifica una proteína proteasa III que tiene actividad reducida de proteasa o es un gen ptr mutado inactivado y/o un OmpT mutado en donde el gen OmpT mutado codifica una proteína OmpT que tiene actividad reducida de proteasa o es un gen OmpT mutado inactivado.

Preferiblemente, en esta realización el genoma de la célula es isogénico con una célula bacteriana de tipo natural excepto por las mutaciones anteriores.

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

a. un gen spr mutado;

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

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

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

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

Preferiblemente, en esta realización el genoma de la célula es isogénico con una célula bacteriana de tipo natural excepto por las mutaciones anteriores.

40 En una realización de la presente invención, la célula lleva un gen DegP mutado. Como se usa en la presente memoria, "DegP" significa un gen que codifica la proteína DegP (también conocida como HtrA), que tiene función doble como una chaperona y como una proteasa ("Families of serine peptidases"; Rawlings ND, Barrett A, J. *Methods Enzymol.* 1994;244:19-61). La secuencia del gen DegP no mutado se muestra en la SEQ ID NO: 7 y la secuencia de la proteína DegP no mutada se muestra en la SEQ ID NO: 8.

45 A bajas temperaturas la DegP funciona como una chaperona y a altas temperaturas la DegP tiene preferencia para funcionar como una proteasa ("A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein". *Cell*, Volume 97, Issue 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 donde la célula comprende la mutación de DegP, la mutación de DegP en la célula proporciona un gen DegP mutado que codifica una proteína DegP que tiene actividad de chaperona pero no actividad de proteasa completa.

5 La expresión "que tiene actividad de chaperona" en el contexto de la presente invención significa que la proteína DegP mutada tiene la misma o sustancialmente la misma actividad de chaperona comparada con la proteína DegP no mutada de tipo natural. Preferiblemente, el gen DegP mutado que 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 chaperona de una proteína DegP no mutada de tipo natural. Más preferiblemente, el gen DegP mutado codifica una proteína DegP que tiene la misma actividad de chaperona comparada con la DegP de tipo natural.

10 La expresión "que tiene actividad reducida de proteasa" en el contexto de la presente invención significa que la proteína DegP mutada no tiene la actividad de proteasa completa comparada con la proteína DegP no mutada de tipo natural. 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 de proteína de una proteína DegP no mutada de tipo natural. Más preferiblemente, el gen DegP mutado codifica una proteína DegP que no tiene actividad de proteasa. La célula no es deficiente en DegP cromosómico, es decir, las secuencias del gen DegP no se han eliminado o mutado para prevenir la expresión de cualquier forma de proteína DegP.

20 Se puede introducir cualquier mutación adecuada en el gen DegP con el fin de producir una proteína que tenga actividad de chaperona y actividad reducida de proteasa. La actividad de proteasa y chaperona de una proteína DegP expresada en una bacteria Gram negativa la puede ensayar fácilmente el experto en la técnica por cualquier método adecuado tal como el método descrito por Spiess et al., en donde se ensayan las actividades de proteasa y chaperona de la DegP en MaIS, un sustrato natural de la DegP ("A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein". *Cell*, Volume 97, Issue 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.

25 La DegP es una serina proteasa y tiene un centro activo que consiste en una triada de restos de aminoácidos catalítica de His105, Asp135 y Ser210 ("Families of serine peptidases", *Methods Enzymol.*, 1994, 244:19-61 Rawlings N y Barrett A). La mutación de DegP para producir una proteína que tiene actividad de chaperona y actividad reducida de proteasa puede comprender una mutación, tal como una mutación de aminoácido en 1, 2 o 3 de His105, Asp135 y Ser210.

30 Por consiguiente, el gen DegP mutado puede comprender:

- una mutación de His105; o
- una mutación de Asp 135; o
- una mutación de Ser210; o
- una mutación de His105 y Asp 135; o
- 35 • una mutación de His105 y Ser210; o
- una mutación de Asp135 y Ser210; o
- una mutación de His105, Asp135 y Ser210.

40 Se pueden mutar 1, 2 o 3 de His105, Asp135 y Ser210 en cualquier aminoácido adecuado que de como resultado una proteína que tenga actividad de chaperona y actividad reducida de proteasa. Por ejemplo, 1, 2 o 3 de His105, Asp135 y Ser210 se pueden mutar en un aminoácido pequeño tal como Gly o Ala. Una mutación adecuada adicional es cambiar 1, 2 o 3 de His105, Asp135 y Ser210 por un aminoácido que tenga propiedades opuestas tal como mutar la Asp135 en Lys o Arg, mutar la His105 polar en un aminoácido no polar tal como Gly, Ala, Val o Leu, y mutar la Ser210 hidrófila pequeña en un resto 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 11c, que se ha encontrado que produce una

45 proteína que tiene actividad de chaperona pero no actividad de proteasa ("A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein". *Cell*, Volume 97, Issue 3, páginas 339 - 347. Spiess C, Beil A, Ehrmann M).

50 DegP tiene dos dominios PDZ, PDZ1 (restos 260-358) y PDZ2 (restos 359-448), que median la interacción proteína-proteína ("A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein". *Cell*, Volume 97, Issue 3, páginas 339 - 347. Spiess C, Beil A, Ehrmann M). En una realización de la presente invención, el gen degP se muta para eliminar el dominio PDZ1 y/o el dominio PDZ2. La eliminación de PDZ1 y PDZ2 da como resultado la pérdida completa de actividad de proteasa de la proteína DegP y menor actividad de chaperona comparada con la proteína degP de tipo natural, mientras que la eliminación de PDZ1 o PDZ2 da como resultado una actividad de proteasa de 5% y actividad de chaperona similar comparada con la

proteína DegP de tipo natural ("A Temperature-Dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein". *Cell*, Volume 97, Issue 3, páginas 339 - 347. Spiess C, Beil A, Ehrmann M).

El gen DegP mutado también puede comprender un sitio de restricción silencioso que no se encuentra de forma natural, tal como Ase I con el fin de ayudar en los métodos de identificación y cribado, por ejemplo, como se muestra en la figura 11c.

La secuencia preferida del gen DegP mutado que comprende la mutación puntual S210A y un sitio de restricción Ase I marcador se proporciona en la SEQ ID NO: 9, y la secuencia de la proteína codificada se muestra en la SEQ ID NO: 10. Las mutaciones que se han hecho en la secuencia de DegP mutada de la SEQ ID NO: 9 se muestran en la figura 11c.

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 chaperona y actividad reducida de proteasa, una o más de las células proporcionadas por la presente invención pueden proporcionar rendimiento mejorado de proteínas plegadas correctamente de la célula con respecto a células mutadas en donde el gen DegP se ha mutado en DegP inactivado que previene la expresión de DegP, tal como DegP deficiente cromosómico. En una célula que comprende un gen DegP mutado inactivado que previene la expresión de DegP, la actividad de chaperona de la DegP se pierde completamente, mientras que en la célula de acuerdo con la presente invención, se retiene la actividad de chaperona de la DegP mientras que se pierde la actividad completa de proteasa. En estas realizaciones, una o más células de acuerdo con la presente invención tienen una actividad de proteasa menor para prevenir la proteólisis de la proteína mientras que mantienen la actividad de chaperona para permitir el plegado correcto y el transporte de la proteína en la célula hospedante.

El experto en la técnica podrá ensayar fácilmente la proteína secretada para ver si la proteína está correctamente plegada, usando métodos bien conocidos en la técnica, tal como HPLC de proteína G, diroísmo circular, RMN, cristalografía de rayos X y métodos de medición de afinidad de epítomos.

En estas realizaciones, una o más células de acuerdo con la presente invención pueden tener crecimiento celular mejorado comparado con células que llevan un gen DegP mutado inactivado que previene la expresión de la DegP. Sin querer estar unido por la teoría, se puede presentar crecimiento celular mejorado debido a la actividad de chaperona que retiene la proteasa DegP que puede aumentar la capacidad de la célula para procesar todas las proteínas que requieren actividad de chaperona. Por lo tanto, se puede aumentar la producción de proteínas correctamente plegadas necesarias para el crecimiento celular y la reproducción en una o más células de la presente invención comparada con células que llevan la mutación de inactivación de DegP, mejorando de esta forma las rutas celulares que regulan el crecimiento. Además, las cepas deficientes en la proteasa DegP conocidas en general son sensibles a la temperatura y típicamente no crecen a temperaturas mayores que 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 usan típicamente en la producción a escala industrial de proteínas a partir de bacterias.

En una realización de la presente invención, la célula lleva un gen ptr mutado. Como se usa en la presente memoria, "gen ptr" significa un gen que codifica la proteasa III, una proteasa que degrada proteínas de alto peso molecular. La secuencia del gen ptr no mutado se muestra en la SEQ ID NO: 4 y la secuencia de la proteína proteasa III no mutada se muestra en la SEQ ID NO: 5.

La referencia al gen ptr mutado o gen ptr mutado que codifica la proteasa III, se refiere a un gen ptr mutado que codifica una proteína proteasa III que tiene actividad reducida de proteasa o un gen ptr mutado inactivado, salvo que se indique otra cosa.

La expresión "gen ptr mutado que codifica una proteína proteasa III que tiene actividad reducida de proteasa" en el contexto de la presente invención, significa que el gen ptr mutado no tiene actividad de proteasa completa comparado con el gen ptr no mutado de tipo natural.

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 de proteasa de una proteína proteasa III no mutada de tipo natural. Más preferiblemente, el gen ptr mutado codifica una proteína proteasa III que no tiene actividad de proteasa. En esta realización, la célula no es deficiente en ptr cromosómica, es decir, la secuencia del gen ptr no se ha eliminado o mutado para prevenir la expresión de cualquier forma de la proteína proteasa III.

Se puede introducir cualquier mutación adecuada en el gen ptr con el fin de producir una proteína proteasa III que tenga actividad reducida de proteasa. La actividad de proteasa de una proteína proteasa III expresada en una bacteria Gram negativa la puede ensayar fácilmente un experto en la técnica, mediante cualquier método adecuado en la técnica.

La expresión "gen ptr mutado inactivado" en el contexto de la presente invención significa que el gen comprende una o más mutaciones haciendo así que no haya expresión de la proteína codificada por el gen para proporcionar una

célula deficiente en la proteína codificada por el gen mutado inactivado. El gen inactivado puede ser parcial o completamente transcrito pero no traducido en la proteína codificada. El gen ptr mutado inactivado se puede mutar de cualquier forma adecuada, por ejemplo, por una o más mutaciones por eliminación, inserción, puntuales, de aminoácido, terminadora y de cambio del marco de lectura, para hacer que no haya expresión de la proteína. Por ejemplo, el gen se puede inactivar por inserción de una secuencia de ADN extraña, tal como un marcador de resistencia a antibiótico, en la secuencia que codifica el gen.

En una realización preferida, el gen no se muta por inserción de una secuencia de ADN extraña, tal como un marcador de resistencia a antibiótico, en la secuencia que codifica el gen. Preferiblemente, el gen de proteasa III comprende una mutación en el codón de inicio del gen y/o uno o más codones de parada situados en la dirección 3' del codón de inicio del gen y en la dirección 5' del codón de parada del gen, de modo que se previene la expresión de la proteína proteasa III.

Una mutación en el codón de parada del gen inactivado diana produce la pérdida de función del codón de inicio y por lo tanto asegura que el gen diana no comprende un codón de inicio adecuado al comienzo de la secuencia codificante. La mutación en el codón de inicio puede ser una mutación de aminoácido de 1, 2 o los tres nucleótidos del codón de inicio. Alternativa o adicionalmente, el codón de inicio se puede mutar por una mutación de cambio del marco de lectura por inserción o eliminación.

En una realización preferida, el gen ptr se muta para cambiar el codón de inicio ATG en ATT, como se muestra en la figura 11a.

El gen ptr mutado inactivado puede comprender alternativa o adicionalmente uno o más codones de parada situados en la dirección 3' del codón de inicio del gen y en la dirección 5' del codón de parada del gen. Preferiblemente, el gen ptr mutado inactivado comprende una mutación de aminoácido en el codón de inicio y uno o más codones de parada insertados.

El uno o más codones de parada insertados preferiblemente son codones de parada dentro del marco de lectura. Sin embargo, el uno o más codones de parada insertados pueden ser alternativa o adicionalmente codones de parada fuera del marco de lectura. Pueden ser necesarios uno o más codones de parada fuera del marco de lectura para detener la traducción cuando un codón de inicio fuera del marco de lectura se cambia a un codón de inicio dentro del marco de lectura por una mutación de cambio del marco de lectura por inserción o eliminación. El uno o más codones de parada se pueden introducir por cualquier mutación adecuada, incluyendo una mutación puntual terminadora y una mutación de cambio del marco de lectura. El uno o más codones de parada preferiblemente se introducen por una mutación de cambio del marco de lectura y/o una mutación por inserción, preferiblemente por 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, que comprende el codón de parada TAA.

En una realización preferida, el gen ptr se muta para insertar un codón de parada en el marco de lectura mediante inserción de un sitio de restricción de Ase I, como se muestra en la figura 11a. En una realización preferida, el gen ptr mutado inactivado tiene la secuencia de ADN de la SEQ ID NO: 6. Las mutaciones que se han hecho en la secuencia del gen ptr mutado inactivado de la SEQ ID NO: 6 se muestran en la figura 11a.

Las mutaciones de inactivación descritas antes son ventajosas porque producen una alteración mínima o no producen alteración del ADN cromosómico en la dirección 5' o dirección 3' del sitio del gen inactivado diana y por lo tanto no requieren la inserción y retención de ADN extraño, tal como marcadores de resistencia a antibióticos, que pueden afectar a la idoneidad de la célula para la expresión de una proteína de interés, en particular proteínas terapéuticas. Por consiguiente, una o más de las células de acuerdo con la presente invención pueden presentar características de crecimiento y/o expresión de proteína mejoradas comparadas con las células en donde el gen de la proteasa se ha inactivado por inserción de ADN extraño en la secuencia codificante del gen.

En una realización, las células de acuerdo con la presente invención llevan un gen OmpT mutado. Como se usa en la presente memoria, "gen OmpT" significa un gen que codifica la proteasa OmpT (proteasa T de membrana externa) que es una proteasa de la membrana externa. La secuencia del gen OmpT no mutado de tipo natural es SWISS-PROT P09169.

La referencia a un gen OmpT mutado o gen OmpT mutado que codifica la OmpT, se refiere a un gen OmpT mutado que codifica una proteína OmpT que tiene actividad reducida de proteasa o un gen OmpT mutado inactivado, salvo que se indique otra cosa.

La expresión "gen OmpT mutado que codifica una proteína OmpT que tiene actividad reducida de proteasa" en el contexto de la presente invención, significa que el gen OmpT mutado no tiene actividad de proteasa completa comparado con el gen OmpT no mutado de tipo natural. 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 de proteasa de una proteína OmpT no mutada de tipo natural. El gen OmpT mutado puede codificar una proteína OmpT que no tiene actividad de proteasa. En esta realización, la célula no es deficiente en OmpT cromosómica, es decir, la secuencia del gen OmpT no se ha eliminado o mutado para prevenir la expresión de cualquier forma de la proteína OmpT.

Se puede introducir cualquier mutación adecuada en el gen OmpT con el fin de producir una proteína que tenga actividad reducida de proteasa. La actividad de proteasa de una proteína OmpT expresada en una bacteria Gram negativa la puede ensayar fácilmente un experto en la técnica, mediante 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).

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) que en la OmpT la sustitución de serinas, histidinas y restos ácidos por alaninas da como resultado una actividad reducida en ~10 veces para Glu27, Asp97, Asp208 o His101, actividad reducida en ~500 veces para Ser99 y actividad reducida en ~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 No 18 5033-5039) describen que tiene un sitio activo que comprende una pareja Asp83-Asp85 y una pareja His212- Asp210 pair. 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) describe que las mutaciones D208A, D210A, H212A, H212N, H212Q, G216K/K217G, K217T y R218L en el bucle L4 dan todas como resultado la pérdida parcial o casi completa de actividad enzimática.

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

Uno o más de E27, D43, D83, D85, D97, S99, H101 E111, E136, E193, D206, D208, D210, H212 G216, K217, R218 y E250 se puede mutar por cualquier aminoácido adecuado que de como resultado una proteína que tenga actividad reducida de proteasa. Por ejemplo, uno o más de E27, D43, D83, D85, D97, S99, H101 E111, E136, E193, D206, D208, D210, H212, G216, K217, R218 y E250 se pueden mutar en alanina. Los 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 y E250A. En una realización, el gen OmpT mutado comprende las mutaciones D210A y H212A. Una secuencia de OmpT mutada que comprende las mutaciones D210A y H212A se muestra en la SEQ ID NO: 23.

La expresión "gen OmpT mutado inactivado" en el contexto de la presente invención significa que el gen comprende una o más mutaciones haciendo así que no haya expresión de la proteína codificada por el gen para proporcionar una célula deficiente en la proteína codificada por el gen mutado inactivado. El gen inactivado puede ser parcial o completamente transcrito pero no traducido en la proteína codificada. El gen OmpT mutado inactivado se puede mutar de cualquier forma adecuada, por ejemplo, por una o más mutaciones por eliminación, inserción, puntuales, de aminoácido, terminadora y de cambio del marco de lectura, para hacer que no haya expresión de la proteína. Por ejemplo, el gen se puede inactivar por inserción de una secuencia de ADN extraña, tal como un marcador de resistencia a antibiótico, en la secuencia que codifica el gen.

En una realización el gen OmpT comprende una mutación en el codón de inicio del gen y/o uno o más codones de parada situados en la dirección 3' del codón de inicio del gen y en la dirección 5' del codón de parada del gen, de modo que se previene la expresión de la proteína OmpT. La mutación en el codón de inicio puede ser una mutación de aminoácido de 1, 2 o los tres nucleótidos del codón de inicio. Se muestra una secuencia de OmpT inactivado mutado adecuada se muestra en la SEQ ID NO: 24. Alternativa o adicionalmente, el codón de inicio se puede mutar por una mutación de cambio del marco de lectura por inserción o eliminación.

En una realización, la célula bacteriana Gram negativa de acuerdo con la presente invención no lleva un gen OmpT mutado inactivado, tal como siendo deficiente en OmpT cromosómico.

En una realización, la célula bacteriana Gram negativa de acuerdo con la presente invención no lleva un gen degP mutado inactivado, tal como siendo deficiente en degP cromosómico. En una realización, la célula bacteriana Gram negativa de acuerdo con la presente invención no lleva un gen degP mutado.

En una realización, la célula bacteriana Gram negativa de acuerdo con la presente invención no lleva un gen ptr mutado inactivado, tal como siendo deficiente en ptr cromosómico.

Muchas mutaciones genéticamente diseñadas que incluyen mutaciones de inactivación implican el uso de marcadores de resistencia a antibióticos que permiten la selección e identificación de células mutadas satisfactoriamente. Sin embargo, como se ha discutido antes, hay una serie de desventajas en el uso de marcadores de resistencia a antibióticos.

Una realización adicional de la presente invención supera las desventajas anteriores del uso de marcadores de resistencia a antibióticos, en donde el gen Tsp mutado, el gen spr mutado y opcionalmente el gen DegP mutado y/o un gen ptr mutado y/o un gen OmpT mutado, se mutan para comprender uno o más sitios de restricción marcadores. Los sitios de restricción se diseñan genéticamente en el gen y no son de tipo natural. Los sitios de restricción

5 marcadores son ventajosos porque permiten el cribado y la identificación de células modificadas correctamente, que comprenden las mutaciones cromosómicas requeridas. Las células que se han modificado para llevar uno o más genes de proteasa mutados se pueden analizar por PCR del ADN genómico de lisatos celulares, usando parejas de oligonucleótidos diseñados para amplificar una región del ADN genómico que comprende un sitio de restricción marcador que no se encuentra de forma natural. Después, el ADN amplificado se puede analizar por electroforesis en gel de agarosa antes y después de incubación con una enzima de restricción adecuada capaz de digerir el ADN en el sitio de restricción marcador que no se encuentra de forma natural. La presencia de fragmentos de ADN después de incubación con la enzima de restricción confirma que las células se han modificado satisfactoriamente para llevar uno o más genes mutados.

10 En la realización en donde la célula lleva un gen ptr mutado inactivado que tiene la secuencia de SEQ ID NO: 6, las secuencias de cebadores oligonucleótidos mostradas en la SEQ ID NO: 17 y SEQ ID NO: 18, se pueden usar para amplificar la región del ADN que comprende el sitio de restricción Ase I que no se encuentra de forma natural, del ADN genómico de las células transformadas. El ADN genómico amplificado después se puede incubar con la enzima de restricción Ase I y después analizar por electroforesis en gel para confirmar la presencia del gen ptr mutado en el ADN genómico.

15 En la realización en donde la célula comprende un gen Tsp mutado inactivado que tiene la secuencia de SEQ ID NO: 3 o los nucleótidos 7 a 2048 de la SEQ ID NO: 3, las secuencias de cebadores oligonucleótidos mostradas en la SEQ ID NO: 15 y SEQ ID NO: 16, se pueden usar para amplificar la región del ADN que comprende el sitio de restricción Ase I que no se encuentra de forma natural, del ADN genómico de las células transformadas. El ADN genómico amplificado después se puede incubar con la enzima de restricción Ase I y después analizar por electroforesis en gel para confirmar la presencia del gen Tsp mutado en el ADN genómico.

20 En la realización en donde la célula comprende un gen DegP mutado que tiene la secuencia de ADN de la SEQ ID NO: 9, las secuencias de cebadores oligonucleótidos mostradas en la SEQ ID NO: 19 y SEQ ID NO: 20, se pueden usar para amplificar la región del ADN que comprende el sitio de restricción Ase I que no se encuentra de forma natural, del ADN genómico de las células transformadas. El ADN genómico amplificado después se puede incubar con la enzima de restricción Ase I y después analizar por electroforesis en gel para confirmar la presencia del gen DegP mutado en el ADN genómico.

25 El uno o más sitios de restricción se pueden introducir mediante cualquier mutación adecuada incluyendo una o más mutaciones por eliminación, inserción, puntuales, de aminoácido, terminadora y de cambio del marco de lectura. Un sitio de restricción se puede introducir mediante la mutación del codón de inicio y/o mutación para introducir uno o más codones de parada, como se ha descrito antes. Esta realización es ventajosa porque el sitio de restricción marcador es un marcador directo y único de las mutaciones de inactivación introducidas.

30 Se puede introducir un sitio de restricción marcador que comprende un codón de parada dentro del marco de lectura, tal como un sitio de restricción Ase I. Esto es particularmente ventajoso porque el sitio de restricción insertado sirve tanto como un sitio de restricción marcador como un codón de parada para prevenir la transcripción completa de la secuencia codificante del gen. Por ejemplo, en la realización en donde se introduce un codón de parada en el gen ptr por introducción de un sitio Ase I, esto también crea un sitio de restricción, como se muestra en la figura 11a. Por ejemplo, en la realización en donde se introduce un codón de parada en el gen Tsp en el codón 21 por introducción de un sitio Ase I, esto también crea un sitio de restricción, como se muestra en la figura 11b.

35 Un sitio de restricción marcador se puede insertar por la mutación en el codón de inicio y opcionalmente una o más mutaciones puntuales adicionales. En esta realización el sitio de restricción marcador preferiblemente es un sitio de restricción de *EcoR I*. Esto es particularmente ventajoso porque la mutación en el codón de inicio también crea un sitio de restricción marcador. Por ejemplo, en la realización en donde el codón de inicio del gen ptr se cambia por ATT, esto crea un sitio marcador *EcoR I*, como se muestra en la figura 11a. Por ejemplo, en la realización en donde el codón de inicio (codón 3) del gen Tsp se cambia de ATG en TCG, como se muestra en la figura 1b, una mutación puntual adicional del codón 2 de AAC a AAT y la mutación del codón 3 ATG en TCG crea un sitio de restricción marcador de *EcoR I*, como se muestra en la figura 11b.

40 En la realización de la presente invención en donde la célula lleva un gen OmpT mutado, el uno o más sitios de restricción se pueden introducir por cualquier mutación adecuada incluyendo una o más mutaciones por eliminación, inserción, puntuales, de aminoácido, terminadora y de cambio del marco de lectura. Por ejemplo, en la realización en donde el gen OmpT comprende las mutaciones D210A y H212A, estas mutaciones introducen un sitio de restricción de HindIII silencioso que se puede usar como un marcador de selección.

45 En el gen DegP o el gen spr, se puede introducir un sitio de restricción marcador usando cambios de codones silenciosos. Por ejemplo, un sitio Ase I se puede usar como un sitio de restricción marcador silencioso, en donde el codón de parada TAA está fuera del marco de lectura, como se muestra en la figura 11c para el gen DegP mutado.

50 En las realizaciones de la presente invención, en donde el gen ptr y/o el gen Tsp se mutan para codificar una proteasa III o Tsp que tienen actividad reducida de proteasa, se pueden introducir uno o más sitios de restricción marcadores usando cambios de codones silenciosos.

La célula bacteriana Gram negativa recombinante de acuerdo con la presente invención se puede producir por cualquier medio adecuado. El experto en la técnica conoce técnicas que se pueden usar para sustituir una secuencia de gen del cromosoma por una secuencia del gen mutada. Se pueden usar vectores adecuados que permiten la integración en el cromosoma del hospedante por recombinación homóloga.

- 5 Los métodos de sustitución de genes adecuados se describen, 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 and 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 homóloga/sustitución es el plásmido pKO3 (Link et al., 1997, *Journal of Bacteriology*, 179, 6228-6237).

Las cepas mutadas satisfactoriamente se pueden identificar usando métodos bien conocidos en la técnica, incluyendo secuenciación de ADN por PCR de colonia y cartografía de enzimas de restricción por PCR de colonia.

En la realización en donde la célula comprende dos o más genes cromosómicos mutados, los genes mutados se pueden introducir en la bacteria Gram negativa en el mismo o diferentes vectores.

- 20 En una realización la célula bacteriana Gram negativa de acuerdo con la presente invención no lleva un gen *OmpT* mutado inactivado, tal como siendo deficiente en *OmpT* cromosómico.

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 hospedante o puede estar en un vector no integrada, típicamente un plásmido.

En una realización, la célula de acuerdo con la presente invención expresa una proteína de interés. "Proteína de interés" en el contexto de la presente memoria descriptiva se pretende que se refiera al polipéptido para la expresión, normalmente un polipéptido recombinante. Sin embargo, la proteína de interés puede ser una proteína endógena expresada a partir de un gen endógeno en la célula hospedante.

- 30 Como se usa en la presente memoria, un "polipéptido recombinante" se refiere a una proteína que se construye o produce usando 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 actividad biológica atenuada, o un fragmento de la misma, expresada a partir de un vector exógeno. Alternativamente, la proteína de interés puede ser cualquier proteína adecuada incluyendo una proteína terapéutica, profiláctica o de diagnóstico.

- 35 En una realización, la proteína de interés es útil en el tratamiento de enfermedades o trastornos incluyendo enfermedades y trastornos inflamatorios, enfermedades y trastornos inmunitarios, trastornos fibróticos y cáncer.

La expresión "enfermedad inflamatoria" o "trastorno" y "enfermedad o trastorno inmunitario" incluye artritis reumatoide, artritis psoriásica, enfermedad de Still, enfermedad de Muckle-Wells, psoriasis, enfermedad de Crohn, colitis ulcerosa, LES (lupus eritematoso sistémico), asma, rinitis alérgica, dermatitis atópica, esclerosis múltiple, vasculitis, diabetes mellitus tipo I, trasplante y enfermedad de injerto contra huésped.

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 relacionada con la edad (DMAE), retinopatía, fibrosis cardíaca reactiva, cicatrices, 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 de epitelio, encontrado en la piel o, más comúnmente, el revestimiento de órganos del cuerpo, por ejemplo: mama, ovario, próstata, pulmón, riñón, páncreas, estómago, vejiga o intestino. Los cánceres tienden a infiltrarse en los tejidos adyacentes y a diseminarse (metastatar) a órganos distantes, por ejemplo: a los huesos, hígado, pulmones o el cerebro.

- 50 La proteína puede ser un polipéptido sensible a proteólisis, es decir, proteínas que tienen tendencia a escindirse, susceptibles de escisión, o escindidas por una o más proteasas de bacterias Gram negativas, tal como *E. coli*, en estado natural o durante la secreción. En una realización, la proteína de interés es sensible a la proteólisis por una proteasa seleccionada de DegP, proteasa III y Tsp. En una realización, la proteína de interés es sensible a la proteólisis por la proteasa Tsp. En una realización, la proteína de interés es sensible a la proteólisis por las proteasas DegP y proteasa III. En una realización, la proteína de interés es sensible a la proteólisis por las proteasas DegP y Tsp. En una realización, la proteína de interés es sensible a la proteólisis por las proteasas Tsp y proteasa

III. En una realización, la proteína de interés es sensible a la proteólisis por 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 usar como la proteína de interés incluyen anticuerpos monoclonales, multivalentes, multiespecíficos, humanizados, completamente humanos o quiméricos. El anticuerpo puede ser de cualquier especie, pero preferiblemente procede de un anticuerpo monoclonal, un anticuerpo humano o un fragmento humanizado. El anticuerpo puede proceder de cualquier clase (p. ej., IgG, IgE, IgM, IgD o IgA) o subclase de molécula de inmunoglobulina y se puede obtener de cualquier especie incluyendo, por ejemplo, ratón, rata, tiburón, conejo, cerdo, hámster, camello, llama, cabra o ser humano. Partes 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.

El anticuerpo puede ser una molécula de anticuerpo completa que tiene las cadenas pesadas y ligeras de longitud completa o un fragmento de la misma, p. ej., VH, VL, VHH, Fab, Fab modificado, Fab', F(ab')₂, Fv, fragmento scFv, Fab-Fv, o anticuerpo de especificidad doble, tal como Fab-dAb, como se describe en el documento PCT/GB2008/003331.

El anticuerpo puede ser específico para cualquier antígeno diana. El antígeno puede ser una proteína asociada a la célula, por ejemplo, una proteína de superficie celular en células tales como células bacterianas, células de levaduras, linfocitos T, células endoteliales o células tumorales, o puede ser una proteína soluble. Los antígenos de interés también pueden ser cualquier proteína relevante en medicina, tal como las proteínas que son reguladas por aumento durante la enfermedad o infección, por ejemplo, receptores y/o sus correspondientes ligandos. Los ejemplos particulares de proteínas de superficie celular incluyen las moléculas de adhesión, por ejemplo, integrinas tales como integrinas β 1, p. ej., 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, CDPC1, CSF1 o CSF1-Receptor, DPCR1, DPCR1, dudulin2, FLJ20584, FLJ40787, HEK2, KIAA0634, KIAA0659, KIAA1246, KIAA1455, LTBP2, LTK, MAL2, MRP2, similar a nectina 2, NKCC1, PTK7, RAIG1, TCAM1, SC6, BCMP101, BCMP84, BCMP11, DTD, antígeno carcinoembrionario (CEA), globulina de la grasa de la leche humana (HMFG1 and 2), antígenos de MHC de clase I y MHC de clase II, KDR y VEGF, y donde sea adecuado, receptores de los mismos.

Los antígenos solubles incluyen interleuquinas 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, tal como IL17A y/o IL17F, antígenos víricos, por ejemplo antígenos del virus sincitial respiratorio o de citomegalovirus, inmunoglobulinas, tal como IgE, interferones tales como interferón α , interferón β o interferón γ , factor de necrosis tumoral TNF (anteriormente conocido como factor de necrosis tumoral- α), factor de necrosis tumoral- β , factores estimuladores de colonias tales como G-CSF o GM-CSF, y factores de crecimiento derivados de plaquetas tales como PDGF- α y PDGF- β , y donde sea adecuado receptores de los mismos. Otros antígenos incluyen antígenos de superficie celular bacteriana, toxinas bacterianas, virus tales como influenza, EBV, HepA, B y C, agentes de bioterrorismo, radionúclidos y metales pesados, veneno de serpiente y araña y toxinas.

En una realización, el anticuerpo se puede usar para alterar funcionalmente la actividad del antígeno de interés. Por ejemplo, el anticuerpo puede neutralizar, antagonizar o hacer de agonista de la actividad de dicho antígeno, directa o indirectamente.

La presente invención también proporciona una célula bacteriana Gram negativa recombinante que comprende un gen Tsp mutado, en donde el gen Tsp mutado codifica una proteína Tsp que tiene actividad reducida de proteasa o es un gen Tsp mutado inactivado, un gen spr mutante que codifica una spr mutante y una secuencia de polinucleótido que codifica un anticuerpo o un fragmento de unión al antígeno del mismo específico para el TNF.

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 (cuyo contenido se incorpora en la presente memoria por referencia).

En una realización, el anticuerpo que tiene especificidad por el TNF α humano, comprende una cadena pesada en donde el dominio variable comprende una CDR que tiene una secuencia mostrada en la SEQ ID NO: 26 para CDRH1, la secuencia mostrada en la SEQ ID NO: 27 o SEQ ID NO:32 para CDRH2 o la secuencia mostrada en la SEQ ID NO: 28 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 la SEQ ID NO: 29 para CDRL1, la secuencia mostrada en la SEQ ID NO: 30 para CDRL2 o la secuencia mostrada en la SEQ ID NO:31 para CDRL3.

Las CDR dadas en las SEQ ID NO: 26 y 28 a 32 a las que se ha hecho referencia antes, proceden de un anticuerpo monoclonal de ratón hTNF40. Sin embargo, la SEQ ID NO: 27 consiste en una CDR híbrida. La CDR híbrida comprende parte de la CDR2 de la cadena pesada del anticuerpo monoclonal de ratón hTNF40 (SEQ ID NO: 32) y

parte de la CDR2 de la cadena pesada de una secuencia de la región V de la línea germinal del grupo 3 humano.

En una realización, el anticuerpo comprende una cadena pesada en donde el dominio variable comprende una CDR que tiene la secuencia mostrada en la SEQ ID NO: 26 para CDRH1, la secuencia mostrada en la SEQ ID NO: 27 o SEQ ID NO: 32 para CDRH2 o la secuencia mostrada en la SEQ ID NO: 28 para CDRH3 y una cadena ligera en donde el dominio variable comprende una CDR que tiene la secuencia mostrada en la SEQ ID NO: 29 para CDRL1, la secuencia mostrada en la SEQ ID NO: 30 para CDRL2 o la secuencia mostrada en la SEQ ID NO: 31 para CDRL3.

En una realización, el anticuerpo comprende la SEQ ID NO: 26 para CDRH1, SEQ ID NO: 27 o SEQ ID NO: 32 para CDRH2, SEQ ID NO: 28 para CDRH3, SEQ ID NO: 29 para CDR1, SEQ ID NO: 30 para CDRL2 and SEQ ID NO: 31 para CDRL3. Preferiblemente, el anticuerpo comprende la SEQ ID NO: 27 para CDRH2.

El anticuerpo anti-TNF preferiblemente es una molécula de anticuerpo con CDR injertada. En una realización preferida, el dominio variable comprende las regiones armazón aceptoras humanas y CDR donadoras no humanas.

Preferiblemente, la molécula de anticuerpo tiene especificidad por el TNF humano (anteriormente conocido como TNF α), 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 preferiblemente es un fragmento Fab o Fab'.

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

Después de la expresión, los fragmentos de anticuerpo pueden ser procesados más, por ejemplo, por conjugación con otra entidad tal como una molécula efectora.

La expresión molécula efectora como se usa 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 anticuerpos, polímeros sintéticos o naturales, ácidos nucleicos y fragmentos de los mismos, p. ej., ADN, ARN y fragmentos de los mismos, radionúclidos, en particular radioyoduro, radioisótopos, metales quelados, nanopartículas y grupos indicadores tales como compuestos fluorescentes o compuestos que se pueden detectar por espectroscopía de RMN o ESR. El efector molecular se puede unir al anticuerpo o fragmento del mismo por cualquier método adecuado, por ejemplo, un fragmento de anticuerpo se puede modificar para unir al menos una molécula efectora, como se describe en los documentos WO05/003171 o WO05/003170 (cuyos contenidos se incorporan en la presente memoria por referencia). Los documentos WO05/003171 o WO05/003170 también describen moléculas efectoras adecuadas.

En una realización, el anticuerpo o fragmento del mismo, tal como Fab, es PEGilado para generar un producto con las propiedades requeridas, por ejemplo similar a los anticuerpos enteros, si se requiere. Por ejemplo, el anticuerpo puede ser un Fab' anti-TNF α PEGilado, como se describe en el documento WO01/094585, preferiblemente teniendo unido a uno de los restos de cisteína en el extremo C-terminal de la cadena pesada un grupo derivado de lisil-maleimida, en donde cada uno de los dos grupos amino del resto lisilo tiene unido covalentemente al mismo un resto metoxipolietilenglicol que tiene un peso molecular de aproximadamente 20.000 Da, de modo que el peso molecular medio total de los restos de metoxipolietilenglicol es aproximadamente 40.000 Da, más preferiblemente el grupo derivado de lisil-maleimida es el [1-[[[2-[[3-(2,5-dioxo-1-pirrolidinil)-1-oxopropil]amino]etil]amino]-carbonil]-1,5-pentanodiol]bis(iminocarbonilo).

La célula también puede comprender secuencias de polinucleótidos adicionales que codifican una o más proteínas de interés adicionales.

En una realización, una o más proteínas del hospedante *E. coli* que en el tipo natural se sabe que copurifican con la proteína recombinante de interés durante la purificación, se seleccionan para la modificación genética, como describen 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 (cuyos contenidos se incorporan en la presente memoria por referencia). El uso de dichas proteínas de hospedante modificadas mejora el procedimiento de purificación de las proteínas de interés, en especial anticuerpos, producidas en *E. coli* alterando las propiedades físicas de las proteínas de *E. coli* seleccionadas, de modo que ya no copurifiquen con el anticuerpo recombinante. Preferiblemente, la proteína de *E. coli* que se altera se selecciona de una o más de la proteína de unión a fosfato (PhoS/PstS), proteína de unión a dipéptido (DppA), proteína de unión a maltosa (MBP) y tioredoxina.

En una realización, se altera una propiedad física de una proteína del hospedante contaminante por la adición de un marcador aminoácido en el extremo C o extremo N. En una realización preferida, la propiedad física que se altera es el punto isoeléctrico y el marcador aminoácido es un marcador poli(ácido aspártico) unido al extremo C. En una realización, las proteínas de *E. coli* alteradas por la adición de dicho marcador son la proteína de unión a dipéptido

(DppA), proteína de unión a maltosa (MBP), tioredoxina y proteína de unión a fosfato (PhoS/PstS). En una realización específica, el pI de la proteína de unión a fosfato (PhoS/PstS) de *E. coli* se reduce de 7,2 a 5,1 por adición de un marcador poli(ácido aspártico) (poliD), que contiene 6 restos ácido aspártico en el extremo C.

5 También se prefiere la modificación de restos específicos de la proteína contaminante de *E. coli*, para alterar sus propiedades físicas, sea sola o en combinación con la adición de marcadores N o C terminales. Dichos cambios pueden incluir inserciones o eliminaciones para alterar el tamaño de la proteína o sustituciones de aminoácidos para alterar el pI o la hidrofobicidad. En una realización, estos restos están situados en la superficie de la proteína. En una realización preferida, se alteran los restos de la superficie de la proteína PhoS con el fin de reducir el pI de la proteína. Preferiblemente, se evitan los restos que se han implicado como importantes en la unión de fosfato (Bass, US5.304.472) con el fin de mantener una proteína PhoS funcional. Preferiblemente, los restos lisina que se proyectan hacia fuera de la superficie de la proteína o están en o cerca de grupos grandes de restos básicos son el objetivo. En una realización, la proteína PhoS tiene un marcador de poli-hexa-(ácido aspártico) unido al extremo C, mientras que los restos de la superficie en el lado opuesto de la molécula son objetivo para la sustitución. Preferiblemente, los restos lisina seleccionados se sustituyen por ácido glutámico o ácido aspártico para conferir un cambio de pI potencial mayor que cuando se cambian restos neutros por ácidos. La designación de 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 natural. El número se refiere a la posición del aminoácido donde se está haciendo la sustitución de aminoácido, y la segunda letra designa el aminoácido que se usa para sustituir el aminoácido de tipo natural. En mutaciones preferidas de PhoS en la presente invención, los restos lisina (K) 275, 107, 109, 110, 262, 265, 266, 309, 313 se sustituyen por ácido glutámico (E) o glutamina (Q), como mutaciones individuales o combinadas, además de que la lisina (K)318 se puede sustituir por ácido aspártico (D) como mutación individual 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 poli(ácido aspártico) (poliD) unido al extremo C y opcionalmente también con la sustitución K318D. En una realización preferida, las mutaciones producen una reducción del pI en 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 de 7,2 a aproximadamente 4,9, aproximadamente 4,8 y aproximadamente 4,5, usando las mutaciones poliD K318D, poliD K265/266E y poliD K110/265/266E, respectivamente.

30 El polinucleótido que codifica la proteína de interés puede ser expresado como una fusión con otro polipéptido, preferiblemente una secuencia señal u otro polipéptido que tenga un sitio de escisión específico en el extremo N del polipéptido maduro. La secuencia señal heteróloga seleccionada debería ser una que sea reconocida y procesada por la célula hospedante. Para células hospedantes procariontas que no reconocen y procesan la secuencia señal del polipéptido natural o uno eucariota, la secuencia señal se sustituye por una secuencia señal procarionta. Las secuencias señal adecuadas incluyen OmpA, PhoA, LamB, PelB, DsbA y DsbC.

35 La construcción de vectores adecuados que contienen uno o más de los componentes citados antes, usa técnicas de ligado convencionales. Los plásmidos o fragmentos de ADN aislados son escindidos, diseñados y vueltos a ligar en la forma deseada para generar los plásmidos requeridos.

40 En una realización, se usa un casete de expresión en la presente invención para llevar el polinucleótido que codifica la proteína de interés, que típicamente comprende una o más secuencias codificantes de proteínas que codifican una o más proteínas de interés y una o más secuencias reguladoras de la expresión. La una o más secuencias reguladoras de la expresión pueden incluir un promotor. La una o más secuencias reguladoras de la expresión también pueden incluir una región 3' no traducida tal como una secuencia de terminación. Los promotores adecuados se describen con más detalle más adelante.

45 En una realización, la célula de acuerdo con la presente invención comprende un vector, tal como un plásmido. El vector preferiblemente comprende uno o más casetes de expresión como se han definido antes.

50 En la realización donde la proteína de interés es un anticuerpo que comprende tanto las cadenas 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.

El vector para usar en la presente invención se puede producir por inserción de un casete de expresión como se ha definido antes, en un vector adecuado. Alternativamente, las secuencias reguladoras de expresión para dirigir la expresión de la secuencia de polinucleótido que codifica una proteína de interés, pueden estar contenidas en el vector y por lo tanto puede ser necesaria solo la región codificante del polinucleótido para completar el vector.

55 Los ejemplos de vectores que se pueden usar para transformar la célula hospedante con un polinucleótido de acuerdo con la invención incluyen:

- un plásmido, tal como pBR322 o pACYC184, y/o
- un vector vírico tal como fago bacteriano

- un elemento genético transponible tal como un transposón

Están disponibles muchas formas de vectores de expresión. Dichos vectores normalmente comprenden un origen de replicación de ADN plasmídico, un marcador seleccionable por antibiótico, un promotor y un terminador de la transcripción separado por un sitio de clonación múltiple (casete de expresión) y una secuencia de ADN que codifica un sitio de unión al ribosoma.

Los promotores usados en la presente invención se pueden unir al polinucleótido relevante directamente o alternativamente se pueden situar en una posición adecuada, por ejemplo, en un vector de modo que cuando el polipéptido relevante se inserta, el promotor relevante puede actuar en el mismo. En una realización, el promotor se coloca antes de la parte codificante del polinucleótido en el que actúa, por ejemplo, un promotor relevante antes de cada parte codificante del polinucleótido. "Antes" como se usa en la presente memoria se pretende que implique que el promotor se coloca en el extremo del cebador 5 con respecto a la parte de polinucleótido codificante.

Los promotores pueden ser endógenos o exógenos para las células hospedantes. Los promotores adecuados incluyen Lac, tac, trp, PhoA, Ipp, Arab, Tet y T7.

Uno o más promotores usados pueden ser promotores inducibles.

Las unidades de expresión para usar en sistemas bacterianos en general también contienen una secuencia de Shine-Dalgarno (S. D.) de ribosoma operativamente unida al ADN que codifica el polipéptido de interés.

En realizaciones de la presente invención en donde una secuencia de polinucleótido comprende dos o más secuencias codificantes para dos o más proteínas de interés, por ejemplo una cadena ligera de anticuerpo y cadena pesada de anticuerpo, la secuencia de polinucleótido puede comprender una o más secuencias de sitios internos de entrada al ribosoma (IRES), que permiten el inicio de la traducción en el medio de un ARNm. Una secuencia de IRES puede estar situada entre secuencias del polinucleótido codificantes para potenciar la traducción separada del ARNm para producir las secuencias de polipéptido codificadas.

El vector de expresión preferiblemente también comprende un mensaje bicistrónico para producir el anticuerpo o fragmento de unión al antígeno del mismo, como se describe en los documentos WO 03/048208 o WO2007/039714 (cuyos contenidos se incorporan en la presente memoria por referencia). Preferiblemente, el cistrón en la dirección 5' contiene ADN que codifica la cadena ligera del anticuerpo y el cistrón en la dirección 3' contiene ADN que codifica la correspondiente cadena pesada, y la secuencia intergénica dicistrónica (IGS) preferiblemente comprende una secuencia seleccionada de IGS1 (SEQ ID NO: 34), IGS2 (SEQ ID NO: 35), IGS3 (SEQ ID NO: 36) y IGS4 (SEQ ID NO: 37).

Los terminadores pueden ser endógenos o exógenos para las células hospedantes. Un terminador adecuado es rrnB.

Se pueden encontrar reguladores transcripcionales adicionales adecuados que incluyen promotores y terminadores y métodos para dirigir proteínas en "Strategies for Achieving High-Level Expression of Genes in Escherichia coli" Sawas C. Makrides, Microbiological Reviews, Sept 1996, pág. 512-538.

La molécula de anticuerpo puede ser secretadas de la célula o dirigida el periplasma mediante secuencias señal adecuadas. Alternativamente, las moléculas de anticuerpo se pueden acumular dentro del citoplasma de la célula. Preferiblemente, la molécula de anticuerpo se dirige al periplasma.

Las realizaciones de la invención descritas en la presente memoria con referencia al polinucleótido se aplican igualmente a realizaciones alternativas de la invención, por ejemplo, vectores, casetes de expresión y/o células hospedantes que comprenden los componentes usados en las mismas, siempre que el aspecto relevante se puede aplicar a las mismas.

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 antes en el primer y segundo aspectos de la presente invención.

La célula bacteriana Gram negativa y la proteína de interés usados preferiblemente en el método de la presente invención, se han descrito antes en detalle.

Cuando el polinucleótido que codifica la proteína de interés es exógeno, el polinucleótido se puede incorporar en la célula hospedante usando cualquier medio adecuado conocido en la técnica. Típicamente, 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.

La secuencia de polinucleótido se puede transformar en una célula usando técnicas convencionales, por ejemplo usando cloruro de rubidio, PEG o electroporación.

5 El método de acuerdo con la presente invención también puede usar un sistema de selección para facilitar la selección de células estables que se han transformado satisfactoriamente con el polinucleótido que codifica la proteína de interés. El sistema de selección típicamente usa la contransformación 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 más marcadores de selección. Por consiguiente, la transformación del polinucleótido que codifica la proteína de interés y el uno o más polinucleótidos que codifican el marcador se produce juntos y el sistema de selección se puede usar para seleccionar aquellas células que producen las proteínas deseadas.

10 Las células que pueden expresar uno o más marcadores pueden sobrevivir/crecer/multiplicarse en determinadas condiciones impuestas artificialmente, por ejemplo, la adición de una toxina o antibiótico, debido a las propiedades de las que están dotados el polipéptido/gen o componente del polipéptido del sistema de selección incorporado en las mismas (p. ej., resistencia a antibiótico). Las células que no pueden expresar el uno o más marcadores no pueden sobrevivir/crecer/multiplicarse en las condiciones impuestas artificialmente. Las condiciones impuestas artificialmente se pueden elegir para que sean más o menos fuertes, según sea necesario.

15 Se puede usar cualquier sistema de selección adecuado en la presente invención. Típicamente, el sistema de selección se puede basar en incluir en el vector uno o más 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 puesto que expresan tanto el gen que da resistencia al antibiótico como la proteína deseada.

20 En una realización, el método de acuerdo con la invención comprende además la etapa de cultivar la célula transformada en un medio para de esta forma expresar la proteína de interés.

Se puede usar un sistema de expresión inducible o un promotor constitutivo en la presente invención para expresar la proteína de interés. Los sistemas de expresión inducibles y promotores constitutivos adecuados son bien conocidos en la técnica.

25 Se puede usar cualquier medio adecuado para cultivar la célula transformada. El medio puede estar adaptado para un sistema de selección específico, por ejemplo, el medio puede comprender un antibiótico, para permitir que crezcan en el medio solo las células que se han transformado satisfactoriamente.

30 Las células obtenidas del medio se pueden someter a cribado y/o purificación adicionales según se requiera. El método puede comprender además una o más etapas para extraer y purificar la proteína de interés, según se requiera.

El polipéptido se puede recuperar de la cepa, incluyendo el citoplasma, periplasma o líquido sobrenadante.

35 El o los métodos específicos usados para purificar una proteína dependen del tipo de proteína. Los métodos adecuados incluyen el fraccionamiento en columnas de inmunoafinidad o de intercambio iónico; precipitación con etanol; HPLC de fase inversa; cromatografía de interacción hidrófoba; cromatografía en sílice; cromatografía en una resina de intercambio iónico tal como S-SEPHAROSE y DEAE; cromatografía de exclusión; precipitación con sulfato amónico; y filtración en gel.

40 Los anticuerpos se pueden separar de forma adecuada del medio de cultivo y/o extraer del citoplasma y/o extraer del periplasma por procedimientos de purificación de anticuerpos convencionales tales como, por ejemplo, proteína A-Sepharose, cromatografía con proteína G, cromatografía con proteína L, resinas tiorflicas, de modo mixto, marcador His, marcador FLAG, cromatografía de hidroxapatita, electroforesis en gel, diálisis, cromatografía de afinidad, sulfato amónico, 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 puede incluir también una etapa adicional de medición de la cantidad de expresión de la proteína de interés y selección de las células que tienen niveles de expresión altos de la proteína de interés.

45 El método puede incluir también una o más etapas de procesamiento corriente abajo, tales como PEGilación de la proteína de interés, tal como un anticuerpo o fragmento de anticuerpo.

Se pueden llevar a cabo una o más etapas del método descrito en la presente memoria en combinación en un recipiente adecuado tal como un biorreactor.

Ejemplos

50 Ejemplo 1 - Generación de cepa de células MXE001 (Δ Tsp)

La cepa MXE001 se generó como sigue:

El casete de Tsp se movió como fragmentos de restricción de Sal I y Not I a plásmidos pKO3 de restricción similar. El plásmido pKO3 usa el mutante sensible a la temperatura del origen de replicación pSC101 (*RepA*) junto con un

marcador de cloranfenicol para forzar y seleccionar los sucesos de integración cromosómica. El gen *sacB* que codifica levansacarosa es letal para el crecimiento de *E. coli* en sacarosa y por lo tanto (junto con el marcador de cloranfenicol y el origen pSC101) se usa para forzar y seleccionar sucesos de desintegración y curado de plásmidos. Esta metodología ha sido descrita previamente (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 hospedante excepto para el gen insertado.

Se construyeron los siguientes plásmidos.

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

Después, el plásmido se transformó en células *E. coli* W3110 competentes electrocompetentes preparadas usando el método encontrado en Miller, E.M. y Nickoloff, J.A., "Escherichia coli electrotransformation," en *Methods in Molecular Biology*, vol. 47, Nickoloff, J.A. (ed.), Humana Press, Totowa, NJ, 105 (1995).

Día 1, se mezclaron 40 µl de células *E. coli* con (10 pg) 1 µl de ADN de pKO3 en una cubeta de electroporación BioRad de 0,2 cm enfriada, a 2500 V, 25 µF y 200 Ω. Se añadieron inmediatamente 1000 µl de 2xPY, las células se recuperaron por agitación a 250 rpm en un incubador a 30°C durante 1 h. Se hicieron diluciones seriadas 1/10 de las células en 2xPY antes de sembrar partes alícuotas de 100 µl en placas de agar 2xPY que contenían cloranfenicol 20 µg/ml previamente calentadas a 30°C y 43°C. Las placas se incubaron durante la noche a 30°C y 43°C.

Día 2, el número de colonias desarrolladas a 30°C daba un cálculo de la eficacia de la electroporación, mientras que las colonias que sobrevivieron al crecimiento a 43°C representan potenciales sucesos de integración. Las colonias individuales de la placa a 43°C se recogieron y se volvieron a suspender en 10 ml de 2xPY. Se sembraron 100 µl de estos en placas de agar 2xPY que contenían sacarosa al 5% (p/v) precalentadas a 30°C para generar colonias individuales. Las placas se incubaron durante la noche a 30°C.

Día 3, las colonias aquí representan los potenciales sucesos de desintegración y curado plasmídico simultáneos. Si los sucesos de desintegración y curado se producían pronto en el crecimiento, entonces el conjunto de la masa de colonias será clónica. Se recogieron colonias individuales y las copias se sembraron en agar 2xPY que contenía cloranfenicol 20 µg/ml o sacarosa al 5% (p/v). Las placas se incubaron durante la noche a 30°C.

Día 4, las colonias que crecían en sacarosa y morían en cloranfenicol representan los sucesos potenciales de sustitución cromosómica y de curado plasmídico. Estas se recogieron y se cribaron por PCR con un oligonucleótido específico de la mutación. Las colonias que generaron una banda de PCR positiva del tamaño correcto se sacaron para producir colonias individuales en agar 2xPY que contenía sacarosa al 5% (p/v) y las placas se incubaron durante la noche a 30°C.

Día 5, las colonias individuales de *E. coli* positivas por PCR, sensibles a cloranfenicol y resistentes a sacarosa, se usaron para hacer células químicamente competentes, preparaciones en glicerol, y actúan como moldes en la PCR para una reacción de PCR con oligonucleótidos flanqueadores 5' y 3' para generar el producto de la PCR para la secuenciación directa de ADN usando la polimerasa Taq.

La cepa celular MXE001 se ensayó para confirmar la modificación satisfactoria del ADN genómico que lleva el gen *Tsp* mutado por amplificación por PCR de la región del gen *Tsp* que comprende un sitio de restricción *Ase* I no natural, como se muestra en las figuras 1a, 1b y 1c, usando cebadores oligonucleótidos. Las regiones amplificadas del ADN después se analizaron por electroforesis en gel antes y después de incubación con la enzima de restricción *Ase* I para confirmar la presencia del sitio de restricción *Ase* I no natural en los genes mutados. Este método se lleva a cabo como sigue:

Se usaron los siguientes oligonucleótidos para amplificar, usando la PCR, el ADN genómico de preparados de lisados celulares de *E. coli* de MXE001 y W3110:

6284 *Tsp* 3' 5'-GCATCATAATTTTCTTTTACCTC-3' (SEQ ID NO: 15)

6283 *Tsp* 5' 5'-GGGAAATGAACCTGAGCAAACGC-3' (SEQ ID NO: 16)

Los lisados se prepararon calentando una colonia individual de células durante 10 min a 95°C en 20 µl de 1x tampón de PCR. La mezcla se dejó enfriar a temperatura ambiente y después se centrifugó a 13.200 rpm durante 10 min. El líquido sobrenadante se separó y se marcó como "lisato celular".

Cada cepa se amplificó usando la pareja de oligonucleótidos de *Tsp*.

El ADN se amplificó usando un procedimiento de PCR convencional.

5 µl tampón x 10 (Roche)

1 µl mezcla de dNTP (Roche, mezcla 10 mM)

- 1,5 µl oligonucleótido 5' (5 pmol)
- 1,5 µl oligonucleótido 3' (5 pmol)
- 2 µl lisato celular
- 0,5 µl ADN polimerasa Taq (Roche 5 U/µl)

5 38,5 µl HO

Ciclo de PCR

94°C 1 minuto

94°C 1 minuto)

55 1 minuto) repetido durante 30 ciclos

10 72 1 minuto)

72 10 minutos

Una vez completadas las reacciones, se separaron 25 µl a un nuevo tubo de microfuga para la digestión con Ase I. A los 25 µl de reacción de PCR se añadieron 19 µl de H₂O, 5 µl de tampón 3 (NEB), 1 µl de Ase I (NEB), se mezclaron y se incubaron a 37°C durante 2 horas.

15 Al resto de la reacción de PCR se añadieron 5 µl de tampón de carga (x6) y se cargaron 200 µl en un gel de agarosa de 200 ml de TAE al 0,8% (Invitrogen) más bromuro de etidio (5 µl de disolución madre de 10 mg/ml) y se ejecutó a 100 voltios durante 1 hora. Se cargaron en la banda final 10 µl de marcador de tamaño (marcador Perfect DNA 0,1-12 kB, Novagen).

20 Una vez completadas las digestiones con Ase I, se añadieron 10 µl de tampón de carga (x6) y se cargaron 200 µl en un gel de agarosa con TAE al 0,8% (Invitrogen) más bromuro de etidio (5 µl de disolución madre de 10 mg/ml) y se ejecutó a 100 voltios durante 1 hora. Se cargaron en la banda final 10 µl de marcador de tamaño (marcador Perfect DNA 0,1-12 kB, Novagen). Ambos geles se visualizaron usando un transluminador de UV.

25 El fragmento genómico amplificado mostró la banda de tamaño correcto de 2,8 kb para la Tsp. Después de digestión con Ase I, esto confirmó la presencia de los sitios Ase I introducidos en la cepa deficiente en Tsp MXE001 pero no en la de control W3110. MXE001: el ADN genómico se amplificó usando el conjunto de cebadores de Tsp y el ADN resultante se digirió con Ase I para producir bandas de 2,2 y 0,6 Kpb.

El ADN amplificado por PCR de W3110 no era digerido por la enzima de restricción Ase I.

Ejemplo 2 - Generación de mutantes de spr

Las mutaciones de spr se generaron y seleccionaron usando un ensayo de complementación.

30 El gen spr se mutó usando el kit de PCR de diversidad de mutagénesis aleatoria de Clontech®, que introdujo de 1 a 2 mutaciones por 1000 pb. El ADN de la PCR de spr mutado se clonó en un vector de expresión inducible [pTTO CDP870] que expresa Fab' CDP870 junto con el mutante de spr. Después este ligado se electrotransformó en una cepa de *E. coli* MXE001 (Δ Tsp) preparada usando el método encontrado en Miller, E.M. y Nickoloff, J.A., "Escherichia coli electrotransformation," en *Methods in Molecular Biology*, vol. 47, Nickoloff, J.A. (ed.), Humana Press, Totowa, NJ, 105 (1995). Se usó el siguiente protocolo, se añadieron 40 µl de MXE001 electrocompetente, 2,5 µl del ligado (100 pg de ADN) a una cubeta de electroporación de 0,2 cm, se llevó a cabo la electrotransformación usando un BioRad Genepulser Xcell con las siguientes condiciones, 2500 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 h con agitación suave.

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

45 Usando este método, se aislaron 8 mutaciones simples, una doble y dos mutaciones múltiples en la proteína spr, que complementaban el fenotipo de Δ Tsp como sigue:

1. V98E

2. D133A

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

Ejemplo 3 - Generación de cepas de células de *E. coli* mutantes que llevan mutaciones de spr

10 Las mutaciones individuales 1 a 5 identificadas en el ejemplo 2 y tres mutaciones de triadas catalíticas de spr (C94A, H145A, H157A) y W174R se usaron para generar nuevas cepas usando la cepa de *E. coli* W3110 de tipo natural (genotipo: FLAM- IN (rrnD-rrnE)1 rph1 (ATCC nº 27325)) para crear cepas con spr mutado o la cepa MXE001 del ejemplo 1 para hacer cepas combinadas de Δ Tsp/spr mutante.

15 Se generaron las siguientes cepas de células de *E. coli* mutantes usando un sistema de vector de sustitución de gen usando el plásmido de recombinación homóloga/sustitución pK03 (Link et al., 1997, *Journal of Bacteriology*, 179, 6228-6237), como se describe en el ejemplo 1 para la generación de MXE001.

Tabla 1

Cepas de células de <i>E. coli</i> mutantes	Genotipo	Vectores de spr
MXE001	Δ Tsp	
MXE008	Δ Tsp, spr D133A	pMXE339, pK03 spr D133A (-Sall)
MXE009	Δ 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	Δ Tsp, spr V135D	pMXE340, pK03 spr V135D (-Sall)
MXE015	Δ Tsp, spr V98E	pMXE342, pK03 spr V98E (-Sall)
MXE016	Δ Tsp, spr C94A	pMXE343, pK03 spr C94A (-Sall)
MXE017	Δ Tsp, spr H145A	pMXE344, pK03 spr H145A (-Sall)
MXE018	Δ Tsp, spr V135G	pMXE341, pK03 spr V135G (-Sall)

20 Los casetes de integración de spr mutante se movieron como fragmentos de restricción de Sal I, Not I en plásmidos pK03 de restricción similar.

25 El plásmido usa el mutante sensible a la temperatura del origen de replicación pSC101 (*RepA*) junto con un marcador de cloranfenicol para forzar y seleccionar los sucesos de integración cromosómica. El gen *sacB* que codifica levansacarosa es letal para el crecimiento de *E. coli* en sacarosa y por lo tanto (junto con el marcador de cloranfenicol y el origen pSC101) se usa para forzar y seleccionar sucesos de desintegración y curado de plásmidos. Esta metodología ha sido descrita previamente (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 hospedante excepto para el gen insertado.

Se construyeron los vectores pK03 citados a continuación que comprenden 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 de

clones.

pMXE336, pK03 spr S95F (-Sall)

pMXE337, pK03 spr Y115F (-Sall)

pMXE338, pK03 spr G147C (-Sall)

5 pMXE339, pK03 spr D133A (-Sall)

pMXE340, pK03 spr V135D (-Sall)

pMXE341, pK03 spr V135G (-Sall)

pMXE342, pK03 spr V98E (-Sall)

pMXE343, pK03 spr C94A (-Sall)

10 pMXE344, pK03 spr H145A (-Sall)

pMXE345, pK03 spr H157A (-Sall)

pMXE346, pK03 spr W174R (-Sall)

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

20 Día 1, se mezclaron 40 μ l de células de *E. coli* electrocompetentes o células MXE001 con (10 pg) 1 μ l de ADN de pK03 en una cubeta de electroporación BioRad de 0,2 cm enfriada, a 2500 V, 25 μ F y 200 Ω . Se añadieron inmediatamente 1000 μ l de 2xPY, las células se recuperaron por agitación a 250 rpm en un incubador a 30°C durante 1 h. Se hicieron diluciones seriadas 1/10 de las células en 2xPY antes de sembrar partes alícuotas de 100 μ l en placas de agar 2xPY que contenían cloranfenicol 20 μ g/ml previamente calentadas a 30°C y 43°C. Las placas se incubaron durante la noche a 30°C y 43°C.

25 Día 2, el número de colonias desarrolladas a 30°C daba un cálculo de la eficacia de la electroporación, mientras que las colonias que sobrevivieron al crecimiento a 43°C representan potenciales sucesos de integración. Las colonias individuales de la placa a 43°C se recogieron y se volvieron a suspender en 10 ml de 2xPY. Se sembraron 100 μ l de estos en placas de agar 2xPY que contenían sacarosa al 5% (p/v) precalentadas a 30°C para generar colonias individuales. Las placas se incubaron durante la noche a 30°C.

30 Día 3, las colonias aquí representan los potenciales sucesos de desintegración y curado plasmídico simultáneos. Si los sucesos de desintegración y curado se producían pronto en el crecimiento, entonces el conjunto de la masa de colonias será clónico. Se recogieron colonias individuales y las copias se sembraron en placa en agar 2xPY que contenía cloranfenicol 20 μ g/ml o sacarosa al 5% (p/v). Las placas se incubaron durante la noche a 30°C.

35 Día 4, las colonias que crecían en sacarosa y morían en cloranfenicol representan los sucesos potenciales de sustitución cromosómica y de curado plasmídico. Estas se recogieron y se cribaron por PCR más digestión de restricción para la pérdida de un sitio Sall. Las colonias que generaban una banda de PCR positiva del tamaño correcto y resistencia a la digestión por Sall se sacaron para producir colonias individuales en agar 2xPY que contenía sacarosa al 5% (p/v) y las placas se incubaron durante la noche a 30°C.

40 Día 5, las colonias individuales de *E. coli* positivas por PCR, sensibles a cloranfenicol y resistentes a sacarosa, se usaron para hacer células químicamente competentes, preparaciones en glicerol, y actúan como moldes en la PCR para una reacción de PCR con oligonucleótidos flanqueadores 5' y 3' para generar el producto de la PCR para la secuenciación directa de ADN usando la polimerasa Taq para confirmar la mutación correcta.

Ejemplo 4 - Expresión de Fab' anti-TNF en las cepas de spr mutante

45 Las cepas de spr mutante MXE008, MXE009, MXE012 y MXE017 proporcionadas en el ejemplo 3 y la cepa MXE001 proporcionada en el ejemplo 1, se transformaron con el plásmido pMXE117 (pTTO CDP870 IGS2), se construyó un vector de expresión para el Fab' CDP870 (un Fab' anti-TNF que tiene una secuencia de la cadena ligera mostrada en la SEQ ID NO: 13 y una secuencia de la cadena pesada mostrada en la SEQ ID NO: 14), usando las metodologías de clonación de 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 IGS17) contenía las siguientes características: un promotor *tac* fuerte y una secuencia de operador *lac*. Los genes de la cadena pesada y ligera de Fab se transcribieron como un mensaje dicistrónico individual. El ADN que codifica el péptido señal de la proteína OmpA de *E. coli* se fusionó en el extremo 5' de las secuencias de genes tanto de la

cadena ligera como de la pesada, que dirigían la traslocación de los polipéptidos al periplasma de *E. coli*. La transcripción se terminó usando un terminador de transcripción doble *rrnB t1t2*. El gen *lacIq* codificaba la proteína represora Lac I expresada de forma constitutiva. Esto reprimía la transcripción a partir del promotor *tac* hasta que se inducía la desrepresión por la presencia de alolactasa o IPTG. El origen de replicación usado era p15A, que mantenía un número de copias bajo. El plásmido contenía un gen de resistencia a tetraciclina para la selección por antibiótico.

La transformación de las cepas se llevó a cabo usando 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 - Expresión de un Fab' anti-TNF α en cepas de *E. coli* mutadas usando cultivos en matraces agitados

10 Las cepas de spr mutante MXE008, MXE009, MXE012 y MXE017 se ensayaron en un experimento de matraz agitado comparando el crecimiento y expresión de un Fab' anti-TNF α frente a W3110 y MXE001.

El protocolo experimental de matraz agitado usado se llevó a cabo como sigue:

Experimento en matraz agitado de 5 ml

15 Se recogió una colonia individual en 5 ml de LB más tetraciclina 10 $\mu\text{g/ml}$ y se cultivó durante la noche a 30°C con agitación a 250 rpm.

El cultivo de la noche se usó para inocular 100 ml más tetraciclina hasta DO600 de 0,1 (es decir para DO de 4, $100/4 \times 0,1 = 2,5$ ml en 100 ml).

Se usaron tubos de cultivo 3x5 ml para cada punto de medición necesario usando este cultivo primario. Se usó 1 cultivo de referencia para ensayar la medición de DO.

20 Los cultivos se agitaron a 30°C a 250 rpm siguiendo el crecimiento primero visualmente y después tomando muestra del cultivo de referencia para tomar los cultivos a DO600 0,5 (normalmente aproximadamente 2 h). Se añadió IPTG a cada tubo de cultivo hasta una concentración 200 μM (25 μl de 0,04 M) una vez que el cultivo había alcanzado una DO mayor que 0,5.

25 Los tubos de cultivo se retiraron en los tiempos de medición requeridos, p. ej. 1h, 2 h, después de inducción y se mantuvieron sobre hielo.

30 Después de centrifugación a 13.200 rpm durante 5 min, el sedimento celular se volvió a suspender en 200 μl de tampón de extracción periplásmico (Tris, Cl 100 mM/EDTA 10 mM pH 7,4). Los extractos periplásmicos se agitaron a 250 rpm durante la noche a 30°C. Al día siguiente, los extractos se centrifugaron durante 10 min a 13.200 rpm, el líquido sobrenadante se separó por decantación y se almacenó a -20°C como "extracto periplásmico". El sedimento celular agotado se descartó.

Cuantificación por ELISA

35 Placas de ELISA de 96 pocillos se revistieron durante la noche a 4°C con AB141 (CH1 de conejo anti-humano, UCB) a 2 $\mu\text{g}\cdot\text{ml}^{-1}$ en PBS. Después de lavar 3x con 300 μl de tampón de muestra/conjugado (PBS, BSA al 0,2% (p/v), Tween 20 al 0,1% (v/v)), se llevaron a cabo diluciones seriadas de $\frac{1}{2}$ de las muestras y las referencias en la placa en 100 μl de tampón de muestra/conjugado, y la placa se agitó a 250 rpm a temperatura ambiente durante 1 h. Después de lavar 3x con 300 μl de tampón de lavado (PBS, Tween 20 al 0,1% (v/v)), se añadieron 100 μl del anticuerpo de revelado 6062 (anticuerpo de conejo anti-IgG kappa humana conjugado con HRP, The Binding Site, Birmingham, Reino Unido), después de dilución 1/1000 en tampón de muestra/conjugado. Después la placa se agitó a 250 rpm a temperatura ambiente durante 1 h. Después de lavar con 3x 300 μl de tampón de lavado, se añadieron 40 100 μl de sustrato MB (mezcla 50:50 de solución de TMB (Calbiochem): dH₂O) y se registró la A_{630} usando un lector de placa automático. La concentración de Fab' en los extractos periplásmicos se calculó por comparación con las referencias de Fab' purificadas del isotipo adecuado.

La figura 1 muestra el crecimiento mejorado de MXE008 y MXE009 comparado con W3110 de tipo natural y MXE001.

45 La figura 2 muestra la expresión mejorada de Fab' a partir de cepas MXE008 y MXE009 comparado con W3110 de tipo natural y MXE001.

La figura 5 muestra el crecimiento mejorado de MXE0012 y MXE017 comparado con W3110 de tipo natural y MXE001.

50 La figura 6 muestra la expresión mejorada de Fab' a partir de cepas MXE001 y MXE017 comparado con W3110 de tipo natural y MXE001.

Ejemplo 6 - Crecimiento de cepas de *E. coli* con spr mutado y expresión de Fab' en cepas de *E. coli* mutadas usando fermentaciones de alta densidad

Las cepas MXE008, MXE009, MXE001 y células W3110 de tipo natural se transformaron con el plásmido pMXE117 ensayado en experimentos de fermentación comparando el crecimiento y la expresión de un Fab' anti-TNF α .

5 Medio de crecimiento

El medio de crecimiento de fermentación se basó en el medio SM6E (descrito en Humphreys et al., 2002, "Protein Expression and Purification", 26, 309-320) con NaH₂PO₄·H₂O 3,86 g/l y glicerol 112 g/l.

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

- 10 Fermentación. Los fermentadores (volumen total de 2,5 litros) se sembraron con cultivo de inóculo a DO₆₀₀ 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 30% de saturación del aire mediante agitación variable y flujo de aire. El pH del cultivo se controló a 7,0 mediante valoración automática con NH₄OH al 15% (v/v) y H₂SO₄ conc. al 10% (v/v). La formación de espuma se controló mediante la adición de solución Struktol J673 al 10% (v/v)
- 15 (Schill and Seilacher). Se hicieron una serie de adiciones en diferentes etapas de la fermentación. Cuando la concentración de biomasa alcanzó aproximadamente DO₆₀₀ 40, se añadieron sales de magnesio y NaH₂PO₄·H₂O. Se hicieron adiciones adicionales de NaH₂PO₄·H₂O antes de 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 reducido (aproximadamente DO₆₀₀ 75) se aplicó una alimentación continua de glicerol al 80% (p/p). En el mismo punto en la
- 20 fermentación, se aplicó una alimentación de IPTG 170 μ M. El inicio de la alimentación de IPTG se consideró como el inicio de la inducción. Las fermentaciones típicamente se llevaron a cabo durante 64-120 horas a velocidades de alimentación de glicerol en el intervalo entre 0,5 y 2,5 ml/h.

Medición de la concentración de biomasa y velocidad de crecimiento. La concentración de biomasa se determinó midiendo la densidad óptica de los cultivos a 600 nm.

- 25 Extracción periplásmica. Las células se recogieron de las muestras de cultivo por centrifugación. La fracción de líquido sobrenadante se retuvo (a -20°C) para el posterior análisis. La fracción de sedimento celular se volvió a suspender al volumen de cultivo original en tampón de extracción (Tris-HCl 100 mM, EDTA 10 mM; pH 7,4). Después de incubación a 60°C durante aproximadamente 16 horas, el extracto se clarificó por centrifugación y la fracción de líquido sobrenadante se retuvo (a -20°C) para el análisis.
- 30 Cuantificación de Fab'. Las concentraciones de Fab' en los extractos periplásmicos y líquidos sobrenadantes de cultivos se determinó mediante ELISA de ensamblaje de Fab' como se describe en Humphreys et al., 2002, "Protein Expression and Purification", 26, 309-320, y usando HPLC de proteína G. Una columna HiTrap Protein-G HP de 1 ml (GE-Healthcare o equivalente) se cargó con el analito (pH aproximadamente neutro, 30°C, filtrado 0,2 μ m) a 2 ml/min, la columna se lavó con fosfato 20 mM, NaCl 50 mM a pH 7,4 y después se eluyó el Fab' usando una
- 35 inyección de glicina 50 mM /HCl a pH 2,7. El Fab' eluido se midió por A280 en un sistema de HPLC Agilent 1100 o 1200 y se cuantificó por referencia a una curva patrón de una proteína Fab' purificada de concentración conocida.

- La figura 3 muestra el perfil de crecimiento de MXE008 y MXE009 comparado con W3110 y MXE001 de control, que muestra que los perfiles de crecimiento son sustancialmente los mismos para MXE001, MXE008 y MXE009 a lo largo de las primeras ~26 horas y son todos mayores comparados con W3110. Después de ~26 horas la velocidad de crecimiento disminuye para MXE001 debido a la lisis celular. Sin embargo, las cepas de células de spr mutante MXE008 y MXE009 continúan mostrando una buena velocidad de crecimiento y no se produce lisis después de 26 horas.

- La figura 4 muestra el rendimiento total de Fab' del periplasma (símbolos sombreados) y del líquido sobrenadante (símbolos blancos) de cepas de *E. coli* MXE008 y MXE009 comparado con W3110 y MXE001 de control. Las cepas MXE008 y MXE009 muestran mayor expresión periplásmica de Fab' comparada con MXE001 y W3110. Además, MXE001 y también MXE008 y MXE009 muestran niveles menores de Fab' en el líquido sobrenadante comparados con MXE001, que muestra lisis celular reducida en MXE008 y MXE009 comparada con MXE001.

- La figura 7 muestra el perfil de crecimiento de MXE001 y MXE008 durante una fermentación de producción de Fab'. Los datos ilustran un pequeño aumento de la velocidad de crecimiento inicial de la cepa MXE008 con mutante spr Δ tsp respecto a la cepa MXE001 Δ tsp durante la acumulación de biomasa y una supervivencia significativamente muy aumentada de MXE008 con respecto a la cepa MXE001 en las últimas ~20 horas de fermentación.

- La figura 8 muestra la acumulación periplásmica de Fab' (líneas y símbolos negros) y la acumulación de Fab' en el medio (líneas de trazos y símbolos blancos) para W3110, MXE001 (Δ tsp) y MXE008 (Δ tsp spr D112A) durante una fermentación de producción de Fab'. Los datos ilustran un pequeño aumento en la acumulación periplásmica de Fab' inicial para la cepa MXE008 con respecto a la cepa MXE001, que se vuelve más pronunciado durante la segunda mitad de la fase de acumulación de Fab'. La adición de la mutación de spr en MXE008 a la mutación Δ tsp en

MXE001 contrarresta sustancialmente el fenotipo de "pérdida" observado con las cepas MXE001 con Δ tsp. Esta característica mejorada produce un mayor rendimiento periplásmico para la cepa MXE008 con respecto a la cepa MXE001 y menor acumulación del Fab' perdido en el medio de cultivo para la cepa MXE008 con respecto a la cepa MXE001.

5 Ejemplo 7 - Determinación de la pérdida de ADN y cantidad de proteínas totales en las cepas

Ensayo de ADNbc

10 La pérdida de ADN bicatenario en el líquido sobrenadante de las cepas W3110, MXE001, MXE008 y MXE012 se determinó usando el kit de ensayo de ADNbc Quant-IT Picogreen (Invitrogen, Ref: P11496). Se preparó una curva patrón diluyendo el patrón de ADN proporcionado en el intervalo de 1-1000 ng/ml. Las muestras se diluyeron en tampón de TE, de modo que la lectura de fluorescencia estuviera en el intervalo lineal del método (de 500 a 1000 veces). En una placa de 96 pocillos, se mezclaron 100 μ l de muestra diluida o de patrón con 100 μ l del reactivo Picogreen, y la placa se incubó durante 5 minutos a temperatura ambiente, protegida de la luz. Los recuentos de fluorescencia se midieron durante 0,1 s usando un filtro de excitación a 485 nm, y un filtro de emisión a 535 nm. Los resultados se muestran en la figura 9.

15 Ensayo de proteínas:

20 La concentración de proteínas totales de las cepas W3110, MXE001, MXE008 y MXE012 se determinó usando el kit de ensayo Coomassie Plus Bradford (Pierce, Ref: 23236). Se hizo una curva patrón diluyendo patrón de albúmina de suero bovino en un intervalo de 25-1000 μ g/ml. Las muestras se diluyeron en agua, de modo que la densidad óptica estuviera dentro del intervalo lineal del método (de 5 a 10 veces), y se mezclaron 33 μ l de la muestra o el patrón con 1 ml de reactivo de Coomassie. Después de incubar durante 10 min a temperatura ambiente, se leyó la DO_{595} nm en un espectrofotómetro con reactivo de Coomassie como blanco. La concentración de proteínas totales se calculó basándose en la curva patrón. Los resultados se muestran en la figura 10.

25 Aunque esta invención se ha mostrado y descrito en particular con referencia a las realizaciones preferidas, los expertos en la técnica entenderán que se pueden hacer diferentes cambios en la forma y el detalle sin salirse del alcance de la invención como se define en las reivindicaciones adjuntas.

LISTADO DE SECUENCIAS

<110> UCB PHARMA S.A.

5 <120> CEPA HOSPEDANTE BACTERIANA

<130> G0104-WO

<150> GB1000587.4

10 <151> 14-01-2010

<160> 37

<170> PatentIn version 3.5

15

<210> 1

<211> 2049

<212> ADN

<213> E. coli

20

<400> 1

atgaacatgt tttttaggct taccgcgta gctggcctgc ttgcaatagc aggccagacc	60
ttcgcctgtag aagatatcac gcgtgctgat caaattccgg tattaaagga agagacgcag	120
catgcgcagc taagtgcgcg cgtaacgtcg cgcttcaccc gttctcatta tcgccagttc	180
gacctcgatc aggcattttc ggccaaaatc tttgaccgct acctgaatct gctcgattac	240
agccacaacg tgctgctggc aagcgcgtgt gaacagttcg cgaaaaagaa aaccgcagtt	300
ggcgcagtaac tgcgcttcagg caaactcgac gttttctacg atctctacaa tctggcgcaa	360
aagcgcgctt ttgagcgta ccagtcgct ttgtcggtag tggaaaagcc gatggatttc	420
accggcaacg acacttataa ccttgaccgc agcaaagcgc cctggccgaa aaacgcaggt	480
gagttgaacg cgctgtggga cagtaaagtc aaattcgacg agttaagcct gaagctgaca	540
ggaaaaacgg ataaagaaat tcgtgaaacc ctgactcgcc gctacaaatt tgccattcgt	600
cgctcggcgc aaaccaacag cgaagatggt ttctcgcctg caatgcagcc gtttgcgct	660
gaaatcgacc cgcataccaa ctatctttcc ccgcgtaata ccgaacagtt caaactgaa	720
atgagtttgt cgctggaagg tattggcgca gtgctgcaaa tggatgatga ctacaccggt	780
atcaattcga tgggtggcagg tgggtccgca gcgaagagta aagctatcag cgttggtag	840
aaaattgtcg gtgttggta aacaggcaag ccgatggtg acgtgattgg ctggcgtctt	900
gatgatgtgg ttgccttaat taaagggccg aagggcagta aagttcgtct ggaaatttta	960
cctgctggtg aagggaccaa gaccctact gtaacgttga cccgtgaacg tattcgtctc	1020
gaagaccgcg cggttaaaat gtcgggtaag accgtcggta aagagaaagt cggcgtgctg	1080
gatattccgg gcttctatgt gggtttgaca gacgatgtca aagtgcaact gcagaaactg	1140
gaaaaacaga atgtcagcag cgtcatcatc gacctgcgta gcaatggcgg tggggcggtta	1200

ES 2 539 931 T3

actgaagccg t atcgtctctc cggctctgttt attcctgagg gtcccattgt tcaggtccgc 1260
gataacaacg gcaaggttcg tgaagatagc gataccgacg gacaggtttt ctataaaggc 1320
ccgctgggtg tgctggttga ccgcttcagt gcttcggcct cagaaatctt tgccgaggca 1380
atgcaggatt acggctcgtg gctggttg ggtgaaccga cgtttggtta aggcaccgtt 1440
cagcaatacc gtccattgaa ccgtatttac gatcagatgt tacgtcctga atggccagcg 1500
ctgggttctg tgcagtacac gatccagaaa ttctatcggc ttaacggcgg cagtacgcaa 1560
cgtaaaggcg taacgccaga catcatcatg ccgacgggta atgaagaaac ggaaacgggt 1620
gagaaattcg aagataacgc gctgccgtgg gatagcattg atgccgagc ttatgtgaaa 1680
tcaggagatt taacggcctt tgaaccggag ctgctgaagg aacataatgc gcgtatcggc 1740
aaagatcctg agttccagaa catcatgaag gatatcggc gcttcaacgc tatgaaggac 1800
aagcgaata tcgtttctct gaattacgct gtgcgtgaga aagagaataa tgaagatgat 1860
gcgacgcgtc tggcgcgttt gaacgaacgc ttaaacgcg aagtaaac ggagttgaag 1920
aaactggatg atctaccgaa agattaccag gagccgatc cttatctgga tgagacggtg 1980
aatatcgcac tcgatctggc gaagcttgaa aaagccagac ccgcggaaca acccgctccc 2040
gtcaagtaa 2049

<210> 2
<211> 682
5 <212> PRT
<213> E. coli

<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

10

ES 2 539 931 T3

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

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

ES 2 539 931 T3

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

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

ES 2 539 931 T3

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

<211> 2048

5 <212> ADN

<213> E. coli

<400> 3

atgaattcgt ttttaggctt accgcgtag ctggcctgct tgcaatagca ggcagacat 60
taattgtaga agatatcacg cgtgctgac aaattccggt attaaaggaa gagacgcagc 120
atgcgacggt aagtgagcgc gtaacgtcgc gcttcaccgc ttctcattat cgcagttcg 180
acctcgatca ggcattttcg gccaaaatct ttgaccgcta cctgaatctg ctcgattaca 240
gccacaacgt gctgctggca agcgatgttg aacagttcgc gaaaaagaaa accgagttag 300
gcgatgaact gcgttcaggc aaactcgacg tttctacga tctctacaat ctggcgcaaa 360
agcgcgcttt tgagcgttac cagtagcgtt tgcggtact ggaaaagccg atggatttca 420
ccggcaacga cacttataac cttgaccgca gcaaagcgcc ctggccgaaa aacgaggctg 480
agttgaacgc gctgtgggac agtaaagta aattcgacga gttaagcctg aagctgacag 540
gaaaaacgga taagaaatt cgtgaaacc tgactcgccg ctacaaattt gccattcgtc 600
gtctggcgca aaccaacagc gaagatgttt tctcgtggc aatgacggcg tttgcgctg 660
aaatgaccc gcataccaac tatctttccc cgogtaatac cgaacagttc aactgaaa 720
tgagtttgtc gctggaaggt attggcgag tgctgcaaat ggatgatgac tacaccgtta 780
10 tcaattcgt ggtggcaggt ggtccggcag cgaagagtaa agctatcagc gttggtgaca 840

ES 2 539 931 T3

aaattgtcgg tgttggtcaa acaggcaagc cgatggttga cgtgattggc tggcgtcttg 900
atgatgtggt tgccttaatt aaagggccga agggcagtaa agttcgtctg gaaattttac 960
ctgctggtaa agggaccaag acccgtactg taacgttgac cctggaacgt attcgtctcg 1020
aagaccgcgc ggttaaaatg tgggtgaaga ccgtcggtaa agagaaagtc ggcgtgctgg 1080
atattccggg cttctatgtg ggtttgacag acgatgtcaa agtgcaactg cagaaactgg 1140
aaaaacagaa tgtcagcagc gtcacatcgc acctgcgtag caatggcggg ggggcgtaa 1200
ctgaagccgt atcgtctctc ggtctgttta ttctcgggg tcccattggt caggctccgcg 1260
ataacaacgg caaggttctg gaagatagcg ataccgacgg acaggttttc tataaaggcc 1320
cgctggtggt gctggttgac cgcttcagtg ctccggcttc agaaatcttt gccgcggcaa 1380
tgcaggatta cggctgtgcg ctggttggtg gtgaaccgac gtttggtaaa ggcacgcttc 1440
agcaataccg 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 ctccaacgct atgaaggaca 1800
agcgcaatat cgtttctctg aattacgctg tgcgtgagaa agagaataat gaagatgatg 1860
cgacgcgtct ggcgcgtttg aacgaacgct ttaaaccgca aggtaaaccg gagttgaaga 1920
aactggatga tctaccgaaa gattaccagg agccggatcc 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 ttggtgtag ttgcctttg ggcaccctta 60
agtcaggcag aaacgggatg gcagccgatt caggaaacca tccgtaaaag tgataaagat 120
aaccgccagt atcaggctat acgtctggat aacggtatgg tggctctgct ggtttctgat 180
ccgcaggcag ttaaactcgt ctccgcgctg gtggtgcccg ttgggtcgct ggaagatccc 240
gaggcgtacc aggggctggc acattacctt gaacatatga gtctgatggg gtcgaaaag 300
taccgcgagg ctgacagtct ggccgaatat ctcaaatgc acggcggtag tcacaatgcc 360

10

ES 2 539 931 T3

agcactgcgc cgtatcgcac ggctttctat ctggaagttg agaacgacgc cttgcctggt 420
 gcggtagacc gcctggccga tgctattgct gaacctttgc tcgacaagaa atatgcogaa 480
 cgtgagcgtg atgcggtgaa cgctgaatta accatggcgc gtacgcgtga cgggatgcgc 540
 atggcacagg tcagcgcaga aaccattaac cgggcacacc cgggttcaaa gttttctggt 600
 ggtaacctcg aaactttaag cgacaaacct ggtaatccgg tgcagcaggc gctgaaagat 660
 ttccacgaga agtactatc cgccaatttg atgaagcggg ttattttacag taataaacgg 720
 ctgcgggagt tggcaaaaat ggcggcggac acctttggtc gcgtgccgaa caaagagagc 780
 aaaaaccgg aaatcaccgt gccggtagtc accgacgcgc aaaagggcat tatcattcat 840
 tacgtccctg cgctgccgag taaagtgttg cgcggtgagt ttcgcatcga taacaactca 900
 gcgaagttcc gtagtaaac cgatgaattg attacctatc tgattggcaa tcgcagccca 960
 ggtacacttt ctgactggct gcaaaagcag ggattagttg agggcattag cgccaactcc 1020
 gatcctatcg tcaacggcaa cagcggcgta ttagcgatct ctgcgtcttt aaccgataaa 1080
 ggctgggcta atcgcgatca ggttgtggcg gcaattttta gctatctcaa tctgttacgt 1140
 gaaaagggca ttgataaaca atacttcgat gaactggcga atgtgctgga tatcgacttc 1200
 cgttatccgt cgatcaccgg tgatatggat tacgtcgaat ggctggcaga taccatgatt 1260
 cgcgttcctg ttgagcatac gctggatgca gtcaatattg ccgatcggta cgatgctaaa 1320
 gcagtaaagg aacgtctggc gatgatgacg ccgcagaatg cgcgtatctg gtatatcagc 1380
 ccgaaagagc cgcacaacaa aacggcttac tttgtcgatg cgccgtatca ggtcgataaa 1440
 atcagcgcac aaactttcgc cgactggcag aaaaaagccg ccgacattgc gctctctttg 1500
 ccagagctta acccttatat tctctgatgat ttctcgtga ttaagtcaga gaagaatac 1560
 gaccatccag agctgattgt tgatgagtcg aatctgcgag tgggtgatgc gccaaagccgt 1620
 tattttgcaa gcgagcccaa agctgatgtc agcctgattt tgcgtaatcc gaaagccatg 1680
 gacagcggcc gcaatcaggt gatgtttgag ctcaatgatt atctcgcagg gctggcggctt 1740
 gatcagttaa gcaaccaggc gtcggttggt ggcataagtt tttccaccaa cgctaacaac 1800
 ggccttatgg ttaatgctaa tggttacacc cagcgtctgc cgcagctggt ccaggcattg 1860
 ctcgaggggt acttttagcta taccgctacg gaagatcagc ttgagcaggc gaagtctctg 1920
 tataaccaga tgatggattc cgcagaaaag ggtaaagcgt ttgagcaggc gattatgcc 1980
 gcgcagatgc tctcgaagt gccgtacttc tcgcgagatg aacggcgtaa aattttgcc 2040
 tccattacgt tgaaagaggt gctggcctat cgcgacgcct taaaatcagg ggctcgacca 2100
 gagtttatgg ttatcgcaa catgaccgag gcccaggcaa caacgctggc acgogatgtg 2160
 caaaaacagt tgggcgctga tggttcagag tgggtcogaa acaaagatgt agtggctgat 2220

ES 2 539 931 T3

aaaaaacaat ccgatcatctt tgaaaaagcc ggtaacagca cggactccgc actggcagcg 2280
 gtatttgtac cgactggcta cgatgaatac accagctcag cctatagctc tctgttgggg 2340
 cagatcgtac agccgtgggt ctacaatcag ttgcgtaccg aagaacaatt gggctatgcc 2400
 gtgtttgcgt ttccaatgag cgtggggcgt cagtggggca tgggcttcct tttgcaaagc 2460
 aatgataaac agccttcatt cttgtgggag cgttacaagg cgtttttccc aaccgcagag 2520
 gcaaaattgc gagcgtgaa gccagatgag tttgcgcaa tccagcaggc ggtaattacc 2580
 cagatgctgc aggcaccgca aacgctcggc gaagaagcat cgaagttaag taaagatttc 2640
 gatcgcggca atatgcgctt cgattcgcgt gataaaatcg tggcccagat aaaactgctg 2700
 acgccgcaaa aacttgctga tttcttccat caggcgtggg tcgagccgca aggcattggc 2760
 attctgtcgc agatttccg cagccagaac gggaaagccg aatatgtaca ccctgaaggc 2820
 tggaaagtgt gggagaacgt cagcgcggtg cagcaacaa tgcccctgat gagtgaaaag 2880
 aatgagtga 2889

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

5

<400> 5

Met Pro Arg Ser Thr Trp Phe Lys Ala 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

10

ES 2 539 931 T3

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

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

ES 2 539 931 T3

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 500 505 510		
Leu Ile Lys Ser Glu Lys Lys Tyr Asp His Pro Glu Leu Ile Val Asp 515 520 525		
Glu Ser Asn Leu Arg Val Val Tyr Ala Pro Ser Arg Tyr Phe Ala Ser 530 535 540		
Glu Pro Lys Ala Asp Val Ser Leu Ile Leu Arg Asn Pro Lys Ala Met 545 550 555 560		
Asp Ser Ala Arg Asn Gln Val Met Phe Ala Leu Asn Asp Tyr Leu Ala 565 570 575		
Gly Leu Ala Leu Asp Gln Leu Ser Asn Gln Ala Ser Val Gly Gly Ile 580 585 590		
Ser Phe Ser Thr Asn Ala Asn Asn Gly Leu Met Val Asn Ala Asn Gly 595 600 605		

ES 2 539 931 T3

Tyr Thr Gln Arg Leu Pro Gln Leu Phe Gln Ala Leu Leu Glu Gly Tyr
610 615 620

Phe Ser Tyr Thr Ala Thr Glu Asp Gln Leu Glu Gln Ala Lys Ser Trp
625 630 635 640

Tyr Asn Gln Met Met Asp Ser Ala Glu Lys Gly Lys Ala Phe Glu Gln
645 650 655

Ala Ile Met Pro Ala Gln Met Leu Ser Gln Val Pro Tyr Phe Ser Arg
660 665 670

Asp Glu Arg Arg Lys Ile Leu Pro Ser Ile Thr Leu Lys Glu Val Leu
675 680 685

Ala Tyr Arg Asp Ala Leu Lys Ser Gly Ala Arg Pro Glu Phe Met Val
690 695 700

Ile Gly Asn Met Thr Glu Ala Gln Ala Thr Thr Leu Ala Arg Asp Val
705 710 715 720

Gln Lys Gln Leu Gly Ala Asp Gly Ser Glu Trp Cys Arg Asn Lys Asp
725 730 735

Val Val Val Asp Lys Lys Gln Ser Val Ile Phe Glu Lys Ala Gly Asn
740 745 750

Ser Thr Asp Ser Ala Leu Ala Ala Val Phe Val Pro Thr Gly Tyr Asp
755 760 765

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

ES 2 539 931 T3

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

5

<400> 6

```

attccccgca gcacctggtt caaagcatta ttgttgtag ttgcccttg ggcacattaa      60
tgtcaggcag aaacgggatg gcagccgatt caggaaacca tccgtaaaag tgataaagat      120
aaccgccagt atcaggctat acgtctggat aacggtatgg tggcttgct ggtttctgat      180
ccgcaggcag ttaaatoctt ctcgcgctg gtggtgcccg ttgggtcgct ggaagatccc      240
gaggcgtacc aggggctggc acattacctt gaacatatga gtctgatggg gtcgaaaaag      300
taccgcagg ctgacagtct ggccgaatat ctcaaatgc acggcggtag tcacaatgcc      360
agcactgctc cgtatcgcac ggctttctat ctggaagttg agaacgacgc cttgcctggt      420
gcggtagacc gcctggccga tgctattgct gaacctttgc tcgacaagaa atatgccgaa      480
cgtgagcgtg atgcggtgaa cgctgaatta accatggcgc gtacgcgtga cgggatgcgc      540
atggcacagg tcagcgcaga aaccattaac ccggcacacc ccggttcaaa gttttctggt      600
ggtaacctcg aaactttaag cgacaaacct ggtaatccgg tgcagcaggc gctgaaagat      660

```

10

ES 2 539 931 T3

ttccacgaga agtactattc cgccaatttg atgaaggcgg ttatttacag taataaacgg 720
 ctgccggagt tggcaaaaat ggcggcggac acctttggtc gcgtgccgaa caaagagagc 780
 aaaaaaccgg aaatcacccg gccggtagtc accgacgcgc aaaagggcat tatcattcat 840
 tacgtccctg cgctgcccg taaagtgttg cgcgttgagt ttcgcatcga taacaactca 900
 gcgaagtcc gtagtaaac cgatgaattg attacctatc tgattggcaa tcgcagccca 960
 ggtacacttt ctgactggct gcaaaagcag ggattagttg agggcattag cgccaactcc 1020
 gatcctatcg tcaacggcaa cagcggcgta ttagcgatct ctgcgtcttt aaccgataaa 1080
 ggcctggcta atcgcgatca ggttgtggcg gcaattttta gctatctcaa tctgttacgt 1140
 gaaaaaggca ttgataaaca atacttcgat gaactggcga atgtgctgga tatcgacttc 1200
 cgttatccgt cgatcacccg tgatatggat tacgtcgaat ggctggcaga taccatgatt 1260
 cgcgttcctg ttgagcatac gctggatgca gtcaatattg ccgatcggta cgatgctaaa 1320
 gcagtaaagg aacgtctggc gatgatgacg ccgcagaatg cgcgtatctg gtatatcagc 1380
 ccgaaagagc cgcacaacaa aacggcttac tttgtcgatg cgccgtatca ggtcgataaa 1440
 atcagcgcac aaactttcgc cgactggcag aaaaaagccg ccgacattgc gctctctttg 1500
 ccagagctta acccttatat tcctgatgat ttctcgctga ttaagtcaga gaagaatac 1560
 gaccatccag agctgattgt tgatgagtcg aatctgcgcg tgggtgatgc gccaaagcgt 1620
 tattttgccg gcgagcccaa agctgatgtc agcctgattt tgcgtaatcc gaaagccatg 1680
 gacagcgcgc gcaatcaggt gatgtttgcg ctcaatgatt atctcgcagg gctggcgctt 1740
 gatcagttaa gcaaccaggc gtcggttggg ggcataagtt tttccaccaa cgctaacaac 1800
 ggccttatgg ttaatgctaa tggttacacc cagcgtctgc cgcagctggt ccaggcattg 1860
 ctcgaggggt actttagcta taccgctacg gaagatcagc ttgagcagge gaagtcctgg 1920
 tataaccaga tgatggattc cgcagaaaag ggtaaaagcgt ttgagcagge gattatgccc 1980
 gcgcagatgc tctcgcaagt gccgtacttc tcgcgagatg aacggcgtaa aattttgccc 2040
 tccattacgt tgaaagaggt gctggcctat cgcgacgcct taaaatcagg ggctcgacca 2100
 gagtttatgg ttatcggcaa catgaccgag gcccaggcaa caacgctggc acgcgatgtg 2160
 caaaaacagt tgggcgctga tggttcagag tgggtgcgaa acaaagatgt agtggtcgat 2220
 aaaaaacaat ccgtcatctt tgaaaaagcc ggtaacagca ccgactccgc actggcagcg 2280
 gtatttgtag cgactggcta 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

ES 2 539 931 T3

gcaaaattgc gagcgatgaa gccagatgag tttgcgcaaa tccagcaggc ggtaattacc 2580
 cagatgctgc aggcaccgca aacgctcggc gaagaagcat cgaagttaag taaagatttc 2640
 gatcgcggca atatgcgctt cgattcgcgt gataaaatcg tggcccagat aaaactgctg 2700
 acgccgcaaa aacttgctga tttcttccat caggcgggtg tgcagccgca aggcattggct 2760
 attctgtcgc agatttccgg cagccagaac gggaaagccg aatatgtaca ccctgaaggc 2820
 tggaaagtgt gggagaacgt cagcgcgttg cagcaacaa tgcccctgat gaggtaaaag 2880
 aatgagtgat gtcgccgaga cactagatcc tttgc 2915

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

5

<400> 7

atgaaaaaaaa ccacattagc actgagtgca ctggctctga gtttaggttt ggcgttatct 60
 ccgctctctg caacggcggc tgagacttct tcagcaacga cagcccagca gatgccaaagc 120
 cttgcaccga tgctcgaaaa ggtgatgcct tcagtggtca gcattaacgt agaaggtagc 180
 acaaccgtta ataccccgcg tatgccgcgt aatttccagc agttcttcgg tgatgattct 240
 ccgttctgcc aggaaggttc tccgttccag agctctccgt tctgccaggg tggccagggc 300
 ggtaatggtg gcggccagca acagaaattc atggcgctgg gttccggcgt catcattgat 360
 gccgataaag gctatgtcgt caccaacaac cacgttgttg ataacgcgac ggtcattaaa 420
 gttcaactga gcgatggccg taagttcgac gcgaagatgg ttggcaaaga tccgcgctct 480
 gatatcgcgc tgatccaaat ccagaaccgg aaaaacctga ccgcaattaa gatggcggat 540
 tctgatgcac tgcgcgtggg tgattacacc gtagcgattg gtaaccogtt tggctctgggc 600
 gagacggtaa cttccgggat tgtctctgcg ctggggcgta gcggcctgaa tgccgaaaaac 660
 tacgaaaact tcatccagac cgatgcagcg atcaaccgtg gtaactccgg tgggtgcgctg 720
 gttaacctga acggcgaact gatcggatc aacaccgcga tcctcgcacc ggacggcggc 780
 aacatcggtg tcggttttgc tatcccagat aacatggtga aaaacctgac ctgcgagatg 840
 gtggaatacg gccaggtgaa acgcggtgag ctgggtatta tggggactga gctgaaactcc 900
 gaactggcga aagcgatgaa agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg 960
 cctaattcct ccgctgcaaa agcgggcatt aaagcgggtg atgtgatcac ctactgaac 1020
 ggtaagccga tcagcagctt tgccgcaact cgtgctcagg tgggtactat gccggtaggc 1080
 agcaaaactga ccctgggctt actgcccagc ggtaagcagg ttaacgtgaa cctggaactg 1140
 cagcagagca gccagaatca ggttgattcc agctccatct tcaacggcat tgaaggcgt 1200
 gagatgagca acaaaggcaa agatcagggc gtggtagtga acaacgtgaa aacgggcaact 1260
 ccggctgcgc agatcggcct gaagaaaggt gatgtgatta ttggcgcgaa ccagcaggca 1320
 gtgaaaaaca tcgctgaact gcgtaaaagt ctgcagacga aaccgtctgt gctggcactc 1380
 aacattcagc gcggcgacag caccatctac ctgttaatgc agtaa 1425

10

ES 2 539 931 T3

<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

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

ES 2 539 931 T3

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

Lys Thr Gly Thr Pro Ala Ala Gln Ile Gly Leu Lys Lys Gly Asp Val

ES 2 539 931 T3

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

5

<400> 9

atgaaaaaaa ccacattagc actgagtgca ctggctctga gtttaggttt ggcgttatct 60
 ccgctctctg caacggcggc tgagacttct tcagcaacga cagcccagca gatgccaagc 120
 cttgcaccga tgctcgaaaa ggtgatgcct tcagtggta gcattaacgt agaaggtagc 180
 acaaccgta ataccccgcg tatgcccggt aatttcagc agttcttcgg tgatgattct 240
 ccgttctgcc aggaaggttc tccgttcag agctctccgt tctgccaggg tggccagggc 300
 ggtaatggtg gcgccagca acagaaattc atggcgctgg gttccggcgt catcattgat 360
 gccgataaag gctatgtcgt caccaacaac cacgttgttg ataacgcgac ggtcattaaa 420
 gttcaactga gcgatggcgg taagtccgac gcgaagatgg ttggcaaga tccgcgctct 480
 gatatcgcg tgatccaaat ccagaacccg aaaacctga ccgcaattaa gatggcggat 540
 tctgatgcac tgccgctggg tgattacacc gtagcgattg gtaaccggt tggctctggc 600
 gagacggtaa cttccgggat tgtctctgcg ctggggcgta gccgcctgaa tgccgaaaa 660
 tacgaaaact tcatccagac cgatgcagcg attaatcgtg gtaacgccgg tggcgcgctg 720
 gttaacctga acggcgaact gatcggatc aacaccgga tcctcgacc ggacggcggc 780
 aacatcggtg tcggttttgc tatcccgagt aacatggtg aaaacctgac ctcgcagatg 840
 gtggaatacg gccaggtgaa acgcggtgag ctgggtatta tggggactga gctgaaactc 900
 gaactggcga aagcgatgaa agttgacgcc cagcgcggtg ctttcgtaag ccaggttctg 960
 cctaattcct ccgctgcaaa agcgggcatt aaagcgggtg atgtgatcac ctactgaac 1020
 ggtaagccga tcagcagctt tgccgactg cgtgctcagg tgggtactat gccggtaggc 1080
 agcaaaactg ccctgggctt actgcgcgac ggtaagcagg ttaacgtgaa cctggaactg 1140
 10 cagcagagca gccagaatca ggttgattcc agctccatct tcaacggcat tgaaggcgt 1200
 gagatgagca acaaaggcaa agatcagggc gtggtagtga acaacgtgaa aacgggcact 1260
 ccggctgcgc agatcggcct gaagaaaggat gatgtgatta ttggcgcgaa ccagcaggca 1320
 gtgaaaaaca tcgctgaact gcgtaaagtt ctgcacagca aaccgtctgt gctggcactc 1380
 aacattcagc gcggcgacag caccatctac ctgttaatgc agtaa 1425

ES 2 539 931 T3

<210> 10
 <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

ES 2 539 931 T3

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 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

ES 2 539 931 T3

<210> 11
 <211> 107
 <212> PRT
 <213> Secuencia Artificial

5

<220>
 <223> hTNF40-gL1

<400> 11

10

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

<210> 12
 <211> 118
 <212> PRT
 <213> Secuencia Artificial

15

<220>
 <223> gh3h TNF40.4

20

<400> 12

ES 2 539 931 T3

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
115

<210> 13

<211> 214

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> Cadena Ligera Injertada

10

<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

ES 2 539 931 T3

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

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
- 5 <212> PRT
- <213> Secuencia Artificial

- <220>
- <223> Cadena Pesada Injertada

- 10 <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

ES 2 539 931 T3

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			
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											

5 <210> 15
 <211> 24
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <223> Cebador oligonucleótido

<400> 15

gcatcataat ttctttta cctc 24

15 <210> 16
 <211> 24
 <212> ADN
 <213> Secuencia Artificial

20 <220>
 <223> Cebador oligonucleótido

<400> 16

gggaaatgaa cctgagcaaa acgc 24
 <210> 17
 <211> 24
 5 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Cebador oligonucleótido
 10 <400> 17
 gtgccaggag atgcagcagc ttgc 24
 15 <210> 18
 <211> 21
 <212> ADN
 <213> Secuencia Artificial
 20 <220>
 <223> Cebador oligonucleótido
 <400> 18
 25 ttgcagcca gtcagaaagt g 21
 <210> 19
 <211> 24
 <212> ADN
 30 <213> Secuencia Artificial
 <220>
 <223> Cebador oligonucleótido
 35 <400> 19
 ctgcctgcga tttcgccgg aacg 24
 40 <210> 20
 <211> 24
 <212> ADN
 <213> Secuencia Artificial
 <220>
 45 <223> Cebador oligonucleótido
 <400> 20
 cgcatgtac gtgccacgat atcc 24
 50 <210> 21
 <211> 188
 <212> PRT
 <213> Escherichia coli
 55 <400> 21

ES 2 539 931 T3

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

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
- <211> 162
- 5 <212> PRT
- <213> Escherichia coli

<400> 22

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

ES 2 539 931 T3

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
- <211> 951
- 5 <212> ADN
- <213> Secuencia Artificial

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

<400> 23

atgcgggcga aacttctggg aatagtctctg acaaccoccta ttgcgatcag ctcttttgct 60
tctaccgaga ctttatcgtt tactcctgac aacataaatg cggacattag tcttggaact 120
ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggaggccg aaaagtcagt 180
caactcgact ggaaattcaa taacgctgca attattaaag gtgcaattaa ttgggatttg 240
atgccccaga tatctatcgg ggctgctggc tggacaactc toggcagccg aggtggcaat 300

ES 2 539 931 T3

atggtcgatc aggactggat ggattccagt aacccccgaa cctggacgga tgaagtaga 360
cacctgata cacaactcaa ttatgccaac gaatttgatc tgaatatcaa aggctggctc 420
ctcaacgaac ccaattaccg cctgggactc atggccggat atcaggaaag ccgttatagc 480
tttacagcca gagtggttc ctatatctac agttctgagg agggattcag agatgatatc 540
ggctccttcc cgaatggaga aagagcaatc ggctacaaac aacgttttaa aatgccctac 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 aaatggagca 900
ggtatagaaa actataactt catcactact gctggtctta agtacacatt t 951

<210> 24

<211> 317

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> Secuencia OmpT mutada

10

<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

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

ES 2 539 931 T3

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

<210> 25
 <211> 954
 5 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Secuencia OmpT mutada

10 <400> 25

ES 2 539 931 T3

attcgggcga aacttctggg aatagtcctg acaaccoccta ttgcgatcag ctcttttgct 60
 tctaccgaga ctttatcggt tactcctgac aacataaatg cggacattag tcttggaact 120
 ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggaggccg aaaagtcagt 180
 caactcgact ggaaattcaa taacgctgca attattaaag gtgcaattaa ttgggatttg 240
 atgccccaga tatctatcgg ggctgctggc tggacaactc tcggcagccg aggtggcaat 300
 atggtcgatc aggactggat ggattccagt aaccccggaa cctggacgga tgaagtaga 360
 caccctgata cacaactcaa ttatgccaac gaatttgatc tgaatatcaa aggttggtc 420
 ctcaacgaac ccaattaccg cctgggactc atggccggat atcaggaaag ccgttatagc 480
 tttacagcca gaggtggttc ctatatctac agttctgagg agggattcag agatgatac 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 ttatgttgaa ggcgatgga atcgggttac gaataaaaaa 840
 ggtaatactt cactttatga tcacaataat aacacttcag actacagcaa aaatggagca 900
 ggtatagaaa actataactt catcactact gctggtctta agtacacatt ttaa 954

<210> 26

<211> 5

5 <212> PRT

<213> Secuencia Artificial

<220>

10 <223> Descripción de la Secuencia Artificial:hTNF40 CDRH1

<400> 26

Asp Tyr Gly Met Asn

1 5

15 <210> 27

<211> 17

<212> PRT

<213> Secuencia Artificial

20 <220>

<223> Descripción de la Secuencia Artificial:hTNF40 CDRH2 híbrida humana

<400> 27

Trp Ile Asn Thr Tyr Ile Gly Glu Pro Ile Tyr Ala Asp Ser Val Lys

25 1 5 10 15

Gly

30 <210> 28

<211> 9

<212> PRT

<213> Secuencia Artificial

<220>

35 <223> Descripción de la Secuencia Artificial:hTNF40 CDRH3

<400> 28

Gly Tyr Arg Ser Tyr Ala Met Asp Tyr
1 5

5 <210> 29
<211> 11
<212> PRT
<213> Secuencia Artificial

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

<400> 29

Lys Ala Ser Gln Asn Val Gly Thr Asn Val Ala
15 1 5 10

<210> 30
<211> 7
<212> PRT
20 <213> Secuencia Artificial

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

25 <400> 30

Ser Ala Ser Phe Leu Tyr Ser
1 5

30 <210> 31
<211> 9
<212> PRT
<213> Secuencia Artificial

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

35 <400> 31
Gln Gln Tyr Asn Ile Tyr Pro Leu Thr
1 5

<210> 32
<211> 17
<212> PRT
40 <213> Secuencia Artificial

<220>
<223> Descripción of Secuencia Artificial:hTNF40 CDRH2

45 <400> 32

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

Gly

50 <210> 33
<211> 84
<212> ADN
<213> Secuencia Artificial

<220>
55 <223> Adaptador oligonucleótido OmpA

<400> 33

ES 2 539 931 T3

tcgagttcta gataacgagg cgtaaaaaat gaaaaagaca gctatcgcaa ttgcagtggc 60
 cttggctctg acgtacgagt cagg 84

5 <210> 34
 <211> 67
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <223> casete-1 IGS
 <400> 34

gagctcacca gtaacaaaaa gttttaatag aggagagtgt taatgaagaa gactgctata 60
 gcaattg 67

15 <210> 35
 <211> 69
 <212> ADN
 <213> Secuencia Artificial

20 <220>
 <223> casete-2 IGS
 <400> 35

gagctcacca gtaacaaaaa gttttaatag aggggagtgt taaaatgaag aagactgcta 60
 tagcaattg 69

25 <210> 36
 <211> 81
 <212> ADN
 <213> Secuencia Artificial

30 <220>
 <223> casete-3 IGS

35 <400> 36

gagctcacca gtaacaaaaa gctttaatag aggagagtgt tgaggaggaa aaaaaaatga 60
 agaaaactgc tatagcaatt g 81

40 <210> 37
 <211> 81
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> casete-4 IGS
 <400> 37

gagctcacca gtaacaaaaa gttttaatag aggagagtgt tgacgaggat tatataatga 60
 agaaaactgc tatagcaatt g 81

50

REIVINDICACIONES

- 1.- Una célula bacteriana Gram negativa recombinante que comprende un gen spr mutante, que codifica una proteína spr UniprotKB/SwissProt P0AFV4, que tiene una mutación en uno o más aminoácidos seleccionada de N31Y, R62C, I70T, Q73R, C94A, S95F, V98E, Q99P, R100G, L108S, Y115F, D133A, V135D, V135G, L136P, G140C, R144C, H145A, G147C y H157A, y en donde la célula tiene actividad reducida de la proteína Tsp, proteasa específica de la cola, comparada con una célula de tipo natural.
- 5 2.- Una célula según la reivindicación 1, en donde el gen spr mutante codifica una proteína spr que tiene las mutaciones S95F y Y115F.
- 3.- Una célula según las reivindicaciones 1 o 2, en donde la célula comprende además uno o más de los siguientes genes mutados:
- 10 a) un gen DegP mutado que codifica una proteína DegP que tiene actividad de chaperona y actividad reducida de proteasa;
- b) un gen ptr mutado, en donde el gen ptr codifica una proteína proteasa III que tiene actividad reducida de proteasa o es un gen ptr mutado inactivado; y
- 15 c) un gen OmpT mutado, en donde el gen OmpT mutado codifica una proteína OmpT que tiene actividad reducida de proteasa o es un gen OmpT mutado inactivado.
- 4.- Una célula según la reivindicación 1, 2 o 3, en donde la célula comprende un gen Tsp mutado, proteasa específica de la cola, que codifica una proteína Tsp que tiene actividad reducida de proteasa o es un gen Tsp mutado inactivado.
- 20 5.- Una célula según la reivindicación 4, en donde el genoma de la célula es isogénico con una célula bacteriana de tipo natural excepto por el gen spr mutado y el gen Tsp mutado.
- 6.- Una célula según una cualquiera de las reivindicaciones 1 a 5, en donde el célula tiene un gen Tsp mutado inactivado, que comprende una mutación en el codón de inicio del gen y/o uno o más codones de parada situados en la dirección 3' del codón de inicio del gen y en la dirección 5' del codón de parada del gen.
- 25 7.- Una célula según la reivindicación 6, en donde el gen Tsp mutado inactivado comprende un sitio de restricción marcador creado por una mutación de aminoácido en el codón de inicio del gen y opcionalmente una o más mutaciones puntuales.
- 8.- Una célula según la reivindicación 7, en donde el gen Tsp mutado inactivado comprende la SEQ ID NO: 3.
- 9.- Una célula según una cualquiera de las reivindicaciones 1 a 8, en donde la célula es E. coli.
- 30 10.- Una célula según cualquier reivindicación precedente, 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 proteína de interés es un anticuerpo o un fragmento de unión al antígeno del mismo.
- 35 12.- Una célula según la reivindicación 11, en donde el anticuerpo o fragmento de unión al antígeno del mismo es específico para el TNF.
- 13.- Un método para producir una proteína de interés que comprende cultivar una célula bacteriana Gram negativa recombinante como se define en una cualquiera de las reivindicaciones 1 a 12, en un medio de cultivo en condiciones eficaces para expresar la proteína recombinante de interés y recuperar la proteína recombinante de interés del periplasma de la célula bacteriana Gram negativa recombinante y/o del medio de cultivo.
- 40 14.- Un método según la reivindicación 13, en donde el método comprende además recuperar la proteína de interés de la célula.
- 15.- Un método según la reivindicación 14, en donde la proteína de interés se recupera del periplasma y/o del líquido sobrenadante.

Figura 1

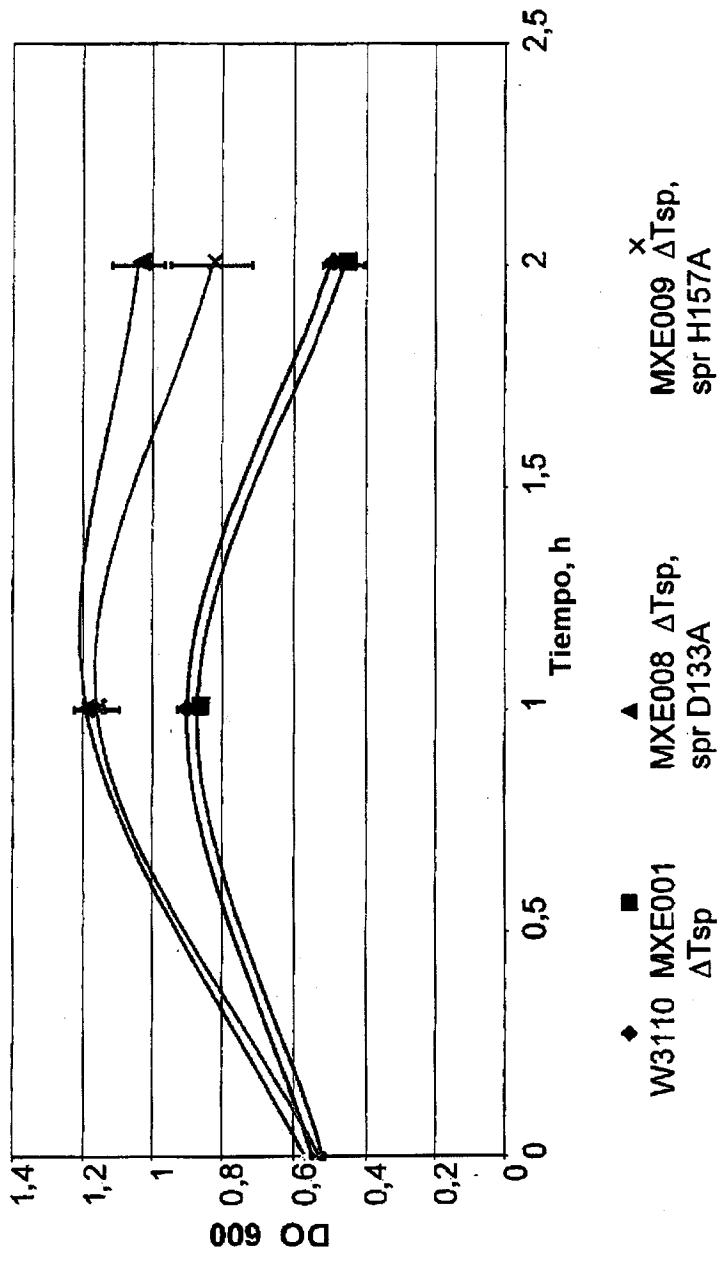


Figura 2

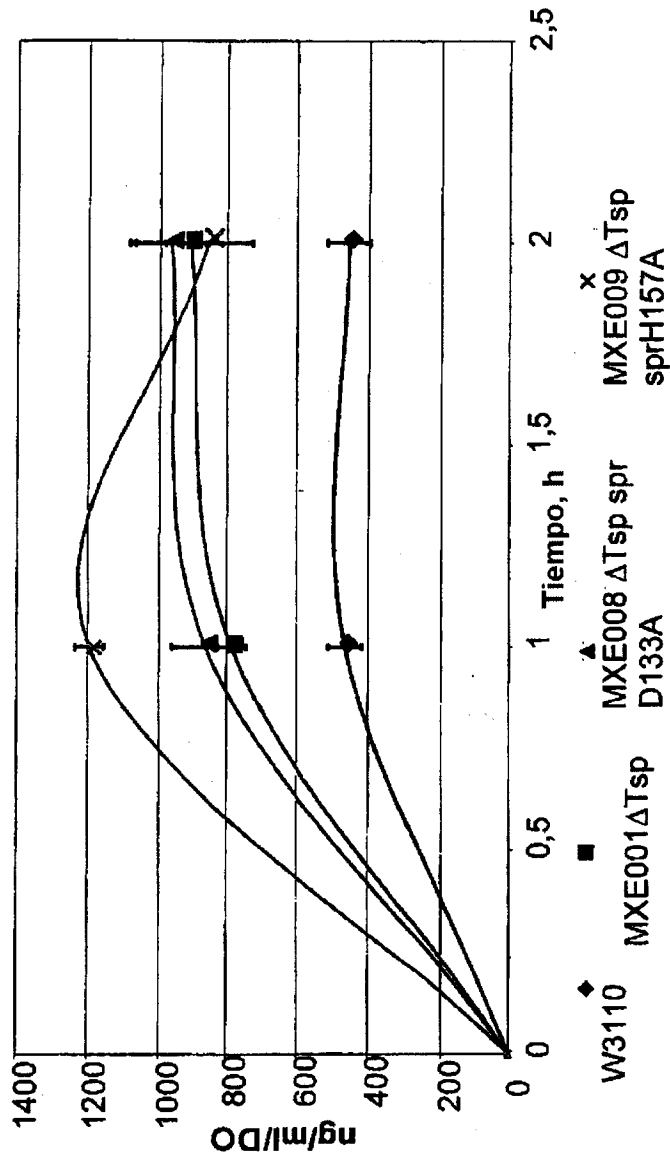


Figura 3

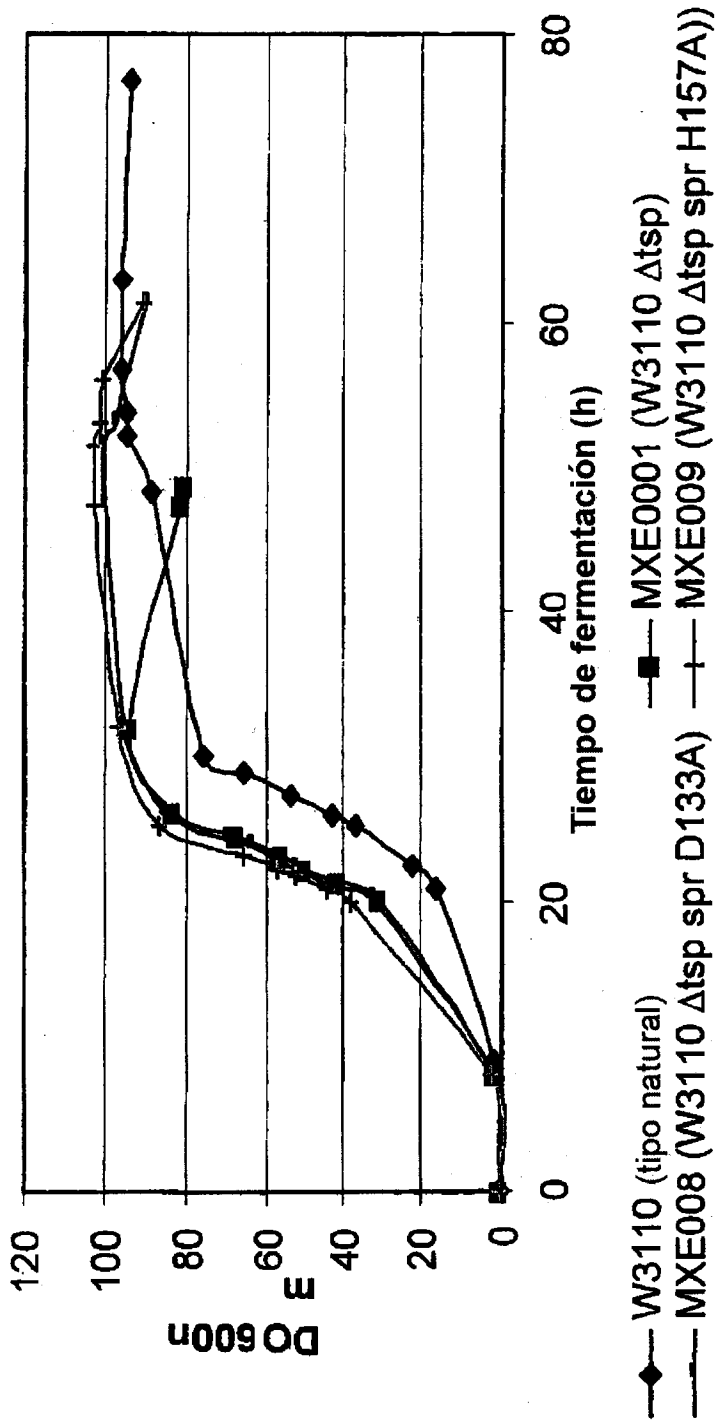


Figura 4

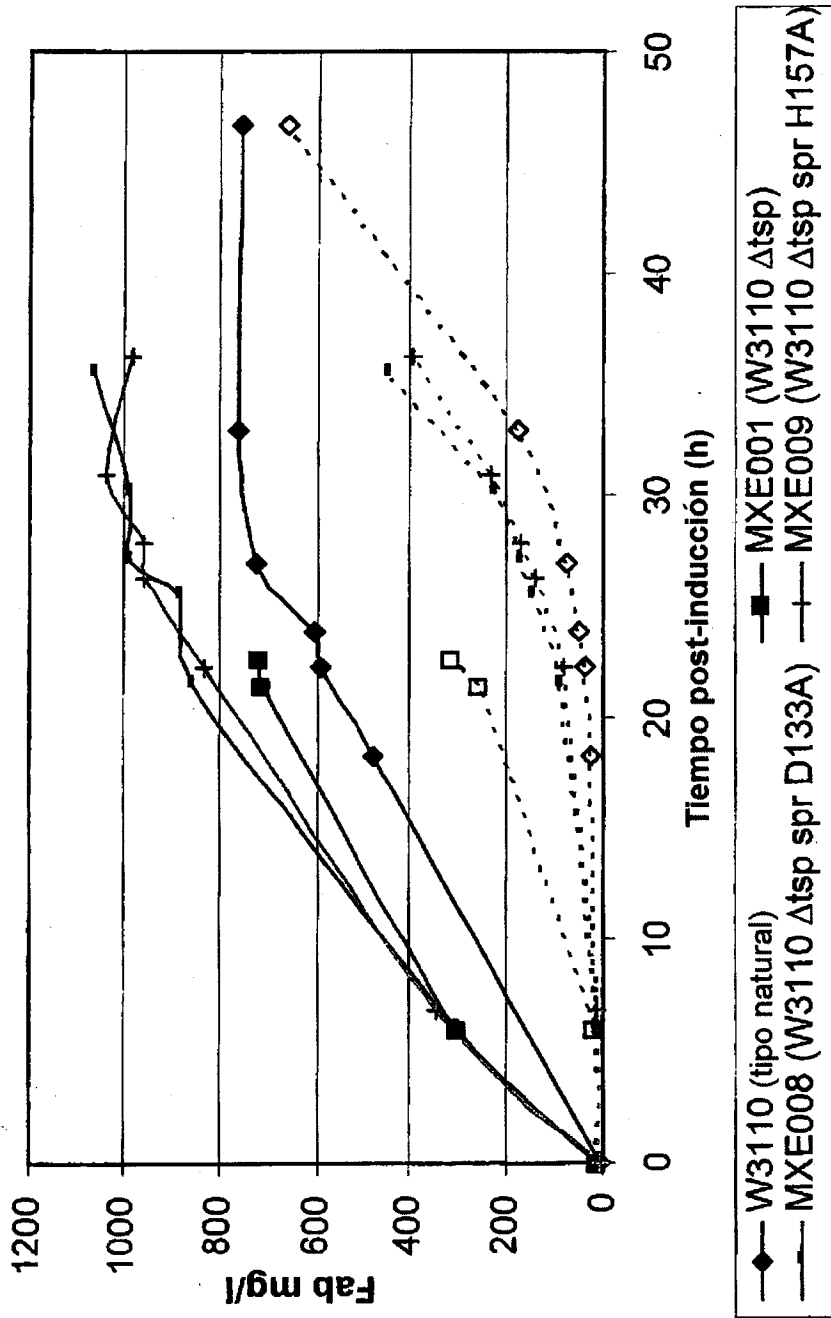


Figura 5

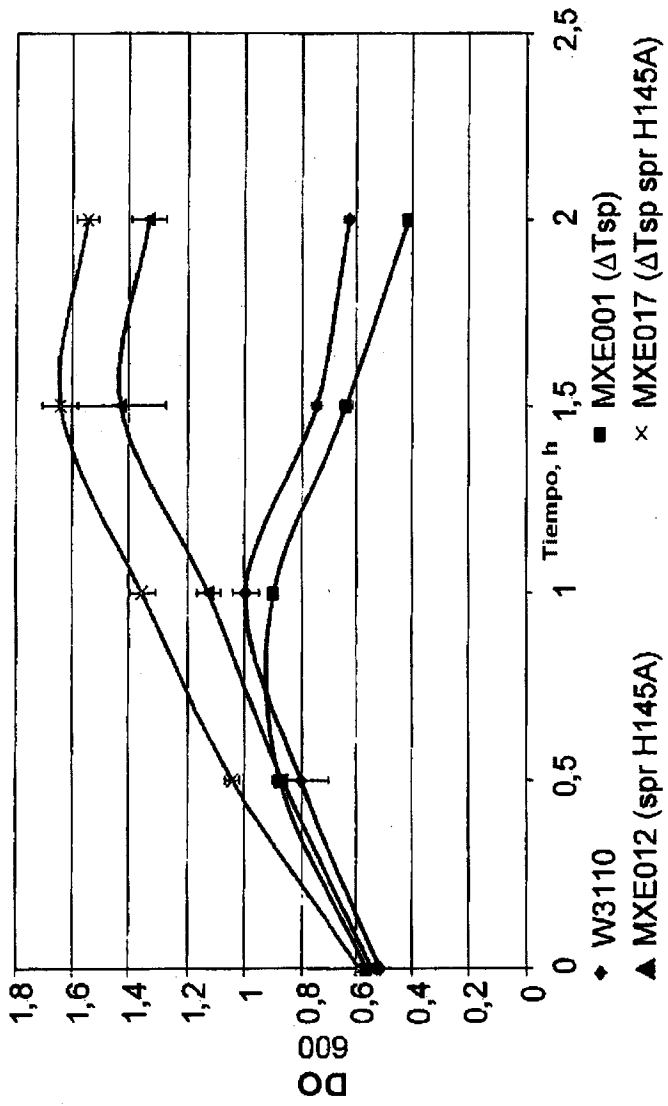


Figura 6

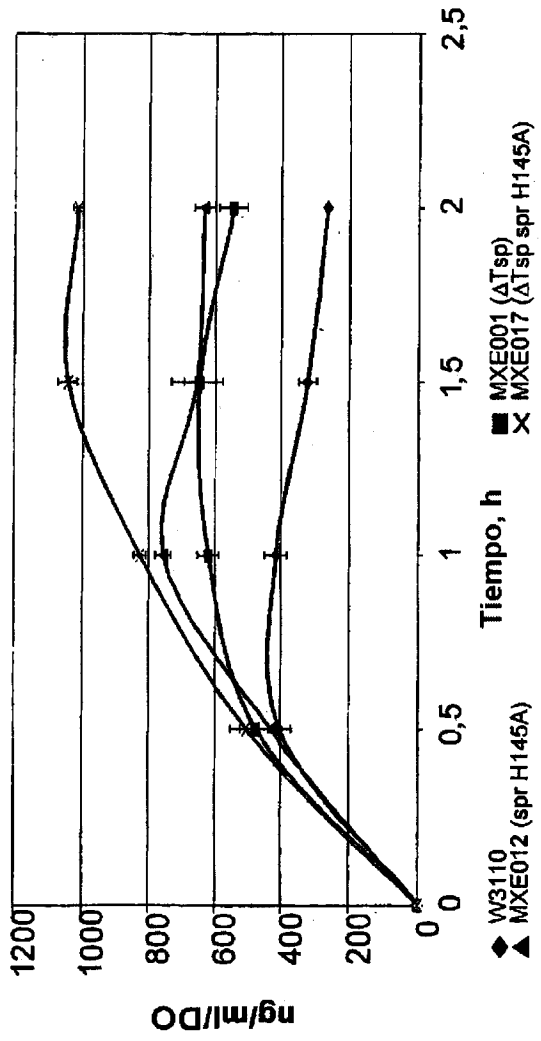


Figura 7

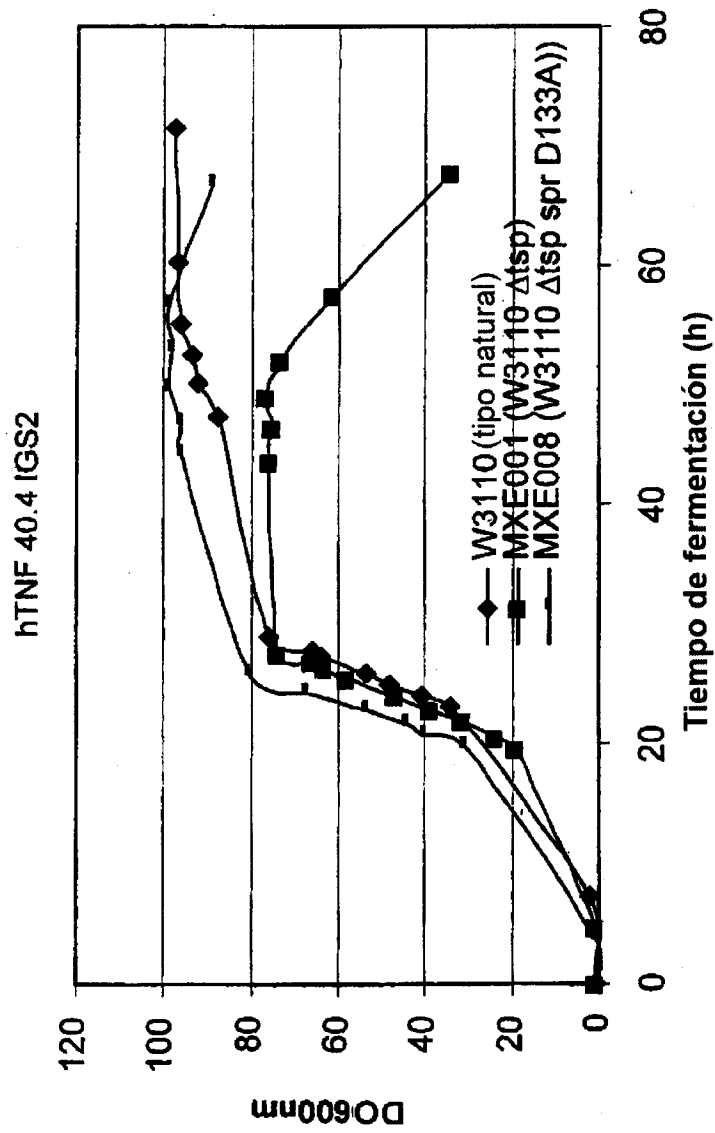


Figura 8

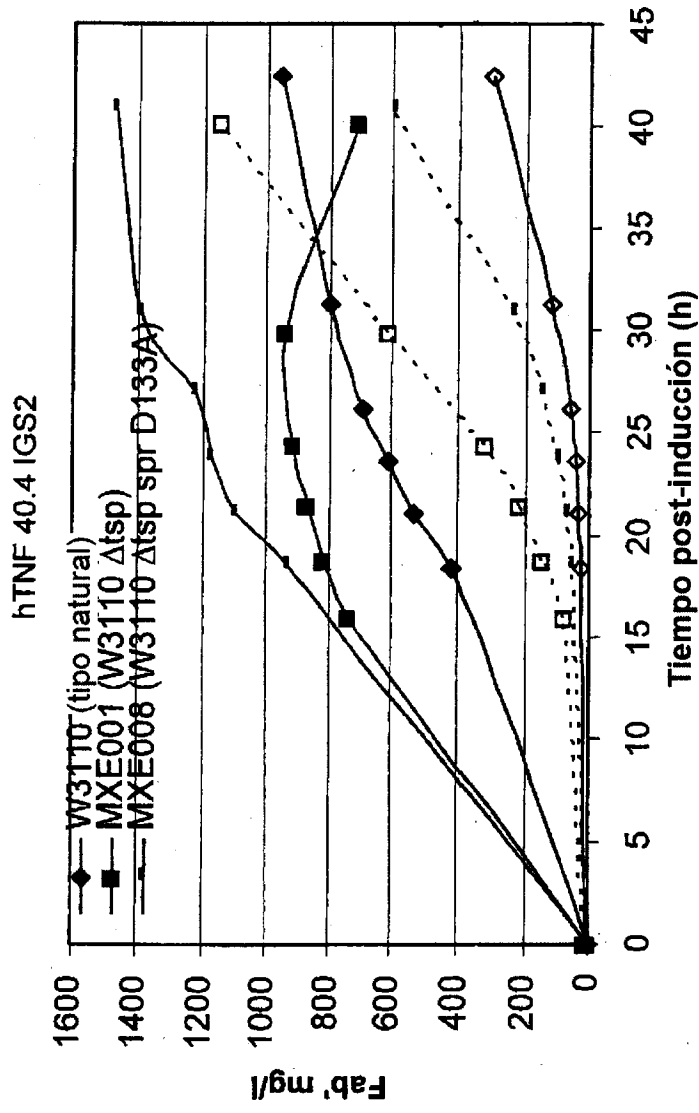


Figura 9

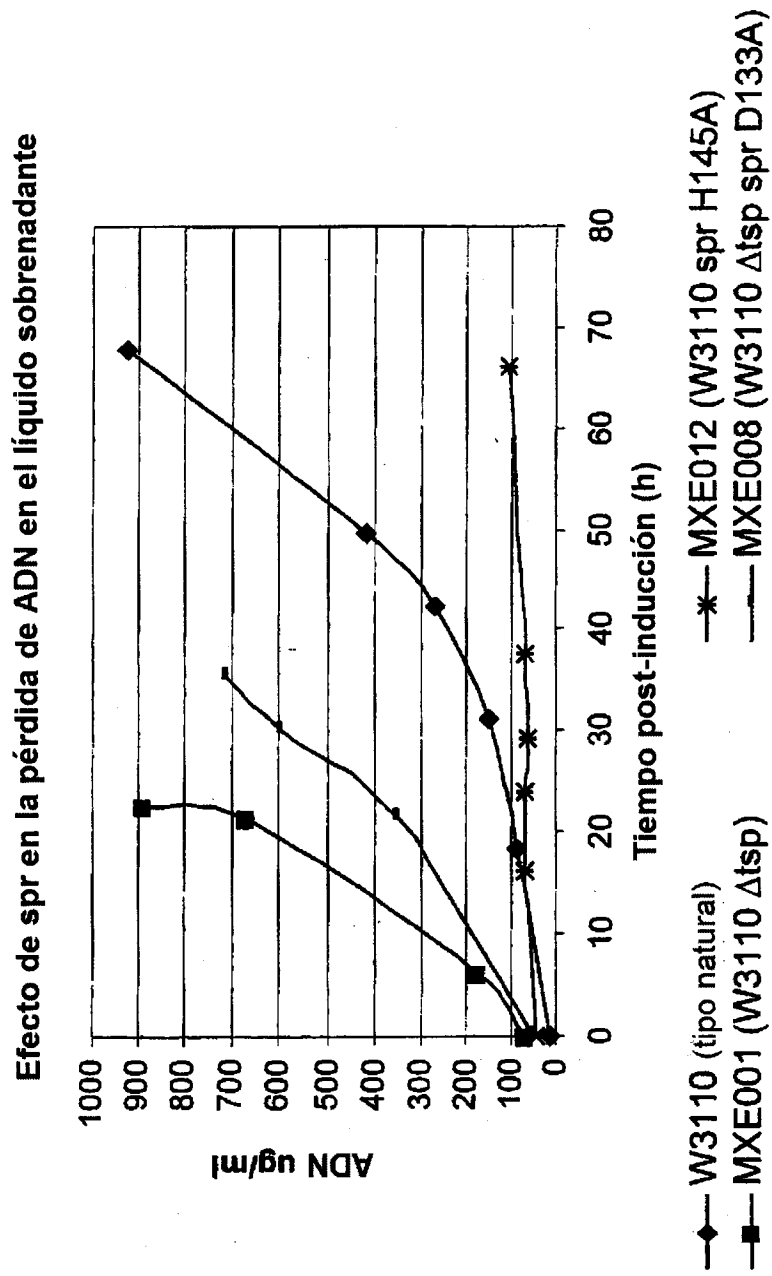


Figura 10

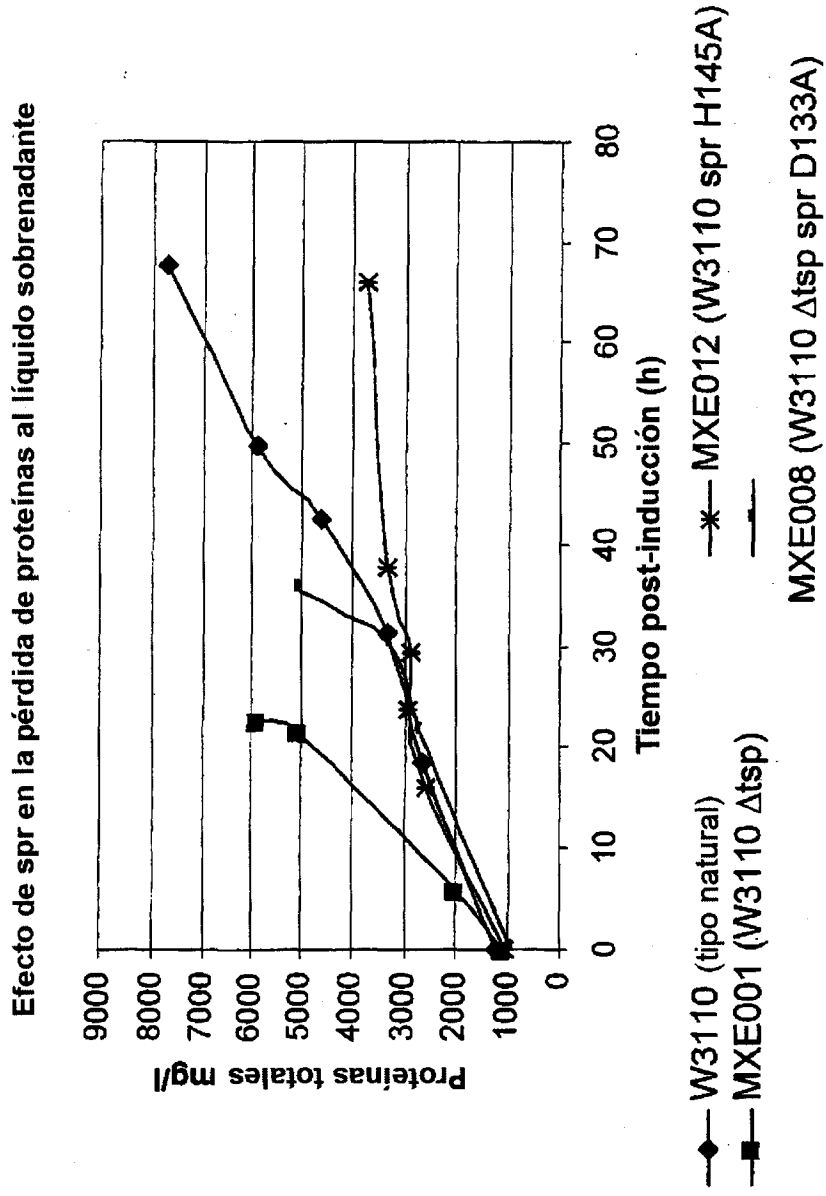


Figura 11a

ptr de tipo natural (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
    
```

ptr mutado Δ (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 11b

Tsp de tipo natural 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
    
```

Tsp mutado Δ 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 11c

DegP de tipo natural

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 mutado 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