



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 303 058**

51 Int. Cl.:
C12N 1/21 (2006.01)
C12N 1/20 (2006.01)
C12N 15/31 (2006.01)
C07K 14/195 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **04715290 .5**
86 Fecha de presentación : **27.02.2004**
87 Número de publicación de la solicitud: **1599579**
87 Fecha de publicación de la solicitud: **30.11.2005**

54 Título: **Un enzima de translocación como marcador de selección.**

30 Prioridad: **04.03.2003 DE 103 09 557**

45 Fecha de publicación de la mención BOPI:
01.08.2008

45 Fecha de la publicación del folleto de la patente:
01.08.2008

73 Titular/es: **Henkel AG. & Co. KGaA**
Henkelstrasse 67
40589 Düsseldorf, DE

72 Inventor/es: **Hintz, Maren;**
Freudl, Roland;
Feesche, Jörg y
Breves, Roland

74 Agente: **Carvajal y Urquijo, Isabel**

ES 2 303 058 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

Un enzima de translocación como marcador de selección.

5 La presente invención se refiere a un sistema de selección para microorganismos, que está basado en la inactivación de un enzima de translocación esencial y en la curación de esta inactivación por medio de un factor equivalente que se pone a disposición de las células correspondientes por medio de un vector.

10 En la actualidad se obtienen por medio de la fermentación de microorganismos, aquellas proteínas que son requeridas en grandes cantidades y, entre éstas, especialmente los enzimas para los campos de aplicación industrial. Frente a tales microorganismos, que forman, de manera natural, las proteínas interesadas, adquieren cada vez mayor significado las cepas de productores genéticamente modificados. Tales procedimientos de ingeniería genética para la obtención de proteínas están establecidos desde hace mucho tiempo en el estado de la técnica. El principio correspondiente consiste en los genes para las proteínas interesadas se incorporan en las células huésped como transgenes, son transcritos por las mismas, traducidos y, en caso dado, son secretados a través de las membranas correspondientes en el periplasma, o bien en el medio circundante. Por lo tanto pueden ser obtenidos a partir de las células correspondientes o bien de los sobrenadantes de cultivo correspondientes.

20 En los procedimientos industriales para la obtención de proteínas se utilizan, en primer lugar, las capacidades naturales de los microorganismos, empleados para la producción, para la síntesis y, en caso dado, para la secreción de las proteínas. Básicamente, se elegirán como sistemas bacterianos para la obtención de las proteínas, aquellos que sean de fermentación económica, que prometan, desde un inicio, una elevada tasa de formación de producto y que garanticen un plegamiento, una modificación, etc. correctos de la proteína que debe ser preparada; esto último es tanto más probable a medida que aumenta el parentesco con el organismo originalmente productor de la proteína interesada. Las células huésped, establecidas de manera especial para esta finalidad, son las bacterias gramnegativas, tales como por ejemplo *Escherichia coli* o *Klebsiella*, o las bacterias grampositivas, tales como, por ejemplo, las especies de los géneros *Staphylococcus* o *Bacillus*.

30 La economía de un procedimiento biotecnológico depende, de manera esencial, del rendimiento en proteína alcanzable. Este rendimiento está determinado, junto con el sistema de expresión empleado para el procedimiento utilizado, de manera especial por los parámetros de la fermentación y por los substratos. Mediante la optimización del sistema de expresión y del proceso de fermentación pueden acrecentarse claramente el potencial de un organismo de producción y el rendimiento que puede ser alcanzado.

35 Por otra parte, se desarrollan y se perfeccionan sistemas de expresión esencialmente a base de dos construcciones genéticas básicamente diferentes. Por un lado se integra el gen de la proteína, que debe ser preparada, en el cromosoma del organismo huésped. Tales constructos son muy estables sin selección con respecto a la presencia de un gen marcador adicional (véase más adelante). El inconveniente consiste en que únicamente está presente en el huésped una copia del gen y en que es muy costosa, desde el punto de vista de la metodología, la integración de otras copias para aumentar la tasa de formación del producto mediante el efecto de dosificación génica. A continuación se ha ilustrado brevemente este estado de la técnica.

45 La patente europea EP 284126 B1 resuelve el problema de una integración múltiple estable mediante la inserción de varias copias génicas en la célula, que contienen entremedias segmentos de ADN endógenos y, esencialmente, cromosómicos.

Otra solución para la integración múltiple estable ha sido divulgada en la solicitud de patente WO 99/41358 A1. Según esta publicación, se integran dos copias del gen interesado en sentidos de transcripción opuestos y se separan entre sí por medio de un segmento de ADN no esencial, para impedir, de este modo, la recombinación homóloga de ambas copias.

55 Se desprende por la solicitud de patente DD 277467 A1 un procedimiento para la obtención de enzimas extracelulares, que está basado en la integración estable, ventajosamente múltiple, en el cromosoma bacteriano del gen que codifica el enzima interesado. La integración se lleva a cabo a través de las regiones homólogas. Un gen de eritromicina, contenido en el plásmido, que se inactiva cuando la integración tenga éxito, sirve para controlar los acontecimientos de integración, que hayan tenido éxito.

60 De conformidad con la publicación DE 4231764 A1 puede llevarse a cabo la integración en el cromosoma a través de un entrecruzamiento sencillo o doble por medio del gen para la timidilato-sintetasa. Éste último permite el control de este modo de proceder puesto que se mantiene la actividad *thy* en el caso de un entrecruzamiento sencillo, mientras que se pierde en el caso de un entrecruzamiento doble, es decir que de este modo se consigue una auxotrofia. En este sistema especial va acompañado un entrecruzamiento sencillo por una resistencia contra el antibiótico constituido por la trimetoprima, y por una sensibilidad correspondiente en el caso de un entrecruzamiento doble.

65 En la solicitud WO 96/23073 A1 se divulga un sistema basado en transposon, para la integración de varias copias de un gen interesado en el cromosoma bacteriano, caracterizado porque el gen marcador del plásmido se somete a delección por medio de la integración y, de este modo, las cepas obtenidas están exentas de un marcador de resistencia.

Asimismo, según esta publicación, únicamente se requiere un marcador para el control de la construcción de la cepa bacteriana correspondiente.

5 De la misma manera, se divulga en la solicitud WO 01/90393 A1 un sistema para aumentar el número de copias de determinados transgenes, integrados en un cromosoma bacteriano.

10 El segundo postulado para la construcción de cepas productoras consiste en transferir el gen interesado a un elemento de replicación autónoma, por ejemplo un plásmido y, de este modo, efectuar su transferencia hasta el organismo huésped. De manera ventajosa se provoca en este caso el número usualmente elevado de copias de plásmidos por célula mediante el efecto de dosificación génica. Constituye un inconveniente el hecho de que tenga que ejercerse una presión de selección durante de todo el tiempo necesario para el cultivo, para mantener los plásmidos en las células. Esto se lleva a cabo, de manera normalizada, a través de la adición de antibióticos en el medio de cultivo, mientras que los genes, que proporcionan resistencia contra la sustancia correspondiente, se presentan sobre los plásmidos. De este modo, serían susceptibles de crecimiento únicamente las células, que tengan los plásmidos en un número suficiente.

15 El empleo de sustancias resistentes a los antibióticos como marcadores de selección se enfrenta, en los últimos años, a una crítica creciente. Por un lado, el empleo de antibióticos es realmente costoso, especialmente cuando la resistencia esté basada en un enzima que degrade al antibiótico y, por lo tanto, tenga que aportarse la sustancia correspondiente durante todo el proceso de cultivo. Por otro lado su empleo cada vez más extendido, especialmente en otros campos industriales, contribuye a la propagación de genes resistentes en otras cepas, incluso en otras cepas patológicas. Esto conduce, por ejemplo, en la higiene medicinal y, de manera especial, en el caso del tratamiento de enfermedades infecciosas, ya a considerables dificultades debidas a las denominadas cepas humanopatógenas multirresistentes.

20 Por lo tanto, se utilizan en grandes cantidades los sistemas precedentemente ilustrados para la integración estable de genes en el cromosoma de las células productoras, puesto que éstas son estables sin una presión de selección permanente. Desde luego, tal como se ha indicado precedentemente, las cepas correspondientes únicamente pueden obtenerse por medio de un gran coste. En la práctica biotecnológica es más rápido y es más cómodo introducir en las células huésped un gen, por ejemplo, recién encontrado o modificado, en un plásmido con marcador de selección y llevar a cabo su expresión de este modo.

25 En el estado de la técnica se han desarrollado, entre tanto, también sistemas de selección exentos de antibióticos. De este modo, se describen, por ejemplo, en la publicación "Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria" de Herrero *et al.* (1990), en J. Bacteriol., tomo 172, páginas 6557-6567, resistencias contra herbicidas y contra metales pesados como marcadores de selección. Sin embargo, contra el empleo de estos compuestos hablan los mismos reparos que en el caso de los antibióticos.

30 En principio la selección funciona, por ejemplo, de manera similar a como lo hace una selección con antibiótico, mediante la auxotrofia, es decir mediante un efecto metabólico específico, que hace que las células correspondientes sean dependientes del aporte de determinados productos del metabolismo. Las cepas auxotrofas contienen entonces un transgen, acoplado con el transgen interesado, que cura esta auxotrofia. En caso de pérdida, éstos perderían simultáneamente su capacidad de supervivencia bajo las condiciones de cultivo correspondientes de tal manera, que se produciría una selección deseada de las cepas productoras auxotrofas. De esta manera, se hace referencia, por ejemplo, en la página 148 de la publicación "Gene cloning in lactic streptococci" de de Vos en Netherlands Milk and Dairy Journal, tomo 40, (1986), páginas 141-154, a diversos marcadores de selección, desarrollados a partir del metabolismo de Lacto-Streptococci; entre los cuales se encuentran los que proceden del metabolismo de la lactosa, los que son resistentes al cobre y los genes resistentes frente a diversas bacteriocinas de Lacto-Streptococci. La patente EP 284126 B1, que se ocupa del problema de la integración estable de genes interesados en cromosomas bacterianos (véase más arriba), reúne los posibles sistemas para la selección, de auxotrofia, resistencia contra los biocidas y resistencia contra las infecciones víricas en la página 7 bajo el concepto de "selección de supervivencia". Como ejemplos de marcadores de selección por auxotrofia se indican en este caso los genes del metabolismo *leu*, *his*, *trp* "o similares"; evidentemente quieren indicarse los que proceden de vías de síntesis de aminoácidos.

35 Sin embargo, el empleo de tales auxotrofias se configura en la práctica, hasta ahora, muy problemático puesto que están disponibles casi todos los substratos necesarios en cantidades suficientes, especialmente, en los medios industriales de fermentación y las células correspondientes pueden compensar la carencia para la síntesis de un compuesto determinado mediante la absorción de este compuesto a partir del medio nutriente.

40 Hasta el presente constituye una excepción, únicamente, la timidina esencial, que está presente, tan solo, en trazas en los medios de fermentación industriales y, por lo tanto, tiene que ser formada por los organismos - proliferantes y, por lo tanto, sintetizadores de ADN - a través de una timidilato-sintasa. De este modo, la solicitud EP 251579 A2 ofrece la solución de emplear como células huésped aquellas que sean deficientes en timidilato-sintasa en lo que se refiere al gen esencial para el metabolismo del nucleótido. Por lo tanto, puede ponerse a disposición el gen precisamente para esta función (*thyA* procedente de *Escherichia coli* K12) a través de un vector y puede curarse el defecto génico. Cuando este vector porte, adicionalmente, el gen para la proteína interesada, se producirá una selección de las células productoras, similar a la que se lleva a cabo con los antibióticos.

ES 2 303 058 T3

En resumen, debe observarse que existe en el estado de la técnica ciertamente una gran experiencia en lo que se refiere a la producción biotecnológica de proteínas y que puede controlarse de manera específica la expresión de los genes interesados mediante la integración cromosómica y la selección por medio de antibióticos, sin embargo no existe prácticamente ninguna alternativa practicable hasta ahora para estos dos sistemas, especialmente no existe ninguna alternativa que sea menos costosa que la integración cromosómica y que, al mismo tiempo, pueda llevarse a cabo sin la selección por medio de un compuesto caro o cuestionable desde el punto de vista ecológico. De manera especial, el postulado de la selección a través de marcadores de auxotrofia ha conducido, hasta ahora, únicamente a resultados muy exigüos en lo que se refiere a los medios nutrientes complejos usuales, en general, en la industria.

Por lo tanto, se plantea la tarea de desarrollar un nuevo sistema de selección, que pueda manipularse de una manera comparativamente más sencilla que la selección a través de un antibiótico, pero que pueda realizarse, sin embargo, sin sustancias caras y, en ciertas circunstancias, dañinas para el medio ambiente. El sistema debería poderse emplear a escala industrial. No debería estar formado sobre un gen esencial, cuyos fallos puedan ser compensados por los productos acompañantes en los medios industriales.

Esta tarea se resuelve, de conformidad con la invención, mediante procedimientos para la selección de un microorganismo, caracterizados porque

(a) se inactiva un gen endógeno presente, que codifique una actividad de translocación esencial y

(b) se cura, por medio de un vector, la actividad de translocación esencial, inactivada según (a).

De manera sorprendente, se ha descubierto que los factores esenciales, que participan en la translocación, son adecuados para una selección. De conformidad con la invención, sirve para la selección un gen, cuya proteína derivada participa en la translocación de la proteína como un factor esencial para la célula correspondiente. Esto significa que la eliminación de este gen es letal y, por lo tanto, posibilita una selección de los microorganismos similar a la que se lleva a cabo con un antibiótico.

Este sistema de selección actúa sin aditivos (tales como, por ejemplo, los antibióticos establecidos en el estado de la técnica) y funciona, en principio, independientemente de la composición de los medios nutrientes. Con esta finalidad se requiere una modificación biológica molecular de los microorganismos correspondientes, que se describe después de las explicaciones relativas a la translocación y que puede verse en los ejemplos de la presente solicitud.

El proceso de translocación consiste en la secreción de las proteínas formadas por las bacterias en el periplasma (en el caso de bacterias gramnegativas), o bien en el medio circundante (tanto en el caso de las bacterias gramnegativas así como, también, en el caso de las bacterias grampositivas). Ésta se ha descrito, por ejemplo, en el artículo de A.J. Driessen (1994): "How proteins cross the bacterial cytoplasmic membrane" en *J. Membr. Biol.*, 142 (2), páginas 145-59. El aparato de secreción, necesario con esta finalidad, está constituido por una serie de proteínas diferentes, fundamentalmente asociadas a la membrana, que se han mostrado en la figura 1 de la presente solicitud. A éstas pertenecen, de manera especial, las proteínas que están perfectamente caracterizadas, por ejemplo, para *Bacillus subtilis* constituidas por SecA, SecD, SecF (en conjunto como complejo SecDF), E, G e Y (van Wely, K.H., Swaving, J., Freudl, R., Driessen, A.J. (2001): "Translocation of proteins across the cell envelope of Gram-positive bacteria", *FEMS Microbiol Rev.* 2001, 25(4), páginas 437-54). Otros factores que pueden asociarse a este sistema son YajC, que entra también en contacto directo con el complejo Sec, y los factores que pueden reconocerse también en la figura 1 Bdb (Dsb), SPasa (para "peptidasa señal"), PrsA y b-SRP (Ffh, Ffs/Scr, SRP-RNA).

El factor, que ha sido citado en último lugar, está constituido por el factor bacteriano, que ejerce, en principio, la misma función que la SRP (partícula de reconocimiento de señal -signal recognition particle-) descrita originalmente en eucariotas. Una subunidad del mismo es el factor Ffh, que está caracterizado tanto a partir de *B. subtilis* así como, también, a partir de *E. coli*. La subunidad siguiente de b-SRP se llama Scr en *B. subtilis* y Ffs en *E. coli*. De igual modo un ARN (SRP-ARN) forma parte del complejo funcional b-SRP. Otro factor funcional, emparentado con el anterior se denomina Srb en *E. coli* y FtsY en *B. subtilis*. Éste corresponde desde el punto de vista funcional a la proteína de acople eucariota.

En otro contexto debe incluirse por ejemplo también el factor PrfB (factor B de desprendimiento de la cadena peptídica -peptide chain release factor B-; denominado también RF2). Éste es conocido por formar parte integrante del aparato para la síntesis de proteínas y, concretamente, tanto en el caso de las bacterias gramasí como, también, en el caso de las bacterias gramnegativas. Éste es responsable, en relación con la traducción, especialmente para el desprendimiento de las proteínas traducidas, acabadas, del ribosoma, es decir que es responsable de la terminación de la traducción. La relación con la translocación precedentemente descrita está únicamente dada de manera indirecta, puesto que el gen *prfB* se transcribe en muchas bacterias simultáneamente con el gen para el factor SecA. Por lo tanto existe una relación reguladora.

La condición previa para la translocación consiste en que la proteína, que debe ser retirada, disponga en el extremo N de un péptido señal (Park, S., Liu, G., Topping, T.B., Cover, W.H., Randall, L.L. (1988): "Modulation of folding pathways of exported proteins by the leader sequence", *Science*, 239, páginas 1033-5). Esto es válido tanto para las proteínas extracelulares así como, también, para las proteínas de la membrana.

ES 2 303 058 T3

Una vez verificada con éxito la traducción en los ribosomas, se mantiene en un estado no plegado la cadena peptídica formada, mediante proteína citoplasmática con función de chaperona, y es transportada hasta la membrana. A continuación se cataliza el transporte del péptido a través de la membrana bajo consumo de ATP (Mitchell, C., Oliver, D. (1993): "Two distinct ATP-binding domains are needed to promote protein export by *Escherichia coli* SecA ATPase", Mol. Microbiol., 10(3), páginas 483-97). La SecA actúa en este caso como componente suministrador de energía (ATPasa) del complejo multienzimático translocasa. Una vez atravesada la membrana, se corta el péptido señal por medio de la actividad de una peptidasa señal y, de este modo, se desprende de la membrana la proteína extracelular. En el caso de las bacterias grampositivas se lleva a cabo de este modo la descarga de la exoproteína directamente en el medio circundante. En el caso de las bacterias gramnegativas las proteínas se encuentran, a continuación, por regla general en el periplasma y son necesarias otras modificaciones para conseguir su liberación en el medio circundante.

Una investigación novedosa de B.v.d.Berg *et al.* (2004; Nature, tomo 427, páginas 36 hasta 44) describe la estructura a los rayos X de la translocasa, es decir del complejo SecY procedente de *Methanococcus jannaschii* como modelo para el complejo correspondiente en otros organismos.

La molécula central de este proceso es, por lo tanto, la preproteína-translocasa, constituida por las subunidades SecA, SecY, SecE, SecD, SecF (SecDF) y SecG. En este caso, tiene una función decisiva para la translocación el factor SecA en todas las ATPasas, que controlan este proceso. Las formas preferentes de realización de la presente invención están caracterizadas por estos factores (véase más adelante).

En la tabla 1 siguiente se han enumerado los factores conocidos hasta ahora, que participan en la translocación y han sido detallados con relación a si son esenciales en uno o en ambos organismos modelo constituidos por *Escherichia coli* (gramnegativo) y *Bacillus subtilis* (grampositivo). Tan pronto como sea esencial uno de los mismos en el correspondiente organismo considerado, será adecuado para la selección de conformidad con la invención. Básicamente se supondrá que existe un modelo correspondiente también en otros organismos, especialmente en las bacterias gramnegativas y en las bacterias grampositivas.

(Tabla pasa a página siguiente)

ES 2 303 058 T3

TABLA 1

Factores proteicos conocidos hasta ahora, que participan en la translocación de las bacterias gramnegativas y grampositivas, habiéndose detallado si son esenciales en estos organismos

5

10

15

20

25

30

35

40

45

50

55

	<i>E. coli</i>	<i>B. subtilis</i>
SecA	esencial	esencial
SecY	esencial	esencial
SecE	esencial	esencial
SecG	no esencial	no esencial
	(fenotipo sensible al frío)	(fenotipo sensible al frío en la superproducción de proteínas de exportación)
SecD, SecF	esencial	no esencial
(SecDF)		(fenotipo sensible al frío)
Peptidasa señal	esencial	no esencial, puesto que está presente de manera redundante
b-SRP (Ffh; Ffs / Scr; SRP-RNA)	esencial	esencial
FtsY / Srb	esencial	esencial
PRSA	no presente	esencial
Bdb / Dsb	no esencial	no esencial
YajC	esencial	no conocido, aún cuando es esencial, sin embargo está presente de manera redundante

Así pues, puede seleccionarse, de conformidad con la invención, en bacterias gramnegativas, especialmente en bacterias coliformes, de manera muy especial en *E. coli*, por medio de los siguientes enzimas de translocación o bien por medio de sus genes correspondientes: SecA, SecY, SecE, SecD, SecF, peptidasa señal, b-SRP (Ffh o Ffs), Srb o YajC.

Así pues, puede seleccionarse, de conformidad con la invención, en bacterias grampositivas, especialmente en *Bacillus*, de manera muy especial en *B. subtilis*, por medio de los siguientes enzimas de translocación o bien por medio de sus genes correspondientes: SecA, SecY, SecE, b-SRP (Ffh o Scr), FtsY o PrsA.

ES 2 303 058 T3

En estas enumeraciones no debe entenderse que está excluido en enlace alternativo. Desde el punto de vista industrial debería ser posible desconectar simultáneamente también varios de los genes correspondientes. De conformidad con la invención es suficiente, sin embargo, elegir para esta finalidad únicamente uno de ellos.

5 Las secuencias individuales respectivas de los genes correspondientes pueden ser obtenidas, por ejemplo, a partir de los bancos de datos siguientes, que son accesibles al público: GenBank (National Center for Biotechnology Information NCBI, National Institutes of Health, Bethesda, MD, USA; <http://www3.ncbi.nlm.nih.gov>); EMBL-European Bioinformatics Institute (EBI) en Cambridge, Gran Bretaña (<http://www.ebi.ac.uk>); Swiss-Prot (Geneva Bioinformatics (GeneBio) S.A., Ginebra, Suiza; <http://www.genebio.com/sprot.html>); “Subtilist” o “Colibri” del Institut Pasteur, 10 25,28 rue du Docteur Roux, 75724 Paris CEDEX 15, Francia para genes y factores procedentes de *B. subtilis* o bien de procedentes de *E. coli* (<http://genolist.pasteur.fr/Subtilist/> y respectivamente <http://genolist.pasteur.fr/Colibri/>). De la misma manera, pueden utilizarse los bancos de datos, que pueden ser visitados por medio de las referencias cruzadas en los bancos de datos citados. De conformidad con la invención, es necesario respectivamente identificar tan sólo un solo gen esencial del aparato de translocación en la cepa considerada para el cultivo y aplicarlo de manera 15 correspondiente.

Otro punto de partida proporciona las secuencias para el factor SecA procedente de diversos microorganismos, que están dadas en el protocolo de secuencias para la presente solicitud. Éstas pueden emplearse bien directamente (véase 20 más adelante: formas de realización preferentes) o pueden emplearse para identificar el homólogo correspondiente en una genoteca, que se ha establecido previamente para el microorganismo interesado.

En primer lugar se entenderán por estos enzimas de translocación o factores de translocación las moléculas de tipo silvestre. Sin embargo, en tanto en cuanto sea posible la preparación de variantes de los mismos, que dispongan en el aparato de translocación, de la misma función, en principio, que el enzima de tipo silvestre, quedarán abarcados en el 25 alcance de la protección los sistemas de selección basados en los mismos.

Para la realización de la presente invención tiene que verificarse con respecto a la cepa que se trata de cultivar, en tanto en cuanto no pertenezca a una de las especies ensayadas en la presente solicitud, si es esencial al menos uno de estos factores. Esto es posible por ejemplo, en el caso más sencillo, retirándose uno de estos genes conocidos 30 de una cepa tan emparentada como sea posible (por ejemplo una bacteria también gramnegativa o respectivamente grampositiva) o sintetizándose por medio de una inscripción en un banco de datos accesible al público y efectuándose su introducción en un vector desactivado (knock-out). Una forma de proceder de este tipo es conocida por el técnico en la materia. Si las transformaciones de este gen, realizadas con este vector y una recombinación homóloga subsiguiente - de manera preferente introducida por separado de la transformación - gen tienen un efecto letal en el genoma de la célula huésped, deberá tratarse como esencial el gen considerado. Este gen, reconocido como esencial, puede 35 emplearse ahora como marcador de selección de conformidad con la invención y, de manera especial, según el modelo de los ejemplos de la presente solicitud.

A título de ejemplo, se lleva a cabo una inactivación, necesaria según la característica (a), por medio de la recombinación homóloga de una copia génica inactivada, que haya sido introducida por ejemplo mediante la transformación con un vector correspondiente en una célula de la cepa del microorganismo interesado. Los métodos para esta finalidad son en sí conocidos. Como resultado del acontecimiento de recombinación se somete a delección total o parcial a la copia cromosómica del gen y, de este modo, es inoperante. Esto puede llevarse a cabo, por ejemplo, a través del mismo gen, con el cual se haya llevado a cabo previamente el ensayo sobre la letalidad. De manera preferente, se empleará 40 sin embargo, en tanto en cuanto sea conocido o pueda aislarse con un coste aceptable, el homólogo endógeno para conseguir una elevada cuota de éxito para la recombinación. Lo decisivo para la realización de la invención consiste en que la inactivación sea realizada con éxito.

Este concepto puede realizarse, por ejemplo, con vectores plásmidos, que tengan un origen de replicación sensible a la temperatura y hayan sido insertados en las regiones de ADN homólogas, consideradas, consideradas, del gen que debe ser sometido a delección (vector de delección). Podría imaginarse, por ejemplo, también una inactivación reversible, por ejemplo mediante la integración de un elemento genéticamente móvil, por ejemplo de un transposon, en el gen 45 diana.

55 En este caso, debe observarse respectivamente la característica (b), en concreto que esté disponible en la célula correspondiente una copia intacta del gen elegido para la selección de conformidad con la invención, ya como paso previo a este episodio de recombinación o bien de este episodio de inactivación o, como más tarde, simultáneamente con el mismo, puesto que la célula no sobreviviría en otro caso la inactivación. De conformidad con la invención, el defecto formado es compensado por medio de un vector, es decir que el vector cura la inactivación. En este caso se 60 emplearán, como se ha indicado precedentemente, de manera preferente los genes endógenos presentes en las células huésped y sometidos a delección según (a). De igual modo, pueden emplearse genes con la misma función procedentes de otros organismos, de manera preferente cepas emparentadas, en tanto en cuanto sean capaces de curar el defecto correspondiente. Así pues, es posible, por ejemplo, curar el defecto de SecA de un *B. subtilis* mediante la preparación de un gen *secA* de *Staphylococcus carnosus*. 65

De igual modo, podría imaginarse aplicar en la célula, a través del vector, otro elemento genético que elimine el primer defecto, por ejemplo el gen de un factor funcionalmente idéntico en principio, pero modificado por mutación.

ES 2 303 058 T3

Por lo tanto, en esta célula reina una situación tal, que se compensa un defecto letal por medio de un elemento genético separado. De igual modo una pérdida de este elemento genético separado sería letal de tal manera que una célula de este tipo está obligada a transmitir este elemento a la generación siguiente en cada división celular.

5 Esto es útil en la característica (c), en una forma preferente de realización (véase más adelante), según la cual el vector, que compensa el defecto génico, porta el transgen, precisamente el gen para la proteína interesada, que debe ser preparado por medio del procedimiento según la invención.

10 En cierta medida, reina una presión de selección endógena sin que sea necesaria la adición de un compuesto de otro tipo desde el exterior, es decir a través del medio nutriente, por ejemplo de un metal pesado o de un antibiótico para impedir la pérdida del vector con el transgen. Por otra parte, quedan eliminadas las modificaciones costosas, descritas al principio, para integrar al transgen propiamente dicho en el ADN cromosómico. De este modo, pueden emplearse para transformaciones, siempre nuevas, con vectores formados de manera similar, por ejemplo una cepa de microorganismos generada una vez, que haya sido preparada para una inactivación definida del aparato de translocación, cuyos
15 vectores pongan a disposición de las mismas, en cada caso, la función que cura el defecto génico, pero que porten en cada caso transgenes diferentes. De esta manera, se dispone de un sistema de selección muy práctico y que puede emplearse de múltiples maneras.

20 El sentido del procedimiento de selección, de conformidad con la invención, consiste en obtener en la célula un elemento genético estable a través de varias generaciones. Este elemento es precisamente el vector, a través del cual se compensa la inactivación de la actividad esencial de translocación, es decir que es curada.

En una forma preferente de realización, los procedimientos de selección, de conformidad con la invención, se caracterizan porque

25 (c) el vector según (b) porta un transgen.

Éste es, precisamente, el campo de aplicación industrial más importante de los sistemas de selección. El elemento genético, estable a través de varias generaciones, es entonces aquél que porta un transgen, especialmente aquél cuyo
30 producto génico es de interés comercial. Formas preferentes de realización a este respecto se explican más adelante.

En las formas preferentes de realización, un procedimiento de selección, de conformidad con la invención, se caracteriza porque la actividad de translocación esencial está constituida por uno de los factores siguientes: SecA, SecY, SecE, SecD, SecF, peptidasa señal, b-SRP (Ffh o Ffs / Scr), FtsY / Srb, PrsA o YajC.

35 Precisamente, como se han agrupado en la tabla 1, estos factores y respectivamente los genes correspondientes están constituidos por aquellos, que se sabe que son esenciales, bien procedentes de *E. coli* o procedentes de *B. subtilis*. Por lo tanto, es evidente que puede establecerse un sistema de selección, de conformidad con la invención, mediante los homólogos correspondientes, especialmente en estos dos organismos o incluso en especies emparentadas o incluso en especies menos emparentadas. Puesto que se sabe, que incluso algunos de estos genes individuales pueden substituir la función correspondiente a través del otro organismo correspondiente, es decir a través de los límites gramnegativo/grampositivo, deberían poderse emplear, de conformidad con la invención, al menos algunos de los genes correspondientes individuales incluso procedentes de especies sólo lejanamente emparentadas.

45 De manera preferente, la actividad de translocación esencial está constituida por una de las siguientes subunidades de la preproteína-translocasa: SecA, SecY, SecE, SecD o SecF, de manera preferente se trata de la subunidad SecA.

Precisamente, estos factores representan, como se ha mostrado en la figura 1, en cierta medida el núcleo funcional del aparato de translocación. Con relación al SecA se ha indicado más arriba que este factor ejerce una posición clave importante a través de la actividad ATPasa. Por este motivo se ha aplicado la invención también en los ejemplos de la presente solicitud por medio del gen *secA*.

55 De manera preferente, los procedimientos de selección, de conformidad con la invención, se caracterizan porque la curación según (b) se lleva a cabo a través de una actividad equivalente a la actividad de translocación esencial existente de manera endógena, inactivada, preferentemente por medio de una actividad genéticamente emparentada, de forma especialmente preferente por medio de la misma actividad.

Aquí se pone nuevamente de manifiesto, que pueden ser empleados de manera prometedora de éxito los genes correspondientes procedentes de especies menos emparentadas, debido a los elevados valores de homología generales entre las especies para los factores correspondientes. Sin embargo son preferentes, evidentemente, los que procedan de especies más emparentadas y, de forma muy preferente, los que procedan de los mismos organismos, puesto que éstos son los más prometedores de éxito en lo que se refiere al necesario entre cruzamiento para la inactivación. Una vez más debe ponerse de manifiesto, que únicamente es suficiente un solo gen adecuado para la inactivación con el fin de realizar una selección de conformidad con la invención.

65 Tal como ya se ha indicado precedentemente, las secuencias de ADN y de aminoácidos correspondientes pueden ser obtenidas a partir de bancos de datos accesibles al público. De este modo, las secuencias que han sido indicadas por ejemplo en el protocolo de secuencias bajo SEQ ID NO. 1 y 2 para la proteína SecA procedente de *B. subtilis*

ES 2 303 058 T3

se ha tomado del banco de datos "Subtilist" del Institut Pasteur (véase más arriba) (situación a: 3.2.2003); éstas son idénticas a las de Swiss-Prot (véase más arriba), que han sido depositadas allí bajo el número de acceso P28366.

5 Las secuencias, indicadas en el protocolo de secuencias bajo SEQ ID NO. 3 y 4, para la proteína SecA procedente de *E. coli* procedente del banco de datos "Colibri" del Institut Pasteur (véase más arriba; situación a: 3.2.2003); éstas son idénticas a las de Swiss-Prot (véase más arriba), que pueden ser consultadas en el mismo bajo el número de acceso P10408.

10 Las SEQ ID NO. 5 y 6 para *B. licheniformis* se obtuvieron, como se ha descrito en el ejemplo 1 de la presente solicitud, a partir de la cepa *B. licheniformis* (DSM13) que puede ser adquirida en el comercio (Colección Alemana de Microorganismos y de Cultivos Celulares -Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH-, MascheroderWeg 1b, 38124 Braunschweig; <http://www.dsmz.de>).

15 Puesto que las especies *Bacillus subtilis*, *Escherichia coli* y *Bacillus licheniformis* constituyen los microorganismos empleados más frecuentemente en la industria, es especialmente importante proporcionar un procedimiento de conformidad con la invención para estas bacterias. Con el protocolo de secuencias de la presente solicitud se ponen a disposición los genes *secA* homólogos procedentes de estos tres organismos, que son los más importantes, sin que tengan que ser aislados para la elaboración final. Mediante estos genes es posible, por ejemplo, identificar los homólogos correspondientes en otros microorganismos, por ejemplo mediante la preparación de una genoteca y el cribado con
20 uno de estos genes a título de sonda. Sin embargo, pueden emplearse de manera especial en especies emparentadas también estos genes propiamente dichos para una inactivación de conformidad con la invención.

Por lo tanto, las formas de realización preferentes se caracterizan porque se lleva a cabo la curación según (b) mediante las regiones del gen *secA*, que restablecen la actividad de translocación, procedente de *Bacillus subtilis*,
25 *Escherichia coli* o de *Bacillus licheniformis*, que se han indicado en el protocolo de secuencias bajo SEQ ID NO. 1, SEQ ID NO. 3 o bien SEQ ID NO. 5.

De la misma manera, los procedimientos preferentes se caracterizan porque la inactivación se lleva a cabo, según (a), de tal manera, que se impida una recombinación entre la región génica inactivada, según (a), y la región homóloga
30 sobre el vector, según (b), de manera preferente con pérdida completa de un segmento génico contenido en el gen cromosómico correspondiente.

Precisamente, cuando el vector estuviese integrado en el cromosoma de la célula huésped, se curaría de manera duradera la mutación letal sin que existiese al mismo tiempo una presión de selección sobre el vector correspondiente.
35 De este modo podría perderse el transgen interesado propiamente dicho a través de la división celular subsiguiente. Esto impide una amplia delección durante la etapa de inactivación (a).

En el estado de la técnica se han descrito procedimientos para la inactivación de genes por medio de un vector de delección, especialmente en la publicación "Genetic manipulation of *Bacillus amyloliquefaciens*" de J. Vehmaanperä *et al.* (1991) en *J. Biotechnol.*, tomo 19, páginas 221-240. Por medio de esta descripción pudo llevarse a cabo con éxito
40 en el ejemplo 3 la delección del gen *secA* de *B. licheniformis*. El origen de la replicación de este vector de delección se caracteriza por su dependencia de la temperatura.

De este modo es posible, de una manera especialmente fácil, efectuar la selección, en primer lugar a temperatura
45 más baja, para obtener una transformación con éxito y, a continuación, ejercer una presión de selección mediante el aumento de la temperatura para conseguir una integración con éxito, es decir una inactivación del gen endógeno. De manera análoga, sería posible por ejemplo también una construcción para el control mediante la adición de compuestos de bajo peso molecular.

50 Los procedimientos preferentes, de conformidad con la invención, se caracterizan, por lo tanto, porque la inactivación, según (a), se lleva a cabo por medio de un vector de delección, preferentemente por medio de un vector de delección con un origen de replicación regulable de manera externa, de manera especialmente preferente por medio de un vector de delección con un origen de la replicación en función de la temperatura.

55 Tal como se ha indicado más arriba, es posible, en principio, que se integre en el cromosoma bacteriano el vector de curación, según (b), con inclusión del transgen. Puesto que, sin embargo, por este motivo existe el peligro de que el transgen se pierda en esta ocasión de manera duradera, los procedimientos preferentes se caracterizan porque el vector según (b) está constituido por un plásmido, que se replica de manera autónoma en el microorganismo. Éste se establece entonces en la línea celular derivada.
60

Es especialmente ventajoso que el plásmido esté constituido por un plásmido que se establezca con un número repetido de copias (por ejemplo entre 2 y 100 plásmidos por célula), de manera preferente con un número múltiple de copias (por encima de 100 plásmidos por célula). Precisamente, cuanto mayor sea el número de plásmidos presentes tanto mayor podrá ser el éxito alcanzado con la curación. De la misma manera, esto aumenta los rendimientos
65 alcanzables en este producto mediante el efecto de dosificación génica en el caso en que el vector porte un transgen que codifique una proteína interesada.

ES 2 303 058 T3

Como consecuencia del gran significado de las cepas bacterianas gramnegativas, especialmente en la clonación y en la caracterización de genes o bien de productos génicos, los procedimientos de selección preferentes se caracterizan porque el microorganismo está constituido por una cepa de bacterias gramnegativas.

5 Bajo este concepto se entenderán de manera especial aquellos procedimientos, caracterizados porque se trata de una cepa bacteriana gramnegativa del género *E. coli* o *Klebsiella*, especialmente se trata de derivados de *Escherichia coli* K12, de *Escherichia coli* B o de *Klebsiella planticola*, y de una manera muy preferente se trata de derivados de las cepas *Escherichia coli* BL21 (DE3), *E. coli* RV308, *E. coli* DH5 α , *E. coli* JM109, *E. coli* XL-1 o *Klebsiella planticola* (Rf). Precisamente, éstos son los organismos que se emplean con mayor frecuencia en la biología molecular.

10 De manera especial, las bacterias grampositivas tienen un significado especial para la obtención de proteínas por fermentación. Esto es válido, de una manera especial para las proteínas secretadas. Por lo tanto, los procedimientos preferentes, de conformidad con la invención, se caracterizan porque el microorganismo está constituido por una cepa de bacterias grampositivas.

15 Entre éstas se han impuesto, especialmente en la industria, las cepas de bacterias grampositivas de los géneros *Staphylococcus*, *Corynebacterien* o *Bacillus*, de manera especial de las especies *Staphylococcus carnosus*, *Corynebacterium glutamicum*, *Bacillus subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, *B. globigii* o *B. lentus*, y de manera muy especial los derivados de las cepas *B. licheniformis* o *B. amyloliquefaciens*, con lo cual éstas caracterizan de
20 manera correspondiente los procedimientos de selección preferentes.

Tienen un interés muy especial los procedimientos, de conformidad con la invención, que estén dirigidos hacia determinados productos, preparados mediante el cultivo de microorganismos. De manera correspondiente son preferentes, por lo tanto, los procedimientos de selección, que se caractericen porque el transgen según (c) esté constituido
25 por aquél que codifique una proteína no enzimática, de manera especial que codifique una proteína relevante desde el punto de vista farmacológico, de manera muy preferente que codifique insulina o calcitonina.

De igual modo, los enzimas tienen un significado industrial. Por este motivo se reivindicarán también, de conformidad con la invención, aquellos procedimientos caracterizados porque el transgen según (c) esté constituido por
30 aquél que codifique un enzima, de manera preferente que codifique un enzima hidrolítico o una oxidoreductasa, de manera especialmente preferente que codifique una proteasa, una amilasa, una hemicelulasa, una celulasa, una lipasa, una cutinasa, una oxidasa, una peroxidasa o una laccasa. A continuación se indican los representantes productores preferentes de las mismas para el empleo en los agentes de lavado y de limpieza.

35 En principio pueden emplearse tales agentes para aumentar el rendimiento de lavado o bien el rendimiento de limpieza de todos los enzimas que se han impuesto en el estado de la técnica. Entre las proteasas son preferentes aquellas del tipo subtilisina, por ejemplo las proteasas alcalinas procedentes de *Bacillus lentus*. Las variantes comercializadas bajo la denominación BLAP[®] se derivan de las proteasas procedentes de *Bacillus lentus* DSM 5483 (WO 91/02792 A1), que han sido descritas, de manera especial, en las publicaciones WO 92/21760 A1, WO 95/23221 A1,
40 WO 02/088340 A2 y WO 03/038082 A2. Otras proteasas, que pueden ser preparadas de conformidad con la invención, procedentes de diversos *Bacillus sp.* y *B. gibsonii* se desprenden de las solicitudes de patente WO 03/054185 A1, WO 03/056017 A2, WO 03/055974 A2 y WO 03/054184 A1.

Ejemplos de las amilasas, que pueden ser preparadas de conformidad con la invención, son las α -amilasas procedentes de *Bacillus licheniformis*, procedentes de *B. amyloliquefaciens* o procedentes de *B. stearothermophilus* así como sus desarrollos perfeccionados especialmente para el empleo en los agentes de lavado y en los agentes de limpieza. De igual manera deben señalarse para esta finalidad las α -amilasas procedentes de *Bacillus sp.* A 7-7 (DSM 12368), que han sido divulgadas en la solicitud WO 02/10356 A2 y la ciclodextrina-glucanotransferasa (CGTasa) procedente de *B. agaradherens* (DSM 9948), que ha sido descrita en la solicitud WO 02/44350 A2. De igual modo, quedan considerados los enzimas amilolíticos en el campo de la presente solicitud, que pertenezcan al campo de secuenciación de las α -amilasas, que se define en la solicitud WO 03/002711 A2 y que han sido descritas en la solicitud WO 03/054177 A2. De la misma manera quedan incluidos los productos de fusión de las moléculas citadas, por ejemplo las que están
50 incluidas en la solicitud DE 10138753 A1.

55 De conformidad con la invención, pueden prepararse también lipasas o cutinasas, por ejemplo las lipasas que pueden ser obtenidas originalmente a partir de *Humicola lanuginosa* (*Thermomyces lanuginosus*), o bien las lipasas perfeccionadas, especialmente aquellas con el intercambio de aminoácidos D96L o las lipasas, o respectivamente las cutinasas, cuyos enzimas de partida hayan sido aislados originalmente a partir de *Pseudomonas mendocina* y *Fusarium solanii*.

60 De la misma manera, se ha pensado en las celulasas, especialmente para el tratamiento industrial de artículos textiles así como también para agentes de lavado, por ejemplo la endoglucanasa y/o las celobiohidrolasas. Ejemplos de las celulasas, que pueden ser preparadas también por medio de productores naturales son aquellas que proceden de *Bacillus sp.* CBS 670.93 y CBS 669.93, como las que han sido divulgadas en la publicación WO 96/34092 A2. De
65 igual modo, pueden prepararse otros enzimas, de conformidad con la invención, que quedan reunidos bajo el concepto de hemicelulasas. A éstas pertenecen, por ejemplo, las mananasas, las xantanliasas, las pectinliasas (= pectinasas), las pectinesterasas, las pectatliasas, las xiloglucanasas (= xilanasas), las pululanastas y las β -glucanasas.

ES 2 303 058 T3

De la misma manera pertenecen a los enzimas para los agentes de lavado y de limpieza las óxidorreductasas, por ejemplo las oxidasas, las oxigenasas, las catalasas, las peroxidasas, tales como las halo-peroxidasas, las cloro-peroxidasas, las bromo-peroxidasas, las lignina-peroxidasas, las glucosa-peroxidasas o las manganeso-peroxidasas, las dioxigenasas o las laccasas (las fenoloxidasas, las polifenoloxidasas) o todos los otros enzimas descritos en el estado de la técnica para este campo de aplicación.

Tal como ya se ha indicado al principio, la tarea, en la que está basada la presente invención, se plantea de una forma especialmente evidente sobre el fundamento de la fermentación a escala industrial. Precisamente, en este campo se presentan los inconvenientes, por un lado, de que una selección a través de antibióticos es cara y, bajo ciertos puntos de vista medioambientales, es problemática, por otro lado tiene que compensarse una auxotrofia, basada en un defecto del metabolismo, debido a la complejidad de los medios industriales.

Así pues, tiene un significado especial la conversión de un procedimiento de selección de conformidad con la invención en un procedimiento a escala industrial, por ejemplo para la obtención de compuestos de bajo peso molecular tales como antibióticos o vitaminas o, de manera muy preferente, para la obtención de proteínas.

En el estado de la técnica se conocen, de manera general, procedimientos para la obtención de una proteína mediante el cultivo de células de una cepa de microorganismo. Éstos están basados en que las células, que producen la proteína interesada de manera natural o después de la transformación con un gen correspondiente, pueden cultivarse de manera adecuada y pueden ser estimuladas de manera preferente para la formación de la proteína interesada.

Otro objeto de la invención consiste, por lo tanto, en un procedimiento para la obtención de una proteína para el cultivo de células de una cepa de microorganismo, caracterizadas por un procedimiento de selección, que ha sido descrito precedentemente. Con esta finalidad, las formas de realización descritas más arriba señalan, de manera especial, que el vector de curación propiamente dicho contiene un transgen y que éste codifica de manera preferente una proteína no enzimática o codifica un enzima. Entre éstas deben entenderse, de manera especial, las proteínas de significación comercial. Así pues se utiliza la insulina, preparada por vía transgénica, para el tratamiento de la diabetes y un amplio espectro de enzimas, de proteasas, de lipasas y de amilasas hasta los enzimas oxidantes se emplean para la obtención de agentes de lavado y de limpieza.

Como consecuencia de la fácil manipulabilidad industrial, cuando se preparen el transgen interesado y el gen de curación sobre un vector, son preferentes aquellos procedimientos de obtención de proteínas, de conformidad con la invención, caracterizados porque el transgen según (c) codifique la proteína preparada por medio de este procedimiento.

En principio, pueden emplearse bacterias sobre una superficie sólida. Esto tiene un significado especial para ensayar sus propiedades metabólicas o para el cultivo duradero a escala de laboratorio. Por el contrario, son preferentes para la producción de proteínas aquellos procedimientos caracterizados porque el cultivo de los microorganismos se lleve a cabo en un medio líquido, preferentemente en un fermentador. Tales técnicas están ampliamente expandidas en el estado de la técnica y se mejoran, de conformidad con la invención, precisamente mediante la transposición a una selección mediante factores de translocación esenciales.

Tienen un significado especial los procedimientos para la obtención de proteínas caracterizados porque la proteína interesada es secretada en el medio circundante. Precisamente, de este modo, se facilita esencialmente la elaboración del producto obtenido. Una posible alternativa, de conformidad con la invención, consiste, sin embargo, en que las correspondientes células, que preparan la proteína, son sometidas a digestión después de la producción propiamente dicha y, de este modo, se obtiene el producto.

En principio se genera una nueva cepa por medio de cada modificación efectuada por biología molecular sobre un microorganismo. Por lo tanto, se preparan también nuevas cepas de microorganismos mediante la realización de la presente invención, precisamente aquellas que se diferencian de la cepa de partida (dicho de una manera más exacta: de las células de partida) mediante la inactivación específica de una actividad de translocación esencial y su curación mediante la preparación de un factor de translocación equivalente. Mediante el empleo de un procedimiento de selección, de conformidad con la invención, se preparan, por lo tanto, microorganismos novedosos.

En el ámbito de la protección de la presente solicitud quedan abarcados, por lo tanto, los microorganismos, que hayan sido obtenidos mediante un procedimiento de selección precedentemente descrito.

Un aspecto especialmente ventajoso consiste en que se obtiene una multitud de microorganismos emparentados cuando se lleve a cabo siempre el mismo tipo de la inactivación y de la curación preparándose, sin embargo, sobre el vector de curación en un transgen diferente en cada caso. De este modo, cuando un procedimiento haya sido aplicado con éxito una vez, podrá extrapolarse, de este modo, a un gran número de otros problemas de selección.

Para la realización, de manera especial, de los procedimientos para la obtención de proteínas, que han sido descritos precedentemente, es necesario que sea expresado el transgen. Los microorganismos preferentes, de conformidad con la invención, están caracterizados de este modo.

ES 2 303 058 T3

Entre éstos son preferentes los microorganismos, de acuerdo con lo que se ha dicho precedentemente con respecto a la obtención de las proteínas, caracterizados porque se secreta el transgen.

De conformidad con lo que se ha dicho precedentemente, tan solo pueden encontrarse los sistemas de selección, de conformidad con la invención, si se reconocen como tales los genes esenciales, que codifiquen actividades de translocación, mediante los cuales sea posible una selección. Un empleo, de este tipo, de estos genes no ha sido imaginado hasta el presente, aún cuando se conoce una pluralidad de los mismos procedentes de un gran número de microorganismos. Precisamente este conocimiento es beneficioso para los sistemas de selección de conformidad con la invención, puesto que, de este modo, pueden definirse prácticamente genes para todos los microorganismos, mediante los cuales sea posible una selección correspondiente. Con esta finalidad únicamente requieren ser inactivados, como se ha indicado precedentemente, y tienen que ser substituidos en las células correspondientes por medio de un homólogo funcional.

Por lo tanto, un objeto de la invención consiste en el empleo de un gen, que codifica una actividad de translocación esencial, para la selección de un microorganismo, caracterizado porque

- (a) se inactiva esta actividad de translocación esencial, existente de manera endógena en el microorganismo y
- (b) se cura, por medio de un vector, la actividad de translocación esencial, inactivada según (a).

Las formas de realización, indicadas a continuación, de este objeto de la invención son preferentes de manera correspondiente, de acuerdo con las explicaciones dadas hasta ahora.

Tiene un interés especial, desde el punto de vista de la genética molecular y desde el punto de vista industrial, que este empleo esté caracterizado porque

- (c) el vector, según (b), porte un transgen.

Un punto clave de la presente invención está constituido por aquél empleo correspondiente caracterizado porque la actividad de translocación esencial esté constituida por uno de los factores siguientes: SecA, SecY, SecE, SecD, SecF, peptidasa señal, b-SRP (Ffh o Ffs / Scr), FtsY / Srb, PrsA o YajC.

En este caso, es preferente aquél empleo que esté basado en la actividad de translocación esencial de una de las subunidades siguientes de la preproteína-translocasa: SecA, SecY, SecE, SecD o SecF, de manera preferente en la subunidad SecA.

Un empleo ventajoso consiste en que la curación según (b) se lleva a cabo a través de una actividad equivalente a la actividad de translocación esencial, presente de manera endógena, inactivada, de manera preferente a través de una actividad genéticamente emparentada, de manera especialmente preferente a través de la misma actividad.

Con esta finalidad, la presente solicitud proporciona un punto de partida correspondientemente preferente, en concreto las citadas posibilidades de aplicación, caracterizadas, de manera especial, porque la curación según (b) se lleva a cabo a través de la actividad de translocación de las regiones restablecedoras del gen *secA* a partir de *Bacillus subtilis*, *Escherichia coli* o de *Bacillus licheniformis*, que se han indicado en el protocolo de secuencias bajo SEQ ID NO. 1, SEQ ID NO. 3 o bien SEQ ID NO. 5.

Las aplicaciones, especialmente ventajosas, se caracterizan porque el vector, según (b), está constituido por un plásmido que se replica de manera autónoma en el microorganismo.

Otras aplicaciones preferentes se caracterizan porque el plásmido está constituido por un plásmido que se establece en varios números de copias, de manera preferente que se establece en un múltiple número de copias.

Finalmente se lleva a cabo la presente invención también mediante la preparación de los vectores correspondientes. En este caso se quiere indicar aquellos vectores que porten un gen para una actividad de translocación esencial y un transgen capaz de llevar a cabo la expresión, que, desde luego, no codifique una resistencia a los antibióticos cuando esté presente como transgen único.

En el sentido de la presente solicitud, no constituye una condición industrial la exclusión de genes para una resistencia a los antibióticos, cuando se trate del transgen único, además del gen que cura la actividad de translocación esencial, inactivada de conformidad con la invención. Únicamente debe tenerse en consideración que éstos hayan sido ya descritos precisamente en el estado de la técnica en relación con la caracterización de las proteínas de translocación, que pueden ser empleadas de conformidad con la invención. Esto comprende, evidentemente, también que hayan sido secuenciados y clonados y, concretamente, por medio de los vectores de clonación más ampliamente extendidos en el estado de la técnica, así como aquellos que contengan como marcadores genes para una resistencia a los antibióticos. De este modo, son conocidos en el estado de la técnica, aquellos vectores que contienen únicamente un gen, que codifica un enzima de translocación esencial y un marcador de los antibióticos.

ES 2 303 058 T3

De acuerdo con lo que se ha descrito precedentemente, un vector, de conformidad con la invención, se caracteriza de manera preferente porque el transgen contenido en el mismo, pensado de manera preferente para la obtención de proteínas, codifique una proteína no enzimática relevante desde el punto de vista farmacológico o que codifique un enzima hidrolítico o que codifique una óxidorreductasa. Éstos se caracterizan porque el gen, que debe ser expresado
5 está dotado con un promotor que funcione. En este caso quedan comprendidos en el ámbito de la protección, evidentemente, todos aquellos constructos, que codifiquen también factores - eventualmente de interés farmacológico -, que puedan promover una resistencia a los antibióticos, en tanto en cuanto la presencia de este vector no sea seleccionada por esta propiedad, sino que sea seleccionada por la actividad de translocación esencial.

10 De acuerdo con las explicaciones correspondientes al sistema de selección, determinados vectores representan también formas preferentes de realización de la presente invención.

A éstas pertenecen los vectores correspondientes caracterizados porque la actividad de translocación, codificada por los mismos, es capaz de curar una actividad de translocación esencial, inactivada, presente de manera endógena
15 en una cepa de microorganismos, preferentemente por medio de una actividad genéticamente emparentada, de manera especialmente preferente por medio de la misma actividad.

De manera más preferente, pertenecen a este respecto los vectores correspondientes caracterizados porque la actividad de translocación esencial está constituida por uno de los siguientes factores: SecA, SecY, SecE, SecD, SecF, peptidasa señal, b-SRP (Ffh o Ffs / Scr), FtsY / Srb, PrsA o YajC.
20

Entre éstos son preferentes aquellos vectores caracterizados porque la actividad de translocación esencial está constituida por una de las siguientes subunidades de la preproteína-translocasa: SecA, SecY, SecE, SecD o SecF, de manera preferente está constituida por la subunidad SecA.
25

De acuerdo con las enseñanzas de la presente solicitud, los vectores preferentes están caracterizados también porque la actividad de translocación esencial está constituida por uno de los genes *secA* de *Bacillus subtilis*, *Escherichia coli* o de *Bacillus licheniformis*, que se han indicado en el protocolo de secuencias bajo SEQ ID NO. 1, SEQ ID NO. 3 o bien SEQ ID NO. 5.
30

De igual modo, constituye una propiedad favorable de los vectores, de conformidad con la invención, que éstos estén constituidos por plásmidos que se repliquen de manera autónoma en el microorganismo empleado.

En este caso es especialmente ventajoso, de manera especial, debido al efecto de dosificación génica, que los plásmidos se caractericen porque estén constituidos en este caso por plásmidos establecidos con un número repetido
35 de copias, preferentemente con un número múltiple de copias.

Ejemplos

40 Todas las fases de trabajo de biología molecular siguen métodos normalizados, como los que se han dado por ejemplo en el manual de los autores Fritsch, Sambrook y Maniatis "Molecular cloning: a laboratory manual", Cold Spring Harbour Laboratory Press, New York, 1989, o en obras especializadas comparables. Se han empleado los enzimas y los estuches (kits) según las indicaciones del fabricante correspondiente.

Ejemplo 1

Aislamiento del gen secA a partir de B. licheniformis

Identificación del locus secA en B. licheniformis

50 Para la identificación del locus *secA/prfB* en *B. licheniformis* se derivó mediante RCP una sonda génica por medio de la secuencia conocida del locus génico *prfB-secA* de *B. subtilis* (banco de datos "Subtilist" del Institut Pasteur, 25,28 rue du Docteur Roux, 75724 Paris CEDEX 15, Francia; <http://genolist.pasteur.fr/Subtilist/>; estado a: 16.8.2002). Este locus génico se ha representado también en la figura 2. La sonda obtenida tenía una longitud de 3113 bp y comprendía además las primeras 451 bp de la región del extremo N del gen *prfB*. A continuación se someten a digestión preparaciones del ADN cromosómico de *B. licheniformis*, que puede adquirirse por ejemplo en la Colección Alemana de Microorganismos y de Cultivos Celulares (Deutschen Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig (<http://www.dsmz.de>) bajo el número de referencia 13, y ADN cromosómico de *B. subtilis*, con fines de control, con diversos enzimas de restricción y se somete a una hibridación
55 de Southern con dicha sonda. Se identificó un único fragmento con un tamaño aproximado de 5,5 kB en el ADN cromosómico de *B. licheniformis*, tratado con el enzima de restricción *MunI*, mientras que la digestión del ADN cromosómico de *B. subtilis* con *MunI* proporcionó los fragmentos esperados para *B. subtilis*.

Clonación de la región identificada de B. licheniformis DSM13

65 Se aisló ADN cromosómico de la misma cepa *B. licheniformis*, se sometió a digestión preparativa con *MunI* y la región de ADN de aproximadamente 5,5 kB se aisló mediante electroforesis con el de agarosa y de igual modo se extrajeron los ácidos nucleicos a partir de la misma con el estuche que puede ser adquirido en el comercio. La

ES 2 303 058 T3

mezcla obtenida de fragmentos de ADN cortados con *MunI* se ligaron en los puntos de corte *EcoRI*, compatibles con *MunI* de vector con un bajo número de copias (Low-copy-number) pHSG575 (descrito en la publicación: "High-copy-number and low-copy-number vectors for lacZ alpha-complementation and chloramphenicol- or kanamycin-resistance selection"; S. Takeshita; M. Sato; M. Toba; W. Masahashi; T. Hashimoto-Gotoh; Gene (1987), tomo 61, páginas 63-74) y se transformó en *E. coli* JM109 (que puede adquirirse en la firma Promega, Mannheim, Alemania).

Se seleccionó con respecto a la resistencia codificada por el vector. Adicionalmente el método del cribado azul/blanco (placas de selección que contenían 80 µg/ml de X-Gal) sirvió para la identificación de los clones, que habían alojado un vector con inserto. De este modo se obtuvieron 200 clones, entre los cuales pudieron identificarse 5 clones mediante la hibridación de colonias, que portaban el gen *B. licheniformis-secA*. Éstos se verificaron mediante un nuevo análisis por transferencia de Southern con la sonda que ha sido descrita precedentemente y se prosiguió con un vector derivado de pHSG575 con el fragmento *MunI* que porta el gen *secA* de *B. licheniformis*, que comprende 5,5 kB, bajo la denominación pMH1.

15 *Análisis por restricción*

La región clonada de 5,5 kB se caracterizó, en primer lugar, mediante mapeado por restricción. Con esta finalidad se llevaron a cabo digestiones simples y dobles con diversos enzimas de pMH1 y se identificaron mediante análisis por transferencia de Southern, aquellos fragmentos, que portan partes del operón *secA/prfB*. El mapa de restricción, resultante, se completó tras la secuenciación completa del fragmento 5,5 kB (véase más adelante) y pueden verse en la figura 3.

25 *Análisis de secuencias*

El fragmento mostrado en la figura 3, con un tamaño de 5,5 kB se secuenció en secuencias parciales según los métodos normalizados. Las secuencias parciales mostraron una fuerte homología con los genes siguientes de *B. subtilis*: *fliT* (que codifica una proteína flagelo), *orf189/yvyD* (función desconocida), *secA* (subunidad enlazadora de translocasa; ATPasa) y *prfB* (factor de desprendimiento de la cadena peptídica 2 -Peptide Chain Release Factor 2-), exactamente en el mismo orden génico que en el caso de *B. subtilis*. Éstos se han mostrado igualmente en la figura 3.

Como consecuencia de estos valores, también puede partirse de la base de que la SecA de *B. licheniformis* ejerce la misma actividad bioquímica que, especialmente, la SecA de *B. subtilis* y, por lo tanto, ejerce la misma función fisiológica. Por lo tanto debe considerarse como enzima esencial, enlazado en la translocación.

La secuencia de ADN y la secuencia de aminoácidos derivada de la misma, determinadas en este ejemplo, se han indicado en el protocolo de secuencias bajo SEQ ID NO. 5 o bien 6. Por lo tanto el inicio de la traducción se encuentra en la posición 154 y el codón de terminación se encuentra en las posiciones 2677 hasta 2679. Como región promotora debe ser considerada probablemente la secuencia parcial de las posiciones 60 hasta 65 o bien 77 hasta 82 y como punto de enlace de los ribosomas la región de las posiciones 138 hasta 144.

45 *Ejemplo 2*

Obtención de un plásmido con gen secA y gen de subtilisina

El gen *secA*, obtenido en el ejemplo 1, se amplificó con su propio promotor mediante RCP, a partir del ADN cromosómico de *B. licheniformis*. Con esta finalidad se eligieron cebadores, como se ha representado en la figura 4, por medio de la secuencia de ADN del gen de *B. licheniformis*, que tenían un punto de corte por restricción *BamHI* respectivamente en el extremo 5'. A través de éste se clonó el fragmento, amplificado con estos cebadores, en el punto de corte del plásmido pCB56C. Éste está descrito en la solicitud WO 91/02792 A1 y contiene el gen para la proteasa alcalina de *B. lentus* (*BLAP*).

Esta estrategia de clonación, que está representada también en la figura 4, proporciona el vector pCB56C*secA*, que tiene un tamaño de 8319 bp, que contiene además de los genes *secA* y *BLAP*, también un gen que codifica una resistencia a la tetraciclina.

Se transformaron en *B. licheniformis* este vector pCB56C*secA* y el pCB56C, para el control del vector de partida, y, concretamente, en el caso de pCB56C en la cepa de tipo silvestre *B. licheniformis* (*secA*) capaz de formar la SecA. En el caso del pCB56C*secA* se llevó a cabo la transformación de tal manera, que se inactivó al mismo tiempo el endógeno *secA*. En el ejemplo 3 se ha descrito la manera de proceder con esta finalidad.

De este modo se obtuvieron las dos cepas *B. licheniformis* (Δ *secA*) pCB56C*secA* y *B. licheniformis* (*secA*) pCB56C que serían capaces ambas de expresar el gen codificado a través del plásmido para la proteasa alcalina. Éstas se investigan adicionalmente en el ejemplo 4.

Ejemplo 3

Obtención de la cepa B. licheniformis ($\Delta secA$) pCB56CsecA

5 La desconexión del gen *secA* se llevó a cabo por medio de un vector de delección. La forma de proceder siguió la descripción de los autores J.Vehmaanperä *et al.* (1991) en J. Biotechn., tomo 19, páginas 221-240.

Como vector para la delección *secA* se eligió el plásmido pE194 descrito en la misma publicación. La ventaja de este vector de delección consiste en que éste tiene un origen de replicación en función de la temperatura. A 33°C puede replicarse el pE194 en la célula de tal manera, que puede seleccionarse a esta temperatura en primer lugar con respecto a una transformación realizada con éxito. A continuación se incuban las células, que contienen el vector, a 42°C. A esta temperatura ya no se replica el vector de delección y se ejerce una presión de selección sobre la integración del plásmido a través de una de las dos regiones homólogas (región en el sentido ascendente 5' o en el sentido descendente 3' -upstream o downstream- de *secA*) en el cromosoma. Otra recombinación homóloga a través de otras (segundas) regiones homólogas conduce entonces a la delección *secA*. De igual modo, sería posible una nueva recombinación de la primera región homóloga. En este caso se recombina el vector de nuevo fuera del cromosoma de tal manera, que se conservaría el *secA* cromosómico. La delección *secA* tiene que demostrarse, por lo tanto, en el análisis por transferencia de Southern tras la restricción del ADN cromosómico con enzimas adecuados o con ayuda de la técnica de la RCP en función del tamaño de la región amplificada.

Para la construcción del vector de delección se amplificaron las regiones en el sentido ascendente 5' y en el sentido descendente 3' (upstream y downstream) de *secA* (figura 5) por medio de la RCP. Los cebadores para la amplificación y los puntos de corte por restricción, relacionados con los mismos, para la clonación subsiguiente (*XbaI* y *EcoRV*) se eligieron por medio de la secuencia de ADN, determinada según el ejemplo 1, del locus *secA/prfB* de *B. licheniformis*. En el caso de la delección *secA* debe tenerse en consideración que el *prfB*, localizado en el sentido descendente 3' (downstream) de *secA*, se encuentra en un operón con *secA*, es decir que no tiene un promotor propio (véase la figura 2). El *prfB* codifica la proteína RF2, que se encarga, en conexión con la biosíntesis de proteína, para desprender la proteína del ribosoma. En el fin de garantizar la transcripción del *prfB*, importante para las biosíntesis de proteína, incluso tras la delección *secA*, se amplificó el *orf189*, que se encuentra por delante del *secA*, con terminador propio y con el promotor *secA*, localizado en el sentido descendente 3' (downstream) de tal manera, que el *prfB* puede transcribirse (figura 5), tras la delección *secA*, directamente a partir del promotor *secA*.

Las regiones amplificadas (*orf189'* y *prfB'*) se sometieron a una clonación intermedia en una etapa de control en el vector pBBRMCS2 de *E. coli*. La secuenciación subsiguiente del constructo *orf189'prfB'* mostró que los fragmentos amplificados se había clonado conjuntamente de manera correcta.

En la siguiente etapa, el constructo *orf189'prfB'* se sometió a una reclonación en *B. subtilis* DB104 en el vector pE 194, elegido para la delección (figura 6). En este caso, se obtuvieron productos de transformación con el método de la transformación de protoplastos según Chang & Cohen, 1979, que portaban el vector de delección pEorfprfB. Todas las etapas de trabajo se llevaron a cabo a 33°C para garantizar una replicación del vector.

En una última etapa se transformó el vector pCB56CsecA, descrito en el ejemplo 2, de igual modo por medio del método de la transformación de protoplastos, en la cepa huésped *B. licheniformis*, que porta el plásmido pEorfprfB. Los productos de transformación obtenidos de es modo e identificados como positivos con métodos usuales, se seleccionaron a continuación a 42°C bajo presión de selección (tetraciclina para pCB56CsecA y eritromicina para pEorfprfB) con respecto a la presencia de ambos plásmidos. A esta temperatura, el vector de delección ya no puede replicarse y únicamente sobreviven aquellas células en las que el vector está integrado en el cromosoma, teniendo lugar esta integración, con la máxima probabilidad, en regiones homólogas o idénticas. Mediante el cultivo a 33°C, sin presión de selección por parte de la eritromicina, puede inducirse a continuación la excisión del vector de delección, retirándose por completo del cromosoma el gen *secA* codificado por vía cromosómica.

En la célula permanece el plásmido pCB56CsecA, que induce la capacidad de la síntesis de la subtilisina así como pone a disposición el factor de translocación esencial SecA. La cepa, obtenida de este modo, se denominó *B. licheniformis* ($\Delta secA$) pCB56CsecA.

Ejemplo 4

Ensayo de la estabilidad del plásmido

Para la determinación de la estabilidad genética del plásmido de subtilisina pCB56CsecA, que porta el *secA*, se ensayaron las dos cepas *B. licheniformis* (*secA*) pCB56C y *B. licheniformis* ($\Delta secA$) pCB56CsecA, obtenidas según los ejemplos 2 y 3, en un experimento en matraz vibrador sin aporte de antibióticos en medio líquido. Con esta finalidad se realizó un cultivo durante la noche, a partir respectivamente de una sola colonia y se inocularon con éste respectivamente 14 ml de medio LB (según receta normalizada) hasta una densidad óptica de 0,05 a 600 nm (OD₆₀₀). El cultivo se llevó a cabo en un matraz vibrador de Erlenmeyer de 100 ml. Respectivamente al cabo de 8 y de 16 horas se sobreinocularon los cultivos con 14 ml de medio fresco y en este caso se ajustó, de nuevo, una OD₆₀₀ de 0,05. El cultivo se llevó a cabo durante 8 días y noches; en este caso se sobreinocularon los cultivos en total 16 veces. Se

ES 2 303 058 T3

distribuyeron a diario sobre placas series de diluciones y se transfirió una elección aleatoria de los clones obtenidos sobre placas de ensayo de proteasa. El resultado se ha representado en la tabla 2 y en la figura 7.

TABLA 2

Estabilidad del plásmido en los productos de transformación B. licheniformis (secA) pCB56C (controles) y B. licheniformis (Δ secA) pCB56CsecA, que pueden detectarse por medio de la proporción correspondiente de los clones con actividad de proteasa

Tiempo [días]	<i>B. licheniformis (secA) pCB56C</i>			<i>B. licheniformis (ΔsecA) pCB56CsecA</i>		
	Número de los clones ensayados		Proporción de los clones con actividad de proteasa [%]	Número de los clones ensayados		Proporción de los clones con actividad de proteasa [%]
	total	sin actividad de proteasa		total	sin actividad de proteasa	
1	72	0	100	72	0	100
2	72	0	100	70	0	100
3	72	1	98,6	49	0	100
4	52	0	100	52	0	100
5	78	0	100	78	0	100
6	78	1	98,7	78	0	100
7	78	1	98,7	77	0	100
8	104	6	94,2	104	0	100

Todos los clones ensayados de la cepa *B. licheniformis (Δ secA) pCB56CsecA* mostraron actividad de proteasa en cualquier instante del cultivo, mientras que en el caso de la cepa *B. licheniformis (secA) pCB56C* algunos clones ya no tenían actividad de proteasa. Esto debe interpretarse como pérdida del plásmido pCB56C; esta pérdida se verificó de manera adicional mediante minipreparación de plásmido.

Por lo tanto, se pone claramente de manifiesto, por medio de estos datos, que el plásmido de subtilisina pCB56CsecA, que porta *secA*, es estable en la cepa Δ secA cuando se verifica el cultivo en medio LB sin aporte de antibióticos, especialmente sin aporte de tetraciclina, para la cual proporcionaría resistencia el plásmido, mientras que el plásmido de subtilisina pCB56C se pierde en la cepa sin delección *secA* cromosómica, en el transcurso del cultivo. Por lo tanto, el gen *secA* procedente de *B. licheniformis* puede curar la deficiencia SecA cromosómica por transferencia a un vector de expresión y, de este modo, puede ser aprovechado para la selección de un cultivo bacteriano, que exprese un gen para otra proteína, en este caso una proteasa alcalina.

Descripción de las figuras

Figura 1

5 *Representación esquemática del aparato de traducción/translocación de bacterias grampositivas*

De conformidad con la publicación de los autores van Wely, K.H., Swaving, J., Freudl, R., Driessen, A.J. (2001); "Translocation of proteins across the cell envelope of Gram-positive bacteria", FEMS Microbiol Rev. 2001, 25(4), S. 437-54.

10

Figura 2

Locus génico de secA en B. subtilis

15 Puede verse que en la región *secA* se encuentra también el gen *prfB* y que se forma un ARNm coherente, de tal manera que puede hablarse también de un operón *secA/prfB*.

Figura 3

20 *Mapa de restricción del locus génico orf189 / secA / prfB en B. licheniformis*

Como se ha representado en el ejemplo 1 se encuentra en la proximidad inmediata con respecto al *secA* el gen *prfB* y un *orf* sobre un fragmento con un tamaño aproximado de 5,5 kB, que puede ser obtenido mediante restricción con *MunI* a partir del ADN genómico de *B. licheniformis*.

25

Figura 4

Obtención de un plásmido con gen secA y gen de subtilisina

30 Como se ha descrito en el ejemplo 2 se amplificó el *secA* mediante a RCP y se clonó un vector, que contiene como transgen la proteasa alcalina procedente de *B. lentus*.

Figura 5

35 *Regiones amplificadas mediante la RCP de secA (sentido ascendente 5' y sentido descendente 3' -upstream y downstream-)*

Amplificación de las regiones en el sentido ascendente 5' y en el sentido descendente 3' (upstream y downstream) de *secA* con los puntos de corte por restricción para elegidos para la clonación como se ha descrito en el ejemplo 3. El extremo 3' de *orf189* se amplifica con el propio terminador y con el promotor *secA* que se encuentra en el sentido 3' (downstream) de tal manera que, tras la delección *secA* puede transcribirse el *prfB* directamente desde el promotor *secA*. Los segmentos derivados *orf189'* y *prfB'* abarcan, respectivamente, 502 bp o bien 546 bp.

40

Figura 6

45 *Construcción del plásmido de delección pEorfprfB*

Las regiones, amplificadas por medio de la RCP, se clonaron de manera intermedia en *E.coli*, se cortaron de nuevo por medio de *XbaI* y de *EcoRV* y, a continuación, se ligaron en los puntos de corte por restricción *XbaI* y *AccI* en el vector pE194.

50

Figura 7

55 *Estabilidad del plásmido en los productos de transformación B. licheniformis (secA) pCB56C (controles) y B. licheniformis (Δ secA) pCB56CsecA*

Tal como se ha descrito en el ejemplo 4, se ha representado respectivamente la proporción de los clones con actividad de proteasa al cabo de un número de días correspondiente.

60

Cuadrados:

B. licheniformis (Δ secA) pCB56CsecA

65

Triángulos:

B. licheniformis (secA) pCB56C (controles).

REIVINDICACIONES

1. Procedimiento para la selección de un microorganismo, **caracterizado** porque

- 5 (a) se inactiva un gen endógeno presente, que codifica una actividad de translocación esencial y
(b) se cura, por medio de un vector, la actividad de translocación esencial, inactivada según (a).

10 2. Procedimiento según la reivindicación 1, **caracterizado** porque

- (c) el vector, según (b), porta un trasgen.

15 3. Procedimiento según la reivindicación 1 o 2, **caracterizado** porque la actividad de translocación esencial está constituida por uno de los factores siguientes: SecA, SecY, SecE, SecD, SecF, peptidasa señal, b-SRP (Ffh o Ffs / Scr), FtsY / Srb, PrsA o YajC.

20 4. Procedimiento según la reivindicación 3, **caracterizado** porque la actividad de translocación esencial está constituida por una de las subunidades siguientes de la preproteína-translocasa: SecA, SecY, SecE, SecD o SecF, de manera preferente está constituida por la subunidad SecA.

25 5. Procedimiento según una de las reivindicaciones 1 a 4, **caracterizado** porque la curación según (b) se lleva a cabo a través de una actividad equivalente a la actividad de translocación esencial, presente de manera endógena, inactivada, preferentemente por medio de una actividad genéticamente emparentada, de manera especialmente preferente por medio de la misma actividad.

30 6. Procedimiento según una de las reivindicaciones 1 a 5, **caracterizado** porque la curación según (b) se lleva a cabo por medio de las regiones, que restablecen la actividad de translocación, del gen *secA* procedente de *Bacillus subtilis*, *Escherichia coli* o de *Bacillus licheniformis*, que se han indicado en el protocolo de secuencias bajo SEQ ID NO. 1, SEQ ID NO. 3 o bien SEQ ID NO. 5.

35 7. Procedimiento según una de las reivindicaciones 2 a 6, **caracterizado** porque la inactivación según (a) se lleva a cabo de tal manera, que se impida una recombinación entre la región génica, inactivada según (a), y la región homóloga sobre el vector según (b), de manera preferente con pérdida completa de un segmento génico, contenido en el gen cromosómico correspondiente.

40 8. Procedimiento según una de las reivindicaciones 1 a 7, **caracterizado** porque la inactivación, según (a), se lleva a cabo por medio de un vector de delección, de manera preferente por medio de un vector de delección con un origen de replicación que puede ser regulado de manera externa, de manera especialmente preferente por medio de un vector de delección con un origen de la replicación dependiente de la temperatura.

45 9. Procedimiento según una de las reivindicaciones 1 a 8, **caracterizado** porque el vector, según (b), está constituido por un plásmido, que se replica de manera autónoma en el microorganismo.

50 10. Procedimiento según la reivindicación 9, **caracterizado** porque el plásmido está constituido por un plásmido, que se establece un en número repetido de copias, preferentemente con un número múltiple de copias.

55 11. Procedimiento según una de las reivindicaciones 1 a 10, **caracterizado** porque el microorganismo está constituido por una cepa bacteriana gramnegativa.

60 12. Procedimiento según la reivindicación 11, **caracterizado** porque se trata de una cepa bacteriana gramnegativa de los géneros *E. coli* o *Klebsiella*, de manera especial se trata de derivados de *Escherichia coli* K12, de *Escherichia coli* B o de *Klebsiella planticola*, y de una manera muy especialmente preferente se trata de derivados de las cepas *Escherichia coli* BL21 (DE3), *E. coli* RV308, *E. coli* DH5 α , *E. coli* JM109, *E. coli* XL-1 o *Klebsiella planticola* (Rf).

65 13. Procedimiento según una de las reivindicaciones 1 a 10, **caracterizado** porque el microorganismo está constituido por una cepa bacteriana grampositiva.

70 14. Procedimiento según la reivindicación 13, **caracterizado** porque se trata de una cepa bacteriana grampositiva de los géneros *Staphylococcus*, *Corynebacterien* o *Bacillus*, de manera especial se trata de las especies *Staphylococcus carnosus*, *Corynebacterium glutamicum*, *Bacillus subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, *B. globigii* o *B. lentus*, y, de manera muy especialmente preferente, se trata de derivados de las cepas *B. licheniformis* o *B. amyloliquefaciens*.

75 15. Procedimiento según una de las reivindicaciones 2 a 14, **caracterizado** porque el transgen, según (c), está constituido por un transgen que codifica una proteína no enzimática, especialmente que codifica una proteína relevante desde el punto de vista farmacológico, de manera especial que codifica insulina o calcitonina.

ES 2 303 058 T3

16. Procedimiento según una de las reivindicaciones 2 a 14, **caracterizado** porque el transgen, según (c), está constituido por un transgen que codifica un enzima, preferentemente un enzima hidrolítico o una óxidoreductasa, de manera especialmente preferente codifica una proteasa, una amilasa, una hemicelulasa, una celulasa, una lipasa, una cutinasa, una oxidasa, una peroxidasa o una laccasa.

5

17. Procedimiento para la obtención de una proteína mediante cultivo de células de una cepa de microorganismo, **caracterizado** por un procedimiento de selección según una de las reivindicaciones 1 a 16.

18. Procedimiento según la reivindicación 17, **caracterizado** por un procedimiento de selección según una de las reivindicaciones 2 a 16, en el que el transgen, según (c), codifica la proteína preparada por medio de este procedimiento.

10

19. Procedimiento según una de las reivindicaciones 17 a 18, **caracterizado** porque se lleva a cabo el cultivo de los microorganismos en un medio líquido, de manera preferente en un fermentador.

15

20. Procedimiento según una de las reivindicaciones 17 a 19, **caracterizado** porque la proteína interesada es secretada en el medio circundante.

21. Microorganismo, obtenido mediante un procedimiento de selección según una de las reivindicaciones 1 a 16.

20

22. Microorganismo según la reivindicación 21, **caracterizado** porque el transgen es expresado.

23. Microorganismo según la reivindicación 22, **caracterizado** porque el transgen es secretado.

25

24. Empleo de un gen, que codifica una actividad de translocación esencial para la selección de un microorganismo, **caracterizado** porque

(a) se inactiva en el microorganismo esta actividad de translocación esencial, presente de manera endógena y

(b) se cura, por medio de un vector, la actividad de translocación esencial, inactivada según (a).

30

25. Empleo según la reivindicación 24, **caracterizado** porque

(c) el vector, según (b), porta un transgen.

35

26. Empleo según la reivindicación 24 o 25, **caracterizado** porque la actividad de translocación esencial está constituida por uno de los factores siguientes: SecA, SecY, SecE, SecD, SecF, peptidasa señal, b-SRP (Ffh o Ffs / Scr), FtsY / Srb, PrsA o YajC.

40

27. Empleo según la reivindicación 26, **caracterizado** porque la actividad de translocación esencial está constituida por una de las siguientes subunidades de la preproteína-translocasa: SecA, SecY, SecE, SecD o SecF, de manera preferente está constituida por la subunidad SecA.

45

28. Empleo según una de las reivindicaciones 24 a 27, **caracterizado** porque la curación, según (b), se lleva a cabo por medio de una actividad equivalente a la actividad de translocación esencial, presente de manera endógena, inactivada, de manera preferente se lleva a cabo por medio de una actividad genéticamente emparentada, de manera especialmente preferente se lleva a cabo por medio de la misma actividad.

50

29. Empleo según una de las reivindicaciones 24 a 28, **caracterizado** porque la curación, según (b), se lleva a cabo por medio de las regiones, que restablecen de la actividad de translocación, del gen *secA* procedente de *Bacillus subtilis*, *Escherichia coli* o de *Bacillus licheniformis*, que se han indicado en el protocolo de secuencias bajo SEQ ID NO. 1, SEQ ID NO. 3 o bien SEQ ID NO. 5.

55

30. Empleo según una de las reivindicaciones 24 a 29, **caracterizado** porque el vector, según (b), está constituido por un plásmido, que se replica de manera autónoma en el microorganismo.

60

31. Empleo según la reivindicación 30, **caracterizado** porque el plásmido está constituido por un plásmido, que se establece con un número repetido de copias, de manera preferente con un número múltiple de copias.

65

32. Vector, que porta un gen para una actividad de translocación esencial y un transgen, que no codifica una resistencia a los antibióticos, cuando esté presente como transgen único.

33. Vector según la reivindicación 32, **caracterizado** porque el transgen codifica una proteína no enzimática farmacológicamente relevante o codifica un enzima hidrolítico o codifica una óxidoreductasa.

70

34. Vector según la reivindicación 32 o 33, **caracterizado** porque la actividad de translocación, codificada por el mismo, es capaz de curar una actividad de translocación esencial inactivada, presente de manera endógena en una cepa de microorganismo, preferentemente por medio de una actividad genéricamente emparentada, de manera especialmente preferente por medio de la misma actividad.

ES 2 303 058 T3

35. Vector según una de las reivindicaciones 32 a 34, **caracterizado** porque la actividad de translocación esencial está constituida por uno de los factores siguientes: SecA, SecY, SecE, SecD, SecF, peptidasa señal, b-SRP (Ffh o Ffs / Scr), FtsY / Srb, PrsA o YajC.

5 36. Vector según la reivindicación 35, **caracterizado** porque la actividad de translocación esencial está constituida por una de las siguientes subunidades de la preproteína-translocasa: SecA, SecY, SecE, SecD o SecF, de manera preferente está constituida por la subunidad SecA.

10 37. Vector según una de las reivindicaciones 32 a 36, **caracterizado** porque la actividad de translocación esencial está constituida por uno de los genes *secA* procedentes de *Bacillus subtilis*, *Escherichia coli* o de *Bacillus licheniformis*, que han sido indicados en el protocolo de secuencias bajo SEQ ID NO. 1, SEQ ID NO. 3 o bien SEQ ID NO. 5.

15 38. Vector según una de las reivindicaciones 32 a 37, **caracterizado** porque se trata de un plásmido, que se replica de manera autónoma en el microorganismo.

39. Vector según la reivindicación 38, **caracterizado** porque el plásmido está constituido por un plásmido, que se establece con un número repetido de copias, de manera preferente se establece con un número múltiple de copias.

20

25

30

35

40

45

50

55

60

65

Figura 1

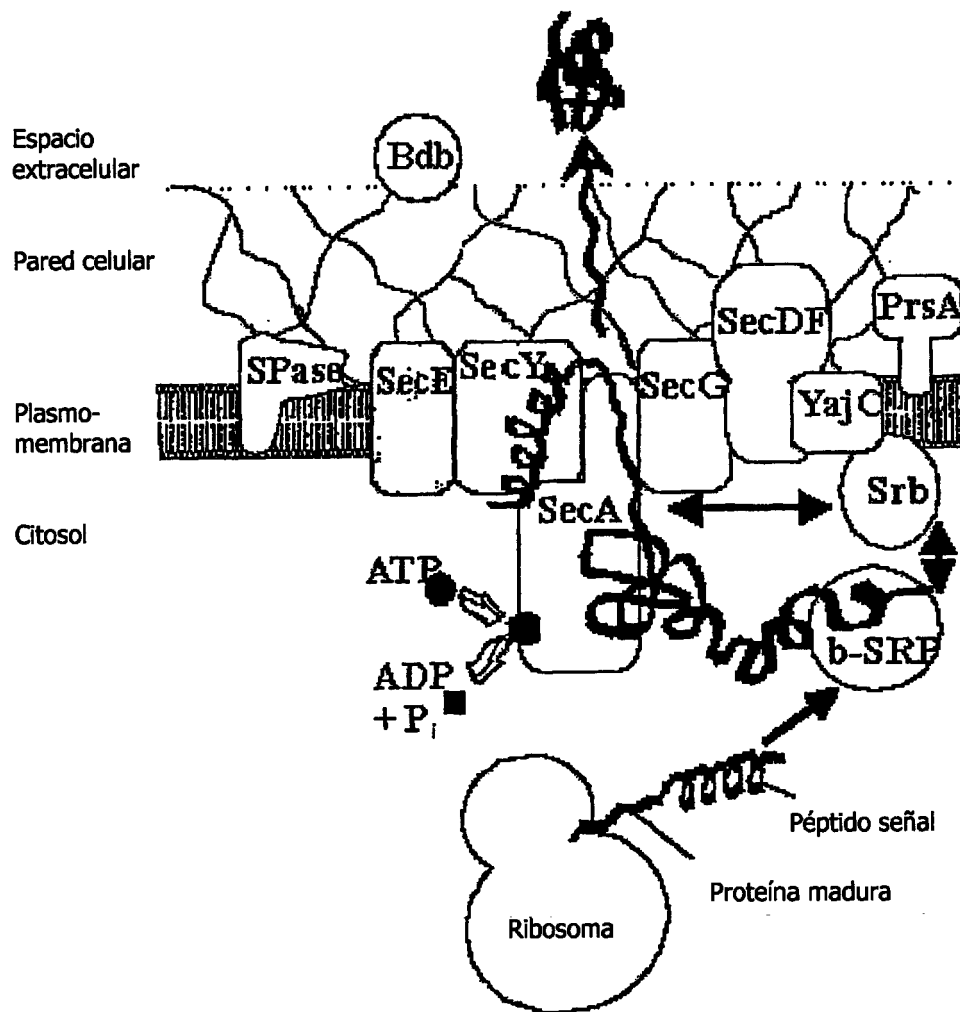


Figura 2

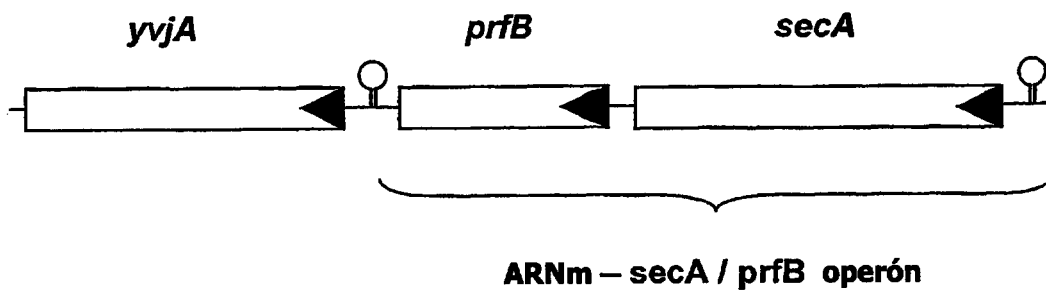


Figura 3

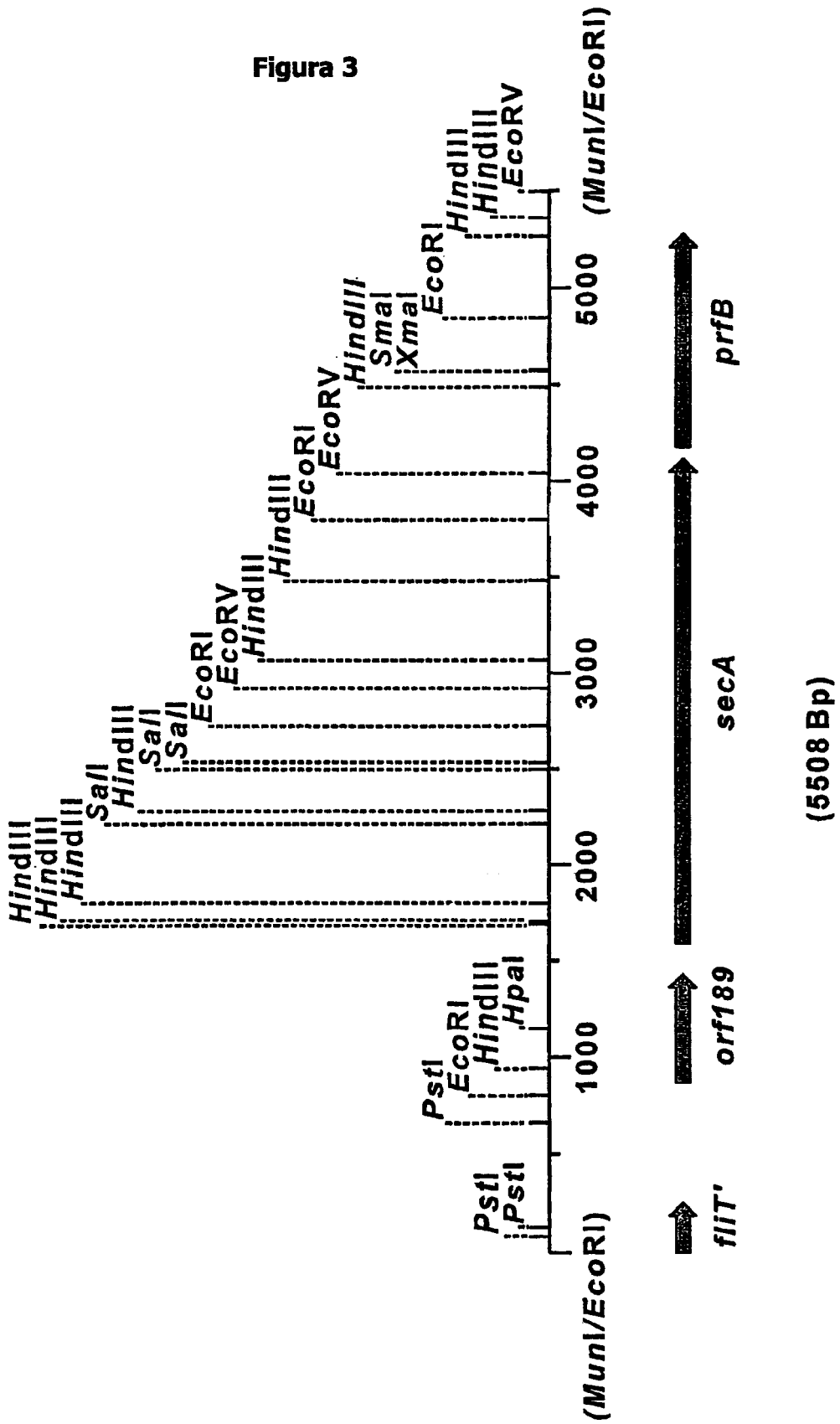


Figura 4

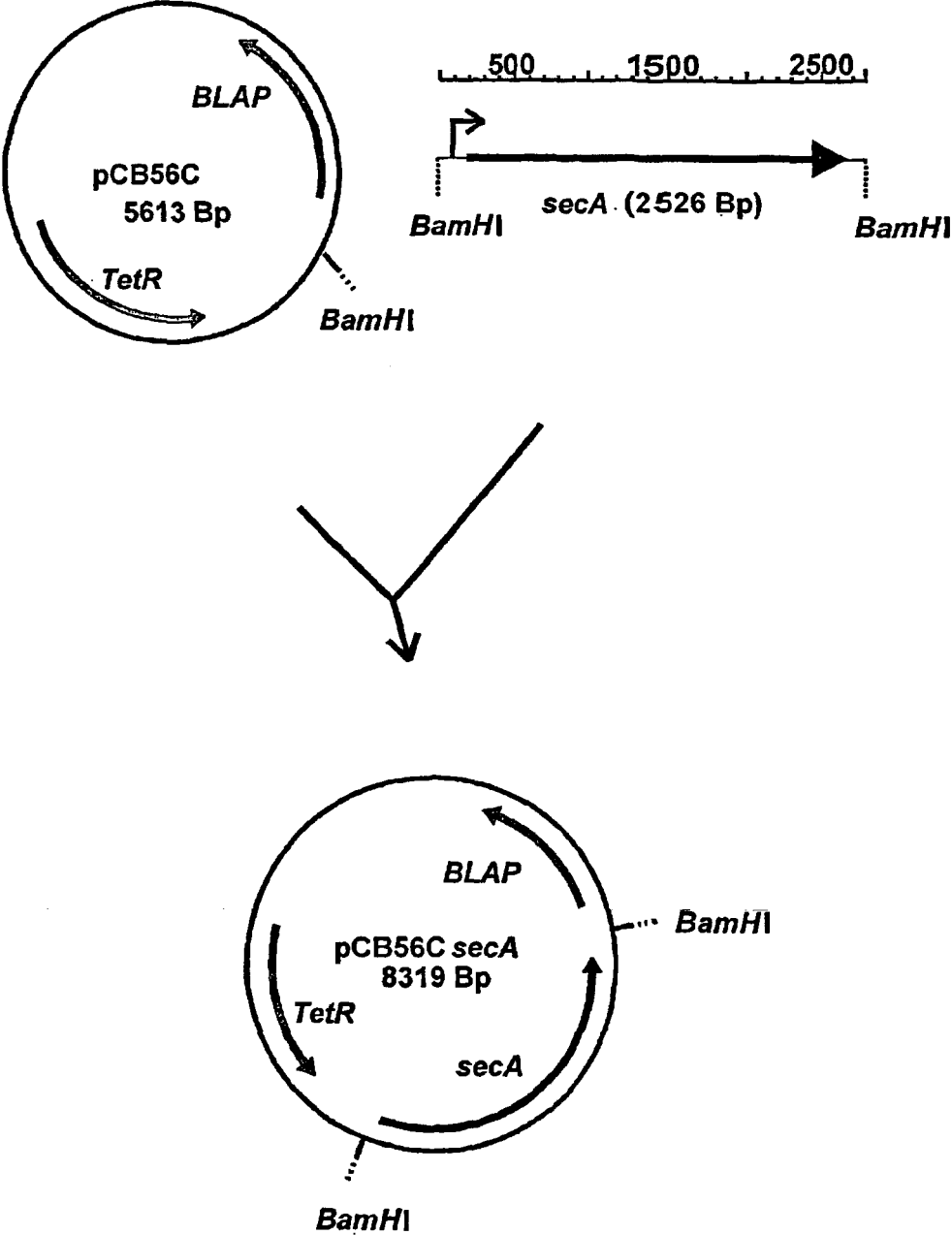


Figura 5

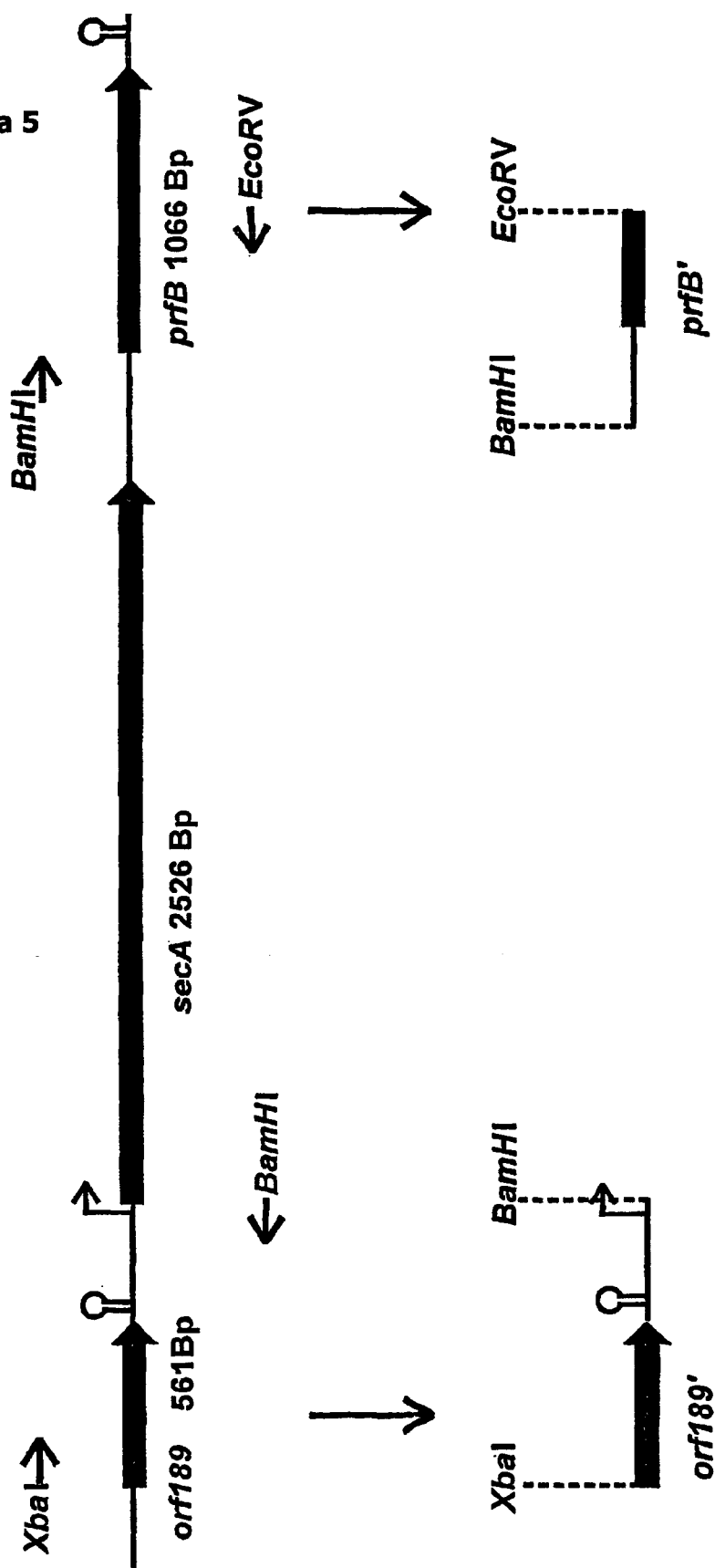


Figura 6

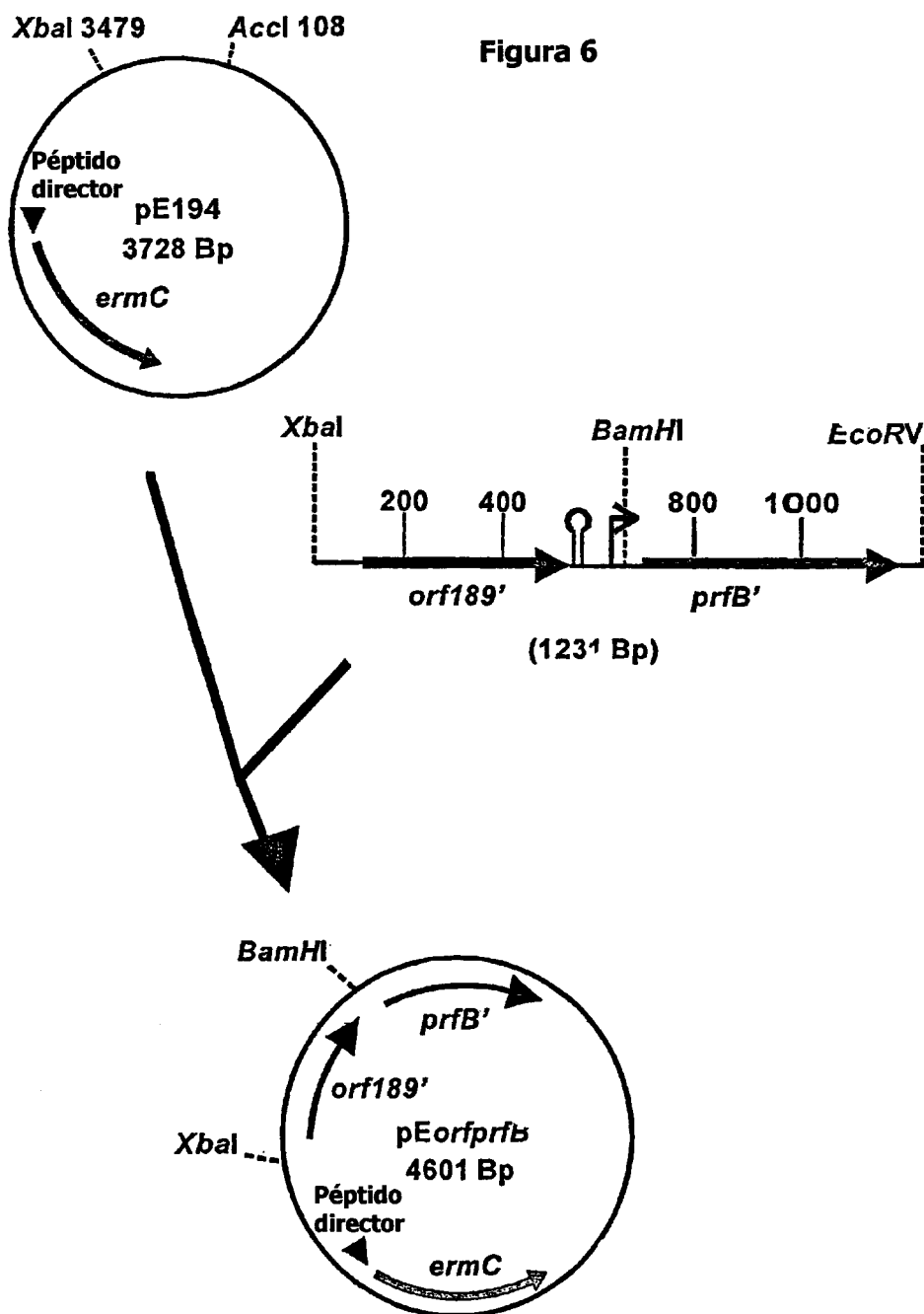
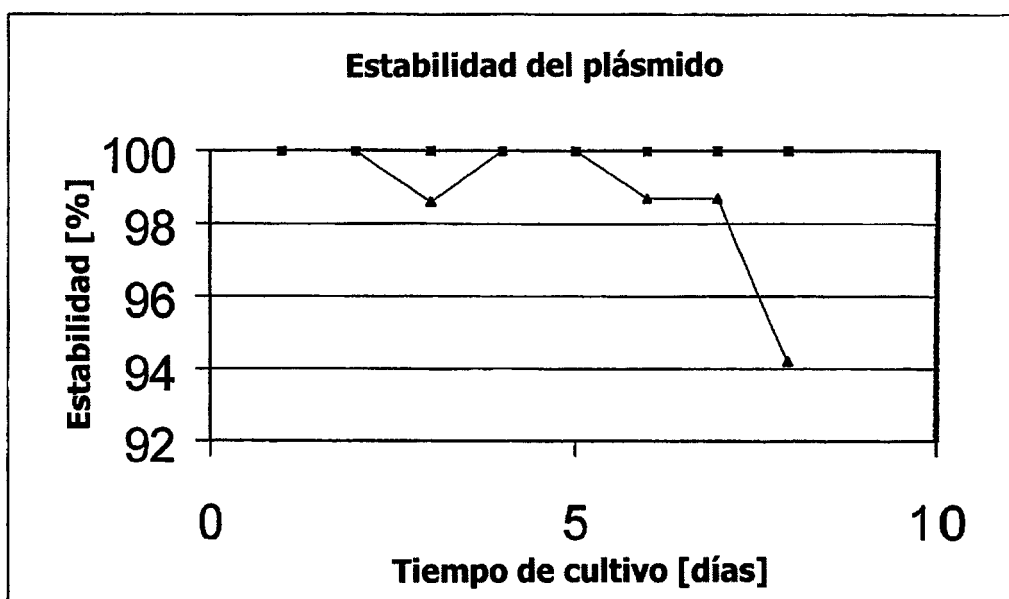


Figura 7



ES 2 303 058 T3

LISTA DE SECUENCIAS

<110> Henkel Kommanditgesellschaft auf Aktien
<120> Un enzima de translocación como marcador de selección
5 <130> H 05599 PCT
<140>
<141>
10 <150> DE10309557.8
<151> 2003-03-04
<160> 6
<170> PatentIn Ver. 2.1
15 <210> 1
<211> 4148
<212> DNA
20 <213> *Bacillus subtilis*
<220>
<221> CDS
<222> (544)..(3069)
25 <400> 1

```
30      gtccgagggtg cataacgagg atatgtacaa cgcaattgat ctcgcaacaa acaaactgga 60
      acgtcaaadc cgtaagcata aaacgaaagt aaaccgtaaa ttccgtgagc agggctctcc 120
      aaaatattta ttggcaaacg gtcttggtc tgatacagat attgcggttc aggatgacat 180
35      agaagaggag gagagcttgg acatcgtccg tcagaaacgc ttaatttaa agccgatgga 240
      tagtgaagaa gcgatcttgc aatgaatat gctcggccat aatttctttg ttttcacaaa 300
      tgcggaaaca aaccttaca atgtcgtgta ccgcagaaat gacgggaaat atggcttaat 360
40      tgaaccgact gaataatgaa gagaagcctt ccgtgatgtc cgcggaaggt ttttgTTTTT 420
      cttatttgca aattctttgg aaataacaaa aggtatgata tgataatgag aggtatacat 480
45      ggactagtaa attatttata catgcctcta aaataggcgt gtgatgatag aggagcgtaa 540
      taa atg ctt gga att tta aat aaa atg ttt gat cca aca aaa cgt acg 588
      Met Leu Gly Ile Leu Asn Lys Met Phe Asp Pro Thr Lys Arg Thr
      1 5 10 15
50      ctg aat aga tac gaa aaa att gct aac gat att gat gcg att cgc gga 636
      Leu Asn Arg Tyr Glu Lys Ile Ala Asn Asp Ile Asp Ala Ile Arg Gly
      20 25 30
55      gac tat gaa aat ctc tct gac gac gca ttg aaa cat aaa aca att gaa 684
      Asp Tyr Glu Asn Leu Ser Asp Asp Ala Leu Lys His Lys Thr Ile Glu
      35 40 45
      ttt aaa gag cgt ctt gaa aaa ggg gcg aca acg gat gat ctt ctt gtt 732
```

65

ES 2 303 058 T3

	Phe	Lys	Glu	Arg	Leu	Glu	Lys	Gly	Ala	Thr	Thr	Asp	Asp	Leu	Leu	Val	
			50					55					60				
5	gaa	gct	ttc	gct	ggt	ggt	cga	gaa	gct	tca	cgc	cgc	gta	aca	ggc	atg	780
	Glu	Ala	Phe	Ala	Val	Val	Arg	Glu	Ala	Ser	Arg	Arg	Val	Thr	Gly	Met	
		65					70					75					
10	ttt	ccg	ttt	aaa	gtc	cag	ctc	atg	ggg	ggc	gtg	gcg	ctt	cat	gac	gga	828
	Phe	Pro	Phe	Lys	Val	Gln	Leu	Met	Gly	Gly	Val	Ala	Leu	His	Asp	Gly	
		80				85					90					95	
15	aat	ata	gcg	gaa	atg	aaa	aca	ggg	gaa	ggg	aaa	aca	tta	acg	tct	acc	876
	Asn	Ile	Ala	Glu	Met	Lys	Thr	Gly	Glu	Gly	Lys	Thr	Leu	Thr	Ser	Thr	
					100					105					110		
20	ctg	cct	ggt	tat	tta	aat	gcg	tta	acc	ggt	aaa	gcg	gta	cac	gtc	gtg	924
	Leu	Pro	Val	Tyr	Leu	Asn	Ala	Leu	Thr	Gly	Lys	Gly	Val	His	Val	Val	
				115					120					125			
25	act	gtc	aac	gaa	tac	ttg	gca	agc	cgt	gac	gct	gag	caa	atg	ggg	aaa	972
	Thr	Val	Asn	Glu	Tyr	Leu	Ala	Ser	Arg	Asp	Ala	Glu	Gln	Met	Gly	Lys	
			130					135					140				
30	att	ttc	gag	ttt	ctc	ggt	ttg	act	gtc	ggt	ttg	aat	tta	aac	tca	atg	1020
	Ile	Phe	Glu	Phe	Leu	Gly	Leu	Thr	Val	Gly	Leu	Asn	Leu	Asn	Ser	Met	
		145					150					155					
35	tca	aaa	gac	gaa	aaa	cgg	gaa	gct	tat	gcc	gct	gat	att	act	tac	tcc	1068
	Ser	Lys	Asp	Glu	Lys	Arg	Glu	Ala	Tyr	Ala	Ala	Asp	Ile	Thr	Tyr	Ser	
						165					170					175	
40	aca	aac	aac	gag	ctt	ggc	ttc	gac	tat	ttg	cgt	gac	aat	atg	ggt	ctt	1116
	Thr	Asn	Asn	Glu	Leu	Gly	Phe	Asp	Tyr	Leu	Arg	Asp	Asn	Met	Val	Leu	
					180					185					190		
45	tat	aaa	gag	cag	atg	ggt	cag	cgc	ccg	ctt	cat	ttt	gcg	gta	ata	gat	1164
	Tyr	Lys	Glu	Gln	Met	Val	Gln	Arg	Pro	Leu	His	Phe	Ala	Val	Ile	Asp	
				195					200					205			
50	gaa	ggt	gac	tct	att	tta	att	gat	gaa	gca	aga	aca	ccg	ctt	atc	att	1212
	Glu	Val	Asp	Ser	Ile	Leu	Ile	Asp	Glu	Ala	Arg	Thr	Pro	Leu	Ile	Ile	
			210					215					220				
55	tct	gga	caa	gct	gca	aaa	tcc	act	aag	ctg	tac	gta	cag	gca	aat	gct	1260
	Ser	Gly	Gln	Ala	Ala	Lys	Ser	Thr	Lys	Leu	Tyr	Val	Gln	Ala	Asn	Ala	
		225				230						235					
60	ttt	gtc	cgc	acg	tta	aaa	gcg	gag	aag	gat	tac	acg	tac	gat	atc	aaa	1308
	Phe	Val	Arg	Thr	Leu	Lys	Ala	Glu	Lys	Asp	Tyr	Thr	Tyr	Asp	Ile	Lys	
					245						250				255		
65	aca	aaa	gct	gta	cag	ctt	act	gaa	gaa	gga	atg	acg	aag	gcg	gaa	aaa	1356
	Thr	Lys	Ala	Val	Gln	Leu	Thr	Glu	Glu	Gly	Met	Thr	Lys	Ala	Glu	Lys	
					260					265					270		
70	gca	ttc	ggc	atc	gat	aac	ctc	ttt	gat	gtg	aag	cat	gtc	gcg	ctc	aac	1404
	Ala	Phe	Gly	Ile	Asp	Asn	Leu	Phe	Asp	Val	Lys	His	Val	Ala	Leu	Asn	
				275					280					285			

ES 2 303 058 T3

	cgt cag gga gac ccg ggg att act caa ttt tat ctt tct atg gaa gat	2172
	Arg Gln Gly Asp Pro Gly Ile Thr Gln Phe Tyr Leu Ser Met Glu Asp	
	530	535
5	gaa ttg atg cgc aga ttc gga gct gag cgg aca atg gcg atg ctt gac	2220
	Glu Leu Met Arg Arg Phe Gly Ala Glu Arg Thr Met Ala Met Leu Asp	
	545	550
10	cgc ttc ggc atg gac gac tct act cca atc caa agc aaa atg gta tct	2268
	Arg Phe Gly Met Asp Asp Ser Thr Pro Ile Gln Ser Lys Met Val Ser	
	560	565
15	cgc gcg gtt gaa tcg tct caa aaa cgc gtc gaa ggc aat aac ttc gat	2316
	Arg Ala Val Glu Ser Ser Gln Lys Arg Val Glu Gly Asn Asn Phe Asp	
	580	585
20	tcg cgt aaa cag ctt ctg caa tat gat gat gtt ctc cgc cag cag cgt	2364
	Ser Arg Lys Gln Leu Leu Gln Tyr Asp Asp Val Leu Arg Gln Gln Arg	
	595	600
25	gag gtc att tat aag cag cgc ttt gaa gtc att gac tct gaa aac ctg	2412
	Glu Val Ile Tyr Lys Gln Arg Phe Glu Val Ile Asp Ser Glu Asn Leu	
	610	615
30	cgt gaa atc gtt gaa aat atg atc aag tct tct ctc gaa cgc gca att	2460
	Arg Glu Ile Val Glu Asn Met Ile Lys Ser Ser Leu Glu Arg Ala Ile	
	625	630
35	gca gcc tat acg cca aga gaa gag ctt cct gag gag tgg aag ctt gac	2508
	Ala Ala Tyr Thr Pro Arg Glu Glu Leu Pro Glu Glu Trp Lys Leu Asp	
	640	645
40	ggt cta gtt gat ctt atc aac aca act tat ctt gat gaa ggt gca ctt	2556
	Gly Leu Val Asp Leu Ile Asn Thr Thr Tyr Leu Asp Glu Gly Ala Leu	
	660	665
45	gag aag agc gat atc ttc ggc aaa gaa ccg gat gaa atg ctt gag ctc	2604
	Glu Lys Ser Asp Ile Phe Gly Lys Glu Pro Asp Glu Met Leu Glu Leu	
	675	680
50	att atg gat cgc atc atc aca aaa tat aat gag aag gaa gag caa ttc	2652
	Ile Met Asp Arg Ile Ile Thr Lys Tyr Asn Glu Lys Glu Glu Gln Phe	
	690	695
55	ggc aaa gag caa atg cgc gaa ttc gaa aaa gtt atc gtt ctt cgt gcc	2700
	Gly Lys Glu Gln Met Arg Glu Phe Glu Lys Val Ile Val Leu Arg Ala	
	705	710
60	ggt gat tct aaa tgg atg gat cat att gat gcg atg gat cag ctc cgc	2748
	Val Asp Ser Lys Trp Met Asp His Ile Asp Ala Met Asp Gln Leu Arg	
	720	725
65	caa ggg att cac ctt cgt gct tac gcg cag acg aac ccg ctt cgt gag	2796
	Gln Gly Ile His Leu Arg Ala Tyr Ala Gln Thr Asn Pro Leu Arg Glu	
	740	745
70	tat caa atg gaa ggt ttt gcg atg ttt gag cat atg att gaa tca att	2844
	Tyr Gln Met Glu Gly Phe Ala Met Phe Glu His Met Ile Glu Ser Ile	

ES 2 303 058 T3

	755	760	765	
5	gag gac gaa gtc gca aaa ttt gtg atg aaa gct gag att gaa aac aat Glu Asp Glu Val Ala Lys Phe Val Met Lys Ala Glu Ile Glu Asn Asn 770	gtg atg aaa gct gag att gaa aac aat Val Met Lys Ala Glu Ile Glu Asn Asn 775	gag att gaa aac aat Glu Ile Glu Asn Asn 780	2892
10	ctg gag cgt gaa gag gtt gta caa ggt caa aca aca gct cat cag ccg Leu Glu Arg Glu Glu Val Val Gln Gly Gln Thr Thr Ala His Gln Pro 785	gag gtt gta caa ggt caa aca aca gct cat cag ccg Val Val Gln Gly Gln Thr Thr Ala His Gln Pro 790	aca gct cat cag ccg Thr Ala His Gln Pro 795	2940
15	caa gaa ggc gac gat aac aaa aaa gca aag aaa gca ccg gtt cgc aaa Gln Glu Gly Asp Asp Asn Lys Lys Ala Lys Lys Ala Pro Val Arg Lys 800	aaa aaa gca aag aaa gca ccg gtt cgc aaa Lys Lys Ala Lys Lys Ala Pro Val Arg Lys 805	gca ccg gtt cgc aaa Ala Pro Val Arg Lys 810	2988
20	gtg gtt gat atc gga cga aat gcc cca tgc cac tgc gga agc ggg aaa Val Val Asp Ile Gly Arg Asn Ala Pro Cys His Cys Gly Ser Gly Lys 820	gga cga aat gcc cca tgc cac tgc gga agc ggg aaa Arg Asn Ala Pro Cys His Cys Gly Ser Gly Lys 825	tgc gga agc ggg aaa Cys Gly Ser Gly Lys 830	3036
25	aaa tat aaa aat tgc tgc ggc cgt act gaa tag ttcgccccgg caagtttact Lys Tyr Lys Asn Cys Cys Gly Arg Thr Glu 835	tgc tgc ggc cgt act gaa tag ttcgccccgg caagtttact Cys Cys Gly Arg Thr Glu 840		3089
30	gaacgcggcg cctgcaggct gcgatctttt aatgaggtga atgaaatgaa ttatcagaaa 3149			3149
35	ttagagcaga gctcgaaaat atggcttctc gtttagcgga ctttaggggg tctctttgac 3209			3209
40	ctcgaatcaa aggaggcccg cattgctgag ctagatgaac aaatggctga tccggaattc 3269			3269
45	tggaatgatc agcaaaaagc tcaaacggtt ataaatgaag caaacggttt aaaggattat 3329			3329
50	gtcaattogt ataaaaaatt gaatgaatcc cacgaagaat tacaaatgac tcatgatctt 3389			3389
55	ttgaaagaag agccggacac tgatctccag cttgagcttg aaaaagaact aaagtcatta 3449			3449
60	acaaaagagt tcaatgagtt tgagcttcag cttcttctca gcgagccgta tgataaaaat 3509			3509
65	aacgcgattt tagaactgca ccctgggtgct ggcggtacag agtcacagga ctggggctct 3569			3569
70	atgcttctta gaatgtatac aagatgggga gagcgcgcg gctttaaagt agagactctc 3629			3629
75	gattaccttc cagggtgacga ggcgggaatc aagtcagtga cattgctcat caaaggacac 3689			3689
80	aacgcttacg ggtatctcaa agcagaaaaa ggtgttcac gtcttgtgcg gatctcacca 3749			3749
85	tttgattcat caggccgccc ccacacatct ttcgtttcat gtgaagtcat gcctgaattt 3809			3809
90	aacgatgaaa ttgatattga tattcgtagc gaggatatta aagttgacac gtaccgtgca 3869			3869
95	agcggcgogg gcggacagca cgtcaatagc acgattcag ccgttcggat tactcacttg 3929			3929
100	ccgacgaacg tagttgtgac atgccaaacg gagcgtcac aaattaaana ccgtgaaaga 3989			3989
105	gccatgaaaa tgetgaaggc caagctgtat cagcgcagaa ttgaagagca gcaggcagag 4049			4049
110	ctggatgaaa ttcgcggtga acaaaaagaa atcggctggg gcagccaaat ccgttcttat 4109			4109
115	gtattccatc cgtattccat ggtaaaagac catcgggac			4148

<210> 2

<211> 841

<212> PRT

<213> *Bacillus subtilis*

ES 2 303 058 T3

<400> 2

5 Met Leu Gly Ile Leu Asn Lys Met Phe Asp Pro Thr Lys Arg Thr Leu
1 5 10 15
Asn Arg Tyr Glu Lys Ile Ala Asn Asp Ile Asp Ala Ile Arg Gly Asp
20 25 30
Tyr Glu Asn Leu Ser Asp Asp Ala Leu Lys His Lys Thr Ile Glu Phe
35 40 45
10 Lys Glu Arg Leu Glu Lys Gly Ala Thr Thr Asp Asp Leu Leu Val Glu
50 55 60
Ala Phe Ala Val Val Arg Glu Ala Ser Arg Arg Val Thr Gly Met Phe
65 70 75 80
15 Pro Phe Lys Val Gln Leu Met Gly Gly Val Ala Leu His Asp Gly Asn
85 90 95
Ile Ala Glu Met Lys Thr Gly Glu Gly Lys Thr Leu Thr Ser Thr Leu
100 105 110
Pro Val Tyr Leu Asn Ala Leu Thr Gly Lys Gly Val His Val Val Thr
115 120 125
20 Val Asn Glu Tyr Leu Ala Ser Arg Asp Ala Glu Gln Met Gly Lys Ile
130 135 140
Phe Glu Phe Leu Gly Leu Thr Val Gly Leu Asn Leu Asn Ser Met Ser
145 150 155 160
Lys Asp Glu Lys Arg Glu Ala Tyr Ala Ala Asp Ile Thr Tyr Ser Thr
165 170 175
25 Asn Asn Glu Leu Gly Phe Asp Tyr Leu Arg Asp Asn Met Val Leu Tyr
180 185 190
Lys Glu Gln Met Val Gln Arg Pro Leu His Phe Ala Val Ile Asp Glu
195 200 205
30 Val Asp Ser Ile Leu Ile Asp Glu Ala Arg Thr Pro Leu Ile Ile Ser
210 215 220
Gly Gln Ala Ala Lys Ser Thr Lys Leu Tyr Val Gln Ala Asn Ala Phe
225 230 235 240
Val Arg Thr Leu Lys Ala Glu Lys Asp Tyr Thr Tyr Asp Ile Lys Thr
245 250 255
35 Lys Ala Val Gln Leu Thr Glu Glu Gly Met Thr Lys Ala Glu Lys Ala
260 265 270
Phe Gly Ile Asp Asn Leu Phe Asp Val Lys His Val Ala Leu Asn His
275 280 285
40 His Ile Asn Gln Ala Leu Lys Ala His Val Ala Met Gln Lys Asp Val
290 295 300
Asp Tyr Val Val Glu Asp Gly Gln Val Val Ile Val Asp Ser Phe Thr
305 310 315 320
Gly Arg Leu Met Lys Gly Arg Arg Tyr Ser Glu Gly Leu His Gln Ala
325 330 335
45 Ile Glu Ala Lys Glu Gly Leu Glu Ile Gln Asn Glu Ser Met Thr Leu
340 345 350
Ala Thr Ile Thr Phe Gln Asn Tyr Phe Arg Met Tyr Glu Lys Leu Ala
355 360 365
Gly Met Thr Gly Thr Ala Lys Thr Glu Glu Glu Glu Phe Arg Asn Ile
370 375 380
50 Tyr Asn Met Gln Val Val Thr Ile Pro Thr Asn Arg Pro Val Val Arg
385 390 395 400
Asp Asp Arg Pro Asp Leu Ile Tyr Arg Thr Met Glu Gly Lys Phe Lys
405 410 415
55 Ala Val Ala Glu Asp Val Ala Gln Arg Tyr Met Thr Gly Gln Pro Val

60

65

ES 2 303 058 T3

			420				425			430						
	Leu	Val	Gly	Thr	Val	Ala	Val	Glu	Thr	Ser	Glu	Leu	Ile	Ser	Lys	Leu
			435					440					445			
5	Leu	Lys	Asn	Lys	Gly	Ile	Pro	His	Gln	Val	Leu	Asn	Ala	Lys	Asn	His
		450				455						460				
	Glu	Arg	Glu	Ala	Gln	Ile	Ile	Glu	Glu	Ala	Gly	Gln	Lys	Gly	Ala	Val
	465					470					475					480
	Thr	Ile	Ala	Thr	Asn	Met	Ala	Gly	Arg	Gly	Thr	Asp	Ile	Lys	Leu	Gly
					485						490					495
10	Glu	Gly	Val	Lys	Glu	Leu	Gly	Gly	Leu	Ala	Val	Val	Gly	Thr	Glu	Arg
				500					505					510		
	His	Glu	Ser	Arg	Arg	Ile	Asp	Asn	Gln	Leu	Arg	Gly	Arg	Ser	Gly	Arg
			515					520					525			
15	Gln	Gly	Asp	Pro	Gly	Ile	Thr	Gln	Phe	Tyr	Leu	Ser	Met	Glu	Asp	Glu
		530				535						540				
	Leu	Met	Arg	Arg	Phe	Gly	Ala	Glu	Arg	Thr	Met	Ala	Met	Leu	Asp	Arg
	545				550						555					560
	Phe	Gly	Met	Asp	Asp	Ser	Thr	Pro	Ile	Gln	Ser	Lys	Met	Val	Ser	Arg
				565						570						575
20	Ala	Val	Glu	Ser	Gln	Lys	Arg	Val	Glu	Gly	Asn	Asn	Phe	Asp	Ser	
				580					585					590		
	Arg	Lys	Gln	Leu	Leu	Gln	Tyr	Asp	Asp	Val	Leu	Arg	Gln	Gln	Arg	Glu
			595					600					605			
25	Val	Ile	Tyr	Lys	Gln	Arg	Phe	Glu	Val	Ile	Asp	Ser	Glu	Asn	Leu	Arg
		610					615					620				
	Glu	Ile	Val	Glu	Asn	Met	Ile	Lys	Ser	Ser	Leu	Glu	Arg	Ala	Ile	Ala
	625				630						635					640
	Ala	Tyr	Thr	Pro	Arg	Glu	Glu	Leu	Pro	Glu	Glu	Trp	Lys	Leu	Asp	Gly
				645						650					655	
30	Leu	Val	Asp	Leu	Ile	Asn	Thr	Thr	Tyr	Leu	Asp	Glu	Gly	Ala	Leu	Glu
			660						665					670		
	Lys	Ser	Asp	Ile	Phe	Gly	Lys	Glu	Pro	Asp	Glu	Met	Leu	Glu	Leu	Ile
			675					680					685			
35	Met	Asp	Arg	Ile	Ile	Thr	Lys	Tyr	Asn	Glu	Lys	Glu	Glu	Gln	Phe	Gly
		690				695						700				
	Lys	Glu	Gln	Met	Arg	Glu	Phe	Glu	Lys	Val	Ile	Val	Leu	Arg	Ala	Val
	705					710					715					720
	Asp	Ser	Lys	Trp	Met	Asp	His	Ile	Asp	Ala	Met	Asp	Gln	Leu	Arg	Gln
				725						730						735
40	Gly	Ile	His	Leu	Arg	Ala	Tyr	Ala	Gln	Thr	Asn	Pro	Leu	Arg	Glu	Tyr
				740					745					750		
	Gln	Met	Glu	Gly	Phe	Ala	Met	Phe	Glu	His	Met	Ile	Glu	Ser	Ile	Glu
			755					760					765			
45	Asp	Glu	Val	Ala	Lys	Phe	Val	Met	Lys	Ala	Glu	Ile	Glu	Asn	Asn	Leu
		770				775							780			
	Glu	Arg	Glu	Glu	Val	Val	Gln	Gly	Gln	Thr	Thr	Ala	His	Gln	Pro	Gln
	785					790						795				800
	Glu	Gly	Asp	Asp	Asn	Lys	Lys	Ala	Lys	Lys	Ala	Pro	Val	Arg	Lys	Val
				805						810					815	
50	Val	Asp	Ile	Gly	Arg	Asn	Ala	Pro	Cys	His	Cys	Gly	Ser	Gly	Lys	Lys
				820					825					830		
	Tyr	Lys	Asn	Cys	Cys	Gly	Arg	Thr	Glu							
			835					840								

55 <210> 3
 <211> 2706
 <212> DNA
 <213> *Escherichia coli*
 60 <220>
 <221> CDS
 <222> (1)..(2706)

65

ES 2 303 058 T3

<400> 3

5 atg cta atc aaa tta tta act aaa gtt ttc ggt agt cgt aac gat cgc 48
Met Leu Ile Lys Leu Leu Thr Lys Val Phe Gly Ser Arg Asn Asp Arg
1 5 10 15

10 acc ctg cgc cgg atg cgc aaa gtg gtc aac atc atc aat gcc atg gaa 96
Thr Leu Arg Arg Met Arg Lys Val Val Asn Ile Ile Asn Ala Met Glu
20 25 30

15 ccg gag atg gaa aaa ctc tcc gac gaa gaa ctg aaa ggg aaa acc gca 144
Pro Glu Met Glu Lys Leu Ser Asp Glu Glu Leu Lys Gly Lys Thr Ala
35 40 45

20 gag ttt cgt gcg cgt ctg gaa aaa ggc gaa gtg ctg gaa aat ctg atc 192
Glu Phe Arg Ala Arg Leu Glu Lys Gly Glu Val Leu Glu Asn Leu Ile
50 55 60

25 ccg gaa gct ttc gcc gtg gtg cgt gag gca agt aag cgc gtc ttt ggt 240
Pro Glu Ala Phe Ala Val Val Arg Glu Ala Ser Lys Arg Val Phe Gly
65 70 75 80

30 atg cgt cac ttc gac gtt cag tta ctc ggc ggt atg gtt ctt aac gaa 288
Met Arg His Phe Asp Val Gln Leu Leu Gly Gly Met Val Leu Asn Glu
85 90 95

35 cgc tgc atc gcc gaa atg cgt acc ggt gaa ggt aaa acc ctg acc gca 336
Arg Cys Ile Ala Glu Met Arg Thr Gly Glu Gly Lys Thr Leu Thr Ala
100 105 110

40 acg ctg cct gct tac ctg aac gca cta acc ggt aaa ggc gta cac gta 384
Thr Leu Pro Ala Tyr Leu Asn Ala Leu Thr Gly Lys Gly Val His Val
115 120 125

45 gtt acc gtc aac gac tac ctg gcg caa cgt gac gcc gaa aac aac cgt 432
Val Thr Val Asn Asp Tyr Leu Ala Gln Arg Asp Ala Glu Asn Asn Arg
130 135 140

50 ccg ctg ttt gaa ttc ctt ggc ctg act gtc ggt atc aac ctg ccg ggc 480
Pro Leu Phe Glu Phe Leu Gly Leu Thr Val Gly Ile Asn Leu Pro Gly
145 150 155 160

55 atg cca gca ccg gca aag cgt gaa gcc tac gct gct gac atc act tac 528
Met Pro Ala Pro Ala Lys Arg Glu Ala Tyr Ala Ala Asp Ile Thr Tyr
165 170 175

60 ggt acg aac aac gaa tac ggc ttt gac tac ctg cgc gac aac atg gca 576
Gly Thr Asn Asn Glu Tyr Gly Phe Asp Tyr Leu Arg Asp Asn Met Ala
180 185 190

65 ttc agc cct gaa gaa cgt gta caa cgt aaa ctg cac tat gcg ctg gtg 624
Phe Ser Pro Glu Glu Arg Val Gln Arg Lys Leu His Tyr Ala Leu Val
195 200 205

ES 2 303 058 T3

	ggc cag ccg gtg ctg gtg ggt aca atc tcc atc gaa aaa tcg gag ctg	1392
	Gly Gln Pro Val Leu Val Gly Thr Ile Ser Ile Glu Lys Ser Glu Leu	
	450 455 460	
5	gtg tca aat gaa ctg acc aaa gcc ggt att aag cac aac gtc ctg aac	1440
	Val Ser Asn Glu Leu Thr Lys Ala Gly Ile Lys His Asn Val Leu Asn	
	465 470 475 480	
10	gcc aaa ttc cat gcc aac gaa gcg gcg att gtt gct cag gca ggt tat	1488
	Ala Lys Phe His Ala Asn Glu Ala Ala Ile Val Ala Gln Ala Gly Tyr	
	485 490 495	
15	ccg gct gcg gtg act atc gcg acc aac atg gcg ggt cgt ggt acc gat	1536
	Pro Ala Ala Val Thr Ile Ala Thr Asn Met Ala Gly Arg Gly Thr Asp	
	500 505 510	
	att gtg ctc ggt ggt agc tgg cag gca gaa gtt gcc gcg ctg gaa aat	1584
	Ile Val Leu Gly Gly Ser Trp Gln Ala Glu Val Ala Ala Leu Glu Asn	
	515 520 525	
20	ccg act gca gag caa att gaa aaa att aaa gcc gac tgg cag gta cgt	1632
	Pro Thr Ala Glu Gln Ile Glu Lys Ile Lys Ala Asp Trp Gln Val Arg	
	530 535 540	
25	cac gat gcg gta ctg gca gca ggt gcc ctg cat atc atc ggt act gaa	1680
	His Asp Ala Val Leu Ala Ala Gly Gly Leu His Ile Ile Gly Thr Glu	
	545 550 555 560	
30	cgt cac gaa tcc cgt cgt atc gat aac cag ctg cgc ggt cgt tct ggt	1728
	Arg His Glu Ser Arg Arg Ile Asp Asn Gln Leu Arg Gly Arg Ser Gly	
	565 570 575	
35	cgt cag ggg gat gct ggt tct tct cgt ttc tac ctg tcg atg gaa gat	1776
	Arg Gln Gly Asp Ala Gly Ser Ser Arg Phe Tyr Leu Ser Met Glu Asp	
	580 585 590	
40	gcg ctg atg cgt att ttt gct tcc gac cga gta tcc gcc atg atg cgt	1824
	Ala Leu Met Arg Ile Phe Ala Ser Asp Arg Val Ser Gly Met Met Arg	
	595 600 605	
45	aaa ctg ggt atg aag cca gcc gaa gcc att gag cac ccg tgg gtg acc	1872
	Lys Leu Gly Met Lys Pro Gly Glu Ala Ile Glu His Pro Trp Val Thr	
	610 615 620	
50	aaa gcg att gcc aac gcc cag cgt aaa gtt gaa agc cgt aac ttc gac	1920
	Lys Ala Ile Ala Asn Ala Gln Arg Lys Val Glu Ser Arg Asn Phe Asp	
	625 630 635 640	
55	att cgt aag caa ctg ctg gaa tat gat gac gtg gct aac gat cag cgt	1968
	Ile Arg Lys Gln Leu Leu Glu Tyr Asp Asp Val Ala Asn Asp Gln Arg	
	645 650 655	
60	cgc gcc att tac tcc cag cgt aac gaa ctg ctg gat gtc agc gat gtg	2016
	Arg Ala Ile Tyr Ser Gln Arg Asn Glu Leu Leu Asp Val Ser Asp Val	
	660 665 670	
65	agc gaa acc atc aac agc att cgt gaa gat gtg ttc aaa gcg acc att	2064
	Ser Glu Thr Ile Asn Ser Ile Arg Glu Asp Val Phe Lys Ala Thr Ile	

ES 2 303 058 T3

	675	680	685	
5	gat gcc tac att ccg cca cag tcg ctg gaa gaa atg tgg gat att ccg Asp Ala Tyr Ile Pro Pro Gln Ser Leu Glu Glu Met Trp Asp Ile Pro 690 695 700			2112
10	ggg ctg cag gaa cgt ctg aag aac gat ttc gac ctc gat ttg cca att Gly Leu Gln Glu Arg Leu Lys Asn Asp Phe Asp Leu Asp Leu Pro Ile 705 710 715 720			2160
15	gcc gag tgg ctg gat aaa gaa cca gaa ctg cat gaa gag acg ctg cgt Ala Glu Trp Leu Asp Lys Glu Pro Glu Leu His Glu Glu Thr Leu Arg 725 730 735			2208
20	gag cgc att ctg gcg cag tcc atc gaa gtg tat cag cgt aaa gaa gaa Glu Arg Ile Leu Ala Gln Ser Ile Glu Val Tyr Gln Arg Lys Glu Glu 740 745 750			2256
25	gtg gtt ggt gct gag atg atg cgt cac ttc gag aaa ggc gtc atg ctg Val Val Gly Ala Glu Met Met Arg His Phe Glu Lys Gly Val Met Leu 755 760 765			2304
30	caa act ctc gac tct ctg tgg aaa gag cac ctg gca gcg atg gac tat Gln Thr Leu Asp Ser Leu Trp Lys Glu His Leu Ala Ala Met Asp Tyr 770 775 780			2352
35	ctg cgt cag ggt atc cac ctg cgt ggc tat gca cag aaa gat ccg aag Leu Arg Gln Gly Ile His Leu Arg Gly Tyr Ala Gln Lys Asp Pro Lys 785 790 795 800			2400
40	cag gaa tac aaa cgt gaa tcg ttc tcc atg ttt gca gcg atg ctg gag Gln Glu Tyr Lys Arg Glu Ser Phe Ser Met Phe Ala Ala Met Leu Glu 805 810 815			2448
45	tcg ttg aaa tat gaa gtt atc agt acg ctg agc aaa gtt cag gta cgt Ser Leu Lys Tyr Glu Val Ile Ser Thr Leu Ser Lys Val Gln Val Arg 820 825 830			2496
50	atg cct gaa gag gtt gag gag ctg gaa caa cag cgt cgt atg gaa gcc Met Pro Glu Glu Val Glu Glu Leu Glu Gln Gln Arg Arg Met Glu Ala 835 840 845			2544
55	gag cgt tta gcg caa atg cag cag ctt agc cat cag gat gac gac tct Glu Arg Leu Ala Gln Met Gln Gln Leu Ser His Gln Asp Asp Asp Ser 850 855 860			2592
60	gca gcc gca gct gca ctg gcg gcg caa acc ggt gaa cgc aaa gta gga Ala Ala Ala Ala Ala Leu Ala Ala Gln Thr Gly Glu Arg Lys Val Gly 865 870 875 880			2640
65	cgt aac gat cct tgc ccg tgt ggt tct ggt aaa aaa tac aag cag tgc Arg Asn Asp Pro Cys Pro Cys Gly Ser Gly Lys Lys Tyr Lys Gln Cys 885 890 895			2688
70	cat ggc cgc ctg caa taa His Gly Arg Leu Gln 900			2706

<210> 4

<211> 901

<212> PRT

<213> *Escherichia coli*

ES 2 303 058 T3

<400> 4

5 Met Leu Ile Lys Leu Leu Thr Lys Val Phe Gly Ser Arg Asn Asp Arg
1 5 10
Thr Leu Arg Arg Met Arg Lys Val Val Asn Ile Ile Asn Ala Met Glu
20 25 30
Pro Glu Met Glu Lys Leu Ser Asp Glu Glu Leu Lys Gly Lys Thr Ala
35 40 45
10 Glu Phe Arg Ala Arg Leu Glu Lys Gly Glu Val Leu Glu Asn Leu Ile
50 55 60
Pro Glu Ala Phe Ala Val Arg Glu Ala Ser Lys Arg Val Phe Gly
65 70 75 80
15 Met Arg His Phe Asp Val Gln Leu Leu Gly Gly Met Val Leu Asn Glu
85 90 95
Arg Cys Ile Ala Glu Met Arg Thr Gly Glu Gly Lys Thr Leu Thr Ala
100 105 110
Thr Leu Pro Ala Tyr Leu Asn Ala Leu Thr Gly Lys Gly Val His Val
115 120 125
20 Val Thr Val Asn Asp Tyr Leu Ala Gln Arg Asp Ala Glu Asn Asn Arg
130 135 140
Pro Leu Phe Glu Phe Leu Gly Leu Thr Val Gly Ile Asn Leu Pro Gly
145 150 155 160
Met Pro Ala Pro Ala Lys Arg Glu Ala Tyr Ala Ala Asp Ile Thr Tyr
165 170 175
25 Gly Thr Asn Asn Glu Tyr Gly Phe Asp Tyr Leu Arg Asp Asn Met Ala
180 185 190
Phe Ser Pro Glu Glu Arg Val Gln Arg Lys Leu His Tyr Ala Leu Val
195 200 205
30 Asp Glu Val Asp Ser Ile Leu Ile Asp Glu Ala Arg Thr Pro Leu Ile
210 215 220
Ile Ser Gly Pro Ala Glu Asp Ser Ser Glu Met Tyr Lys Arg Val Asn
225 230 235 240
Lys Ile Ile Pro His Leu Ile Arg Gln Glu Lys Glu Asp Ser Glu Thr
245 250 255
35 Phe Gln Gly Glu Gly His Phe Ser Val Asp Glu Lys Ser Arg Gln Val
260 265 270
Asn Leu Thr Glu Arg Gly Leu Val Leu Ile Glu Glu Leu Leu Val Lys
275 280 285
40 Glu Gly Ile Met Asp Glu Gly Glu Ser Leu Tyr Ser Pro Ala Asn Ile
290 295 300
Met Leu Met His His Val Thr Ala Ala Leu Arg Ala His Ala Leu Phe
305 310 315 320
Thr Arg Asp Val Asp Tyr Ile Val Lys Asp Gly Glu Val Ile Ile Val
325 330 335
45 Asp Glu His Thr Gly Arg Thr Met Gln Gly Arg Arg Trp Ser Asp Gly
340 345 350
Leu His Gln Ala Val Glu Ala Lys Glu Gly Val Gln Ile Gln Asn Glu
355 360 365
50 Asn Gln Thr Leu Ala Ser Ile Thr Phe Gln Asn Tyr Phe Arg Leu Tyr
370 375 380
Glu Lys Leu Ala Gly Met Thr Gly Thr Ala Asp Thr Glu Ala Phe Glu
385 390 395 400
Phe Ser Ser Ile Tyr Lys Leu Asp Thr Val Val Val Pro Thr Asn Arg
405 410 415
55 Pro Met Ile Arg Lys Asp Leu Pro Asp Leu Val Tyr Met Thr Glu Ala

60

65

ES 2 303 058 T3

420 425 430
 Glu Lys Ile Gln Ala Ile Ile Glu Asp Ile Lys Glu Arg Thr Ala Lys
 435 440 445
 Gly Gln Pro Val Leu Val Gly Thr Ile Ser Ile Glu Lys Ser Glu Leu
 450 455 460
 Val Ser Asn Glu Leu Thr Lys Ala Gly Ile Lys His Asn Val Leu Asn
 465 470 475 480
 Ala Lys Phe His Ala Asn Glu Ala Ala Ile Val Ala Gln Ala Gly Tyr
 485 490 495
 Pro Ala Ala Val Thr Ile Ala Thr Asn Met Ala Gly Arg Gly Thr Asp
 500 505 510
 Ile Val Leu Gly Gly Ser Trp Gln Ala Glu Val Ala Ala Leu Glu Asn
 515 520 525
 Pro Thr Ala Glu Gln Ile Glu Lys Ile Lys Ala Asp Trp Gln Val Arg
 530 535 540
 His Asp Ala Val Leu Ala Ala Gly Gly Leu His Ile Ile Gly Thr Glu
 545 550 555 560
 Arg His Glu Ser Arg Arg Ile Asp Asn Gln Leu Arg Gly Arg Ser Gly
 565 570 575
 Arg Gln Gly Asp Ala Gly Ser Ser Arg Phe Tyr Leu Ser Met Glu Asp
 580 585 590
 Ala Leu Met Arg Ile Phe Ala Ser Asp Arg Val Ser Gly Met Met Arg
 595 600 605
 Lys Leu Gly Met Lys Pro Gly Glu Ala Ile Glu His Pro Trp Val Thr
 610 615 620
 Lys Ala Ile Ala Asn Ala Gln Arg Lys Val Glu Ser Arg Asn Phe Asp
 625 630 635 640
 Ile Arg Lys Gln Leu Leu Glu Tyr Asp Asp Val Ala Asn Asp Gln Arg
 645 650 655
 Arg Ala Ile Tyr Ser Gln Arg Asn Glu Leu Leu Asp Val Ser Asp Val
 660 665 670
 Ser Glu Thr Ile Asn Ser Ile Arg Glu Asp Val Phe Lys Ala Thr Ile
 675 680 685
 Asp Ala Tyr Ile Pro Pro Gln Ser Leu Glu Glu Met Trp Asp Ile Pro
 690 695 700
 Gly Leu Gln Glu Arg Leu Lys Asn Asp Phe Asp Leu Asp Leu Pro Ile
 705 710 715 720
 Ala Glu Trp Leu Asp Lys Glu Pro Glu Leu His Glu Glu Thr Leu Arg
 725 730 735
 Glu Arg Ile Leu Ala Gln Ser Ile Glu Val Tyr Gln Arg Lys Glu Glu
 740 745 750
 Val Val Gly Ala Glu Met Met Arg His Phe Glu Lys Gly Val Met Leu
 755 760 765
 Gln Thr Leu Asp Ser Leu Trp Lys Glu His Leu Ala Ala Met Asp Tyr
 770 775 780
 Leu Arg Gln Gly Ile His Leu Arg Gly Tyr Ala Gln Lys Asp Pro Lys
 785 790 795 800
 Gln Glu Tyr Lys Arg Glu Ser Phe Ser Met Phe Ala Ala Met Leu Glu
 805 810 815
 Ser Leu Lys Tyr Glu Val Ile Ser Thr Leu Ser Lys Val Gln Val Arg
 820 825 830
 Met Pro Glu Glu Val Glu Glu Leu Glu Gln Gln Arg Arg Met Glu Ala
 835 840 845
 Glu Arg Leu Ala Gln Met Gln Gln Leu Ser His Gln Asp Asp Asp Ser
 850 855 860
 Ala Ala Ala Ala Ala Leu Ala Ala Gln Thr Gly Glu Arg Lys Val Gly
 865 870 875 880
 Arg Asn Asp Pro Cys Pro Cys Gly Ser Gly Lys Lys Tyr Lys Gln Cys
 885 890 895
 His Gly Arg Leu Gln
 900

60

<210> 5
 <211> 2706
 <212> DNA
 <213> *Bacillus licheniformis*

ES 2 303 058 T3

<220>

<221> CDS

<222> (154)..(2679)

5

<400> 5

```

gatccccctc ccggatcttc cgcagagggg attttttccg ttcccccgcg gtaaattggt 60
10
tggaatgac aaaaggtatg atatgatatt gcatatataa aaattactgt ttactcatgc 120
ttaaacaagg aaattaaaga ggagcgttat tct atg ctt gga att tta aat aaa 174
Met Leu Gly Ile Leu Asn Lys
15
1
5
gtg ttt gat ccg aca aaa cgc acg ctc agc cgt tat gaa aag aaa gcg 222
Val Phe Asp Pro Thr Lys Arg Thr Leu Ser Arg Tyr Glu Lys Lys Ala
10 15 20
20
aac gag att gat gcg ctc aag gca gat ata gag aag ctt tca gac gaa 270
Asn Glu Ile Asp Ala Leu Lys Ala Asp Ile Glu Lys Leu Ser Asp Glu
25 30 35
25
gct ttg aag caa aag acg atc gag ttc aaa gag cgc ctt gaa aaa ggc 318
Ala Leu Lys Gln Lys Thr Ile Glu Phe Lys Glu Arg Leu Glu Lys Gly
40 45 50 55
30
gaa acg gtt gac gat ctt ttg gtt gaa gcg ttt gcc gtt gtc agg gaa 366
Glu Thr Val Asp Asp Leu Leu Val Glu Ala Phe Ala Val Val Arg Glu
60 65 70
35
gct tcc cgg cgc gtg aca ggc atg ttt ccg ttt aag gtt cag ctg atg 414
Ala Ser Arg Arg Val Thr Gly Met Phe Pro Phe Lys Val Gln Leu Met
75 80 85
40
ggg ggc gtc gcc ctt cat gaa ggg aat atc gcc gaa atg aaa acg ggg 462
Gly Gly Val Ala Leu His Glu Gly Asn Ile Ala Glu Met Lys Thr Gly
90 95 100
45
gaa ggt aaa acg ctg act tcc aca atg ccc gtt tac ttg aac gct ctg 510
Glu Gly Lys Thr Leu Thr Ser Thr Met Pro Val Tyr Leu Asn Ala Leu
105 110 115
50
tca ggg aaa ggc gtt cac gtc gtg acg gtc aac gaa tac ctg gcg agc 558
Ser Gly Lys Gly Val His Val Val Thr Val Asn Glu Tyr Leu Ala Ser
120 125 130 135
55
cgc gac gct gaa gag atg ggg aaa atc ttt gag ttt ctc ggg ctg acg 606
Arg Asp Ala Glu Glu Met Gly Lys Ile Phe Glu Phe Leu Gly Leu Thr
140 145 150

```

55

60

65

ES 2 303 058 T3

	gtc ggc cta aac ctg aac agc ctg tca aaa gac gag aag cgt gaa gcc	654
	Val Gly Leu Asn Leu Ser Leu Ser Lys Asp Glu Lys Arg Glu Ala	
	155 160 165	
5	tat gca gca gat att acg tat tct acg aat aat gag ctt ggc ttt gac	702
	Tyr Ala Ala Asp Ile Thr Tyr Ser Thr Asn Asn Glu Leu Gly Phe Asp	
	170 175 180	
10	tac ttg cgc gac aac atg gtg ctt tat aaa gag cag atg gtt cag cgc	750
	Tyr Leu Arg Asp Asn Met Val Leu Tyr Lys Glu Gln Met Val Gln Arg	
	185 190 195	
15	ccg ctt cat ttt gcg gtc atc gat gaa gtc gac tcc att ttg atc gat	798
	Pro Leu His Phe Ala Val Ile Asp Glu Val Asp Ser Ile Leu Ile Asp	
	200 205 210 215	
20	gaa gca aga acg ccg ctc atc att tct gga caa gcg gcc aaa tcc acc	846
	Glu Ala Arg Thr Pro Leu Ile Ile Ser Gly Gln Ala Ala Lys Ser Thr	
	220 225 230	
25	aag ctt tat gtt cag gcc aat gcg ttt gtc cgc acg cta aaa gcg gat	894
	Lys Leu Tyr Val Gln Ala Asn Ala Phe Val Arg Thr Leu Lys Ala Asp	
	235 240 245	
30	cag gac tac aca tac gat gtg aaa aca aaa ggc gtt cag ctg act gaa	942
	Gln Asp Tyr Thr Tyr Asp Val Lys Thr Lys Gly Val Gln Leu Thr Glu	
	250 255 260	
35	gag ggg atg aca aaa gct gaa aag gca ttt ggc atc gaa aac ttg ttt	990
	Glu Gly Met Thr Lys Ala Glu Lys Ala Phe Gly Ile Glu Asn Leu Phe	
	265 270 275	
40	gac gtc cgc cat gtc gcc tta aac cat cat att gcc cag gcg ctg aaa	1038
	Asp Val Arg His Val Ala Leu Asn His His Ile Ala Gln Ala Leu Lys	
	280 285 290 295	
45	gcc cat gcg gcg atg cat aaa gac gtc gac tac gtc gtc gaa gac ggt	1086
	Ala His Ala Ala Met His Lys Asp Val Asp Tyr Val Val Glu Asp Gly	
	300 305 310	
50	cag gtc gtt atc gtc gac tct ttt aca ggc cgt ttg atg aaa ggc cgc	1134
	Gln Val Val Ile Val Asp Ser Phe Thr Gly Arg Leu Met Lys Gly Arg	
	315 320 325	
55	cgc tac agc gac gga ctt cac cag gcc att gaa gcg aag gaa ggc ctt	1182
	Arg Tyr Ser Asp Gly Leu His Gln Ala Ile Glu Ala Lys Glu Gly Leu	
	330 335 340	
60	gag atc caa aat gag agc atg acg ctc gcg acg atc acc ttc cag aac	1230
	Glu Ile Gln Asn Glu Ser Met Thr Leu Ala Thr Ile Thr Phe Gln Asn	
	345 350 355	
65	tat ttc cga atg tat gaa aaa ttg gct gga atg acg ggt acc gca aaa	1278
	Tyr Phe Arg Met Tyr Glu Lys Leu Ala Gly Met Thr Gly Thr Ala Lys	
	360 365 370 375	
70	acg gaa gaa gaa gaa ttc cgc aac atc tac aac atg cag gtt gtt acg	1326
	Thr Glu Glu Glu Glu Phe Arg Asn Ile Tyr Asn Met Gln Val Val Thr	
	380 385 390	

ES 2 303 058 T3

5	<p>att ccg acc aac aag ccg att gcc cgc gat gac cga ccg gat tta att 1374 Ile Pro Thr Asn Lys Pro Ile Ala Arg Asp Asp Arg Pro Asp Leu Ile 395 400 405</p>	
10	<p>tac cgg acc atg gaa gga aaa ttt aaa gct gtt gca gag gat gtc gcc 1422 Tyr Arg Thr Met Glu Gly Lys Phe Lys Ala Val Ala Glu Asp Val Ala 410 415 420</p>	
15	<p>cag cgc tat atg gtc gga cag ccg gta ctt gtc ggt acg gtt gcg gtt 1470 Gln Arg Tyr Met Val Gly Gln Pro Val Leu Val Gly Thr Val Ala Val 425 430 435</p>	
20	<p>gaa aca tct gaa ttg ata tca agg ctc ctt aaa aat aaa gga atc ccg 1518 Glu Thr Ser Glu Leu Ile Ser Arg Leu Leu Lys Asn Lys Gly Ile Pro 440 445 450 455</p>	
25	<p>cat caa gtg ttg aac gcg aaa aac cat gag cgg gaa gct cag att atc 1566 His Gln Val Leu Asn Ala Lys Asn His Glu Arg Glu Ala Gln Ile Ile 460 465 470</p>	
30	<p>gaa gat gcc ggg caa aaa ggc gcg gtc acc atc gcg acc aac atg gcg 1614 Glu Asp Ala Gly Gln Lys Gly Ala Val Thr Ile Ala Thr Asn Met Ala 475 480 485</p>	
35	<p>ggc cgc gga acg gac atc aag ctt ggc gaa ggt gta aaa gag ctt ggc 1662 Gly Arg Gly Thr Asp Ile Lys Leu Gly Glu Gly Val Lys Glu Leu Gly 490 495 500</p>	
40	<p>gga ctg gcc gtc atc ggt acg gaa cgc cat gaa tca agg cgg att gac 1710 Gly Leu Ala Val Ile Gly Thr Glu Arg His Glu Ser Arg Arg Ile Asp 505 510 515</p>	
45	<p>aac cag ctg cgc gga cgt tca ggc cgt cag ggg gac cct ggt atc acc 1758 Asn Gln Leu Arg Gly Arg Ser Gly Arg Gln Gly Asp Pro Gly Ile Thr 520 525 530 535</p>	
50	<p>caa ttt tat ctg tcc atg gaa gat gaa tta atg aaa cgc ttc ggc gca 1806 Gln Phe Tyr Leu Ser Met Glu Asp Glu Leu Met Lys Arg Phe Gly Ala 540 545 550</p>	
55	<p>gag cgg acg atg gcg atg ctt gac cgc ttc gga atg gac gat tcg acg 1854 Glu Arg Thr Met Ala Met Leu Asp Arg Phe Gly Met Asp Asp Ser Thr 555 560 565</p>	
60	<p>ccg ata cag agc aag atg gtt tca aga gcg gtc gaa tct tca cag aag 1902 Pro Ile Gln Ser Lys Met Val Ser Arg Ala Val Glu Ser Ser Gln Lys 570 575 580</p>	
65	<p>cgt gtg gaa ggc aac aac ttt gat gcc cgt aag cag ctt ctg caa tac 1950 Arg Val Glu Gly Asn Asn Phe Asp Ala Arg Lys Gln Leu Leu Gln Tyr 585 590 595</p>	
70	<p>gat gac gtg ctc cgc cag cag cgc gaa gtc atc tat aaa cag cgc ttt 1998 Asp Asp Val Leu Arg Gln Gln Arg Glu Val Ile Tyr Lys Gln Arg Phe 600 605 610 615</p>	
75	<p>gag gtc atc gat tcc gat aac ctc cgc tcc atc gtc gaa aat atg att 2046 Glu Val Ile Asp Ser Asp Asn Leu Arg Ser Ile Val Glu Asn Met Ile</p>	

ES 2 303 058 T3

	620	625	630	
5	aaa gct tca ctc gag cgg gct gtt gct tca tat acg ccg aag gaa gat Lys Ala Ser Leu Glu Arg Ala Val Ala Ser Tyr Thr Pro Lys Glu Asp 635	640	645	2094
10	ctg cct gaa gag tgg aat ctt gac ggc ctt gtg gag ctt gta aat gcg Leu Pro Glu Glu Trp Asn Leu Asp Gly Leu Val Glu Leu Val Asn Ala 650	655	660	2142
15	aat ttc ctt gat gaa ggt gga gtg gag aaa agc gac att ttc gga aaa Asn Phe Leu Asp Glu Gly Gly Val Glu Lys Ser Asp Ile Phe Gly Lys 665	670	675	2190
20	gag ccc gag gag att aca gag ctc att tac gac cgc atc aaa acg aaa Glu Pro Glu Glu Ile Thr Glu Leu Ile Tyr Asp Arg Ile Lys Thr Lys 680	685	690	2238
25	tac gat gag aaa gaa gag cgg tac ggc tct gaa caa atg cgc gaa ttt Tyr Asp Glu Lys Glu Glu Arg Tyr Gly Ser Glu Gln Met Arg Glu Phe 700	705	710	2286
30	gag aaa gtc atc gtt ctc cgc gaa gtg gat acg aaa tgg atg gat cac Glu Lys Val Ile Val Leu Arg Glu Val Asp Thr Lys Trp Met Asp His 715	720	725	2334
35	atc gat gcg atg gat cag ctg cgg caa gga att cat ctg cgc gct tat Ile Asp Ala Met Asp Gln Leu Arg Gln Gly Ile His Leu Arg Ala Tyr 730	735	740	2382
40	gct cag aca aac ccg ctc cgc gag tat cag atg gaa ggc ttt gca atg Ala Gln Thr Asn Pro Leu Arg Glu Tyr Gln Met Glu Gly Phe Ala Met 745	750	755	2430
45	ttt gaa aac atg atc gcg gcg att gaa gat gat gta gcc aaa ttc gtc Phe Glu Asn Met Ile Ala Ala Ile Glu Asp Asp Val Ala Lys Phe Val 760	765	770	2478
50	atg aag gct gaa atc gaa aac aac ctt gag cgc gaa gag gtc att caa Met Lys Ala Glu Ile Glu Asn Asn Leu Glu Arg Glu Glu Val Ile Gln 780	785	790	2526
55	gga cag acg aca gcc cat cag ccg aaa gaa ggc gat gag gaa aaa caa Gly Gln Thr Thr Ala His Gln Pro Lys Glu Gly Asp Glu Glu Lys Gln 795	800	805	2574
60	gcg aag aaa aaa ccg gtc cgc aaa gcg gtg gat atc gga cgc aat gat Ala Lys Lys Lys Pro Val Arg Lys Ala Val Asp Ile Gly Arg Asn Asp 810	815	820	2622
65	cct tgc tac tgc gga agc gga aaa aaa tat aaa aac tgc tgc gga aga Pro Cys Tyr Cys Gly Ser Gly Lys Lys Tyr Lys Asn Cys Cys Gly Arg 825	830	835	2670
	aca gaa taa aaagaggtgc acgcctcttt ttatttg Thr Glu 840			2706

<210> 6

<211> 841

<212> PRT

<213> *Bacillus licheniformis*

ES 2 303 058 T3

<400> 6

5 Met Leu Gly Ile Leu Asn Lys Val Phe Asp Pro Thr Lys Arg Thr Leu
1 5 10
Ser Arg Tyr Glu Lys Lys Ala Asn Glu Ile Asp Ala Leu Lys Ala Asp
20 25 30
Ile Glu Lys Leu Ser Asp Glu Ala Leu Lys Gln Lys Thr Ile Glu Phe
35 40 45
10 Lys Glu Arg Leu Glu Lys Gly Glu Thr Val Asp Asp Leu Leu Val Glu
50 55 60
Ala Phe Ala Val Val Arg Glu Ala Ser Arg Arg Val Thr Gly Met Phe
65 70 75 80
15 Pro Phe Lys Val Gln Leu Met Gly Gly Val Ala Leu His Glu Gly Asn
85 90 95
Ile Ala Glu Met Lys Thr Gly Glu Gly Lys Thr Leu Thr Ser Thr Met
100 105 110
Pro Val Tyr Leu Asn Ala Leu Ser Gly Lys Gly Val His Val Val Thr
115 120 125
20 Val Asn Glu Tyr Leu Ala Ser Arg Asp Ala Glu Glu Met Gly Lys Ile
130 135 140
Phe Glu Phe Leu Gly Leu Thr Val Gly Leu Asn Leu Asn Ser Leu Ser
145 150 155 160
25 Lys Asp Glu Lys Arg Glu Ala Tyr Ala Ala Asp Ile Thr Tyr Ser Thr
165 170 175
Asn Asn Glu Leu Gly Phe Asp Tyr Leu Arg Asp Asn Met Val Leu Tyr
180 185 190
Lys Glu Gln Met Val Gln Arg Pro Leu His Phe Ala Val Ile Asp Glu
195 200 205
30 Val Asp Ser Ile Leu Ile Asp Glu Ala Arg Thr Pro Leu Ile Ile Ser
210 215 220
Gly Gln Ala Ala Lys Ser Thr Lys Leu Tyr Val Gln Ala Asn Ala Phe
225 230 235 240
Val Arg Thr Leu Lys Ala Asp Gln Asp Tyr Thr Tyr Asp Val Lys Thr
245 250 255
35 Lys Gly Val Gln Leu Thr Glu Glu Gly Met Thr Lys Ala Glu Lys Ala
260 265 270
Phe Gly Ile Glu Asn Leu Phe Asp Val Arg His Val Ala Leu Asn His
275 280 285
40 His Ile Ala Gln Ala Leu Lys Ala His Ala Ala Met His Lys Asp Val
290 295 300
Asp Tyr Val Val Glu Asp Gly Gln Val Val Ile Val Asp Ser Phe Thr
305 310 315 320
Gly Arg Leu Met Lys Gly Arg Arg Tyr Ser Asp Gly Leu His Gln Ala
325 330 335
45 Ile Glu Ala Lys Glu Gly Leu Glu Ile Gln Asn Glu Ser Met Thr Leu
340 345 350
Ala Thr Ile Thr Phe Gln Asn Tyr Phe Arg Met Tyr Glu Lys Leu Ala
355 360 365
Gly Met Thr Gly Thr Ala Lys Thr Glu Glu Glu Glu Phe Arg Asn Ile
370 375 380
50 Tyr Asn Met Gln Val Val Thr Ile Pro Thr Asn Lys Pro Ile Ala Arg
385 390 395 400
Asp Asp Arg Pro Asp Leu Ile Tyr Arg Thr Met Glu Gly Lys Phe Lys
405 410 415
55 Ala Val Ala Glu Asp Val Ala Gln Arg Tyr Met Val Gly Gln Pro Val

60

65

ES 2 303 058 T3

420 425 430
 Leu Val Gly Thr Val Ala Val Glu Thr Ser Glu Leu Ile Ser Arg Leu
 435 440 445
 5 Leu Lys Asn Lys Gly Ile Pro His Gln Val Leu Asn Ala Lys Asn His
 450 455 460
 Glu Arg Glu Ala Gln Ile Ile Glu Asp Ala Gly Gln Lys Gly Ala Val
 465 470 475 480
 Thr Ile Ala Thr Asn Met Ala Gly Arg Gly Thr Asp Ile Lys Leu Gly
 485 490 495
 10 Glu Gly Val Lys Glu Leu Gly Gly Leu Ala Val Ile Gly Thr Glu Arg
 500 505 510
 His Glu Ser Arg Arg Ile Asp Asn Gln Leu Arg Gly Arg Ser Gly Arg
 515 520 525
 15 Gln Gly Asp Pro Gly Ile Thr Gln Phe Tyr Leu Ser Met Glu Asp Glu
 530 535 540
 Leu Met Lys Arg Phe Gly Ala Glu Arg Thr Met Ala Met Leu Asp Arg
 545 550 555 560
 Phe Gly Met Asp Asp Ser Thr Pro Ile Gln Ser Lys Met Val Ser Arg
 565 570 575
 20 Ala Val Glu Ser Ser Gln Lys Arg Val Glu Gly Asn Asn Phe Asp Ala
 580 585 590
 Arg Lys Gln Leu Leu Gln Tyr Asp Asp Val Leu Arg Gln Gln Arg Glu
 595 600 605
 25 Val Ile Tyr Lys Gln Arg Phe Glu Val Ile Asp Ser Asp Asn Leu Arg
 610 615 620
 Ser Ile Val Glu Asn Met Ile Lys Ala Ser Leu Glu Arg Ala Val Ala
 625 630 635 640
 Ser Tyr Thr Pro Lys Glu Asp Leu Pro Glu Glu Trp Asn Leu Asp Gly
 645 650 655
 30 Leu Val Glu Leu Val Asn Ala Asn Phe Leu Asp Glu Gly Gly Val Glu
 660 665 670
 Lys Ser Asp Ile Phe Gly Lys Glu Pro Glu Glu Ile Thr Glu Leu Ile
 675 680 685
 Tyr Asp Arg Ile Lys Thr Lys Tyr Asp Glu Lys Glu Glu Arg Tyr Gly
 690 695 700
 35 Ser Glu Gln Met Arg Glu Phe Glu Lys Val Ile Val Leu Arg Glu Val
 705 710 715 720
 Asp Thr Lys Trp Met Asp His Ile Asp Ala Met Asp Gln Leu Arg Gln
 725 730 735
 40 Gly Ile His Leu Arg Ala Tyr Ala Gln Thr Asn Pro Leu Arg Glu Tyr
 740 745 750
 Gln Met Glu Gly Phe Ala Met Phe Glu Asn Met Ile Ala Ala Ile Glu
 755 760 765
 Asp Asp Val Ala Lys Phe Val Met Lys Ala Glu Ile Glu Asn Asn Leu
 770 775 780
 45 Glu Arg Glu Glu Val Ile Gln Gly Gln Thr Thr Ala His Gln Pro Lys
 785 790 795 800
 Glu Gly Asp Glu Glu Lys Gln Ala Lys Lys Lys Pro Val Arg Lys Ala
 805 810 815
 50 Val Asp Ile Gly Arg Asn Asp Pro Cys Tyr Cys Gly Ser Gly Lys Lys
 820 825 830
 Tyr Lys Asn Cys Cys Gly Arg Thr Glu
 835 840

55

60

65