(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 650 799**

(51) Int. CI.:

C12N 9/10 (2006.01)
C12N 9/90 (2006.01)
C12P 19/04 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **24.04.2008 PCT/JP2008/058323**

(87) Fecha y número de publicación internacional: **06.11.2008 WO08133350**

(96) Fecha de presentación y número de la solicitud europea: **24.04.2008 E 08740988 (4)**

(97) Fecha y número de publicación de la concesión europea: **01.11.2017 EP 2142643**

(54) Título: **Bacteria que produce condroitina y procedimiento de producción de condroitina**

(30) Prioridad:

24.04.2007 US 913713 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
22.01.2018

(73) Titular/es:

SEIKAGAKU CORPORATION (100.0%)
6-1, MARUNOUCHI 1-CHOME
CHIYODA-KU TOKYO 100-0005, JP

(72) Inventor/es:

SUZUKI, KIYOSHI;
MIYAMOTO, KENTARO y
KASEYAMA, HIROMI

(74) Agente/Representante:

DURAN-CORRETJER, S.L.P

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

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

Bacteria que produce condroitina y procedimiento de producción de condroitina

Sector técnico

La presente invención se refiere a una bacteria que produce condroitina y a un procedimiento de producción de condroitina.

Antecedentes técnicos

La condroitina es un polisacárido que comprende una estructura repetida de disacáridos de un resto de ácido glucurónico (GlcUA) y un resto de N-acetil-D-galactosamina (GalNAc) (-GlcUAβ (1-3)-GalNAcβ (1-4)-; en la presente memoria también se denomina un esqueleto de hidratos de carbono con condroitina). El sulfato de condroitina es un polisacárido que consiste en la sulfatación de la condroitina.

De manera convencional, la condroitina y el sulfato de condroitina se extraen y se purifican a partir de cartílagos, órganos y similares de animales. Sin embargo, en los últimos años, se ha estudiado una técnica para sintetizar artificialmente un esqueleto de hidratos de carbono común a la condroitina y al sulfato de condroitina, debido a la escasez de materiales, tales como cartílagos y órganos.

Se han descrito condroitina sintasas, que producen condroitina mediante la transferencia alternativa de GlcUA y GalNAc desde sus sustratos donantes a oligosacáridos aceptores y se ha propuesto un procedimiento de producción de condroitina utilizando la enzima.

El documento J. Biol. Chem. 275(31), 24124-24129 (2000) da a conocer una condroitina sintasa derivada de *Pasteurella multocida*.

Además, los documentos WO 2003/102193 y WO 2003/102194 dan a conocer condroitina sintasas derivadas de ser humano.

Además, el documento US 2003-0109693 (JP 2003-199583 A) da a conocer una condroitina sintasa novedosa (KfoC) producida por la cepa K4 de *Escherichia coli*.

Sin embargo, en el caso de producir condroitina mediante los procedimientos enzimáticos, es necesario preparar materiales costosos, tales como oligosacáridos como aceptores y nucleótidos de azúcar para los donantes de GlcUA y GalNAc, y, de este modo, se ha deseado un procedimiento de producción de condroitina que utilice materiales más económicos.

Es conocido que la cepa K4 de *Escherichia coli* produce un polisacárido que tiene una estructura de esqueleto de condroitina como una cápsula. Sin embargo, su estructura consiste en una unidad de repetición que comprende trisacáridos de un resto de GalNAc, un resto de GlcUA y un resto de fructosa que está unida a un grupo hidroxilo en C3 de un resto de GlcUA. Además, se han detectado más de 100 polisacáridos capsulares químicamente diferentes en *Escherichia coli*. Por ejemplo, la cepa K5 de *Escherichia coli* produce un polisacárido capsular K5 que tiene un esqueleto de hidratos de carbono con heparina/sulfato de heparán (J. Biol. Chem., 272 (5), pág. 2682-2687, 1997). Sin embargo, la existencia de *Escherichia coli* que produce condroitina en sí no es conocida. Por lo tanto, se desconoce un procedimiento de producción de condroitina utilizando una bacteria de *Escherichia*.

Aunque la solicitud de patente publicada de Estados Unidos 2007-0281342 da a conocer un procedimiento de producción de condroitina mediante la utilización de una bacteria *Bacillus* gram-positiva recombinante introducida con un gen de condroitina sintasa derivado de *Pasteurella multocida*, se ha deseado un mayor desarrollo en la producción fermentativa de condroitina. Otro documento del estado de la técnica que proporciona información de base relevante para la presente solicitud es Ninomiya y otros 2002 (J. Biol. Chem, vol. 277, No. 24, páginas 21575-21575), que da a conocer el procedimiento de clonación y caracterización de condroitina polimerasa. El documento WO0180810 da a conocer un procedimiento para aislar y producir un gen de condroitina sintasa a partir de *Pasteurella multocida* y la utilización de condroitina no sulfatada. Lidholt y otros 1997 (J. Biol. Chem, vol. 272, no. 5, páginas 2682-2687) da a conocer la síntesis del polisacárido capsular de *E. coli* K4. El documento EP1283259 da a conocer un procedimiento para producir una condroitina polimerasa, y Rodríguez y otros 1988 (Eur. J. Biochem. Vol. 177, páginas 117-124) da a conocer las características estructurales y serológicas del antígeno K4 capsular de *E. coli*.

Descripción de la invención

Un objetivo de la presente invención es dar a conocer un nuevo microorganismo capaz de producir condroitina y un nuevo procedimiento de producción de condroitina utilizando el microorganismo.

Los inventores de la presente invención han realizado amplios estudios y, como resultado, han descubierto que una bacteria productora de ácido glucurónico-UDP, tal como la cepa K5 de *Escherichia coli*, que se introduce con genes kfoA y kfoC derivados de la cepa K4 de *Escherichia coli,* produce condroitina con alta eficiencia, realizando, de este modo, la presente invención.

Un objetivo de la presente invención es dar a conocer una bacteria productora de ácido glucurónico-UDP, en la que se introduce un gen kfoA de la cepa K4 de *Escherichia coli* y un gen kfoC de la cepa K4 de *Escherichia coli*, en la que dicha bacteria tiene la capacidad de producir condroitina como un polisacárido capsular y en la que el gen kfoA codifica una proteína seleccionada del grupo que comprende (A) y (B): (A) una proteína que comprende la secuencia de aminoácidos de la SEQ ID NO: 2; y (B) una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 2, que incluye la sustitución, deleción, inserción o adición de 1 a 20 aminoácidos; y en la que el gen kfoC codifica una proteína seleccionada del grupo que comprende (C) a (F); (C) una proteína que comprende la secuencia de aminoácidos de la SEQ ID NO: 4; (D) una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 4, que incluye la sustitución, deleción, inserción o adición de 1 a 20 aminoácidos; (E) una proteína que comprende la secuencia de aminoácidos de la SEQ ID NO: 6; y (F) una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 6, que incluye la sustitución, deleción, inserción o adición de 1 a 20 aminoácidos.

El gen kfoA derivado de la cepa K4 de *Escherichia coli* puede ser, de manera preferente, un ADN seleccionado del grupo que comprende los siguientes (a) y (b):

(a) un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 1; y

(b) un ADN que hibrida con un ADN que comprende la secuencia de nucleótidos complementaria a la SEQ ID NO: 1 en condiciones rigurosas y que codifica una proteína que comprende la actividad de UDP-glucosa-4-epimerasa.

El gen kfoC derivado de la cepa K4 de *Escherichia coli* puede ser, de manera preferente, un ADN seleccionado del grupo que comprende los siguientes (c) y (d):

(c) un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 3; y

(d) un ADN que hibrida con un ADN que comprende la secuencia de nucleótidos complementaria a la SEQ ID NO: 3 en condiciones rigurosas y que codifica una proteína que tiene actividad de condroitina sintasa.

El gen kfoC derivado de la cepa K4 de *Escherichia coli* puede ser, de manera preferente, un ADN seleccionado del grupo que comprende los siguientes (e) y (f):

(e) un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 5; y

(f) un ADN que hibrida con un ADN que comprende la secuencia de nucleótidos complementaria a la SEQ ID NO: 5 en condiciones rigurosas y codifica una proteína que tiene actividad de condroitina sintasa.

La bacteria productora de ácido glucurónico-UDP es, de manera preferente, la cepa K5 de *Escherichia coli*.

Otro objetivo de la presente invención es dar a conocer un procedimiento de producción de condroitina que comprende, como mínimo, las siguientes etapas (1) y (2):

(1) cultivar la bacteria, tal como se ha descrito anteriormente; y

(2) recoger la condroitina del cultivo.

Otro objetivo de la presente invención es dar a conocer un procedimiento de producción de sulfato de condroitina que comprende: producir condroitina mediante el procedimiento, tal como se ha descrito anteriormente; y a continuación, sulfatar la condroitina para producir sulfato de condroitina.

Descripción breve de los dibujos

La figura 1 es un dibujo que muestra una estructura de un vector para expresar los genes kfoA y kfoC.

La figura 2 es un dibujo (fotografía) que muestra la expresión de las proteínas KfoA y kfoC. El carril 1 muestra una cepa introducida sin plásmido (control), el carril 2 muestra una cepa introducida con pTrcHis-kfoCA y M muestra un marcador de peso molecular.

La figura 3 muestra los resultados de análisis de composiciones de disacáridos utilizando un sistema de HPLC fluorescente para la fracción L tratada con condroitinasa ABC (cABC) (A), la fracción L tratada sin cABC (B) y la fracción L tratada con cABC inactivada por calor (C).

La figura 4 muestra los resultados de análisis de disacáridos por fluorescencia para la fracción E tratada con cABC (A), la fracción E tratada sin cABC (B) y la fracción E tratada con cABC inactivada por calor (C).

La figura 5 muestra los resultados de análisis de disacáridos por fluorescencia para la fracción S tratada con cABC (A), la fracción S tratada sin cABC (B) y la fracción S tratada con cABC inactivada por calor (C).

La figura 6 muestra los resultados de análisis de disacáridos por fluorescencia para el patrón de condroitina tratado con cABC (A) y una mezcla del patrón de condroitina tratado con cABC y la fracción L tratada con cABC (B).

La figura 7 es un dibujo (fotografía) que muestra la expresión de proteínas KfoA, kfoC y Kfo⊿ C utilizando (A) un anticuerpo contra KfoC (EK-C) como primer anticuerpo, (B) un anticuerpo contra KfoA (EK-A) como primer anticuerpo y (C) suero de conejo normal. Los carriles 1 y 5 muestran el marcador molecular, patrón MagicMarkXP (Invitrogen), el carril 2 muestra una cepa introducida con pTrcHis-kfoCA, el carril 3 muestra una cepa introducida con

pTrcHis-kfo⊿CA y el carril 4 muestra una cepa introducida con pTrcHis-kfoA.

Descripción de las realizaciones preferentes

En lo sucesivo, la presente invención se describirá en detalle.

<1> Bacteria de la presente invención

La bacteria de la presente invención es una bacteria productora de ácido glucurónico-UDP introducida con los genes kfoA y kfoC derivados de la cepa K4 de *Escherichia coli* y que tiene capacidad de producir condroitina.

La "bacteria productora de ácido glucurónico-UDP" proporciona ácido glucurónico-UDP, que se utiliza para la producción de un polisacárido compuesto por ácido glucurónico. La "bacteria productora de ácido glucurónico-UDP" no está limitada, siempre y cuando produzca ácido glucurónico-UDP y entre los ejemplos de las mismas se incluyen una bacteria que pertenece al género *Gluconacetobacter*, tal como *Gluconacetobacter hansenii* y *Gluconacetobacter xylinus*, una bacteria que pertenece al género *Rhizobium*, tal como *Rhizobium meffloti*, una bacteria que pertenece al género *Acetobacter*, tal como *Acetobacter xylinum*, una bacteria que pertenece al género *Erwinia*, tal como *Erwinia amylovora*, una bacteria que pertenece al género *Thiobacillus*, tal como *Thiobacillus ferrooxidans*, una bacteria que pertenece al género *Xylella*, tal como *Xylella fastidiosa*, una bacteria que pertenece al género *Sinorhizobium*, tal como *Sinorhizobium meliloti*, una bacteria que pertenece al género *Rhodococcus*, tal como *Rhodococcus rhodochrous*, una bacteria que pertenece al género *Klebsiella*, tal como *Klebsiella aerogenes*, una bacteria que pertenece al género *Enterobacter*, tal como *Enterobacter aerogenes* y una bacteria que pertenece al género *Escherichia*, tal como *Escherichia coli*.

La "bacteria productora de ácido glucurónico-UDP" es, de manera preferente, una bacteria que pertenece al género *Escherichia* y es capaz de producir ácido glucurónico-UDP y, de manera más preferente, una bacteria que pertenece a *Escherichia coli* y es capaz de producir ácido glucurónico-UDP.

Un ejemplo específico de la misma incluye la cepa K5 de *Escherichia coli*. La cepa K5 se ha depositado en la American Type Culture Collection (ATCC: PO Box 1549, Manassas, VA 20108, Estados Unidos de América) bajo el número de acceso ATCC23506 y está disponible en el catálogo o en la página principal de la ATCC.

La cepa a introducir con los genes kfoA y kfoC puede ser una cepa que es un derivado de la cepa K5 de *Escherichia coli* y dichas cepas derivadas se pueden obtener mediante la introducción de una mutación génica en la cepa K5 de *Escherichia coli* o la introducción de un gen en la cepa K5 de *Escherichia coli* mediante recombinación genética. Es decir, un aspecto de la bacteria de la presente invención incluye: una cepa obtenida mediante la introducción de los genes kfoA y kfoC derivados de la cepa K4 de *Escherichia coli* en la cepa K5 de *Escherichia coli*; una cepa obtenida mediante la introducción de los genes kfoA y kfoC derivados de la cepa K4 de *Escherichia coli* en la cepa K5 de *Escherichia coli* y la introducción adicional de una mutación génica; y una cepa obtenida mediante la introducción de los genes kfoA y kfoC derivados de la cepa K4 de *Escherichia coli* en la cepa K5 de *Escherichia coli* y la introducción adicional de otro gen.

Entre los ejemplos del gen kfoA derivado de la cepa K4 de *Escherichia coli* se incluyen un ADN que codifica una proteína que comprende la secuencia de aminoácidos de la SEQ ID NO: 2 y un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 1.

En general, la sustitución, deleción, inserción o adición en uno o varios aminoácidos que constituyen una proteína no tiene ningún efecto sobre la actividad de la proteína, en muchos casos y, por lo tanto, el gen kfoA derivado de la cepa K4 de *Escherichia coli* puede ser un gen que codifica una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 2, incluyendo la sustitución, deleción, inserción o adición de uno o varios aminoácidos y que tiene actividad de UDP-glucosa-4-epimerasa.

La frase "uno o varios aminoácidos", tal como se utiliza en el presente documento, se refiere al número de aminoácidos que pueden provocar una sustitución, deleción, inserción o adición sin afectar a la actividad de UDP-glucosa-4-epimerasa. Específicamente, el número es, por ejemplo, un número entero de 1 a 20, de manera preferente, un número entero de 1 a 10, de manera más preferente, un número entero de 1 a 5.

Paralelamente, el gen kfoA derivado de la cepa K4 de *Escherichia coli* puede ser un gen que codifica una proteína que comprende una secuencia de aminoácidos que es idéntica en no menos del 90%, de manera preferente, no menos del 95%, de manera más preferente, no menos del 98% a la secuencia completa de la SEQ ID NO: 2 y que tiene actividad de UDP-glucosa-4-epimerasa.

El gen kfoA derivado de la cepa K4 de *Escherichia coli* se puede obtener mediante PCR utilizando un ADN cromosómico de la cepa K4 de *Escherichia coli* como una plantilla.

La cepa K4 de *Escherichia coli* se ha depositado en la American Type Culture Collection (ATCC) con el número de

acceso ATCC23502 y está disponible en el catálogo o en la página principal de la ATCC.

El gen que codifica una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 2, incluyendo la sustitución, deleción, inserción o adición de uno o varios aminoácidos o una secuencia de aminoácidos que es idéntica en no menos del 90% a la secuencia completa de la SEQ ID NO: 2 y que tiene actividad de UDP-glucosa-4-epimerasa, se puede obtener mediante la modificación de la secuencia de nucleótidos de la SEQ ID NO: 1, de manera que se introduce la sustitución, deleción, inserción o adición de uno o varios aminoácidos en la secuencia de aminoácidos de la SEQ ID NO: 2 mediante mutagénesis específica de sitio (Kramer, W. y Frits, H.J., Meth. in Enzymol., 154, 350 (1987); Kunkel, T.A. y otros, Meth. in Enzymol., 154, 367 (1987)) o similares.

El gen también se puede obtener mediante: cribado de un ADN que codifica una proteína que tiene actividad UDP-glucosa-4-epimerasa mediante hibridación con un ADN que comprende la secuencia de nucleótidos complementaria a la secuencia de nucleótidos de la SEQ ID NO: 1 o una secuencia parcial de la misma en condiciones rigurosas.

La expresión "condiciones rigurosas" se refiere a condiciones en las que se forma el denominado híbrido específico y no se forma el híbrido no específico (véase Sambrook, J. y otros, Molecular Cloning. A Laboratory Manual, segunda edición, Cold Spring Harbor Laboratory Press (1989), etc.). Entre los ejemplos específicos de las "condiciones rigurosas" se incluyen las condiciones de hibridación a 42ºC en una solución que contiene formamida al 50%, 4 x SSC, HEPES 50 mM (pH 7,0), 10 x solución de Denhardt, ADN de esperma de salmón a 100 µg/ml y lavado con 2 X SSC, solución de SDS al 0,1% a temperatura ambiente y, a continuación, 0,1 X SSC, solución de SDS al 0,1% a 60ºC.

Si el gen obtenido, tal como se ha descrito anteriormente, codifica una proteína que tiene actividad UDP-glucosa-4-epimerasa puede evaluarse mediante la introducción del gen resultante en un huésped apropiado para expresar la proteína y la medición a continuación de la actividad de UDP-glucosa-4-epimerasa, según el procedimiento descrito en el documento J. Biol. Chem., 277 (24), pág. 21567-21575, 2002.

Entre los ejemplos del gen kfoC derivado de la cepa K4 de *Escherichia coli* se incluyen un ADN que codifica una proteína que comprende la secuencia de aminoácidos de la SEQ ID NO: 4 y un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 3. Entre los ejemplos del gen kfoC derivado de la cepa K4 de *Escherichia coli* también se incluyen un ADN que codifica una proteína que comprende la secuencia de aminoácidos de la SEQ ID NO: 6 y un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 5.

El gen kfoC derivado de la cepa K4 de *Escherichia coli* puede ser un gen que codifica una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 4 ó 6, que incluye la sustitución, deleción, inserción o adición de uno o varios aminoácidos y que tiene actividad de condroitina sintasa.

El gen kfoC derivado de la cepa K4 de *Escherichia coli* puede ser un gen que codifica una proteína que comprende una secuencia de aminoácidos que es idéntica en no menos del 90%, de manera preferente, no menos del 95%, de manera más preferente, no menos del 98% a la secuencia completa de la SEQ ID NO: 4 ó 6 y que tiene actividad de condroitina sintasa.

La expresión "actividad de condroitina sintasa", tal como se utiliza en el presente documento, se refiere a una actividad para transferir alternativamente GlcUA de un donante de GlcUA o GalNAc de un donante de GalNAc al extremo no reducido de una cadena de azúcares.

El gen kfoC derivado de la cepa K4 de *Escherichia coli* se puede obtener mediante PCR utilizando un ADN cromosómico de la cepa K4 de *Escherichia coli* como plantilla. Además, el gen kfoC que codifica la secuencia de aminoácidos de la SEQ ID NO: 6, tal como un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 5, se puede obtener según el procedimiento descrito en el documento WO2007/145197.

El gen que codifica una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 4 ó 6, incluyendo la sustitución, deleción, inserción o adición de uno o varios aminoácidos o una secuencia de aminoácidos que es idéntica en no menos del 90% a la secuencia completa de la SEQ ID NO: 4 ó 6 y que tiene actividad de condroitina sintasa, se puede obtener mediante la modificación de la secuencia de nucleótidos de la SEQ ID NO: 3 ó 5, de manera que la sustitución, deleción, inserción o adición de uno o varios aminoácidos se introduce en la secuencia de aminoácidos de la SEQ ID NO: 4 ó 6 mediante mutagénesis específica de sitio o similar.

El gen también se puede obtener mediante: cribado de un ADN que codifica una proteína que tiene actividad de condroitina sintasa mediante hibridación con un ADN que comprende la secuencia de nucleótidos complementaria a la secuencia de nucleótidos de la SEQ ID NO: 3 ó 5 o una secuencia parcial de la misma en condiciones rigurosas.

Las definiciones de las expresiones "uno o varios" y "condiciones rigurosas" son las mismas que se han descrito anteriormente.

Si el gen obtenido, tal como se ha descrito anteriormente, codifica una proteína que tiene actividad de condroitina sintasa puede evaluarse mediante la introducción del gen resultante en un huésped apropiado para expresar la proteína y a continuación la medición de la actividad de condroitina sintasa según el procedimiento descrito en el documento US 2003-0109693 (JP 2003-199583 A).

La bacteria de la presente invención se puede obtener mediante la introducción de un gen kfoA derivado de la cepa K4 de *Escherichia coli* y de un gen kfoC derivado de la cepa K4 de *Escherichia coli* en una bacteria productora de ácido glucurónico-UDP, tal como se ha descrito anteriormente.

El término "introducción" utilizado en la presente invención incluye la introducción de ambos genes en una bacteria productora de ácido glucurónico-UDP utilizando un vector, tal como un plásmido o un bacteriófago, y la introducción de ambos genes en un cromosoma de la bacteria productora de ácido glucurónico-UDP mediante recombinación homóloga o similar.

El vector, tal como un plásmido o un bacteriófago, a utilizar en el presente documento no está particularmente limitado, siempre y cuando pueda utilizarse para la introducción de genes en *Escherichia coli*, y entre los ejemplos de los mismos se incluyen pTrcHis (Invitrogen Corporation), el vector pET (Novagen) y el vector pGEX (Amersham Pharmacia).

En la introducción de un gen kfoA derivado de la cepa K4 de *Escherichia coli* y de un gen kfoC derivado de la cepa K4 de *Escherichia coli*, se pueden utilizar promotores nativos de los genes, pero es preferente utilizar un promotor que sea potente en *Escherichia coli*, tales como el promotor lac, el promotor trp, el promotor trc, el promotor PR o el promotor lacUV.

El gen kfoA y el gen kfoC derivados de la cepa K4 de *Escherichia coli* se pueden introducir uno por uno o de manera simultánea utilizando un único vector. Es más preferente utilizar un vector, tal como un plásmido o un bacteriófago, que contiene ambos genes kfoA y kfoC, ya que la introducción y la expresión de los genes se pueden realizar fácilmente.

Los genes kfoA y kfoC derivados de la cepa K4 de *Escherichia coli* se pueden introducir de manera que se expresen como proteínas de fusión con otros péptidos.

Entre los ejemplos de los otros péptidos se incluyen etiqueta de polihistidina y etiqueta de GST (glutatión-S-transferasa). La expresión de los productos génicos como proteínas de fusión es preferente porque la detección (confirmación de la expresión) de los productos génicos se puede realizar fácilmente.

La introducción de un gen se puede realizar mediante un procedimiento de transformación conocido. Entre los ejemplos del procedimiento se incluyen el procedimiento de electroporación, el procedimiento de DEAE-dextrano y el procedimiento de fosfato de calcio.

La introducción de un gen se puede confirmar mediante: un procedimiento de detección de un gen, tal como RT-PCR o transferencia Northern; un procedimiento de expresión de detección particular de una proteína recombinante, tal como transferencia de Western; y un procedimiento de medición de la actividad, tal como se ha descrito anteriormente.

Si el gen kfoA derivado de la cepa K4 de *Escherichia coli* y el gen kfoC derivado de la cepa K4 de *Escherichia coli* introducidos funcionan en una bacteria productora de ácido glucurónico-UDP, la bacteria adquiere la capacidad de producir condroitina como un polisacárido capsular.

<2> Procedimiento de producción de condroitina

El procedimiento de producción de condroitina de la presente invención comprende, como mínimo, las siguientes etapas (1) y (2).

(1) cultivar una bacteria de la presente invención.
(2) recoger la condroitina del cultivo.

El cultivo se puede realizar mediante un procedimiento general para el cultivo de bacterias que pertenecen al género *Escherichia*.

El medio de cultivo no está particularmente limitado, siempre y cuando se pueda utilizar para el cultivo de bacterias que pertenecen al género *Escherichia* y entre los ejemplos preferentes del mismo se incluyen medio LB (medio de Luria-Bertani) (que contiene 10,0 g de triptona Bacto, 5,0 g de extracto de levadura Bacto y 5,0 g de NaCl por litro) y medio CYG (ácido casamino al 2,0%, extracto de levadura al 0,5% y glucosa al 0,2%, ajustar el pH a 7,0 antes de utilizar el autoclave). En el caso en el que se introduce un gen utilizando un vector que contiene un gen resistente a antibióticos, el medio contiene, de manera preferente, un antibiótico correspondiente al gen.

Las condiciones de cultivo no están particularmente limitadas, siempre y cuando las bacterias que pertenecen al género *Escherichia* puedan crecer y, a efectos de producir condroitina con alta eficacia, el cultivo se realiza, de manera preferente, a una temperatura de 20ºC a 40ºC durante de 8 a 72 horas.

A efectos de producir la condroitina con alta eficacia, una bacteria se puede precultivar en una placa o medio líquido.

El procedimiento de recogida de la condroitina del cultivo no está particularmente limitado, siempre y cuando la condroitina se pueda recoger, y entre los ejemplos se incluye el procedimiento, tal como se describe en los ejemplos mostrados a continuación. Específicamente, las células recombinantes recogidas se suspenden en PBS (solución salina tamponada con fosfato) o similares; se tratan las células con lisozima, ADNasa I y proteinasa K; y se extraen las proteínas. La condroitina se puede detectar, por ejemplo, mediante el tratamiento de la fracción con condroitinasa y la realización de un análisis de la composición de disacárido.

<3> Procedimiento de producción de sulfato de condroitina

El sulfato de condroitina se puede producir mediante la sulfatación adicional de condroitina preparada a partir de las bacterias recombinantes descritas en la presente invención.

La sulfatación se puede realizar mediante un procedimiento conocido de sulfatación de glicosaminoglicanos, que no está particularmente limitado, y entre los ejemplos del mismo se incluye el procedimiento descrito en el documento JP 61-47701 A.

Además, la sulfatación se puede realizar utilizando una enzima para transferir un grupo sulfato a la condroitina (sulfotransferasa). Bajo la existencia del 3'-fosfoadenosina 5'-fosfosulfato (PAPS) como donante de sulfato, las sulfotransferasas pueden producir sulfato de condroitina mediante reacción a condroitina producida por las bacterias recombinantes descritas en la presente invención. Entre los ejemplos de una sulfotransferasa conocida para la transferencia de un grupo sulfato a la condroitina se incluyen la transferasa de grupo 6-O-sulfato de condroitina (J. Biol. Chem., 275 (28), 21075-21080 (2000)) y la transferasa del grupo 4-sulfato de galactosaminoglicano (JP 2000-4877 A).

EJEMPLOS

En lo sucesivo, la presente invención se describirá en detalle haciendo referencia a los ejemplos.

Ejemplo 1: Preparación del vector de coexpresión de kfoC/KfoA

<Amplificación del gen kfoA>

Se realizó una PCR utilizando un ADN cromosómico de la cepa K4 de *E. coli* como plantilla, según el procedimiento descrito en el documento J. Biol. Chem., Vol. 277, Edición 24, 21567-21575, y el producto de PCR resultante se insertó en pTrcHis (vector de expresión de la proteína de fusión con histidina, Invitrogen), para producir de este modo pTrcHis-kfoA.

Los cebadores utilizados son los siguientes:
K4A-SP: 5'-CGGGATCCCGATGAGTATTCTTAATCAAGC-3' (SEQ ID NO: 7) y
K4A-AS: 5'-GGAATTCCGGCCAGTCTACATGTTTATCAC-3' (SEQ ID NO: 8)

<Amplificación del gen kfoC>

Se realizó una PCR utilizando un ADN cromosómico de la cepa K4 de *E. coli* como plantilla, según el procedimiento descrito en el documento JP 2003-199583 A, y el producto de PCR resultante se insertó en pTrcHis, para producir de este modo pTrcHis-kfoC.

Los cebadores utilizados son los siguientes:
K4C-SP: 5'-CGGGATCCCGATGAGTATTCTTAATCAAGC-3' (SEQ ID NO: 9) y
K4C-AS: 5'-GGAATTCCGGCCAGTCTACATGTTTATCAC-3' (SEQ ID NO: 10).

<Preparación de pTrcHis-kfoCA>

El pTrcHis-kfoCA se preparó a partir de los pTrcHis-kfoA y pTrcHis-kfoC mencionados anteriormente de la siguiente manera.

En primer lugar, se introdujo el sitio NotI-SalI en la secuencia C-terminal de pTrcHis-kfoC mediante PCR. La PCR se realizó utilizando los cebadores mostrados a continuación para amplificar la longitud completa de pTrcHis-kfoC desde el lado C-terminal de kfoC, y el producto de PCR resultante se autoligó para producir de este modo un plásmido pTrcHis-kfoC-NotI-SalI.

Los cebadores utilizados son los siguientes:
NotSal-pTrcHisC-rev: 5'-GCGGCCGCAAAACAGCCAAGCTTCGAATTC-3' (SEQ ID NO: 11) y
NotSal-pTrcHisC-for: 5'-ACGCGTCGACGGCGGATGAGAGAAGATTTTCA-3' (SEQ ID NO: 12)

A continuación, se diseñaron los cebadores (NotI-RBS-KfoA-N: 5'-GCGGCCGCAAAATTAAAGAGGTATATATTAATGTATCGA-3' (SEQ ID NO: 13) y SalI-KfoA-C: 5'-GTCGACCTCTCATCCGCCAAAACA-3' (SEQ ID NO: 14)) en dirección 5' del sitio de unión al ribosoma (RBS) y la región C-terminal de kfoA de pTrcHis-kfoA, respectivamente, y se utilizaron para PCR utilizando pTrcHis-kfoA como plantilla. A continuación, el producto de PCR resultante se insertó en pCR4-TOPO (Invitrogen) para producir de este modo pCR4-TOPO-kfoA.

Se escindió un inserto que contenía kfoA de pCR4-TOPO-kfoA con NotI-SalI y se introdujo en el sitio NotI-SalI de pTrcHis-kfoC-NotI-SalI, y el plásmido resultante se denominó pTrcHis-kfoCA (figura 1).

Ejemplo 2

Coexpresión de la kfoC y KfoA recombinante

Se introdujo el vector de expresión, pTrcHis-kfoCA, en la cepa K5 de *Escherichia coli* mediante electroporación (100 µl de células, 200 Q, 25 µF, 2,5 kV, cubeta de 0,1 cm) para producir de este modo la cepa K5/pTrcHis-kfoCA de *E. coli*. Las colonias se transfirieron a medio LB complementado con 100 ppm de ampicilina y, a continuación, el cultivo de siembra se incubó a 37ºC durante una noche. Se transfirió un ml del cultivo de siembra a 20 ml de medio CYG complementado con 100 ppm de ampicilina. Las bacterias recombinantes se cultivaron a 37ºC durante 3 horas (DO600 = 1,8, 1,7-2,0) y, a continuación, se añadió IPTG al cultivo a una concentración final de 1,0 mM. El cultivo resultante se cultivó a 37ºC durante 5 horas adicionales para inducir la expresión de las proteínas recombinantes.

Se recogieron las células bacterianas de 900 µl del cultivo mediante centrifugación. Las células obtenidas se suspendieron en 90 µl de tampón de Laemmli (x1). Después de calentar en un baño de agua hirviendo durante 10 minutos, se sometieron 10 µl del sobrenadante a SDS-PAGE en un gel de gradiente 5/20% y se transfirieron a una membrana de PVDF, seguido por análisis de Western para detectar la expresión de proteínas recombinantes. Se utilizó anti-Penta-His-HRP (QIAGEN) diluido a 1/2.000 para la detección de proteínas de fusión con histidina recombinantes. Como resultado, se confirmaron las expresiones de las proteínas recombinantes KfoA y kfoC (figura 2).

Ejemplo 3

Preparación de polisacárido y su detección mediante análisis de disacárido

<Preparación de células bacterianas>

Se cultivó una cepa de K5/pTrcHis-kfoCA de *E. coli* en medio LB complementado con 100 ppm de ampicilina como cultivo de siembra. El cultivo de siembra (750 µl) se transfirió a medio CYG complementado con 100 ppm de ampicilina (15 ml) y, a continuación, se cultivó a 37ºC durante 3 horas. Al cultivo se le añadió IPTG (concentración final: 1 mM), y se realizó el cultivo durante 5 horas adicionales. Las células bacterianas se recogieron por centrifugación.

<Preparación del sobrenadante de cultivo>

El cultivo de siembra (1 ml) se transfirió a medio CYG complementado con 100 ppm de ampicilina (20 ml) y, a continuación, se cultivó a 37ºC durante 3 horas. Al cultivo se le añadió IPTG (concentración final: 0,1 mM) y se realizó el cultivo durante 5 horas adicionales. El sobrenadante de cultivo se recogió por centrifugación.

<Preparación de la fracción de polisacárido>

Los polisacáridos se prepararon a partir de las células recombinantes y el sobrenadante de cultivo mediante los tres procedimientos siguientes.

Preparación 1) Las células bacterianas obtenidas descritas anteriormente se resuspendieron en PBS y se trataron con lisozima y ADNasa I a 37ºC durante 1 hora, seguido de tratamiento con proteinasa K a 37ºC durante 1 hora. A continuación, las enzimas se sometieron a inactivación por calor y las proteínas se extrajeron mediante precipitación con sulfato amónico al 70%. La fracción sobrenadante resultante se dializó contra Tris-HCl 10 mM (pH 8,0) para producir la muestra L.

Preparación 2) Las células bacterianas obtenidas descritas anteriormente se trataron con los agentes para purificar

el plásmido: P1 (que contiene lisozima), P2, y P3 (QIAGEN) secuencialmente, según las instrucciones del fabricante. Después de recogerlo por centrifugación, el sobrenadante se sometió a precipitación con etanol. Los precipitados se recogieron por centrifugación y se resuspendieron en 100 µl de agua esterilizada. La suspensión se trató con ADNasa I a 37ºC durante 1 hora para digerir los ADN, seguido de tratamiento con proteinasa K a 37ºC durante 1 hora. A continuación, las enzimas se sometieron a inactivación por calor y las proteínas se extrajeron por precipitación con sulfato amónico al 70%. La fracción sobrenadante resultante se dializó contra Tris-HCl 10 mM (pH 8,0) para producir la muestra E.

Preparación 3) El sobrenadante de cultivo (20 ml) se evaporó a sequedad. Los materiales secados se disolvieron en 2 ml de agua destilada. La solución se dializó contra Tris-HCl 10 mM (pH 8,0) para producir la muestra S.

<Detección de condroitina mediante análisis de disacáridos>

Las muestras L, E y S (100 µl cada una) fueron tratadas con cABC (condroitinasa ABC; fabricada por Seikagaku Corporation) en tampón Tris-HCl 50 mM (pH 8,0) que contenía NaOAc 50 mM a 37ºC durante toda la noche.

Como controles, se prepararon muestras L, E y S tratadas sin cABC y muestras L, E y S tratadas con cABC inactivada por calor. Se realizó ultrafiltración para extraer las macromoléculas con tamaño igual o superior a 10.000 y las muestras se analizaron mediante análisis de la composición de disacáridos utilizando un sistema de HPLC fluorescente (J Biol Chem. Julio 2000 21; 275 (29): 2269 a 2275) (figuras 3, 4 y 5). Se utilizó Senshu Pak Decosil C22 (D.I. 4,6 x 150 mm) como una columna de separación en el sistema HPLC y las longitudes de onda de excitación y emisión del detector de fluorescencia se fijaron a 346 nm y 410 nm, respectivamente.

Como resultado, sólo en los casos de las muestras tratadas con cABC, se detectaron picos específicos. El perfil de la composición de disacáridos de cada muestra se correspondió estrechamente con el perfil de la muestra obtenida mediante el tratamiento de la condroitina patrón con cABC, tal como se muestra en la figura 6(A). Paralelamente, en el caso en el que la condroitina patrón tratada con cABC se mezcló con la muestra L tratada con cABC, sólo se detectó un pico y, de este modo, se confirmó que la muestra L contenía condroitina (figura 6(B)). En el caso de la muestra E y S, se obtuvieron los mismos resultados (datos no mostrados).

Ejemplo 4

Preparación del vector de coexpresión Kfo⊿C/KfoA

<Amplificación del Kfo⊿c>

Según el procedimiento descrito en el documento WO2007/145197, se clonó un ADN que codificaba KfoC truncada en N-terminal que comprendía la secuencia de nucleótidos de la SEQ ID NO: 5 (gen kfo⊿C). El gen kfo⊿C se insertó en pTrcHis para obtener pTrcHis-kfo⊿C.

<Preparación de pTrcHis-kfo⊿CA>

El plásmido pTrcHis-kfoCA preparado en el ejemplo 1 se escindió utilizando las enzimas de restricción SphI y SmaI. Se obtuvo un fragmento de ADN (2,9 kDa) mediante digestión del plásmido pTrcHis-kfo⊿C utilizando las mismas enzimas de restricción. Según el procedimiento convencional, el fragmento de ADN se insertó en el sitio SphI-SmaI del plásmido escindido y, a continuación, el plásmido resultante se denominó pTrcHis-kfo⊿CA.

Ejemplo 5

Preparación de anticuerpos contra KfoA y KfoC

<Preparación de anticuerpos contra KfoA>

Se sintetizó el oligopéptido que comprendía la secuencia peptídica de la SEQ ID No: 15 (CIVSR RDGDI AESWS SPEKA NK, pureza: >70%) y 4 mg del oligopéptido se conjugaron con KLH (hemocianina de lapa californiana) mediante el procedimiento convencional. Después de preparar la solución de antígeno mediante la mezcla con adyuvante completo, se llevó a cabo la inmunización 5 veces en el intervalo de 2 semanas. Se recogió la sangre completa, después de confirmar el título de anticuerpos. A partir de la sangre de cada conejo, se obtuvieron 48 ml (muestra A) y 60 ml (muestra B) de suero contra KfoA, respectivamente. Después de la purificación utilizando columna de Proteína A, se obtuvieron 32 ml de la fracción EK-A de IgG (concentración de IgG: 5,30 mg/ml) a partir de la muestra A.

<Preparación de anticuerpos contra KfoC>

Se sintetizó el oligopéptido que comprendía la secuencia de péptido de la SEQ ID No: 16 (CQEPP GKENE TDRAA GK, pureza: >70%) y 4 mg del oligopéptido se conjugaron con KLH (hemocianina de lapa californiana) mediante el

procedimiento convencional. Después de preparar la solución de antígeno mediante la mezcla con adyuvante completo, se llevó a cabo la inmunización 5 veces en el intervalo de 2 semanas. Se recogió la sangre completa, después de confirmar el título de anticuerpos. A partir de la sangre de cada conejo, se obtuvieron 65 ml (muestra A) y 39 ml (muestra B) de suero contra KfoA, respectivamente. Después de la purificación utilizando columna de Proteína A, se obtuvieron 39 ml de la fracción EK-A de IgG (concentración de IgG: 4,41 mg/ml) a partir de la muestra A.

Ejemplo 6

De la misma manera que en el ejemplo 2, se introdujo el plásmido pTrcHis-kfo⊿CA en la cepa K5 de *E. coli* para obtener la cepa K5/pTrcHis-kfo⊿CA de E. coli y la cepa recombinante se cultivó para evaluar la expresión de las proteínas recombinantes. El cultivo de siembra desarrollado en medio LB complementado con 100 ppm de ampicilina (0,5 ml) se transfirió a medio CYG complementado con 100 ppm de ampicilina (10 ml) y, a continuación, el cultivo se llevó a cabo a 37ºC durante 1,5 horas. Después de la adición de IPTG (concentración final: 1 mM), se continuó el cultivo durante 4 horas más. Las células bacterianas se recogieron a partir de 900 µl del cultivo y se suspendieron en 90 µl de tampón de Laemmli (x1). Después de calentar en un baño de agua hirviendo durante 10 min, se sometieron 10 µl del sobrenadante a SDS-PAGE en un gel al 7,5% y se transfirieron a una membrana de PVDF, seguido por análisis de Western para detectar la expresión de proteínas recombinantes. Se utilizaron anticuerpos contra KfoA (EK-A) y anticuerpos contra KfoC (EK-C) diluidos a 1/1.000 como primer anticuerpo para la detección de las proteínas recombinantes KfoA, KfoC y Kfo⊿C. Se utilizaron inmunoglobulinas HRP contra conejo (DAKO, # p0448) como segundo anticuerpo. Además, se llevaron a cabo los mismos exámenes utilizando el cultivo de K5/pTrcHis-kfoCA de *E. coli* y K5/pTrcHis-kfoA de *E. coli* para comparar el nivel de expresión de proteínas recombinantes. Como resultado, no hubo una gran diferencia en el nivel de expresión de KfoA y KfoC (Kfo⊿C) entre K5/pTrcHis-kfoCA de *E. coli* y K5/pTrcHis-kfoA de *E. coli* (figura 7).

Ejemplo 7

Preparación de polisacárido y su detección mediante análisis de disacáridos

<Preparación de célula bacteriana>

De la misma manera que en el ejemplo 3, se prepararon células bacterianas y sobrenadante del cultivo de las cepas de K5/pTrcHis-kfo⊿CA de *E. coli*, K5/pTrcHis-kfoCA de *E. coli* y K5/pTrcHis-kfoA de *E. coli* después de tratamiento en autoclave a 121ºC durante 5 minutos.

<Preparación de la fracción de polisacáridos>

El sobrenadante (1 ml) se dializó frente a agua corriente durante una noche y, a continuación, la solución dializada se liofilizó. Los materiales liofilizados se disolvieron en 100 µl de Tris-HCl 50 mM (pH 8,0) para utilización como una fracción de polisacáridos.

<Detección de condroitina mediante análisis de disacáridos>

De la misma manera que en el ejemplo 3, se analizó la fracción obtenida mediante el análisis de la composición de disacáridos. Se analizó la solución patrón de condroitina (200 µg/ml y 100 µg/ml) mediante HPLC fluorescente para elaborar la curva de calibración de condroitina. La relación entre la concentración de condroitina y el área de pico de azúcar doble insaturado,⊿Di-0S, se mostró en la ecuación-1.

[Condroitina; µg/ml de sobrenadante] = (0,000591 x [área del pico] + 1,219)/[índice de concentración] (r^2 = 0,9984) (Ecuación-1)

El pico de acuerdo con el disacárido insaturado (⊿Di-0S) se detectó en las muestras preparadas a partir de K5/pTrcHis-kfo⊿CA de *E. coli* y K5/pTrcHis-kfoCA de *E. coli* (datos no mostrados). La tabla 1 muestra el resultado de la concentración de condroitina calcula utilizando la ecuación-1 en los cultivos recombinantes.

Tabla 1 Comparación del contenido de condroitina en el cultivo de las cepas recombinantes.

K5 de E. coli	Concentración	Área de pico	Condroitina [µg/ml]
K5/pTrcHis-kfoA	10	10.187,4	0,7
K5/pTrcHis-kfoCA	10	505.032,3	30,0
K5/pTrcHis-kfo⊿CA	10	888.677,6	52,6

Mediante la comparación de la productividad de condroitina entre K5/pTrcHis-kfo⊿CA de E. coli y K5/pTrcHis-kfoCA de *E. coli*, la cepa recombinante K5/pTrcHis-kfo⊿CA de *E. coli* es más adecuada para la producción de condroitina

mediante fermentación. Este resultado sugirió que la producción de condroitina se podría mejorar mediante la utilización del gen kfo⊿C que codifica KfoC truncada en N-terminal.

Aplicabilidad industrial

Mediante la utilización de la bacteria de la presente invención, se pueden producir condroitina y sulfato de condroitina con una alta eficiencia y a bajo coste.

LISTADO DE SECUENCIAS

<110> Seikagaku Corporation

<120> Bacteria que produce condroitina y procedimiento de producción de condroitina

<130> F200717C7102

<150> US 60/913.713
<151> 24-04-2007

<160> 16

<170> PatentIn versión 3.3

<210> 1
<211> 1020
<212> ADN
<213> Escherichia coli

<220>
<221> CDS
<222> (1).. (1017)

<400> 1

```
atg aat ata tta gtt aca ggt gga gca ggc tat att ggc tcg cat act       48
Met Asn Ile Leu Val Thr Gly Gly Ala Gly Tyr Ile Gly Ser His Thr
1               5                  10                  15
agt tta tgt ctt ctg aat aaa ggt tac aat gtt gta atc att gac aac       96
Ser Leu Cys Leu Leu Asn Lys Gly Tyr Asn Val Val Ile Ile Asp Asn
            20                  25                  30
tta att aat tca tct tgc gag agc att cga agg att gaa tta ata gct      144
Leu Ile Asn Ser Ser Cys Glu Ser Ile Arg Arg Ile Glu Leu Ile Ala
        35                  40                  45
aaa aaa aaa gtt act ttc tat gag ttg aac atc aac aat gaa aaa gaa      192
Lys Lys Lys Val Thr Phe Tyr Glu Leu Asn Ile Asn Asn Glu Lys Glu
    50                  55                  60
gtt aat caa att cta aaa aaa cac aaa ttt gat tgt ata atg cat ttt      240
Val Asn Gln Ile Leu Lys Lys His Lys Phe Asp Cys Ile Met His Phe
65                  70                  75                  80
gcc ggt gca aag tct gtt gct gaa tct tta ata aaa ccc att ttt tat      288
Ala Gly Ala Lys Ser Val Ala Glu Ser Leu Ile Lys Pro Ile Phe Tyr
```

```
                85                  90                  95
tat gat aat aat gtt tca ggg acg ttg caa tta att aat tgc gct ata      336
Tyr Asp Asn Asn Val Ser Gly Thr Leu Gln Leu Ile Asn Cys Ala Ile
            100                 105                 110
aaa aac gat gtg gct aat ttt att ttt agc tct tct gca acg gtt tat      384
Lys Asn Asp Val Ala Asn Phe Ile Phe Ser Ser Ser Ala Thr Val Tyr
        115                 120                 125
ggt gaa agc aaa ata atg cct gta aca gaa gat tgc cat ata gga gga      432
Gly Glu Ser Lys Ile Met Pro Val Thr Glu Asp Cys His Ile Gly Gly
    130                 135                 140
aca tta aat cca tat ggt aca tca aag tat ata tca gaa ttg atg att      480
Thr Leu Asn Pro Tyr Gly Thr Ser Lys Tyr Ile Ser Glu Leu Met Ile
145                 150                 155                 160
aga gat att gca aaa aaa tat agc gat act aat ttt ttg tgt ctg aga      528
Arg Asp Ile Ala Lys Lys Tyr Ser Asp Thr Asn Phe Leu Cys Leu Arg
                165                 170                 175
tat ttt aac cca aca ggt gct cac gag tcg gga atg atc ggt gaa agt      576
Tyr Phe Asn Pro Thr Gly Ala His Glu Ser Gly Met Ile Gly Glu Ser
            180                 185                 190
ccc gct gat ata cca agc aat tta gtt cct tat ata tta caa gtt gct      624
Pro Ala Asp Ile Pro Ser Asn Leu Val Pro Tyr Ile Leu Gln Val Ala
        195                 200                 205
atg ggt aaa cta gaa aaa ctt atg gtg ttt ggg ggg gat tac cct aca      672
Met Gly Lys Leu Glu Lys Leu Met Val Phe Gly Gly Asp Tyr Pro Thr
    210                 215                 220
aag gat gga acc ggt gtt cgt gat tat ata cac gta atg gat tta gcg      720
Lys Asp Gly Thr Gly Val Arg Asp Tyr Ile His Val Met Asp Leu Ala
225                 230                 235                 240
gaa ggg cat gtg gct gct tta tct tac ctt ttc cgt gat aat aac act      768
Glu Gly His Val Ala Ala Leu Ser Tyr Leu Phe Arg Asp Asn Asn Thr
                245                 250                 255
aat tat cat gtt ttt aat tta ggt act ggt aaa gga tat tct gtt tta      816
Asn Tyr His Val Phe Asn Leu Gly Thr Gly Lys Gly Tyr Ser Val Leu
            260                 265                 270
gag ctg gtt tct acc ttt gaa aaa ata tct ggg gtt aga att cca tat      864
Glu Leu Val Ser Thr Phe Glu Lys Ile Ser Gly Val Arg Ile Pro Tyr
        275                 280                 285
gaa att gtt tcg aga aga gat ggg gat att gct gaa agt tgg tca tca      912
Glu Ile Val Ser Arg Arg Asp Gly Asp Ile Ala Glu Ser Trp Ser Ser
    290                 295                 300
cca gaa aaa gca aat aag tat ctc aat tgg aaa gct aaa agg gaa ttg      960
Pro Glu Lys Ala Asn Lys Tyr Leu Asn Trp Lys Ala Lys Arg Glu Leu
305                 310                 315                 320
```

```
gaa aca atg ctt gag gat gcc tgg cgc tgg caa atg aaa aac cca aat     1008
Glu Thr Met Leu Glu Asp Ala Trp Arg Trp Gln Met Lys Asn Pro Asn
                325                 330                 335
ggt tat att taa                                                  1020
Gly Tyr Ile
```

<210> 2
<211> 339
<212> PRT
<213> Escherichia coli

<400> 2

```
Met Asn Ile Leu Val Thr Gly Gly Ala Gly Tyr Ile Gly Ser His Thr
1               5                  10                  15
Ser Leu Cys Leu Leu Asn Lys Gly Tyr Asn Val Val Ile Ile Asp Asn
            20                  25                  30
Leu Ile Asn Ser Ser Cys Glu Ser Ile Arg Arg Ile Glu Leu Ile Ala
        35                  40                  45
Lys Lys Lys Val Thr Phe Tyr Glu Leu Asn Ile Asn Asn Glu Lys Glu
    50                  55                  60
Val Asn Gln Ile Leu Lys Lys His Lys Phe Asp Cys Ile Met His Phe
65                  70                  75                  80
Ala Gly Ala Lys Ser Val Ala Glu Ser Leu Ile Lys Pro Ile Phe Tyr
                85                  90                  95
Tyr Asp Asn Asn Val Ser Gly Thr Leu Gln Leu Ile Asn Cys Ala Ile
            100                 105                 110
Lys Asn Asp Val Ala Asn Phe Ile Phe Ser Ser Ser Ala Thr Val Tyr
        115                 120                 125
Gly Glu Ser Lys Ile Met Pro Val Thr Glu Asp Cys His Ile Gly Gly
    130                 135                 140
Thr Leu Asn Pro Tyr Gly Thr Ser Lys Tyr Ile Ser Glu Leu Met Ile
145                 150                 155                 160
Arg Asp Ile Ala Lys Lys Tyr Ser Asp Thr Asn Phe Leu Cys Leu Arg
                165                 170                 175
Tyr Phe Asn Pro Thr Gly Ala His Glu Ser Gly Met Ile Gly Glu Ser
            180                 185                 190
Pro Ala Asp Ile Pro Ser Asn Leu Val Pro Tyr Ile Leu Gln Val Ala
        195                 200                 205
Met Gly Lys Leu Glu Lys Leu Met Val Phe Gly Gly Asp Tyr Pro Thr
    210                 215                 220
Lys Asp Gly Thr Gly Val Arg Asp Tyr Ile His Val Met Asp Leu Ala
```

```
225                 230                 235                 240
Glu Gly His Val Ala Ala Leu Ser Tyr Leu Phe Arg Asp Asn Asn Thr
                245                 250                 255
Asn Tyr His Val Phe Asn Leu Gly Thr Gly Lys Gly Tyr Ser Val Leu
            260                 265                 270
Glu Leu Val Ser Thr Phe Glu Lys Ile Ser Gly Val Arg Ile Pro Tyr
        275                 280                 285
Glu Ile Val Ser Arg Arg Asp Gly Asp Ile Ala Glu Ser Trp Ser Ser
    290                 295                 300
Pro Glu Lys Ala Asn Lys Tyr Leu Asn Trp Lys Ala Lys Arg Glu Leu
305                 310                 315                 320
Glu Thr Met Leu Glu Asp Ala Trp Arg Trp Gln Met Lys Asn Pro Asn
                325                 330                 335
Gly Tyr Ile
```

<210> 3
<211> 2061
<212> ADN
<213> Escherichia coli

<220>
<221> CDS
<222> (1).. (2058)

<400> 3

```
atg agt att ctt aat caa gca ata aat tta tat aaa aac aaa aat tat      48
Met Ser Ile Leu Asn Gln Ala Ile Asn Leu Tyr Lys Asn Lys Asn Tyr
1               5                   10                  15
cgc caa gct tta tct ctt ttt gag aag gtt gct gaa att tat gat gtt      96
Arg Gln Ala Leu Ser Leu Phe Glu Lys Val Ala Glu Ile Tyr Asp Val
            20                  25                  30
agt tgg gtc gaa gca aat ata aaa tta tgc caa acc gca ctc aat ctt     144
Ser Trp Val Glu Ala Asn Ile Lys Leu Cys Gln Thr Ala Leu Asn Leu
        35                  40                  45
tct gaa gaa gtt gat aag tta aat cgt aaa gct gtt att gat att gat     192
Ser Glu Glu Val Asp Lys Leu Asn Arg Lys Ala Val Ile Asp Ile Asp
    50                  55                  60
gca gca aca aaa ata atg tgt tct aac gcc aaa gca att agt ctg aac     240
Ala Ala Thr Lys Ile Met Cys Ser Asn Ala Lys Ala Ile Ser Leu Asn
65                  70                  75                  80
```

```
gag gtt gaa aaa aat gaa ata ata agc aaa tac cga gaa ata acc gca      288
Glu Val Glu Lys Asn Glu Ile Ile Ser Lys Tyr Arg Glu Ile Thr Ala
                85                  90                  95
aag aaa tca gaa cgg gcg gag tta aag gaa gtc gaa ccc att cct tta      336
Lys Lys Ser Glu Arg Ala Glu Leu Lys Glu Val Glu Pro Ile Pro Leu
            100                 105                 110
gat tgg cct agt gat tta act tta ccg ccg tta cct gag agc aca aac      384
Asp Trp Pro Ser Asp Leu Thr Leu Pro Pro Leu Pro Glu Ser Thr Asn
        115                 120                 125
gat tat gtt tgg gcg ggg aaa aga aaa gag ctt gat gat tat cca aga      432
Asp Tyr Val Trp Ala Gly Lys Arg Lys Glu Leu Asp Asp Tyr Pro Arg
    130                 135                 140
aaa cag tta atc att gac ggg ctt agt att gta att cct aca tat aat      480
Lys Gln Leu Ile Ile Asp Gly Leu Ser Ile Val Ile Pro Thr Tyr Asn
145                 150                 155                 160
cga gca aaa ata ctt gca att aca ctt gct tgt ctt tgt aac caa aag      528
Arg Ala Lys Ile Leu Ala Ile Thr Leu Ala Cys Leu Cys Asn Gln Lys
                165                 170                 175
acc ata tac gac tat gaa gtt att gtt gcc gat gat gga agt aaa gaa      576
Thr Ile Tyr Asp Tyr Glu Val Ile Val Ala Asp Asp Gly Ser Lys Glu
            180                 185                 190
aat att gaa gaa ata gta aga gaa ttt gaa agt tta tta aat ata aaa      624
Asn Ile Glu Glu Ile Val Arg Glu Phe Glu Ser Leu Leu Asn Ile Lys
        195                 200                 205
tat gta cgt cag aag gat tat gga tat caa ctg tgt gct gtt aga aat      672
Tyr Val Arg Gln Lys Asp Tyr Gly Tyr Gln Leu Cys Ala Val Arg Asn
    210                 215                 220
ctt ggg ctt agg gct gca aag tat aat tat gtt gca att ctg gat tgt      720
Leu Gly Leu Arg Ala Ala Lys Tyr Asn Tyr Val Ala Ile Leu Asp Cys
225                 230                 235                 240
gat atg gct ccg aac cca cta tgg gtt cag tca tat atg gaa cta tta      768
Asp Met Ala Pro Asn Pro Leu Trp Val Gln Ser Tyr Met Glu Leu Leu
                245                 250                 255
gcg gtg gac gat aat gtt gct cta att ggc cct aga aaa tat ata gat      816
Ala Val Asp Asp Asn Val Ala Leu Ile Gly Pro Arg Lys Tyr Ile Asp
            260                 265                 270
aca agc aag cat aca tat tta gat ttc ctt tcc caa aaa tca cta ata      864
Thr Ser Lys His Thr Tyr Leu Asp Phe Leu Ser Gln Lys Ser Leu Ile
        275                 280                 285
aat gaa att cct gaa atc att act aat aat cag gtt gca ggc aag gtt      912
Asn Glu Ile Pro Glu Ile Ile Thr Asn Asn Gln Val Ala Gly Lys Val
    290                 295                 300
gag caa aac aaa tca gtt gac tgg cga ata gaa cat ttc aaa aat acc      960
```

```
Glu Gln Asn Lys Ser Val Asp Trp Arg Ile Glu His Phe Lys Asn Thr
305                 310                 315                 320
gat aat cta aga tta tgc aac aca cca ttt cga ttt ttt agc gga ggt    1008
Asp Asn Leu Arg Leu Cys Asn Thr Pro Phe Arg Phe Phe Ser Gly Gly
                325                 330                 335
aat gtc gct ttt gcg aaa aaa tgg ctt ttc cgt gca gga tgg ttt gat    1056
Asn Val Ala Phe Ala Lys Lys Trp Leu Phe Arg Ala Gly Trp Phe Asp
            340                 345                 350
gaa gag ttt acg cat tgg ggg ggg gag gat aat gag ttt gga tat cgt    1104
Glu Glu Phe Thr His Trp Gly Gly Glu Asp Asn Glu Phe Gly Tyr Arg
        355                 360                 365
ctc tac aga gaa gga tgt tac ttt cgg tct gtt gaa gga gca atg gca    1152
Leu Tyr Arg Glu Gly Cys Tyr Phe Arg Ser Val Glu Gly Ala Met Ala
    370                 375                 380
tat cat caa gaa cca ccc ggg aaa gaa aac gag acg gat cgt gcg gca    1200
Tyr His Gln Glu Pro Pro Gly Lys Glu Asn Glu Thr Asp Arg Ala Ala
385                 390                 395                 400
ggg aaa aat att act gtt caa ttg tta cag caa aaa gtt cct tat ttc    1248
Gly Lys Asn Ile Thr Val Gln Leu Leu Gln Gln Lys Val Pro Tyr Phe
                405                 410                 415
tat aga aaa aaa gaa aaa ata gaa tcc gcg aca tta aaa aga gta cca    1296
Tyr Arg Lys Lys Glu Lys Ile Glu Ser Ala Thr Leu Lys Arg Val Pro
            420                 425                 430
cta gta tct ata tat att ccc gcc tat aac tgc tct aaa tat att gtt    1344
Leu Val Ser Ile Tyr Ile Pro Ala Tyr Asn Cys Ser Lys Tyr Ile Val
        435                 440                 445
cgt tgt gtt gaa agc gcc ctt aat cag aca ata act gac tta gaa gta    1392
Arg Cys Val Glu Ser Ala Leu Asn Gln Thr Ile Thr Asp Leu Glu Val
    450                 455                 460
tgc ata tgc gat gat ggt tcc aca gat gat aca ttg cgg att ctt cag    1440
Cys Ile Cys Asp Asp Gly Ser Thr Asp Asp Thr Leu Arg Ile Leu Gln
465                 470                 475                 480
gag cat tat gca aac cat cct cga gtt cgt ttt att tca caa aaa aac    1488
Glu His Tyr Ala Asn His Pro Arg Val Arg Phe Ile Ser Gln Lys Asn
                485                 490                 495
aaa gga att ggt tca gca tct aat aca gca gtt aga ttg tgt cgg gga    1536
Lys Gly Ile Gly Ser Ala Ser Asn Thr Ala Val Arg Leu Cys Arg Gly
            500                 505                 510
ttc tat ata ggt cag tta gac tct gat gac ttt ctt gaa cca gat gct    1584
Phe Tyr Ile Gly Gln Leu Asp Ser Asp Asp Phe Leu Glu Pro Asp Ala
        515                 520                 525
gtt gaa cta tgt cta gat gaa ttt aga aaa gat cta tca ttg gca tgt    1632
Val Glu Leu Cys Leu Asp Glu Phe Arg Lys Asp Leu Ser Leu Ala Cys
```

```
    530                 535                 540
gtt tat aca act aac cgt aat ata gat cgt gaa ggt aat ttg ata tca     1680
Val Tyr Thr Thr Asn Arg Asn Ile Asp Arg Glu Gly Asn Leu Ile Ser
545                 550                 555                 560
aat ggc tat aat tgg ccc att tat tcg cga gaa aaa ctt act agt gca     1728
Asn Gly Tyr Asn Trp Pro Ile Tyr Ser Arg Glu Lys Leu Thr Ser Ala
                565                 570                 575
atg ata tgt cat cat ttc agg atg ttc aca gca aga gca tgg aac cta     1776
Met Ile Cys His His Phe Arg Met Phe Thr Ala Arg Ala Trp Asn Leu
            580                 585                 590
act gaa ggt ttc aac gaa tcg atc agc aac gca gtt gat tac gat atg     1824
Thr Glu Gly Phe Asn Glu Ser Ile Ser Asn Ala Val Asp Tyr Asp Met
        595                 600                 605
tat tta aaa ctt agt gaa gtt gga ccg ttc aag cat ata aac aaa att     1872
Tyr Leu Lys Leu Ser Glu Val Gly Pro Phe Lys His Ile Asn Lys Ile
    610                 615                 620
tgt tat aat cgc gta ttg cat ggt gaa aat acg tct ata aaa aag ttg     1920
Cys Tyr Asn Arg Val Leu His Gly Glu Asn Thr Ser Ile Lys Lys Leu
625                 630                 635                 640
gat att caa aag gaa aat cat ttt aaa gtt gtt aac gaa tca tta agt     1968
Asp Ile Gln Lys Glu Asn His Phe Lys Val Val Asn Glu Ser Leu Ser
                645                 650                 655
agg cta ggc ata aaa aaa tat aaa tat tca cca tta act aat ttg aat     2016
Arg Leu Gly Ile Lys Lys Tyr Lys Tyr Ser Pro Leu Thr Asn Leu Asn
            660                 665                 670
gaa tgt aga aaa tat acc tgg gaa aaa ata gag aat gat tta taa         2061
Glu Cys Arg Lys Tyr Thr Trp Glu Lys Ile Glu Asn Asp Leu
        675                 680                 685
```

<210> 4
<211> 686
<212> PRT
<213> Escherichia coli

<400> 4

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

50 55 60

Ala Ala Thr Lys Ile Met Cys Ser Asn Ala Lys Ala Ile Ser Leu Asn
65 70 75 80

Glu Val Glu Lys Asn Glu Ile Ile Ser Lys Tyr Arg Glu Ile Thr Ala
85 90 95

Lys Lys Ser Glu Arg Ala Glu Leu Lys Glu Val Glu Pro Ile Pro Leu
100 105 110

Asp Trp Pro Ser Asp Leu Thr Leu Pro Pro Leu Pro Glu Ser Thr Asn
115 120 125

Asp Tyr Val Trp Ala Gly Lys Arg Lys Glu Leu Asp Asp Tyr Pro Arg
130 135 140

Lys Gln Leu Ile Ile Asp Gly Leu Ser Ile Val Ile Pro Thr Tyr Asn
145 150 155 160

Arg Ala Lys Ile Leu Ala Ile Thr Leu Ala Cys Leu Cys Asn Gln Lys
165 170 175

Thr Ile Tyr Asp Tyr Glu Val Ile Val Ala Asp Asp Gly Ser Lys Glu
180 185 190

Asn Ile Glu Glu Ile Val Arg Glu Phe Glu Ser Leu Leu Asn Ile Lys
195 200 205

Tyr Val Arg Gln Lys Asp Tyr Gly Tyr Gln Leu Cys Ala Val Arg Asn
210 215 220

Leu Gly Leu Arg Ala Ala Lys Tyr Asn Tyr Val Ala Ile Leu Asp Cys
225 230 235 240

Asp Met Ala Pro Asn Pro Leu Trp Val Gln Ser Tyr Met Glu Leu Leu
245 250 255

Ala Val Asp Asp Asn Val Ala Leu Ile Gly Pro Arg Lys Tyr Ile Asp
260 265 270

Thr Ser Lys His Thr Tyr Leu Asp Phe Leu Ser Gln Lys Ser Leu Ile
275 280 285

Asn Glu Ile Pro Glu Ile Ile Thr Asn Asn Gln Val Ala Gly Lys Val
290 295 300

Glu Gln Asn Lys Ser Val Asp Trp Arg Ile Glu His Phe Lys Asn Thr
305 310 315 320

Asp Asn Leu Arg Leu Cys Asn Thr Pro Phe Arg Phe Phe Ser Gly Gly
325 330 335

Asn Val Ala Phe Ala Lys Lys Trp Leu Phe Arg Ala Gly Trp Phe Asp
340 345 350

Glu Glu Phe Thr His Trp Gly Gly Glu Asp Asn Glu Phe Gly Tyr Arg
355 360 365

Leu Tyr Arg Glu Gly Cys Tyr Phe Arg Ser Val Glu Gly Ala Met Ala
370 375 380

Tyr His Gln Glu Pro Pro Gly Lys Glu Asn Glu Thr Asp Arg Ala Ala
385 390 395 400

Gly Lys Asn Ile Thr Val Gln Leu Leu Gln Gln Lys Val Pro Tyr Phe
405 410 415
Tyr Arg Lys Lys Glu Lys Ile Glu Ser Ala Thr Leu Lys Arg Val Pro
420 425 430
Leu Val Ser Ile Tyr Ile Pro Ala Tyr Asn Cys Ser Lys Tyr Ile Val
435 440 445
Arg Cys Val Glu Ser Ala Leu Asn Gln Thr Ile Thr Asp Leu Glu Val
450 455 460
Cys Ile Cys Asp Asp Gly Ser Thr Asp Asp Thr Leu Arg Ile Leu Gln
465 470 475 480
Glu His Tyr Ala Asn His Pro Arg Val Arg Phe Ile Ser Gln Lys Asn
485 490 495
Lys Gly Ile Gly Ser Ala Ser Asn Thr Ala Val Arg Leu Cys Arg Gly
500 505 510
Phe Tyr Ile Gly Gln Leu Asp Ser Asp Asp Phe Leu Glu Pro Asp Ala
515 520 525
Val Glu Leu Cys Leu Asp Glu Phe Arg Lys Asp Leu Ser Leu Ala Cys
530 535 540
Val Tyr Thr Thr Asn Arg Asn Ile Asp Arg Glu Gly Asn Leu Ile Ser
545 550 555 560
Asn Gly Tyr Asn Trp Pro Ile Tyr Ser Arg Glu Lys Leu Thr Ser Ala
565 570 575
Met Ile Cys His His Phe Arg Met Phe Thr Ala Arg Ala Trp Asn Leu
580 585 590
Thr Glu Gly Phe Asn Glu Ser Ile Ser Asn Ala Val Asp Tyr Asp Met
595 600 605
Tyr Leu Lys Leu Ser Glu Val Gly Pro Phe Lys His Ile Asn Lys Ile
610 615 620
Cys Tyr Asn Arg Val Leu His Gly Glu Asn Thr Ser Ile Lys Lys Leu
625 630 635 640
Asp Ile Gln Lys Glu Asn His Phe Lys Val Val Asn Glu Ser Leu Ser
645 650 655
Arg Leu Gly Ile Lys Lys Tyr Lys Tyr Ser Pro Leu Thr Asn Leu Asn
660 665 670
Glu Cys Arg Lys Tyr Thr Trp Glu Lys Ile Glu Asn Asp Leu
675 680 685

<210> 5
<211> 1890
<212> ADN
<213> Escherichia coli

<220>
<221> CDS
<222> (1) .. (1887)

<400> 5

```
aaa gct gtt att gat att gat gca gca aca aaa ata atg tgt tct aac      48
Lys Ala Val Ile Asp Ile Asp Ala Ala Thr Lys Ile Met Cys Ser Asn
1               5                  10                  15
gcc aaa gca att agt ctg aac gag gtt gaa aaa aat gaa ata ata agc      96
Ala Lys Ala Ile Ser Leu Asn Glu Val Glu Lys Asn Glu Ile Ile Ser
            20                  25                  30
aaa tac cga gaa ata acc gca aag aaa tca gaa cgg gcg gag tta aag     144
Lys Tyr Arg Glu Ile Thr Ala Lys Lys Ser Glu Arg Ala Glu Leu Lys
        35                  40                  45
gaa gtc gaa ccc att cct tta gat tgg cct agt gat tta act tta ccg     192
Glu Val Glu Pro Ile Pro Leu Asp Trp Pro Ser Asp Leu Thr Leu Pro
    50                  55                  60
ccg tta cct gag agc aca aac gat tat gtt tgg gcg ggg aaa aga aaa     240
Pro Leu Pro Glu Ser Thr Asn Asp Tyr Val Trp Ala Gly Lys Arg Lys
65                  70                  75                  80
gag ctt gat gat tat cca aga aaa cag tta atc att gac ggg ctt agt     288
Glu Leu Asp Asp Tyr Pro Arg Lys Gln Leu Ile Ile Asp Gly Leu Ser
                85                  90                  95
att gta att cct aca tat aat cga gca aaa ata ctt gca att aca ctt     336
Ile Val Ile Pro Thr Tyr Asn Arg Ala Lys Ile Leu Ala Ile Thr Leu
            100                 105                 110
gct tgt ctt tgt aac caa aag acc ata tac gac tat gaa gtt att gtt     384
Ala Cys Leu Cys Asn Gln Lys Thr Ile Tyr Asp Tyr Glu Val Ile Val
        115                 120                 125
gcc gat gat gga agt aaa gaa aat att gaa gaa ata gta aga gaa ttt     432
Ala Asp Asp Gly Ser Lys Glu Asn Ile Glu Glu Ile Val Arg Glu Phe
    130                 135                 140
gaa agt tta tta aat ata aaa tat gta cgt cag aag gat tat gga tat     480
Glu Ser Leu Leu Asn Ile Lys Tyr Val Arg Gln Lys Asp Tyr Gly Tyr
145                 150                 155                 160
caa ctg tgt gct gtt aga aat ctt ggg ctt agg gct gca aag tat aat     528
Gln Leu Cys Ala Val Arg Asn Leu Gly Leu Arg Ala Ala Lys Tyr Asn
                165                 170                 175
tat gtt gca att ctg gat tgt gat atg gct ccg aac cca cta tgg gtt     576
Tyr Val Ala Ile Leu Asp Cys Asp Met Ala Pro Asn Pro Leu Trp Val
            180                 185                 190
cag tca tat atg gaa cta tta gcg gtg gac gat aat gtt gct cta att     624
Gln Ser Tyr Met Glu Leu Leu Ala Val Asp Asp Asn Val Ala Leu Ile
```

```
        195                 200                 205
ggc cct aga aaa tat ata gat aca agc aag cat aca tat tta gat ttc      672
Gly Pro Arg Lys Tyr Ile Asp Thr Ser Lys His Thr Tyr Leu Asp Phe
    210                 215                 220
ctt tcc caa aaa tca cta ata aat gaa att cct gaa atc att act aat      720
Leu Ser Gln Lys Ser Leu Ile Asn Glu Ile Pro Glu Ile Ile Thr Asn
225                 230                 235                 240
aat cag gtt gca ggc aag gtt gag caa aac aaa tca gtt gac tgg cga      768
Asn Gln Val Ala Gly Lys Val Glu Gln Asn Lys Ser Val Asp Trp Arg
                245                 250                 255
ata gaa cat ttc aaa aat acc gat aat cta aga tta tgc aac aca cca      816
Ile Glu His Phe Lys Asn Thr Asp Asn Leu Arg Leu Cys Asn Thr Pro
            260                 265                 270
ttt cga ttt ttt agc gga ggt aat gtc gct ttt gcg aaa aaa tgg ctt      864
Phe Arg Phe Phe Ser Gly Gly Asn Val Ala Phe Ala Lys Lys Trp Leu
        275                 280                 285
ttc cgt gca gga tgg ttt gat gaa gag ttt acg cat tgg ggg ggg gag      912
Phe Arg Ala Gly Trp Phe Asp Glu Glu Phe Thr His Trp Gly Gly Glu
    290                 295                 300
gat aat gag ttt gga tat cgt ctc tac aga gaa gga tgt tac ttt cgg      960
Asp Asn Glu Phe Gly Tyr Arg Leu Tyr Arg Glu Gly Cys Tyr Phe Arg
305                 310                 315                 320
tct gtt gaa gga gca atg gca tat cat caa gaa cca ccc ggg aaa gaa     1008
Ser Val Glu Gly Ala Met Ala Tyr His Gln Glu Pro Pro Gly Lys Glu
                325                 330                 335
aac gag acg gat cgt gcg gca ggg aaa aat att act gtt caa ttg tta     1056
Asn Glu Thr Asp Arg Ala Ala Gly Lys Asn Ile Thr Val Gln Leu Leu
            340                 345                 350
cag caa aaa gtt cct tat ttc tat aga aaa aaa gaa aaa ata gaa tcc     1104
Gln Gln Lys Val Pro Tyr Phe Tyr Arg Lys Lys Glu Lys Ile Glu Ser
        355                 360                 365
gcg aca tta aaa aga gta cca cta gta tct ata tat att ccc gcc tat     1152
Ala Thr Leu Lys Arg Val Pro Leu Val Ser Ile Tyr Ile Pro Ala Tyr
    370                 375                 380
aac tgc tct aaa tat att gtt cgt tgt gtt gaa agc gcc ctt aat cag     1200
Asn Cys Ser Lys Tyr Ile Val Arg Cys Val Glu Ser Ala Leu Asn Gln
385                 390                 395                 400
aca ata act gac tta gaa gta tgc ata tgc gat gat ggt tcc aca gat     1248
Thr Ile Thr Asp Leu Glu Val Cys Ile Cys Asp Asp Gly Ser Thr Asp
                405                 410                 415
gat aca ttg cgg att ctt cag gag cat tat gca aac cat cct cga gtt     1296
Asp Thr Leu Arg Ile Leu Gln Glu His Tyr Ala Asn His Pro Arg Val
            420                 425                 430
```

```
cgt ttt att tca caa aaa aac aaa gga att ggt tca gca tct aat aca     1344
Arg Phe Ile Ser Gln Lys Asn Lys Gly Ile Gly Ser Ala Ser Asn Thr
        435                 440                 445
gca gtt aga ttg tgt cgg gga ttc tat ata ggt cag tta gac tct gat     1392
Ala Val Arg Leu Cys Arg Gly Phe Tyr Ile Gly Gln Leu Asp Ser Asp
    450                 455                 460
gac ttt ctt gaa cca gat gct gtt gaa cta tgt cta gat gaa ttt aga     1440
Asp Phe Leu Glu Pro Asp Ala Val Glu Leu Cys Leu Asp Glu Phe Arg
465                 470                 475                 480
aaa gat cta tca ttg gca tgt gtt tat aca act aac cgt aat ata gat     1488
Lys Asp Leu Ser Leu Ala Cys Val Tyr Thr Thr Asn Arg Asn Ile Asp
                485                 490                 495
cgt gaa ggt aat ttg ata tca aat ggc tat aat tgg ccc att tat tcg     1536
Arg Glu Gly Asn Leu Ile Ser Asn Gly Tyr Asn Trp Pro Ile Tyr Ser
            500                 505                 510
cga gaa aaa ctt act agt gca atg ata tgt cat cat ttc agg atg ttc     1584
Arg Glu Lys Leu Thr Ser Ala Met Ile Cys His His Phe Arg Met Phe
        515                 520                 525
aca gca aga gca tgg aac cta act gaa ggt ttc aac gaa tcg atc agc     1632
Thr Ala Arg Ala Trp Asn Leu Thr Glu Gly Phe Asn Glu Ser Ile Ser
    530                 535                 540
aac gca gtt gat tac gat atg tat tta aaa ctt agt gaa gtt gga ccg     1680
Asn Ala Val Asp Tyr Asp Met Tyr Leu Lys Leu Ser Glu Val Gly Pro
545                 550                 555                 560
ttc aag cat ata aac aaa att tgt tat aat cgc gta ttg cat ggt gaa     1728
Phe Lys His Ile Asn Lys Ile Cys Tyr Asn Arg Val Leu His Gly Glu
                565                 570                 575
aat acg tct ata aaa aag ttg gat att caa aag gaa aat cat ttt aaa     1776
Asn Thr Ser Ile Lys Lys Leu Asp Ile Gln Lys Glu Asn His Phe Lys
            580                 585                 590
gtt gtt aac gaa tca tta agt agg cta ggc ata aaa aaa tat aaa tat     1824
Val Val Asn Glu Ser Leu Ser Arg Leu Gly Ile Lys Lys Tyr Lys Tyr
        595                 600                 605
tca cca tta act aat ttg aat gaa tgt aga aaa tat acc tgg gaa aaa     1872
Ser Pro Leu Thr Asn Leu Asn Glu Cys Arg Lys Tyr Thr Trp Glu Lys
    610                 615                 620
ata gag aat gat tta taa                                            1890
Ile Glu Asn Asp Leu
625
```

<210> 6
<211> 629
<212> PRT
<213> Escherichia coli

<400> 6

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

305 310 315 320
Ser Val Glu Gly Ala Met Ala Tyr His Gln Glu Pro Pro Gly Lys Glu
325 330 335
Asn Glu Thr Asp Arg Ala Ala Gly Lys Asn Ile Thr Val Gln Leu Leu
340 345 350
Gln Gln Lys Val Pro Tyr Phe Tyr Arg Lys Lys Glu Lys Ile Glu Ser
355 360 365
Ala Thr Leu Lys Arg Val Pro Leu Val Ser Ile Tyr Ile Pro Ala Tyr
370 375 380
Asn Cys Ser Lys Tyr Ile Val Arg Cys Val Glu Ser Ala Leu Asn Gln
385 390 395 400
Thr Ile Thr Asp Leu Glu Val Cys Ile Cys Asp Asp Gly Ser Thr Asp
405 410 415
Asp Thr Leu Arg Ile Leu Gln Glu His Tyr Ala Asn His Pro Arg Val
420 425 430
Arg Phe Ile Ser Gln Lys Asn Lys Gly Ile Gly Ser Ala Ser Asn Thr
435 440 445
Ala Val Arg Leu Cys Arg Gly Phe Tyr Ile Gly Gln Leu Asp Ser Asp
450 455 460
Asp Phe Leu Glu Pro Asp Ala Val Glu Leu Cys Leu Asp Glu Phe Arg
465 470 475 480
Lys Asp Leu Ser Leu Ala Cys Val Tyr Thr Thr Asn Arg Asn Ile Asp
485 490 495
Arg Glu Gly Asn Leu Ile Ser Asn Gly Tyr Asn Trp Pro Ile Tyr Ser
500 505 510
Arg Glu Lys Leu Thr Ser Ala Met Ile Cys His His Phe Arg Met Phe
515 520 525
Thr Ala Arg Ala Trp Asn Leu Thr Glu Gly Phe Asn Glu Ser Ile Ser
530 535 540
Asn Ala Val Asp Tyr Asp Met Tyr Leu Lys Leu Ser Glu Val Gly Pro
545 550 555 560
Phe Lys His Ile Asn Lys Ile Cys Tyr Asn Arg Val Leu His Gly Glu
565 570 575
Asn Thr Ser Ile Lys Lys Leu Asp Ile Gln Lys Glu Asn His Phe Lys
580 585 590
Val Val Asn Glu Ser Leu Ser Arg Leu Gly Ile Lys Lys Tyr Lys Tyr
595 600 605
Ser Pro Leu Thr Asn Leu Asn Glu Cys Arg Lys Tyr Thr Trp Glu Lys
610 615 620
Ile Glu Asn Asp Leu
625

<210> 7
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 7
cgggatcccg atgagtattc ttaatcaagc 30

<210> 8
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 8
ggaattccgg ccagtctaca tgtttatcac 30

<210> 9
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 9
cgggatcccg atgagtattc ttaatcaagc 30

<210> 10
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 10
ggaattccgg ccagtctaca tgtttatcac 30

<210> 11
<211> 30
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 11
gcggccgcaa aacagccaag cttcgaattc 30

<210> 12
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 12
acgcgtcgac ggcggatgag agaagatttt ca 32

<210> 13
<211> 39
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 13
gcggccgcaa aattaaagag gtatatatta atgtatcga 39

<210> 14
<211> 24
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador

<400> 14
gtcgacctct catccgccaa aaca 24

<210> 15
<211> 22
<212> PRT
<213> Secuencia artificial

<220>
<223> péptido parcial de KfoA

<400> 15

Cys Ile Val Ser Arg Arg Asp Gly Asp Ile Ala Glu Ser Trp Ser Ser
1 5 10 15
Pro Glu Lys Ala Asn Lys
20

<210> 16
<211> 17
<212> PRT
<213> Secuencia artificial

<220>
<223> péptido parcial de KfoC

<400> 16

Cys Gln Glu Pro Pro Gly Lys Glu Asn Glu Thr Asp Arg Ala Ala Gly
1 5 10 15
Lys

REIVINDICACIONES

1. Bacteria productora de ácido glucurónico-UDP, en la que se introducen un gen kfoA de la cepa K4 de *Escherichia coli* y un gen kfoC de la cepa K4 de *Escherichia coli*, en la que dicha bacteria tiene la capacidad de producir condroitina como un polisacárido capsular,
y en la que el gen kfoA codifica una proteína seleccionada del grupo que comprende (A) y (B)
(A) una proteína que comprende la secuencia de aminoácidos de la SEQ ID NO: 2; y
(B) una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 2, que incluye la sustitución, deleción, inserción o adición de 1 a 20 aminoácidos y que tiene actividad de UDP-glucosa-4-epimerasa; y
en la que el gen kfoC codifica una proteína seleccionada del grupo que comprende (C) a (F);
(C) una proteína que comprende la secuencia de aminoácidos de la SEQ ID NO: 4;
(D) una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 4, que incluye la sustitución, deleción, inserción o adición de 1 a 20 aminoácidos y que tiene actividad de condroitina sintasa;
(E) una proteína que comprende la secuencia de aminoácidos de la SEQ ID NO: 6; y
(F) una proteína que comprende una secuencia de aminoácidos de la SEQ ID NO: 6, que incluye la sustitución, deleción, inserción o adición de 1 a 20 aminoácidos y que tiene actividad de condroitina sintasa.

2. Bacteria, según la reivindicación 1, en la que el gen kfoA es un ADN seleccionado del grupo que comprende (a) y (b):
(a) un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 1; y
(b) un ADN que hibrida con un ADN que comprende la secuencia de nucleótidos complementaria a la SEQ ID NO: 1 en condiciones rigurosas y que codifica una proteína que tiene actividad de UDP-glucosa-4-epimerasa.

3. Bacteria, según la reivindicación 1 ó 2, en la que el gen kfoC es un ADN seleccionado del grupo que comprende (c) y (d):
(c) un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 3; y
(d) un ADN que hibrida con un ADN que comprende la secuencia de nucleótidos complementaria a la SEQ ID NO: 3 en condiciones rigurosas y que codifica una proteína que tiene actividad de condroitina sintasa.

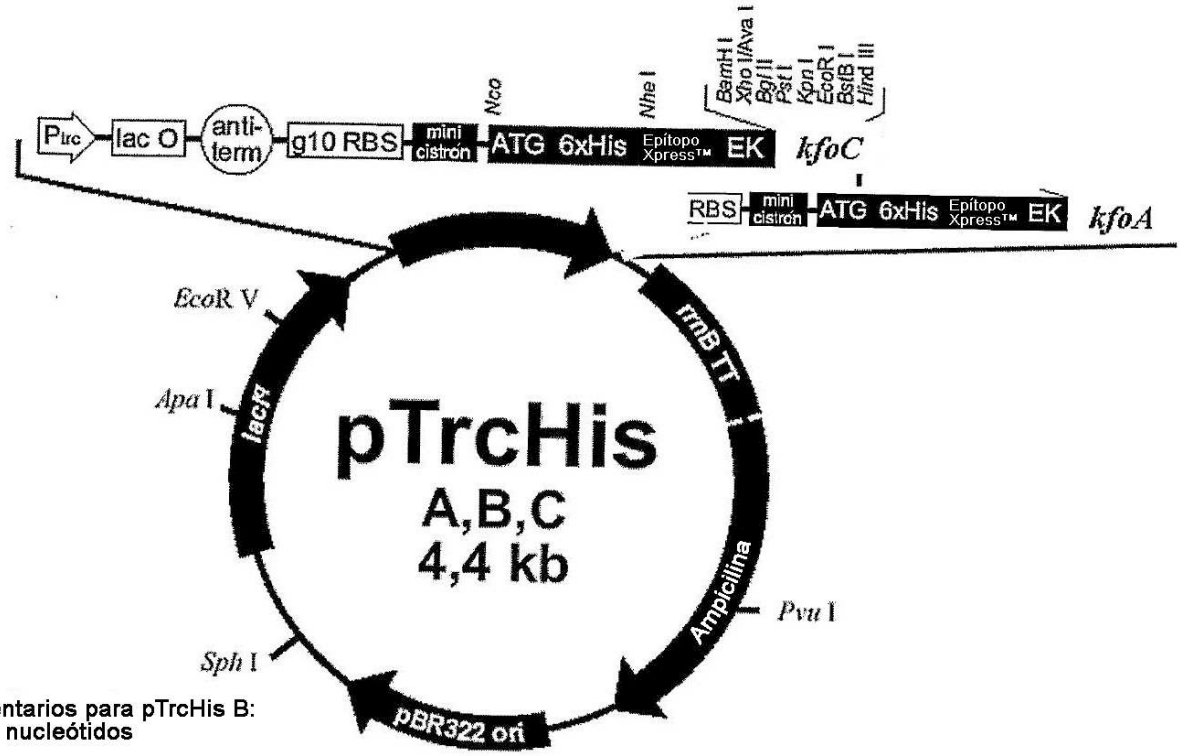
4. Bacteria, según la reivindicación 1 ó 2, en la que el gen kfoC es un ADN seleccionado del grupo que comprende (e) y (f):
(e) un ADN que comprende la secuencia de nucleótidos de la SEQ ID NO: 5; y
(f) un ADN que hibrida con un ADN que comprende la secuencia de nucleótidos complementaria a la SEQ ID NO: 5 en condiciones rigurosas y que codifica una proteína que tiene actividad de condroitina sintasa.

5. Bacteria, según cualquiera de las reivindicaciones 1 a 4, en la que dicha bacteria es la cepa K5 de *Escherichia coli*.

6. Procedimiento de producción de condroitina que comprende, como mínimo, las siguientes etapas (1) y (2):
(1) cultivar la bacteria, según cualquiera de las reivindicaciones 1 a 5; y
(2) recoger la condroitina del cultivo.

7. Procedimiento de producción de sulfato de condroitina que comprende: producir condroitina mediante el procedimiento, según la reivindicación 6; y sulfatar la condroitina para producir sulfato de condroitina.

Fig. 1



Comentarios para pTrcHis B:
4.404 nucleótidos

Fig. 2

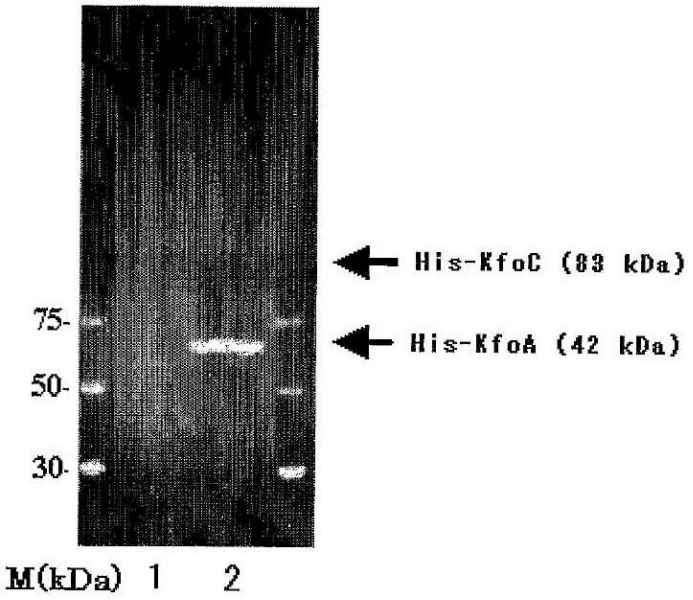
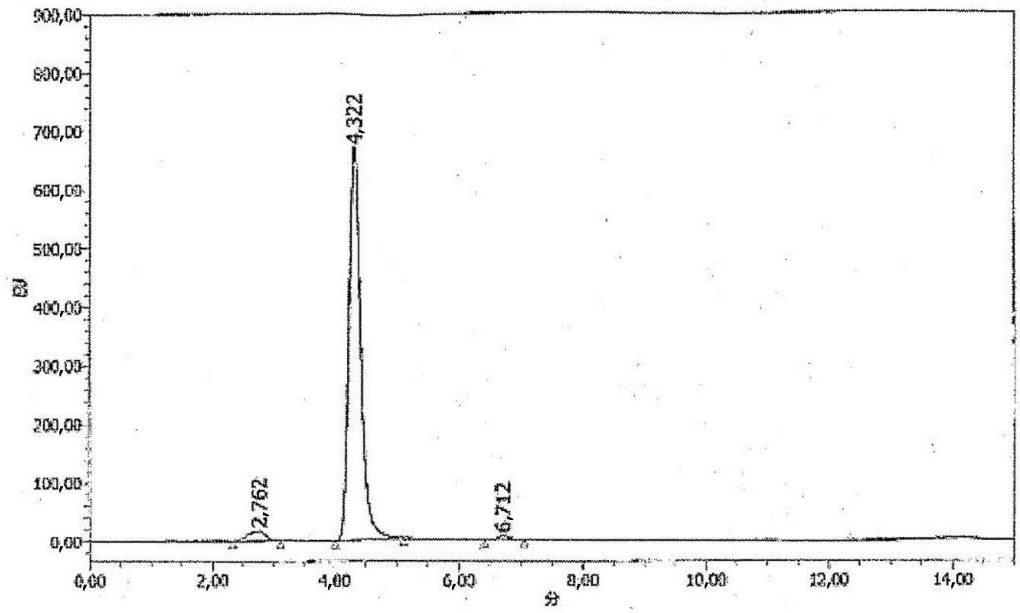
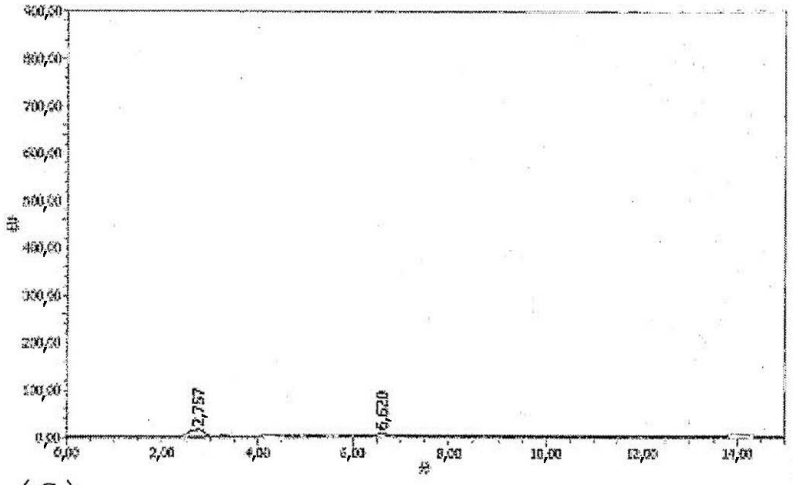


Fig. 3

(A)


900,00
800,00
700,00
600,00
500,00
400,00
300,00
200,00
100,00
0,00
mV
2,762
4,322
6,712
0,00
2,00
4,00
6,00
8,00
10,00
12,00
14,00
分

(B)


mV
2,757
6,628
0,00
2,00
4,00
6,00
8,00
10,00
12,00
14,00
分

(C)

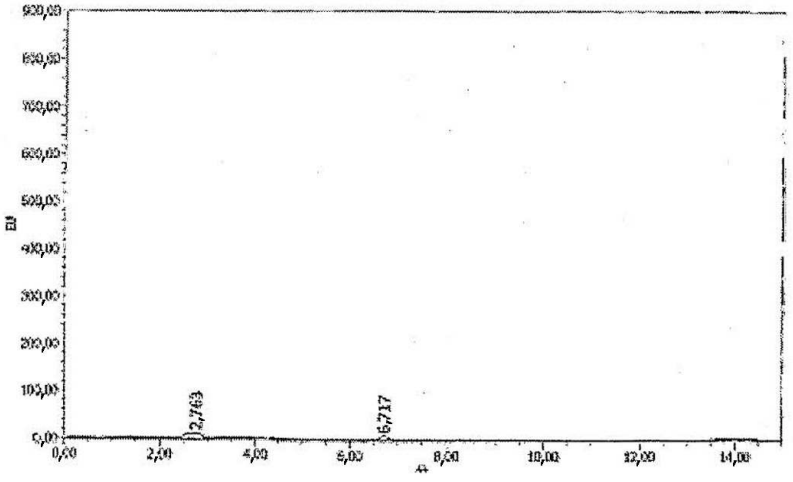
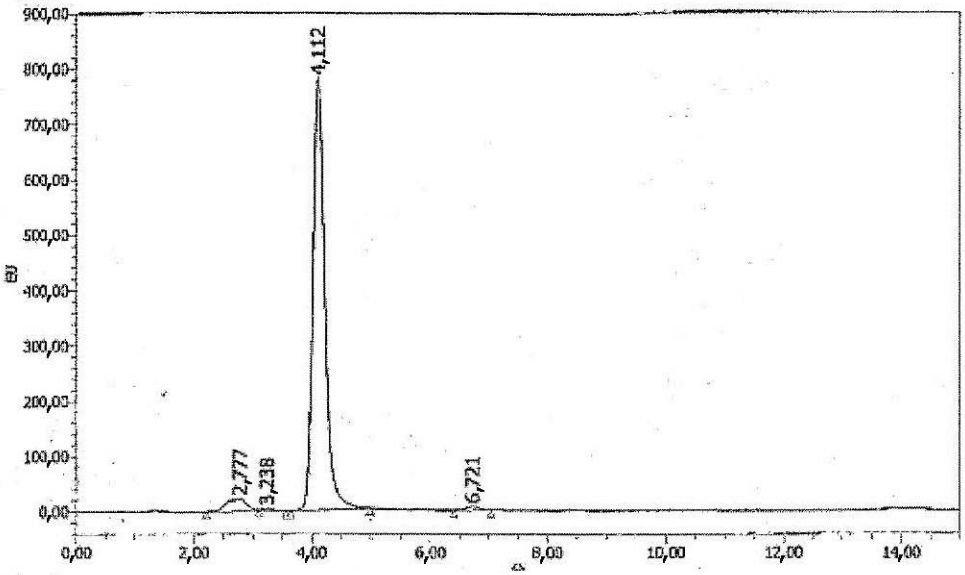
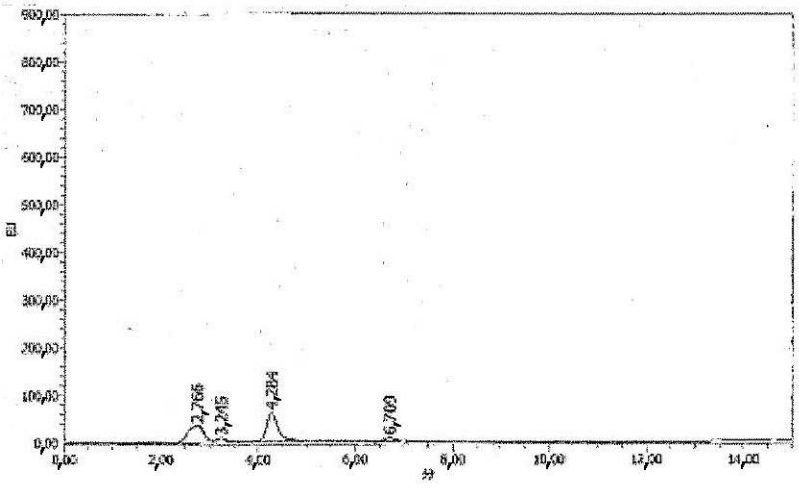

mV
2,763
6,717
0,00
2,00
4,00
6,00
8,00
10,00
12,00
14,00
分

Fig. 4

(A)


4,112
2,777
3,238
6,721
EU
0,00
2,00
4,00
6,00
8,00
10,00
12,00
14,00

(B)


2,766
3,245
4,284
6,709
EU

(C)

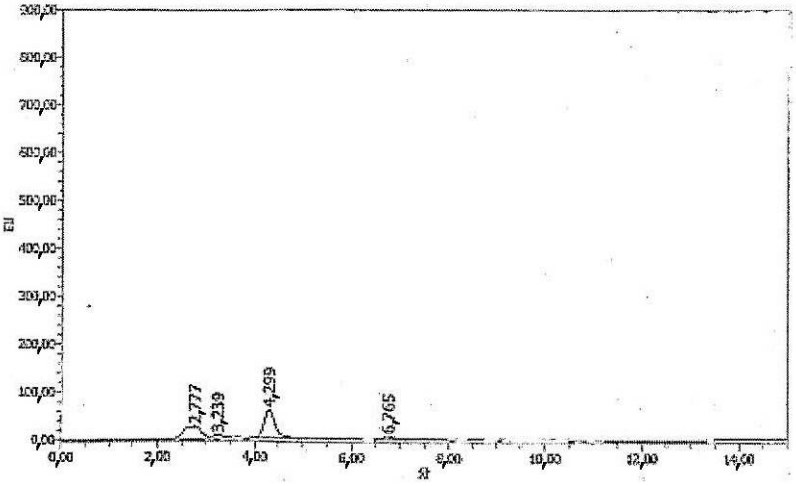
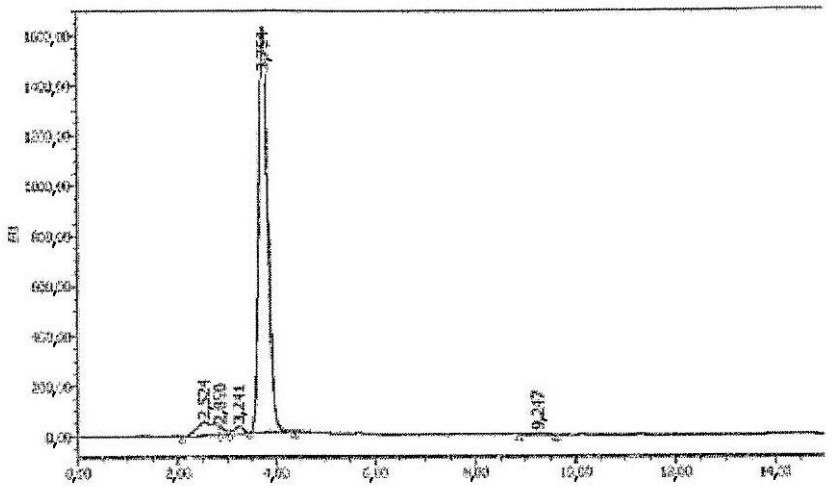
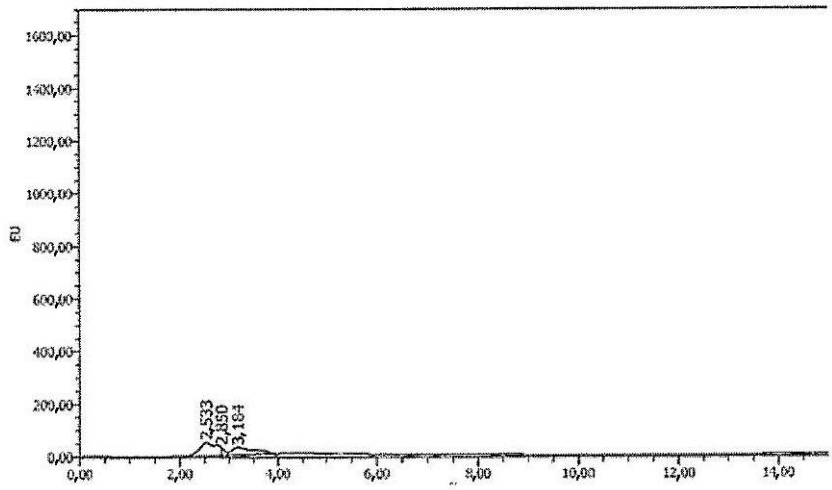

2,777
3,239
4,299
6,765
EU

Fig. 5

(A)


EU
2,524
2,850
3,241
9,247
0,00
2,00
4,00
6,00
8,00
10,00
12,00
14,00

(B)


EU
1600,00
1400,00
1200,00
1000,00
800,00
600,00
400,00
200,00
0,00
2,533
2,850
3,184
0,00
2,00
4,00
6,00
8,00
10,00
12,00
14,00

(C)

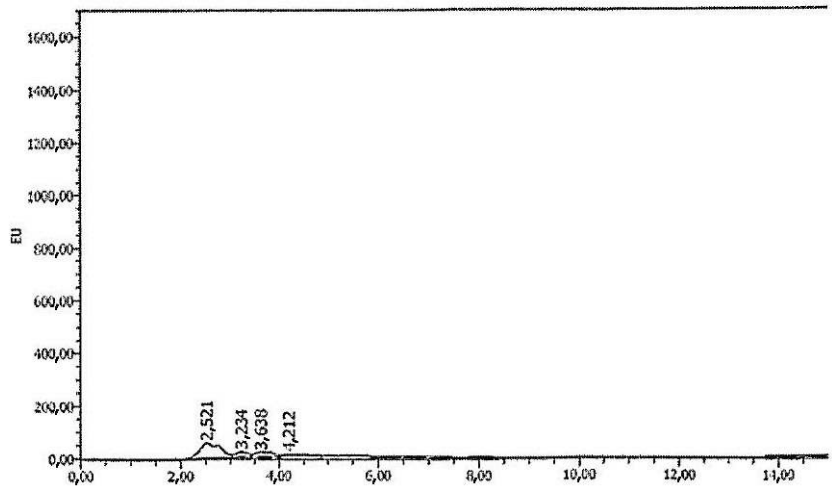
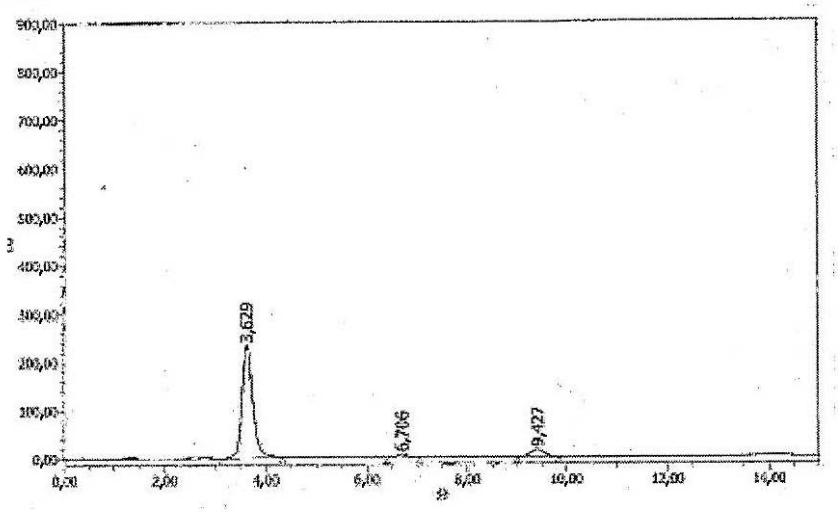

EU
1600,00
1400,00
1200,00
1000,00
800,00
600,00
400,00
200,00
0,00
2,521
3,234
3,638
4,212
0,00
2,00
4,00
6,00
8,00
10,00
12,00
14,00

Fig. 6

(A)



(B)

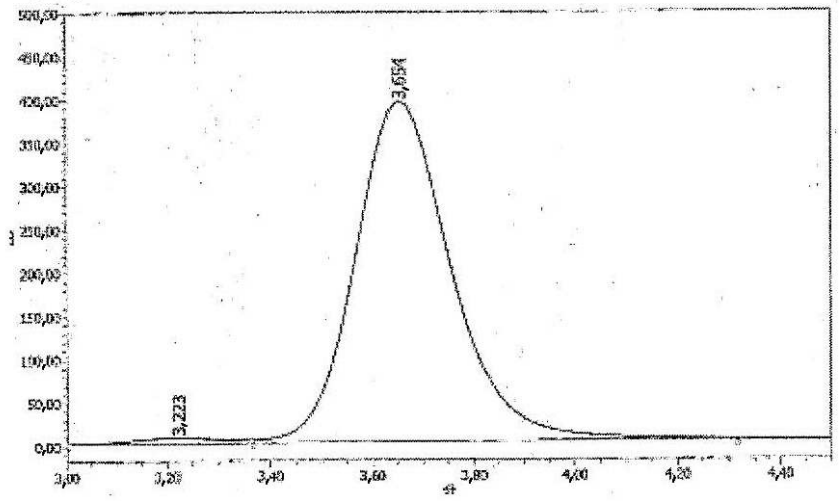
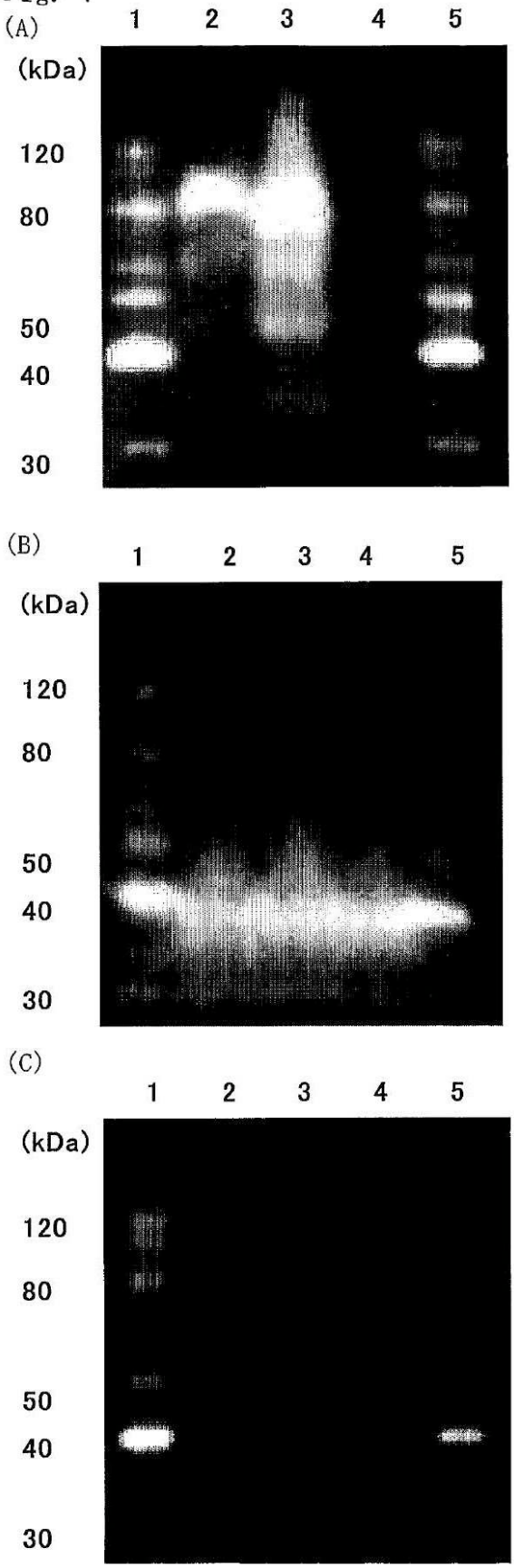


Fig. 7


(A)
1 2 3 4 5
(kDa)
120
80
50
40
30
(B)
1 2 3 4 5
(kDa)
120
80
50
40
30
(C)
1 2 3 4 5
(kDa)
120
80
50
40
30