



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 365 388**

51 Int. Cl.:

C12N 9/10 (2006.01)

C12P 19/18 (2006.01)

C12N 15/31 (2006.01)

C12N 15/54 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **02017764 .8**

96 Fecha de presentación : **09.08.2002**

97 Número de publicación de la solicitud: **1283259**

97 Fecha de publicación de la solicitud: **12.02.2003**

54 Título: **Condroitina-polimerasa y ADN que la codifica.**

30 Prioridad: **10.08.2001 JP 2001-244685**
22.10.2001 JP 2001-324127
05.04.2002 JP 2002-103136

45 Fecha de publicación de la mención BOPI:
03.10.2011

45 Fecha de la publicación del folleto de la patente:
03.10.2011

73 Titular/es: **SEIKAGAKU CORPORATION**
6-1, Marunouchi, 1-chome
Chiyoda-ku, Tokyo 100-0005, JP

72 Inventor/es: **Ninomiya, Toshio;**
Sugiura, Nobuo y
Kimata, Koji

74 Agente: **Curell Aguilá, Marcelino**

ES 2 365 388 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

Condroitina-polimerasa y ADN que la codifica.

5 Antecedentes de la presente invención

1. Campo de la presente invención

10 La presente invención se refiere a una nueva condroitina-polimerasa (condroitina-sintasa), a un ADN que la codifica, a un procedimiento para producir la condroitina-polimerasa, a un procedimiento para producir una cadena de azúcar que contiene la unidad de condroitina disacarídica repetitiva, a una sonda de hibridación para la condroitina-polimerasa y similares.

2. Breve descripción de la técnica anterior

15 En primer lugar, se describen las abreviaturas utilizadas frecuentemente en la presente memoria.

En las fórmulas y similares, "GlcUA", "GalNAc", "GlcNAc", "UDP" y "-" representan ácido D-glucurónico, N-acetil-D-galactosamina; N-acetil-D-glucosamina; uridina 5'-difosfato y un enlace glucosídico, respectivamente.

20 Condroitina es una cadena de azúcar que comprende una estructura disacarídica repetitiva del resto GlcUA y del resto GalNAc (-GlcUA β (1-3)-GalNAc β (1-4)); denominada también en lo sucesivo "estructura de condroitina") y una cadena de azúcar en la que la condroitina que está más sulfatada es el sulfato de condroitina.

25 En cuanto a una enzima que sintetiza condroitina procedente de un donante de GlcUA y de un donante de GalNAc transfiriendo alternativamente GlcUA y GalNAc a un aceptor (condroitina-polimerasa o condroitina-sintasa) y ADN que codifica la misma, solamente se conoce una condroitina-sintasa de *Pasteurella multocida* (*J. Biol. Chem.*, 215(31), 24124-24129 (2000)).

30 Además, determinadas cepas de *Escherichia coli* (serotipo 05:K4(L):H4 de *Escherichia coli*, denominado en lo sucesivo "cepa K4 de *Escherichia coli*") producen un polisacárido con estructura de condroitina, como antígeno capsular, pero su estructura es una estructura trisacarídica repetitiva en la que la fructosa está unida a una cadena lateral del resto GlcUA en la posición β 2-3. Por consiguiente, no estaba claro si la cepa K4 de *Escherichia coli* realmente tiene una condroitina-polimerasa como su propio sistema de síntesis del antígeno capsular.

Sumario de la invención

35 Un objetivo de la presente invención consiste en proporcionar una nueva condroitina-polimerasa, un ADN que codifica la misma, un procedimiento para producir la condroitina-polimerasa, un procedimiento para producir la cadena de azúcar que contiene la unidad de condroitina polisacarídica repetitiva, una sonda de hibridación para la condroitina-polimerasa y similares.

40 Este y otros objetivos de la presente invención se alcanzan mediante una nueva condroitina-polimerasa con las propiedades específicas descritas a continuación.

Breve descripción de los dibujos

45 La figura 1 presenta una cartografía de restricción de los clones del fago λ que contienen una parte de la zona R-I o de la zona R-III del grupo génico del antígeno K de la cepa K4 de *Escherichia coli*.

50 La figura 2 presenta el marco abierto de lectura (ORF) de la región R-II del grupo génico del antígeno K de la cepa K4 de *Escherichia coli*.

55 La figura 3 es un gráfico que presenta la transferencia de GalNAc al hexasacárido del sulfato de condroitina C por la enzima de la presente invención.

La figura 4 es un gráfico que presenta la transferencia de GalNAc al hexasacárido del sulfato de condroitina C por la enzima de la presente invención y las dimensiones de la cadena de azúcar producida.

60 La figura 5 es un gráfico que presenta la transferencia de cada monosacárido cuando se utilizó como donante UDP-GlcUA, UDP-GlcNAc o UDP-GalNAc, y se utilizó como aceptor hexasacárido o heptasacárido de sulfato de condroitina C.

65 La figura 6 es un gráfico que presenta la influencia de la temperatura sobre la actividad de la enzima de la presente invención.

La figura 7 es un gráfico que presenta los modelos de filtración en gel de los productos de la reacción enzimática después de varios periodos de reacción enzimática.

La figura 8 es un gráfico que presenta la relación entre el periodo de reacción enzimática y la cantidad de incorporación de radioactividad.

La figura 9 es un gráfico que presenta la relación entre la radioactividad incorporada (V) y la concentración en el sustrato de UDP-azúcar (S).

La figura 10 es un gráfico que presenta representaciones recíprocas dobles.

Descripción detallada de la invención

Los presentes inventores han llevado a cabo intensos estudios y han descubierto una nueva condroitina-polimerasa producida por microorganismos específicos (cepa 4 de *Escherichia coli* (serotipo 05:K4(L):H4 de *Escherichia coli*, ATCC 23502)), ADNc aislado que codifica la condroitina-polimerasa, y ha logrado preparar la condroitina-polimerasa utilizando el ADNc. De este modo, se ha realizado la presente invención.

Además, este y otros objetivos de la presente invención se han alcanzado mediante un procedimiento para producir la condroitina-polimerasa aislando el ADNc que codifica la condroitina-polimerasa y utilizando el ADNc. La expresión "síntesis de condroitina" tal como se utiliza en la presente memoria es un concepto que incluye el alargamiento de la cadena de azúcar de condroitina transfiriendo y agregando monosacáridos a una cadena de azúcar tal como condroitina. Por consiguiente, la reacción para alargar la cadena de azúcar de condroitina transfiriendo alternativamente y añadiendo los monosacáridos (GlcUA y GalNAc) a la cadena de azúcar para sintetizar la condroitina está incluida en un concepto de "síntesis de condroitina".

En la presente memoria se describe una condroitina-polimerasa (en lo sucesivo denominada también "enzima de la presente invención") que tiene las propiedades siguientes:

(1) Acción:

la polimerasa transfiere GlcUA y GalNAc alternativamente a un terminal no reducido de una cadena de azúcar de un donante de GlcUA y de un donante de GalNAc, respectivamente;

(2) Especificidad de sustrato:

la polimerasa transfiere GlcUA a un oligosacárido que tiene GalNAc en su terminal no reducido y una estructura de condroitina de un donante de GlcUA, pero no transfiere sustancialmente GalNAc al oligosacárido de un donante de GalNAc;

la polimerasa transfiere GalNAc a un oligosacárido que tiene GlcUA en su terminal no reducido y una estructura de condroitina de un donante de GalNAc, pero no transfiere sustancialmente GlcUA al oligosacárido de un donante de GlcUA;

(3) Influencia de los iones metálicos y similares:

La polimerasa actúa en presencia de ión Mn^{2+} pero no actúa sustancialmente en presencia de ión Ca^{2+} , ión Cu^{2+} o ácido etilendiamintetraacético.

La enzima descrita en la presente memoria procede preferentemente de *Escherichia coli*.

La presente invención se refiere a una proteína seleccionada de entre el grupo constituido por los siguientes apartados (A) y (B) (denominada en lo sucesivo "proteína de la presente invención"):

(A) una proteína que comprende la secuencia de aminoácidos representada por SEC. ID. nº 2;

(B) una proteína que comprende la secuencia de aminoácidos en la que 1 a 30 resto(s) de aminoácidos en la secuencia de aminoácidos representada por SEC. ID. nº 2 se eliminan, sustituyen, insertan o transponen, y que tiene actividad de condroitina-polimerasa.

La presente invención se refiere a un ADN que comprende uno de los siguientes apartados (a) a (c) (denominado en lo sucesivo "ADN de la presente invención"):

(a) ADN que codifica una proteína constituida por la secuencia de aminoácidos representada por SEC. ID. nº 2;

(b) ADN que codifica una proteína constituida por la secuencia de aminoácidos en la que 1 a 30 resto(s) de

aminoácidos en la secuencia de aminoácidos representada por SEC. ID. nº 2 se eliminan, sustituyen, insertan o transponen, y que tiene actividad de condroitina-polimerasa;

(c) ADN que se hibrida con

(i) el ADN en (a), o

(ii) el ADN complementario al ADN en (a) en condiciones severas en las que la hibridación se realiza a 42°C en una solución que contiene 50% de formamida, 4 x SSC, HEPES 50 mM (pH 7,0), 10 x solución de Denhardt y 100 µg/ml de ADN de esperma de salmón, y que codifica una proteína que tiene actividad de condroitina-polimerasa.

El ADN en el apartado (a) está representado preferentemente por SEC. ID. nº 1.

La presente invención se refiere a un vector que comprende el ADN de la presente invención (denominado en lo sucesivo "el vector de la presente invención").

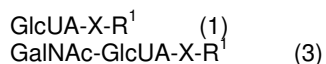
El vector de la presente invención es preferentemente un vector de expresión.

La presente invención se refiere a un transformante en el que se transforma un hospedador con el vector de la presente invención (en lo sucesivo denominado también "transformante de la presente invención").

La presente invención se refiere a un procedimiento para producir una condroitina-polimerasa, que comprende: cultivar el transformante de la presente invención; y recoger la condroitina-polimerasa del material cultivado (denominado en lo sucesivo "procedimiento de producción enzimática de la presente invención").

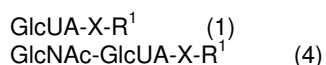
La presente invención se refiere a un agente para la síntesis de la cadena de azúcar, que comprende la proteína de la presente invención y que tiene actividad de condroitina-polimerasa.

La presente invención se refiere a un procedimiento para producir una cadena de azúcar representada por la fórmula (3) siguiente, que comprende por lo menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GalNAc y una cadena de azúcar representada por la fórmula (1) siguiente (denominado en lo sucesivo "procedimiento 1 de producción de la cadena de azúcar de la presente invención"):



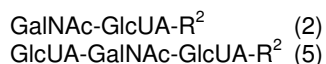
en las que X representa GalNAc o GlcNAc; R¹ representa un grupo cualquiera; y los demás símbolos tienen los mismos significados descritos anteriormente.

La presente invención se refiere a un procedimiento para producir una cadena de azúcar representada por la fórmula (4) siguiente, que comprende por lo menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GlcNAc y una cadena de azúcar representada por la fórmula (1) siguiente (denominado en lo sucesivo "procedimiento 2 de producción de la cadena de azúcar de la presente invención"):



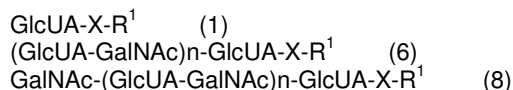
en las que todos los símbolos tienen los mismos significados descritos anteriormente.

La presente invención se refiere a un procedimiento para producir una cadena de azúcar representada por la fórmula (5) siguiente, que comprende por lo menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GlcUA y una cadena de azúcar representada por la fórmula (2) siguiente (denominado en lo sucesivo "procedimiento 3 de producción de la cadena de azúcar de la presente invención"):



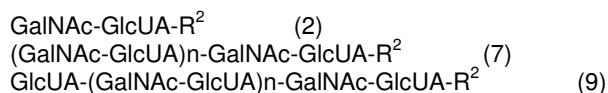
en las que R² representa un grupo cualquiera; y los demás símbolos tienen los mismos significados descritos anteriormente.

La presente invención se refiere a un procedimiento para producir una cadena de azúcar representada por la fórmula (6) y (8) siguientes, que comprende por lo menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GalNAc un donante de GlcUA y una cadena de azúcar representada por la fórmula (1) siguiente (denominado en lo sucesivo "procedimiento 4 de producción de la cadena de azúcar de la presente invención"):



en las que n es un número entero de 1 ó más, y los demás símbolos tienen los mismos significados descritos anteriormente.

La presente invención se refiere a un procedimiento para producir una cadena de azúcar representada por las fórmulas (7) y (9) siguientes, que comprende por lo menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GalNAc, un donante de GlcUA y una cadena de azúcar representada por la fórmula (2) siguiente (denominado en lo sucesivo "procedimiento 5 de producción de la cadena de azúcar de la presente invención"):



en las que todos los símbolos tienen los mismos significados descritos anteriormente.

La presente invención se refiere a una sonda de hibridación que comprende una secuencia nucleotídica complementaria a la secuencia nucleotídica representada por la SEC. ID. nº 1 (denominada también en los sucesivos "sonda de la presente invención").

La presente invención se refiere a un catalizador de transferencia de glucosilo (denominado en los sucesivos "catalizador de la presente invención") que comprende la proteína de la presente invención y que tiene actividad de condroitina-polimerasa.

La presente invención se explica a continuación con más detalle.

<1> Enzima de la presente invención y proteína de la presente invención.

La enzima descrita en la presente memoria es una condroitina-polimerasa que tiene las propiedades (1) a (3) siguientes.

(1) Acción:

La enzima descrita en la presente memoria transfiere GlcUA y GalNAc alternativamente a un terminal no reducido de una cadena de azúcar de un donante de GlcUA y de un donante de GalNAc, respectivamente.

Como donante de GlcUA, se prefiere un difosfato de nucleósido-GlcUA, y en particular se prefiere UDP-GlcUA. Además, como donante de GalNAc, se prefiere un difosfato de nucleósido-GalNAc, y en particular se prefiere UDP-GalNAc.

La enzima descrita en la presente memoria transfiere GlcUA y GalNAc alternativamente a un terminal no reducido de una cadena de azúcar (aceptor) de estos respectivos donantes de sacáridos. Por ejemplo, cuando GlcUA se transfiere en primer lugar a un terminal no reducido de una cadena de azúcar (aceptora), los monosacáridos se transfieren de tal manera que a continuación se transfiere GalNAc, a continuación se transfiere GlcUA, a continuación se transfiere GalNAc y así sucesivamente. De la misma manera, cuando GalNAc se transfiere en primer lugar a un terminal no reducido de una cadena de azúcar (aceptora), los monosacáridos se transfieren de tal manera que a continuación se transfiere GlcUA, a continuación se transfiere GalNAc, a continuación se transfiere GlcUA y así sucesivamente. Como resultado, una estructura de disacárido repetitiva de resto de GlcUA y de resto de GalNAc, es decir un esqueleto de condroitina, es sintetizada por la enzima de la presente invención.

Como cadena de azúcar que se convierte en aceptora de monosacáridos, es preferible una cadena de azúcar que tenga un esqueleto de condroitina. Como cadena de azúcar que tiene un esqueleto de condroitina, se pueden poner como ejemplos sulfato de condroitina y condroitina. Entre los sulfatos de condroitina, es preferible un sulfato de condroitina que está compuesto principalmente por un esqueleto de 6-sulfato de condroitina y también contiene una pequeña cantidad de una estructura de 4-sulfato-condroitina (denominada en lo sucesivo "sulfato C de condroitina").

Además, la cadena de azúcar que se convierte en un aceptor es más preferentemente un oligosacárido. El tamaño del oligosacárido no está específicamente restringido, pero cuando el aceptor es un oligosacárido de sulfato C de condroitina, es preferible el hexasacárido o heptasacárido, y el tetrasacárido o el hexasacárido es preferible cuando es un oligosacárido de condroitina.

Además, es preferible que la enzima descrita en la presente memoria sea capaz de transferir más GalNAc a una

cadena de azúcar que tiene una estructura de ácido hialurónico (una estructura de disacárido repetitiva del resto GlcUA y del resto GlcNAc) procedente de un donante de GalNAc. Es preferible que la cadena de azúcar que tiene una estructura de ácido hialurónico sea también un oligosacárido. El tamaño del oligosacárido no está particularmente restringido, pero los que están compuestos de aproximadamente hexasacáridos son particularmente preferibles.

(2) Especificidad de sustrato:

La enzima descrita en la presente memoria transfiere GlcUA a un oligosacárido que tiene GalNAc en su terminal no reducido y una estructura de condroitina procedente de un donante de GlcUA, pero no transfiere sustancialmente GalNAc al oligosacárido procedente de un donante de GalNAc.

La enzima descrita en la presente memoria transfiere GalNAc a un oligosacárido que tiene GlcUA en su terminal no reducida y un esqueleto de condroitina procedente de un donante de GalNAc, pero no transfiere sustancialmente GlcUA al oligosacárido procedente de un donante de GlcUA.

Es preferible que la enzima descrita en la presente memoria además no transfiera sustancialmente GlcNAc desde un donante de GlcNAc a un oligosacárido que tiene GalNAc en su terminal no reducido y además tiene un esqueleto de condroitina. Además, es preferible que la enzima descrita en la presente memoria transfiera más GlcNAc desde un donante de GlcNAc a un oligosacárido que tiene GlcUA en su terminal no reducido y también tiene un esqueleto de condroitina. Sin embargo, es preferible que GlcUA no sea transferido sustancialmente desde un donante de GlcUA a un oligosacárido producido por la transferencia de GlcNAc.

Además, es preferible que la enzima descrita en la presente memoria transfiera GlcNAc desde un donante de GlcNAc a un oligosacárido que tiene GlcUA en su terminal no reducido y que tiene también un esqueleto de ácido hialurónico, pero no transfiere sustancialmente GlcUA desde un donante de GlcUA.

(3) Influencia de los iones metálicos y similares:

La enzima descrita en la presente memoria actúa en presencia de ión Mn^{2+} pero no actúa sustancialmente en presencia de ión Ca^{2+} , ión Cu^{2+} o ácido etilendiamintetraacético.

Además, es preferible que la enzima descrita en la presente memoria actúe además en presencia del ión Fe^{2+} o Mg^{2+} . Por otra parte, es preferible que el grado de actuación (actividad enzimática) de la enzima descrita en la presente memoria en presencia de ión Mn^{2+} sea superior a su grado de actuación (actividad enzimática) en presencia del ión Fe^{2+} o Mg^{2+} .

Además, se observó que cuando una reacción se realiza utilizando la enzima descrita en la presente memoria a una temperatura de 25°C o más, el tamaño del producto de reacción (cadena de condroitina) se vuelve pequeño a medida que aumenta la temperatura de reacción (véanse, los ejemplos mostrados a continuación). Por consiguiente, se considera que la actividad enzimática de la enzima descrita en la presente memoria disminuye a medida que la temperatura de reacción aumenta a 25°C o más en las condiciones de reacción descritas en los siguientes ejemplos.

Es preferible que la enzima descrita en la presente memoria proceda de *Escherichia coli*. En particular, es preferible la cepa de *Escherichia coli* que tiene un gen relacionado con la producción del polisacárido capsular, y es más preferible la cepa de *Escherichia coli*, cuyo antígeno capsular (K) es "K4".

Como cepa de *Escherichia coli* cuyo serotipo del antígeno capsular es "K4", se puede citar preferentemente la cepa K4 de *Escherichia coli* (serotipo 05:K4(L):B4 de *Escherichia coli*), y más específicamente, se pueden citar preferentemente ATCC 23502, NCDC U1-41, número 2616 de la colección Freiburg y similares.

La enzima de la presente invención es una proteína seleccionada de entre los apartados (A) y (B) siguientes:

(A) una proteína que comprende la secuencia de aminoácidos representada por SEC. ID. nº 2;

(B) una proteína que comprende la secuencia de aminoácidos en la que 1 a 30 resto(s) de aminoácidos en la secuencia de aminoácidos representada por SEC. ID. nº 2 se eliminan, sustituyen, insertan o transponen, y que tiene actividad de condroitina-polimerasa.

Aunque la mutación tal como la sustitución, eliminación, inserción, transposición y similares puede suceder en las secuencias de aminoácidos de las proteínas existentes en la naturaleza producida por las reacciones de modificación dentro de las células o durante la purificación de proteínas después de su formación, además del polimorfismo y mutación de las moléculas de ADN que las codifican, es sabido que determinadas proteínas con dichas mutaciones presentan actividades fisiológicas y biológicas que son sustancialmente idénticas a las correspondientes proteínas que no tienen mutaciones. Dichas proteínas que presentan ligeras diferencias estructurales pero no diferencias significativas en sus funciones se describen también en la presente memoria. Es el

mismo caso en el que la mutación anterior se introduce artificialmente en la secuencia de aminoácidos de la proteína. En dicho caso, es posible preparar variedades mayores de mutantes. Por ejemplo, es sabido que un polipéptido preparado sustituyendo un resto de cisteína en la secuencia de aminoácidos de la interleucina 2 humana (IL-2) por un resto de serina mantiene la actividad de la interleucina 2 (*Science*, 224, 1431 (1984)). Además, es sabido que determinada proteína tiene una zona peptídica que no es esencial para su actividad. Por ejemplo, un péptido señal que existe en una proteína que se segrega en el resto extracelular y una pro-secuencia que puede encontrarse en un precursor de proteasa o similar corresponde a este caso, y la mayoría de estas zonas se eliminan después de la traducción de las proteínas o durante su conversión en proteínas activas. Dichas proteínas son proteínas que están presentes de diferentes formas desde el punto de vista de su estructura primaria pero finalmente tienen funciones similares.

La expresión "unos pocos restos de aminoácidos" tal como se utiliza en la presente memoria significa el número de aminoácidos que pueden producir mutación en tal grado que la actividad de la condroitina-polimerasa no se pierde, y en el caso de la proteína compuesta por 600 restos de aminoácidos por ejemplo, significa el número de aproximadamente 2 a 30, preferentemente de 2 a 15, y más preferentemente de 2 a 8.

Además, la proteína de la presente invención puede contener una secuencia de aminoácidos de otra proteína o péptido, con tal de que contenga la secuencia de aminoácidos del apartado (A) o (B) anterior. Es decir, la proteína de la presente invención puede ser una proteína de fusión con otra proteína o péptido.

Por ejemplo, las proteínas de fusión de una proteína que comprende la secuencia de aminoácidos descrita en el apartado (A) o (B) anterior con un péptido marcador y similar están también incluidas en la proteína de la presente invención. Dichas proteínas de la presente invención tienen la ventaja de que su purificación puede realizarse fácilmente. Ejemplos del péptido marcador anterior incluyen la proteína A, la secuencia señal de la insulina, His, FLAG, CBP (proteína de unión a la calmodulina), GST (glutación S-transferasa) y similares. Por ejemplo, su proteína de fusión con la proteína A puede purificarse convenientemente por cromatografía de afinidad utilizando una fase sólida unida a IgG. De la misma manera, una fase sólida a la que el níquel magnético se une puede utilizarse para una proteína de fusión con la etiqueta His, y una fase sólida a la que un anticuerpo anti-FLAG está unido puede utilizarse para una proteína de fusión con FLAG. Además, ya que una proteína de fusión con señal de insulina se segrega en el resto extracelular (un medio o similar), las operaciones de extracción tales como la disgregación celular y similares se vuelven innecesarias. Es preferible que la proteína de la presente invención (la enzima de la presente invención) sea soluble.

Los ejemplos preferidos incluyen una proteína de fusión con un péptido (etiqueta His) representado por la secuencia de aminoácidos representada por la SEC. ID. nº 11. Es preferible realizar la fusión de esta etiqueta de His continuamente en una posición justo antes de la secuencia de aminoácidos representada por SEC. ID. nº 2. La proteína de fusión puede producirse expresando un ADN en el que la secuencia nucleotídica representada por la SEC. ID. nº 4 está conectada continuamente a una posición justo antes de la secuencia nucleotídica representada por la SEC. ID. nº 1. La proteína de fusión es soluble.

La "actividad de la condroitina-polimerasa" puede detectarse según un método analítico de glucosiltransferasa utilizado generalmente. Específicamente, puede detectarse como actividad para sintetizar transfiriendo GlcUA y GalNAc alternativamente a un terminal no reducido de una cadena de azúcar (aceptor), utilizando un donante de GlcUA, un donante de GalNAc y una cadena de azúcar que se convierte en el aceptor.

Por ejemplo, cuando GlcUA se transfiere desde un donante de GlcUA a una cadena de azúcar que tiene GalNAc en su terminal reducido, se transfiere desde un donante de GalNAc a una cadena de azúcar que tiene un GlcUA en su terminal no reducido, puede interpretarse que tiene actividad para transferir GlcUA y GalNAc alternativamente a una terminal no reducido de la cadena de azúcar, a saber actividad de condroitina-polimerasa. Es preferible emplear el método de medición de actividad enzimática descrito en los Ejemplos siguientes. Mediante dicho método puede seleccionarse fácilmente, la eliminación, sustitución, inserción o transposición de aminoácidos que conservan la actividad de condroitina-polimerasa.

Los procedimientos para la producción de la enzima de la presente invención y de la proteína de la presente invención no están particularmente restringidos, y pueden producirse expresando el ADN de la presente invención descrito a continuación en células apropiadas. Además, las que se aíslan de los recursos naturales y se producen por síntesis química y similares están incluidas en la enzima de la presente invención y en la proteína de la presente invención como algo corriente. Los procedimientos de producción de la enzima de la presente invención (la proteína de la presente invención) que utilizan en ADN de la presente invención se describirán a continuación.

<2> ADN de la presente invención

(a) ADN que codifica una proteína constituida por la secuencia de aminoácidos representada por SEC. ID. nº 2;

(b) ADN que codifica una proteína constituida por la secuencia de aminoácidos en la que 1 a 30 resto(s) de aminoácidos en la secuencia de aminoácidos representada por SEC. ID. nº 2 se eliminan, sustituyen, insertan o

transponen, y que tiene actividad de condroitina-polimerasa;

(c) ADN que se hibrida con

(i) el ADN en (a), o

(ii) el ADN complementario al ADN en (a) en condiciones severas en las que la hibridación se realiza a 42°C en una solución que contiene 50% de formamida, 4 x SSC, HEPES 50 mM (pH 7,0), 10 x solución de Denhardt y 100 µg/ml de ADN de esperma de salmón, y que codifica una proteína que tiene actividad de condroitina-polimerasa.

El ADN está representado preferentemente por la SEC. ID. nº 1.

Las "condiciones severas" tal como se utiliza en la presente memoria significa las condiciones bajo las cuales se forma un denominado híbrido específico pero no se forma un híbrido no específico (véase Sambrook, J. *et al.*, Molecular Cloning, A Laboratory Manual, segunda edición, Cold Spring Harbor Laboratory Press (1989) y similares). Ejemplo de "condiciones severas" incluyen las condiciones en las que se lleva a cabo la hibridación a 42°C en una solución que contiene formamida al 50%, 4 x SSC, HEPES 50 mM (pH 7,0), 10 x solución de Denhardt y 100 µg/ml de ADN de esperma de salmón, y se lava el producto con 2 x SSC, solución SDS al 0,1% a temperatura ambiente y a continuación con 0,1 x SSC, solución SDS al 0,1% a 50°C.

El ADN de la presente invención se obtiene generalmente a partir de la cepa de *Escherichia coli* que tiene el antígeno K4, pero las muestras de ADN obtenidas de otras especies de organismo transformadas y producidas por síntesis química y similares están también incluidas en la presente memoria como algo corriente. Los procedimientos para la producción del ADN de la presente invención no están tampoco específicamente limitados, pero es preferible utilizar, por ejemplo, el procedimiento descrito en los ejemplos siguientes.

Los expertos en la materia entienden fácilmente que los ADN que tienen varias secuencias nucleotídicas diferentes debido a la degeneración del código genético están presentes como ADN de la presente invención.

<3> Vector de la presente invención

El vector de la presente invención es un vector que comprende el ADN de la presente invención. Preferentemente el ADN en el vector de la presente invención es el mismo descrito en el apartado <2> anterior. Además, ya que el vector de la presente invención se utiliza preferentemente en el proceso de producción de enzimas de la presente invención que se describirá después, es preferentemente un vector de expresión.

El vector de la presente invención puede prepararse insertando el ADN de la presente invención en un vector conocido.

Como vector en el que se inserta el ADN de la presente invención puede utilizarse, por ejemplo, un vector apropiado que puede expresar el ADN introducido (un vector de fago, un vector de plásmido o similar) y puede seleccionarse opcionalmente en respuesta a cada célula hospedadora en la que el vector de la presente invención está insertado. Ejemplos del sistema de vector hospedador incluyen una combinación de una célula de mamífero tal como célula COS, 3LL-HK46 o similar con un vector de expresión para célula de mamífero tal como pGIR201 (Kitagawa, H. y Paulson, J. C., *J. Biol. Chem.*, 269, 1394-1401 (1994)), pEF-BOS (Mizushima, S. y Nagata, S., *Nucleic Acid Res.*, 18, 5322 (1990)), pCXN2 (Niwa, H., Yamamura, K. y Miyazaki, J., *Gene*, 108, 193-200 (1991)), pCMV-2 (preparado por Eastman Kodak), pCEV18, pME18S (Maruyama *et al.*, *Med. Immunol.*, 20, 27 (1990)), pSVL (preparado por Pharmacia Biotech) o similares y una combinación de *Escherichia coli* con un vector de expresión para células procarióticas tal como pTrcHis (fabricado por Invitrogen), pGEX (preparado por Pharmacia Biotech), pTrc99 (preparado por Pharmacia Biotech), pKK233-3 (preparado por Pharmacia Biotech), pEZZZ18 (preparado por Pharmacia Biotech), pCH110 (preparado por Pharmacia Biotech), pET (elaborado por Stratagene), pBAD (elaborado por Invitrogen), pRSET (elaborado por Invitrogen), pSE420 (elaborado por Invitrogen) o similares. Además, una célula de insecto, levadura, *Bacillus subtilis* y similares pueden servir como ejemplo también como la célula hospedadora y varios vectores correspondientes a la misma. Entre los sistemas de vector para hospedador anteriores, es preferible una combinación de *Escherichia coli* con pTrcHis.

Además, a medida que el vector en el que se inserta el ADN de la presente invención, un vector construido de tal manera que expresa una proteína de fusión de la proteína de la presente invención (enzima de la presente invención) con un péptido marcador, puede utilizarse también, el cual, como se describió en el apartado <1> anterior, es particularmente preferible cuando se purifica la condroitina-polimerasa expresada utilizando el vector de la presente invención. Específicamente, es preferible un vector que comprende una secuencia nucleotídica que expresa His (por ejemplo, la secuencia nucleotídica representada por la SEC. ID. nº 4).

Cuando se utiliza alguno de los vectores anteriores, el ADN de la presente invención puede estar conectado con el vector después de tratar a ambos con enzimas de restricción y similares y opcionalmente llevando a cabo la

terminación suave y la conexión de un terminal cohesivo de modo que la conexión del ADN de la presente invención con el vector se hace posible.

Como procedimiento de producción del vector de la presente invención puede utilizarse y es preferible, por ejemplo, el procedimiento descrito en los siguientes ejemplos.

<4> Transformante de la presente invención

El transformante de la presente invención es un transformante en el que un hospedador se transforma con el vector de la presente invención.

El "hospedador" tal como se utiliza en la presente memoria puede ser cualquier hospedador en el que puede realizarse la recombinación por el vector de la presente invención pero es preferible uno que pueda ejercer la función del ADN de la presente invención o un vector recombinante en el que se inserta el ADN de la presente invención. Ejemplos de hospedador incluyen las células animales, células vegetales y células microbianas, y están incluidas las células de mamíferos tales como las células COS (célula COS-1, células CO-7 y similares), célula 3LL-HK46, etc., *Escherichia coli*, células de insecto, levaduras, *Bacillus subtilis* y similares. El hospedador puede seleccionarse opcionalmente en respuesta a cada vector de la presente invención, pero, por ejemplo, cuando se utiliza un vector de la presente invención preparado basándose en pTrcHis, es preferible seleccionar la cepa de *Escherichia coli*.

El hospedador puede ser transformado por el vector de la presente invención de la manera habitual. Por ejemplo, el hospedador puede transformarse introduciendo el vector de la presente invención en el hospedador por un procedimiento que utiliza un reactivo de transfección disponible en el mercado, un procedimiento con DEAE-dextrano, electroporación o similares.

El transformante de la presente invención obtenido de esta manera puede utilizarse en el proceso de producción enzimática de la presente invención descrito a continuación.

<5> Proceso de producción enzimática de la presente invención

El proceso de producción enzimática de la presente invención es un proceso para producir una condroitina-polimerasa, que comprende cultivar el transformante de la presente invención; y recolectar una condroitina-polimerasa del material de cultivo.

El término "cultivo" tal como se utiliza en la presente memoria significa una idea general que incluye el cultivo de células o de microorganismos como el propio transformante de la presente invención y el cultivo de un animal, insecto o similar en el que se incorporan las células como el transformante de la presente invención. Además, la expresión "material de cultivo" tal como se utiliza en la presente memoria significa un concepto que incluye un medio (sobrenadante del medio de cultivo) y células hospedadoras cultivadas después del cultivo del transformante de la presente invención, la materia segregada, la materia excretada y similares.

Las condiciones de cultivo (medio, condición de cultivo y similares) se seleccionan apropiadamente basándose en el hospedador que va a utilizarse.

Según el proceso de producción enzimática de la presente invención, pueden producirse varias formas de condroitina-polimerasa, basándose en el transformante que va a utilizarse.

Por ejemplo, cuando se cultiva un transformante transformado con un vector de expresión construido para expresar una proteína de fusión con un péptido marcador, se produce una condroitina-polimerasa fusionada con el péptido marcador. Específicamente, por ejemplo, una condroitina-polimerasa fusionada con la etiqueta His se produce para cultivar un transformante transformado con un vector de expresión construido para efectuar la expresión de una proteína en la que la secuencia de aminoácidos representada por SEC. ID. nº 11 se fusiona continuamente a una posición justo antes de la secuencia de aminoácidos representada por SEC. ID. nº 2. Particularmente, es preferible utilizar un transformante transformado con un vector de expresión construido conectando la secuencia nucleotídica representada por SEC. ID. nº 4 continuamente a una posición justo antes de la secuencia nucleotídica representada por SEC. ID. nº 1.

La condroitina-polimerasa puede recogerse del material cultivado por procedimientos de extracción y purificación de proteínas basados en la forma de la condroitina-polimerasa producida.

Por ejemplo, cuando se produce la condroitina-polimerasa en forma soluble segregada en un medio (sobrenadante del medio de cultivo), el medio puede recogerse y utilizarse directamente como condroitina-polimerasa. Además, cuando la condroitina-polimerasa se produce en forma soluble segregada en el citoplasma o en forma insoluble (fijación a la membrana), la condroitina-polimerasa puede extraerse por disgregación celular tal como un procedimiento que utiliza un aparato de cavitación de nitrógeno, homogeneización, molino de bolas de vidrio,

tratamiento con ultrasonidos, un método de choque osmótico, congelación-descongelación, etc., extracción con tensioactivo, una combinación de los mismos o similares, y el extracto puede utilizarse directamente como condroitina-polimerasa.

5 La condroitina-polimerasa puede purificarse más en el medio o extractos, lo cual es preferible. La purificación puede ser purificación imperfecta (purificación parcial) o purificación perfecta, que puede seleccionarse de manera apropiada basándose, por ejemplo, en el objeto que utiliza la condroitina-polimerasa, y similares.

10 Ejemplos del procedimiento de purificación incluyen el lavado con sulfato amónico, sulfato sódico, etc., centrifugación, diálisis, ultrafiltración, cromatografía de adsorción, cromatografía de intercambio iónico, cromatografía hidrófoba, cromatografía en fase inversa, filtración en gel, cromatografía de penetración en gel, cromatografía por afinidad, electroforesis, una combinación de los mismos y similares.

15 La producción de la condroitina-polimerasa puede confirmarse analizando su secuencia de aminoácidos, acciones, especificidad del sustrato y similares.

<6> Agente sintetizador de la presente invención

20 El agente sintetizador de la presente invención es un agente sintetizador de la cadena de azúcar, que comprende la proteína de la presente invención y que tiene actividad de condroitina-polimerasa.

25 El agente sintetizador de la presente invención es el resultado de aplicar la "acción de transferencia de GalNAc", "acción de transferencia de GlcNAc" y de "acción de transferencia de GlcUA" de la enzima de la presente invención y la proteína de la presente invención como agente sintetizador de la cadena de azúcar.

30 El agente sintetizador de la presente invención se utiliza para la síntesis de las cadenas de azúcar. La expresión "síntesis de las cadenas de azúcar", tal como se utiliza en la presente memoria significa un concepto que incluye el alargamiento de una determinada cadena de azúcar transfiriendo y añadiendo un monosacárido a dicha cadena. Por ejemplo, un concepto de transferencia y adición de un monosacárido tal como GlcUA, GalNAc, GlcNAc o similares a una cadena de azúcar tal como condroitina, sulfato de condroitina, ácido hialurónico y similares para alargar la cadena de azúcar está incluido en la expresión "síntesis de la cadena de azúcar" tal como se utiliza en la presente memoria.

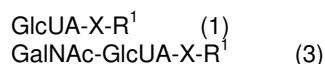
35 La forma del agente sintetizador de la presente invención no está limitada, y puede ser cualquiera de entre una forma de solución, una forma congelada, una forma liofilizada y una forma de enzima inmovilizada en la que está unida a un vehículo. Además, puede contener otros compuestos (por ejemplo, vehículo farmacéuticamente aceptable, vehículo que es aceptable para el reactivo, etc.), con tal de que no tenga influencia sobre la actividad de la condroitina-polimerasa.

40 <7> Procedimiento de producción de la cadena de azúcar de la presente invención

45 El procedimiento de producción de la cadena de azúcar de la presente invención utiliza el agente sintetizador de la presente invención y se divide en los cinco métodos siguientes en respuesta a los sustratos donantes y aceptores del azúcar.

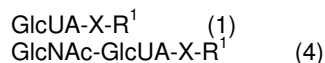
(1) Procedimiento de producción de la cadena de azúcar de la presente invención 1

50 Procedimiento para producir una cadena de azúcar representada por la fórmula (3) siguiente, que comprende por lo menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GalNAc y una cadena de azúcar representada por la fórmula (1) siguiente :



55 (2) Procedimiento de producción de la cadena de azúcar de la presente invención 2

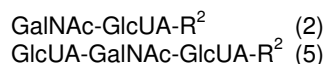
60 Procedimiento para producir una cadena de azúcar representada por la fórmula (4) siguiente, que comprende por lo menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GlcNAc y una cadena de azúcar representada por la fórmula (1) siguiente:



65 (3) Procedimiento de producción de la cadena de azúcar de la presente invención 3

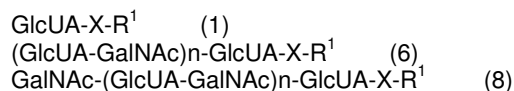
Procedimiento para producir una cadena de azúcar representada por la fórmula (5) siguiente, que comprende por lo

menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GlcUA y una cadena de azúcar representada por la fórmula (2) siguiente:



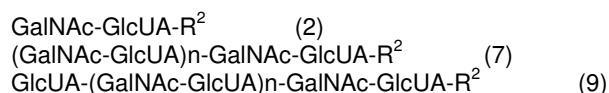
(4) Procedimiento de producción de la cadena de azúcar de la presente invención 4

Procedimiento para producir una cadena de azúcar representada por la fórmula (6) y (8) siguientes, que comprende por lo menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GalNAc y un donante de GlcUA y una cadena de azúcar representada por la fórmula (1) siguiente:



(5) Procedimiento de producción de la cadena de azúcar de la presente invención 5

Procedimiento para producir una cadena de azúcar representada por las fórmulas (7) y (9) siguientes, que comprende por lo menos una etapa que permite al agente sintetizador de la presente invención ponerse en contacto con un donante de GalNAc, un donante de GlcUA y una cadena de azúcar representada por la fórmula (2) siguiente:



En cada una de las fórmulas, X representa GalNAc o GlcNAc, y R^1 y R^2 representa cada uno cualquier grupo. R^1 y R^2 son iguales o diferentes entre sí.

Los ejemplos de R^1 y R^2 incluyen una cadena de azúcar que tiene un esqueleto de condroitina, una cadena de azúcar que tiene un esqueleto de ácido hialurónico y similares.

La cadena de azúcar representada por la fórmula (1) es preferentemente sulfato de condroitina (específicamente sulfato de condroitina C), condroitina o ácido hialurónico que tiene GlcUA en su terminal no reducido, o un oligosacárido de los mismos.

La cadena de azúcar representada por la fórmula (2) es preferentemente sulfato de condroitina (específicamente sulfato de condroitina C) o condroitina que tiene GalNAc en su terminal no reducido, o un oligosacárido de los mismos.

Como donante de GalNAc, es preferible un nucleósido difosfato de GalNAc, y en particular es preferible UDP-GalNAc. Además, como donante de GalNAc, es preferible un nucleósido difosfato de GlcNAc, y en particular es preferible UDP-GlcNAc. Por otra parte, como donante de GlcUA, es preferible el nucleósido difosfato de GlcUA, y en particular es preferible UDP-GlcUA.

El procedimiento para realizar el "contacto" no está particularmente limitado, siempre que la reacción enzimática se genere por contacto mutuo de las moléculas de la enzima de la presente invención (o la proteína de la presente invención) contenida en el agente sintetizador de la presente invención, un donante y un aceptor (cadena de azúcar). Por ejemplo, el contacto puede realizarse en una solución en la que se disuelven los tres componentes. Además, la reacción enzimática puede realizarse continuamente utilizando una enzima inmovilizada en la que la condroitina-polimerasa contenida en el agente sintetizador de la presente invención está unido a una fase sólida apropiada (perlas o similares) o a un reactor de tipo membrana que utiliza membrana de ultrafiltración, membrana de diálisis o similares. Además, la reacción enzimática puede realizarse conectando el aceptor a una fase sólida similar al método descrito en el documento WO 00/27437. Además, un birreactor que regenera (sintetiza) el donante puede utilizarse en combinación.

Además, en los procedimientos (4) y (5) anteriores, no es siempre necesario poner en contacto el donante GalNAc y el donante GlcUA simultáneamente con el agente sintetizador de la presente invención y la cadena de azúcar representadas por las fórmulas (1) ó (2), y los donantes pueden entrar en contacto alternativamente.

Las condiciones para llevar a cabo la reacción enzimática no están específicamente limitadas, siempre y cuando sean condiciones bajo las cuales la enzima de la presente invención (o la proteína de la presente invención) pueden funcionar, pero es preferible llevar a cabo la reacción a pH aproximadamente neutro (por ejemplo, aproximadamente pH 7,0 a 7,5) y es más preferible llevar a cabo la reacción en un tampón con una acción tamponante bajo el pH. Además, la temperatura en este caso no está específicamente limitada, siempre y cuando la actividad de la enzima de la presente invención (o la proteína de la presente invención) se mantenga, pero una temperatura de

aproximadamente 25°C a 30°C puede servir como ejemplo. Además, cuando existe una sustancia que aumenta la actividad de la enzima de la presente invención (o la proteína de la presente invención), puede añadirse la sustancia. Por ejemplo, es preferible dejar que coexista Mn^{2+} y similares. La duración de la reacción puede ser determinada opcionalmente por los expertos en la materia en respuesta a las cantidades del agente sintetizador de la presente invención, del donante y del aceptor que va a utilizarse y de otras condiciones de la reacción.

El aislamiento y similares de la cadena de azúcar procedente del producto formado puede llevarse a cabo por procedimientos conocidos.

Además, puede producirse un sacárido sulfatado tal como sulfato de condroitina o similar utilizando el agente sintetizador de la presente invención (condroitin-polimerasa) y una sulfato transferasa en combinación.

Por ejemplo, puede producirse un sacárido sulfatado tal como sulfato de condroitina o similar llevando a cabo simultáneamente la formación de condroitina y la transferencia del grupo sulfato en el proceso de producción de la cadena de azúcar anterior, permitiendo además que coexista un donante del grupo sulfato (5'-fosfosulfato de 3'-fosfoadenosina (PAPS) o similar) y una sulfotransferasa. La sulfotransferasa puede utilizarse como enzima inmovilizada uniéndola a una fase sólida apropiada (perlas o similar) similar al caso anterior o se permite reaccionar continuamente utilizando un reactor de tipo membrana que utiliza membrana de ultrafiltración, membrana de diálisis o similar. En este caso, puede utilizarse en combinación un biorreactor que regenera (sintetiza) el donante del grupo sulfato.

La sulfotransferasa que puede utilizarse en la presente memoria puede ser cualquier enzima que pueda transferir un grupo sulfato a la condroitina y puede seleccionarse apropiadamente a partir de las enzimas conocidas basándose en el tipo de sulfato de condroitina deseado. Además, dos o más tipos de sulfotransferasas que tienen posiciones diferentes de transferencia del grupo sulfato pueden utilizarse en combinación.

La condroitina 6-O-sulfotransferasa (*J. Biol. Chem.*, 215 (28), 21075-21080 (2000)) puede servir como ejemplo como sulfotransferasa. Sin embargo, no hay ninguna limitación a ésta y pueden utilizarse también otras enzimas.

<8> Sonda de la presente invención

La sonda de la presente invención es una sonda de hibridación que comprende una secuencia nucleotídica complementaria con la secuencia nucleotídica representada por la SEC. ID. nº 1.

La sonda de la presente invención puede obtenerse preparando un oligonucleótido que comprende una secuencia nucleotídica complementaria con la secuencia nucleotídica representada por la SEC. ID. nº 1, y marcándola con un marcador adecuado para hibridación (por ejemplo, radioisótopo).

El tamaño del oligonucleótido se selecciona apropiadamente basándose en las condiciones y similares de la hibridación que utiliza la sonda de la presente invención.

Es de esperar que la sonda de la presente invención llegue a ser una herramienta útil para examinar las funciones biológicas del sulfato de condroitina. Esto es debido a que el sulfato de condroitina se expresa ampliamente y desempeña una función importante en un gran número de tejidos, particularmente en el cerebro. Se considera que la sonda sirve también para la búsqueda de la relación entre genes y enfermedades.

<9> Catalizador de la presente invención

El catalizador de la presente invención es un catalizador de transferencia de glucosilo que comprende la proteína de la presente invención y que tiene una actividad de condroitina polimerasa.

El catalizador de la presente invención es un resultado de aplicar la "acción de transferencia de GalNAc", "la acción de transferencia de GlcNAc", y "acción de transferencia de GlcUA" de la enzima de la presente invención y de la proteína de la presente invención como catalizador de la transferencia de glucosilo.

El catalizador de la presente invención puede utilizarse para la transferencia de GlcUA, GalNAc o GlcNAc. Por ejemplo, puede utilizarse para transferir un monosacárido, tal como GlcUA, GalNAc o GlcNAc o similar a un terminal no reducido de una cadena de azúcar tal como condroitina, sulfato de condroitina, ácido hialurónico y similares.

La forma del catalizador de la presente invención no está restringida, y puede ser cualquiera de una forma en solución, una forma congelada, una forma liofilizada y una forma de enzima inmovilizada en la que está unido a un vehículo. Además, puede contener otros componentes (por ejemplo, vehículo farmacéuticamente aceptable, vehículo que es aceptable para reactivo, etc.), siempre y cuando no tenga influencia sobre la actividad de transferencia de GlcUA, GalNAc o GlcNAc.

Dado que la enzima de la presente invención y la proteína de la presente invención pueden transferir GlcUA y

GalNAc alternativamente, como una molécula única, son notablemente útiles como herramientas para la producción a gran escala de la cadena de azúcar que tiene un esqueleto de condroitina (condroitina, sulfato de condroitina, y similares), como ingredientes activos del agente sintetizador de la presente invención y el catalizador de la presente invención, reactivos o similares. Además, el ADN de la presente invención es muy útil como herramienta para la producción a gran escala de dicha enzima de la presente invención y de la proteína de la presente invención. El vector de la presente invención es muy útil, porque puede retener el ADN de la presente invención de manera estable y funcionar con eficacia y eficiencia. El transformante de la presente invención es además muy útil, porque no solamente puede retener el vector de la presente invención de manera estable y funcionar eficaz y eficientemente, sino que también puede utilizarse como tal para la producción a gran escala de la enzima de la presente invención y de la proteína de la presente invención. Además, el proceso de producción de la enzima de la presente invención es muy útil para la producción a gran escala de la enzima de la presente invención y de la proteína de la presente invención. También, el proceso de producción de la cadena de azúcar de la presente invención es muy útil para la producción a gran escala de la cadena de azúcar que tiene el esqueleto de condroitina (condroitina, sulfato de condroitina, y similares). El agente sintetizador de la presente invención y el catalizador de la presente invención son muy útiles, porque pueden utilizarse en el proceso de producción de la cadena de azúcar de la presente invención.

Ya que la alta calidad y la condroitina-polimerasa uniforme pueden producirse de manera conveniente y a gran escala mediante la presente invención, pueden proporcionarse productos a bajo coste para el campo industrial y por consiguiente la presente invención tiene muy alta disponibilidad.

La presente invención se explica con detalle basándose en los ejemplos, sin embargo, la presente invención no se limita a los mismos.

UDP-[¹⁴C]GlcUA, UDP-[³H]GalNAc y UDP-[¹⁴C]GlcNAc utilizados en los ejemplos se adquirieron en NEN Life Sciences. Además, UDP-GlcUA, UDP-GalNAc y UDP-GlcNAc se adquirieron en Sigma.

Ejemplo 1

Clonación del gen de condroitina-polimerasa:

(1) Preparación del banco de ADN

La cepa K4 (serotipo O5:K4(L):H4, ATCC 23502) de *Escherichia coli*, se cultivó a 37°C durante la noche en 50 ml de medio LB. Se recogieron las células por centrifugación (3.800 rpm, 15 minutos), se pusieron suspensión en suspensión en 9 ml de tampón Tris-HCl 10 mM (pH 8,0) que contiene ácido etilendiamintetraacético 1 mM (EDTA) (denominado en lo sucesivo "TE") y a continuación se trata a 37°C durante 1 hora, añadiendo 0,5 ml de SDS al 10% y 50 µl de proteinasa K (20 mg/ml, Boehringer Mannheim). A la suspensión, se añadieron 10 ml de solución PCI (fenol:cloroformo:alcohol isoamílico = 25:24:1), seguido de agitación durante 30 minutos, y la mezcla resultante se centrifugó (3.800 rpm, 15 minutos) para recoger la capa de agua y la materia insoluble de la capa intermedia y se centrifugó (10.000 rpm, 30 minutos) otra vez. Se recuperó el sobrenadante y se añadieron 50 µl de ARNasa A (20 mg/ml, Sigma) a ésta para la reacción a 37°C durante 1 hora. A la solución tratada se añadieron 10 ml de solución PCI, seguido de agitación durante 30 minutos, y la mezcla resultante se centrifugó (3.800 rpm, 15 minutos), para recoger la capa de agua y se centrifugó (10.000 rpm, 30 minutos) otra vez. Se recuperó el sobrenadante y se dializó frente a 2.000 ml de TE a 4°C durante la noche, y la solución (7,5 ml) dializada de este modo se utilizó como solución de ADN cromosómico (concentración de ADN, 0,9 mg/ml). La solución de ADN cromosómico (120 µl) procedente de la cepa K4 obtenido de este modo se digirió utilizando una enzima de restricción *Sau3A1* (4 unidades: NEB) a 37°C durante 30 minutos y a continuación se sometió a electroforesis en gel de agarosa al 0,3%, y después se cortó el gel de agarosa correspondiente al ADN de aproximadamente 7 a 11 kpb. El gel cortado de este modo se colocó en un tubo de 1,5 ml de capacidad con un orificio en su fondo picado con una aguja y, junto con el tubo, se insertó en un tubo de 2 ml de capacidad y se centrifugó (8.000 rpm, 1 minuto) para romper el gel. El fenol neutralizado en casi el mismo volumen de gel se añadió a éste, seguido de agitación intensa y a continuación la mezcla resultante se congeló a -80°C. Treinta minutos después, la temperatura se volvió a la temperatura ambiente y se fundió la mezcla, seguido de centrifugación (13.000 rpm, 5 minutos). La capa acuosa resultante se recogió, se añadió a ésta el mismo volumen de solución de PCI seguido de agitación, y a continuación se centrifugó la mezcla (13.000 rpm, 5 minutos). Se recogió la capa acuosa, 1/10 de volumen de solución de acetato sódico 3 M y el mismo volumen de 2-propanol se añadieron a esto para precipitar el ADN, y el precipitado se recogió a continuación por centrifugación (13.000 rpm, 30 minutos). Al precipitado recogido de este modo, se añadió solución de etanol al 70%, seguido de centrifugación (13.000 rpm, 5 minutos), y a continuación el precipitado resultante se disolvió añadiendo 100 µl de TE. Para concentrar la solución resultante, se precipitó el ADN añadiendo 10 µl de solución de acetato sódico 3 M y 300 µl de etanol y a continuación se recuperó por centrifugación (13.000 rpm, 20 minutos). Al precipitado recogido de este modo, se añadió solución de etanol al 70%, seguido de centrifugación (13.000 rpm, 5 minutos), y el precipitado resultante se disolvió en 4 µl de agua purificada para obtener una solución de fragmento de ADN cromosómico. La solución del fragmento de ADN (2 l) se insertó en un vector de fago λ (λ EMBL3: STRATAGENE) que se había tratado con enzimas de restricción (*Bam*Hi (80 unidades: NEB) y *Eco*RI (80 unidades: NEB)) y se sometió a embalaje utilizando un kit de embalaje (Gigapack III Gold Packaging Extract, STRATAGENE)

según las instrucciones del fabricante, y a continuación la cepa de *Escherichia coli* (NM 539) se infectó con el fago λ y se propagó para preparar el banco de ADN cromosómico K4.

(2) Preparación de la sonda

Entre 3 regiones del resto del grupo génico del antígeno K de la cepa K5 de *Escherichia coli* (serotipo 010:K5(L):H4, ATCC 23506) que tienen secuencias conocidas (*Mol. Microbiol.*, 17(4), 611-620 (1995)), a la vez que se interpone la región R-II específica del polisacárido del antígeno K (banco de genes nº de registro X77617), una serie de cebador (CS-S 5'-ACCCAACACTGCTACAACCTATATCGG-3' (SEC. ID. nº 5); CS-AS 5'-GCGTCTTCACCAATAAATACCCACGAACT-3' (SEC. ID. nº 6)) para obtener un fragmento de ADN de aproximadamente 1 kpb del terminal 3' de la región R-I (banco de genes nº de registro X74567) y otra serie de cebador (TM-S 5'-CGAGAAATACGAACACGCTTTGGTAA-3' (SEC. ID. nº 7); TM-AS 5'-ACTCAATTTTCTTTCAGCTCTTCTTG-3' (SEC. ID. nº 8)) para obtener el fragmento de ADN de aproximadamente 1 kpb procedente del terminal 5' de la región R-III (banco de genes nº de registro X53819) se seleccionaron y prepararon.

Utilizando los respectivos conjuntos de cebador para respuesta inmunitaria para R-I y R-III y utilizando, como plantilla, fragmentos de ADN genómico de la cepa K4 extraída y purificada tras el tratamiento con Sau3AI y posterior electroforesis en gel de agarosa en el apartado (1) anterior, se llevó a cabo la reacción en cadena de la polimerasa (PCR) (94°C, 1 min.-(94°C, 45 s-47°C, 30 s-72°C, 5 min.) 30 ciclos-72°C, 10 min. (para R-I), 94°C, 1 min.-(94°C, 45 s-50°C, 30 s-72°C, 5 min.) 30 ciclos-72°C, 10 min. (para R-III) para obtener el fragmento de ADN de la región R-I de 1,3 kpb (K4RI3) procedente de K4 de la región R-III de 1,0 kpb (K4RII5). Se determinaron las secuencias nucleotídicas de los fragmentos de ADN obtenidas de este modo utilizando el analizador genético ABI PRISM 310 (Perkin-Elmer). La homología con la cepa K5 del ADN en las mismas posiciones genéticas fue 96% y 95%, respectivamente.

(3) Clonación genética de la región K4R-II

Utilizando como sondas los fragmentos (K4RI3 y K4RII5) de ADN de la región R-I y de la región R-III respectivas, se identificaron los grupos génicos del antígeno K4 del banco de ADN cromosómico K4 obtenido en el apartado (1) anterior. El cultivo de *Escherichia coli* (cepa NM539) (30 μ l) se infectó con el banco de ADN cromosómico K4 (40 μ l de fago λ) (37°C, 15 minutos), 10 ml de agarosa de la parte superior se añadieron a esto, la mezcla resultante se extendió sobre el medio en la placa LB contenido en cinco placas Petri de 10 x 14 cm, seguido de cultivo a 37°C por 9,5 horas para formar placas. Se prepararon dos membranas de 9 x 13 cm (Hybond-N+: Amersham Pharmacia Biotech) para cada placa, y la primera y segunda membranas se colocaron en el medio durante 1 minuto y 3 minutos, respectivamente. Después de eliminar el exceso de humedad, cada membrana se empapó durante 2 minutos en NaOH 0,5 M que contenía NaCl 1,5 M para llevar a cabo el tratamiento de desnaturalización y a continuación se empapó durante 3 minutos en Tris-HCl 1 M (pH 7,4) que contenía NaCl 1,5 M para llevar a cabo el tratamiento de neutralización. Después de secar, la membrana se desecó a 80°C durante 2 horas para preparar un filtro. El filtro se sometió a tratamiento de prehibridación a 65°C durante 1 hora, se hibridó con K4RI3 marcado con [³²P] a 64°C durante la noche (15 horas) en tampón fosfato Church 0,5 M (pH 7,2), EDTA 1 mM y SDS al 7%, y a continuación se lavó tres veces con tampón fosfato Church 40 mM (7,2) que contenía SDS al 1% (65°C, cada uno durante 15 minutos). El filtro se secó y después se expuso a una película de rayos X para obtener de este modo 30 placas positivas. La presencia de K4RI3 se confirmó para cada una de ellas por PCR, y 7 placas de entre ellas se sometieron a una segunda identificación. A continuación, el filtro hibridado con K4RI3 se hirvió en solución SDS al 0,5% durante 3 minutos para eliminar K4RI3 y después se secó para utilizarlo como filtro de hibridación K4RII5. El filtro se sometió a tratamiento de prehibridación a 65°C durante 1 hora, se hibridó con K4RII5 marcado con [³²P] a 64°C durante la noche y a continuación se lavó tres veces con tampón fosfato Church 40 mM que contenía SDS al 1%. Se secó el filtro y a continuación se expuso a una película de rayos X para obtener de este modo 29 placas positivas. Se confirmó la presencia de K4RII5 para cada una de ellas por PCR, y 18 placas de entre ellas se sometieron a una segunda identificación. En la segunda identificación, se utilizó medio de la placa LB de Φ 9 cm, y se obtuvieron placas positivas por el mismo procedimiento de la primera identificación.

Tras la primera y segunda identificación, se obtuvieron clones del fago λ de la región R-I y 10 clones de la región R-III. Cada uno de los clones se sometió a tratamiento enzimático con *Eco*RI (10 unidades: NEB), *S*all (10 unidades: NEB) y *B*amHI (10 unidades: NEB) cada una independiente o simultáneamente en varias combinaciones, y se prepararon sus cartografías de restricción basándose en el tamaño de los fragmentos observados por electroforesis (figura 1).

Entre estos clones, un clon (CS23, región de inserción de 15,4 kpb) es un clon de ADN preparado basándose en la sonda de la región R-III, pero dado que también presentaba una reacción lábil con la sonda de la región R-I, se examinó la secuencia del terminal en 5' de la región de inserción para hallar una secuencia que coincidía completamente con el extremo 3' de la sonda con la región R-I. Dado que ambos fragmentos de ADN de la región R-I y de la región R-III estaban contenidos en la región de inserción, se consideró el clon como un clon que contiene toda la región R-II del grupo genético del antígeno K de la cepa K4.

(4) Análisis genético de la región R-II de K4

Se realizó la subclonación del clon SC23 anterior para llevar a cabo su secuenciado. En primer lugar, cada uno, de los fragmentos de aproximadamente 3 kpb y 8 kpb, obtenidos tratando el clon SC23 con *EcoRI* y los fragmentos de ADN de 2 kpb, 5 kpb y 7 kpb, obtenidos tratándolo con *SalI* se ligó con un vector de clonación (pBluescript II KS(-)) y se integró en la cepa (XLI-Blue) de *Escherichia coli* para obtener un clon con diferente dirección de inserción. Repitiendo el "tratamiento de los puntos de multiclonación del vector con la transformación por ligadura de varias enzimas de restricción" en el clon, se obtuvieron 22 plásmidos que tienen fragmentos parciales de ADN de la región R-II. Realizando un secuenciado de los fragmentos de inserción del ADN y conectándolos, se determinó la secuencia genética completa de la región R-II de K4 (SEC. ID. nº 3).

(5) Identificación del gen de condroitin-polimerasa

Como resultado del análisis de la secuencia de ADN de la región R-II de la cepa K4, se predijo la presencia de 8 marcos de lectura abierta (ORF) (figura 2).

Entre ellos, el ORF en tercera posición contando a partir del lado R-III (2.061 pb (nucleótidos números 3.787 a 5.847 en la SEC. ID. nº 3, la secuencia de 2058 pb excluyendo el codón de terminación se muestra en la SEC. ID. nº 1), 686 como número de aminoácidos, peso molecular obtenido por cálculo 79.276 (SEC. ID. nº 2)) presentó el 59% de homología con una ácido hialurónico sintasa de *Pasteurella multocida* (tipo pmHAS de clase 2; *J. Biol. Chem.*, 273(14), 8454-8458 (1998)). Además, el recuento de ORF en la primera posición a partir del lado R-III (1.017 pb (nucleótidos números 643 a 1.649 en la SEC. ID. nº 3, 339 como número de aminoácidos) presentaba el 60% de homología con UDP-4-epimerasa de *Pasteurella multocida* (propuesta (29-OCT.-1996) Genetics and Microbiology, Universidad Autónoma de Barcelona, Edificio C, Bellaterra, BCN 08193, España), ORF en cuarta posición (1.332 pb (nucleótidos números 5.877 a 7.207 en la SEC. ID. nº 3)) presentaba gran homología (98%) con la secuencia 2 de inserción (*Nucleic Acids Res.*, 6(3), 1.111-1.122 (1979)) y el ORF en séptima posición (1.167 pb (nucleótidos números 11.453 a 12.619 en la SEC. ID. nº 3), 389 como número de aminoácidos) presentaba 65% de homología con el gen *kfiD* (*Mol. Microbiol.*, 174(4), 611-620 (1995)) (codifica la UDP-glucosa-deshidrogenasa) de la cepa (K5) de *Escherichia coli*. Además, dado que el ORF en octava posición (1.035 pb (nucleótidos números 13.054 a 14.088 en la SEC. ID. nº 3), 345 como número de aminoácidos) contenía un motivo DXD común a la glucosiltransferasa, se consideró que tiene actividad de transferir azúcar. En cuanto a los tres ORF restantes (nº 2, nº 5 y nº 6 (nucleótidos números 1.849 a 3.486, 7.210 a 8.673 y 9.066 a 10.631, respectivamente, en la SEC. ID. nº 3), no se encontró ninguna homología.

Ejemplo 2**Expresión y actividad enzimática de la proteína condroitina-polimerasa**

(1) Para confirmar que el ORF de la región R-II de K4 (nº 3) es un gen de condroitina-polimerasa, se prepararon los cebadores con puntos de corte de la enzima de restricción y que interponen el correspondiente resto de ORF (K4C-SP 5'-CGGGATCCCGATGAGTATTCTTAATCAAGC-3' (SEC. ID. nº 9); K4C-AS 5'-GGAATTCGGGCCAGTCTACATGTTTATCAC-3' (SEC. ID. nº 10)) y se llevó a cabo la PCR (94°C, 1 min.-(94°C, 30 s-59°C, 30 s-74°C, 3 min.) 20 ciclos-74°C, 10 min.). El producto de la PCR se sometió a electroforesis en gel de agarosa al 0,7% y se extrajo y purificó utilizando un kit de extracción en gel (QIAGEN). Después del tratamiento con las enzimas de restricción (*BamHI* y *EcoRI*), el producto se sometió de nuevo a electroforesis en gel de agarosa al 0,7% y se extrajo y purificó de la misma manera para utilizarse como inserción.

La inserción preparada en el párrafo anterior se insertó en un vector de expresión (pTrcHisC: Invitrogen; que contenía la secuencia nucleotídica representada por SEC. ID. nº 4) que se había tratado con las enzimas de restricción (*BamHI* y *EcoRI*) y CIP, a 16°C pasando 1 hora en presencia de T4 ADN ligasa, y se transformó en la cepa (TOP10) de *Escherichia coli*. Cultivando la cepa de *Escherichia coli* (medio de la placa LB que contenía ampicilina, 37°C, durante la noche), se obtuvieron 7 colonias. Un clon que contenía un plásmido en el que se insertó correctamente la inserción anterior se seleccionó de entre ellos. Se cultivó *Escherichia coli* (a 37°C durante la noche) en 1,5 ml de medio LB que contenía ampicilina (100 µg/ml) y 50 µl de la suspensión de las células cultivadas se inoculó en 50 µl de medio LB que contenía ampicilina (100 µg/ml) y se cultivó a 37°C hasta que la DO₆₀₀ fue de 0,6. Al cultivo se añadió 1 ml de isopropil 1-tio-β-D-galactósido (IPTG) 0,5 M (concentración final: 1 mM) y la inducción se realizó a 37°C durante 3 horas. Se recogieron las células por centrifugación (10.000 rpm, 30 minutos) y se pusieron en suspensión añadiendo 4 ml de un tampón de lisis (Na₂B₄O₇ 50 mM (pH 8,0) que contenía NaCl 300 mM e imidazol 10 mM). A la suspensión, se agregaron 4 mg de lisozima (Sigma), la mezcla resultante se dejó en reposo en hielo durante 30 minutos y a continuación se destruyeron las células tres veces por exposición a ultrasonidos, cada una durante 10 segundos, utilizando un baño con ultrasonidos. Se recogió el sobrenadante por centrifugación (10.000 rpm, 30 minutos) y se aplicó a la columna de agarosa con Ni-NTA (1 ml de vehículo, equilibrado con tampón de lisis, QIAGEN), seguido de agitación con un agitador C durante 1 hora. El vehículo se sumergió instalando la columna y a continuación la columna se lavó dos veces utilizando 4 ml para cada uno, uno era un tampón (NaH₂PO₄ 50 mM (pH 8,0) que contenía NaCl 300 mM e imidazol 20 mM). Después, se eluyeron las proteínas pasando 4 veces 0,5 ml de cada una de un tampón de elución (NaH₂PO₄ 50 mM (pH 8,0) que contenía NaCl 300 mM e imidazol

250 mM). El eluido que contenía la proteína de interés (1 ml) se dializó a 4°C durante 2 días frente a 500 ml de PBS (solución salina tamponada con fosfato) que contenía glicerol al 20% para obtener de este modo aproximadamente 0,5 ml (contenido de proteína 0,4 mg/ml) de una solución dializada (solución de la enzima de la presente invención (proteína de la presente invención)).

La inmunotransferencia Western de la proteína obtenida de este modo se llevó a cabo por electroforesis en dodecil sulfato sódico con gel de poliacrilamida (SDS-PAGE). En la SDS-PAGE se utilizó gel al 10%. La proteína se detectó por tinción con azul brillante de Coomassie. Se realizó la inmunotransferencia Western transfiriendo la proteína al gel SDS-PAGE en una membrana de nitrocelulosa, bloqueando la membrana con leche descremada al 5% (disuelta en Tris-HCl 25 mM (pH 7,5) que contenía NaCl 150 mM y Tween 20 al 0,05% (esta solución se denominó TBS-T) y a continuación tratándola con anticuerpo anti-tetra-His (Qiagen). Después de lavar varias veces con TBS-T, esta membrana se trató con IgG anti-ratón unida a peroxidasa. Después de lavar con TBS-T, la proteína reaccionada se detectó por el sistema de detección ECL (Amersham).

Como resultado, la proteína presentaba una banda a alrededor de 80 kDa por análisis de inmunotransferencia Western utilizando SDS-PAGE y anticuerpo anti-tetra-His. Por otra parte, en el extracto de cultivo de una referencia no se detectó una banda que reacciona inmunológicamente (vector de expresión sin inserción).

(2) Análisis de actividad enzimática (análisis de actividad de transferencia de GalNAc)

La enzima de la presente invención (2 µg) hexasacárido de sulfato C de condroitina del cartílago del tiburón, purificado por degradación con hialuronidasa testicular, como aceptor (70 pmol) y UDP-GalNAc (3 nmoles), UDP-GlcUA (3 nmoles) y UDP-[³H]GalNAc (0,1 nmol, 0,1 µCi) se añadieron como donantes a Tris-HCl 50 mM (pH 7,2) que contenía MnCl₂ 20 mM, (NH₄)₂SO₄ 0,1 M y etilenglicol 1 M y el volumen total se ajustó a 50 µl, y a continuación, se llevó a cabo la reacción a 30°C durante 30 minutos y la enzima se inactivó térmicamente. A la solución de reacción, se añadieron 3 volúmenes de etanol al 95% que contenían acetato de potasio al 1,3%, y la muestra se centrifugó a 10.000 x g durante 20 minutos. El precipitado se disolvió en 50 µl de agua destilada y se aplicó a una columna de péptido Superdex (300 x Φ10 mm: Amersham Biosciences, condiciones de cromatografía; tampón: NaCl 0,2 M, caudal: 0,5 ml/min) y el eluido se fraccionó a 0,5 ml y se midió la radioactividad (recuento de [³H]) de cada fracción utilizando un contador de centelleo. Se determinó la actividad sintetizadora de la condroitina calculando la cantidad de radioactividad incorporada en las fracciones de peso molecular más alto que el sustrato del aceptor. Los resultados se muestran en la figura 3. En la figura 3, los cuadrados en negro indican radioactividad cuando se utilizó el hexasacárido del sulfato C de condroitina como aceptor, los triángulos en blanco indican la radioactividad cuando se trató el producto de la reacción enzimática con condroitinasa ABC y los círculos en negro indican la referencia (en la que se utilizó la enzima inactivada térmicamente de la presente invención).

Como resultado, la posición de la elución de la radioactividad apareció en el lado del peso molecular más alto que el del hexasacárido de sulfato de condroitina C (el pico ancho que tiene su parte superior a alrededor de 5.000 Da) (cuadrados en negro en la figura 3). Además, cuando el producto de la reacción enzimática se trató con condroitinasa ABC, el pico del peso molecular alto se desplazó a una posición correspondiente a la fracción de disacárido (triángulos en blanco en la figura 3). Cuando la composición del disacárido de este digesto de condroitinasa ABC se analizó utilizando una cromatografía líquida de alta resolución (HPLC), solamente se detectó un disacárido (ΔdiOS) de condroitina insaturado y no sulfatado.

Además, el producto de la reacción enzimática fue digerido completamente por tratamiento con condroitinasa ACII también, pero no fue digerido por hialuronidasa de *Streptomyces* ni por heparitinasa I.

Basándose en lo anterior, se demostró que la enzima de la presente invención obtenida de este modo por lo menos transfiere GalNAc al hexasacárido del sulfato de condroitina C desde UDP-GalNAc (donante). Esta actividad específica fue de 3,25 ± 0,64 nmoles de GalNAc/min/mg de proteína.

(3) Análisis del tamaño del producto de la reacción enzimática

El tamaño del producto de la reacción enzimática se examinó llevando a cabo la reacción enzimática y la cromatografía de la misma manera que en el "apartado (2) Análisis de actividad enzimática" anterior. Los resultados se presentan en la figura 4.

Se confirmó a partir de la figura 4 que la enzima de la presente invención obtenida anteriormente por lo menos transfiere GalNAc a un aceptor (oligosacárido que tiene un esqueleto de condroitina (hexasacárido de sulfato de condroitina C preparado utilizando hialuronidasa testicular)) del donante GalNAc (UDP-GalNAc) para formar de este modo condroitina que tiene un peso molecular de 10.000 a 20.000 o más.

(4) Análisis de especificidad del sustrato del donante

Utilizando UDP-[¹⁴C]GlcUA, UDP-[¹⁰C]GlcNAc o UDP-[³H]GalNAc como donante, se examinó la actividad de transferencia de los siguientes aceptores según el procedimiento descrito en el "apartado (2) Medición de la

actividad enzimática" anterior. Los productos después de la reacción enzimática se analizaron por filtración en gel.

Los resultados de la utilización de los hexasacárido o heptasacárido de sulfato de condroitina C como aceptor se muestran en la figura 5. Además, el círculo en negro en la figura 5 indica una referencia (un caso en el que se utilizó la enzima inactivada térmicamente de la presente invención).

(A) Hexasacárido de sulfato de condroitina C (producto purificado al degradar sulfato de condroitina C del cartílago de tiburón con hialuronidasa testicular, y el terminal no reducido es GlcUA).

Se sintetizó el heptasacárido solo cuando se utilizaba UDP-GalNAc solo como donante (pastillas en negro en la figura 5).

La transferencia sustancial no se encontró cuando se utilizaba UDP-GlcUA solo como donante (triángulo en negro en la figura 5).

Aunque se encontró muy poca transferencia de GlcNAc cuando se utilizaba UDP-GlcNAc solo como donante. Esta actividad de transferencia era el 6,3% basada en la actividad al 100% en el caso de la utilización de UDP-GalNAc como donante (cuadrado en negro en la figura 5). Sin embargo, los polisacáridos que tienen un tamaño de octasacárido o más no se obtuvieron aun cuando se utilizaba tanto UDP-GlcNAc como UDP-GlcUA junto con éste.

Además, ya que la incorporación de la radioactividad no se encontraba en ausencia del sustrato aceptor, se sugirió que el sustrato aceptor es esencial para la síntesis de la condroitina por esta enzima.

(B) Heptasacárido de sulfato de condroitina C (producto obtenido dejando que la enzima de la presente invención reaccione con el anterior hexasacárido de sulfato de condroitina C y uniendo de este modo un resto de GalNAc al terminal no reducido del hexasacárido).

Se sintetizó octasacárido solo cuando se utilizaba UDP-GlcNAc como donante (triángulo en blanco en la figura 5).

Cuando se utilizaba UDP-GalNAc o UDP-GlcNAc solo como donante, no se encontró transferencia sustancial en cada caso (respectivamente, pastilla en blanco y cuadrado en blanco en la figura 5).

Además, en la Tabla 1 se muestran los resultados de llevar a cabo pruebas similares independientemente cada una. Además, el término "SC" en la Tabla 1 significa sulfato de condroitina C. Además, los paréntesis en la Tabla muestran la longitud de la cadena de azúcar después de la reacción de la enzima y "-" significa que el azúcar marcado con UDP no se transfirió al sulfato de aceptor correspondiente.

Tabla 1

Sustrato del donante		Actividad específica de condroitina-polimerasa (nmoles/min/mg de proteína)	
Azúcar marcado con UDP	Azúcar sin marcar con UDP	Hexasacárido de SC	Heptasacárido de SC
UDP-[³ H]GalNAc UDP-	ninguno	0,59 ± 0,16 (hepta)	0,0 ± 0,0 (-)
[¹⁴ C]GlcNAc UDP-	ninguno	0,04 ± 0,02 (hepta)	0,0 ± 0,0 (-)
[¹⁴ C]GlcUA	ninguno	0,0 ± 0,0 (-)	0,53 ± 0,08 (octa)
UDP-[³ H]GalNAc UDP-	UDP-GlcUA	3,25 ± 0,64 (poli)	no medido
[¹⁴ C]GlcUA UDP-	UDP-GalNAc	2,75 ± 0,28 (poli)	no medido
[¹⁴ C]GlcNAc UDP-	UDP-GlcUA	0,05 ± 0,02 (poli)	no medido
[¹⁴ C]GlcUA	UDP-GlcNAc	0,0 ± 0,0 (-)	no medido

A partir de los resultados anteriores, se demostró que la enzima de la presente invención obtenida en lo anterior transfiere GalNAc desde UDP-GlcNAc a una cadena de azúcar (aceptor) que tiene un esqueleto de condroitina cuyo terminal no reducido es GlcUA. Además, se demostró que cuando se utiliza el aceptor, la enzima de la presente invención obtenida en lo anterior presenta la actividad para transferir GlcNAc desde UDP-GlcNAc, pero la actividad es mucho menor que su actividad de transferencia de GalNAc. Además, se demuestra que cuando se utiliza el aceptor, la enzima de la presente invención obtenida en lo anterior no tiene sustancialmente actividad para transferir GlcUA desde UDP-GlcUA. Basándose en estos resultados, se demostró que la enzima anterior de la presente invención no es capaz de transferir GlcUA al terminal GlcUA no reducido pero es capaz de transferir un resto de GalNAc (o, bastante ligero, GlcNAc).

Además, se demostró que la enzima de la presente invención obtenida en lo anterior transfiere GlcUA desde UDP-GlcUA a una cadena de azúcar (aceptor) que tiene un esqueleto de condroitina cuyo terminal no reducido es GalNAc. Además, se demostró que cuando se utiliza el aceptor, la enzima de la presente invención obtenida en lo anterior sustancialmente no tiene actividades para transferir GalNAc desde UDP-GalNAc y para transferir GlcNAc

desde UDP-GlcNAc. Basándose en estos resultados, se demuestra que la enzima anterior de la presente invención no es capaz de transferir GalNAc al terminal GalNAc no reducido pero es capaz de transferir un resto de GlcUA.

Basándose en lo anterior, se demostró que la enzima anterior de la presente invención es capaz de transferir GlcUA y GalNAc alternativamente, desde un donante GlcUA y un donante GalNAc, respectivamente, al terminal no reducido de la cadena de azúcar.

(5) Análisis de especificidad del sustrato aceptor

Utilizando tetrasacárido (140 pmoles) o hexasacárido (140 pmoles) de sulfato de condroitina C degradado y purificado que utiliza hialuronidasa testicular, tetrasacárido (260 pmoles) o hexasacárido (175 pmoles) de condroitina degradada y purificada por el método de Nagasawa (*Carbohydrate Research*, 97, 263-278 (1981)), hexasacárido (175 pmoles) de ácido hialurónico degradado y purificado utilizando hialuronidasa testicular, sulfato de condroitina C (peso molecular 20.000), condroitina (peso molecular 10.000), sulfato de dermatán (peso molecular 15.000), ácido hialurónico (peso molecular 20.000), o heparina (peso molecular 10.000) como aceptor, se examinó la actividad de transferencia por el método siguiente. Además, las cadenas de azúcar se adquirieron en Seikagaku Corporation.

La enzima de la presente invención (2 µg), UDP-GalNAc (Sigma) (60 pmoles), UDP-GlcUA (Sigma) (0,6 nmoles) y UDP-[³H]GalNAc (0,1 nmoles, 0,1 µCi) como donantes de cada una de las anteriores cadenas de azúcar como aceptor se añadieron a Tris-HCl 50 mM que contenía MnCl₂ 20 mM, 0,1 M (NH₄)₂SO₄ y etilenglicol 1 M, y el volumen total se ajustó a 50 µl, y a continuación se llevó a cabo la reacción enzimática a 30°C durante 30 minutos y la enzima se inactivó térmicamente. Se aplicó la solución de reacción a una columna Superdex Peptide (300 x Φ10 mm: Amersham Biosciences, condiciones de la cromatografía; tampón: NaCl 0,2 M, caudal: 0,5 ml/min) y se fraccionó el eluido en 0,5 ml y se midió la radioactividad (recuento de [³H]) de cada fracción utilizando un contador de centelleo. Los resultados se presentan en la Tabla 2. Los numerales entre paréntesis en la Tabla son valores relativos cuando la cantidad de radioactividad (cantidad de GalNAc transferida) mediante la utilización del hexasacárido del sulfato de condroitina C como aceptor se definió como 100%.

Tabla 2

Sustrato del aceptor		Actividad específica de incorporación de [³ H]	
		nmoles/min/mg de proteína	%
Sulfato de condroitina C	Tetrasacárido	1,44 ± 0,24	43,0
	Hexasacárido	3,34 ± 0,50	100,0
	Polisacárido (PM 20.000)	3,41 ± 0,48	100,0
Condroitina	Tetrasacárido	1,12 ± 0,21	33,5
	Hexasacárido	1,24 ± 0,45	37,0
	Polisacárido (PM 10.000)	0,53 ± 0,13	15,8
Ácido hialurónico	Hexasacárido	0,80 ± 0,15	24,0
	Polisacárido (PM 20.000)	0,27 ± 0,02	8,2
Sulfato de dermatán	Polisacárido (PM 15.000)	0,06 ± 0,02	1,9
Heparina	Polisacárido (PM 10.000)	0,0 ± 0,0	0,0

Basándose en los resultados anteriores, se demostró que el hexasacárido de sulfato de condroitina C se vuelve el sustrato aceptor más adecuado. El hexasacárido de condroitina también funcionó como un sustrato aceptor, pero su actividad fue baja (37%) en comparación con el caso del hexasacárido de sulfato de condroitina C. La incorporación en el tetrasacárido del sulfato de condroitina o en tetrasacárido de condroitina fue la misma (43% y 33,5%, respectivamente). Para sorpresa de los autores, el hexasacárido del ácido hialurónico y el ácido hialurónico (peso molecular 20.000) funcionaron también como sustratos aceptores. El nivel de incorporación del sulfato de condroitina C (peso molecular 20.000) fue similar al del hexasacárido de condroitina C. La incorporación no fue tan alta en el caso de la condroitina (peso molecular 10.000).

Resumiendo los resultados anteriores, se demostró que la enzima anterior de la presente invención utiliza oligosacáridos y polisacáridos que tienen esqueleto de condroitina (por lo menos tetrasacárido, hexasacárido y heptasacárido de sulfato de condroitina C, tetrasacárido y hexasacárido de condroitina, sulfato de condroitina C (peso molecular 20.000) y condroitina (peso molecular 10.000)) y oligosacáridos y polisacáridos de ácido hialurónico (por lo menos hexasacárido de ácido hialurónico y ácido hialurónico (peso molecular 20.000)) como aceptores.

Por otra parte, la incorporación de la radioactividad no se encontró en el sulfato de dermatán (peso molecular 15.000) ni heparina (peso molecular 10.000) lo que demuestra que no funcionan sustancialmente como sustratos

aceptores.

(6) Análisis de la influencia de la temperatura

5 La reacción enzimática se realizó de la misma manera que en el "apartado (2) Análisis de actividad enzimática" anterior, cambiando la temperatura de la reacción enzimática a 25, 30, 37, 40, 45 ó 100°C, y la solución después de la reacción se aplicó a una columna Superdex 74 (30 x Φ10 mm: Amersham Biosciences, condiciones de la cromatografía; tampón: NaCl 0,2 M, caudal: 0,5 ml/min). Se fraccionó el eluido en porciones de 1 ml y se midió la radioactividad (recuento de [³H]) de cada fracción utilizando un contador de centelleo. Los resultados se muestran en la figura 6. Además, las marcas de la pastilla, cuadrado, triángulo, signo x y * en la figura 6 presentan los resultados de 25, 30, 37, 40, 45 ó 100°C, respectivamente. Además, el círculo en la figura 6 presenta el resultado de una referencia (en la que se utilizó la enzima inactivada térmicamente de la presente invención).

15 Como se muestra en la figura 6, en las condiciones de reacción y en el intervalo de temperatura examinado esta vez, el peso molecular del producto de reacción se volvió pequeño a medida que se aumentaba la temperatura. La incorporación más alta se encontró a 30°C, pero el producto con el peso molecular mayor se obtuvo a 25°C.

20 Se considera a partir de los resultados que la actividad enzimática de la enzima anterior de la presente invención disminuye a medida que aumenta la temperatura de reacción comenzando a 25°C en las condiciones de reacción de esta vez.

(7) Análisis de la influencia de los iones metálicos y similares

25 La reacción enzimática se realizó de la misma manera que en el "apartado (2) Análisis de actividad enzimática" anterior, excepto que cada una de las diversas sales metálicas (MnCl₂, FeCl₂, MgCl₂, CaCl₂ o CuCl₂) o se añadió EDTA en lugar de MnCl₂, y la solución después de la reacción se aplicó a una columna Superdex 75 (300 x Φ100 mm. Amersham Biosciences, condiciones de la cromatografía; tampón: NaCl 0,2 M, caudal: 0,5 ml/min). Se fraccionó el eluido en 1 ml y se midió la radioactividad (recuento de [³H]) de cada fracción utilizando un contador de centelleo. Los valores relativos de radioactividad cuando está presente MnCl₂ se definió como 100% se presentan a continuación.

Tabla 3

Sal metálica	Valor relativo de radioactividad (%)
MnCl ₂	100,0
FeCl ₂	30,6
MgCl ₂	30,7
CaCl ₂	0,0
CuCl ₂	0,0
EDTA	0,0

35 Basándose en estos resultados, la enzima anterior de la presente invención presentaba la actividad mayor en presencia de ión Mn²⁺. Además, en presencia del ión Fe²⁺ o Mg²⁺, presentaba aproximadamente el 30% de actividad en comparación con el caso de la presencia del ión Mn²⁺. Además, se demostró que no actúa sustancialmente en presencia del ión Ca²⁺ o Cu²⁺ o del ácido etilendiamintetraacético.

40 Además, se demostró que la enzima anterior de la presente invención actúa en presencia del ión Fe²⁺ o Mg²⁺ también, y su actividad enzimática en presencia del ión Mn²⁺ es mayor que la actividad enzimática en presencia del ión Fe²⁺ o Mg²⁺.

(8) pH óptimo de reacción

45 Cuando el pH óptimo de reacción de la enzima de la presente invención se examinó cambiando el pH del "apartado (2) Análisis de actividad enzimática" anterior a varios niveles, era de pH 7,0 a 7,5.

(9) Relación con el tiempo de reacción enzimática

50 Ajustando el tiempo de reacción enzimática en 10 minutos, 30 minutos, 1 hora, 3 horas, 6 horas o 18 horas, la reacción enzimática se realizó de la misma manera que en el "apartado (2) Análisis de actividad enzimática" anterior, y se analizó la radioactividad incorporada en el producto de la reacción enzimática. Los patrones de filtración en gel de [³H]GalNAc después de varios periodos de reacción se presentan en la figura 7, y las cantidades totales de incorporación de radioactividad después de varios periodos de reacción enzimática en la figura 8. El círculo blanco, el círculo negro, el triángulo blanco, el triángulo negro, el cuadrado blanco y el cuadrado negro, presentan los resultados después de 10 minutos, 30 minutos, 1 hora, 3 horas, 6 horas y 18 horas, respectivamente. Además, la flecha con "20 k", "10 k", "5 k", "14", "8" ó "6" presenta la posición de elución del peso molecular 20.000, 10.000, 5.000, tetradecasacárido (peso molecular: aproximadamente 2.800), octasacárido (peso molecular: 1.600) o

hexasacárido (peso molecular: aproximadamente 1.200) de ácido hialurónico (patrón), respectivamente.

Se demostró a partir de los resultados de la figura 8 que en las condiciones de ensayo, la incorporación rápida de [³H]GalNAc se encuentra después de 3 horas y de 6 horas, pero la incorporación se vuelve lenta a medida que la reacción se acerca a 20 horas.

Además, a partir de los resultados de la figura 7, se demostró que la incorporación aumenta y se obtiene un producto de reacción de alto peso molecular después de un largo periodo de tiempo de reacción.

Además, un producto de peso molecular elevado se obtuvo rápidamente cuando el sustrato aceptor (hexasacárido) se puso a concentración más baja, y un producto de bajo peso molecular se obtuvo cuando el sustrato aceptor (hexasacárido) se puso a alta concentración.

(10) Determinación de la constante de Michaelis (K_m)

La radioactividad incorporada en el producto de la reacción enzimática se midió según el "apartado (2) Análisis de actividad enzimática" anterior, colocando la cantidad de la enzima de la presente invención a 1,3 µg, que contenía el sustrato donante (UDP-azúcar; UDP-GlcUA o UDP-GalNAc) en una concentración fija (240 µM) y añadiendo a ésta el otro sustrato donante radiomarcado (radiación UDP-azúcar; UDP-[³H]GalNAc o UDP-[¹⁴C]GlcUA) con varias concentraciones (0,6 a 200 µM). Se realizaron ensayos independientes tres veces, y el valor promedio se utilizó como valor medido.

La relación entre la radioactividad incorporada (V) y la relación entre sustrato (S) de UDP-azúcar se muestra en la figura 9, y su doble representación recíproca en la figura 10. El círculo negro y el cuadrado blanco en los dibujos presenta los resultados sobre UDP-GlcUA y UDP-GalNAc, respectivamente.

Como resultado, la K_m para UDP-GlcUA y la K_m para UDP-GalNAc se calculó que eran 3,44 µM y 31,6 µM, respectivamente.

Listado de secuencias

<110> Seikagaku Corporation

<120> CONDROITINA-POLIMERASA Y ADN QUE LA CODIFICA

<150> JP 2001-244685

<151> 2001-08-10

<150> JP 2001-324127

<151> 2001-10-22

<150> JP 2002-103136

<151> 2002-04-05

<160> 11

<170> PatentIn Ver. 2.1

<210> 1

<211> 2058

<212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

<222> (1).. (2058)

<400> 1

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
 y
 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 1243
 Gly Lys Asn Ile Thr Val Gln Leu Leu 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 2058
 Glu Cys Arg Lys Tyr Thr Trp Glu Lys Ile Glu Asn Asp Leu
 675 680 685

<210> 2

<211> 686

<212> PRT

<213> *Escherichia coli*

<400> 2

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

5

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> 3

5 <211> 14483

<212> ADN

<213> *Escherichia coli*

<220>

10 <221> CDS

<222> (3787)..(5847)

<400> 3

15

```

cgacattatg tctttsagaa tttaaatatt gaaatccctt caggaaaaag tgttgccctt 60
attgggtcgt atgggtcggg taatcaacg ttactgagaa tgattgggtg cattgaccgc 120
cccgatagcg gaagatcat caccaatnaa acgatatacat ggccagtcgg ccttgtaggt 180
ggatttcagg gaagttaa acggacgcgaa aatgtaaaat ttgtcgcgag gttatacgcg 240
aagcaagaag aactgaaaga gaaaattgag tttgttgaa aatttgccga actcggcaag 300
tattttgata tgcgatcaa aacttactcc tctggtatgc gatctgcctt aggccttggg 360
ttaagtatgg catttaaat tgattattat atcgtcgatg aagtaaccgc agtcggtgat 420
gccagggtta aagaanaatg cgtcaattg tttaaagaaa ggcataaaga atctagtttt 480
ttaatgggtt cacatagttt gaattcatig aaagagtttt gtgatgtggc cattgttttt 540
aaggatgaca atgctggttag tttcatgag gatgttcagg aggggataga agagtatata 600
acggaacaaa ataattactg atgttgtttt caagggtgaa aaatgaatat attagttaca 660
ggtggagcag gctatatagg ctcgcatact agtttatgtc ttctgaataa aggttacaat 720
gttgtaatca ttgacaactt aattaattca tcttgcgaga gcattcgaag gattgaatta 780
atagctaaaa aaaaagtac tttctatgag ttgaacatca acaatgaaaa agaagttaat 840
cazattctaa aaaaacacaa atttgattgt ataatgcatt ttgccgtgac aaagtctgtt 900
gctgaatcct taataaaacc cattttttat tatgataata atgtttcagg gacgttgcaa 960
ttaattaatt gcgctataaa aaacgatgtg gctaatttta tttttagctc ttctgcaacg 1020
gtttatgggt aagcaaaaat aatgcctgtt acagaagatt gccatatagg aggaacatta 1080
aatccatatg gtacatcaaa gtatatatca gaattgatga ttagagatat tgcaaaaaaa 1140
tatagcgata ctaatttttt cgtctgaga tattttaacc caacagggtc tcacgagtcg 1200
ggaatgatcg gtgaaagtcc cgtgatata ccaagcaatt tagttcctta tatattacaa 1260
gttgctatgg gtaaaactaga aaaacttatg gtgtttgggg gggattaccc tacaaaggat 1320
ggaaccgggt ttcgtgatta tatacacgta atggatttag cggaagggca tgtggctgct 1380
ttatcttacc tttccgtga taataacact aattatcatg tttttaattt aggtactggt 1440
aaaggatatt ctgttttaga gctggtttct accttgaaa aaatatctgg ggttagaatt 1500
ccatattgaa ttgtttcgag aagagatggg gatattgctg aaagttagtc atcaccagaa 1560
aaagcaataa agtatctcaa ttggaagctt aaaaggggat tggaaacaat gcttgaggat 1620
gcctggcgct ggcaaatgaa aaacccaaat ggttatattt aatcatgcat aaggtaaagc 1680
aaaattttac gataacttct atactatttg gcatcttaat aatcgaaatt tagtgtggtg 1740
aattggttta ataaattcgt agacactaaa aagtatggtg ggtgttaatt aaagctaact 1800
aaacaaatgg cgaataaagg gttttctgcc ttggggggat aaacatcat gaatagacta 1860
gtaatagttg gtcacccgag ctctaattat caaattgtag aagaactttt gcatcaaga 1920
ggaatgaatt ctctatgtcc atcgaagcga gataatttaa gccccaaaga tatcactcag 1980
acgttctgta aggcataatc atcccctgat atatacactg taacagatag tgotgatttc 2040
gaaccattac acgtttctac tgtctggaat ggtatagccc ttgacttgat gcttagtaat 2100
ttgaaccaaa aattgtgcgg atggtcggat cctaattgcaa tccatacatt agaattattg 2160
aagagtgttg atgaaaacat tacatttatt ctaatttacg atcatccaaa gtctatttta 2220
acaaattatt tttcagatca aaatatatcg tccaactata cgtcagaaca ttttaattaa 2280

```

aactggcttg cctataatac agcattgtta cacttttttc ttaataatcg cggtaggtgt 2340
ttattagtaa gctcagaaca agtcaaacgt aatgcagagg attgcataca acaacttcaa 2400
cataagotta agttgaaatt tggctctttca ttttcaataa cgattaatca ttactagaa 2460
caatctgtaa atgattttta gacggctgaa gcttcgatta ctctggaana agagoatcag 2520
gaaataatgt ctctatcagg tattgatata ggaaccggag atattatatt caaacagagt 2580
gaaacagagg aatattttaat tttcaatgta ttgaatgatt atccagattg taaagagctt 2640
tattttgaat tacaatctaa tgcaaatact ccgcttaggg ttttagasaa ggaaaattac 2700
aagccttcc tttatggga gacatttata aaacaacgoc aaataacatt agatattgtc 2760
aatggattat atcagtcctc taaaaaata attttagata acgagttaca cacatcaaaa 2820
caattaaatg catatcaggc tattttaaaa gaattgagtg actctaaaga agaattgatt 2880
caatatgatt taataataaa aaataagaca atacaagttc aggagcttga atgcgccata 2940
gaaaattttg aatcgctgtt aaagaagaa caaataaaaa atgaattgca acaacaaga 3000
ttggaaaaat taagtgtgta aaaageacta ctgcttaatc aattgcattt agtacaacaa 3060
aaacttgaac aatattttat tgataatcaa cgattagaaa aaaaaaacct tccagaatta 3120
tatgggtcag cagaacgtat aacacaggac attggatctc gcctgggtgc tgaatgggtt 3180
agtcgttcaa aaacattttt aggettaatt agcattcctt ttgctttgat atccgaatgg 3240
cgaacatgga aaaagaaata cgaatgtgaa tatcaagtct ctctgccatc aatatttctt 3300
tatgctgata aacatgagggc agaaagggtg aaaaaacatt tatcctatca attaggaaaa 3360
ttaatcataa atcaaatca ttttccacta gggttgatat ctttgccatt ttcgatatac 3420
agaacaatac gtcaattcaa aagaacaaaa aataattctc aggtaggggt taaatactgt 3480
ggaaaataaa tccaggctac taatatataa gttaaaatat cggctataaa tgcgtgcgaa 3540
ttaatagtgaa aaattttctt agttaagtga aatagctttt tictaattgt ttaagtcata 3600
gtggtttacg ctttatttta ttaaaaaat aaaataataa attaaaaata cgattctcaa 3660
tatttcttca agtatcaatt aggatttaat ggggcaagat tatgatatcg catgaaaaata 3720
tatatatagg gacatgatta ttatgcgtgt tgatacttta attatactag attagggtga 3780
aataat atg agt att ctt aat caa gca ata aat tta tat aaa aac aaa 3828
Met Ser Ile Leu Asn Gln Ala Ile Asn Leu Tyr Lys Asn Lys
1 5 10
aat tat cgc caa gct tta tct ctt ttt gag aag gtt gct gaa att tat 3876
Asn Tyr Arg Gln Ala Leu Ser Leu Phe Glu Lys Val Ala Glu Ile Tyr
15 20 25 30
gat gtt agt tgg gtc gaa gca aat ata aaa tta tgc caa acc gca ctc 3924
Asp Val Ser Trp Val Glu Ala Asn Ile Lys Leu Cys Gln Thr Ala Leu
35 40 45
5 aat ctt tct gaa gaa gtt gat aag tta aat cgt aaa gct gtt att gat 3972
Asn Leu Ser Glu Glu Val Asp Lys Leu Asn Arg Lys Ala Val Ile Asp
50 55 60
att gat gca gca aca aaa ata atg tgt tct aac gcc aaa gca att agt 4020
Ile Asp Ala Ala Thr Lys Ile Met Cys Ser Asn Ala Lys Ala Ile Ser
65 70 75
ctg aac gag gtt gaa aaa aat gaa ata ata agc aaa tac cga gaa ata 4068
Leu Asn Glu Val Glu Lys Asn Glu Ile Ile Ser Lys Tyr Arg Glu Ile
80 85 90
acc gca aag aaa tca gaa cgg cgc gag tta aag gaa gtc gaa ccc att 4116
Thr Ala Lys Lys Ser Glu Arg Ala Glu Leu Lys Glu Val Glu Pro Ile
95 100 105 110
cct tta gat tgg cct agt gat tta act tta ccg ccg tta cct gag agc 4164
Pro Leu Asp Trp Pro Ser Asp Leu Thr Leu Pro Pro Leu Pro Glu Ser
115 120 125
aca aac gat tat gtt tgg cgc ggg aaa aga aaa gag ctt gat gat tat 4212
Thr Asn Asp Tyr Val Trp Ala Gly Lys Arg Lys Glu Leu Asp Asp Tyr
130 135 140

cca aga aaa cag tta atc att gac ggg ctt agt att gta att cct aca 4260
 Pro Arg Lys Gln Leu Ile Ile Asp Gly Leu Ser Ile Val Ile Pro Thr
 145 150 155
 tat aat cga gca aaa ata ctt gca att aca ctt gct tgt ctt tgt aac 4308
 Tyr Asn Arg Ala Lys Ile Leu Ala Ile Thr Leu Ala Cys Leu Cys Asn
 160 165 170
 caa aag acc ata tac gac tat gaa gtt att gtt gcc gat gat gga agt 4356
 Gln Lys Thr Ile Tyr Asp Tyr Glu Val Ile Val Ala Asp Asp Gly Ser
 175 180 185 190
 aaa gaa aat att gaa gaa ata gta aga gaa ttt gaa agt tta tta aat 4404
 Lys Glu Asn Ile Glu Glu Ile Val Arg Glu Phe Glu Ser Leu Leu Asn
 195 200 205
 ata aaa tat gta cgt cag aag gat tat gga tat caa ctg tgt gct gtt 4452
 Ile Lys Tyr Val Arg Gln Lys Asp Tyr Gly Tyr Gln Leu Cys Ala Val
 210 215 220
 aga aat ctt ggg ctt agg gct gca aag tat aat tat gtt gca att ctg 4500
 Arg Asn Leu Gly Leu Arg Ala Ala Lys Tyr Asn Tyr Val Ala Ile Leu
 225 230 235
 gat tgt gat atg gct ccg aac cca cta tgg gtt cag tca tat atg gaa 4548
 Asp Cys Asp Met Ala Pro Asn Pro Leu Trp Val Gln Ser Tyr Met Glu
 240 245 250
 cta tta gcg gtg gac gat aat gtt gct cta att ggc cct aga aaa tat 4596
 Leu Leu Ala Val Asp Asp Asn Val Ala Leu Ile Gly Pro Arg Lys Tyr
 255 260 265 270
 ata gat eca agc aag cat aca tat tta gat ttc ctt tcc caa aaa tca 4644
 Ile Asp Thr Ser Lys His Thr Tyr Leu Asp Phe Leu Ser Gln Lys Ser
 275 280 285
 cta ata aat gaa att cct gaa atc att act aat aat cag gtt gca ggc 4692
 Leu Ile Asn Glu Ile Pro Glu Ile Ile Thr Asn Asn Gln Val Ala Gly
 290 295 300
 aag gtt gag caa aac aaa tca gtt gac tgg cga ata gaa cat ttc aaa 4740
 Lys Val Glu Gln Asn Lys Ser Val Asp Trp Arg Ile Glu His Phe Lys
 305 310 315
 aat acc gat aat cta aga tta tgc aac aca cca ttt cga ttt ttt agc 4788
 Asn Thr Asp Asn Leu Arg Leu Cys Asn Thr Pro Phe Arg Phe Phe Ser
 320 325 330
 gga ggt aat gtc gct ttt gcg aaa aaa tgg ctt ttc cgt gca gga tgg 4836
 Gly Gly Asn Val Ala Phe Ala Lys Lys Trp Leu Phe Arg Ala Gly Trp
 335 340 345 350
 ttt gat gaa gag ttt acg cat tgg ggg gag gat aat gag ttt gga 4884
 Phe Asp Glu Glu Phe Thr His Trp Gly Gly Glu Asp Asn Glu Phe Gly
 355 360 365
 tat cgt ctc tac aga gaa gga tgt tac ttt cgg tct gtt gaa gga gca 4932
 Tyr Arg Leu Tyr Arg Glu Gly Cys Tyr Phe Arg Ser Val Glu Gly Ala
 370 375 380
 atg gca tat cat caa gaa cca ccc ggg aaa gaa aac gag acg gat cgt 4980
 Met Ala Tyr His Gln Glu Pro Pro Gly Lys Glu Asn Glu Thr Asp Arg
 385 390 395

gcg gca ggg aaa aat att act gtt caa ttg tta cag caa aaa gtt cct 5028
 Ala Ala Gly Lys Asn Ile Thr Val Gln Leu Leu Gln Gln Lys Val Pro
 400 405 410

tat ttc tat aga aaa aaa gaa aaa ata gaa tcc gcg aca tta aaa aga 5076
 Tyr Phe Tyr Arg Lys Lys Glu Lys Ile Glu Ser Ala Thr Leu Lys Arg
 415 420 425 430

gta cca cta gta tct ata tat att ccc gcc tat aac tgc tct aaa tat 5124
 Val Pro Leu Val Ser Ile Tyr Ile Pro Ala Tyr Asn Cys Ser Lys Tyr
 435 440 445

att gtt cgt tgt gtt gaa agc gcc ctt aat cag aca ata act gac tta 5172
 Ile Val Arg Cys Val Glu Ser Ala Leu Asn Gln Thr Ile Thr Asp Leu
 450 455 460

gaa gta tgc ata tgc gat gat ggt tcc aca gat gat aca ttg cgg att 5220
 Glu Val Cys Ile Cys Asp Asp Gly Ser Thr Asp Asp Thr Leu Arg Ile
 465 470 475

ctt cag gag cat tat gca aac cat cct cga gtt cgt ttt att tca caa 5268
 Leu Gln Glu His Tyr Ala Asn His Pro Arg Val Arg Phe Ile Ser Gln
 480 485 490

aaa aac aaa gga att ggt tca gca tct aat aca gca gtt aga ttg tgt 5316
 Lys Asn Lys Gly Ile Gly Ser Ala Ser Asn Thr Ala Val Arg Leu Cys
 495 500 505 510

cgg gga ttc tat ata ggt cag tta gac tct gat gac ttt ctt gaa cca 5364
 Arg Gly Phe Tyr Ile Gly Gln Leu Asp Ser Asp Asp Phe Leu Glu Pro
 515 520 525

gat gct gtt gaa cta tgt cta gat gaa ttt aga aaa gat cta tca ttg 5412
 Asp Ala Val Glu Leu Cys Leu Asp Glu Phe Arg Lys Asp Leu Ser Leu
 530 535 540

gca tgt gtt tat aca act aac cgt aat ata gat cgt gaa ggt aat ttg 5460
 Ala Cys Val Tyr Thr Thr Asn Arg Asn Ile Asp Arg Glu Gly Asn Leu
 545 550 555

ata tca aat ggc tat aat tgg ccc att tat tgc cga gaa aaa ctt act 5508
 Ile Ser Asn Gly Tyr Asn Trp Pro Ile Tyr Ser Arg Glu Lys Leu Thr
 560 565 570

agt gca atg ata tgt cat cat ttc agg atg ttc aca gca aga gca tgg 5556
 Ser Ala Met Ile Cys His His Phe Arg Met Phe Thr Ala Arg Ala Trp
 575 580 585 590

aac cta act gaa ggt ttc aac gaa tgc atc agc aac gca gtt gat tac 5604
 Asn Leu Thr Glu Gly Phe Asn Glu Ser Ile Ser Asn Ala Val Asp Tyr
 595 600 605

gat atg tat tta aaa ctt agt gaa gtt gga ccg ttc aag cat ata aac 5652
 Asp Met Tyr Leu Lys Leu Ser Glu Val Gly Pro Phe Lys His Ile Asn
 610 615 620

aaa att tgt tat aat cgc gta ttg cat ggt gaa aat acg tct ata aaa 5700
 Lys Ile Cys Tyr Asn Arg Val Leu His Gly Glu Asn Thr Ser Ile Lys
 625 630 635

aag ttg gat att caa aag gaa aat cat ttt aaa gtt gtt aac gaa tca 5748
 Lys Leu Asp Ile Gln Lys Glu Asn His Phe Lys Val Val Asn Glu Ser
 640 645 650

tta agt agg cta ggc ata aaa aaa tat aaa tat tca cca tta act aat 5796
 Leu Ser Arg Leu Gly Ile Lys Lys Tyr Lys Tyr Ser Pro Leu Thr Asn
 655 660 665 670

ttg aat gaa tgt aga aaa tat acc tgg gaa aaa ata gag aat gat tta 5844
 Leu Asn Glu Cys Arg Lys Tyr Thr Trp Glu Lys Ile Glu Asn Asp Leu
 675 680 685

taa ttattgatat attacaagt ataaacatgt agactggccc cctgaatctc 5897
 cagacaacca atatcactta aataagtgt agtcttaata ctagtittta gactagtcac 6957
 tggagaacag atgattgatg tcttagggcc ggagaaacgc agacggogta ctacacagga 6017
 aaagatcgct atcggtcagc agagcttga accgggaatg acggtctccc ttgttgcccg 6077
 gcaacatggt gtggcagcca gccagctatt tctctggcgt aagcaatacc aggaggggag 6137
 tcttaotgct gtggctgccg gagaacaggt cgttctgcc tctgaacttg ctgccgccat 6197
 gaagcagatt aaagaactcc agcgtotgct cggcaaaaaa acgatggaaa atgaactcct 6257
 taaagaagcc gttgaatatg ggcgagcaaa aaagtggata gcgcacgcgc cttattgcc 6317
 cggggatggg gaggtaagctt cgtcagccgt tgtctccggg tgcgcgtgc gcagttgcac 6377
 gtcattctca gacgagccga tgactggaag gacggccgcc gcagccgtca cacggatgat 6437
 acggatgtgc ttcgccgtat acaccatggt atcgagagag tgcacacgta tggttatcgt 6497
 cgggtatggg cgtgcttcg cagacaaaca gaacttgatg gtatgcctgc gatcaatgcc 6557
 aaacgtggtt accggaccat gcgccagaat gcgctgttgc ttgagcgaaa acccgctgta 6617
 ccgccatcga aacgggcaca taccggcaga gtggctgtga aagaaaagtaa tcagcgatgg 6677
 tgcctgacg ggtttgagtt ccgctgtgat aacggagaaa aactgcgggt cacttgcg 6737
 ctggactgct gtgaccgtga ggcactgcac tggcggtca caacgggtgg cttcgacagt 6797
 gaaacagtac aggcagtcac gctgggagca gtggaaacgcc gctttggcag cgagcttccg 6857
 gcgtctccag tggagtggct gacggataat ggttcatgct accgggcgaa tgaacacagt 6917
 cagttcgcca ggtggttggg acttgaaccg aagaacacgg cagtgcggag tccggagagt 6977
 aacgggaatag cagagagctt cgtgaaacg ataaagcgtg actacataag tatcatgcc 7037
 aaaccagacg ggttaacggc agcaaaagac cttgcagagg cgttcgagca ttataacgaa 7097
 tggcatccgc atagtgcgct gggttatcgc tcgccacggg aatatctgcg gcagtgggcc 7157
 agtgatgggt taagtataaa caggtatcgt gaaacatagg ggcaaatcca acatgctaaa 7217
 aaacttaaca ttgatcata tattaagtct ttcaaaagaa gaagataaaa ttaacttgt 7277
 acaattaat gttaactcatt tagatgagag aacattaago tgtataaaaa atactctac 7337
 tggtaaggga tttaatgctc atttaaaaa acttgaactt ttgacctat ggttgagtga 7397
 gtattttgaa tatattatta tacctaataa gttaagcaat gcagggactt tttattttgc 7457
 attctttttt cccgagtttt atattaaaag attcaataag aataatactg atctttctc 7517
 gttaggagac acatctttta aacgacttat gactcgacca catataccaa actatgttta 7577
 caatcttggt ataaactcta atggatgtac ttttaactcc attaaattat tattgctgg 7637
 tcttagtcta acatcaaaaa ggttttatga aacacctcag caagaacgta attttttgtg 7697
 tcatataaat gaaattgtot tggctaastgc tgacgaatac tccggtatca tttcttgcat 7757
 tataaaaaagt agaataatcgt taattgatga ctttatitca agtaatgttt cattaaatac 7817
 aaacaggcaa atagctttat tcataactgg acaatcaaga ggttttatag atgccctaco 7877
 taatctcgta agtgaataaa cgaattcctc tgatgttgat gtttttatta gtacatggaa 7937
 agatctcgga cacacacagt tatctaagga aagaatatgt aggatatttg acagcgaggc 7997
 tgcacaatat gtttcagagc cagataacta ctggtttgta gacgaacact atgatgaat 8057
 aaaagacttg agcttaagtt catataaaaa taataattta gagggagat atcaagttt 8117
 tttctctgga tgaactcag ttttaataaa cattaaagat gatggggaat atccataaa 8177
 taaaatgagt aatgcagaaa aaatgtatta ccataattca ttttggttct gtagtcttaa 8237
 aaatcataat tgggataaat ataggtgcac tataagata aggcctgatg ctttattgca 8297
 agtggataat gtgacaatta atgataga tgtagatgat tctgtttatt gtgaggatag 8357
 taatgggtgg atatttagag aatgggggtt tggcataggc gaccaattat ttacggcga 8417
 tctgatata atgaaaaagc tgatgtgtgt ccatggttta gaaaaaatat atagtcagct 8477
 aacatccttg atctcaagtt ctaatgttta ttattcaggt catattaatg tagggttatg 8537
 tgccttgggt aatgtatatg attgccaggt ttctaattta aaaataaaaa atattgttag 8597
 cgcttcgaaa aatatcgcta gaacaaatac ttcttttgcg ggaatgagtt atttattgca 8657
 aggtatagat atcacttaac attgaaagat gcactatatg aaaaaataa ttgtagatt 8717
 agtaaacacc atatctttta tttatcagg aaaaatttca catgcaactc caaacaaaa 8777
 actaatlgaa aaattgtatg aatataagct taatggtttt tatattgtta tttttacagc 8837
 aaggaaatag aggcacataa aagaaaatat aggttaagatt aatattcata cattaccagt 8897
 tataattgat tggttgaatg aaaatagagt cccttatgat gaggtgattg ttggtaaac 8957
 ttgggttgga gatgaggggt tttatgttga tgatagagct attcgaccat cagaacttg 9017
 caatatgacc ttagaagaga ttctaatat gtagaagac gagaaaaaat gcttctaata 9077
 atgtctggtt cctatgttca acaagaatta ggggccgaat ttggttctat tctccaagc 9137
 ttcttctctt tagctaataa acgattatit aagcatcaag tatctttagg gcagatggt 9197
 catgcaatat atctgggttt accggaagat tttgtgtttg acaaacatga ttatgaatgg 9257
 ttgcttcgta ataaagtaac aatgatccct gtcgatagta acttgacatt agggcaagcg 9317
 atagttaccg catggaattt aataggagat aaagatgaca aaggcttaca attattgtt 9377
 ggcatcac tcttataaaa aattcctgca gggggggaat tagtagcaaa aagtcacct 9437
 gatgaaaatt atcaatgggc cattttttac gaaacagagt taagagccgt cagtagagaa 9497
 gataataaaa atgtaatttg tgggtatttt tcttttagaa aaccgaattt tttattagg 9557
 gaattagta cttcaaaatt tgattttacg gcggcactta aaaagtatca cgacagctat 9617
 agtttagct ctatatacgt gtctgattgg cttgattttg gacatattaa tacatactat 9677
 aagtcaaaag tacaatacac aaccagcgt gcatttaagt aattatgcat tacaacaaa 9737
 tccgttatca aatcaagttc aatgaaagt aaaattgaag ctgaatcaaa atggtttgaa 9797

actattcccc gagaattaaa gatctatact ccaatgttat tggaaaccgtt tgatcatatc 9867
 agaagagatt ataagcttga atattttatat aatacgacgt taaatgaatt atttgttttt 9917
 tctcgccctac caaataatat ttttaacaaat atattaataa gttgttttaga cttcatcgat 9977
 ctgtgcacaa aatatcattc aattgatact gacaaaaata tactgcaaga tttattttat 10037
 gaaaaaacga ttgagcgggt tagcaagtag ataacagatt taaatattga tccaaatgca 10097
 aaatggaaat ttaataataa tataagcgtt tcaattaatg atattcctta tgatactaata 10157
 aaattiatcc caagtgaact gcaatataaa actattatgc atggcgattt atgctttagt 10217
 aatataattt ttaacttttag aactggtaga atacaagttt ttgatcccag aggattgaac 10277
 cactctggag aaataagtat ttatgggtat tttcgttatg atatagctaa attatcacat 10337
 tcaatactag ggctctatga ttggataatt gcaggatatt atataataaa taaaaaaaat 10397
 aaactcata gttattgaatt caaaattaat attgataata aattgtttga aattcaatca 10457
 acatttgttt ctataataaa agagaataat tcaatctccg aaaaatcatt gtatgcgatg 10517
 caaatacatt tatttttatc aatgcttccc cttcattccg atgacaaaaa aaggcaagat 10577
 gcactatttg ctaatgcatt tagattatat gaaattttta aggaggctgc agtatgatta 10637
 taattccaat ggccgggagtg agttcgcgtt ttttcaaggc tggatattcc aaaccaaaat 10697
 atatgcttga attgaatggg gaggtttctat tcgatctatg ttgaaaagt ttcaaatatt 10757
 attttgagac tgaacatttt gtctttatcc ttagggtatg tttcaatagc aagtcttttg 10817
 tattacaaag aatagcatct ttagggtatta atagctacac cttgattact cttgataaag 10877
 aaactcgggg gcagcagaa acagtatat tggctatatc aaaattattt aatatagaac 10937
 aaccaatcac tatttttaac attgatacaa ttaggcctaa ttttatattt actaagttcg 10997
 aaggggaaaa tgaatgttat attgaagtat ttcgaggaga tggggataac tggctctttg 11057
 ttatgccatc aaatgatgta aaaaatgagg tcatgtctac tagcgaaaaa aaacaaattt 11117
 ctaacttatg ctgcacagga ttatatcatt tttctacaa taaaaatttt atttcagcat 11177
 atgaacatta taaaaatcta cctcaagaaa attgggatgc tgggagagt atatatagcc 11237
 ccaatataca attatctaata tagtaatggg atcaaatgtt attatacaga aataaataag 11297
 tctgatgtta ttttttgttg tactccatga gaatatgaaa atttgcaagg aaaaaataaa 11357
 aaattagggt tggcctaata aatctgataa ttattgtta tcaatgotta atcacattto 11417
 tttgatttta tacacgggaat taaatataat ctattatgaa aattgcagtt gctgggttag 11477
 gatattgttg tatatcaatt gctatattac tttcacaaaa acatgatatt atcgctctcg 11537
 atatagatcc taagaaagtt cagttgatta ataaaaaat atcaccaata tgtgatcccg 11597
 aaatacaaaa atttttatct aatagaaaat taaacctata tgcatacaaa gaaaaatacg 11657
 aagcgtatag agatgctgat tatgttataa tgcgaacacc aaccaattat gatoccatia 11717
 ataataactt cgatcacatc tcagtagaat cagtagcatg tgacgtacta agtataaatc 11777
 ctaatgcaac tatcataatt aaatctacag tccccgtcgg atttactgaa cgactaaaac 11837
 gcgatctaaa cacgaataat attatctttt cccagaatt tttacgtgaa ggtaaaagtc 11897
 tttatgacaa cctatatcca tctcgtatag ttgtgggaga gagtagcgaa cgagcaagna 11957
 agttcgcaga gcttctcagt gaaggcgcta taaaaaaga tattccaata ttgttaacgg 12017
 atagccctga agctgaagcc attaaacttt ttgcaaatat ttaccttgca atgaggattg 12077
 cttatttcaa tgaattggat acttatgcct ccgttcattg tttagataca aagcaaatat 12137
 tagagggtgt tagtttagat cctagaattg gtcaacatta taataatcct tcttttggtt 12197
 atggagggtta ctgcttacct aaggatacca agcaatcact cgcaaatat cgtgatgttc 12257
 cgagaaactt aatccaggct attgtcagtg ccaatactac ccgaaaagac tttgttcgag 12317
 aggatataat aagtcgtaaa ccaaaagtgt taggaatota tgcctcata atgaaagcag 12377
 gtagtgataa cttagagca agtagtatte aaggtgtaat gaaacgactc aaagccaaag 12437
 gaattgagat agttgtatat gaacctgtac taaaagagcc ttatttcttt ggttcttatg 12497
 ttgagcgtga tattaattct tttaaagaac gtgttgatgt tatagtagcc aatcgccgca 12557
 cgtcagaatt agaagatgta agtgaaaaag tttatccgag agattttatt ggtgtcgact 12617
 cttgattatg ttcaataatt taaaattttt atggctatta aagaaaagtc gatattgaca 12677
 tgcttttagct gcaatacaag atgactgccg attttggcag tcaaaacgta tattggcaat 12737
 gtacaggctt aatatgtatt ggtcattaca taactctact gatcacccgt cagattggcg 12797
 gtgtaaatta gcaatcaaaa tagctaaaaa tgcctgttgt gcataagct taactccgga 12857
 attactcatg gaggtttaag acgagttcac agcacacacat caaaaagtgt agttagcaaa 12917
 aaccttagca tcatactctc ctacgttttc attatcatta ttagataatg tgcataattg 12977
 tcccttagac ttgtatcagc tcttcaatta agaactcggg taactcaaaa agctatatca 13037
 acactcgcto agattgatgc cagtgatatt gtatatccc ctgatataat actcttgcaa 13097
 aaataatgct ttcagagaaa cggcagaat ctctgttaat agacttaacg aatactataa 13157
 gtacttttgt ttatctccgg tcgcactgac agataactca tctcctttgt caccctgtaa 13217
 tattattaca tcatctctt atcctgcccc aacgggcccc ctgatttcta ttttaatgac 13277
 aacatacaat accggtaggo gggtagaaaa tgcagtaata tcattgctta atcaaacata 13337
 ccgttcattt gagctaatta ttgtggatga tgccagcacc gatgatacgc tattccgtct 13397
 tcagagatta gcactcaaa atactcgaat aaaaattatt agcctgccac aaaaattgtg 13457
 aacatatgct gcaaaacgaa taggcttaat acaggcuuag ggagagtttg tgacatgcca 13517

```

tgactcagat gactgggtccc atcctgaaaa attatttaga cagatatac ctttattgtt 13577
aaaccctaaa ctaatttgtt cgatttctga ttgggtaagg ttgcaagata atgggatttt 13637
ctatgcgcgt gcggtctatc cactaaaaag actgaatcct tcttctctgt tgtttagaag 13697
agcggtgtga ttgcaaaaag cgggcgtttg ggactgtgtt aaaacggggg ctgatatga 13757
attcattgct cgacttaagc taatttttgg tgattccact gtacatcgta ttaaatggc 13817
tttgacgcta ggaagccatc gtaccgactc gttaatgaat tcacctacaa caggatatac 13877
atctcagga atttoaccag atcgccaaaa atattgggat tcctggctgc gatggcaccat 13937
tcaggcgta agaaataaag aaagtcttta cataggaaat tctgatttca ctataaaaa 13997
tcgaccattt totgcgcccg actcaatatt agtagatact aatgccatca aaactgcatt 14057
acaaagtgtc catgttaact ttaccagtat ataacttctc actaaatgta taatctataa 14117
tatttatttt aataatttat tgtgttttct aattatagta tgtaatcat ttatttaatg 14177
aatgggagtt taagaatggg tatgttaatg ccattcattg tgatgtttt ttgatagcat 14237
aatacaatct tttttatctt tttttatatt tttttatctt gttaaaaato atctcccat 14297
aataaaccgc attaacctgc gtcttaacca ataaataccc tcgaaacttc ttaaacagtt 14357
tcattatcgg tttaaaatcg gcctgccaga actgggtgcaa atgcccttga tacgtcaage 14417
ctttgatgtc gtacagggca ttgcccatca ctttgagtgg tttgttatga atcagcgcag 14477
agatcc 14483

```

<210> 4

5 <211> 108

<212> ADN

<213> Secuencia artificial

<220>

10 <221> CDS

<222> (1).. (108)

<223> Descripción de la secuencia artificial: secuencia que codifica la secuencia de aminoácidos His-tag

<400> 4

15

```

atg ggg ggt tct cat cat cat cat cat cat ggt atg gct agc atg act 48
Met Gly Gly Ser His His His His His His Gly Met Ala Ser Met Thr
  1             5             10             15

```

```

ggg gga cag caa atg ggt cgg gat ctg tac gac gat gac gat aag gat 96
Gly Gly Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp
      20             25             30

```

```

cga tgg atc ccg 108
Arg Trp Ile Pro
  35

```

<210> 6

<211> 27

20 <212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: secuencia del cebador CS-S

25

<400> 5

accaaacact gctacaacct atacgg

27

30 <210> 6

<211> 30

<212> ADN

<213> Secuencia artificial

35 <220>

<223> Descripción de la secuencia artificial: secuencia del cebador CS-AS

<400> 6

40 gcgtcttcac caataaatac ccacgaaact

30

<210> 7

	<211> 26	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> Descripción de la secuencia artificial: secuencia del cebador TM-S	
	<400> 7	
10	cgagaaatac gaacacgctt tggtaa	26
	<210> 8	
	<211> 28	
	<212> ADN	
15	<213> Secuencia artificial	
	<220>	
	<223> Descripción de la secuencia artificial: secuencia del cebador TM-AS	
20	<400> 8	
	actcaatttt ctcttcagc tcttcttg	28
	<210> 9	
25	<211> 30	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
30	<223> Descripción de la secuencia artificial: secuencia del cebador K4C-SP	
	<400> 9	
	cgggatcccg atgagtattc ttaatcaagc	30
35	<210> 10	
	<211> 30	
	<212> ADN	
	<213> Secuencia artificial	
40	<220>	
	<223> Descripción de la secuencia artificial: secuencia del cebador K4C-AS	
	<400> 10	
45	ggaattccgg ccagctcaca tgtttatcac	30
	<210> 11	
	<211> 36	
50	<212> PRT	
	<213> Secuencia artificial	
	<220>	
	<223> Descripción de la secuencia artificial: secuencia de aminoácidos His-tag	
55	<400> 11	
	Met Gly Gly Ser His His His His His Gly Met Ala Ser Met Thr	
	1 5 10 15	
	Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp	
	20 25 30	
	Arg Trp Ile Pro	
	35	

REIVINDICACIONES

1. Proteína aislada seleccionada de entre el grupo constituido por los siguientes apartados (A) y (B):

(A) una proteína que comprende la secuencia de aminoácidos representada por la SEC. ID. nº 2;

(B) una proteína que comprende la secuencia de aminoácidos en la que 1 a 30 resto(s) de aminoácido(s) en la secuencia de aminoácidos representada por SEC. ID. nº 2 se eliminan, sustituyen, insertan o transponen, y que tiene actividad de condroitina-polimerasa.

2. Proteína según la reivindicación 1, en la que la proteína está en forma soluble.

3. ADN que comprende uno de los siguientes apartados (a) a (c):

(a) ADN que codifica una proteína constituida por la secuencia de aminoácidos representada por la SEC. ID. nº 2;

(b) ADN que codifica una proteína constituida por la secuencia de aminoácidos en la que 1 a 30 resto(s) de aminoácidos en la secuencia de aminoácidos representada por SEC. ID. nº 2 se eliminan, sustituyen, insertan o transponen, y que tiene actividad de condroitina-polimerasa;

(c) ADN que se hibrida a

(i) el ADN en el apartado (a), o

(ii) un ADN complementario al ADN en el apartado (a) en condiciones severas en las que la hibridación se realiza a 42°C en una solución que contiene 50% de formamida, 4 x SSC, HEPES 50 mM (pH 7,0), 10 x solución de Denhardt y 100 µg/ml de ADN de esperma de salmón, y que codifica una proteína que tiene actividad de condroitina-polimerasa.

4. ADN según la reivindicación 3, en el que el ADN en el apartado (a) está representado por la SEC. ID. nº 1.

5. Vector que comprende el ADN de la reivindicación 3 ó 4.

6. Vector según la reivindicación 5, que es un vector de expresión.

7. Transformante en el que un hospedador se transforma con el vector de la reivindicación 5 ó 6.

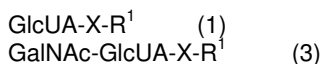
8. Procedimiento para producir una condroitina-polimerasa, que comprende:

cultivar el transformante de la reivindicación 7; y

recoger la condroitina-polimerasa del material cultivado.

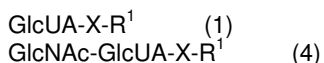
9. Agente para la síntesis de la cadena de azúcar, que comprende la proteína de la reivindicación 1 ó 2 y que tiene actividad de condroitina-polimerasa.

10. Procedimiento para producir una cadena de azúcar representada por la fórmula (3) siguiente, que comprende por lo menos una etapa que permite al agente sintetizador de la reivindicación 9 ponerse en contacto con un donante de GalNAc y una cadena de azúcar representada por la fórmula (1) siguiente:



en las que GlcUA representa ácido D-glucurónico; GalNAc representa N-acetil-D-glucosamina; X representa GalNAc o GlcNAc en la que GalNAc representa N-acetil-D-glucosamina; - representa un enlace glucosídico y R¹ representa cualquier grupo.

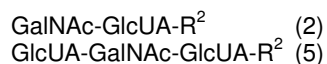
11. Procedimiento para producir una cadena de azúcar representada por la fórmula (4) siguiente, que comprende por lo menos una etapa que permite al agente sintetizador de la reivindicación 9 ponerse en contacto con un donante de GalNAc y una cadena de azúcar representada por la fórmula (1) siguiente:



en las que GlcUA representa ácido D-glucurónico; GalNAc representa N-acetil-D-glucosamina; X representa GalNAc o GlcNAc en la que GalNAc representa N-acetil-D-glucosamina; - representa un enlace glucosídico y R¹ representa

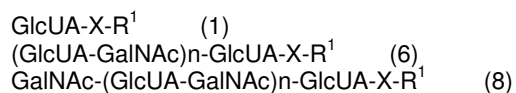
cualquier grupo.

12. Procedimiento para producir una cadena de azúcar representada por la fórmula (5) siguiente, que comprende por lo menos una etapa que permite al agente sintetizador de la reivindicación 9 ponerse en contacto con un donante de GlcUA y una cadena de azúcar representada por la fórmula (2) siguiente:



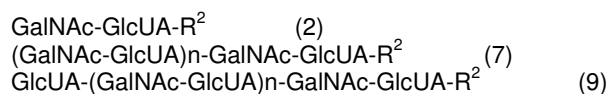
en las que GlcUA representa ácido D-glucurónico; GalNAc representa N-acetil-D-glucosamina; X representa GalNAc o GlcNAc en la que GlcNAc representa N-acetil-D-glucosamina; - representa un enlace glucosídico y R^2 representa cualquier grupo.

13. Procedimiento para producir una cadena de azúcar seleccionada de entre las fórmulas (6) y (8) siguientes, que comprende por lo menos una etapa que permite al agente sintetizador de la reivindicación 9 ponerse en contacto con un donante de GalNAc, con un donante de GlcUA y con una cadena de azúcar representada por la fórmula (1) siguiente:



en las que GlcUA representa ácido D-glucurónico; GalNAc representa N-acetil-D-glucosamina; X representa GalNAc o GlcNAc en la que GlcNAc representa N-acetil-D-glucosamina; - representa un enlace glucosídico; R^1 representa cualquier grupo; y n es un número entero de 1 ó más.

14. Procedimiento para producir una cadena de azúcar seleccionada de entre las fórmulas (7) y (9) siguientes, que comprende por lo menos una etapa que permite al agente sintetizador de la reivindicación 9 ponerse en contacto con un donante de GalNAc, con un donante de GlcUA y con una cadena de azúcar representada por la fórmula (2) siguiente:



en las que GlcUA representa ácido D-glucurónico; GalNAc representa N-acetil-D-glucosamina; - representa un enlace glucosídico; R^2 representa cualquier grupo; y n es un número entero de 1 ó más.

15. Sonda de hibridación que comprende una secuencia nucleotídica complementaria de la secuencia nucleotídica representada por SEC. ID. nº 1.

16. Catalizador de transferencia de glucosilo que comprende la proteína de la reivindicación 1 ó 2 y que tiene actividad de condroitina-polimerasa.

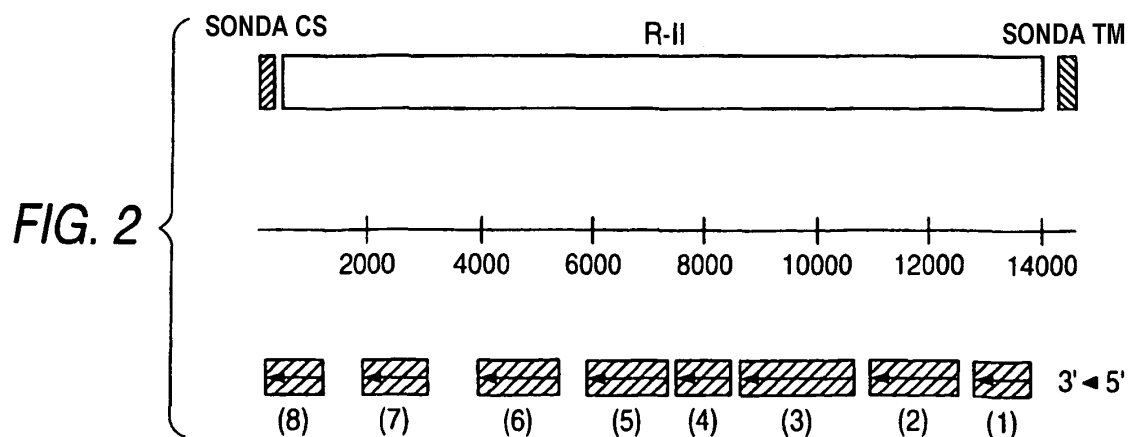
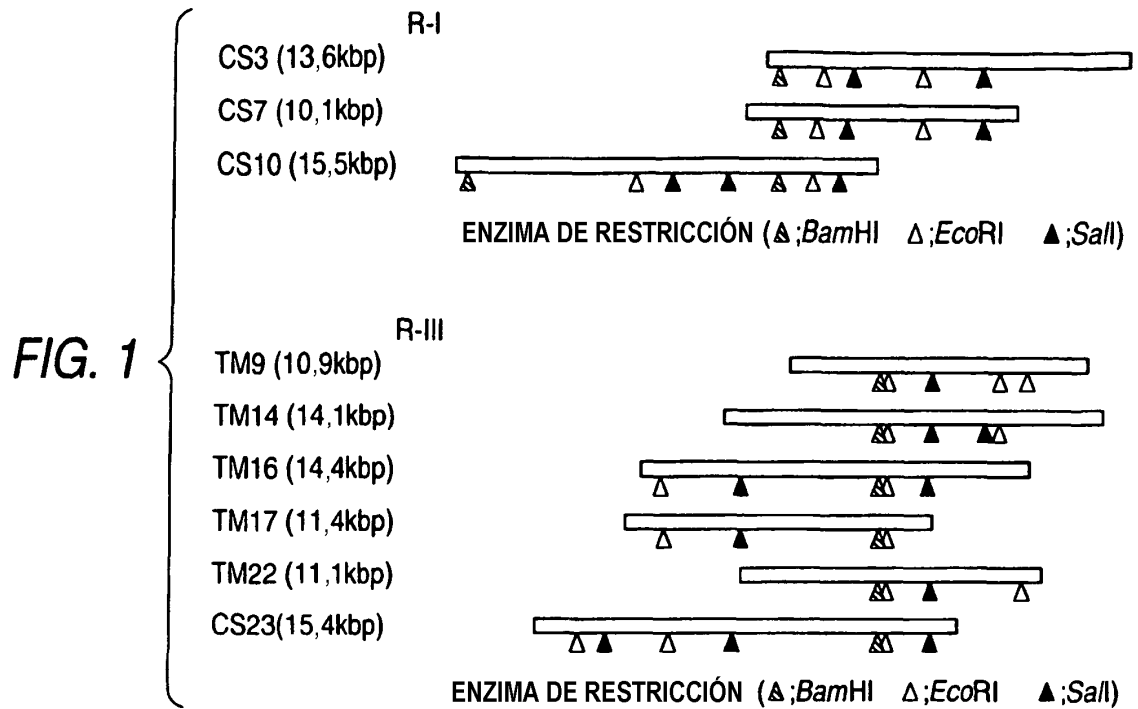


FIG. 3

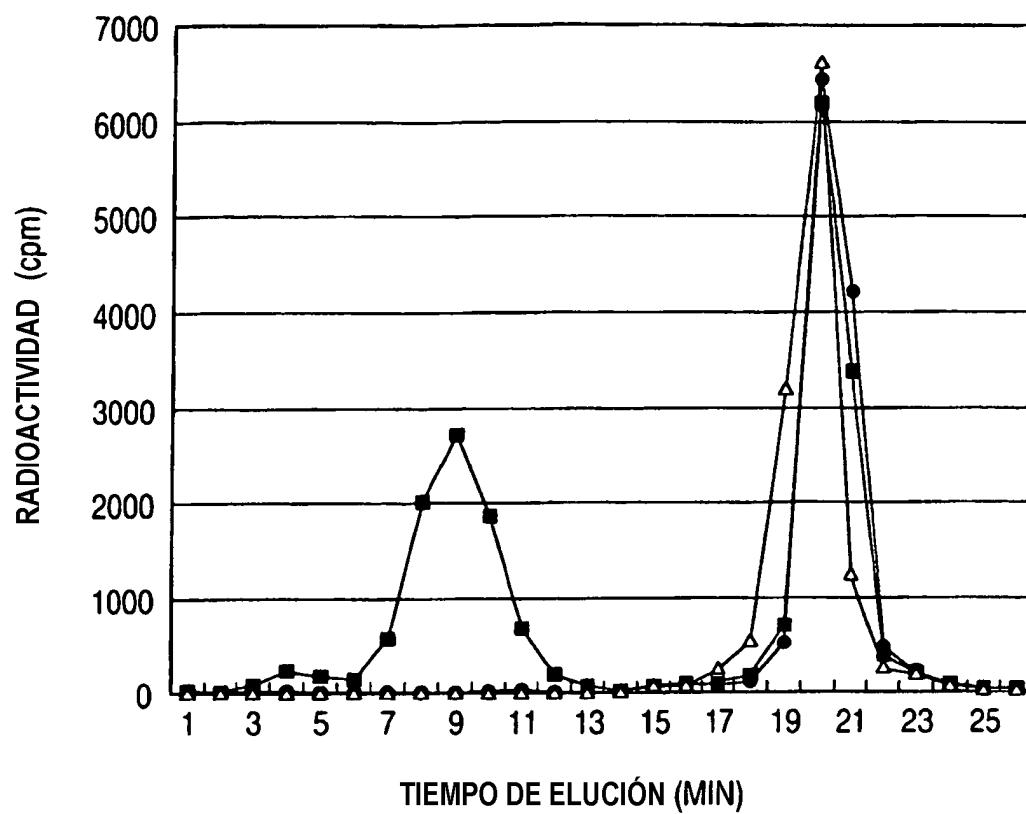


FIG. 4

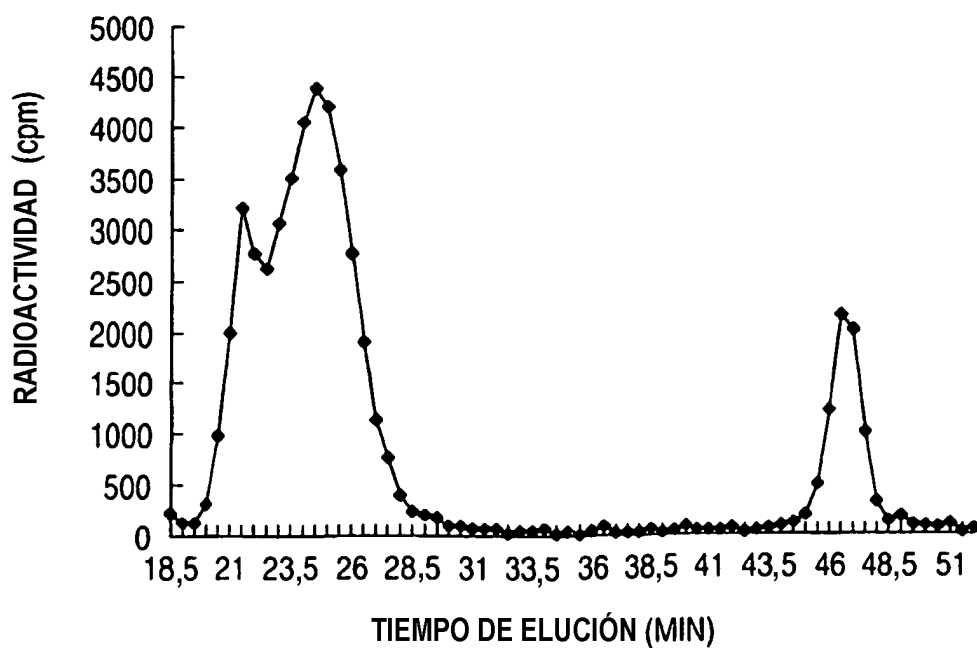


FIG. 5

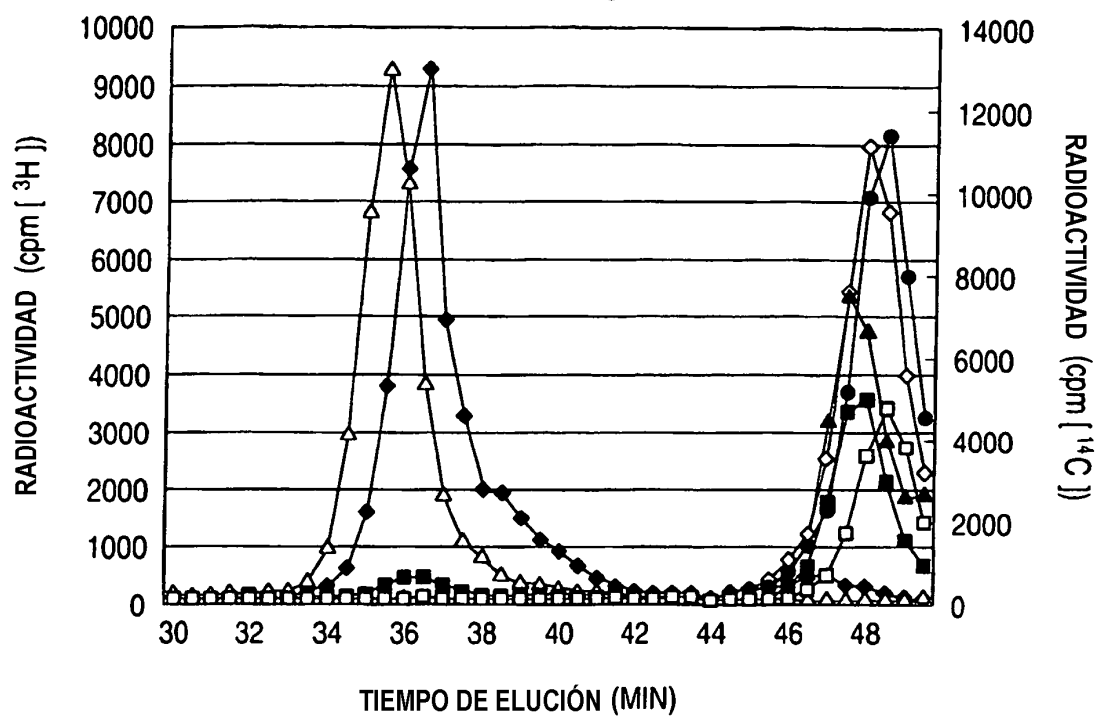


FIG. 6

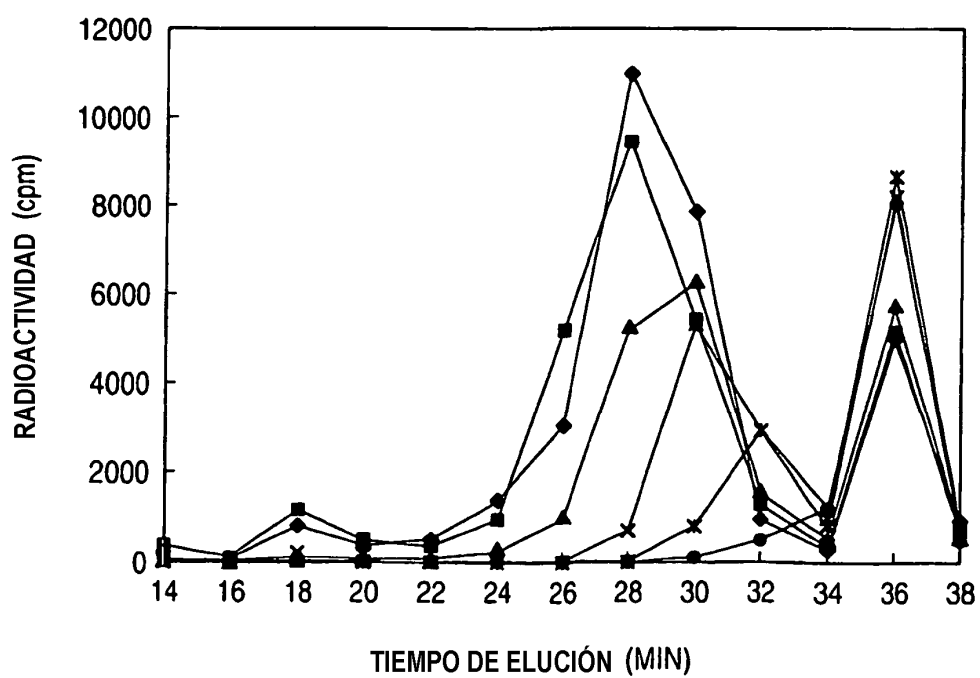


FIG. 7

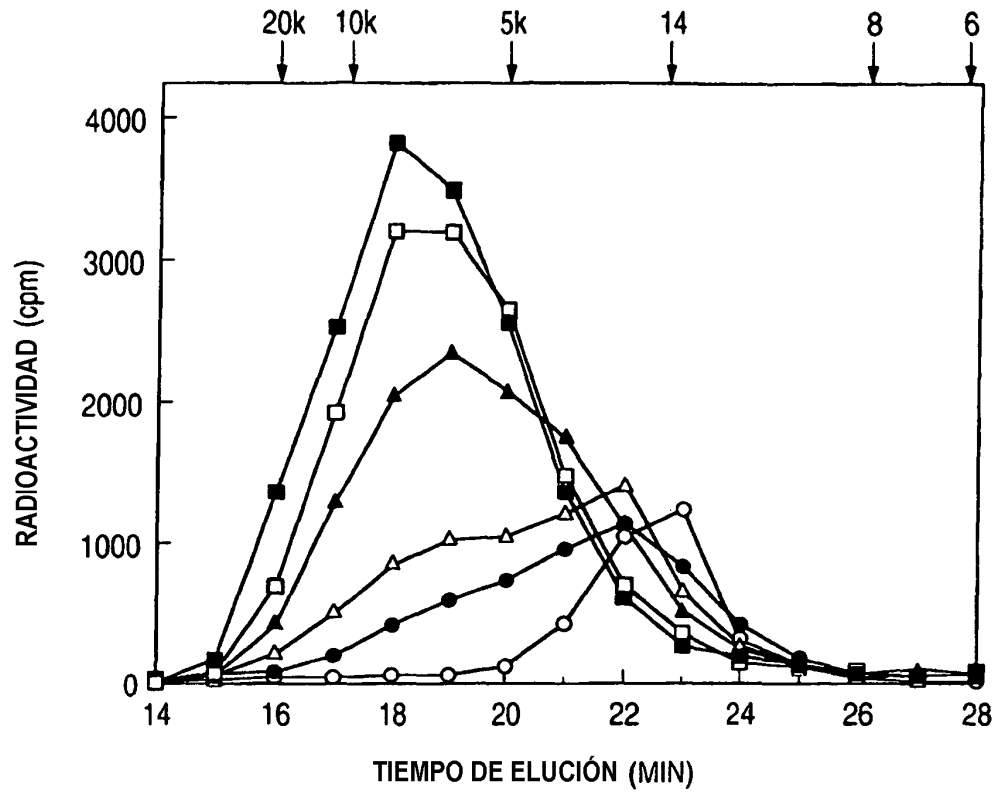


FIG. 8

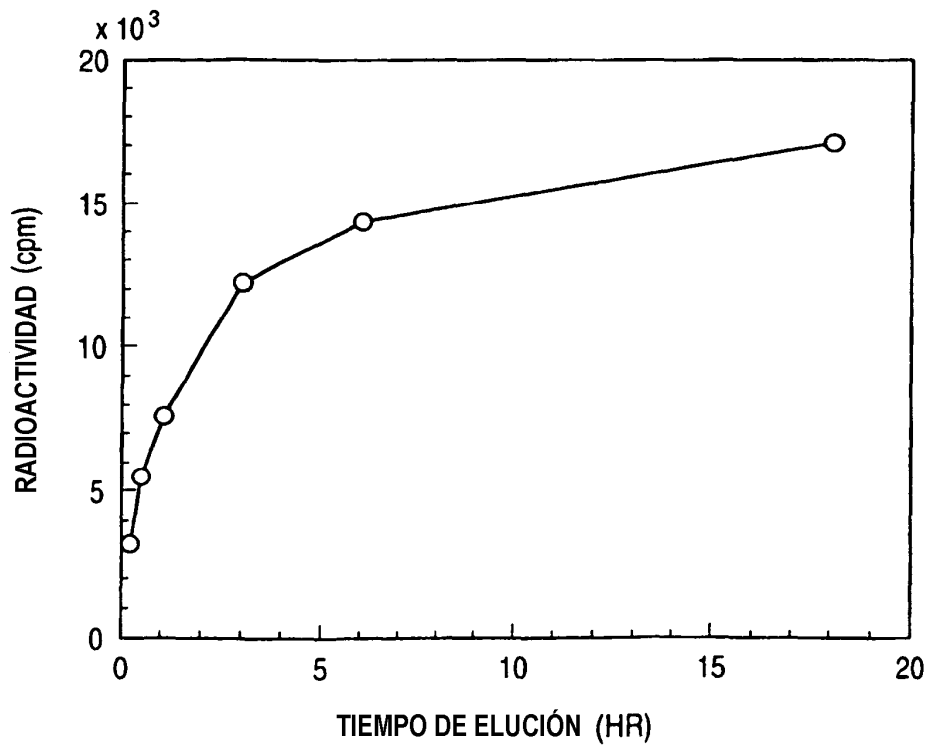


FIG. 9

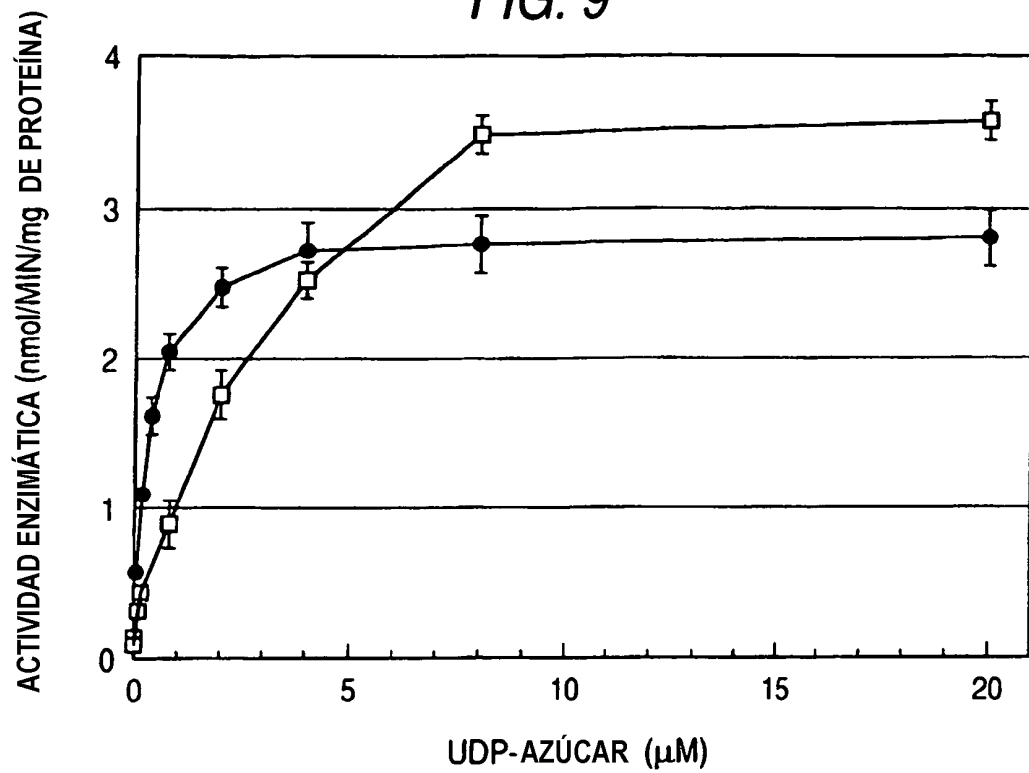


FIG. 10

