

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 381 504**

51 Int. Cl.:
C12P 21/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **07762625 .7**
- 96 Fecha de presentación: **01.02.2007**
- 97 Número de publicación de la solicitud: **1984013**
- 97 Fecha de publicación de la solicitud: **29.10.2008**

54 Título: **Modificaciones de CST-II para una expresión de proteína aumentada**

30 Prioridad:
01.02.2006 US 764499 P
01.02.2006 US 764625 P
15.02.2006 US 774088 P
15.02.2006 US 773941 P

45 Fecha de publicación de la mención BOPI:
28.05.2012

45 Fecha de la publicación del folleto de la patente:
28.05.2012

73 Titular/es:
BioGeneriX AG
Janderstrasse 3
68199 Mannheim , DE

72 Inventor/es:
JOHNSON, Karl, F.;
MOBASSERI, Aliakbar;
BAI, Xiaomei y
KINEALY, Kyle

74 Agente/Representante:
Arias Sanz, Juan

ES 2 381 504 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

Modificaciones de CST-II para una expresión de proteína aumentada

5 **Campo de la invención**

La presente invención proporciona proteínas sialiltransferasa de *Campylobacter* que muestran expresión aumentada comparadas con su forma sin modificar. También se incluyen los ácidos nucleicos que codifican las proteínas sialiltransferasa, como lo son los métodos para producir y usar las proteínas sialiltransferasas.

10

Antecedentes de la invención

Se reconoce ahora que los hidratos de carbono son de una importancia principal en muchos sucesos de reconocimiento célula-célula, notablemente la adhesión de bacterias y virus a células de mamífero en patogénesis y la interacción leucocito-célula endotelial a través de selectinas en inflamación (Varki (1993) *Glycobiology* 3: 97-130). Además, se piensa que los gliconjugados sialilados que se encuentran en bacterias (Preston *et al.* (1996) *Crit. Rev. Microbiol.* 22:139-180; Reuter *et al.* (1996) *Biol. Chem. Hoppe-Seyler* 377:325-342) imitan a oligosacáridos encontrados en glicolípidos de mamíferos para evadir la respuesta inmune del huésped (Moran *et al.* (1996) *FEMS Immunol. Med. Microbiol.* 16:105-115). Se considera que la imitación molecular de estructuras del huésped por la parte sacárido del lipopolisacárido (LPS) es un factor de virulencia de varios patógenos de las mucosas, que usan esta estrategia para evadir una respuesta inmune del huésped (Moran. *et al.* (1996) *FEMS Immunol. Med. Microbiol.* 16: 105-115; Moran *et al.* (1996) *J. Endotoxin Res.* 3: 521-531).

Las estructuras de oligosacáridos implicados en estos y otros procesos son potenciales agentes terapéuticos, pero se necesita mucho tiempo y son caros para hacerlos por medios químicos tradicionales. Una vía muy prometedora para la producción de estructuras de oligosacáridos específicos es mediante el uso de enzimas que los hacen *in vivo*, las glicosiltransferasas. Tales enzimas se pueden usar como catalizadores regio- y estereoselectivos para la síntesis *in vitro* de oligosacáridos (Ichikawa *et al.* (1992) *Anal. Biochem.* 202: 215-238). Las sialiltransferasas son un grupo de glicosiltransferasas que transfieren ácido siálico de un nucleótido azúcar activado a oligosacáridos aceptores encontrados en glicoproteínas, glicolípidos o polisacáridos. El gran número de estructuras de oligosacárido sialiladas ha llevado a la caracterización de muchas sialiltransferasas diferentes implicadas en la síntesis de varias estructuras. Se han aislado y caracterizado sialiltransferasas de mamíferos y otros eucariotas y de microbios incluyendo *C. jejuni*, *Neisseria*, *Haemophilus*, y *E. coli*. (Tsuji *et al.* (1996) *Glycobiology* 6:v-vii; patentes en EE UU No. 6.503.744; 6.699.705; 6.096.529; 6.210.933; y Weisgerber *et al.* (1991) *Glycobiol.* 1:357-365).

La síntesis enzimática de oligosacáridos a gran escala depende de la disponibilidad de cantidades suficientes de las glicosiltransferasas requeridas. Sin embargo, la producción de glicosiltransferasas en cantidades suficientes para su uso en la preparación de estructuras de oligosacáridos ha sido problemática. La presente invención resuelve este y otros problemas.

40

Breve compendio de la invención

En un aspecto, la presente invención proporciona una proteína sialiltransferasa con etiqueta que incluye una secuencia de aminoácidos con al menos el 95% de identidad a SEQ ID NO:4 y una etiqueta MGS en el extremo amino de SEQ ID NO:4. La proteína sialiltransferasa con etiqueta tiene actividad sialiltransferasa.

45

En otro aspecto, la presente invención proporciona un ácido nucleico que codifica la proteína sialiltransferasa con etiqueta anterior. En una forma de realización, la invención proporciona un vector de expresión que incluye el ácido nucleico. En otra forma de realización, la presente invención proporciona un vector de expresión de SEQ ID NO:3 para expresar una proteína sialiltransferasa con etiqueta. En una forma de realización más, la invención proporciona una célula huésped que comprende el vector de expresión.

50

En otro aspecto la invención proporciona un método de hacer una proteína sialiltransferasa con etiqueta, cultivando la célula huésped anterior en condiciones adecuadas para la expresión de la proteína sialiltransferasa con etiqueta. El método de hacer la proteína sialiltransferasa con etiqueta también puede incluir un paso de purificación de la proteína sialiltransferasa con etiqueta.

55

En otro aspecto, la presente invención proporciona un método *in vitro* para producir un producto de oligosacárido sialilado, poniendo en contacto un sustrato aceptor que comprende un oligosacárido con una sialiltransferasa con etiqueta de antes, y un sustrato donante de ácido siálico; y después permitir la transferencia de ácido siálico del sustrato donante al sustrato aceptor para producir el producto de oligosacárido sialilado. La producción del producto de oligosacárido sialilado ocurre *in vitro*.

60

Breve descripción de las figuras

- 5 La figura 1 proporciona un mapa de restricción del plásmido pNT1, el vector pcwin2 construido para expresar una proteína CSTII con etiqueta MGS. El recuadro muestra las secuencias de aminoácidos N-terminales y de ácido nucleico 5' de la proteína CSTII en los vectores de expresión pNT1 y pNT2.
- 10 La figura 2 proporciona la secuencia de aminoácidos traducida completa de la proteína CSTII expresada del vector de expresión pNT1. La etiqueta MGS amino-terminal está entre paréntesis.
- La figura 3 proporciona los componentes de medios rico y simple.
- La figura 4 proporciona un esquema del proceso de purificación de CST-II.
- 15 La figura 5 proporciona un gradiente en pasos en Q-Sepharosa. Columna: XK26 con resina Q-Sepharosa FF (40 ml). Equilibrio: 5 volúmenes de columna (VC) Tris 20 mM, pH 8,3 (tampón A). Carga de 130 ml de homogenizado clarificado (13 g de precipitado sin EDTA). La columna se lavó con 5 VC de tampón A. Elución en gradiente en pasos con tampón B (NaCl 1 M en tampón A): 4 VC de tampón B al 15%, 4 VC de tampón B al 25%, 4 VC de tampón B al 50%. Velocidad de flujo: 115 cm/h. Absorbancia a 280 nm. Se tomaron muestras de las fracciones para SDS-PAGE (figura 4) y ensayo enzimático de CST-II. El histograma de la actividad enzimática de CST-II se representa en el cromatograma en rojo. Se juntó la fracción 5 de elución de Q-Sepharosa para una purificación posterior. FT = flujo que pasa (Flow Through).
- 20 La figura 6 proporciona el análisis por SDS-PAGE de las fracciones del gradiente por pasos de Q-Sepharosa. Se muestran las muestras de las fracciones de Q-Sepharosa de la figura 3. Est = See Blue Plus 2 (15 µl). Carril 1: Homogenizado de CST-II (10 µl). Carril 2: fracción 1 del flujo que pasa (10 µl). Carril 3: fracción 2 del flujo que pasa (10 µl). Carril 4: fracción 3 del flujo que pasa (10 µl). Carril 5: fracción 4 del flujo que pasa (10 µl). Carril 6: fracción 5 del flujo que pasa (10 µl). Carril 7: fracción 6 del flujo que pasa (10 µl). Carril 8: Mezcla flujo que pasa/lavado (22,5 µl). Carril 9: Fracción de elución 1 (14 µl). Carril 10: Fracción de elución 2 (2,8 µl). Carril 11: Fracción de elución 3 (2,4 µl). Carril 12: Fracción de elución 4 (3 µl). Carril 13: Fracción de elución 5 (1,5 µl). Carril 14: Fracción de elución 4 (4,4 µl). Carril 15: Fracción de elución 7 (4,4 µl). Carril 16: Fracción de elución 8 (11,8 µl). Carril 17: Fracción de elución 9 (3 µl). Carril 18: Fracción de elución 10 (1,8 µl). Los geles de SDS PAGE de Tris-glicina del 4-20% se tiñeron mediante Simply Blue Safe Stain. La fracción de elución 5 de Q-Sepharosa (mostrada en el carril 13) se recogió para purificación posterior.
- 25 La figura 7 proporciona los gradientes en pasos en Q-Sepharosa con y sin dilución de la solución de carga. Columna: XK16 Q-Sepharosa FF (20 ml). Homogenizado clarificado cargado (2 g con EDTA 5 mM, 20 ml). Figura 7A: sin diluir, conductividad: 4,66 ms/cm. Figura 7B: diluido 1:1 con tampón A (Tris 50 mM, pH 8,3), conductividad: 3,3 ms/cm. Las columnas se lavaron con 5 CV de tampón A. Elución de gradiente en pasos con tampón B (NaCl 1 M en tampón A): 5 VC de tampón B al 15%, 5 VC de tampón B al 100%. Velocidad de flujo: 115 cm/h. Absorbancia a 280 nm. El histograma de la actividad enzimática de CST-II se representa en el cromatograma en rojo. Las fracciones de CST-II se juntaron como se indica.
- 30 La figura 8 proporciona el análisis por SDS-PAGE de las mezclas de fracciones de las purificaciones de CST-II en Q-Sepharosa. Est = See Blue Plus 2 (15 µl). Carril 1: mezcla de fracción del gradiente en pasos en Q-Sepharosa del homogenizado sin diluir (figura 5A) (10 µg). Carril 2: mezcla de fracción del gradiente en pasos en Q-Sepharosa del homogenizado diluido (figura 5B) (10 µg). Carril 3: mezcla de fracción del gradiente lineal en Q-Sepharosa del homogenizado diluido (figura 7) (10 µg). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain.
- 35 La figura 9 proporciona elución con gradiente lineal en Q-Sepharosa. Columna: XK16 con resina Q-Sepharosa FF (20 ml). Equilibrio: 5 volúmenes de columna (VC) Tris 50 mM, pH 8,3 (tampón A). Homogenizado clarificado (2 g de precipitado con EDTA, 20 ml diluido 1:1 con tampón A, conductividad: 3,3 ms/cm). Columna lavada con 5 VC de tampón A. Gradiente de elución lineal con tampón B (NaCl 1 M en tampón A): tampón B del 0-15% durante 10 VC seguido por 5 VC con tampón B al 100%. Velocidad de flujo: 115 cm/h. Absorbancia a 280 nm. Las fracciones se juntaron como se muestra. El histograma de la actividad enzimática de CST-II se representa en el cromatograma en rojo.
- 40 La figura 10 proporciona SP-Sepharosa HP- pH 6,8. Columna: XK16 con resina SP-Sepharosa HP (20 ml). Equilibrio: 5 volúmenes de columna (VC) de fosfato de sodio 25 mM, pH 6,8 (tampón A). CST-II purificada (Q-Sepharosa de la figura 3) (24 ml) diluida y ajustada el pH a 6,5 y filtrada por 0,2 micrómetros. Se cargó la CST-II condicionada (82 ml, 3,8 ms/cm, A280: 0,855 UA). La columna se lavó con 5 VC de tampón A. Gradiente de elución con tampón B (NaCl 1 M en tampón A): tampón B del 0-100% durante 10 VC. Velocidad de flujo: equilibrio, carga y lavado a 150 cm/h, elución a 80 cm/h. Absorbancia a 280 nm. Se tomaron muestras del flujo que pasa (FT) y las fracciones de elución 1-3 para SDS-PAGE.
- 45
- 50
- 55
- 60
- 65

5 La figura 11 proporciona el análisis por SDS-PAGE de las fracciones de SP-Sepharosa HP - pH 6,8. Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada por Q-Sepharosa (de la figura 3) (10 µg). Carril 2: Carga en SP-Sepharosa HP (figura 8) (19,8 µl). Carril 3: Mezcla flujo que pasa/lavado de SP-Sepharosa HP (23,6 µl). Carril 4: mezcla de elución 1 de SP-Sepharosa HP concentrada 13 x (30 µl). Carril 5: mezcla de elución 2 de SP-Sepharosa HP concentrada 4,8 x (30 µl). Carril 6: mezcla de elución 3 de SP-Sepharosa HP concentrada 3,3 x (30 µl). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain.

10 La figura 12 proporciona SP-Sepharosa HP - pH 6,0. Columna: XK16 con resina SP-Sepharosa HP (20 ml). Equilibrio: 5 volúmenes de columna (VC) de MES 25 mM, pH 6,0 (tampón A). La CST-II purificada (Q-Sepharosa de la figura 3) (20 ml) se diluyó, el pH se ajustó a 6,0 y se filtró por 0,2 micrómetros. Se cargó la CST-II condicionada (40 ml, 2,4 ms/cm, A280: 1,20 UA). La columna se lavó con 5 VC de tampón A. Gradiente de elución con tampón B (NaCl 1 M en tampón A): tampón B del 0-100% durante 10 VC. Velocidad de flujo: equilibrio, carga y lavado a 150 cm/h, elución a 80 cm/h. Absorbancia a 280 nm. Se tomaron muestras del flujo que pasa (FT)/lavado y las fracciones de elución 1-5 para SDS-PAGE (figura 11). Las fracciones 2-5 se juntaron para una purificación posterior.

15 La figura 13 proporciona el análisis por SDS-PAGE de las fracciones de SP-Sepharosa HP - pH 6,0. Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada por Q-Sepharosa (10 µg). Carril 2: Carga en SP-Sepharosa HP (figura 10) (10 µl). Carril 3: Mezcla flujo que pasa/lavado de SP-Sepharosa HP (30 µl). Carril 4: Fracción 1 de SP-Sepharosa HP (10 µg). Carril 5: Fracción 2 de SP-Sepharosa HP (10 µg). Carril 6: Fracción 3 de SP-Sepharosa HP (10 µg). Carril 7: Fracción 4 de SP-Sepharosa HP (10 µg). Carril 8: Fracción 5 de SP-Sepharosa HP (10 µg). Carril 9: Mezcla de elución de las fracciones 2-5 de SP-Sepharosa HP (10 µg). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain.

20 La figura 14 proporciona una comparación de hidroxiapatita de tipo I y tipo II. Columna: Tricorn 5 empaquetada con figura 14A: hidroxiapatita de tipo I (1 ml) – 40 micrómetros, figura 14B: hidroxiapatita de tipo II (1 ml – 40 micrómetros). Equilibrio: 5 volúmenes de columna (VC) de fosfato de sodio 5 mM, pH 6,5 (tampón A). Se diluyó CST-II purificada por Q-Sepharosa y SP-Sepharosa HP (de la figura 10) (1 ml, glicerol al 50%, A280: 1,22 UA) con fosfato de sodio 5 mM, pH 6,5 (3 ml) y se cargó. La columna se lavó con 10 VC de tampón A. Gradiente de elución de tampón B (NaCl 1,5 M en tampón A): B del 0-100% durante 20 VC. Velocidad de flujo: 153 cm/h (0,5 ml/min). Absorbancia a 280 nm. Se tomaron muestras del flujo que pasa (FT)/lavado y las mezclas de fracciones 1 y 2 para SDS-PAGE.

25 La figura 15 proporciona el análisis por SDS-PAGE de las mezclas de fracciones de CST-II de la cromatografía en hidroxiapatita (HA) de tipo I y tipo II (40 micrómetros). Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada por Q-Sepharosa (figura 10) (1,0 µg). Carril 2: Mezcla flujo que pasa/lavado de HA de tipo I (20 µl). Carril 3: mezcla de elución 1 de HA de tipo I (1,0 µg). Carril 4: mezcla de elución 2 de HA de tipo I (1,0 µg). Carril 5: Mezcla flujo que pasa/lavado de HA de tipo II (20 µl). Carril 6: mezcla de elución 1 de HA de tipo II (1,0 µg). Carril 7: mezcla de elución 2 de HA de tipo II (1,0 µg). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó con el kit de tinción de plata Wako.

30 La figura 16 proporciona cromatografía en hidroxiapatita de tipo II (40 micrómetros) – 10 ml. Columna: XK16 hidroxiapatita de tipo II – 40 micrómetros (10 ml). Equilibrio: 5 VC de fosfato de sodio 5 mM, pH 6,5 (tampón A). CST-II purificada por Q-Sepharosa (30 ml, glicerol al 50%, A280: 5,56 UA) diluida con fosfato de sodio 5 mM, pH 6,5 (120 ml). El pH se ajustó a 6,0 y se filtró por 0,2 micrómetros. Se cargó la CST-II condicionada (150 ml, conductividad: 1,65 ms/cm, A280: 1,043 UA). La columna se lavó con 5 VC de tampón. Elución en gradiente con tampón B (NaCl 1,5 M en tampón A): B del 0-100% durante 20 VC, seguido por 5 VC de fosfato de sodio 500 mM, pH 6,5). Velocidad de flujo: 119 cm/h (4 ml/min). Absorbancia a 280 nm. Se tomaron muestras de mezclas de fracciones para SDS-PAGE. FT = flujo que pasa.

35 La figura 17 proporciona el análisis por SDS-PAGE de las fracciones hidroxiapatita (HA) de tipo II (40 micrómetros, 10 ml). Est = See Blue Plus 2 (15 µl). Carril 1: Homogenizado de CST-II (10 µl). Carril 2: CST-II purificada por Q-Sepharosa (10 µg). Carril 3: Carga de HA (figura 14) (10 µl). Carril 4: Fracción 1 del flujo que pasa en HA (10 µl). Carril 5: Fracción 2 del flujo que pasa en HA (10 µl). Carril 6: Mezcla flujo que pasa/lavado de HA (12,7 µl). Carril 7: fracción de elución 1 de HA (20 µl). Carril 8: fracción de elución 2 de HA (30 µl). Carril 9: fracción de elución 3 de HA (30 µl). Carril 10: fracción de elución 4 de HA (30 µl). Carril 11: = See Blue Plus 2 (15 µl). Carril 12: fracción de elución 5 de HA (30 µl). Carril 13: fracción de elución 6 de HA (10 µg). Carril 14: fracción de elución 7 de HA (10 µg). Carril 15: fracción de elución 8 de HA (10 µg). Carril 16: fracción de elución 9 de HA (10 µg). Carril 17: fracción de elución 10 de HA (30 µl). Carril 18: Mezcla de elución de las fracciones 7-9 de HA (10 µg). Carril 19: fracción de elución 11 de HA (10 µg). Los geles de SDS PAGE de Tris-glicina del 4-20% se tiñeron mediante Simply Blue Safe Stain.

40 La figura 18 proporciona cromatografía en fenil Sepharosa de alta sustitución. Columna: Fenil Sepharosa de alta sus. (1,0 ml). Absorbancia a 280 nm. Muestra: CST-II purificada por Q-Sepharosa (condicionada en tampón B: NaCl 2 M, Tris-HCl 25 mM, pH 7,2), 1,0 ml, 4,2 mg. Gradiente: de 0-10 minutos, B al 100%, 10-40 minutos, B del 100-0%;

40-60 minutos, tampón A al 100% (Tris-HCl 25 mM, pH 7,2). Las fracciones 1-5 se recogieron y analizaron por SDS-PAGE. Las fracciones se almacenaron a 4°C.

5 La figura 19 proporciona el análisis por SDS-PAGE de las fracciones de la cromatografía en fenil Sepharosa (alta sustitución). Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada por Q-Sepharosa (10 µl). Carril 2: Fracción 1 de fenil de alta sus (figura 16) (5,0 µl). Carril 3: Fracción 2 de fenil de alta sus (5,0 µl). Carril 4: Fracción 3 de fenil de alta sus (5,0 µl). Carril 5: Fracción 4 de fenil de alta sus (5,0 µl). Carril 6: Fracción 5 de fenil de alta sus (5,0 µl). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó con el kit de tinción de plata Wako.

10 La figura 20 proporciona cromatografía en fenil Sepharosa (baja sustitución). Columna: Fenil Sepharosa de baja sus. (1,0 ml). Absorbancia a 280 nm. Muestra: CST-II purificada por Q-Sepharosa (condicionada en tampón B: NaCl 2 M, Tris-HCl 25 mM, pH 7,2), 1,0 ml, 4,2 mg. Gradiente: de 0-10 minutos, B al 100%, 10-25 minutos, B del 100-0%; 25-50 minutos, tampón A al 100% (Tris-HCl 25 mM, pH 7,2). Las fracciones 1-4 se recogieron y analizaron por SDS-PAGE. Las fracciones se almacenaron a 4°C.

15 La figura 21 proporciona el análisis por SDS-PAGE de las fracciones de la cromatografía en fenil Sepharosa (baja sustitución). Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada por Q-Sepharosa (10 µl). Carril 2: Fracción 1 de fenil de baja sus (figura 18) (5,0 µl). Carril 3: Fracción 2 de fenil de baja sus (5,0 µl). Carril 4: Fracción 3 de fenil de baja sus (5,0 µl). Carril 5: Fracción 4 de fenil de baja sus (5,0 µl). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó con el kit de tinción de plata Wako.

20 La figura 22 proporciona cromatografía en fenil Sepharosa (baja sustitución) con etilenglicol al 20%. Columna: Fenil Sepharosa de baja sus. (1,0 ml). Absorbancia a 280 nm. Muestra: CST-II purificada por Q-Sepharosa (condicionada en tampón B: NaCl 2 M, Tris-HCl 25 mM, pH 7,2), 1,0 ml, 4,2 mg. Gradiente: de 0-10 minutos, B al 100%, 10-25 minutos, B del 100-0%; 25-50 minutos, tampón A al 100% (Tris-HCl 25 mM, etilenglicol al 20%, pH 7,2). Las fracciones 1-4 se recogieron y analizaron por SDS-PAGE. Las fracciones se almacenaron a 4°C.

25 La figura 23 proporciona el análisis por SDS-PAGE de las fracciones de la cromatografía en fenil Sepharosa (baja sustitución) (con etilenglicol al 20%). Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada por Q-Sepharosa (10 µl). Carril 2: Fracción 1 de fenil Sepharosa de baja sus (etilenglicol) (figura 20) (5,0 µl). Carril 3: Fracción 2 de fenil Sepharosa de baja sus (etilenglicol) (5,0 µl). Carril 4: Fracción 3 de fenil Sepharosa de baja sus (etilenglicol) (5,0 µl). Carril 5: Fracción 4 de fenil Sepharosa de baja sus (etilenglicol) (5,0 µl). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó con el kit de tinción de plata Wako.

30 La figura 24 proporciona cromatografía en butil Sepharosa con etilenglicol al 20%. Columna: Butil Sepharosa (1,0 ml). Absorbancia a 280 nm. Muestra: CST-II purificada por Q-Sepharosa (condicionada en tampón B: NaCl 3 M, Tris-HCl 25 mM, pH 7,2), 1,0 ml, 4,2 mg. Gradiente: de 0-10 minutos, B al 50% de, 10-20 minutos, B del 50-0%; 20-30 minutos, tampón A al 100% (Tris-HCl 25 mM, etilenglicol al 20%, pH 7,2). Las fracciones 1 y 2 se recogieron y analizaron por SDS-PAGE (figura 23). Las fracciones se almacenaron a 4°C.

35 La figura 25 proporciona el análisis por SDS-PAGE de las fracciones de la cromatografía en butil Sepharosa (con etilenglicol al 20%). Est = See Blue Plus 2 (10 µl). Carril 1: CST-II purificada por Q-Sepharosa (10 µl). Carril 2: Fracción 1 de butil Sepharosa con etilenglicol (figura 22) (5,0 µl). Carril 3: Fracción 2 de butil Sepharosa con etilenglicol (5,0 µl). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó con el kit de tinción de plata Wako.

40 La figura 26 proporciona el análisis por SDS-PAGE del cribado de excipientes de CST-II (acetato de celulosa). Se intercambió el tampón de CST-II a tampones de fosfato de sodio que contenían varios excipientes (descritos en la tabla 2) usando filtros de centrifugación de acetato de celulosa. Se tomó una muestra de cada solución de CST-II con el tampón intercambiado (30 µl) para comparar y confirmar la recuperación de proteína. Est = See Blue Plus 2 (15 µl). Carril 1: Dilución y ajuste de pH iniciales de CST-II (30 µl). Carril 2: Tampón de CST-II intercambiado con fosfato de sodio 10 mM, NaCl 10 mM, pH 6,5 (30 µl). Carril 3: Tampón de CST-II intercambiado con NaCl 0,1 M, fosfato de sodio 10 mM, pH 6,5 (30 µl). Carriles 4, 6, 8, 11, 13, 15, 17: blancos. Carril 5: Tampón de CST-II intercambiado con NaCl 0,2 M, fosfato de sodio 10 mM, pH 6,5 (30 µl). Carril 7: Tampón de CST-II intercambiado con sacarosa 0,5 M, fosfato de sodio 10 mM, NaCl 10 mM, pH 6,5 (30 µl). Carril 9: Tampón de CST-II intercambiado con manitol 0,1 M, fosfato de sodio 10 mM, NaCl 10 mM, pH 6,5 (30 µl). Carril 10: Tampón de CST-II intercambiado con manitol 0,1 M, sacarosa 0,5 M, fosfato de sodio 10 mM, NaCl 10 mM, pH 6,5 (30 µl). Carril 12: Tampón de CST-II intercambiado con glicerol al 10%, fosfato de sodio 10 mM, NaCl 10 mM, pH 6,5 (30 µl). Carril 14: Tampón de CST-II intercambiado con sorbitol 0,1 M, fosfato de sodio 10 mM, NaCl 10 mM, pH 6,5 (30 µl). Carril 16: Tampón de CST-II intercambiado con Tween-20 al 0,02%, fosfato de sodio 10 mM, NaCl 10 mM, pH 6,5 (30 µl). Carril 18: Tampón de CST-II intercambiado con trehalosa 0,1 M, fosfato de sodio 10 mM, NaCl 10 mM, pH 6,5 (30 µl). Los gels de SDS PAGE de Tris-glicina del 4-20% se tiñeron mediante Simply Blue Safe Stain.

45 La figura 27 proporciona la purificación en Q-Sepharosa de CST-II (elución en gradiente por pasos) con y sin Tween-20. Superposición de cromatogramas que muestran la purificación en Q-Sepharosa con (cromatograma azul) y sin Tween-20 (cromatograma rojo de puntos) en las soluciones de carga y elución. Columna: XK26 con resina Q-Sepharosa FF (50 ml). Tampón A: Tris 50 mM, pH 8,3 (cromatograma azul: sin Tween-20, cromatograma rojo de

puntos: Tween-20 al 0,005%). Tampón B: NaCl 1 M en tampón A. Equilibrio de la columna: 5 volúmenes de columna (VC) de tampón A. Homogenizado clarificado cargado (5 g de precipitado con EDTA 10 mM (y para el cromatograma azul solo: Tween-20 al 0,02%), 50 ml, diluido con 150 ml de tampón A, conductividad: 3,12 ms/cm). Columna lavada con 5 VC de tampón A. Elución con gradiente por pasos con tampón B: 4 VC de tampón B al 15%, 4 VC de tampón B al 100%. Velocidad de flujo de 113 cm/h (10 ml/min). Absorbancia a 280 nm. Se recogieron mezclas de fracción 1 (Tween-20) y 2 (sin Tween-20) y se tomaron muestras para SDS-PAGE.

La figura 28 proporciona un análisis por SDS-PAGE de las purificaciones de CST-II en Q-Sepharosa con y sin Tween-20. Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada por Q-Sepharosa con Tween-20 (5,0 µg). Carril 2: CST-II purificada por Q-Sepharosa sin Tween-20 (5,0 µg). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain.

La figura 29 proporciona la purificación de CST-II en hidroxiapatita de tipo II – 20 micrómetros (azul) y 40 micrómetros (rojo). Superposición de cromatogramas que muestran la purificación de CST-II (después de Q Sepharosa con Tween-20, figura 25) en hidroxiapatita (HA) de tipo II de tamaño de partícula de 20 micrómetros (trazo azul) o tamaño de partícula de 40 micrómetros (trazo rojo de puntos). Columna: columna XK16 con resina hidroxiapatita de tipo II (10 ml) (cromatograma azul: 20 micrómetros, cromatograma rojo de puntos: 40 micrómetros). Equilibrio: 5 volúmenes de columna (VC) fosfato de sodio 5 mM, Tween-20 al 0,005%, pH 6,5 (tampón A). Carga: CST-II purificada por Q Sepharosa (5 ml de la figura 25). Columna lavada con 10 VC de tampón A. Elución en gradiente desde el 0-100% de tampón B (NaCl 1,5 M en tampón A) durante 25 VC, seguido por 5 VC de tampón B al 100% y VC de fosfato de sodio 500 mM, pH 6,5. Velocidad de flujo 4,0 ml/min (120 cm/h). Absorbancia a 280 nm. Se recogieron mezclas de fracciones 1 (20 micrómetros, azul) y 3 (40 micrómetros, rojo) como se muestra y se analizaron por SDS-PAGE.

La figura 30 proporciona un análisis por SDS-PAGE de las mezclas de fracciones de hidroxiapatita de tipo II (20 micrómetros y 40 micrómetros) antes y después de cromatografía en SP Sepharosa HP. Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada por Q Sepharosa e hidroxiapatita (HA) de tipo II (20 micrómetros) (figura 27, mezcla de fracciones 1, 5,0 µg). Carril 2: CST-II purificada por Q Sepharosa, hidroxiapatita (HA) de tipo II (20 micrómetros) y SP Sepharosa HP (figura 29, mezcla de fracciones 2, 5,0 µg). Carril 3: CST-II purificada por Q Sepharosa e hidroxiapatita (HA) de tipo II (40 micrómetros) (figura 27, mezcla de fracciones 3, 5,0 µg). Carril 4: CST-II purificada por Q Sepharosa, hidroxiapatita (HA) de tipo II (40 micrómetros) y SP Sepharosa HP (figura 29, mezcla de fracciones 4, 5,0 µg). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain.

La figura 31 proporciona cromatografía en SP-Sepharosa HP de CST-II parcialmente purificada por Q Sepharosa e hidroxiapatita de tipo II (20 micrómetros y 40 micrómetros). El cromatograma de la purificación en SP-Sepharosa HP de CST-II parcialmente purificada por Q Sepharosa e hidroxiapatita de tipo II (20 micrómetros) se muestra en azul. El cromatograma superpuesto (líneas rojas) representa la purificación en SP-Sepharosa HP de CST-II parcialmente purificada por Q Sepharosa e hidroxiapatita de tipo II (40 micrómetros). Columna: XK16 resina SP-Sepharosa HP (10 ml). Equilibrio: 5 volúmenes de columna (VC) con MES 25 mM, pH 6,0 que contenía Tween-20 al 0,005% (tampón A). Carga: CST-II parcialmente purificada (de la figura 27) ajustada al Tween al 0,02% y dializada contra MES 25 mM, NaCl 10 mM, pH 6,0 y filtrada por 0,2 micrómetros (42,3 ml de cada una, pH 6,0, conductividad: 3,6 ms/cm, A280: 0,261 UA). La columna se lavó con 5 VC de tampón A. Elución en gradiente con tampón B (NaCl 1 M en tampón A): tampón B del 0-50% durante 15 VC, tampón B al 50% durante 1 VC, 5 VC de tampón B al 100%. Velocidad de flujo 120 cm/h (4,0 ml/min). Absorbancia a 280 nm. Se recogieron muestras de mezclas de fracciones de CST-II 2 (HA 20 micrómetros, azul) y 4 HA (40 micrómetros, rojo) para SDS-PAGE.

La figura 32 proporciona la cromatografía en Superdex 200 de CST-II purificada. Columna: Superdex 200 10/30 GL. CST-II purificada (Q Sepharosa, hidroxiapatita de tipo II, SP Sepharosa HP, filtrada (Mustang E) y formulada en glicerol al 50%) almacenada a -20°C. Inyectados 250 µl (711 µg) en la columna. Fase móvil: fosfato de sodio 50 mM, NaCl 150 mM, pH 7,2. Velocidad de flujo 0,5 ml/min. Absorbancia a 280 nm. Se recogieron las mezclas de fracciones 1-3 y se analizaron por SDS-PAGE.

La figura 33 proporciona un análisis por SDS-PAGE de fracciones de Superdex 200 de CST-II. Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada: carga en Superdex 200 (10 µg). Carril 2: fracción 1 de Superdex 200 (figura 30) concentrada 17,6 x (30 µl). Carril 3: fracción 2 de Superdex 200 (10 µg). Carril 4: fracción 3 de Superdex 200 concentrada 16,3 x (30 µl). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain.

La figura 34 proporciona un análisis por SDS-PAGE para la actividad proteolítica en mezclas de purificación de CST-II. Se incubaron las mezclas de purificación de CST-II a 32°C durante 14 horas y se analizó la evidencia de actividad proteolítica por gel de SDS-PAGE (aumento en el número o intensidad de fragmentos proteolíticos). Est = See Blue Plus 2 (15 µl). Carril 1: CST-II purificada por Q Sepharosa, almacenada a -20°C (10 µg). Carril 2: CST-II purificada por Q Sepharosa, incubada a 32°C (10 µg). Carril 3: CST-II purificada por Q Sepharosa y HA de tipo II, almacenada a -20°C (10 µg). Carril 4: CST-II purificada por Q Sepharosa y HA de tipo II, incubada a 32°C (10 µg). Carril 5: CST-II purificada por Q Sepharosa, HA de tipo II y SP Sepharosa HP, almacenada a -20°C (10 µg). Carril 6: CST-II purificada por Q Sepharosa, HA de tipo II y SP Sepharosa HP, incubada a 32°C (10 µg). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain.

La figura 35 proporciona una cromatografía de exclusión molecular de CST-II homogenizada en varias condiciones (prevención de proteólisis). Columna: Superdex 200 10/30GL. Fase móvil: tampón de prueba de homogenización (descrito posteriormente) con NaCl 150 mM. Absorbancia a 280 nm. Muestra: inyectados 250 μ l de CST-II (filtrada por 0,2 micrómetros) homogenizada en los siguientes tampones: cromatograma de pH 4: acetato de sodio 50 mM, pH 4,0. Cromatograma de pH 6,0: MES 50 mM, pH 6,0. Cromatograma de pH 8,3: Tris 50 mM, pH 8,3. Cromatograma de pH 8, EDTA: Tris 50 mM, EDTA 10 mM, pH 8,3. Cromatograma de pH 8,3, mezcla de proteasas: Tris 50 mM, pH 8,3 más mezcla de inhibidores de proteasas. Las fracciones de elución de CST-II recogidas (mostradas en una caja) se analizaron por SDS-PAGE.

La figura 36 proporciona un análisis por SDS-PAGE de las fracciones en Superdex 200 del experimento de homogenización de CST-II (prevención de proteólisis). Est = See Blue Plus 2 (15 μ l). Carril 1: Homogenizado de CST-II en Tris 50 mM, pH 8,3 (10 μ l). Carril 2: Homogenizado de CST-II en Tris 50 mM, EDTA 10 mM, pH 8,3 (10 μ l). Carril 3: Homogenizado de CST-II en MES 50 mM, pH 6,0 (10 μ l). Carril 4: Homogenizado de CST-II en Tris 50 mM, pH 8,3 más mezcla de inhibidores de proteasas (10 μ l). Carril 5: Homogenizado de CST-II en Tris 50 mM, pH 8,3 más fracción SEC (10 μ l). Carril 6: Fracción SEC de homogenizado de CST-II en Tris 50 mM, EDTA 10 mM, pH 8,3 (10 μ l). Carril 7: Fracción SEC de homogenizado de CST-II en MES 50 mM, pH 6,0 más (10 μ l). Carril 8: Fracción SEC de homogenizado de CST-II en Tris 50 mM, pH 8,3 más mezcla de inhibidores de proteasas (10 μ l). La muestra homogenizada a pH 4,0 no se corrió en el gel, ya que no se obtuvo pico de CST-II de la purificación de Superdex 200. El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain.

La figura 37 proporciona una purificación preparativa en Q-Sepharosa FF de CST-II. Columna: Vantage S con resina Q-Sepharosa FF (1,5 L). Equilibrio: 5 volúmenes de columna (VC) de Tris 50 mM, Twen-20 al 0,005%, pH 8,3 (tampón A). Carga: Homogenizado clarificado (150 g de precipitado con EDTA 10 mM, 1,5 L, conductividad 6,1 ms/cm). La columna se lavó con 5 VC de tampón A. Elución en gradiente por pasos con tampón B (NaCl 1 M en tampón A): 5 VC de tampón B al 15%, 5 VC de tampón B al 100%. Velocidad de flujo: 100 ml/min (49 cm/h). Absorbancia a 280 nm. Se tomaron muestras de las fracciones para SDS-PAGE (figura 36) y ensayo de CST-II (tabla 5). La fracción 2 de elución de Q-Sepharosa (600 ml) se recogió para purificación adicional. FT = flujo que pasa.

La figura 38 proporciona un análisis por SDS-PAGE de las fracciones de la purificación preparativa en Q-Sepharosa FF de CST-II. Est = See Blue Plus 2 (15 μ l). Carril 1: Homogenizado de CST-II (5 μ l). Carril 2: Fracción 1 del flujo que pasa en Q-Sepharosa (figura 35) (15 μ l). Carril 3: Fracción 2 del flujo que pasa en Q-Sepharosa (15 μ l). Carril 4: Fracción 3 del flujo que pasa en Q-Sepharosa (15 μ l). Carril 5: Fracción 4 del flujo que pasa en Q-Sepharosa (15 μ l). Carril 6: Fracción 5 del flujo que pasa en Q-Sepharosa (15 μ l). Carril 7: Fracción de elución 1 de Q-Sepharosa (30 μ l). Carril 8: Fracción de elución 2 de Q-Sepharosa (2,5 μ l, 10 μ g). Carril 9: Fracción de elución 3 de Q-Sepharosa (15 μ l). Carril 10: Fracción de elución 4 de Q-Sepharosa (10 μ l). Carril 11: Fracción de elución 5 de Q-Sepharosa (1 μ l). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain. La fracción de elución 2 de Q-Sepharosa (mostrada en el carril 8) se preparó para purificación posterior.

La figura 39 proporciona purificación preparativa en hidroxiapatita de tipo II (20 micrómetros) de CST-II. Columna: XK50 hidroxiapatita de tipo II – 20 micrómetros (320 ml). Equilibrio: 5 volúmenes de columna (VC) de fosfato de sodio 10 mM, Tween-20 al 0,005%, pH 6,5 (tampón A). Carga: Fracción de elución 2 de Q-Sepharosa (600 ml, figura 35) (pH ajustado, diluida, ajustada con Tween-20 y filtrada por 0,2 micrómetros: 4,2 L, pH = 6,54, conductividad = 4,0 ms/cm, A280 = 1,004 UA). La columna se lavó con 5 VC de tampón A. Elución en gradiente de tampón B desde el 0-100% durante 25 VC (tampón B = cloruro de sodio 1,5 M, fosfato de sodio 10 mM, Tween-20 al 0,005%, pH 6,5), 5 VC de tampón B al 100%, seguido por 5 VC de fosfato de sodio 500 mM, pH 6,5. Velocidad de flujo: 39,3 ml/min (120 cm/h). Absorbancia a 280 nm. Se tomaron muestras de las fracciones para SDS-PAGE (figura 38) y ensayo de CST-II. Se recogió la mezcla de fracción de elución 5 para purificación posterior. FT = flujo que pasa.

La figura 40 proporciona un análisis por SDS-PAGE de las fracciones de la purificación preparativa en hidroxiapatita (HA) de tipo II (20 micrómetros) de CST-II. Est = See Blue Plus 2 (15 μ l). Carril 1: Fracción de elución 2 de CST-II de Q-Sepharosa (10 μ g). Carril 2: Fracción del flujo que pasa 1 de HA (figura 39) (20 μ l). Carril 3: Fracción del flujo que pasa 2 de HA (20 μ l). Carril 4: Fracción del flujo que pasa 3 de HA (20 μ l). Carril 5: Fracción del flujo que pasa 4 de HA (20 μ l). Carril 6: Fracción del flujo que pasa 5 de HA (20 μ l). Carril 7: Mezcla de elución de HA 1 (20 μ l). Carril 8: Mezcla de elución de HA 2 (20 μ l). Carril 9: Mezcla de elución de HA 3 (20 μ l). Carril 10: Fracción de elución de HA 4 (20 μ l). Carril 11: Fracción de elución de HA 5 (10 μ g). Carril 12: Fracción de elución de HA 6 (10 μ l). Carril 13: Fracción de elución de HA 7 (20 μ l). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain. La fracción de elución 5 de hidroxiapatita de tipo II (mostrada en el carril 11) se preparó para purificación posterior.

La figura 41 proporciona la purificación preparativa en SP Sepharosa HP de CST-II. Columna: XK50 SP-Sepharosa HP (300 ml). Equilibrio: 5 volúmenes de columna (VC) de MES 25 mM, Tween-20 al 0,005% (tampón A). Carga: CST-II parcialmente purificada por Q Sepharosa e hidroxiapatita de tipo II (20 micrómetros) (Mezcla de elución de HA 5, figura 37, 1,65 L) (dializada, diluida, filtrada por 0,2 micrómetros, 6 L, pH = 6,03, conductividad = 3,9 ms/cm A280: 0,291 UA). La columna se lavó con 5 VC de tampón A. Gradiente de elución de tampón B desde el 0-50%

durante 15 VC (Tampón B = NaCl 1 M en tampón A), mantener tampón B al 15% durante 1 VC, 5 VC de tampón B al 100%. Velocidad de flujo 39,3 ml/min (120 cm/h). Absorbancia a 280 nm. Se tomaron muestras de las fracciones para SDS-PAGE (figura 40) y ensayo de CST-II (tabla 7). Mezcla de elución 2 de SP Sepharosa HP elegida para procesamiento y formulación adicional.

La figura 42 proporciona el análisis por SDS-PAGE de las fracciones de la purificación preparativa en SP-Sepharosa HP de CST-II. Est = See Blue Plus 2 (15 μ l). Carril 1: CST-II purificada por Q-Sepharosa e hidroxapatita de tipo II (20 micrómetros) (fracción de elución 5, 19,4 μ l, 10 μ g). Carril 2: Mezcla flujo que pasa/lavado de SP-Sepharosa HP (30 μ l). Carril 3: Mezcla de elución 1 de SP-Sepharosa HP (10 μ g). Carril 4: Mezcla de elución 2 de SP-Sepharosa HP (10 μ g). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Safe Stain. La mezcla de elución 2 de SP-Sepharosa HP (mostrada en el carril 4) se eligió para la formulación final.

La figura 43 proporciona el análisis por SDS-PAGE de CST-II formulada. Gel teñido con Simply Blue Safe (A): Est = See Blue Plus 2 (15 μ l). Carril 1: mezcla de fracciones 2 de SP-Sepharosa HP (figura 39, 10 μ g). Carril 2: CST-II formulada (10 μ g). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó mediante Simply Blue Stain Safe. Gel teñido con plata (B): Est = See Blue Plus 2 (15 μ l). Carril 1: CST-II formulada (8 μ g). El gel de SDS PAGE de Tris-glicina del 4-20% se tiñó con el kit de tinción de plata Wako.

La figura 44 proporciona el análisis de pureza por HPLC de CST-II formulada. Condiciones de RP HPLC C₃: Columna; Zorbax 300SB-C3 (150 x 2,1 mm, 5 micrómetros), P/N 883750-909; S/N USL B001354. Detector: Beckman System Gold 168, haz de diodo, UV 214 nm. Temperatura de la columna; 45°C. Velocidad de flujo: 0,6 ml/min. Eluyentes: A: TFA al 0,1% en agua; B: TFA al 0,09% en acetonitrilo. Gradiente: B al 40% de durante 1 minuto, después B del 40% al 60% en 20 minutos(1%/min), del 60% al 80% en 2 minutos, isocrático durante 1 minuto, del 80% al 40% en 2 minutos, isocrático de B al 40% durante 6 minutos. Inyección: 4 μ g de CST-II purificada. El pico pequeño que eluye a 14,3 minutos es un fragmento de proteólisis de CST-II (~10 kDa).

Descripción detallada de la invención

I. Introducción

La presente invención proporciona una proteína sialiltransferasa con etiqueta según las reivindicaciones adjuntas que se expresa en bacterias a niveles que son mayores que los de la proteína sialiltransferasa parental sin etiqueta. La invención también proporciona vectores de expresión y células huésped que se pueden usar para obtener los niveles mayores de producción de proteína.

II. Definiciones

Como se usa en el presente documento "polipéptido sialiltransferasa" se refiere a una proteína que tiene actividad sialiltransferasa y que comprende el motivo A de sialiltransferasa y el motivo B de sialiltransferasa como se divulga en el presente documento. Como se usa en el presente documento "motivo A de sialiltransferasa" se refiere a una secuencia de aminoácidos encontrada en los polipéptidos sialiltransferasa, es decir, DVFRCNQFYFED/E, (SEQ ID NO:11), y variantes conservativamente modificadas de esa secuencia. Por tanto, el motivo A de sialiltransferasa se refiere a DVFRCNQFYFED, (SEQ ID NO:13) y DVFRCNQFYFEE, (SBQ ID NO:14), y variantes conservativamente modificadas de esas secuencias, también. Como se usa en el presente documento "motivo B de sialiltransferasa" se refiere a una secuencia de aminoácidos encontrada en polipéptidos sialiltransferasa, es decir, RITSGVYMC, (SEQ ID NO:12) y variantes conservativamente modificadas de esa secuencia. En general, el motivo A de sialiltransferasa se encuentra amino terminal respecto a sialiltransferasa B en un polipéptido sialiltransferasa. El espaciamiento entre los dos motivos de sialiltransferasa no es crítico. En algunas formas de realización, aproximadamente 30, 35, 40, 44, 45, 50, 55, 60, 65, 70, 75, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 105 o 110 residuos de aminoácidos separan los dos motivos. Típicamente, el espaciamiento entre los dos motivos es entre, por ejemplo, 80 y 100 residuos o entre 90 y 95 residuos, y para algunas formas de realización habitualmente es, por ejemplo, 91, 92 o 93 residuos de aminoácidos. Los motivos de sialiltransferasa se describen adicionalmente en la solicitud internacional PCT/CA2005/001432, presentada el 16 de septiembre de 2005. Las proteínas sialiltransferasas pueden ser de organismos procariotas, tales como bacterias. En algunas formas de realización, las proteínas sialiltransferasa son de *Campylobacter*. En una forma de realización más, las proteínas sialiltransferasa son de *Campylobacter jejuni*. Las proteínas sialiltransferasa de *Campylobacter* incluyen, por ejemplo, las proteínas CstI, CstII y CstIII.

En algunas formas de realización, la proteína sialiltransferasa se refiere a un polipéptido que tiene al menos el 95%, el 96%, el 97%, el 98% o el 99% de identidad a SEQ ID NO:4. El polipéptido sialiltransferasa tiene actividad sialiltransferasa, es decir, la proteína cataliza la transferencia de un sustrato donante, tal como una molécula activada de ácido siálico, a un sustrato aceptor, tal como un oligosacárido, glicolípido o glicoproteína. Este grupo de polipéptidos sialiltransferasa incluye proteínas que catalizan la adición del residuo de ácido siálico en un enlace α 2,3, proteínas que catalizan la adición del residuo de ácido siálico en un enlace α 2,8, y proteínas de función dual que catalizan la adición del residuo de ácido siálico en un enlace α 2,3 y un enlace α 2,8. En una forma de realización preferida, el polipéptido sialiltransferasa es de *Campylobacter jejuni*, por ejemplo, un polipéptido CSTII. Los ejemplos

de proteínas sialiltransferasa de *Campylobacter* se encuentran en, por ejemplo, la patente en EE UU No. 6.503.744 concedida el 7 de enero de 2003 y la patente en EE UU No. 6.699.705 concedida el 2 de marzo de 2004, que se incorporan ambas en el presente documento mediante referencia; y secuencias divulgadas en los siguientes números de acceso: CAA40567, CAB73395, AAL09368, AAL36462, ZP_00322176, ZP_00321441, ZP_00155359; 5 ZP_00156191, AAL05990, AAG43979, AAK03258, AAF13495, AAK96001, AAK91725, AAL06004, CAB73395, AAL09368, NP_245125, y AAL36462.

Como se usa en el presente documento, un "polipéptido sialiltransferasa truncado" o variantes gramaticales, se refiere a un polipéptido sialiltransferasa que se ha manipulado para eliminar al menos un residuo de aminoácido, 10 relativo al polipéptido sialiltransferasa salvaje que se da en la naturaleza, siempre que el polipéptido sialiltransferasa truncado retenga actividad enzimática. Por ejemplo, los polipéptidos Cst-I de *C. jejuni* que comprenden los aminoácidos 1 a 285 son activos; los polipéptidos Cst-II de *C. jejuni* que comprenden los aminoácidos 1 a 255 son activos; y los polipéptidos Cst-III de *C. jejuni* que comprenden los aminoácidos 1 a 255 son activos.

Como se usa en el presente documento, una "proteína o polipéptido sialiltransferasa con etiqueta" se refiere a una sialiltransferasa que está codificado por una secuencia de ácido nucleico que incluye la etiqueta de aminoácidos MGS en el extremo amino de la proteína. En algunas formas de realización una proteína expresada por ese ácido 15 nucleico empieza con la metionina del MGS. En otras formas de realización, se puede producir la degradación proteolítica de la proteína expresada y la proteína producida es una mezcla de proteínas, incluyendo proteínas que empiezan con la metionina, glicina o serina de la etiqueta MGS o la metionina nativa de la sialiltransferasa. En una forma de realización más, la mayor parte de la proteína producida empieza con la glicina de la secuencia MGS. En una forma de realización, entre el 50 y el 99% de la proteína sialiltransferasa con etiqueta empieza con la glicina de la secuencia MGS. En otra forma de realización, el 50%, el 51%, el 55%, el 60%, el 65%, el 70%, el 75%, el 80%, el 85%, el 90%, el 91%, el 92%, el 93%, el 94%, el 95%, el 96%, el 97%, el 98%, o el 99% de la proteína 20 sialiltransferasa con etiqueta empieza con la glicina de la secuencia MGS.

"Variantes conservadoramente modificadas" se aplica tanto a secuencias de aminoácidos como de ácidos nucleicos. Con respecto a secuencias particulares de ácidos nucleicos, variantes conservadoramente modificadas se refiere a esos ácidos nucleicos que codifican secuencias de aminoácidos idénticas o sustancialmente idénticas, o donde el 30 ácido nucleico no codifica una secuencia de aminoácidos, a secuencias esencialmente idénticas. Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier proteína determinada. Por ejemplo, los codones GCA, GCC, GCG y GCU, todos codifican el aminoácido alanina. Por tanto, en cada posición donde una alanina se especifica por un codón, el codón se puede cambiar a cualquiera de los codones correspondientes descritos sin cambiar el polipéptido codificado. Tales variaciones de ácidos nucleicos son "variaciones silenciosas", que son una especie de variaciones conservadoramente modificadas. Cada secuencia de ácido nucleico en el presente documento que codifica un polipéptido también describe cada posible variación silenciosa del ácido nucleico. El experto en la materia reconocerá que cada codón en un ácido nucleico (excepto AUG, que es normalmente el único codón para metionina, y TGG que es normalmente el único codón para triptófano) se puede modificar para dar una molécula funcionalmente idéntica. Según esto, cada variación silenciosa 35 de un ácido nucleico que codifica un polipéptido está implícita en cada secuencia descrita con respecto al producto de expresión, pero no con respecto a las secuencias sonda reales.

Con respecto a las secuencias de aminoácidos, el experto en la materia reconocerá que las sustituciones, 45 delecciones o adiciones individuales a una secuencia de ácido nucleico, péptido, polipéptido o proteína que cambia, añade o deleciona un único aminoácido o un pequeño porcentaje de aminoácidos en la secuencia codificada es una "variante conservadoramente modificada" donde el cambio produce la sustitución de un aminoácido por un aminoácido químicamente similar. Las tablas de sustituciones conservadoras que proporcionan aminoácidos funcionalmente similares se conocen bien en la técnica. Tales variantes conservadoramente modificadas son además de y no excluyen variantes polimórficas, homólogos interespecie y alelos de la invención.

Los expertos en la materia reconocerán que muchos aminoácidos se pueden sustituir entre sí en una proteína sin 50 afectar la función de la proteína, es decir, una sustitución conservadora puede ser la base de una variante conservadoramente modificada de una proteína tales como las sialiltransferasas divulgadas. Sigue una lista incompleta de sustituciones conservadoras de aminoácidos. Los siguientes ocho grupos contienen cada uno aminoácidos que son sustituciones conservadoras entre sí: 1) alanina (A), glicina (G); 2) ácido aspártico (D), ácido glutámico (E); 3) asparragina (N), glutamina (Q); 4) Arginina (R), lisina (K); 5) isoleucina (I), leucina (L), metionina (M), valina (V), alanina (A); 6) fenilalanina (F), tirosina (Y), triptófano (W); 7) serina (S), treonina (T), cisteína (C); y 8) cisteína (C), metionina (M) (véase, por ejemplo, Creighton, *Proteins* (1984)).

Las células y métodos de la invención son útiles para producir un producto sialilado, en general transfiriendo un grupo ácido siálico de un sustrato donante a una molécula aceptora. Las células y métodos de la invención también son útiles para producir un azúcar producto sialilado que comprende residuos adicionales de azúcar, en general transfiriendo un monosacárido adicional o un grupo sulfato desde un sustrato donante a una molécula aceptora. La adición en general tiene lugar en el extremo no reductor de un oligosacárido, polisacárido (por ejemplo, heparina, carragenina, y similares) o un grupo glucídico sobre un glicolípido o glicoproteína, por ejemplo, una biomolécula. Las biomoléculas como se definen en el presente documento incluyen, pero no están limitadas a, molécula 65

biológicamente significativas tales como hidratos de carbono, oligosacáridos, proteínas (por ejemplo, glicoproteínas) y lípidos (por ejemplo, glicolípidos, fosfolípidos, esfingolípidos y gangliósidos).

En el presente documento se usan las siguientes abreviaturas:

5 Ara = arabinosilo;
 Fru = fructosilo;
 Fuc = fucosilo;
 Gal = galactosilo;
 10 GalNAc = N-acetilgalactosaminilo;
 Glc = glucosilo;
 GlcNAc = N-acetilglucosaminilo;
 Man = manosilo; y
 15 NeuAc = sialilo (N-acetilneuraminilo).

El término “ácido siálico” o “grupo ácido siálico” se refiere a cualquier miembro de una familia de azúcares carboxilados de nueve carbonos. El miembro más común de la familia del ácido siálico es el ácido N-acetilneuramínico (ácido 2-ceto-5-acetamido-3,5-dideoxi-D-glicero-D-galactononulopiranos-1-onico (con frecuencia abreviado como Neu5Ac, NeuAc o NANA). Un segundo miembro de la familia es ácido N-glicolilneuramínico (Neu5Gc o NeuGc), en el que el grupo N-acetilo de NeuA está hidroxilado. Un tercer miembro de la familia del ácido siálico es ácido 2-ceto-3-desoxi-nonulosónico (KDN) (Nadano *et al.* (1986) *J. Biol. Chem.* **261**: 11550-11557; Kanamori *et al.*, *J. Biol. Chem.* **265**: 21811-21819 (1990)). También se incluyen ácidos siálicos sustituidos en 9 tal como un 9-O-acilo de C₁-C₆-Neu5Ac como 9-O-lactil-Neu5Ac o 9-O-acetil-Neu5Ac, 9-desoxi-9-fluoro-Neu5Ac y 9-azido-9-desoxi-Neu5Ac. Para una revisión de la familia del ácido siálico, véase, por ejemplo, Varki, *Glycobiology* **2**: 25-40 (1992); *Sialic Acids: Chemistry, Metabolism and Function*, R. Schauer, Ed. (Springer-Verlag, Nueva York (1992)). La síntesis y uso de los compuestos de ácido siálico en un procedimiento de sialilación se divulga en la solicitud internacional WO 92/16640, publicada el 1 de octubre de 1992.

Un “sacárido producto sialilado” se refiere a un oligosacárido, polisacárido (por ejemplo, heparina, carragenina, y similares) o un grupo glucídico, sin conjugar o conjugado a un glicolípido o glicoproteína, por ejemplo, una biomolécula, que incluye un grupo ácido siálico. Se puede usar cualquiera de los grupos de ácido siálico anteriores así como derivados de ácido siálico PEGilados. En algunas formas de realización también se añaden otros grupos glucídicos, por ejemplo, fucosa, galactosa, glucosa, GalNAc o GluNAc al sustrato aceptor para producir el sacárido producto sialilado. Los ejemplos de sacáridos productos sialilados incluyen, por ejemplo, sialilactosa.

El término “PEG” se refiere a poli(etilenglicol). PEG es un polímero ejemplar que se ha conjugado a péptidos. Se ha demostrado que el uso de PEG para derivar agentes terapéuticos peptídicos reduce la inmunogenicidad de los péptidos y prolonga el tiempo de depuración de la circulación. Por ejemplo, la patente en EE UU No. 4.179.337 (Davis *et al.*) se refiere a péptidos no inmunogénicos, tales como enzimas y hormonas peptídicas acoplados a polietilenglicol (PEG) o polipropilenglicol. Se usan entre 10 y 100 moles de polímero por mol de péptido y se mantiene al menos el 15% de la actividad fisiológica.

Un “sustrato aceptor” o un “sacárido aceptor” para una glicosiltransferasa, por ejemplo, una sialiltransferasa, es un grupo oligosacárido que puede actuar como un aceptor para una glicosiltransferasa particular. Cuando el sustrato aceptor se pone en contacto con la correspondiente glicosiltransferasa y un sustrato donante de azúcar, y otros componentes de la mezcla de reacción necesarios, y la mezcla de reacción se incuba durante un periodo de tiempo suficiente, la glicosiltransferasa transfiere residuos de azúcar del sustrato donante de azúcar al sustrato aceptor. El sustrato aceptor puede variar para diferentes tipos de una glicosiltransferasa particular. Según esto, el término “sustrato aceptor” se toma en contexto con la glicosiltransferasa particular de interés para una aplicación particular. En el presente documento se describen sustratos aceptores para sialiltransferasas y glicosiltransferasas adicionales.

Un “sustrato donante” para glicosiltransferasas es un nucleótido azúcar activado. Los sustratos donantes para sialiltransferasas, por ejemplo, son nucleótidos azúcar activados que comprenden el ácido siálico deseado. Por ejemplo, en el caso de NeuAc, el azúcar activado es CMP-NeuAc. Los sistemas bacterianos, vegetales y fúngicos, algunas veces pueden usar otros nucleótidos azúcar activados.

Se considera que los oligosacáridos tienen un extremo reductor y un extremo no reductor, sea o no el sacárido en el extremo reductor de hecho un azúcar reductor. Según la nomenclatura aceptada, los oligosacáridos se representan en el presente documento con el extremo no reductor a la izquierda y el extremo reductor a la derecha. Todos los oligosacáridos descritos en el presente documento se describen con el nombre a abreviatura para el sacárido no reductor (por ejemplo, Gal), seguido por la configuración del enlace glicosídico (α o β), el enlace de anillo, la posición del anillo del sacárido reductor implicado en el enlace, y después el nombre o abreviatura del sacárido reductor (por ejemplo, GlcNAc). El enlace entre dos azúcares se puede expresar, por ejemplo, como 2,3, 2→3, o (2,3). Cada sacárido es una piranosa o furanosa.

El término “poner en contacto” se usa en el presente documento de forma intercambiable con los siguientes: combinado con, añadido a, mezclado con, pasado sobre, incubado con, deslizado sobre, etc.

5 Se pueden encontrar gran parte de la nomenclatura y procedimientos generales de laboratorio requeridos en esta solicitud en Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual* (2^a Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York, 1989. El manual se denomina de aquí en adelante “Sambrook *et al.*”

10 “Escala comercial” se refiere a producción a escala de gramos de un producto sialilado en una única reacción. En formas de realización preferidas, escala comercial se refiere a la producción de más de aproximadamente 50, 75, 80, 90, 100, 125, 150, 175 o 200 gramos de producto sialilado.

15 Las proteínas recombinantes de la invención se pueden construir y expresar como una proteína de fusión con una “etiqueta de purificación” molecular en un extremo, lo que facilita la purificación o identificación de la proteína. Tales etiquetas también se pueden usar para la inmovilización de una proteína de interés durante la reacción de glicosilación. Las etiquetas adecuadas incluyen “etiquetas epítomos”, que son una secuencia de proteína que es específicamente reconocida por un anticuerpo. Las etiquetas epítomo en general se incorporan a proteínas de fusión para permitir el uso de un anticuerpo fácilmente disponible para detectar de forma inequívoca o aislar la proteína de fusión. Una “etiqueta FLAG” es una etiqueta epítomo comúnmente usada, reconocida específicamente por un anticuerpo monoclonal anti-FLAG, que consiste en la secuencia AspTyrLysAspAspAspLys o una variante sustancialmente idéntica de la misma. Los expertos en la materia conocen otras etiquetas adecuadas, e incluyen, por ejemplo, una etiqueta de afinidad tal como un péptido hexahistidina, que se unirá a iones metálicos tales como iones níquel o cobalto o una etiqueta myc. Las proteínas que comprenden etiquetas de purificación se pueden purificar usando un compañero de unión que se une a la etiqueta de purificación, por ejemplo, anticuerpos contra la etiqueta de purificación, iones o resinas de níquel o cobalto, y amilosa, maltosa o ciclodextrina. Las etiquetas de purificación también incluyen dominios de unión a maltosa y dominios de unión a almidón. Los expertos en la materia conocen la purificación de proteínas con dominios de unión a maltosa. Los dominios de unión a almidón se describen en el documento WO 99/15636. La purificación de afinidad de una proteína de fusión que comprende un dominio de unión a almidón usando una resina de betaciclodextrina (BCD) derivada se describe en el documento USSN 60/468.374, presentado el 5 de mayo de 2003.

30 El término “ácido nucleico” se refiere a un polímero de desoxirribonucleótidos o ribonucleótidos en forma mono- o bicatenaria, y a menos que se limite de otra manera, abarca análogos conocidos de nucleótidos naturales que hibridan con ácidos nucleicos de una manera similar a los nucleótidos naturales. A menos que se indique de otra manera, una secuencia de ácido nucleico particular incluye la secuencia complementaria de la misma. Los términos ácido nucleico, “secuencia de ácido nucleico” y “polinucleótido” se usan de forma intercambiable en el presente documento.

35 El término “operativamente unido” se refiere a la unión funcional entre una secuencia de control de la expresión de ácido nucleico (tal como un promotor, secuencia señal, o haz de sitios de unión de factores de transcripción) y una segunda secuencia de ácido nucleico, en donde la secuencia de control de expresión afecta a la transcripción y/o traducción del ácido nucleico correspondiente a la segunda secuencia.

40 El término “recombinante” cuando se usa con referencia a una célula indica que la célula replica un ácido nucleico heterólogo, o expresa un péptido o proteína codificado por un ácido nucleico heterólogo. Las células recombinantes pueden contener genes que no se encuentran en la forma nativa (no recombinante) de la célula. Las células recombinantes también pueden contener genes encontrados en la forma nativa de la célula en donde los genes se modifican y reintroducen en la célula por medios artificiales. El término también abarca células que contienen un ácido nucleico endógeno a la célula que se ha modificado sin eliminar el ácido nucleico de la célula; tales modificaciones incluyen las obtenidas por reemplazo génico, mutación específica de sitio y técnicas relacionadas.

45 Un “ácido nucleico recombinante” se refiere a un ácido nucleico que se construyó artificialmente (por ejemplo, formado por unión de dos fragmentos de ácido nucleico naturales o sintéticos). Este término también aplica a ácidos nucleicos que se producen por replicación o transcripción de un ácido nucleico que se construyó de forma artificial. Un “polipéptido recombinante” se expresa por transcripción de un ácido nucleico recombinante (es decir, un ácido nucleico que no es nativo a la célula o que se ha modificado de su forma natural), seguido por la traducción del transcrito resultante.

50 Un “polinucleótido heterólogo” o un “ácido nucleico heterólogo”, como se usa en el presente documento, es uno que se origina de una fuente exógena a la célula huésped particular, o, si es de la misma fuente, está modificado de su forma original. Por tanto, un gen de glicosiltransferasa heterólogo en una célula huésped procariota incluye un gen de glicosiltransferasa que es endógena a la célula huésped particular pero que se ha modificado. La modificación de la secuencia heteróloga se puede producir, por ejemplo, tratando el ADN con una enzima de restricción para generar un fragmento de ADN que pueda estar operativamente unido a un promotor. Las técnicas tales como la mutagénesis dirigida también son útiles para modificar una secuencia heteróloga.

65

Una “subsecuencia” se refiere a una secuencia de ácido nucleico o aminoácidos que comprende una parte de una secuencia más larga de ácido nucleico o aminoácidos (por ejemplo, polipéptido), respectivamente.

Un “casete de expresión recombinante” o simplemente un “casete de expresión” es una construcción de ácido nucleico, generada de forma recombinante o sintética, con elementos de ácido nucleico que pueden afectar la expresión de un gen estructural en huéspedes compatibles con tales secuencias. Los casetes de expresión incluyen al menos promotores y opcionalmente, señales de terminación de la transcripción. Típicamente, el casete de expresión recombinante incluye un ácido nucleico que se va a transcribir (por ejemplo, un ácido nucleico que codifica un polipéptido deseado) y un promotor. También se pueden usar factores adicionales necesarios o útiles en efectuar la expresión como se describe en el presente documento. Por ejemplo, un casete de expresión también puede incluir secuencias de nucleótidos que codifican una secuencia señal que dirige la secreción de una proteína expresada de la célula huésped. También se pueden incluir en un casete de expresión señales de terminación de la transcripción, potenciadores y otras secuencias de ácido nucleico que influyen la expresión génica.

Un “polipéptido de fusión sialiltransferasa” o un “polipéptido de fusión glicosiltransferasa” de la invención es un polipéptido que contiene un dominio catalítico de glicosiltransferasa y un segundo dominio catalítico de una enzima accesoria (por ejemplo, una CMP-Neu5Ac sintetasa). El polipéptido de fusión es capaz de catalizar la síntesis de un nucleótido azúcar (por ejemplo, CMP-Neu5Ac) así como la transferencia del residuo de azúcar del nucleótido azúcar a una molécula aceptora. Típicamente, los dominios catalíticos de los polipéptidos de fusión serán al menos sustancialmente idénticos a los de glicosiltransferasas y las proteínas de fusión de las que derivan los dominios catalíticos. En algunas formas de realización, un polipéptido CMP-ácido siálico sintasa y un polipéptido sialiltransferasa se fusionan para formar un único polipéptido. Los expertos en la materia conocen muchas enzimas sialiltransferasa y se pueden usar en los métodos de la invención. Por ejemplo, se describe una fusión entre un polipéptido CMP-ácido siálico sintasa de *Neisseria* y una proteína sialiltransferasa de *Neisseria* en, por ejemplo, el documento WO99/31224 y Gilbert *et al.*, *Nat Biotechnol.* 16:769-72 (1998). Se pueden usar otras fusiones en la invención, por ejemplo, entre un polipéptido CMP-ácido siálico sintasa de *Neisseria* y una sialiltransferasa de *Campylobacter*.

Una “enzima accesoria”, como se denomina en el presente documento, es una enzima que está implicada en catalizar una reacción que, por ejemplo, forma un sustrato u otro reactivo para una reacción de glicosiltransferasa. Una enzima accesoria puede, por ejemplo, catalizar la formación de un nucleótido azúcar que se usa como grupo donante de azúcar por una glicosiltransferasa. Una enzima accesoria también puede ser una que se usa en la generación de un nucleótido trifosfato que se requiere para la formación de un nucleótido azúcar, o en la generación del azúcar que se incorpora en el nucleótido azúcar.

Un “dominio catalítico” se refiere a una parte de una enzima que es suficiente para catalizar una reacción enzimática que normalmente se lleva a cabo por la enzima. Por ejemplo, un dominio catalítico de una sialiltransferasa incluirá una parte suficiente de la sialiltransferasa para transferir un residuo de ácido siálico de un donante de azúcar a un sacárido aceptor. Un dominio catalítico puede incluir una enzima entera, una subsecuencia de la misma, o puede incluir secuencias adicionales de aminoácidos que no están unidas a la enzima o subsecuencia como se encuentra en la naturaleza.

El término “aislado” se refiere a material que está sustancial o esencialmente libre de componentes que interfieren con la actividad de una enzima. Para células, sacáridos, ácidos nucleicos y polipéptidos de la invención, el término “aislado” se refiere a material que está sustancial o esencialmente libre de componentes que normalmente acompañan al material como se encuentra en su estado nativo. Típicamente, los sacáridos, proteínas o ácidos nucleicos aislados de la invención son al menos aproximadamente el 50%, el 55%, 60%, el 65%, el 70%, el 75%, el 80% o el 85% puros, habitualmente al menos aproximadamente el 90%, el 91%, el 92%, el 93%, el 94%, el 95%, el 96%, el 97%, el 98% o el 99% puros, medido por la intensidad de la banda en un gel teñido con plata u otro método para determinar la pureza. Se puede indicar la pureza u homogeneidad por un número de medios bien conocidos en la técnica, tales como electroforesis en gel de poliacrilamida de una muestra de proteína o ácido nucleico, seguido por visualización tras la tinción. Para ciertos fines se necesitará alta resolución y se utilizará HPLC o un medio similar para la purificación. Para oligonucleótidos u otros productos sialilados, la pureza se puede determinar usando, por ejemplo, cromatografía en capa fina, HPLC o espectroscopia de masas.

Los términos “idéntico” o porcentaje de “identidad”, en el contexto de dos o más secuencias de ácido nucleico o polipéptidos, se refiere a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son iguales, cuando se comparan y alinean para la máxima correspondencia, medido usando uno de los siguientes algoritmos de comparación de secuencias o por inspección visual.

La frase “sustancialmente idénticos”, en el contexto de dos ácidos nucleicos o polipéptidos, se refiere a dos o más secuencias o subsecuencias que tienen al menos el 60%, preferiblemente el 80% o el 85%, lo más preferiblemente al menos el 90%, el 91%, el 92%, el 93%, el 94%, el 95%, el 96%, el 97%, el 98% o el 99% de identidad de nucleótidos o residuos de aminoácidos, cuando se comparan y alinean para la máxima correspondencia, medido usando uno de los siguientes algoritmos de comparación de secuencias o por inspección visual. Preferiblemente, la

identidad sustancial existe sobre una región de la secuencia que tiene al menos aproximadamente 50 residuos de longitud, más preferiblemente sobre una región de al menos aproximadamente 100 residuos, y lo más preferiblemente las secuencias son sustancialmente idénticas sobre al menos aproximadamente 150 residuos. En una forma de realización más preferida, las secuencias son sustancialmente idénticas sobre la longitud entera de las regiones codificantes.

Para la comparación de secuencias, típicamente una secuencia actúa como una secuencia de referencia, a la que se comparan secuencias de prueba. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de prueba y referencia se introducen en un ordenador, se designan coordenados de subsecuencia, si es necesario, y se designan parámetros de programa de algoritmo de secuencia. El algoritmo de comparación de secuencia calcula entonces el porcentaje de identidad de secuencia para la(s) secuencia(s) de prueba, relativo a la secuencia de referencia, basado en los parámetros del programa designados.

El alineamiento óptimo de secuencias para la comparación se puede realizar, por ejemplo, por el algoritmo de homología local de Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), por el algoritmo de alineamiento de homología de Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), por la búsqueda para el método de similitud Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), por implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el paquete de software genético de Wisconsin, Genetics Computer Group, 575 Science Dr., Madison, WI), o por inspección visual (véase, en general, *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., Current Protocols, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (1995 Suplemento) (Ausubel)).

Los ejemplos de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-410 y Altschuel *et al.* (1977) *Nucleic Acids Res.* 25: 3389-3402, respectivamente. El software para realizar los análisis de BLAST está públicamente disponible a través del Centro Nacional para Información de Biotecnología (www.ncbi.nlm.nih.gov/). Este algoritmo implica primero identificar pares de secuencias de alta puntuación (HSP) identificando palabras cortas de longitud W en la secuencia problema, que coincide o satisface alguna puntuación umbral de valor positivo T cuando se alinea con una palabra de la misma longitud en una base de datos de secuencias. T se denomina umbral de puntuación de palabra vecina (Altschul *et al.*, anteriormente). Estas coincidencias iniciales de palabras vecinas actúan como semillas para iniciar búsquedas para encontrar HSP más largas que las contengan. Las coincidencias de palabras se extienden en ambas direcciones a lo largo de cada secuencia durante tanto como se pueda aumentar la puntuación acumulativa de alineamiento. Las puntuaciones acumulativas se calculan usando, para secuencias de nucleótidos, los parámetros M (puntuación recompensa para un par de residuos coincidentes; siempre > 0) y N (puntuación de penalización para residuos mal emparejados; siempre < 0). Para secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de las coincidencias de palabra en cada dirección se paran cuando: la puntuación acumulativa de alineamiento desciende en la cantidad X de su valor máximo alcanzado; la puntuación acumulativa se hace cero o menor, debido a la acumulación de uno o más alineamientos de residuos de puntuación negativa; o se alcanza el final de cualquier secuencia. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad del alineamiento. El programa BLASTN (para secuencias de nucleótidos) usa por defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, M=5, N=-4 y una comparación de ambas hebras. Para secuencias de aminoácidos, el programa BLASTP usa por defecto una longitud de palabra (W) de 3, una expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA*, 89:10915 (1989)).

Además de calcular el porcentaje de identidad de secuencia, el algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). Una medida de similitud proporcionada por el algoritmo BLAST es la menor suma de probabilidad (P(N)), que proporciona una indicación de la probabilidad por la que se produjera una coincidencia entre dos secuencias de nucleótidos o aminoácidos al azar. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la menor suma de probabilidad en una comparación del ácido nucleico de prueba con el ácido nucleico de referencia es menor de aproximadamente 0,1, más preferiblemente menor de aproximadamente 0,01 y los más preferiblemente menor de aproximadamente 0,001.

Una indicación adicional de que dos secuencias de ácido nucleico o polipéptidos son sustancialmente idénticos es que el polipéptido codificado por el primer ácido nucleico tenga inmunológicamente reactividad cruzada con el polipéptido codificado por el segundo ácido nucleico, como se describe posteriormente. Por tanto, típicamente un polipéptido es sustancialmente idéntico a un segundo polipéptido, por ejemplo, donde los dos polipéptidos se diferencian solo por sustituciones conservadoras. Otra indicación de que dos secuencias de ácido nucleico son sustancialmente idénticas es que las dos moléculas hibridan entre sí en condiciones rigurosas, como se describe a continuación.

La frase "que hibrida específicamente con", se refiere a la unión, formación de dúplex o hibridación de una molécula solo a una secuencia de nucleótidos particular en condiciones rigurosas cuando esa secuencia está presente en una mezcla compleja de ADN o ARN (por ejemplo, celular total).

El término “condiciones rigurosas” se refiere a las condiciones en las que una sonda hibridará con su subsecuencia diana, pero no con otras secuencias. Las condiciones rigurosas dependen de la secuencia y serán diferentes en circunstancias diferentes. Las secuencias más largas hibridan específicamente a temperaturas más altas. En general, las condiciones rigurosas se seleccionan para que sean aproximadamente 5°C menor que el punto de fusión térmico (T_m) para la secuencia específica a una fuerza iónica y pH definidos. La T_m es la temperatura (a una fuerza iónica, pH y concentración de ácido nucleico definidos) a la que el 50% de las sondas complementarias a la secuencia diana hibrida con la secuencia diana en equilibrio. (Como las secuencias diana en general están presentes en exceso, a la T_m, el 50% de las sondas están ocupadas a equilibrio). Típicamente, las condiciones rigurosas serán esas en las que la concentración de sal es menor de aproximadamente 1,0 N de ión Na⁺, típicamente aproximadamente de 0,01 a 1,0 M de concentración de ión Na⁺ (u otras sales) a pH de 7,0 a 8,3 y la temperatura es al menos aproximadamente 30°C para sondas cortas (por ejemplo, de 10 a 50 nucleótidos) y al menos aproximadamente 60°C para sondas largas (por ejemplo, mayores de 50 nucleótidos). Las condiciones rigurosas también se pueden alcanzar con la adición de agentes desestabilizantes tales como formamida. Para amplificación por PCR de alta rigurosidad, es típica una temperatura de aproximadamente 62°C, aunque las temperaturas de hibridación muy rigurosas pueden variar desde aproximadamente 50°C hasta aproximadamente 65°C, dependiendo de la longitud y especificidad del cebador. Las condiciones de ciclo típicas para amplificaciones tanto muy como poco rigurosas incluyen una fase de desnaturalización de 90-95°C durante 30-120 segundos, una fase de hibridación que dura 30-120 segundos y una fase de extensión a aproximadamente 72°C durante 1-2 minutos. Los protocolos y directrices para reacciones de amplificación poco y muy rigurosas están disponibles, por ejemplo, en Innis, et al. (1990) *PCR Protocols: A Guide to Methods and Applications* Academic Press, N.Y.

Las frases “se une específicamente a” o “específicamente inmunorreactivo con”, cuando se refieren a un anticuerpo se refiere a una reacción de unión que es determinante de la presencia de la proteína u otro antígeno en presencia de una población heterogénea de proteínas, sacáridos y otros productos biológicos. Por tanto, en condiciones de inmunoensayo designadas, los anticuerpos especificados se unen preferentemente a un antígeno particular y no se unen en cantidad significativa a otras moléculas presentes en la muestra. La unión específica a un antígeno en tales condiciones requiere un anticuerpo que se selecciona por su especificidad para un antígeno particular. Se puede usar una variedad de formatos de inmunoensayos para seleccionar anticuerpos específicamente inmunorreactivos con un antígeno particular. Por ejemplo, los inmunoensayos ELISA en fase sólida se usan rutinariamente para seleccionar anticuerpos monoclonales específicamente inmunorreactivos con un antígeno. Véase, Harlow y Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, Nueva York, para una descripción de formatos y condiciones de inmunoensayos que se pueden usar para determinar la inmunorreactividad específica.

“Anticuerpo” se refiere a un polipéptido que comprende una región marco de un gen de inmunoglobulina o fragmentos de los mismos que se une específicamente y reconoce un antígeno. Los genes reconocidos de inmunoglobulinas incluyen los genes de las regiones constantes kappa, lambda, alfa, gamma, delta, épsilon y mu, así como la miríada de genes de la región variable de la inmunoglobulina. Las cadenas ligeras se clasifican como kappa o lambda. Las cadenas pesadas se clasifican como gamma, mu, alfa, delta o épsilon, que a su vez definen las clases de inmunoglobulinas IgG, IgM, IgA, IgD e IgE, respectivamente. Típicamente, la región de unión a antígeno de un anticuerpo será la más crítica en la especificidad y afinidad de unión.

Una unidad estructural de inmunoglobulina (anticuerpo) ejemplar comprende un tetrámero. Cada tetrámero está compuesto de dos pares idénticos de cadenas polipeptídicas, cada par tiene una cadena “ligera” (aproximadamente 25 kDa) y una “pesada” (aproximadamente 50-70 kDa). El extremo N de cada cadena define una región variable de aproximadamente 100 a 110 aminoácidos o más principalmente responsables para el reconocimiento del antígeno. Los términos cadena ligera variable (V_L) y cadena pesada variable (V_H) se refieren a estas cadenas ligera y pesada, respectivamente.

Los anticuerpos existen, por ejemplo, como inmunoglobulinas intactas o como un número de fragmentos bien caracterizados producidos por la digestión con varias peptidasas. Por tanto, por ejemplo, la pepsina digiere un anticuerpo por debajo de los enlaces disulfuro en la región bisagra para producir F(ab)₂, un dímero de Fab que él mismo es una cadena ligera unida a V_H-C_H1 por un puente disulfuro. El F(ab)₂ se puede reducir en condiciones moderadas para romper el enlace disulfuro en la región bisagra, convirtiendo de esta manera el dímero F(ab)₂ en un monómero Fab'. El monómero Fab' es esencialmente Fab con parte de la región bisagra (véase, *Fundamental Immunology* (Paul ed., 3ª ed. 1993). Mientras que se definen varios fragmentos de anticuerpo en términos de la digestión de un anticuerpo intacto, el experto apreciará que tales fragmentos se pueden sintetizar de nuevo químicamente o usando metodología de ADN recombinante. Por tanto, el término anticuerpo, como se usa en el presente documento, también incluye fragmentos de anticuerpo producidos por la modificación de los anticuerpos completos o los sintetizados de nuevo usando metodologías de ADN recombinante (por ejemplo, Fv de cadena única) o los identificados usando librerías de presentación en fagos (véase, por ejemplo, McCafferty *et al.*, *Nature* 348:552-554 (1990)).

Para la preparación de anticuerpos, por ejemplo, anticuerpos recombinantes, monoclonales o policlonales, se pueden usar muchos métodos conocidos en la técnica (véase, por ejemplo, Kohler & Milstein, *Nature* 256:495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pp. 77-96 en *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985); Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *Antibodies, A*

Laboratory Manual (1988); y Goding, *Monoclonal Antibodies: Principles and Practice* (2ª ed. 1986)). Los genes que codifican las cadenas pesada y ligera de un anticuerpo de interés se pueden clonar de una célula, por ejemplo, los genes que codifican un anticuerpo monoclonal se pueden clonar de un hibridoma y usarse para producir un anticuerpo monoclonal recombinante. También se pueden hacer genotecas que codifican cadenas pesadas y ligeras de anticuerpos monoclonales a partir de hibridomas o células plasmáticas. Las combinaciones aleatorias de productos de genes de cadenas pesadas y ligeras generan un gran conjunto de anticuerpos con diferentes especificidad antigénica (véase, por ejemplo, Kubly, *Immunology* (3ª ed. 1997)). Las técnicas para la producción de anticuerpos de cadena única o anticuerpos recombinantes (patente en EE UU 4.946.778, patente en EE UU No. 4.816.567) se pueden adaptar para producir anticuerpos contra polipéptidos de esta invención. Además, se pueden usar ratones transgénicos u otros organismos tales como otros mamíferos, para expresar anticuerpos humanizados o humanos (véase, por ejemplo, las patentes en EE UU No. 5.545.807; 5.545.806; 5.569.825; 5.625.126; 5.633.425; 5.661.016, Marks *et al.*, *Bio/Technology* 10:779-783 (1992); Lonberg *et al.*, *Nature* 368:856-859 (1994); Morrison, *Nature* 368:812-13 (1994); Fishwild *et al.*, *Nature Biotechnology* 14:845-51 (1996); Neuberger, *Nature Biotechnology* 14:826 (1996); y Lonberg & Huszar, *Intern. Rev. Immunol.* 13:65-93 (1995)). De forma alternativa, se puede usar tecnología de presentación en fagos para identificar anticuerpos y fragmentos Fab heteroméricos que se unen específicamente a antígenos seleccionados (véase, por ejemplo, McCafferty *et al.*, *Nature* 348:552-554 (1990); Marks *et al.*, *Biotechnology* 10:779-783 (1992)). Los anticuerpos también se pueden hacer biespecíficos, es decir capaces de reconocer dos antígenos diferentes (véase, por ejemplo, el documento WO 93/08829, Trauneker *et al.*, *EMBO J.* 10:3655-3659 (1991); y Suresh *et al.*, *Methods in Enzymology* 121:210 (1986)). Los anticuerpos también pueden ser heteroconjugados, por ejemplo, dos anticuerpos o inmunotoxinas unidos covalentemente (véase, por ejemplo, la patente en EE UU No. 4.676.980, documentos WO 91/00360; WO 92/200373; y EP 03089).

En una forma de realización, el anticuerpo se conjuga a un grupo "efector". El grupo efector puede ser cualquier número de moléculas, incluyendo grupos marcadores tales como marcadores radioactivos o marcadores fluorescentes para su uso en ensayos diagnósticos.

La frase "se une específicamente (o selectivamente)" a un anticuerpo o "específicamente (o selectivamente) inmunorreacciona con", cuando se refiere a una proteína o péptido, se refiere a una reacción de unión que es determinante de la presencia de la proteína, con frecuencia en una población heterogénea de proteínas y otros productos biológicos. Por tanto, en condiciones de inmunoensayo designadas, los anticuerpos especificados se unen a una proteína particular al menos dos veces el nivel de fondo y más típicamente más de 10 a 100 veces el nivel de fondo. La unión específica a un anticuerpo en tales condiciones requiere un anticuerpo que se selecciona por su especificidad para una proteína particular. Por ejemplo, los anticuerpos policlonales hechos contra la proteína IgE, variantes polimórficas, alelos, ortólogos y variante conservadoramente modificadas, o variantes de ajuste, o partes de las mismas, se pueden seleccionar para obtener solo esos anticuerpos policlonales que son específicamente inmunorreactivos con proteínas IgE y no con otras proteínas. Esta selección se puede lograr restando los anticuerpos que dan reacción cruzada con otras moléculas. Se pueden usar una variedad de formatos de inmunoensayo para seleccionar anticuerpos específicamente inmunorreactivos con una proteína particular. Por ejemplo, los inmunoensayos ELISA en fase sólida se usan rutinariamente para seleccionar anticuerpos monoclonales específicamente inmunorreactivos con un antígeno (véase, por ejemplo, Harlow & Lane, *Antibodies, A Laboratory Manual*, (1988) para una descripción de formatos y condiciones de inmunoensayos que se pueden usar para determinar la inmunorreactividad específica).

Un "antígeno" es una molécula que es reconocida y se une a un anticuerpo, por ejemplo, péptidos, hidratos de carbono, moléculas orgánicas o moléculas más complejas tales como glicolípidos y glicoproteínas. La parte del antígeno que es la diana de la unión del anticuerpo es un determinante antigénico y un grupo funcional pequeño que corresponde a único determinante antigénico se llama hapteno.

Un "marcador" es una composición detectable por medios espectroscópicos, fotoquímicos, bioquímicos, inmunoquímicos o químicos. Por ejemplo, los marcadores útiles incluyen, ³²P, ¹²⁵I, colorantes fluorescentes, reactivos densos a los electrones, enzimas (por ejemplo, comúnmente usadas en ELISA), biotina, digoxigenina o haptenos y proteínas para los que están disponibles antisueros o anticuerpos monoclonales (por ejemplo, el polipéptido de SEQ ID NO: 3 se puede hacer detectable, por ejemplo, incorporando un radiomarcador en el péptido, y usarse para detectar anticuerpos específicamente reactivos con el péptido).

El término "inmunoensayo" es un ensayo que usa un anticuerpo para unirse específicamente a un antígeno. El inmunoensayo se caracteriza por el uso de propiedades de unión específicas de un anticuerpo particular para aislar, dirigirse a y/o cuantificar el antígeno.

El término "molécula soporte" significa una molécula inmunogénica que contiene determinantes antigénicos reconocidos por células T. Una molécula soporte puede ser una proteína o puede ser un lípido. Una proteína soporte se conjuga a un polipéptido para hacer el polipéptido inmunogénico. Las proteínas soporte incluyen hemocianina de lapa californiana, hemocianina de cangrejo herradura y seroalbúmina bovina.

El término “adyuvante” significa una sustancia que aumenta de forma no específica la respuesta inmune al antígeno. Los adyuvantes incluyen adyuvante de Freund, completo o incompleto, adyuvante de oro Titermax, alumbre y LPS bacteriano.

5 III. Polipéptidos sialiltransferasa

Los polipéptidos sialiltransferasa de la invención típicamente comprenden dos motivos: motivo A de sialiltransferasa, DVFRCNQFYFED/E, y variantes conservadoramente modificadas de esa secuencia y motivo B de sialiltransferasa, RITSGVYMC y variantes conservadoramente modificadas de esa secuencia. En algunas formas de realización, los polipéptidos sialiltransferasa comprenden el motivo A de sialiltransferasa DVFRCNQFYFED o DVFRCNQFYFEE, y el motivo B de sialiltransferasa RITSGVYMC. Los polipéptidos sialiltransferasa de la invención catalizan la transferencia de un grupo de ácido siálico de un sustrato donante a un sustrato aceptor.

En algunas formas de realización, los polipéptidos sialiltransferasa también comprenden otros residuos de aminoácidos que parecen ser importantes para la actividad enzimática. Por ejemplo, se ha resuelto la estructura de Cst-II de la cepa OH4384 de *Campylobacter jejuni* (véase, por ejemplo, Chiu *et al.*, *Nat. Struc. Mol. Biol.* 11:163-170 (2004)). El análisis mutacional de la enzima Cst-II demostró que, por ejemplo el residuo de arginina del motivo B de sialiltransferasa se requiere para la actividad. El residuo de arginina del motivo B de sialiltransferasa se denomina R129 en Cst-II y se correlaciona con R165 de la secuencia consenso de sialiltransferasa de la figura 1. Otros residuos de aminoácidos que parecen ser importantes para la actividad catalítica incluyen Y156 de Cst-II, Y162 de Cst-II y H188 de Cst-II. Por tanto, en algunas formas de realización, los polipéptidos sialiltransferasa comprenden el motivo A de sialiltransferasa, el motivo B de sialiltransferasa y un residuo de aminoácido correspondiente a Y192 consenso; o el motivo A de sialiltransferasa, el motivo B de sialiltransferasa y un residuo de aminoácido correspondiente a Y199 o H226 consenso; o el motivo A de sialiltransferasa, el motivo B de sialiltransferasa y un residuo de aminoácido correspondiente a Y199 consenso; o el motivo A de sialiltransferasa, el motivo B de sialiltransferasa y un residuo de aminoácido correspondiente a Y199 consenso y un residuo de aminoácido correspondiente a H226 consenso; o el motivo A de sialiltransferasa, el motivo B de sialiltransferasa y un residuo de aminoácido correspondiente a H226 consenso; o el motivo A de sialiltransferasa, el motivo B de sialiltransferasa y un residuo de aminoácido correspondiente a Y192 consenso, un residuo de aminoácido correspondiente a Y199 y un residuo de aminoácido correspondiente a H226 consenso.

Otros residuos de aminoácidos pueden ser importantes para la actividad enzimática, basados en los datos estructurales y se pueden incluir en los polipéptidos sialiltransferasa con motivos A y B de sialiltransferasa, por ejemplo, residuos de aminoácidos correspondientes a los residuos consenso N44, N86, Q93, D190, F191, S198, F215 o Y222. Los expertos en la materia notarán al revisar la figura 1, que en los residuos consenso S198, Y222 y F215, se pueden tolerar otros aminoácidos. Además, N86 y Q93 se delecionan de polipéptidos sialiltransferasa, por ejemplo, de algunos polipéptidos sialiltransferasa de *H. influenzae*. Los residuos de aminoácidos anteriores se pueden incluir en un polipéptido sialiltransferasa, es decir, un polipéptido de forma única o en cualquier combinación, incluyendo combinaciones con residuos de aminoácidos correspondientes a Y192, Y199 o H226 consenso.

Los polipéptidos sialiltransferasa también se pueden modificar, siempre que mantengan actividad sialiltransferasa. Las modificaciones incluyen truncamientos, descritos anteriormente, y, en algunas formas de realización, mutagénesis dirigida de la proteína.

Se puede usar la mutagénesis dirigida para cambiar la especificidad del aceptor de un polipéptido sialiltransferasa que comprende motivos de secuencia conservados. Algunos polipéptidos sialiltransferasas pueden sialilar una molécula aceptora formando enlaces α 2,3 y/o α 2,8. Por ejemplo, las enzimas CstII de las cepas OH4382, OH4384, O:10 y O:41 de *C. jejuni* son todas capaces de formar enlaces α 2,3 y/o α 2,8. La mutación de Asn51 a un residuo de treonina eliminó la capacidad de CstII de OH4382, OH4384 de añadir ácido siálico en un enlace α 2,8. Sin embargo, la mutación de Thr51 a asparragina en una cepa monofuncional O:19, produjo una enzima que era capaz de añadir un ácido siálico tanto en enlace α 2,3 así como α 2,8 (véase, por ejemplo, Gilbert *et al.*, *J. Biol. Chem.* 277:327-337 (2002)).

La mutagénesis dirigida se puede usar para cambiar la actividad enzimática o estabilidad global de proteína. Además, una mutación del residuo Ile53 a una serina o glicina en enzimas CstII de las cepas OH4382, OH4384, de *C. jejuni* produjo aumentos grandes de actividad enzimática. Véase, por ejemplo, Chiu *et al.*, *Nat. Struc. Mol. Biol.* 11:163-170 (2004).

60 IV. Aislamiento de ácidos nucleicos que codifican polipéptidos sialiltransferasa

Los expertos en la materia conocen los ácidos nucleicos que codifican los polipéptidos sialiltransferasa para su uso en la presente invención. Los ácidos nucleicos adecuados (por ejemplo, ADNc, genómico o subsecuencias (sondas)) se pueden clonar o amplificar por métodos *in vitro* tales como la reacción en cadena de la polimerasa (PCR), la reacción en cadena de la ligasa (LCR), el sistema de amplificación basado en transcripción (TAS) o el sistema de replicación de secuencia autosostenida (SSR). Los expertos en la materia conocen muy bien de una amplia variedad

de metodologías de amplificación *in vitro*. Se encuentran ejemplos de estas técnicas e instrucciones suficientes para dirigir a los expertos en la materia a través de muchos ejercicios de clonación en Berger y Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology* 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook *et al.* (1989) *Molecular Cloning - A Laboratory Manual* (2ª ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY, (Sambrook *et al.*); *Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., Current Protocols, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (1994 Suplemento) (Ausubel); Cashion *et al.*, patente en EE UU número 5.017.478; y Carr, Patente Europea No. 0.246.864.

Se puede preparar un ADN que codifica un polipéptido sialiltransferasa por cualquier método adecuado descrito anteriormente, incluyendo, por ejemplo, clonación y restricción de secuencias apropiadas con enzimas de restricción. En una forma de realización preferida, los ácidos nucleicos que codifican polipéptidos sialiltransferasa se aíslan por métodos de clonación rutinarios. Se puede usar una secuencia de nucleótidos de un polipéptido sialiltransferasa como se proporciona en, por ejemplo, SEQ ID NO:4 y 5, u otra base de datos de secuencias (véase anteriormente) para proporcionar sondas que hibriden específicamente con un gen que codifica un polipéptido sialiltransferasa; o con un ARNm que codifica un polipéptido sialiltransferasa (por ejemplo, en una transferencia de tipo Southern o Northern). Una vez que se identifica el ácido nucleico diana que codifica el polipéptido sialiltransferasa, se puede aislar según métodos estándar que conocen los expertos en la materia (véase, por ejemplo, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, 2ª Ed., Vols. 1-3, Cold Spring Harbor Laboratory; Berger y Kimmel (1987) *Methods in Enzymology*, Vol. 152: *Guide to Molecular Cloning Techniques*, San Diego: Academic Press, Inc.; o Ausubel *et al.* (1987) *Current Protocols in Molecular Biology*, Greene Publishing y Wiley-Interscience, Nueva York). Además, los ácidos nucleicos aislados se pueden cortar con enzimas de restricción para crear ácidos nucleicos que codifican el polipéptido sialiltransferasa de longitud total, o subsecuencias del mismo, por ejemplo, que contienen subsecuencias que codifican al menos una subsecuencia de un dominio catalítico de un polipéptido sialiltransferasa. Estos fragmentos de enzimas de restricción, que codifican un polipéptido sialiltransferasa o subsecuencias del mismo, se pueden ligar después, por ejemplo, para producir un ácido nucleico que codifica una proteína sialiltransferasa.

Un ácido nucleico que codifica un polipéptido sialiltransferasa c, o una subsecuencia del mismo, se puede caracterizar ensayando el producto expresado. Se pueden usar ensayos basados en la detección de las propiedades físicas, químicas o inmunológicas de las proteínas expresadas. Por ejemplo, se puede identificar una sialiltransferasa clonada por la capacidad de una proteína codificada por el ácido nucleico para catalizar la transferencia de un grupo ácido siálico de un sustrato donante a un sustrato aceptor. En un método, se emplea electroforesis capilar para detectar los productos de reacción. Este ensayo muy sensible implica el uso de derivados aminofenolis de sacárido o disacárido que están marcados con fluoresceína como se describe en Wakarchuk *et al.* (1996) *J. Biol. Chem.* 271 (45): 28271-276. Para ensayar para α 2,3-sialiltransferasa, se usa Lac-FCHASE como sustrato. Para ensayar para α 2,8-sialiltransferasa, se usa GM3-FCHASE como sustrato. (Véase, por ejemplo, la patente en EE UU No. 6.503.744). Se pueden detectar los productos de reacción de otras glicosiltransferasas usando electroforesis capilar, por ejemplo, para ensayar una enzima *IgtC* de *Neisseria*, se pueden usar FCHASE-AP-Lac o FCHASE-AP-Gal, mientras que para la enzima *IgtB* de *Neisseria* un reactivo apropiado es FCHASE-AP-GlcNAc (Wakarchuk, anteriormente). Otros métodos para la detección de productos de reacción oligosacáridos incluyen cromatografía en capa fina y CG/EM y se divulgan en la patente en EE UU No. 6.503.744.

Además, un ácido nucleico que codifica un polipéptido sialiltransferasa o una subsecuencia del mismo, se puede sintetizar químicamente. Los métodos adecuados incluyen el método del fosfotriéster de Narang *et al.* (1979) *Meth. Enzymol.* 68: 90-99; el método del fosfodiéster de Brown *et al.* (1979) *Meth. Enzymol.* 68: 109-151; el método de la dietilfosforamidita de Beaucage *et al.* (1981) *Tetra. Lett.*, 22: 1859-1862; y el método del soporte sólido de la patente de EE UU No. 4.458.066. La síntesis química produce un oligonucleótido monocatenario. Este se puede convertir en ADN bicatenario mediante hibridación con una secuencia complementaria, o mediante polimerización con una ADN polimerasa que usa la cadena sencilla como molde. El experto en la materia reconoce que mientras que la síntesis química de ADN con frecuencia está limitada a secuencias de aproximadamente 100 bases, se pueden obtener secuencias más largas por ligación de secuencias más cortas.

Los ácidos nucleicos que codifican polipéptidos sialiltransferasas o subsecuencias de los mismos, se pueden clonar usando métodos de amplificación de ADN tal como la reacción en cadena de la polimerasa (PCR). Por tanto, por ejemplo, la secuencia o subsecuencia del ácido nucleico se amplifica por PCR, usando un cebador sentido que contiene un sitio de enzima de restricción (por ejemplo, NdeI) y un cebador antisentido que contiene otro sitio de enzima de restricción (por ejemplo, HindIII). Esto producirá un ácido nucleico que codifica el polipéptido o subsecuencia sialiltransferasa deseado y que tiene sitios de enzimas de restricción terminales. Este ácido nucleico se puede ligar después fácilmente en un vector que contiene un ácido nucleico que codifica la segunda molécula y que tiene los sitios apropiados de enzimas de restricción correspondientes. El experto en la materia puede determinar los cebadores de PCR adecuados usando la información de secuencias proporcionada en GenBank u otras fuentes. También se pueden añadir sitios apropiados de enzimas de restricción al ácido nucleico que codifica la proteína sialiltransferasa o una subsecuencia de proteína de la misma por mutagénesis dirigida. El plásmido que contiene la secuencia o subsecuencia de nucleótidos que codifica la sialiltransferasa se corta con las endonucleasas de restricción apropiadas y después se liga en un vector apropiado para amplificación y/o expresión según métodos

estándar. Se encuentran ejemplos de técnicas suficientes para dirigir a expertos en la materia a través de métodos de amplificación *in vitro* en Berger, Sambrook y Ausubel, así como Mullis *et al.*, (1987) patente en EE UU No. 4,683,202; *PCR Protocols A Guide to Methods and Applications* (Innis *et al.*, eds) Academic Press Inc. San Diego, CA (1990) (Innis); Arnheim & Levinson (1 de Octubre, 1990) *C&EN* 36-47; *The Journal Of NIH Research* (1991) 3: 81-94; (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86: 1173; Guatelli *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87, 1874; Lomell *et al.* (1989) *J. Clin. Chem.*, 35: 1826; Landegren *et al.*, (1988) *Science* 241: 1077-1080; Van Brunt (1990) *Biotechnology* 8: 291-294; Wu y Wallace (1989) *Gene* 4: 560; y Barringer *et al.* (1990) *Gene* 89: 117.

Algunos ácidos nucleicos que codifican proteínas sialiltransferasas se pueden amplificar usando cebadores de PCR basados en la secuencia de proteínas sialiltransferasa previamente identificadas, por ejemplo, Cst-I (véase, por ejemplo, la patente en EE UU No. 6.689.604); Cst-II (véase, por ejemplo, la patente en EE UU No. 6.503.744) y Cst-III. Los ejemplos de cebadores de PCR que se pueden usar para amplificar ácidos nucleicos que codifican proteínas sialiltransferasas incluyen los siguientes pares de cebadores:

Para ácidos nucleicos de Cst-I:

CJ18F: 5' (41-mero, sitio Ndel en cursiva)

5' C TTA GGA GGT CAT ATG ACA AGG ACT AGA ATG GAA AAT GAA C 3'

y

CJ40R: 3' con cola de 6 His (60-mero, sitio Sall en cursiva, etiqueta (His)₆ en negrita)

5' CC TAG GTC GAC TCA TTA **GTG GTG ATG GTG GTG ATG** TTC CCC TTT CTC AAA CTC TCT CTT C 3';

Para ácidos nucleicos de Cst-II:

CJ-131: 5' CTTAGGAGGTCATATGAAAAAGTTATTATTGCTGGAAATG 3' y

CJ-132: 5' CCTAGGTCGACTTATTTTCCTTTGAAATAATGCTTTATATC 3';

Para ácidos nucleicos de Cst-III:

CstH-5p: 5' GGGGGGCATATGAGTATGAATATTAATGCTTTG 3' y

CstH-3p: 5' GGGGGGTCGACTCATTATCTATTTTTATTGTCATATTTTC 3'

En algunas bacterias, los ácidos nucleicos que codifican la proteína sialiltransferasa se pueden aislar amplificando un locus cromosómico específico, por ejemplo, el locus LOS de *C. jejuni*, y después identificar una sialiltransferasa típicamente encontrada en ese locus (véase, por ejemplo, la patente en EE UU No. 6.503.744). Los ejemplos de cebadores de PCR que se pueden usar para amplificar un locus LOS que comprende los ácidos nucleicos que codifican la proteína sialiltransferasa incluyen los siguientes pares de cebadores:

CJ42: cebador en heptosilTasa-II

5' GC CAT TAC CGT ATC GCC TAA CCA GG 3' 25-mero

CJ43: cebador en heptosilTasa-I

5' AAA GAA TAC GAA TTT GCT AAA GAG G 3' 25-mero.

Se pueden comparar otras propiedades físicas de un polipéptido sialiltransferasa recombinante expresado de un ácido nucleico particular a las propiedades de sialiltransferasas conocidas para proporcionar otro método de identificar secuencias o dominios adecuados del polipéptido sialiltransferasa que son determinantes de especificidad del sustrato aceptor y/o actividad catalítica. De forma alternativa, se puede mutar un gen de polipéptido sialiltransferasa putativo o gen de polipéptido sialiltransferasa recombinante y establecer su papel como una sialiltransferasa, o el papel de secuencias o dominios, detectando una variación en la estructura de un hidrato de carbono normalmente producido por el polipéptido sialiltransferasa, sin mutar, natural o control. Los expertos en la materia reconocerán que la mutación o modificación de polipéptidos sialiltransferasa de la invención puede estar facilitada por técnicas de biología molecular para manipular los ácidos nucleicos que codifican los polipéptidos sialiltransferasa, por ejemplo, PCR.

Se pueden identificar los dominios funcionales de polipéptidos sialiltransferasa recién identificados usando métodos estándar para mutar o modificar los polipéptidos y probar las actividades tales como actividad de sustrato aceptor y/o actividad catalítica, como se describe en el presente documento. Los dominios funcionales de los varios polipéptidos sialiltransferasa se pueden usar para construir ácidos nucleicos que codifican sialiltransferasas y los dominios funcionales de uno o más polipéptidos sialiltransferasa. Se puede probar después el sustrato aceptor deseado o la actividad catalítica de estas proteínas de fusión multi-sialiltransferasa.

En un planteamiento ejemplar para clonar ácidos nucleicos que codifican proteínas sialiltransferasa, las secuencias conocidas de ácido nucleico o aminoácidos de sialiltransferasas clonadas se alinean y comparan para determinar la cantidad de identidad de secuencia entre varias sialiltransferasas. Esta información se puede usar para identificar y seleccionar dominios de proteína que confieren o modulan actividades sialiltransferasa, por ejemplo, actividad de sustrato aceptor y/o actividad catalítica basándose en la cantidad de identidad de secuencia entre las sialiltransferasas de interés. Por ejemplo, se pueden usar dominios que tienen identidad de secuencia entre las

sialiltransferasas de interés y que se asocian con una actividad conocida, para construir proteínas sialiltransferasas que contienen ese dominio y que tienen la actividad asociada con ese dominio (por ejemplo, especificidad de sustrato aceptor y/o actividad catalítica).

5 V. Expresión de polipéptidos sialiltransferasa en células huésped

En preferidas los polipéptidos sialiltransferasa con etiqueta de la invención se expresan en células huéspedes *E. coli*. En una forma de realización más preferida, se usan las cepas JM109 o BNN93 de *E. coli* como células huéspedes.

10 En otra forma de realización preferida, se usa el vector de expresión pcWin2 para expresar la proteína sialiltransferasa con etiqueta en una célula huésped *E. coli*. El vector pcWin2 es conocido. Véase, por ejemplo, el documento WO/2005/067601 (2005).

15 Las proteínas sialiltransferasa de la invención se pueden expresar en una variedad de células huésped, incluyendo *E. coli*, otros huéspedes bacterianos y levadura. Las células huéspedes son preferiblemente microorganismos, tales como, por ejemplo, células de levadura, células bacterianas o células fúngicas filamentosas. Los ejemplos de células huéspedes adecuadas incluyen, por ejemplo, *Azotobacter sp.* (por ejemplo, *A. vinelandii*), *Pseudomonas sp.*, *Rhizobium sp.*, *Erwinia sp.*, *Escherichia sp.* (por ejemplo, *E. coli*), *Bacillus*, *Pseudomonas*, *Proteus*, *Salmonella*, *Serratia*, *Shigella*, *Rhizobia*, *Vitreoscilla*, *Paracoccus* y *Klebsiella sp.*, entre muchas otras. Las células pueden ser de cualquiera de varios géneros, incluyendo *Saccharomyces* (por ejemplo, *S. cerevisiae*), *Candida* (por ejemplo, *C. utilis*, *C. parapsilosis*, *C. krusei*, *C. versatilis*, *C. lipolytica*, *C. zeylanoides*, *C. guilliermondii*, *C. albicans* y *C. humicola*), *Pichia* (por ejemplo, *P. farinosa* y *P. ohmeri*), *Torulopsis* (por ejemplo, *T. candida*, *T. sphaerica*, *T. xylinus*, *T. famata* y *T. versatilis*), *Debaryomyces* (por ejemplo, *D. subglobosus*, *D. cantarellii*, *D. globosus*, *D. hansenii* y *D. japonicus*), *Zygosaccharomyces* (por ejemplo, *Z. rouxii* y *Z. bailii*), *Kluyveromyces* (por ejemplo, *K. marxianus*), *Hansenula* (por ejemplo, *H. anomala* y *H. jadinii*), y *Brettanomyces* (por ejemplo, *B. lambicus* y *B. anomalus*). Ejemplos de bacterias útiles incluyen, pero no están limitados a, *Escherichia*, *Enterobacter*, *Azotobacter*, *Erwinia*, *Klebsiella*, *Bacillus*, *Pseudomonas*, *Proteus*, y *Salmonella*.

30 Una vez expresados en una célula huésped, los polipéptidos sialiltransferasa se pueden usar para producir productos sialilados. Por ejemplo, los polipéptidos sialiltransferasa se pueden aislar usando técnicas de purificación de proteínas estándar y usarse en reacciones *in vitro* descritas en el presente documento para hacer productos sialilados. También se pueden usar los polipéptidos sialiltransferasa parcialmente purificados en reacciones *in vitro* para hacer productos sialilados como se pueden las células huésped permeabilizadas. También se pueden usar las células huéspedes en un sistema *in vitro* (por ejemplo, producción fermentativa) para producir productos sialilados.

Típicamente, el polinucleótido que codifica los polipéptidos sialiltransferasa se coloca bajo el control de un promotor que es funcional en la célula huésped deseada. Se conocen bien una variedad extremadamente amplia de promotores y se pueden usar en los vectores de expresión de la invención, dependiendo de la aplicación particular. Normalmente, el promotor seleccionado depende de la célula en que el promotor tiene que ser activo. También se incluyen, opcionalmente, otras secuencias de control de la expresión, tales como sitios de unión a ribosomas, sitios de terminación de la transcripción y similares. Las construcciones que incluyen una o más de estas secuencias de control se denominan "casetes de expresión". Según esto, la invención proporciona casetes de expresión en los que se incorporan los ácidos nucleicos que codifican proteínas de fusión para expresión a alto nivel en una célula huésped deseada.

Las secuencias de control de expresión que son adecuadas para su uso en una célula huésped particular con frecuencia se obtienen clonando un gen que se expresa en esa célula. Las secuencias de control procariontas comúnmente usadas, que se definen en el presente documento para que incluyan promotores para el inicio de la transcripción, opcionalmente con un operador, junto con secuencias de sitios de unión a ribosomas, incluyen tales promotores comúnmente usados como los sistemas de promotor de beta-lactamasa (penicilinasa) y lactosa (*lac*) (Change *et al.*, *Nature* (1977) 198: 1056), el sistema del promotor de triptófano (*trp*) (Goeddel *et al.*, *Nucleic Acids Res.* (1980) 8: 4057), el promotor *tac* (DeBoer, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* (1983) 80:21-25); y el promotor derivado de lambda P_L y el sitio de unión al ribosoma gen N (Shimatake *et al.*, *Nature* (1981) 292: 128). El sistema de promotor particular no es crítico para la invención, se puede usar cualquier promotor disponible que funcione en procariontas.

Para la expresión de las proteínas sialiltransferasas en células procariontas diferentes de *E. coli*, se requiere un promotor que funcione en la especie procarionta particular. Tales promotores se pueden obtener de genes que se han clonado de la especie, o se pueden usar promotores heterólogos. Por ejemplo, el promotor híbrido *trp-lac* funciona en *Bacillus* además de *E. coli*.

Se incluye convenientemente un sitio de unión a ribosomas (RBS) en los casetes de expresión de la invención. Un RBS en *E. coli*, por ejemplo, consiste en una secuencia de nucleótidos de 3-9 nucleótidos de longitud localizada de 3-11 nucleótidos en 5' del codón de iniciación (Shine y Dalgarno, *Nature* (1975) 254: 34; Steitz, *En Biological regulation and development: Gene expression* (ed. R.F. Goldberger), vol. 1, p. 349, 1979, Plenum Publishing, NY).

Se pueden usar promotores constitutivos o regulados en la presente invención. Los promotores regulados pueden ser ventajosos porque las células huésped se pueden hacer crecer hasta altas densidades antes de que se induzca la expresión de las proteínas de fusión. En algunas situaciones el alto nivel de expresión de proteínas heterólogas reduce el crecimiento celular. Un promotor inducible es un promotor que dirige la expresión de un gen donde el nivel de expresión es alterable por factores medioambientales o de desarrollo tales como, por ejemplo, temperatura, pH, condiciones anaerobias o aerobias, luz, factores de transcripción y agentes químicos. Tales promotores se denominan en el presente documento promotores "inducibles", que permiten controlar el tiempo de expresión de la glicosiltransferasa o enzima implicada en la síntesis del nucleótido azúcar. Los expertos en la materia conocen promotores inducibles para *E. coli* y otras células huéspedes bacterianas. Estos incluyen, por ejemplo, el promotor *lac*, el promotor del bacteriófago lambda P_L, el promotor híbrido *trp-lac* (Amann *et al.* (1983) *Gene* 25: 167; de Boer *et al.* (1983) *Proc. Nat'l. Acad. Sci. USA* 80: 21), y el promotor del bacteriófago T7 (Studier *et al.* (1986) *J. Mol. Biol.*; Tabor *et al.* (1985) *Proc. Nat'l. Acad. Sci. USA* 82: 1074-8). Estos promotores y su uso se discuten en Sambrook *et al.*, anteriormente. Un promotor inducible particularmente preferido para la expresión en procariontes es un promotor dual que incluye un componente del promotor *tac* unido a un componente promotor obtenido de un gen o genes que codifican enzimas implicadas en el metabolismo de la galactosa (por ejemplo, un promotor de un gen de UDP galactosa 4-epimerasa (*galE*)). El promotor dual *tac-gal*, que se describe en la publicación de la solicitud de patente PCT No. WO98/20111.

Una construcción que incluye un polinucleótido de interés operativamente unido a señales de control de expresión génicas que, cuando se coloca en una célula huésped apropiada, dirige la expresión del polinucleótido se denomina un "casete de expresión". Los casetes de expresión que codifican las proteínas de fusión de la invención con frecuencia se colocan en vectores de expresión para su introducción en la célula huésped. Los vectores típicamente incluyen, además del casete de expresión, una secuencia de ácido nucleico que permite que el vector se replique independientemente en una o más células huéspedes seleccionadas. En general, esta secuencia es una que permite que el vector se replique independientemente del ADN cromosómico del huésped e incluye orígenes de replicación o secuencias de replicación autónoma. Tales secuencias se conocen bien para una variedad de bacterias. Por ejemplo, el origen del replicación del plásmido pBR322 es adecuado para la mayoría de las bacterias gram negativas. De forma alternativa, el vector se puede replicar integrándose en el complemento genómico de la célula huésped y replicándose según la célula experimenta replicación de ADN. Un vector de expresión preferido para la expresión de las enzimas en células bacterianas es pTGK, que incluye un promotor dual *tac-gal* y que se describe en la publicación de la solicitud de patente PCT No. WO98/20111.

La construcción de construcciones de polinucleótidos en general requiere el uso de vectores capaces de replicarse en bacterias. Una plétora de kits están comercialmente disponibles para la purificación de plásmidos de bacterias (véase, por ejemplo, EasyPrepJ, FlexiPrepJ, ambos de Pharmacia Biotech; StrataCleanJ, de Stratagene; y el sistema de expresión QIAexpress, Qiagen). Los plásmidos aislados y purificados se pueden manipular después adicionalmente para producir otros plásmidos y usar para transfectar células. También es posible la clonación en *Streptomyces* o *Bacillus*.

Con frecuencia se incorporan marcadores de selección en los vectores de expresión usados para expresar los polinucleótidos de la invención. Estos genes pueden codificar un producto génico, tal como una proteína, necesario para la supervivencia o crecimiento de células huésped transfectadas cultivadas en un medio de cultivo selectivo. Las células huésped no transformadas con el vector que contiene el gen de selección no sobrevivirán en el medio de cultivo. Los genes de selección típicos codifican proteínas que confieren resistencia a antibióticos u otras toxinas, tales como ampicilina, neomicina, kanamicina, cloranfenicol o tetraciclina. De forma alternativa, los marcadores de selección pueden codificar proteínas que complementan deficiencias auxotróficas o suministran nutrientes críticos no disponibles de medios complejos, por ejemplo, el gen que codifica D-alanina racemasa para bacilos. Con frecuencia, el vector tendrá un marcador de selección que es funcional en, por ejemplo, *E. coli* u otras células en las que el vector se replica antes de ser introducido en la célula huésped. Los expertos en la materia conocen un número de marcadores de selección y se describen por ejemplo en Sambrook *et al.*, anteriormente.

La construcción de vectores adecuados que contienen uno o más de los componentes enumerados anteriormente emplea técnicas de ligación estándar como se describe en las referencias citadas anteriormente. Los plásmidos aislados o fragmentos de ADN se cortan, adaptan y vuelven a ligar en la forma deseada para generar los plásmidos requeridos. Para confirmar las secuencias correctas en los plásmidos construidos, los plásmidos se pueden analizar por técnicas estándar tales como digestión con endonucleasas de restricción y/o secuenciación según métodos conocidos. Las técnicas de clonación molecular para alcanzar estos extremos se conocen en la técnica. Los expertos en la materia conocen bien una amplia variedad de métodos de clonación y amplificación *in vitro* adecuados para la construcción de ácidos nucleicos recombinantes. Se encuentran ejemplos de estas técnicas e instrucciones suficientes para dirigir a expertos en la materia a través de muchos ejercicios de clonación en Berger y Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology* 152 Academic Press, Inc., San Diego, CA (Berger); y Current Protocols in Molecular Biology, F.M. Ausubel *et al.*, eds., *Current Protocols*, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (1994 Suplemento) (Ausubel).

Se conocen bien en la técnica una variedad de vectores comunes adecuados para su uso como materiales de partida para construir los vectores de expresión de la invención. Para clonación en bacterias, los vectores comunes incluyen vectores derivados de pBR322 tales como pBLUESCRIPT™ y vectores derivados del fago λ. En levaduras, los plásmidos incluyen vectores de integración en levadura (por ejemplo, YIp5) y plásmidos de replicación en levadura (los plásmidos de la serie YRp) y pGDP-2. La expresión en células de mamífero se puede alcanzar utilizando una variedad de plásmidos comúnmente disponibles, incluyendo pSV2, pBC12BI y p91023, así como vectores de virus líticos (por ejemplo, virus vaccinia, adenovirus y baculovirus), vectores de virus episomales (por ejemplo, virus del papiloma bovino) y vectores retrovíricos (por ejemplo, retrovirus murinos).

Los métodos para introducir los vectores de expresión en una célula huésped elegida no son particularmente críticos, y tales métodos los conocen los expertos en la materia. Por ejemplo, los vectores de expresión se pueden introducir en células procariontas, incluyendo *E. coli*, por transformación con cloruro de calcio, y en células eucariotas por tratamiento con fosfato de calcio o electroporación. Otros métodos de transformación también son adecuados.

Se puede usar acoplamiento de traducción para aumentar la expresión. La estrategia usa un marco abierto de lectura corto en 5' derivado de gen nativo muy expresado al sistema de traducción, que se coloca en 3' del promotor, y un sitio de unión a ribosomas seguido después de unos pocos codones de aminoácidos por un codón de terminación. Justo antes del codón de terminación está un segundo sitio de unión a ribosomas y después del codón de terminación está un codón para la iniciación de la traducción. El sistema disuelve la estructura secundaria en el ARN, lo que permite el inicio eficaz de la traducción. Véase, Squires, et. al. (1988), *J. Biol. Chem.* 263: 16297-16302.

Los polipéptidos sialiltransferasa se pueden expresar intracelularmente, o se pueden secretar de la célula. La expresión intracelular con frecuencia produce altos rendimientos. Si es necesario, la cantidad de proteína de fusión activa, soluble se puede aumentar realizando procedimientos de replegamiento (véase, por ejemplo, Sambrook *et al.*, anteriormente; Marston *et al.*, *Bio/Technology* (1984) 2: 800; Schoner *et al.*, *Bio/Technology* (1985) 3: 151). En formas de realización en las que los polipéptidos sialiltransferasa se secretan de la célula, en el periplasma o en el medio extracelular, la secuencia de ADN está unida a una secuencia de péptido señal cortable. La secuencia señal dirige la translocación de la proteína de fusión a través de la membrana celular. Un ejemplo de un vector adecuado para su uso en *E. coli* que contiene una unidad promotor-secuencia señal es pTA1529, que tiene el promotor y la secuencia señal de *phoA* de *E. coli* (véase, por ejemplo, Sambrook *et al.*, anteriormente; Oka *et al.*, *Proc. Natl. Acad. Sci. USA* (1985) 82: 7212; Talmadge *et al.*, *Proc. Natl. Acad. Sci. USA* (1980) 77: 3988; Takahara *et al.*, *J. Biol. Chem.* (1985) 260: 2670). En otra forma de realización, las proteínas de fusión se fusionan a una subsecuencia de proteína A o seroalbúmina bovina (BSA), por ejemplo, para facilitar la purificación, secreción o estabilidad.

Se puede expresar más de una proteína recombinante en una única célula huésped colocando múltiples casetes transcripcionales en un único vector de expresión, o utilizando diferentes marcadores de selección para cada uno de los vectores de expresión que se emplean en la estrategia de clonación.

VI. Purificación de los polipéptidos sialiltransferasa con etiqueta

Las proteínas sialiltransferasa de la presente invención se pueden expresar como proteínas intracelulares o como proteínas que se secretan de la célula, y se pueden usar en esta forma, en métodos de la presente invención. Por ejemplo, se puede usar un extracto celular crudo que contiene el polipéptido sialiltransferasa expresado intracelular o secretado en los métodos de la presente invención.

De forma alternativa, el polipéptido sialiltransferasa se puede purificar según procedimientos estándar de la técnica, incluyendo precipitación con sulfato de amonio, columnas de afinidad, cromatografía en columna, electroforesis en gel y similares (véase, en general, R. Scopes, *Protein Purification*, Springer-Verlag, N.Y. (1982), Deutscher, *Methods in Enzymology Vol. 182: Guide to Protein Purification.*, Academic Press, Inc. N.Y. (1990)). Se prefieren composiciones sustancialmente puras de al menos aproximadamente el 70, el 75, el 80, el 85, el 90% de homogeneidad, y el 92, el 95, del 98 al 99% o más de homogeneidad son más preferidas. Las proteínas purificadas también se pueden usar, por ejemplo, como inmunógenos para la producción de anticuerpos.

Para facilitar la purificación de los polipéptidos sialiltransferasa de la invención, los ácidos nucleicos que codifican las proteínas también pueden incluir una secuencia codificante para un segundo epítipo o "etiqueta epítipo" para el que está disponible un reactivo de unión de afinidad, es decir, una etiqueta de purificación. Ejemplos de epítopos adecuados incluyen los genes indicadores *myc* y V-5; los vectores de expresión útiles para la producción recombinante de proteínas de fusión que tienen estos epítopos están comercialmente disponibles (por ejemplo, los vectores pcDNA3.1/Myc-His y pcDNA3.1/V5-His de Invitrogen (Carlsbad CA) son adecuados para la expresión en células de mamífero). Los expertos en la materia conocen vectores de expresión adicionales adecuados para unir una etiqueta al polipéptido sialiltransferasa de la invención y los sistemas de detección correspondientes, y varios están comercialmente disponibles (por ejemplo, FLAG" (Kodak, Rochester NY). Otro ejemplo de una etiqueta adecuada es una secuencia de polihistidina, que es capaz de unirse a ligando de afinidad de quelatos metálicos. Típicamente, se usan seis histidinas adyacentes, aunque se pueden usar más o menos de seis. Los ligandos de afinidad de quelatos metálicos adecuados que pueden servir como el grupo de unión para una etiqueta de polihistidina incluyen ácido nitrilo-triacético (NTA) (Hochuli, E. (1990) "Purification of recombinant proteins with metal

chelating adsorbents" En Genetic Engineering: Principles and Methods, J.K. Setlow, Ed., Plenum Press, NY; comercialmente disponible de Qiagen (Santa Clarita, CA)).

5 Las etiquetas de purificación también incluyen dominios de unión a maltosa y dominios de unión a almidón. Los expertos en la materia conocen la purificación de proteínas con dominios de unión a maltosa. Los dominios de unión a almidón se describen en el documento WO 99/15636. La purificación de afinidad de una proteína de fusión que comprende un dominio de unión a almidón usando una resina de betaciclodextrina (BCD) derivada se describe en el documento USSN 60/468.374, presentada el 5 de mayo, 2003.

10 Los expertos en la materia conocen otros haptenos que son adecuados para su uso como etiquetas y se describen, por ejemplo, en Handbook of Fluorescent Probes and Research Chemicals (6ª Ed., Molecular Probes, Inc., Eugene OR). Por ejemplo, dinitrofenol (DNP), digoxigenina, barbituratos (véase, por ejemplo, la patente en EE UU No. 5.414.085) y varios tipos de fluoróforos son útiles como haptenos, como lo son derivados de estos compuestos. Hay kits comercialmente disponibles para unir haptenos y otros grupos a proteínas y otras moléculas. Por ejemplo, cuando el hapteno incluye un tiol, se puede usar un enlazador heterobifuncional tal como SMCC para unir la etiqueta a un residuo de lisina presente en el reactivo de captura.

20 El experto en la materia reconocerá que se pueden hacer modificaciones a los dominios catalítico o funcional del polipéptido sialiltransferasa sin disminuir su actividad biológica. Se pueden hacer algunas modificaciones para facilitar la clonación, expresión o incorporación del dominio catalítico en una proteína de fusión. Los expertos en la materia conocen bien tales modificaciones e incluyen, por ejemplo, la adición de codones en cualquier extremo del polinucleótido que codifica el dominio catalítico para proporcionar, por ejemplo, una metionina añadida en el extremo amino para proporcionar un sitio de iniciación, o aminoácidos adicionales (por ejemplo, poli His) colocados en cualquier extremo para crear sitios de enzimas de restricción convenientemente colocados o codones de terminación o secuencias de purificación.

25 Una vez purificadas, las proteínas sialiltransferasa de la invención típicamente se almacenan a baja temperatura, por ejemplo, -20°C. Una manera de aumentar la estabilidad de y disminuir la agregación de las proteínas sialiltransferasa con etiqueta es almacenar las proteínas en soluciones que incluyen concentraciones relativamente altas de glicerol. Por ejemplo, las proteínas sialiltransferasas con etiqueta de la invención se pueden almacenar en soluciones que comprenden entre el 20% y el 95% de glicerol. Por ejemplo, las soluciones de almacenamiento pueden incluir, por ejemplo, el 20%, el 25%, el 30%, el 35%, el 40%, el 45%, el 50%, el 55%, el 60%, el 65%, el 70%, el 75%, el 80%, el 85%, el 90% o el 95% de glicerol. En una forma de realización preferida, la solución de almacenamiento incluye entre el 40% y el 60% de glicerol. En otra forma de realización preferida, la solución de almacenamiento incluye glicerol aproximadamente al 50%. Si es necesario, las proteínas sialiltransferasa con etiqueta purificadas se pueden introducir en un tampón o solución de almacenamiento que incluya la concentración de glicerol deseada. Los expertos en la materia son conscientes de métodos para cambiar tampones o soluciones de almacenamiento, por ejemplo, cromatografía de exclusión molecular, diálisis y filtración, por ejemplo, filtración de flujo tangencial. Estos métodos de intercambio de tampón se pueden practicar en cualquier paso del protocolo de purificación de proteínas, no solo como un paso final antes del almacenamiento de la proteína purificada.

VI. Proteínas de fusión de sialiltransferasa

45 En algunas formas de realización, las células recombinantes de la invención expresan proteínas de fusión que tienen más de una actividad enzimática que está implicada en la síntesis de un oligosacárido sialilado deseado. Los polipéptidos de fusión pueden estar compuestos de, por ejemplo, un polipéptido sialiltransferasa con etiqueta que se une a una enzima accesoria, por ejemplo CMP-ácido siálico sintasa. Las proteínas de fusión también se pueden hacer usando dominios catalíticos u otros truncamientos de las enzimas. Por ejemplo, un polinucleótido que codifica un polipéptido sialiltransferasa se puede unir, en el mismo marco de lectura, a un polinucleótido que codifica una enzima implicada en la síntesis de CMP-ácido siálico. La proteína de fusión resultante puede catalizar después no solo la síntesis de la molécula de ácido siálico activada, sino también la transferencia del grupo ácido siálico a la molécula aceptora. La proteína de fusión puede ser dos o más enzimas del ciclo del ácido siálico unidas en una secuencia de nucleótidos expresable. Los polipéptidos de fusión de sialiltransferasa de la presente invención se pueden diseñar y fabricar fácilmente utilizando varias técnicas de ADN recombinante que conocen bien los expertos en la materia. Se describen proteínas de fusión ejemplares en la solicitud de patente PCT PCT/CA98/01180, que se publicó como WO99/31224 el 24 de junio, 1999 y que divulga CMP-ácido siálico sintasa de *Neisseria* fusionada con una α 2,3 sialiltransferasa de *Neisseria*. Los expertos en la materia reconocerán que se pueden fusionar muchos otros polipéptidos CMP-ácido siálico sintasa y sialiltransferasas para su uso en la invención. En algunas formas de realización, se fusiona una CMP-ácido siálico sintasa de *Neisseria* a una sialiltransferasa de *C. jejuni*. La sialiltransferasa de *C. jejuni* (Cst) puede ser una enzima CstI, CstII o CstIII. Se puede usar una versión de longitud completa o truncada del polipéptido sialiltransferasa de *C. jejuni* en la proteína de fusión de sialiltransferasa. En algunas formas de realización, se expresa más de un polipéptido de fusión de sialiltransferasa en la célula.

65 En algunas formas de realización, las células recombinantes de la invención expresan proteínas de fusión que tienen más de una actividad enzimática que está implicada en la adición de al menos un residuo de azúcar adicional, por

ejemplo, un residuo no de ácido siálico. Estos polipéptidos de fusión pueden estar compuestos de, por ejemplo, un dominio catalítico de una glicosiltransferasa, por ejemplo, no una sialiltransferasa, que se une a un dominio catalítico de una enzima accesoria. El dominio catalítico de la enzima accesoria puede, por ejemplo, catalizar un paso en la formación de un nucleótido azúcar que es un donante para la glicosiltransferasa o una reacción implicada en un ciclo de glicosiltransferasa. Por ejemplo, un polinucleótido que codifica una glicosiltransferasa se puede unir, en el mismo marco de lectura, a un polinucleótido que codifica una enzima implicada en la síntesis de un nucleótido azúcar. La proteína de fusión resultante puede catalizar después no solo la síntesis del nucleótido azúcar, sino también la transferencia del grupo azúcar a la molécula aceptora. La proteína de fusión puede ser dos o más enzimas del ciclo unidas en una secuencia de nucleótidos expresable. Los polipéptidos de la presente invención se pueden diseñar y fabricar fácilmente utilizando varias técnicas de ADN recombinante que conocen bien los expertos en la materia. Se describen proteínas de fusión adecuadas en la solicitud de patente PCT PCT/CA98/01180, que se publicó como WO99/31224 el 24 de junio, 1999, e incluye, por ejemplo, una UDP glucosa epimerasa fusionada en el mismo marco de lectura a una galactosiltransferasa.

15 VIII. Sustratos donantes y sustratos aceptores

Los sustratos donantes adecuados usados por los polipéptidos sialiltransferasa con etiqueta y otras glicosiltransferasas en los métodos de la invención incluyen, pero no están limitados a, UDP-Glc, UDP-GlcNAc, UDP-Gal, UDP-GalNAc, GDP-Man, GDP-Fuc, UDP-GlcNAc y CMP-ácido siálico y grupos de ácido siálico activado. Guo *et al.*, *Applied Biochem. and Biotech.* **68**: 1-20 (1997).

Típicamente, los sustratos aceptores incluyen un residuo de galactosa terminal para la adición de un residuo de ácido siálico mediante un enlace α 2,3. Para la adición de un residuo de ácido siálico en un enlace α 2,8, se une un segundo residuo de ácido siálico a un primer ácido siálico mediante un enlace α 2,8. Los ejemplos de aceptores adecuados incluyen una Gal terminal que se une a GlcNAc o Glc por un enlace β 1,4 y una Gal terminal que se une por β 1,3 a GlcNAc o GalNAc. Los aceptores adecuados incluyen, por ejemplo, aceptores galactosilo tales como Gal β 1,4GlcNAc, Gal β 1,4GalNAc, Gal β 1,3GalNAc, lacto-N-tetraosa, Gal β 1,3GlcNAc, Gal β 1,3Ara, Gal β 1,6GlcNAc, Gal β 1,4Glc (lactosa) y otros aceptores que conocen los expertos en la materia (véase, por ejemplo, Paulson *et al.*, *J. Biol. Chem.* **253**: 5617-5624 (1978)). El residuo terminal al que se une el ácido siálico puede estar él mismo unido a, por ejemplo, H, un sacárido, oligosacárido o un grupo aglucona que tiene al menos un átomo glucídico. En algunas formas de realización, el residuo aceptor es una parte de un oligosacárido que está unido a una proteína, lípido o proteoglicano, por ejemplo.

Los sustratos aceptores adecuados usados por los polipéptidos sialiltransferasa c y los métodos de la invención incluyen, pero no están limitados a, polisacáridos y oligosacáridos. Por ejemplo, la lactosa se puede sialilar para formar una sialilactosa, por ejemplo, 3' sialilactosa. Las sialiltransferasas descritas en el presente documento también se pueden usar en sistemas multienzimáticos para producir un producto deseado a partir de un material de partida conveniente.

Los sustratos aceptores adecuados usados por los polipéptidos sialiltransferasa y los métodos de la invención incluyen, pero no están limitados a, proteínas, lípidos, gangliósidos y otras estructuras biológicas (por ejemplo, células enteras) que se pueden modificar por los métodos de la invención. Estos sustratos aceptores típicamente comprenderán las moléculas de polisacáridos u oligosacáridos descritas anteriormente. Las estructuras ejemplares, que se pueden modificar por los métodos de la invención incluyen cualquiera de un número de estructuras de glicolípidos, glicoproteínas e hidratos de carbono en células que conocen los expertos en la materia.

Los ejemplos de sustratos aceptores adecuados usados en reacciones catalizadas por sialiltransferasa, y los ejemplos de sustratos aceptores adecuados usados en reacciones catalizadas por sialiltransferasa se describen en Guo *et al.*, *Applied Biochem. and Biotech.* **68**: 1-20 (1997), pero no están limitados a los mismos.

La presente invención proporciona polipéptidos sialiltransferasa que se seleccionan por su capacidad para producir oligosacáridos, glicoproteínas y glicolípidos que tienen los grupos de oligosacáridos deseados. De forma similar, si están presentes, las enzimas accesorias se eligen basándose en un sustrato azúcar activado deseado o en un azúcar encontrado en el oligosacárido producto.

Para la síntesis de glicoproteínas, se pueden identificar fácilmente polipéptidos sialiltransferasa adecuados haciendo reaccionar varias cantidades de un polipéptido sialiltransferasa de interés (por ejemplo, 0,01-100 mU/mg de proteína) con una glicoproteína (por ejemplo, de 1-10 mg/ml) a la que se une un oligosacárido que tiene un potencial sitio aceptor para glicosilación por la sialiltransferasa de interés. Se comparan las capacidades de proteínas sialiltransferasa recombinantes de la presente invención para añadir un residuo de azúcar en el sitio aceptor deseado, y se selecciona un polipéptido sialiltransferasa que tiene la propiedad deseada (por ejemplo, especificidad de sustrato aceptor o actividad catalítica).

En general, la eficacia de la síntesis enzimática de oligosacáridos, glicoproteínas y glicolípidos, que tienen los grupos oligosacáridos sialilados deseados, se puede aumentar mediante el uso de polipéptidos sialiltransferasa recombinantemente producidos de la presente invención. Las técnicas recombinantes permiten la producción de

polipéptidos sialiltransferasa recombinantes en las grandes cantidades que se requieren para la modificación de glicoproteínas y glicolípidos a gran escala *in vitro*.

5 En algunas formas de realización, los oligosacáridos, glicoproteínas y glicolípidos adecuados para su uso por los polipéptidos sialiltransferasa y los métodos de la invención pueden ser glicoproteínas y glicolípidos inmovilizados sobre un soporte sólido durante la reacción de glicosilación. El término "soporte sólido" también abarca soportes semisólidos. Preferiblemente, la glicoproteína o glicolípido diana se inmoviliza de forma reversible de modo que la glicoproteína o glicolípido respectivos se puedan liberar después de completar la reacción de glicosilación. Los expertos en la materia conocen muchas matrices adecuadas. Se pueden emplear, por ejemplo, intercambio iónico para inmovilizar temporalmente una glicoproteína o glicolípido sobre una resina apropiada mientras sigue la reacción de glicosilación. También se puede usar un ligando que se une específicamente a la glicoproteína o glicolípido de interés para la inmovilización por afinidad. Por ejemplo, los anticuerpos que se unen específicamente a una glicoproteína son adecuados. Además, donde la glicoproteína de interés es ella misma un anticuerpo o contiene un fragmento del mismo, se puede usar proteína A o G como la resina de afinidad. También son adecuados colorantes y otras moléculas que se unen específicamente a una glicoproteína o glicolípido de interés.

Preferiblemente, cuando el sacárido aceptor es una versión truncada de la glicoproteína de longitud total, preferiblemente incluye la subsecuencia biológicamente activa de la glicoproteína de longitud total. Subsecuencias biológicamente activas ejemplares incluyen, pero no están limitadas a, sitios activos de enzimas, sitios de unión a receptores, sitios de unión a ligandos, regiones determinantes de complementariedad de anticuerpos y regiones antigénicas de antígenos.

En una forma de realización las proteínas sialiltransferasa con etiqueta de la invención se usan para añadir residuos de ácido siálico a una proteína eritropoyetina. Los residuos de ácido siálico añadidos pueden estar modificados, por ejemplo, un residuo de ácido siálico PEGilado. Las proteínas sialiltransferasa con etiqueta se pueden purificar antes de usarlas en una reacción de sialilación. Los métodos enzimáticos de adición de residuos de ácido siálico, incluyendo ácido siálico PEGilado, a proteínas tales como eritropoyetina, se divulgan en el documento WO 2006/127910.

30 IX. Producción de productos sialilados

Los polipéptidos sialiltransferasa con etiqueta se pueden usar para hacer productos sialilados en mezclas de reacción *in vitro* o por reacciones *in vivo*, por ejemplo, por crecimiento fermentativo de microorganismos recombinantes que comprenden los nucleótidos que codifican los polipéptidos sialiltransferasa con etiqueta.

35 A. Reacciones *in vitro*

Los polipéptidos sialiltransferasa con etiqueta se pueden usar para hacer productos sialilados en mezclas de reacción *in vitro*. Las mezclas de reacción *in vitro* pueden incluir microorganismos permeabilizados que comprenden los polipéptidos sialiltransferasa, polipéptidos sialiltransferasa parcialmente purificados o polipéptidos sialiltransferasa purificados; así como sustratos donantes, sustratos aceptores y tampones de reacción apropiados. Para las reacciones *in vitro*, las proteínas glicosiltransferasa recombinantes, tales como polipéptidos sialiltransferasa, sustratos aceptores, sustratos donantes y otros ingredientes de la mezcla de reacción se combinan mediante mezcla en un medio de reacción acuoso. Se pueden usar glicosiltransferasas adicionales en combinación con los polipéptidos sialiltransferasa, dependiendo del producto sialilado deseado. El medio generalmente tiene un valor de pH desde aproximadamente 5 hasta aproximadamente 8,5. La selección de un medio se basa en la capacidad del medio de mantener el valor de pH al nivel deseado. Por tanto, en algunas formas de realización, el medio está tamponado a un valor de pH de aproximadamente 7,5. Si no se usa un tampón, el pH del medio se debe mantener desde aproximadamente 5 hasta 8,5 dependiendo de la glicosiltransferasa particular usada. Para fucosiltransferasas, el intervalo de pH preferiblemente se mantiene desde aproximadamente 6,0 hasta 8,0. Para sialiltransferasas, el intervalo es preferiblemente desde aproximadamente 5,5 hasta aproximadamente 8,0.

Las cantidades o concentraciones de enzima se expresan en unidades de actividad, que una medida de la velocidad inicial de catálisis. Una unidad de actividad cataliza la formación de 1 μmol de producto por minuto a una temperatura (típicamente 37°C) y valor de pH (típicamente 7,5) determinados. Por tanto, 10 unidades de una enzima es una cantidad catalítica de esa enzima donde 10 μmoles de sustrato se convierten en 10 μmoles de producto en un minuto a una temperatura de 37°C y un valor de pH de 7,5.

La mezcla de reacción puede incluir cationes metálicos divalentes (Mg^{2+} , Mn^{2+}), El medio de reacción también puede comprender detergentes solubilizantes (por ejemplo, Triton o SDS) y solventes orgánicos tales como metanol o etanol, si es necesario. Las enzimas se pueden utilizar libres en solución o se pueden unir a un soporte tal como un polímero. Por tanto, la mezcla de reacción es sustancialmente homogénea al principio, aunque se puede formar algo de precipitado durante la reacción.

65 La temperatura a la que lleva a cabo un proceso anterior puede variar desde justo por encima de la congelación a la temperatura a la que la enzima más sensible se desnaturaliza. Ese intervalo de temperatura es preferiblemente

desde aproximadamente 0°C hasta aproximadamente 45°C, y más preferiblemente desde aproximadamente 20°C hasta aproximadamente 37°C.

La mezcla de reacción así formada se mantiene durante un periodo de tiempo suficiente para obtener el alto rendimiento deseado de determinantes de oligosacáridos deseados presentes en los grupos oligosacáridos unidos a la proteína que se va a glicosilar. Para preparaciones a gran escala, con frecuencia se dejará que la reacción siga durante aproximadamente 0,5-240 horas, y más típicamente entre aproximadamente 1-18 horas.

B. Reacciones in vivo

Los polipéptidos sialiltransferasa con etiqueta se pueden usar para hacer productos sialilados en reacciones in vivo, por ejemplo, crecimiento fermentativo de microorganismos recombinantes que comprenden los polipéptidos sialiltransferasa. El crecimiento fermentativo de microorganismos recombinantes se puede producir en presencia de medio que incluye un sustrato aceptor y un sustrato donante o un precursor de un sustrato donante, por ejemplo, ácido siálico. Véase, por ejemplo, Priem *et al.*, *Glycobiology* 12:235-240 (2002). El microorganismo absorbe el sustrato aceptor y el sustrato donante o el precursor de un sustrato donante y la adición del sustrato donante al sustrato aceptor tiene lugar en la célula viva. El microorganismo se puede alterar para facilitar la absorción del sustrato aceptor, por ejemplo, expresando una proteína transportadora de azúcar. Por ejemplo, donde la lactosa es el sacárido aceptor, se pueden usar células de *E. coli* que expresan la permeasa LacY. Se pueden usar otros métodos para disminuir la rotura de un sacárido aceptor o para aumentar la producción de un sacárido donante o un precursor del sacárido donante. En algunas formas de realización, la producción de productos sialilados aumenta por la manipulación del microorganismo huésped. Por ejemplo, en *E. coli*, se puede minimizar la rotura del ácido siálico usando un cepa huésped que carece de CMP-sialato sintasa (*NanA*). (En *E. coli*, la CMP-sialato sintasa parece ser una enzima catabólica). También en *E. coli*, cuando la lactosa es, por ejemplo, el sacárido aceptor o un intermedio en la síntesis del producto sialilado, se puede minimizar la rotura de la lactosa usando células huésped que son *LacZ*⁻. Los métodos para la síntesis in vivo de oligosacáridos, incluyendo oligosacáridos que contienen ácido siálico, se encuentran en, por ejemplo, Samain y Priem WO/2001/004341 (2001) y Johnson *et al.*, WO/2006/034225 (2006).

C. Caracterización de y aislamiento de productos sialilados

La producción de productos sialilados se puede seguir, por ejemplo, determinando que la producción de los productos deseados se ha producido o determinando que un sustrato tal como el sustrato aceptor se ha agotado. Los expertos en la materia reconocerán que los productos sialilados tal como un oligosacárido, se pueden identificar usando técnicas tales como cromatografía, por ejemplo usando papel o placas de TLC, o por espectrometría de masas, por ejemplo, espectrometría MALDI-TOF o por espectroscopía de RMN. Los expertos en la materia conocen los métodos de identificación de los productos sialilados y se encuentran, por ejemplo, en la patente en EE UU No. 6.699.705 y en Varki *et al.*, *Preparation and Analysis of Glycoconjugates*, en Current Protocols in Molecular Biology, Capítulo 17 (Ausubel *et al.* eds, 1993).

En algunas formas de realización, los polipéptidos sialiltransferasa con etiqueta y los métodos de la presente invención se usan para sintetizar enzimáticamente una glicoproteína o glicolípido que tiene un patrón de glicosilación sustancialmente uniforme. Las glicoproteínas y glicolípidos incluyen un sacárido u oligosacárido que se une a una proteína, glicoproteína, lípido o glicolípido para el que se desea una alteración de glicoforma. El sacárido u oligosacárido incluye una estructura que puede funcionar como un sustrato aceptor para una glicosiltransferasa. Cuando el sustrato aceptor se glicosila, el grupo oligosacárido deseado se forma. El grupo oligosacárido deseado es uno que imparte la actividad biológica deseada a la glicoproteína o glicolípido al que se une. En las composiciones de la invención, el residuo de sacárido preseleccionado se une a al menos aproximadamente el 30% de los potenciales sitios aceptores de interés. Más preferiblemente, el residuo de sacárido preseleccionado se une a al menos aproximadamente el 50% de los potenciales sustratos aceptores de interés y aún más preferiblemente a al menos el 70% de los potenciales sustratos aceptores de interés. En situaciones en las que la glicoproteína o glicolípido de partida muestra heterogeneidad en el grupo de oligosacárido de interés (por ejemplo, algunos de los oligosacáridos en la glicoproteína o glicolípido de partida ya tienen el residuo de sacárido preseleccionado unido al sustrato aceptor de interés), los porcentajes descritos incluyen tales residuos de sacáridos preunidos.

El término "alterado" se refiere a la glicoproteína o glicolípido de interés que tienen un patrón de glicosilación que, después de la aplicación de los polipéptidos sialiltransferasa y los métodos de la invención, es diferente del observado en la glicoproteína originalmente producida. Un ejemplo de tales glicoconjugados son glicoproteínas en las que las glicoformas de las glicoproteínas son diferentes de las encontradas en la glicoproteína cuando se produce por células del organismo al que la glicoproteína es nativa. También se proporcionan polipéptidos sialiltransferasa y métodos de usar tales proteínas para sintetizar enzimáticamente glicoproteínas y glicolípidos en los que el patrón de glicosilación de estos glicoconjugados se modifica comparado con el patrón de glicoconjugados originalmente producidos por una célula huésped, que puede ser de la misma especie o de una diferente que las células donde se producen los glicoconjugados nativos.

Se pueden evaluar diferencias en los patrones de glicosilación no solo mediante un análisis estructural de las glicoproteínas y glicolípidos, sino también mediante la comparación de una o más actividades biológicas de los glicoconjugados. Por ejemplo, una glicoproteína que tiene una “glicofoma alterada” incluye una que exhibe una mejora en una o más actividades biológicas de la glicoproteína después de la reacción de glicosilación comparada con la glicoproteína sin modificar. Por ejemplo, un glicoconjugado alterado incluye uno que, después de aplicar los polipéptidos sialiltransferasa y los métodos de la invención, muestra mayor afinidad de unión a un ligando o receptor de interés, una mayor semivida terapéutica, antigenicidad reducida, y direccionamiento a tejidos específicos. La cantidad de mejora observada preferiblemente es estadísticamente significativa, y es más preferiblemente al menos aproximadamente un 25% de mejora, y aún más preferiblemente es al menos aproximadamente el 30%, el 40%, el 50%, el 60%, el 70% e incluso aún más preferiblemente es al menos el 80%, el 90% o el 95%.

Los productos producidos usando polipéptidos sialiltransferasa se pueden usar sin purificación. Sin embargo, se pueden usar técnicas estándar, bien conocidas, por ejemplo, cromatografía en capa fina o gruesa, cromatografía de intercambio iónico o filtración en membrana, para la recuperación de sacáridos glicosilados. Además, por ejemplo, se puede usar la filtración en membrana, que utiliza una membrana de nanofiltración u osmótica inversa como se describe en la patente en AU No. 735695 cedida en común. Como un ejemplo adicional, se puede usar la filtración en membrana en donde las membranas tienen un límite de peso molecular de aproximadamente 1000 hasta aproximadamente 10.000 Dalton para eliminar proteínas. Como otro ejemplo, se puede usar después nanofiltración u ósmosis inversa para eliminar sales. Las membranas nanofiltro son una clase de membranas de ósmosis inversa que pasan sales monovalentes pero retienen sales polivalentes y solutos sin carga mayores de aproximadamente 200 hasta aproximadamente 1000 dalton, dependiendo de la membrana usada. Por tanto, por ejemplo, los oligosacáridos producidos por las composiciones y métodos de la presente invención pueden ser retenidos en la membrana y las sales contaminantes la atravesarán.

Se debe advertir que como se usa en el presente documento y en las reivindicaciones adjuntas, las formas singulares “un/una” y “el/la” incluyen referentes plurales a menos que el contexto claramente dicte otra cosa. Por tanto, por ejemplo, la referencia a “un ácido nucleico” incluye una pluralidad de tales ácidos nucleicos y la referencia a “el polipéptido” incluye referencias a uno o más polipéptidos y equivalentes del mismo que conocen los expertos en la materia, etcétera.

Las publicaciones discutidas en el presente documento se proporcionan solamente por su divulgación antes de la fecha de presentación de la presente solicitud. Nada en el presente documento se debe interpretar como una admisión de que la presente invención no tiene derecho a preceder tal publicación en virtud de una invención anterior. Además, las fechas de publicación proporcionadas pueden ser diferentes de las fechas de publicación reales que puede ser necesario confirmar independientemente.

Ejemplos

Ejemplo 1: Construcción de vectores de expresión de CSTII

Se aisló el ácido nucleico de la sialiltransferasa II de *Campylobacter* (CSTII) del plásmido pcWori/CSTII. Véase, por ejemplo, Gilbert, *et al.*, *J. Biol. Chem.*, 275: 3896-3906 (2000). Las células huésped de *E. coli* se cultivaron en 250 ml de medio L Martone con ampicilina 10 mg/ml durante aproximadamente 16 horas. Las células se recogieron después por centrifugación a 5000 RPM durante 15 minutos a 4°C. El ADN de plásmido se aisló como se describe en el manual que acompaña el kit Qiagen HiSpeed Plasmid Maxi.

Usando la técnica de PCR, se añadió un sitio BamHI inmediatamente antes del ATG de iniciación de CstII. Se introdujeron codones de terminación dobles seguidos por un sitio XhoI en el extremo 3' del gen CstII. La proteína codificada por esta construcción abarca la secuencia de la proteína natural para Cst-II con tres aminoácidos adicionales (MGS) en el extremo amino (SEQ ID NO:2). Después de subclonar el fragmento BamHI-XhoI en el vector de expresión pcWin2, se transformaron células *E. coli* competentes (células TOP10, Invitrogen) con la mezcla de ligación.

Se seleccionaron transformantes en placas de agar de LB de Martone que contenían 50 microgramos/ml de kanamicina. Se seleccionaron colonias de las placas de kanamicina y se cultivaron en medio líquido con kanamicina. Se aisló el ADN de plásmido usando un kit de miniprep Qiaprep de Qiagen. El ADN de plásmido se incubó con enzimas de digestión de restricción NdeI y XhoI, ambas de New England Biolabs. El ADN digerido se analizó en un gel de agarosa para determinar el patrón de restricción del plásmido. Se seleccionó un plásmido con el patrón de restricción esperado para la inserción de CSTII en pcWin2 y se secuenció en ambas hebras, confirmando la secuencia de ácido nucleico de CSTII esperada.

Se construyó un segundo plásmido de CSTII (pNT2) para eliminar el codón de iniciación en 5'. pNT2 se generó de pNT1 usando PCR y los cebadores apropiados. La secuencia de pNT2 se determinó por secuenciación de doble cadena. Las secuencias de ácido nucleico de CSTII 5' de pNT1 y pNT2 se muestran en la figura 1, como lo son las secuencias de aminoácidos N-terminales.

Ejemplo 2: Expresión de la proteína CSTII

- Se probó la expresión de la proteína CSTII de los plásmidos pNT1 y pNT2 cultivando células BNN93 transformadas a una escala de 10 L y aislando proteína de las células. El plásmido de expresión pNT2 no expresó proteína CSTII.
- 5 El plásmido de expresión pNT1 si expresó proteína CSTII (datos no mostrados). Se analizó la secuencia de la proteína CSTII expresada del plásmido pNT1. La proteína CSTII expresada empieza con MGS MKK y también se expresa establemente con GSMKK y MKK en el extremo N de la proteína, confirmado por secuenciación N-terminal de la proteína.
- 10 Se determinaron las condiciones de fermentación para expresión mejorada de CSTII a partir de pNT1. Los parámetros probados incluían: 1) fermentación en medio simple (designado E1.0) o medio rico (designado E6.0); 2) selección de la línea de células huésped K-12 de *E. coli* usando cepas JM109 o BNN93 transformadas con pNT1; y 3) modo de regulación del oxígeno disuelto (OD) durante la fermentación.
- 15 Se cultivó un cultivo de inoculación de células de *E. coli* transformadas con pNT1 en medio LB soja con kanamicina. El cultivo de inoculación se usó para inocular un cultivo de producción de diez litros de, por ejemplo, medio simple o rico. Se siguió el crecimiento bacteriano leyendo la densidad óptica del cultivo a 600 nm.
- 20 Cuando la DO600 del cultivo de producción estaba entre 0,6 y 1,0, se indujo la expresión de CSTII mediante la adición de IPTG para alcanzar una concentración final de IPTG de 1,0 mM. Los cultivos inducidos se incubaron durante la noche a 37°C para una incubación total de veinticuatro horas. Se retiraron alícuotas de cincuenta o doscientos ml de los cultivos de producción empezando cuatro horas después de la adición de IPTG y se siguió a intervalos hasta la recogida final.
- 25 En la recogida, se midió la DO de los cultivos de *E. coli*. Se recogieron las células de *E. coli* del cultivo de producción por centrifugación a 5000 rpm durante veinte minutos. Los sobrenadantes se desecharon; los precipitados celulares se pesaron y almacenaron a -80°C.
- 30 Los precipitados se descongelaron y resuspendieron en Tris HCl 20 mM, pH 8,5 a una relación de 10,0 ml/g de peso húmedo. Las células se rompieron mediante dos pases a través de un Microfluidizador empaquetado en hielo a 14.000 PSI. Las células lisadas se centrifugaron a 5000 RPM durante 20 minutos. Los sobrenadantes se separaron de los precipitados y se filtró 1,0 ml de sobrenadante a través de filtro jeringuilla de 0,2 micrómetros. Las células lisadas se usaron tanto para análisis por SDS-PAGE y determinaciones de la actividad enzimática.
- 35 Se corrieron muestras de proteína en geles Nu-Page usando geles de Bis-Tris al 12% ya polimerizados de Novex en Novex XCELL. Las muestras se prepararon mezclando 25 microlitros de solución de proteína con 25 microlitros de tampón de carga 2X y 4 microlitros de DTT 1 M seguido por calentamiento a 98°C durante 6 minutos. Se cargaron 10 microlitros de cada muestra en el gel y se corrió a una diferencia de potencial constante de 120 V. La carrera de la electroforesis se paró cuando el colorante marcador alcanzó la parte inferior del gel. Los geles se lavaron con agua dos veces durante un minuto cada vez y después se metieron en el microondas durante un minuto a potencia alta. Los geles se tiñeron con Simply Blue Safestain en el microondas durante un minuto a alta potencia seguido por incubación durante quince minutos a temperatura ambiente con agitación moderada. El gel se destiñó con agua para clarificar el fondo.
- 40
- 45 Se usó un ensayo enzimático basado en electroforesis capilar (CE) para medir la actividad de la CSTII expresada. El sustrato aceptor fue sialilactosa marcada con FCHASE [éster succinimidílico del ácido (6-fluorescein-5-carboxamido) hexanoico]. El sustrato donante fue CMP-NeuAc. El ensayo enzimático se conocía. Véase, por ejemplo, Wakarchuk, W. W. y Cunningham, A. M., *Methods Mol. Biol.*, 213:263-274 (2003)). Después de parar la reacción, la mezcla se diluyó y se aplicó directamente a electroforesis capilar controlado por un detector LIF (fluorescencia inducida por laser). El producto se eluyó medio minuto después que el aceptor y el resultado se calcula y expresa en unidades/litro. El ensayo control incluyó la mezcla de tampones reactivos sin enzima.
- 50

Conclusiones

- 55 Se expresó una proteína CSTII con etiqueta (MGS) del plásmido de expresión pNT1. No se expresó una proteína CSTII sin etiqueta del plásmido de expresión pNT2. Usando el vector de expresión pNT1 se produjo CSTII de *Campylobacter jejuni* en las cepas de *E. coli* JM109 y BNN93. La producción de CSTII se determinó tanto por análisis de expresión de proteína como actividad enzimática.
- 60 Se obtuvieron mayores niveles de actividad CSTII usando medio rico comparado con medio sencillo. Los valores de la actividad enzimática variaron desde 200-500 U/litro de medio de fermentación E6.0, mientras que el medio de producción E1.0 dio 100 U/litro de medio de fermentación.
- 65 Se obtuvieron mayores niveles de actividad CSTII en la cepa huésped de *E. coli* BNN93 comparada con la cepa JM109. Los valores de la actividad enzimática variaron desde 200-300 U/litro de medio de fermentación para JM109 en medio E6.0, comparado con 200-500 U/litro de medio de fermentación para BNN93 en medio E6.0.

La fermentación de *E. coli* que expresa CSTII con la prioridad de la cascada de agitación a flujo de aire dio un título mayor de actividad enzimática (500 U/l comparada con 360 U/l) y fue menos problemática en términos de formación de espuma que el modo cascada de flujo de aire a agitación para regular el oxígeno disuelto.

Ejemplo 3: Purificación de proteína CSTII expresada de pNT

Las células de *E. coli* que expresan el plásmido pNT1 se cultivaron como anteriormente y después se recogieron por centrifugación y se almacenaron a -80°C. Los precipitados celulares se descongelaron y se resuspendieron en Tris 50 mM, EDTA 5 mM, pH 8,3 (10 ml/g). Los precipitados celulares se homogenizaron con dos pases a través de un microfluidizador. El lisado se centrifugó a 15.000 x G durante veinte minutos y el sobrenadante se pasó a través de un filtro de 0,2 µm. El lisado se usó para los métodos de purificación descritos a continuación. Se proporciona un protocolo de purificación ejemplar en la figura 4.

Métodos

Determinación de la concentración de proteína (A280). La concentración de proteína se determinó usando un espectrofotómetro a una absorbancia fijada de 280 nm (A280). Antes de la medida a A280, el espectrofotómetro se ajustó a cero (autocero) usando el tampón apropiado. La concentración de proteína se determinó usando un coeficiente de extinción molar de 1 UA = 1,012 mg de CST-II.

Análisis por SDS-Page. Se usaron geles de tris glicina del 4-20% de Invitrogen. Todas las muestras se prepararon en condiciones reductoras (mezcladas 1:1 con tampón de muestra con SDS que contenía DTT 1 M), (geles no reductores: la muestra se mezcló 1:1 con tampón de muestra con SDS sin DTT). Todas las muestras se calentaron a 85°C durante 5 minutos a menos que se especifique de otra manera. Los geles se corrieron a una diferencia de potencial constante de 200 V durante 60 minutos. Después de la electroforesis, las proteínas se tiñeron con una solución Simply Blue Safestain durante 60 minutos. Los geles teñidos con plata se prepararon usando el kit de tinción con plata de Wako según el protocolo del fabricante. Se usó el marcador de peso molecular See Blue Plus 2 como patrón en cada gel, que contenía: miosina (250 kDa), fosforilasa (148 kDa), BSA (98 kDa), glutámico deshidrogenasa (64 kDa), alcohol deshidrogenasa (50 kDa), anhidrasa carbónica (36 kDa), lisozima (22 kDa), aprotinina (16 kDa) y cadena B de insulina (4 kDa). Las bandas de proteína se visualizaron y escanearon con un Scanjet 7400C de HP y la imagen del gel se optimizó con el programa Precision Scan de HP.

Pureza de proteína por HPLC. Se determinó la pureza de proteína por cromatografía de RP HPLC C3 (Zorbax 300SB-C3, 150 x 2,1 mm, 5 µm) usando las siguientes dos soluciones; A, TFA al 0,1% en agua, y B, TFA al 0,09% en ACN. La fase móvil se corrió como un gradiente desde el 40-60% de B durante 20 minutos a 0,5 ml/min a 40°C. La columna se reequilibró corriendo el gradiente desde el 95-40% de B durante 1 minuto, y después isocrático durante 10 minutos.

Ensayo de actividad enzimática de CST-II. Típicamente se diluyeron muestras de homogenizados de CST-II clarificados (10 µl) con HEPES 50 mM, Tween-20 al 0,02%, pH 7,5 (tampón de dilución) para el ensayo de actividad (~ 1/120 a 1/140: 10 µl de homogenizado + 1190 µl de tampón de dilución o 10 µl + 1390 µl de tampón de dilución). Se diluyeron muestras de CST-II purificada (10 µl) hasta 0,08-0,012 mg de proteína/ml con tampón de dilución. Se ensayaron las muestras de enzima CST-II diluida según el protocolo de ensayo de CST-II previamente descrito (véase, Neose Research Report, Capillary Electrophoresis Based Sialyltransferase Assay to Monitor CST-II Enzyme Activity, 17 de febrero, 2006).

Rotura de células mediante microfluidización. Se almacenaron los precipitados celulares de cosechas de fermentación de BNN93 CST-II a -80°C. En experimentos iniciales, el precipitado celular se resuspendió en Tris 20 mM, pH 8,3 (1 g/7,5 ml de tampón). Todos los experimentos posteriores utilizaron EDTA 5 mM o 10 mM en Tris 50 mM, pH 8,3. Tras resuspender, el precipitado se homogenizó con dos pases a través de un microfluidizador ajustado a 16.000-18.000 PSI. El sistema se enjuagó con 2,5 ml adicionales de tampón por gramo de precipitado para recoger cualquier material que permaneciera en el sistema después de terminar la homogenización del precipitado celular. La concentración final del homogenizado fue 1 g de precipitado celular/10 ml de tampón.

Clarificación. El homogenizado de células BNN93 se centrifugó inmediatamente a 12000 x G durante 60 minutos en una centrífuga equipada con un rotor de ángulo fijo (Sorval) a 4°C. El sobrenadante se filtró usando un filtro AcroPak-20 (0,2 micrómetros). El homogenizado filtrado se almacenó solo brevemente a 4°C hasta la purificación.

Secuenciación N-terminal de aminoácidos. Se analizaron muestras de CST-II purificada por secuenciación N-terminal para los primeros 10 (o más) aminoácidos. La CST-II purificada se mezcló 1:1 con tampón de muestra Tris-glicina con SDS con o sin algún agente reductor (DTT) y se desnaturizó por calor a 85°C durante 6 minutos. La muestra desnaturizada se cargó en múltiples carriles de geles de Tris-glicina con SDS del 4-20%. Los geles se corrieron en condiciones desnaturizantes reductoras o no reductoras a 125 V durante 118 minutos y se dejaron sin teñir para electrotransferencia a una membrana de PVDF. Los geles de SDS-PAGE se electrotransferieron usando un Mini Trans-Blot Cell de Biorad. El tampón de transferencia fue CAPS 10 mM (pH = 11,0) con metanol al 10%. Las

proteínas se transfirieron a una membrana de PVDF ProBlott de Applied Biosystems con tamaño de poro de 0,1-0,2 µm. La transferencia se corrió durante 158 minutos a 110 voltios para cada gel. Las bandas transferidas se visualizaron con tinción de proteínas azul de Coomassie. Las bandas se cortaron de las membranas de PVDF para análisis de secuencia en un secuenciador de aminoácidos Procise 492HT de Applied Biosystems. Las muestras se analizaron con un método de PVDF de pulso líquido. Los datos se analizaron con el software SequencPro versión 2.1 suministrado por Applied Biosystems.

La tabla 12 proporciona los resultados de la secuenciación de aminoácidos de la proteína CST-II purificada. La figura demuestra que la mayor parte de la proteína CST-II empezaba GSMK..., empezando con el residuo de glicina de la etiqueta MGS. Una pequeña cantidad de la proteína CST-II purificada empezaba con la secuencia nativa MKK...

Análisis de múltiples picos de HPLC de CST-II. Aislamiento de fracciones de CST-II de RP-HPLC. Las fracciones de CST-II se aislaron por RPHPLC usando las condiciones descritas en el presente documento. Las fracciones se recogieron de cuatro inyecciones (total ~ 100 µg de CST II). Cada fracción se concentró por SpeedVac, reduciendo el volumen a la mitad y eliminando la mayor parte del acetonitrilo. Una pequeña parte (~ el 5%) del volumen restante se analizó en RP-HPLC para comparar el perfil de la fracción purificada al del material de partida. Las muestras restantes se liofilizaron a sequedad. Puesto que la fracción 3 no estaba aparentemente enriquecida de forma adecuada después del primer aislamiento, se volvió a purificar disolviendo el material liofilizado en agua, inyectándolo todo en RP-HPLC y recogiendo 3 fracciones de nuevo. Una pequeña parte de la fracción 3 repurificada se diluyó con agua y se analizó por RP-HPLC. La fracción 3 repurificada restante se liofilizó a sequedad. Las fracciones liofilizadas se analizaron después mediante EM QToF (como se describe a continuación) y SDS-PAGE.

Análisis de EM QToF. Se inyectó la muestra proteica de CST-II para análisis de RP-HPLC/EM q-TOF sin preparación anterior de la muestra. Se usó un sistema microcapilar LC-Packings (Sunnyvale, CA) para la inyección y análisis de cromatografía. Un autoinyector de microinyección Famos™ inyectó 1,3 microlitros de la muestra. Se usaron los solventes A (ácido fórmico al 0,1% (v/v) en H₂O) y B (ácido fórmico al 0,1% (v/v) en acetonitrilo) como eluyentes. Una bomba capilar Ultimate™ se ajustó a una velocidad de flujo de 4 microlitros/minutos y bombeó un gradiente que empezaba desde el 25% de B hasta el 95% de B en 43 minutos. Se conectó un espectrómetro de masas QStar-XL (Applied Biosystems, Foster City, CA) en serie con el sistema LC-Packings. Se operó en modo estándar de ión positivo con una fuente de ionización de electrospray (ESI). El analizador de tiempo de vuelo (TOF) adquirió un intervalo de m/z de 400 a 3000. Los datos sin procesar para la proteína intacta se deconvolucionaron con software Analyst QSTM sobre un intervalo de masa de 25.000 a 35.000 Da. Determinación de endotoxina. Se determinó la contaminación con endotoxina usando el ensayo de lisado de amebocitos de Limulus (LAL) (Bio Whittaker, kit Kinetic-QCL, N°. de catálogo: 50-650U) (Neose QC160). Brevemente, la reacción de LAL es un ensayo enzimático, y la endotoxina cataliza la activación de una proenzima en el LAL. La velocidad inicial de activación se determina por la concentración de endotoxina presente. La enzima activada cataliza la liberación de pntroanilina (pNA) a partir del sustrato incoloro Ac-Ile-Ala-Arg-pNA. La liberación de pNA se mide fotométricamente a 405 nm. La concentración de endotoxina en una muestra se calcula a partir de su tiempo de reacción por comparación con el tiempo de reacción de soluciones que contienen cantidades conocidas de patrón de endotoxina.

Cromatografía de intercambio aniónico

Cromatografía en Q-Sepharosa – Gradiente en pasos. Se empaquetó una columna de cromatografía XK26 con resina Q-Sepharosa Fast Flow (Q-FF) (40 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de Tris 20 mM, pH 8,3 (tampón A). Se cargó un homogenizado de *E. coli* clarificado que contenía CST-II (un precipitado celular de 13 gramos preparado como se describe anteriormente sin la adición de EDTA, 130 ml) en la columna preparada a una velocidad de flujo de 115 cm/h. El material no unido se lavó de la columna usando 5 VC de tampón A. La CST-II se eluyó a una velocidad de flujo de 115 cm/h con el siguiente gradiente en pasos usando tampón B (NaCl 1 M en tampón A): 4 VC de tampón B al 15%, 4 VC de tampón B al 25%, 4 VC de tampón B al 50% (figura 5). Se tomaron muestras de las fracciones para análisis por SDS-PAGE y ensayos de actividad enzimática de CST-II (figuras 5 y 6). Se encontró la actividad CST-II purificada en el paso de elución de B al 15%. Las fracciones que contenían CST-II se juntaron, se formularon con glicerol al 50% y se almacenaron a -20°C para conservar la actividad.

Cromatografía en Q-Sepharosa – Gradiente en pasos (Sin dilución de la solución de carga). Se empaquetó una columna de cromatografía XK16 con resina Q-Sepharosa Fast Flow (Q-FF) (20 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de Tris 50 mM, pH 8,3 (tampón A). La conductividad de homogenizado de *E. coli* clarificado que contenía CST-II (un precipitado celular de 2 gramos como se prepara anteriormente con EDTA 5 mM, 20 ml) fue 4,66 ms/cm y la muestra se cargó en la columna preparada a una velocidad de flujo de 115 cm/h. El material no unido se lavó de la columna usando 5 VC de tampón A. La CST-II se eluyó a una velocidad de flujo de 115 cm/h con el siguiente gradiente en pasos usando tampón B (NaCl 1 M en tampón A): 5 VC de tampón B al 15%, 5 VC de tampón B al 100% (figura 7A). Las fracciones se analizaron por SDS-PAGE y ensayo de actividad enzimática de CST-II (figuras 7A y 8, tabla 1). Se encontró la actividad CST-II purificada en el paso de elución de B al 15%. Las fracciones que contenían CST-II se juntaron, se formularon con glicerol al 50% y se almacenaron a -20°C para conservar la actividad.

Tabla 1. Recuperación de la actividad enzimática CST-II de Q-Sepharosa FF

| Elución de Q-Sepharosa | Carga (unidades) | FT/Lavado (unidades) | Elución (unidades) | Elución (mg de proteína) | Actividad específica (U/mg de proteína) | Recuperación (%) |
|--|------------------|----------------------|--------------------|--------------------------|---|------------------|
| Gradiente por pasos (Solución de carga sin diluir) | 40,8 | 12,7 | 19,5 | 154,5 | 0,126 | 47,8 |
| Gradiente por pasos (Solución de carga diluida) | 40,8 | 4,6 | 27,6 | 194 | 0,142 | 67,6 |
| Gradiente lineal (Solución de carga diluida) | 40,8 | 4,6 | 13,1 | 93,4 | 0,149 | 32,1 |

Recuperación; unidades totales de CST-II recuperadas de la columna divididas por las unidades totales en el homogenizado clarificado cargado x 100. FT = Flujo que pasa.

5 Cromatografía en Q-Sepharosa – Gradiente en pasos (Dilución de la solución de carga). Se empaquetó una columna de cromatografía XK16 con resina Q-Sepharosa Fast Flow (Q-FF) (20 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de Tris 50 mM, pH 8,3 (tampón A). La conductividad de homogenizado de *E. coli* clarificado que contenía CST-II (un precipitado celular de 2 gramos como se prepara anteriormente con EDTA 5 mM, 20 ml) fue 4,66 ms/cm. El homogenizado clarificado se diluyó 1:1 con tampón A (20 ml), la conductividad era 3,3 ms/cm y la muestra se cargó en la columna preparada a una velocidad de flujo de 115 cm/h. El material no unido se lavó de la columna usando 5 VC de tampón A. La CST-II se eluyó a una velocidad de flujo de 115 cm/h con el siguiente gradiente en pasos con tampón B (NaCl 1 M en tampón A): 5 VC de tampón B al 15%, 5 VC de tampón B al 100% (figura 7B). Las fracciones se analizaron por SDS-PAGE y ensayo de actividad enzimática de CST-II (figuras 7B y 8, tabla 1). Se encontró la actividad CST-II purificada en el paso de elución de B al 15%. Las fracciones que contenían CST-II se juntaron, se formularon con glicerol al 50% y se almacenaron a -20°C para conservar la actividad.

20 Cromatografía en Q-Sepharosa – Gradiente lineal. Se empaquetó una columna de cromatografía XK16 con resina Q-Sepharosa Fast Flow (Q-FF) (20 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de Tris 50 mM, pH 8,3 (tampón A). La conductividad de homogenizado de *E. coli* clarificado que contenía CST-II (un precipitado celular de 2 gramos como se prepara anteriormente con EDTA 5 mM, 20 ml) fue 4,66 ms/cm. El homogenizado clarificado se diluyó 1:1 con tampón A (20 ml), la conductividad era 3,3 ms/cm y la muestra se cargó en la columna preparada a una velocidad de flujo de 115 cm/h. El material no unido se lavó de la columna usando 5 VC de tampón A. La CST-II se eluyó a una velocidad de flujo de 115 cm/h con el siguiente gradiente lineal con tampón B (NaCl 1 M en tampón A): tampón B del 0-15% durante 10 VC seguido por 5 VC de tampón B al 100% (figura 9). Las fracciones se analizaron por SDS-PAGE y ensayo de actividad enzimática de CST-II. Las fracciones que contenían CST-II se juntaron, se formularon con glicerol al 50% y se almacenaron a -20°C para conservar la actividad.

30 Resultados de Q-Sepharosa. Se evaluó Q-Sepharosa FF como un paso de captura inicial de CST-II. CST-II se unió eficazmente a Q-Sepharosa y se pudo eluir usando un gradiente en pasos (NaCl). La enzima purificada carecía de muchas de las proteínas de *E. coli* presentes en el homogenizado crudo. Cuando se usaron concentraciones bajas de EDTA para reducir la proteólisis, se requirió la dilución de la solución de carga para reducir la conductividad y para maximizar la capacidad de unión de la resina Q-Sepharosa. También se pudo usar una elución con un gradiente lineal con NaCl para purificar la CST-II en Q-Sepharosa. La cromatografía que usa resina 15Q de fuente también fue eficaz (cromatograma no mostrado).

40 **Cromatografía de intercambio catiónico**

45 Cromatografía en SP-Sepharosa de alta eficacia (HP) – pH 6,8. Se empaquetó una columna de cromatografía XK16 con resina SP-Sepharosa HP (20 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de fosfato de sodio 25 mM, pH 6,8 (tampón A).

50 Se diluyó una muestra de CST-II parcialmente purificada por Q-Sepharosa y almacenada a -20°C (24 ml, A280: 0,434 UA, glicerol al 50%) con 24 ml de fosfato de sodio 100 mM, pH 6,8 para ajustar el pH y después se diluyó además con fosfato de sodio 2,5 mM, pH 6,8 (82 ml) para reducir la conductividad a 3,8 ms/cm (A280: 0,855 UA). La muestra condicionada se filtró usando un filtro Nalgene de 250 ml (0,2 micrómetros) y se cargó en la columna preparada a una velocidad de flujo de 150 cm/h y se lavó con tampón A hasta que la absorbancia a 280 volvió a valores basales. La CST-II se eluyó a 80 cm/h con un gradiente con tampón B (NaCl 1 M en tampón A): tampón B del 0-100% durante 10 VC (figura 10). Las fracciones se analizaron por SDS-PAGE (figura 11). Las fracciones que contenían CST-II se juntaron, se formularon con glicerol al 50% y se almacenaron a -20°C para conservar la actividad.

55

Cromatografía en SP-Sepharosa HP – pH 6,0. Se empaquetó una columna de cromatografía XK16 con resina SP-Sepharosa HP (20 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de MES 25 mM, pH 6,0 (tampón A). Se diluyó una muestra de CST-II parcialmente purificada por Q-Sepharosa y almacenada a -20°C (20 ml, A280: 0,434 UA, glicerol al 50%) con 20 ml de MES 100 mM, pH 6,0 para ajustar el pH y después se diluyó además con MES 25 mM, pH 6,0 (40 ml) para reducir la conductividad a 2,4 ms/cm (A280: 1,20 UA). La muestra condicionada se filtró a través de una unidad de filtro Nalgene de 250 ml (0,2 micrómetros) y se cargó en la columna preparada a una velocidad de flujo de 150 cm/h y se lavó con tampón A hasta que la absorbancia a 280 volvió a valores basales. La CST-II se eluyó a 80 cm/h con un gradiente con tampón B (NaCl 1 M en tampón A): tampón B del 0-100% durante 10 VC (figura 12). Las fracciones se analizaron por SDS-PAGE (figura 13). Las fracciones que contenían CST-II se juntaron, se formularon con glicerol al 50% y se almacenaron a -20°C para conservar la actividad.

Resultados de la cromatografía de intercambio catiónico. Se evaluó SP-Sepharosa HP como un segundo paso de cromatografía para seguir la purificación con Q-Sepharosa FF. Este paso se introdujo para eliminar impurezas de *E. coli* (gliceraldehído-3-fosfato deshidrogenasa y proteínas osmóticamente inducibles: proteína de resistencia, resultados de secuenciación N-terminal no mostrados) (figuras 12 y 13). Se observó muy poca unión de la CST-II a las resinas SP usando condiciones de cromatografía de pH 6,8 incluso aunque este pH estaba por debajo del pl teórico de CST-II (7,5). CST-II se unió eficazmente a la resina SP a pH 6,0 en tampón MES 25 mM y se pudo eluir usando un gradiente de sal.

Paso de purificación en hidroxiapatita

Cromatografía de hidroxiapatita de tipo I (40 micrómetros) (1 ml). Se empaquetó una columna Tricorn 5 con resina de hidroxiapatita de tipo I (40 micrómetros) (1 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 VC de fosfato de sodio 5 mM, pH 6,5 (tampón A). Se diluyó una muestra de CST-II que se había purificado parcialmente por Q-Sepharosa y SP-Sepharosa HP y almacenado a -20°C (1 ml, glicerol al 50%, A280: 1,22 UA) con fosfato de sodio 5 mM, pH 6,5 (3 ml). La muestra se cargó en la columna preparada a 153 cm/h (0,5 ml/min). El material no unido se lavó de la columna con 10 VC de tampón A. La CST-II se eluyó con un gradiente de tampón B (NaCl 1,5 M en tampón A): tampón B del 0-100% durante 20 VC (figura 14A). Se tomaron muestras de las fracciones para SDS-PAGE (figura 15). Las fracciones que contenían CST-II se juntaron, se formularon con glicerol al 50% y se almacenaron a -20°C para conservar la actividad.

Cromatografía de hidroxiapatita de tipo II (40 micrómetros) (1 ml). Se empaquetó una columna Tricorn 5 con resina de hidroxiapatita de tipo II (40 micrómetros) (1 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 VC de fosfato de sodio 5 mM, pH 6,5 (tampón A). Se diluyó una muestra de CST-II que se había purificado parcialmente por Q-Sepharosa y SP-Sepharosa HP y almacenado a -20°C (1 ml, glicerol al 50%, A280: 1,22 UA) con fosfato de sodio 5 mM, pH 6,5 (3 ml). La muestra se cargó en la columna preparada a 153 cm/h (0,5 ml/min). El material no unido se lavó de la columna con 10 VC de tampón A. La CST-II se eluyó con un gradiente de tampón B (NaCl 1,5 M en tampón A): B del 0-100% durante 20 VC (figura 14B). Las fracciones se analizaron por SDS-PAGE (figura 15). Las fracciones que contenían CST-II se juntaron, se formularon con glicerol al 50% y se almacenaron a -20°C para conservar la actividad.

Cromatografía de hidroxiapatita de tipo II (40 micrómetros) (10 ml). Se empaquetó una columna XK16 con resina de hidroxiapatita de tipo II (40 micrómetros) (10 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 VC de fosfato de sodio 5 mM, pH 6,5 (tampón A). Se diluyó una muestra de CST-II que se había purificado parcialmente por Q-Sepharosa y SP-Sepharosa HP y almacenado a -20°C (30 ml, glicerol al 50%, A280: 5,56 UA) con fosfato de sodio 5 mM, pH 6,5 (120 ml). El pH se ajustó a 6,0 con HCl 1 M y después se filtró a través de una unidad de filtro Nalgene (0,2 micrómetros). La muestra (150 ml, conductividad; 1,65 ms/cm, A280: 1,043) se cargó en la columna preparada a 119 cm/h (4 ml/min). El material no unido se lavó de la columna con 5 VC de tampón A hasta que la absorbancia alcanzó valores basales. La CST-II se eluyó con un gradiente de tampón B (NaCl 1,5 M en tampón A): B del 0-100% durante 20 VC, seguido por 5 VC de fosfato de sodio 500 mM, pH 6,5 (figura 16). Las fracciones se analizaron por SDS-PAGE (figura 17). Las fracciones que contenían CST-II se juntaron, se formularon con glicerol al 50% y se almacenaron a -20°C para conservar la actividad.

Resultados de cromatografía de hidroxiapatita (HA). La cromatografía en hidroxiapatita, de tipo I o tipo II, se usó para purificar CST-II. La resina de tipo II ofrecía una resolución ligeramente mejor y eliminó algunas de las otras impurezas principales de *E. coli* que habían coeluido en Q-Sepharosa (transportador de oligopéptidos: proteína de unión periplásmica, resultados de secuenciación N-terminal no mostrados).

Cromatografía de interacción hidrofóbica (HIC)

Condicionamiento de alimentación de CST-II. Se usó para estos experimentos CST-II, parcialmente purificada por Q-Sepharosa, formulada con glicerol al 50% y almacenada a -20°C. Se diluyó CST-II purificada por Q-Sepharosa (1 ml)

en NaCl 2 M, Tris-HCl 25 mM, pH 7,2 (tampón B, 14 ml) y se concentró hasta 0,1 ml en un filtro de centrifugación Centricon Plus-20 (5000 Da de MWCO). La CST-II se diluyó de nuevo con tampón B (10 ml) y se concentró hasta 0,2 ml. La CST-II concentrada se diluyó con tampón B (0,8 ml) hasta un volumen final de 1,0 ml (A280: 4,23 mg).

5 Cromatografía en fenil-Sepharosa (alta sus.). Se conectó una columna ya empaquetada de fenil-Sepharosa de alta sustitución (1 ml, GE Healthcare) a un sistema de HPLC de Varian que controlaba la absorbancia a 280 nm. La columna se lavó con 30 ml de Tris-HCl 25 mM, pH 7,2 (tampón A) y se equilibró con 50 ml de NaCl 2 M, Tris-HCl 25 mM, pH 7,2 (tampón B) a una velocidad de flujo de 0,5 ml/min. La CST-II purificada por Q-Sepharosa condicionada (1,0 ml como se ha descrito anteriormente) se cargó en la columna preparada. La columna se lavó durante 10 minutos con tampón B (5 VC) a 0,5 ml/min para eliminar el material no unido. La proteína unida se eluyó usando el siguiente gradiente: 10-40 minutos, B del 100-0%, 40-60 minutos, tampón A al 100%. Se recogieron cinco fracciones y se analizaron por SDS-PAGE (figuras 18-19). Las fracciones se almacenaron a 4°C.

15 Cromatografía en fenil-Sepharosa (baja sus.). Se conectó una columna ya empaquetada de fenil-Sepharosa de baja sustitución (1 ml, GE Healthcare) a un sistema de HPLC de Varian que controlaba la absorbancia a 280 nm. La columna se lavó con 30 ml de Tris-HCl 25 mM, pH 7,2 (tampón A) y se equilibró con 50 ml de NaCl 2 M, Tris-HCl 25 mM, pH 7,2 (tampón B) a una velocidad de flujo de 0,5 ml/min. La CST-II purificada por Q-Sepharosa condicionada (1,0 ml como se ha descrito anteriormente) se cargó en la columna preparada. La columna se lavó durante 10 minutos con tampón B (5 VC) a una velocidad de flujo de 0,5 ml/min para eliminar el material no unido. La proteína unida se eluyó usando el siguiente gradiente: 10-25 minutos, tampón B del 100-0%, 25-50 minutos, tampón A al 100%. Se recogieron cinco fracciones y se analizaron por SDS-PAGE (figuras 20-21). Las fracciones se almacenaron a 4°C.

25 Cromatografía en fenil-Sepharosa (baja sus.) con etilenglicol al 20%. Se conectó una columna ya empaquetada de fenil-Sepharosa de baja sustitución (1 ml, GE Healthcare) a un sistema de HPLC de Varian que controlaba la absorbancia a 280 nm. La columna se lavó con Tris-HCl 25 mM, pH 7,2, etilenglicol al 20% (tampón A, 30 ml) y se equilibró con tampón B al 67% (NaCl 3 M, Tris-HCl 25 mM, pH 7,2) en tampón A (50 ml) a una velocidad de flujo de 1,0 ml/min. La CST-II purificada por Q-Sepharosa condicionada (1,0 ml como se ha descrito anteriormente) se cargó en la columna. La columna se lavó y eluyó usando el siguiente gradiente: 0-10 minutos, tampón B al 67% en Tris-HCl 25 mM con etilenglicol al 20%, pH 7,2 (tampón A); 10-25 minutos, tampón B del 67-0%, 25-50 minutos, tampón A al 100%. Se recogieron cuatro fracciones y se analizaron por SDS-PAGE (figuras 22-23). Las fracciones se almacenaron a 4°C.

35 Cromatografía en butil-Sepharosa con etilenglicol al 20%. Se conectó una columna ya empaquetada de butil-Sepharosa (1 ml, GE Healthcare) a un sistema de HPLC de Varian que controlaba la absorbancia a 280 nm. La columna se lavó con Tris-HCl 25 mM, pH 7,2, etilenglicol al 20% (tampón A, 30 ml) y se equilibró con tampón B al 50% (NaCl 3 M, Tris-HCl 25 mM, pH 7,2) en tampón A (50 ml) a una velocidad de flujo de 1,0 ml/min. La CST-II purificada por Q-Sepharosa condicionada (1 ml como se ha descrito anteriormente) se inyectó en la columna. La columna se lavó y el producto se eluyó usando el siguiente gradiente: 0-10 minutos, tampón B al 50% en Tris-HCl 25 mM con etilenglicol al 20%, pH 7,2 (tampón A); 10-20 minutos, tampón B del 50-0%, 20-30 minutos, tampón A. Se recogieron dos fracciones y se analizaron por SDS-PAGE (figuras 24-25). Las fracciones se almacenaron a 4°C.

45 Resultados de la cromatografía de interacción hidrofóbica (HIC). Se evaluó la cromatografía en fenil Sepharosa (HIC) como una alternativa a los pasos de purificación en hidroxipatita o SP-Sepharosa HP. Los experimentos iniciales con fenil-Sepharosa (alta sustitución) revelaron que la CST-II se unía muy fuertemente y no se eluía fácilmente. Cuando se usó fenil Sepharosa (baja sustitución) la recuperación solo era ligeramente superior, pero el material que eluía estaba muy purificado. La adición de etilenglicol al 20% al tampón de elución de HIC mejoraba mucho la recuperación de CST-II al tiempo que mantenía la resolución de purificación (figuras 20 y 21). La butil Sepharosa proporcionó incluso mejores recuperaciones y purificación de la CST-II cuando se usaba etilenglicol al 20% en el tampón de elución.

Otros métodos de purificación

55 Efectos de excipientes de tampones en la estabilidad de CST-II durante la filtración en acetato de celulosa. Se intercambió el tampón de la CST-II purificada por Q-Sepharosa usando filtros de centrifugación de acetato de celulosa en presencia de un panel de excipientes y se controló la evidencia de pérdida de proteína o agregación midiendo la absorbancia a 280 nm y el % de transmitancia a 600 nm. La CST-II purificada por Q-Sepharosa (5,2 ml, Tris 25 mM, NaCl 100 mM, glicerol al 50%, pH 8,3, A280: 0,909 UA) se diluyó con fosfato de sodio 50 mM, pH 6,5 (5,2 ml) y el pH se ajustó a 6,5 con tres gotas de fosfato de sodio monobásico 1 M. Se midieron la absorbancia a 280 nm y el % de transmitancia a 600 nm de la solución de CST-II con el pH ajustado usando un espectrofotómetro (GE Healthcare). La solución de CST-II se dividió en 10 alícuotas (1 ml). Cada alícuota de CST-II se diluyó con 3 ml de una de las siguientes diez soluciones de excipientes: 1. NaCl 0,01 M, fosfato de sodio 0,01 M, pH 6,5; 2. NaCl 0,1 M, fosfato de sodio 0,01 M, pH 6,5; 3. NaCl 0,2 M, fosfato de sodio 0,01 M, pH 6,5; 4. sacarosa 0,5 M, NaCl 0,01 M, fosfato de sodio 0,01 M, pH 6,5; 5. manitol 0,1 M, NaCl 0,01 M, fosfato de sodio 0,01 M, pH 6,5; 6. manitol 0,1 M, sacarosa 0,5 M, NaCl 0,01 M, fosfato de sodio 0,01 M, pH 6,5; 7. glicerol al 10% v/v, NaCl 0,01 M, fosfato de sodio

0,01 M, pH 6,5; 8. sorbitol 0,1 M, NaCl 0,01 M, fosfato de sodio 0,01 M, pH 6,5; 9. Tween-20 al 0,02%, NaCl 0,01 M, fosfato de sodio 0,01 M, pH 6,5; 10. trehalosa 0,1 M, NaCl 0,01 M, fosfato de sodio 0,01 M, pH 6,5.

Las soluciones de CST-II diluidas fueron concentradas cada una hasta 0,5 ml usando filtros de centrifugación Amicon Ultra 4 (10 kDa de MWCO) separados a 4°C. Las muestras concentradas (0,5 ml) se volvieron a diluir con el tampón correspondiente (0,5 ml) y concentradas de nuevo hasta 0,5 ml. Las muestras se volvieron a diluir (0,5 ml) y concentrar (0,5 ml) de nuevo dos veces más para un total de 4 pasos de concentración. La solución final de CST-II en los filtros se retiró, cada filtro se lavó con su tampón respectivo (0,5 ml), se combinaron y el volumen final de la solución de CST-II se ajustó a exactamente 1,0 ml con el tampón apropiado. Se midió inmediatamente en cada muestra el % de transmitancia (600 nm) y después se centrifugaron a 5000 x G durante 20 minutos (4°C) para sedimentar cualquier materia particulada. Se midió la absorbancia a 280 nm de cada solución usando el tampón correspondiente como blanco para cada lectura (% de T600 y A280) (tabla 2). Se midió la absorbancia (280 nm) de cada solución de excipiente contra agua como control. Cada solución de CST-II centrifugada se analizó por SDS-PAGE (30 µl) (figura 26). Efectos de los excipientes de tampones en la estabilidad de CST-II durante la filtración en poliétersulfona (Biomax). Se intercambió el tampón de la CST-II purificada por Q-Sepharosa usando filtros de centrifugación de poliétersulfona (PES, Biomax) en presencia de excipientes seleccionados (véase posteriormente) y se controló la evidencia de pérdida de proteína o agregación midiendo la absorbancia a 280 nm y el % de transmitancia a 600 nm. La CST-II purificada por Q-Sepharosa (4,5 ml, Tris 25 mM, NaCl 100 mM, glicerol al 50%, pH 8,3, A280: 0,909 UA) se diluyó con fosfato de sodio 50 mM, pH 6,5 (4,5 ml) y el pH se ajustó a 6,5 con tres gotas de fosfato de sodio monobásico 1 M. Se midieron la absorbancia a 280 nm y el % de transmitancia a 600 nm de la solución de CST-II con el pH ajustado. Se diluyeron alícuotas de solución de CST-II con el pH ajustado (2,0 ml) con 4,0 ml de una de las siguientes soluciones de excipientes (6,0 ml de volumen total) en filtros de centrifugación Centricon Plus-20 (Biomax 5, 5000 Da de MWCO): 1. fosfato de sodio 0,01 M, NaCl 0,01 M, pH 6,5; 2. Tween-20 al 0,02%, NaCl 0,1 M, fosfato de sodio 0,01 M, pH 6,5; 3. sacarosa 0,5 M, Tween-20 al 0,02%, NaCl 0,1 M, fosfato de sodio 0,01 M, pH 6,5. Las soluciones de CST-II diluidas se concentraron hasta aproximadamente 2 ml cada una a 4°C. Las muestras concentradas (2 ml) se volvieron a diluir con el tampón correspondiente (2 ml) y concentrar de nuevo hasta 2 ml. Las muestras se volvieron a diluir y concentrar de nuevo dos veces más para un total de 4 pasos de concentración. Durante el paso final de concentración, las muestras se concentraron hasta aproximadamente 0,5 ml y el retenido se retiró de la unidad. Las unidades de Centricon se enjuagaron con aproximadamente 1,5 ml del tampón apropiado y los enjuagues se juntaron con las soluciones concentradas para dar un volumen final medido de 2,0 ml. Se midió el % de transmitancia (600 nm) y la absorbancia (280 nm) de cada solución concentrada y control de tampón exactamente como se ha descrito anteriormente (tabla 3).

Tabla 2. Efectos de excipientes en la recuperación de CST-II usando filtros de centrifugación de acetato de celulosa

| Condición* | A280 del Excipiente | CST-II con tampón intercambiado | | |
|--------------------------------|---------------------|---------------------------------|-------|------------------|
| | | % de T600 | A280 | Recuperación (%) |
| Tampón | 0,000 | 96,6 | 0,143 | 33,5 |
| NaCl 0,1 M | -0,029 | 97,8 | 0,163 | 38,2 |
| NaCl 0,2 M | -0,021 | 99,4 | 0,149 | 34,9 |
| Sacarosa 0,5 M | 0,001 | 94,2 | 0,181 | 42,4 |
| Manitol 0,1 M | 0,004 | 96,0 | 0,144 | 33,7 |
| Sacarosa 0,5 M + Manitol 0,1 M | 0,012 | 95,6 | 0,186 | 43,6 |
| Glicerol al 10% | -0,030 | 96,4 | 0,169 | 39,6 |
| Sorbitol 0,1 M | 0,085 | 94,8 | 0,160 | 37,5 |
| Tween-20 al 0,02% | -0,022 | 98,0 | 0,221 | 51,8 |
| Trehalosa 0,1 M | 0,059 | 97,4 | 0,581 | 136 |

- o Se tomaron las lecturas de absorbancia (280 nm) y % de transmitancia (600 nm) después de la concentración/intercambio de tampón de CST-II (purificada por Q-Sepharosa) en tampón fosfato de sodio 10 mM, cloruro de sodio 10 mM, pH 6,5 que contenía el excipiente mostrado. Se usó la solución de excipiente correspondiente como blanco para cada lectura de A280 % de T600 en el espectrofotómetro.
- o A280 del excipiente: Absorbancia (280 nm) de la solución del excipiente frente a fosfato de sodio 10 mM, cloruro de sodio 10 mM, pH 6,5.

Recuperación (%) calculada como el valor de absorbancia (280 nm) dividido por la absorbancia (280 nm) de CST-II antes de la concentración/intercambio de tampón x 100.

Tabla 3. Efectos de excipientes en la recuperación de CST-II usando filtros de centrifugación de poliétersulfona (Biomax)

| Condición | % de T600 | A280 | Recuperación (%) |
|---|-----------|-------|------------------|
| Tampón | 97,2 | 0,139 | 32,3 |
| NaCl 0,1 M, Tween-20 al 0,02% | 98,2 | 0,18 | 41,8 |
| NaCl 0,1 M, Tween-20 al 0,02%, Sacarosa 0,5 M | 97,6 | 0,185 | 43 |

- o Se tomaron las lecturas de absorbancia (280 nm) y % de transmitancia (600 nm) después de la concentración/intercambio de tampón de CST-II (purificada por Q-Sepharosa) en fosfato de sodio 10 mM, cloruro de sodio 10 mM, pH 6,5 que contenía el excipiente mostrado. Se usó la solución de excipiente correspondiente como blanco para cada lectura de A280 % de T600 en el espectrofotómetro.
- 5 Recuperación (%) calculada como el valor de absorbancia (280 nm) dividido por la absorbancia (280 nm) de CST-II antes de la concentración/intercambio de tampón x 100.

10 *Efectos de diálisis en la estabilidad de CST-II.* Se intercambió el tampón de la CST-II purificada por Q-Sepharosa por diálisis y se controló la evidencia de pérdida de proteína midiendo la absorbancia a 280 nm. Se diluyeron dos alícuotas de CST-II purificada por Q-Sepharosa (1,0 ml, Tris 25 mM, NaCl 100 mM, glicerol al 50%, pH 8,3, A280: 0,909 UA) con fosfato de sodio 10 mM, NaCl 10 mM, pH 6,5 (1 ml) o Tris-HCl 10 mM, NaCl 10 mM, pH 8,5 (1,0 ml). Se midió la absorbancia (280 nm) de cada solución usando un espectrofotómetro (GE Healthcare). Cada solución de CST-II diluida se inyectó en un casete de diálisis Slide-A-Lyzer (3500 Da de MWCO, Pierce) según las instrucciones del fabricante. Los casetes se dializaron contra los tampones correspondientes durante 8 horas con agitación moderada a 4°C. Se retiraron las muestras de CST-II dializadas de los casetes de Slide-A-Lyzer y se midieron el volumen, absorbancia (280 nm) y % de transmitancia (600 nm). Se calcularon las concentraciones de proteína y la masa de proteína total. Se determinó la cantidad de CST-II recuperada de este proceso dividiendo la masa de proteína de la solución de CST-II dializada por la de la solución de CST-II diluida original (tabla 4).

20 **Tabla 4. Recuperación de actividad enzimática CST-II después de diálisis a pH 6,5 y 8,5**

| Condición de diálisis | Pre-diálisis | | Post-diálisis | | | Recuperación (%) |
|-----------------------|--------------|-------|---------------|-----------|-------|------------------|
| | Volumen (ml) | A280 | Volumen (ml) | % de T600 | A280 | |
| pH 6,5 | 2 | 0,438 | 3,7 | 98,2 | 0,157 | 66,3 |
| pH 8,5 | 2 | 0,456 | 3,8 | 97,2 | 0,162 | 67,5 |

- o La absorbancia a 280 y el % de transmitancia a 600 nm se midieron después de la concentración/intercambio de tampón de CST-II (purificada por Q-Sepharosa). Se usó como blanco el tampón de diálisis correspondiente para cada lectura de A280 y % de T600 en el espectrofotómetro.
- 25 o Recuperación (%) calculada como (valor de absorbancia a 280 nm multiplicado por el volumen después de la diálisis) dividido por (absorbancia (280 nm) multiplicada por la concentración antes de la diálisis) x 100.

Combinación de métodos

30 Se realizaron purificaciones de CST-II a pequeña escala para determinar qué combinación de métodos de purificación (cromatografías de Q-Sepharosa FF- elución en gradiente, hidroxiapatita de tipo II (40 micrómetros) y SP-Sepharosa HP) daba mejores resultados. Después del paso de cromatografía en Q-Sepharosa, se pudo conseguir la misma pureza de CST-II combinando hidroxiapatita y SP-Sepharosa en cualquier orden (datos no mostrados).

35 Se realizó después un proceso completo de purificación a pequeña escala de CST-II usando cromatografía en Q-Sepharosa FF – elución en gradiente, hidroxiapatita de tipo II (40 micrómetros) y SP-Sepharosa HP como se muestra en la figura 4, excepto que entre cada paso de cromatografía se intercambié el tampón CST-II y/o se concentró usando filtros de centrifugación según fuera necesario para preparar para la siguiente cromatografía.

40 Se ensayó la actividad y concentración de la proteína medida por A280 de la CST-II en cada fase (tablas 5 y 6). Este proceso produjo una CST-II de gran pureza, sin embargo, la actividad enzimática recuperada global y las recuperaciones de masa fueron extremadamente bajas como resultado de la baja recuperación de la elución en gradiente de la cromatografía de Q-Sepharosa combinada con pérdidas de proteína durante los pasos de intercambio de tampón y concentración (tabla 6). Además, se observó un precipitado blanco durante la carga de la columna de hidroxiapatita que produjo una contrapresión significativa. Se emprendió un breve estudio para determinar los métodos para prevenir la pérdida de CST-II resultante de la agregación o desnaturalización proteica durante estos pasos del proceso.

50 **Tabla 5. Resultados del desarrollo del proceso del proceso preliminar de purificación de CST-II**

| Paso del proceso | Recuperación en el paso (%) | Recuperación global (%) |
|-----------------------------------|-----------------------------|-------------------------|
| Recogida | 100% | 100% |
| Q-Sepharosa | 16,2 | 16,2 |
| Hidroxiapatita de tipo II | 5,8* | 0,93* |
| SP-Sepharosa HP/Formulación final | 56,2 | 2,3 |

- o Recuperación en el paso (%): calculado como unidades totales de CST-II en la mezcla de fracciones dividido por las unidades totales de CST-II en la fracción de elución del paso anterior x 100. Recuperación

global (%) calculada como la actividad total de CST-II (unidades) recuperada del paso de purificación dividida por las unidades totales al inicio del proceso (homogenizado celular) x 100.

- o * La alta concentración de cloruro de sodio en la mezcla de elución de hidroxapatita inhibe el ensayo de actividad de CST-II lo que da resultados inexactos. La concentración de enzima no era lo suficientemente alta (debido a la baja recuperación) para reducir con éxito la concentración de sal por dilución y obtener un valor de actividad de CST-II preciso.

Tabla 6. Recuperación de masa de CST-II después de los pasos de intercambio de tampón/concentración entre cromatografías del proceso preliminar

| Intercambio de tampón de la mezcla de elución | Recuperación en el paso (%) |
|---|-----------------------------|
| Recogida | na |
| Intercambio de tampón de la mezcla de elución de Q-Sepharosa | 45 |
| Intercambio de tampón de la mezcla de elución de Hidroxapatita de tipo II | 87 |
| Intercambio de tampón de la mezcla de elución de SP-Sepharosa HP | 72 |

- o Determinada la cantidad de proteína perdida durante el condicionamiento (intercambio de tampón/concentración) de CST-II en preparación para cargar la siguiente columna.
- o Recuperación en el paso: mg totales de proteína recuperados de un paso de intercambio de tampón/concentración dividido por las unidades totales de CST-II antes del intercambio de tampón/concentración x 100.

Se concentró la CST-II purificada por Q-Sepharosa y se intercambió el tampón usando dos tipos diferentes de membranas de filtro de centrifugación (acetato de celulosa y poliéter sulfona). Se añadieron excipientes a los tampones para determinar si estos reactivos eran capaces de prevenir o reducir la pérdida de proteína CST-II durante este proceso juzgado por medidas de absorbancia (280 nm) y SDS PAGE (tablas 2 y 3 y figura 26). La trehalosa (trehalosa 0,1 M) parecía prometedora aunque un potencial contaminante que absorbe (280 nm) hizo la interpretación difícil. Tween-20 era el excipiente más beneficioso, aunque aún se observó una pérdida significativa de CST-II (mediante absorbancia a 280 nm).

Se evaluó la diálisis como una alternativa al intercambio de tampón usando filtros de centrifugación. Se probaron dos tampones y para ambos las recuperaciones de CST-II (parcialmente purificada por Q-Sepharosa) fueron mayores del 66% (basadas en la absorbancia a 280 nm) (tabla 4). Estos resultados fueron muy superiores a las recuperaciones típicas para unidades de filtración por centrifugación (33%), por tanto, se eligió diálisis como un proceso sustituto para cualquier paso de intercambio de tampón.

Aunque no se obtuvieron la recuperación de la actividad enzimática para los experimentos de diálisis y filtración por centrifugación, el descenso observado en la cantidad total de proteína, reflejado por las lecturas de absorbancia a 280 nm, son consistentes con la pérdida de proteína y actividad enzimática CST-II durante la purificación a pequeña escala de la enzima (tabla 6). Se perdió proteína durante los pasos de concentración y filtración y se observó en las membranas de filtración como un precipitado floculante blanco (filtración de flujo tangencial, datos no mostrados). Esta pérdida disminuyó significativamente la recuperación total de CST-II (tabla 5).

Producción de CST-II (pequeña escala)

Rotura celular por microfluidización. Se homogenizó el precipitado celular de recogida de fermentación de BNN93 CST-II (10 g, almacenado a -80°C) en Tris 50 mM, EDTA 10 mM, pH 8,3 y se clarificó como se ha descrito previamente. El homogenizado se dividió en dos muestras de 50 ml. Se añadió Tween-20 (0,02%) a una de las muestras de homogenizado de 50 ml y la otra se dejó sin tratar.

Clarificación. Los homogenizados de células BNN93 (2 x 50 ml) se centrifugaron inmediatamente a 15000 x G durante 20 minutos en una centrífuga equipada con un rotor de ángulo fijo (Sorval) a 4°C. Los sobrenadantes se filtraron usando un filtro Nalgene de 75 mm (de una unidad de filtro de 1 litro, 0,2 micrómetros) y se recogieron en una botella Nalgene de 150 ml. Las muestras de homogenizado filtradas se mantuvieron a 4°C brevemente hasta su purificación de Q-Sepharosa FF.

Cromatografía en Q-Sepharosa usando un gradiente en pasos (sin Tween-20). Se empaquetó una columna de cromatografía XK26 con resina Q-Sepharosa FF (50 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de Tris 50 mM, pH 8,3 (tampón A). El homogenizado de *E. coli* clarificado que contenía CST-II (se preparó un precipitado celular de 5 gramos como se describe anteriormente que contenía EDTA 10 mM, 50 ml) se diluyó con tampón A (150 ml) y se midió la conductividad (3,12 ms/cm). La muestra se cargó en la columna a una velocidad de flujo de 113 cm/h (10 ml/min). El material no unido se lavó de la columna con 5 VC de tampón A. La CST-II se eluyó a una velocidad de flujo de 113 cm/h (10 ml/min) con el siguiente gradiente en pasos con tampón B (NaCl 1 M en tampón A): 4 VC de tampón B al 15%, 4 VC de tampón B al 100% (figura 27). Se tomaron muestras de las fracciones para SDS-PAGE (figura 28). Se encontró la actividad CST-II purificada en el paso de elución de tampón B al 15%. Las fracciones que

contenían CST-II se juntaron y se formuló una muestra de 0,5 ml con glicerol al 50% y se almacenó a -20°C para conservar la actividad.

5 Cromatografía en Q-Sepharosa usando un gradiente en pasos (con Tween-20). Se empaquetó una columna de cromatografía XK26 con resina Q-Sepharosa FF (50 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de Tris 50 mM, pH 8,3 que contenía Tween-20 al 0,005% (tampón A). El homogenizado de *E. coli* clarificado que contenía CST-II (se preparó un precipitado celular de 5 gramos como se describe anteriormente que contenía EDTA 10 mM, 50 ml) y Tween-20 (0,02%) se diluyó con tampón A (150 ml) y se midió la conductividad (3,12 ms/cm). La CST-II diluida se cargó en la columna preparada a una velocidad de flujo de 113 cm/h (10 ml/min). El material no unido se lavó de la columna con 5 VC de tampón A. La CST-II se eluyó a una velocidad de flujo de 113 cm/h (10 ml/min) con el siguiente gradiente en pasos con tampón B (NaCl 1 M en tampón A): 4 VC de tampón B al 15%, 4 VC de tampón B al 100% (figura 27). Se analizaron las fracciones por SDS-PAGE (figura 26). Se encontró la actividad CST-II purificada en el paso de elución de tampón B al 15%. Las fracciones que contenían CST-II se juntaron y se formuló una muestra de 0,5 ml con glicerol al 50% y se almacenó a -20°C para conservar la actividad. El resto de la mezcla de fracciones de CST-II se almacenó a 4°C antes de la purificación adicional por cromatografía de hidroxiapatita de tipo II.

20 Cromatografía de hidroxiapatita de tipo II (40 micrómetros) (con Tween-20). Se empaquetó una columna XK16 con resina de hidroxiapatita de tipo II (40 micrómetros) (10 ml, 6,0 g) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 VC de fosfato de sodio 5 mM, pH 6,5 que contenía Tween-20 al 0,005% (tampón A). Se diluyó lentamente una muestra de 5,0 ml de CST-II que se había purificado parcialmente por Q-Sepharosa que contenía Tween-20, véase anteriormente (figura 27, fracción 1) y almacenado a 4°C con fosfato de sodio 50 mM, pH 6,5 (5 ml) mientras se controlaba el pH. El pH se ajustó de 6,8 a 6,5 con 0,5 ml adicionales de tampón fosfato de sodio 50 mM, pH 6,5. La muestra (10,5 ml) se diluyó además con fosfato de sodio 10 mM, pH 6,5 que contenía Tween-20 al 0,005% (29,5 ml) para reducir la concentración de fosfato. La CST-II diluida se filtró a través de un filtro de 0,2 micrómetros, se midieron el pH y la conductividad (pH 6,53, 4,99 ms/cm, A280 = 1,022 UA) y la solución se cargó en la columna preparada a 120 cm/h (4 ml/min). El material no unido se lavó de la columna con 5 VC de tampón A. La CST-II se eluyó con un gradiente de tampón B (NaCl 1,5 M en tampón A): tampón B del 0-100% durante 25 VC, seguido por 5 VC de fosfato de sodio 500 mM, pH 6,5 (figura 29). Las fracciones se analizaron por SDS-PAGE (figura 30). Las fracciones que contenían CST-II se juntaron y se formuló una muestra de 0,5 ml con glicerol al 50% y se almacenó a -20°C para conservar la actividad. El resto de la mezcla de fracciones de CST-II se almacenó a 4°C antes de la purificación adicional por cromatografía de SP-Sepharosa HP.

35 Cromatografía de hidroxiapatita de tipo II (20 micrómetros) (con Tween-20). Se empaquetó una columna XK16 con resina de hidroxiapatita de tipo II (20 micrómetros) (10 ml, 6,0 g) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 VC de fosfato de sodio 5 mM, pH 6,5 que contenía Tween-20 al 0,005% (tampón A). Se diluyó lentamente una muestra de 5,0 ml de CST-II que se había purificado parcialmente por Q-Sepharosa que contenía Tween-20, véase anteriormente (figura 27, fracción 1) y almacenado a 4°C con fosfato de sodio 50 mM, pH 6,5 (5 ml) mientras se controlaba el pH. El pH se ajustó de 6,8 a 6,5 con 0,5 ml adicionales de tampón fosfato de sodio 50 mM, pH 6,5. La muestra (10,5 ml) se diluyó además con fosfato de sodio 10 mM, pH 6,5 que contenía Tween-20 al 0,005% (29,5 ml) para reducir la concentración de fosfato. La CST-II diluida se filtró a través de un filtro de 0,2 micrómetros, se midieron el pH y la conductividad (pH 6,53, 4,99 ms/cm, A280 = 1,022 UA) y la solución se cargó en la columna preparada a 120 cm/h (4 ml/min). El material no unido se lavó de la columna con 5 VC de tampón A. La CST-II se eluyó con un gradiente de tampón B (NaCl 1,5 M en tampón A): tampón B del 0-100% durante 25 VC, seguido por 5 VC de fosfato de sodio 500 mM, pH 6,5 (figura 29). Las fracciones se analizaron por SDS-PAGE (figura 30). Las fracciones que contenían CST-II se juntaron y se formuló una muestra de 0,5 ml con glicerol al 50% y se almacenó a -20°C para conservar la actividad. El resto de la mezcla de fracciones de CST-II se almacenó a 4°C antes de la purificación adicional por cromatografía de SP-Sepharosa HP.

55 Cromatografía en SP-Sepharosa HP – pH 6,0 (fracciones de HA de 40 micrómetros). Se empaquetó una columna de cromatografía XK16 con resina SP-Sepharosa HP (10 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de MES 25 mM, pH 6,0, que contenía Tween-20 al 0,005% (tampón A). Una muestra de CST-II parcialmente purificada por Q-Sepharosa e hidroxiapatita de tipo II (40 micrómetros), almacenada a 4°C, véase anteriormente (35 ml, Tween-20 al 0,005%) se ajustó a Tween-20 al 0,02% y se dializó durante la noche frente a MES 25 mM, NaCl 10 mM, pH 6,0 (4,0 L, conductividad; 1,91 ms/cm) en bolsa de diálisis Snakeskin a 4°C (véase anteriormente). La muestra dializada se filtró a través de un filtro Tube Top de 50 ml (0,2 micrómetros) y el volumen final fue 42,3 ml, pH 6,0, conductividad: 3,6 ms/cm, A280: 0,261.

65 La muestra filtrada se cargó en la columna preparada a una velocidad de flujo de 120 cm/h (4 ml/min) y se lavó con 5 VC de tampón A. La CST-II se eluyó a 120 cm/h (4 ml/min) con un gradiente con tampón B (NaCl 1 M en tampón A): tampón B del 0-50% durante 15 VC, mantenido a tampón B al 50% durante 1 VC y después cualquier material unido estrechamente se eluyó con 5 VC de tampón B al 100% (figura 31). Las fracciones se analizaron por SDS-

PAGE (figura 30). Las fracciones que contenían CST-II se juntaron y se formuló una muestra de 0,5 ml con glicerol al 50% y se almacenó a -20°C para conservar la actividad. El resto de la mezcla de fracciones de CST-II (25 ml) se formuló.

5 Cromatografía en SP-Sepharosa HP – pH 6,0 (fracciones de HA de 20 micrómetros). Se purificó la CST-II purificada por HA de tipo II (20 micrómetros) usando SP-Sepharosa HP de la misma manera que se ha descrito anteriormente (cromatografía en SP-Sepharosa HP – 20 micrómetros) (figura 31). Las fracciones que contenían CST-II se combinaron después del análisis por SDS-PAGE (figura 30). Las fracciones del producto juntas se formularon como se describe a continuación.

10 Formulación de CST-II. El resto de las mezclas de fracciones de CST-II (25 ml, cada una) de los pasos de cromatografía en SP-Sepharosa HP, véase anteriormente, se concentraron hasta 5 ml en un concentrador de centrifugación Amicon Ultra-15 (10 kDa de MWCO) y se diluyeron con tampón de formulación (MES 25 mM, NaCl 200 mM, pH 6,0, 10 ml) y se volvieron a concentrar hasta un volumen final de 5 ml, cada una. Las mezclas de CST-II con tampón intercambiado se filtraron a través de unidades Mustang E Aerodisc que se habían precondicionado con tampón de formulación. Después de la filtración los filtros se enjuagaron con 1,5 ml adicionales de tampón de formulación y los filtrados se juntaron y diluyeron con volúmenes iguales de glicerol (6,5 ml, cada una) para dar CST-II en una formulación final de MES 12,5 mM, NaCl 100 mM, Tween-20 al 0,0125%, pH 6,0. Las muestras se almacenaron inmediatamente a -20°C.

20 Resultados de la producción de CST-II (pequeña escala). Se desarrolló un esquema de purificación basado en los resultados de los pasos del proceso descritos anteriormente. Estos incluyen el uso de cromatografías en Q-Sepharosa FF, hidroxiapatita de tipo II y SP-Sepharosa HP para purificar CST-II (figura 4). Se realizó un proceso de purificación de prueba a pequeña escala para determinar la durabilidad del proceso y determinar las recuperaciones de producto. La CST-II se homogenizó en tampón Tris a pH 8,3 con EDTA 10 mM y se clarificó en presencia y ausencia de Tween-20. La CST-II fue capturada en Q-Sepharosa (con y sin Tween-20) y se eluyó con el gradiente en pasos de NaCl para maximizar la recuperación (figuras 27 y 28).

30 La mezcla de CST-II parcialmente purificada que contenía Tween-20 se diluyó, separó y purificó sobre una resina de hidroxiapatita de tipo II (40 micrómetros) o una de hidroxiapatita de tipo II (20 micrómetros) (figuras 29 y 30). Se encontró que la resina de 20 micrómetros proporcionaba una resolución ligeramente mejorada entre la CST-II y una impureza proteica de E. coli que eluye a 60 kDa (transportador de oligopéptido: proteína de unión periplásmica, resultado de secuenciación terminal no mostrado). Las mezclas de hidroxiapatita de 40 micrómetros y 20 micrómetros se dializaron individualmente y se purificaron sobre SP-Sepharosa HP (figuras 30 y 31). Se estimó que
35 ambas recuperaciones de CST-II eran aproximadamente del 22% basándose en la actividad enzimática.

Análisis de CST-II durante la purificación

40 Proteólisis de CST-II. Intento de separación de fragmentos proteolíticos por exclusión molecular (SEC). Se inyectó una muestra de CST-II purificada (Q-Sepharosa, hidroxiapatita de tipo II y SP-Sepharosa HP) y formulada como se ha descrito anteriormente (250 µl, 711 µg de proteína) en una columna de Superdex 200 10/30GL unida a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La muestra se cromatografió con fosfato de sodio 50 mM, NaCl 150 mM, pH 7,2 a una velocidad de flujo de 0,5 ml/min (figura 32). Las fracciones se analizaron por SDS-PAGE (figura 33).

45 Prueba para actividad proteolítica. Se recogió CST-II después de cada paso de cromatografía del proceso a pequeña escala descrito anteriormente (Q-Sepharosa, HA de tipo II (20 micrómetros) y SP-Sepharosa HP) y se almacenaron inmediatamente con glicerol al 50% en tampón de elución a -20°C en los siguientes tampones: 1. Q-Sepharosa, Tris-HCl 25 mM, NaCl 75 mM, glicerol al 50%, pH 8,3; 2. HA-tipo II (20 micrómetros), fosfato de sodio 5 mM, NaCl 550 mM, glicerol al 50%, pH 6,5; y 3. SP-Sepharosa HP, MES 12,5 mM, NaCl 85 mM, glicerol al 50%, pH 6,0. Cada muestra (10 µg de cada una) se incubó 14 horas a 32°C y se analizó por SDS-PAGE para determinar si se observaba proteólisis adicional de la CST-II relativa a las muestras almacenadas a -20°C (figura 34).

55 Prevención de la proteólisis de CST-II durante la homogenización del precipitado celular. Se resuspendieron muestras de precipitado celular (2,0 g cada una) de recogida de fermentación de BNN93 CST-II (almacenadas a -80°C) en uno de los siguientes tampones de prueba (20 ml): Condición 1. Tris-HCl 50 mM, pH 8,3; Condición 2. Tris-HCl 50 mM, EDTA 10 mM, pH 8,3; Condición 3. MES 50 mM, pH 6,0; Condición 4. acetato de sodio 50 mM, pH 4,0; Condición 5. Tris-HCl 50 mM, pH 8,3 con 500 µl del conjunto II de mezcla de inhibidores de proteasas de Calbiochem (que contiene AEBSF 20 mM, bestatina 1,7 mM, E-64 200 micromolar, EDTA 85 mM, pepstatina 2 mM).

60 Las células se microfluidizaron y clarificaron como se ha descrito previamente. Se filtraron muestras de los homogenizados clarificados a través de filtros de jeringuilla Pall (0,2 micrómetros) y se inyectaron en una columna de Superdex 200 10/30GL que se había equilibrado previamente con el tampón de homogenización correspondiente para cada muestra que contenía además NaCl 150 mM. Las muestras se cromatografiaron en un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm (figura 35). Se recogieron mezclas de picos de CST-II y se analizaron por SDS-PAGE (figura 36).

65

Resultados de la caracterización de CST-II. CST-II es una proteína tetramérica. La secuenciación N-terminal confirmó que se purificaron dos especies de proteína de la fermentación células BNN93: la especie de 260 aminoácidos esperada (MKK...NINF) y una especie de 262 aminoácidos (GSMKK...NINF) que se piensa surge de una iniciación de traducción alternativa del vector de expresión pNT1. Se encontró que la relación de la especie GSMKK (262 aminoácidos) respecto a la especie MKK (260 aminoácidos) en la proteína purificada era aproximadamente 90:10.

Análisis de la proteólisis de CST-II. Durante el desarrollo de un proceso de purificación, se observó proteólisis de CST-II. El análisis de estos fragmentos por secuenciación N-terminal y CL/EM Q-ToF confirmó la presencia de dos fragmentos proteolíticos de CSTII en la muestra purificada. La adición de EDTA al tampón de homogenización redujo ligeramente, pero no previno del todo, la proteólisis.

Los fragmentos de proteólisis de CST-II no eran separables por ninguno de los métodos de cromatografía utilizados. Usando métodos analíticos, se pudo observar el fragmento C-terminal proteolizado por HPLC de fase inversa. Se sabe que CST-II forma una especie oligomérica consistente con un tetrámero en solución. Por tanto, no fue sorprendente que la cromatografía de exclusión molecular no separara tampoco las impurezas proteolíticas. Se probó la actividad proteolítica residual de cada una de las mezclas de elución de CST-II (Q-Sepharosa, hidroxiapatita, SP-Sepharosa HP, almacenadas a -20°C) mediante incubación ampliada a 32°C. El análisis de SDS PAGE reveló que no se había producido proteólisis adicional de CST-II en ninguna de las mezclas de elución de cromatografía, lo que indica que la(s) proteasa(s) se había(n) eliminado durante el paso de cromatografía en Q-Sepharosa (figura 34).

Como resultado de los datos anteriores, se examinaron los pasos de homogenización y clarificación de *E. coli* como la fuente potencial de la proteólisis. Se idearon una serie de experimentos para reducir o eliminar la proteólisis durante la homogenización de los precipitados celulares de *E. coli* (almacenados a -80°C). Los precipitados celulares se homogenizaron por microfluidización usando un panel de tampones desde pH 4 hasta pH 8,3 en presencia y ausencia de EDTA y una mezcla de inhibidores de proteasas (Calbiochem). Las muestras homogenizadas y clarificadas se purificaron por SEC (figura 35) y las fracciones que contenían CST-II se analizaron por SDS PAGE (figura 36). Las variaciones en el pH no redujeron la proteólisis, sin embargo, se encontró que CST-II era inestable usando las condiciones de pH 4. Solo la presencia de EDTA produjo un descenso observable en la proteólisis. Por tanto, se incluyó EDTA en todos los pasos posteriores de rotura celular y clarificación. Los dos principales fragmentos de proteólisis parecían estar presentes independientemente del método utilizado para prevenir la proteólisis incluyendo el uso de una mezcla de proteasas. Estos resultados sugieren que el suceso de la proteólisis es difícil de prevenir y puede suceder durante la fermentación o recogida de la CST-II.

Análisis de múltiples picos de HPLC de CST-II. Mientras se exploraban métodos de purificación, se observó que CST-II purificada usando una combinación de Q-Sepharosa, hidroxiapatita y SP-Sepharosa HP contenía tres bandas que corrían muy próximas a pesos moleculares muy similares en un gel de SDS-PAGE usando condiciones no reductoras. Esto está en contraste con una única banda que se observó de la misma CST-II purificada en geles de SDS PAGE usando condiciones reductoras. Las secuencias de aminoácidos N-terminales (secuenciación de Edman) de las tres bandas del gel de SDS PAGE no reductor eran idénticas al extremo amino de CST-II. Además, se observaron tres picos para la CST-II purificada cuando se analizó por RP HPLC usando condiciones no reductoras. Cada uno de los picos de RP-HPLC se aisló y volvió a analizar por RP-HPLC, SDS PAGE y QToF/EM (EM: picos 1 y 2 solo). El análisis de EM confirmó (dentro de la variabilidad experimental) que las especies en las dos fracciones principales representaban la proteína CST-II intacta. Los datos de los análisis de RP-HPLC y SDS PAGE de cada pico aislado por RP-HPLC mostraron sorprendentemente que las tres especies se interconvertían tras el almacenamiento y manejo. El patrón de múltiples bandas observado para CST-II por SDS-PAGE (condiciones no reductoras) y los múltiples picos observados por RP-HPLC se pueden deber a la presencia de diferentes formas de especies con puentes disulfuro intramoleculares de CST-II. CST-II tiene 6 residuos de cisteína no apareados que se había mostrado previamente que estaban en un estado reducido por cristalografía de rayos X (C. P. Chiu, *et. al. Nature Structural and Molecular Biology*. 11, 163-170, Febrero 2004). Cuatro de estos residuos de cisteína tienen distancias intramoleculares que pueden permitir la formación de puentes disulfuro intramoleculares tras el manejo en un ambiente aerobio.

Purificación de CST-II (gran escala)

Rotura celular por microfluidización. Se almacenaron precipitados celulares de recogidas de fermentación de BNN93 CST-II a -80°C. El precipitado celular (150 g) se resuspendió en Tris 50 mM, EDTA 10 mM, pH 8,3 (1 g/7,5 ml de tampón). Tras la resuspensión, el precipitado se homogenizó con dos pases sobre un microfluidizador ajustado a 16.000-18.000 PSI. El sistema se enjuagó con 2,5 ml adicionales de tampón por gramo de precipitado (375 ml) para recoger cualquier material en el sistema debido al volumen de retención. La concentración final del homogenizado fue 1 g de precipitado celular/10 ml de tampón (1,5 L). Se añadió Tween-20 al homogenizado (al 0,02% v/v) antes de la clarificación.

Clarificación. El homogenizado celular de BNN93 (1,5 L) se centrifugó inmediatamente a 5000 x G durante 120 minutos en una centrífuga equipada con un rotor basculante (Sorval) a 4°C. El sobrenadante se filtró usando un Cuno Zeta Plus Biocap (0,2 micrómetros). Después de la filtración, se midieron el pH y la conductividad; pH = 8,3, y conductividad = 6,1 ms/cm. El homogenizado filtrado se almacenó a 4°C durante menos de 1 hora antes de la purificación.

Cromatografía en Q-Sepharosa. Se empaquetó una columna de cromatografía Amicon Vantage S (12,5 cm de di) con resina Q-Sepharosa FF (1,5 L) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 volúmenes de columna (VC) de Tris 50 mM, Tween-20 al 0,005%, pH 8,3 (tampón A). El homogenizado clarificado (1,5 L) se diluyó con tampón A (4,5 L) para reducir la conductividad; (después de la dilución: pH = 8,34, 3,2 ms/cm). El homogenizado diluido (6 L) se aplicó a la columna preparada a 49 cm/h (100 ml/min). Cuando se completó la carga, el material no unido se lavó del gradiente de tampón B (NaCl 1 M en tampón A): 4 VC de tampón B al 15%, 4 VC de tampón B al 100% (figura 37). Se analizaron las fracciones por SDS-PAGE y se determinó la actividad enzimática de CST-II (figura 38, tabla 7). Las fracciones de CST-II purificada se juntaron (fracción de elución 2, figura 35, 600 ml) y se tomó una alícuota de 0,5 ml y se formuló con glicerol al 50% y se almacenó a -20°C para conservar la actividad. La CST-II restante se almacenó a 4°C durante aproximadamente 36 horas hasta la purificación adicional en hidroxiapatita.

Tabla 7. Recuperación de la actividad enzimática CST-II después de la purificación en Q-Sepharosa FF

| Fracción mezcla | Actividad (U/ml) | Volumen (ml) | Unidades totales | Recuperación del paso (%) |
|--------------------------------------|------------------|--------------|------------------|---------------------------|
| Homogenizado | 1,72 | 1500 | 2790 | 100 |
| Fracción FT 1 de Q-Sepharosa | 0,011 | 2000 | 22 | 0,8 |
| Fracción FT 2 de Q-Sepharosa | 0,071 | 2000 | 142 | 5,1 |
| Fracción FT 3 de Q-Sepharosa | 0,124 | 2000 | 248 | 8,9 |
| Fracción FT 4 de Q-Sepharosa | 0,03 | 2000 | 60 | 2,2 |
| Fracción FT 5 de Q-Sepharosa | 0,09 | 2000 | 180 | 6,5 |
| Mezcla FT/lavado de Q-Sepharosa | 0,049 | 10000 | 490 | 17,6 |
| Fracción de elución 1 de Q-Sepharosa | 0,078 | 550 | 43 | 1,5 |
| Fracción de elución 2 de Q-Sepharosa | 4,017 | 600 | 2410 | 86,4 |
| Fracción de elución 3 de Q-Sepharosa | 0,553 | 550 | 304 | 10,9 |
| Fracción de elución 4 de Q-Sepharosa | 0,07 | 600 | 42 | 1,5 |
| Fracción de elución 5 de Q-Sepharosa | 0,087 | 600 | 52 | 1,9 |

- o El homogenizado se prepara a partir de 150 gramos de precipitado celular (15 L de fermentación de CST-II produjeron 320,7 gramos de precipitado celular).
- o Recuperación del paso (%): calculada como unidades totales de CST-II en la mezcla individual de fracción dividida por las unidades totales de CST-II en el homogenizado x 100.
- o La fracción de elución 2 de Q-Sepharosa se recogió y preparó para purificación posterior. Fracción FT = fracción del flujo que pasa.

Cromatografía de hidroxiapatita de tipo II (20 micrómetros). Se empaquetó una columna XK50 (GE Healthcare) con resina de hidroxiapatita de tipo II (Biorad, 20 micrómetros) (320 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 VC de fosfato de sodio 5 mM, Tween-20 al 0,005%, pH 6,5. Se añadió lentamente la CST-II purificada por Q-Sepharosa FF (figura 37, fracción de elución 2) (600 ml) a fosfato de sodio monobásico 50 mM (600 ml) mientras se controlaba el pH continuamente. El pH se ajustó a 6,65 con la adición de 60 ml adicionales de fosfato de sodio monobásico 50 mM. La solución de CST-II (660 ml) se diluyó además con fosfato de sodio 10 mM, Tween-20 al 0,005%, pH 6,5 (tampón A, 3,6L) para reducir la concentración de fosfato de sodio total (volumen final 4,2 L). La solución de CST-II condicionada se filtró (botellas filtro Nalgene de 1 L, 0,2 micrómetros) y se midieron el pH, la conductividad y la absorbancia a 280 nm; pH 6,54; conductividad, 4,0 ms/cm, A280 = 1,004 UA. La CST-II filtrada se cargó en la columna preparada a 120 cm/h (39,3 ml/min). El material no unido se lavó de la columna con 5 VC de tampón A. La CST-II se eluyó a 120 cm/h (39,3 ml/min) con un gradiente de tampón B (NaCl 1,5 M en tampón A): tampón B del 0-100% en 25 VC, seguido por 5 VC de tampón B al 100%, seguido por 5 VC de fosfato de sodio 500 mM, pH 6,5 (figura 39). Se tomaron muestras de las fracciones para SDS-PAGE y ensayo de CST-II (figura 40, tabla 8). Las fracciones de CST-II purificada se juntaron (fracción de elución 5, figura 39, 1,65 L) y se tomó una alícuota de 0,5 ml y se formuló con glicerol al 50% y se almacenó a -20°C para conservar la actividad. La CST-II restante se almacenó a 4°C durante aproximadamente 36 hasta la purificación adicional en SP-Sepharosa HP.

Tabla 8. Recuperación de la actividad enzimática CST-II después de la purificación en hidroxiapatita de tipo II (20 micrómetros)

| Fracción | Actividad (U/ml) | Volumen (ml) | Unidades totales | Recuperación del paso (%) |
|---|-----------------------|--------------|------------------|---------------------------|
| Fracción de elución 2 de Q-Sepharosa | 4,017 | 600 | 2410 | 100 |
| Carga de hidroxiapatita | 0,477 | 4200 | 2003 | 83,1 |
| Mezcla FT/lavado de hidroxiapatita | por debajo del límite | 4500 | 0 | 0 |
| Fracción de elución 1 de hidroxiapatita | por debajo del límite | 750 | 0 | 0 |
| Fracción de elución 2 de hidroxiapatita | 0,013 | 560 | 7 | 0,3 |
| Fracción de elución 3 de hidroxiapatita | 0,279 | 600 | 167 | 6,9 |
| Fracción de elución 4 de hidroxiapatita | 0,425 | 100 | 43 | 1,8 |
| Fracción de elución 5 de hidroxiapatita | 0,918 | 1650 | 1515 | 62,9 |
| Fracción de elución 6 de hidroxiapatita | 0,151 | 1150 | 174 | 7,2 |
| Fracción de elución 7 de hidroxiapatita | por debajo del límite | 250 | 0 | 0 |

- 5
- Recuperación del paso (%): calculada como unidades totales de CST-II en la mezcla individual de fracción dividida por las unidades totales de CST-II en la fracción de elución 2 de Q-Sepharosa FF (QFF) x 100.
 - La fracción de elución 5 de hidroxiapatita de tipo II se recogió y preparó para purificación posterior. FT = fracción del flujo que pasa.

10 *Cromatografía en SP-Sepharosa HP.* Se empaquetó una columna de cromatografía XK50 (GE Healthcare) con resina SP-Sepharosa HP (300 ml) y se unió a un sistema AKTA Explorer 100 LC que controla continuamente la absorbancia a 280 nm. La columna se equilibró con 5 VC de MES 25 mM, Tween-20 al 0,005%, pH 6,0 (tampón A). La CST-II purificada en hidroxiapatita de tipo II (20 micrómetros) (figura 39, fracción de elución 5, 1,65 L, Tween-20 al 0,005%) se ajustó a Tween-20 al 0,02% y se dializó durante la noche frente a MES 25 mM, NaCl 10 mM, pH 6,0 (19 L, conductividad; 1,9 ms/cm) a 4°C. El volumen de la solución de CST-II dializada era de 2,2 L y se midieron el pH y la conductividad (pH 6,3; conductividad 11,3 ms/cm). La CST-II dializada se diluyó adicionalmente con tampón A (3,8 L) y se filtró (botellas filtro Nalgene de 1 L, 0,2 micrómetros). Se midieron el pH, conductividad y absorbancia (280 nm) de la solución filtrada (pH 6,03, conductividad: 3,9 ms/cm, A280 = 0,291 UA) y la solución (6 L) se cargó en la columna preparada a 120 cm/h (39,3 ml/min). El material no unido se lavó de la columna a la misma velocidad de flujo con 5 VC de tampón A. La CST-II se eluyó de la columna con un gradiente de tampón B (NaCl 1 M en tampón A): tampón B del 0-50% durante 15 VC, mantenido a tampón B al 15% durante 1 VC y después cualquier material unido estrechamente se eluyó con 5 VC de tampón B al 100% (figura 41). Las fracciones se analizaron por SDS-PAGE (figura 42) y se determinó la actividad enzimática CST-II (tabla 9). Las fracciones de CST-II purificada se juntaron (mezcla de elución 2, figura 41, 648 ml), se tomó una alícuota de 0,5 ml y se formuló con glicerol al 50% y se almacenó a -20°C para conservar la actividad. El resto de CST-II se almacenó a 4°C durante aproximadamente 16 horas antes de la formulación final.

Tabla 9. Recuperación de la actividad enzimática CST-II después de la cromatografía en SP-Sepharosa HP

| Fracción | Actividad (U/ml) | Volumen (ml) | Unidades totales | Recuperación del paso (%) |
|---|-----------------------|--------------|------------------|---------------------------|
| Fracción de elución 5 de hidroxiapatita | 0,918 | 1650 | 1515 | 100 |
| Mezcla FT/lavado de SP-Sepharosa HP | por debajo del límite | 6000 | 0 | 0 |
| Mezcla de elución 2 de SP-Sepharosa HP | 1,67 | 648 | 1082 | 71,4 |
| Mustang E/formulación final | 0,563 | 1460 | 822 | 54,3 |

- 30
- Recuperación del paso (%): calculada como unidades totales de CST-II en la mezcla de fracciones dividida por las unidades totales de CST-II en la fracción de elución 5 de hidroxiapatita (HA) x 100. Se recogió y formuló la mezcla de elución 2 de SP-Sepharosa HP. FT = fracción del flujo que pasa.

35 *Formulación final.* La mezcla de CST-II de SP-Sepharosa HP (mezcla de elución 2, 648 ml, A280: 2,19 UA) estaba compuesta de fracciones que eluían en tampón B entre el 14,7-22% (MES 25 mM, NaCl 1M, Tween-20 al 0,005%, pH 6,0) que corresponde a una concentración media de NaCl de 183,5 mM (tampón B al 18,35%) (figura 41). La concentración de cloruro de sodio se ajustó a 200 mM mediante la adición de NaCl 5 M (2,14 ml). La concentración de Tween-20 aumentó del 0,005% al 0,02% mediante la adición de Tween-20 al 10% (0,975 ml). Se preconditionó un filtro cápsula Mustang E (10 ml) con MES 25 mM, pH 6,0 (130 ml). La solución de CST-II ajustada (650 ml) se pasó a través del filtro de cápsula Mustang E a una velocidad de flujo de 100 ml/min y el filtrado se recogió un envase Nalgene estéril (2 L). La cápsula se enjuagó con MES 25 mM, NaCl 200 mM, Tween-20 al 0,02%, pH 6,0 (aproximadamente 100 ml) para reducir la pérdida de la retención del filtro. La solución de CST-II se pasó a través de una unidad de filtración Nalgene (1 L, 0,2 micrómetros) para eliminar una cantidad vestigial de precipitado fibroso que se observó. La solución de CST-II filtrada (730 ml, A280 = 1,825 UA) se diluyó (1:1) con glicerol (730 ml) en una botella Nalgene estéril (2 L) para dar CST-II en una formulación final de MES 12,5 mM, NaCl 100 mM, Tween-20 al

0,01%, glicerol al 50%, pH 6,0 (1,43 L). La muestra se almacenó inmediatamente a -20°C. La CST-II formulada se analizó por SDS-PAGE (figura 43), RP-HPLC (figura 44), actividad enzimática CST-II (tablas 9 y 10) y concentración de endotoxina.

5 **Tabla 10. Resumen del proceso de CST-II**

| Paso del proceso | Proteína total (mg) | Unidades totales | Endotoxina (UE/mg) | Factor de purificación | Recuperación (%) |
|---------------------------|---------------------|------------------|--------------------|------------------------|------------------|
| Homogenizado (150 g) | na | 2790 | 192200* | na | 100 |
| Q Sepharosa | 4772,7 | 2410 | 303 | na | 86,4 |
| Hidroxiapatita de tipo II | 1679,3 | 1515 | < 49 | 1,79 | 54,3 |
| SP Sepharosa HP | 1402,3 | 1082 | 2,98 | 0,86 | 38,8 |
| Formulación final | 1316,5 | 822 | 0,987 | 0,81 | 29,5 |

Nota: * La endotoxina se describe como UE/mg en la recogida de la enzima. La actividad específica de la CST-II purificada es 0,624 U/mg de proteína. Recuperación (%) calculada como actividad de CST-II recuperada del paso de purificación dividido por las unidades totales al inicio del proceso (2790 unidades) x 100.

10 *Resultados de la purificación de CST-II a gran escala.* Se homogenizó CST-II de 150 gramos de precipitado celular de *E. coli* BNN93 (menos de la mitad de una fermentación única de 15 L) en presencia de EDTA y se clarificó en presencia de Tween-20. El homogenizado clarificado se capturó en una columna de 1,5 L de Q-Sepharosa (Fast Flow) y se eluyó con un gradiente en pasos de NaCl en presencia de Tween-20 (figuras 35 y 36). La actividad CST-II se recuperó del paso de Q-Sepharosa en un rendimiento del 86% (recuperación del paso) basado en la actividad enzimática (tabla 8). La mezcla de elución de Q-Sepharosa se diluyó en tampón fosfato de sodio y se capturó en resina de hidroxiapatita de tipo II (20 micrómetros). La CST-II se eluyó en presencia de Tween-20 en una recuperación del paso mayor del 75% de actividad enzimática (figuras 39 y 38, tabla 8). La mezcla de elución de hidroxiapatita se dializó en tampón MES con Tween-20 y se purificó por cromatografía de SP-Sepharosa HP (figuras 41 y 42, tabla 9). La CST-II se eluyó con NaCl y se recogió una mezcla de fracción estrecha (mezcla de elución 2) a una concentración alta de la enzima. Se usó una mezcla de fracción más estrecha para mantener la alta concentración de proteína después de este paso, eliminado de esta manera la necesidad de concentrar la CST-II antes de la formulación (mezcla de elución 2 frente a mezcla de elución 1) (figura 41, tabla 11). La CST-II purificada se filtró para reducir la endotoxina y la carga biológica y se formuló en glicerol al 50%. La recuperación global de la actividad enzimática para el proceso entero fue casi del 30% (tabla 10). La enzima CST-II resultante tenía una actividad específica de 0,624 U/mg y una pureza de más del 98% por análisis de RP-HPLC (tabla 12, figuras 43 y 44).

30 **Tabla 11. Comparación de dos mezclas de elución del producto de SP Sepharosa HP**

| SP Sepharosa HP | Volumen (ml) | A280 | mg totales |
|---------------------|--------------|------|------------|
| Mezcla de elución 1 | 1128 | 1,40 | 1560,5 |
| Mezcla de elución 2 | 648 | 2,19 | 1402,3 |

Tabla 12. Secuencia de aminoácidos N-terminales de CST-II purificada en condiciones reductoras

| Banda | Secuencia de aminoácidos N-terminal | | | | | | | | | |
|---------------------|-------------------------------------|---|---|---|---|---|---|---|---|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| CST-II (principal) | G | S | M | K | K | V | I | I | A | G |
| CST-II (secundaria) | M | K | K | V | I | I | A | G | N | G |
| A (principal) | G | S | M | K | K | V | I | I | A | G |
| A (secundaria) | M | K | K | V | I | I | A | G | N | G |
| B (principal) | G | S | M | K | K | V | I | I | A | G |
| B (secundaria) | M | K | K | V | I | - | - | - | - | - |
| C (principal) | N | L | L | K | L | A | - | - | - | - |
| C (secundaria) | L | A | P | N | F | K | - | - | - | - |
| C (secundaria) | F | K | N | D | N | S | - | - | - | - |

- 35
- Se cortaron las bandas de proteína identificadas en la electrotransferencia (PVDF) de un gel de SDS-PAGE (figura 50) para el análisis individual como se describe en la sección de métodos. Se muestran las secuencias N-terminales de las dos especies esperadas de CST-II. Solo se detectó proteína CST-II.
 - CST-II principal; la proteína predominante secuenciada para CST-II total purificada.
 - CST-II secundaria; la secuencia de CST-II secundaria detectada estimada que es aproximadamente el 10% de la CST-II purificada total.
 - Banda A: CST-II (intacta), extremos N esperados principal y secundario observados.
 - Banda B: Fragmento N-terminal de CST-II, extremos N esperados principal y secundarios observados.
- 40

- o Banda C: Fragmento interno de CST-II con extremo irregular que empieza en N¹⁷², L¹⁷⁶ y F¹⁸⁰ (basado en la especie GSMKK).

Conclusión

5 Se evaluó un panel extenso de métodos de cromatografía para la purificación de CST-II del plásmido pNT1 expresado en células de *E. coli* BNN93. Se identificaron cuatro metodologías de purificación eficaces que incluían captura en Q Sepharosa, hidroxapatita de tipo II, SP Sepharosa HP y butil Sepharosa. Usando estos métodos de cromatografía, se desarrolló un método escalable de 3 pasos (Q Sepharosa FF, hidroxapatita de tipo II y SP Sepharosa HP) para la producción de CST-II (figura 4). Se utilizaron pasos de dilución y diálisis en lugar de filtración con intercambio de tampón para minimizar la pérdida de CST-II. Este proceso produjo con éxito 822 unidades de CSTII a partir de 150 gramos de precipitado celular (2786 unidades de actividad enzimática cruda) a una pureza mayor del 98% con un rendimiento total del 29,5% basándose en la actividad enzimática (tabla 10).

15 Lista de secuencias informal

SEQ ID NO: 1

Secuencia de nucleótidos del inserto de CSTII en pNT1 (cepa OH4384):

atgaaaaaagtattattgctggaaatggaccaagttaaaa gaaattgattattcaagactaccaaagatttgatgtatttagatgtaatc
aattttatttgaagataaatactatcttggtaaaaaatgcaaggcagattttacaatcctagcttttttgaacaatactacactttaaaca
ttaaaccaaaatcaagaatatgagaccgaactaattatgtgttaatacaaccaagctcatctagaaaatgaaaatttgaaaaactttt
acgattatttctgatgctcatttgggatatgatttttcaacaactaaagatttaaatgctattttaaattcagaaaattattcaatcaa
gaattacctcaggggtctatatgtgtgcagtagccatagccctaggatacaaagaaattatcttccgggaattgattttatcaaaatgggt
catcttatgcttttgatactaaacaaaaaatcttttaaattggctcctaattttaaataatgataattcacactatattggacatagtaaaaata
cagatataaaaagcttttagaatttctagaaaaaaacttaaaaaataaactatattgcttatgtcctaaca gtcttttagcaaatttatagaacta
gcgccaatttaaatcaaatttatcatacaagaaaaaaataactacactaaagatatactcataccttagtgaggcttatggaaaatttt
caaaaaatattaatttttaa

20

SEQ ID NO: 2

Secuencia de proteína del inserto de CSTII en pNT1:

mgs mkkvii agngpslkeidysrlpndfdvfrncqfyfedkyylgkckkavfynpslffe qyytlkhliqnqeyetelimcsnyn
qahlenenfvktfydyfpdahlgydffkqlkdfnayfkfheiyfnqritsgvymcavaialgykeiylsgidfyngssyafdkq
knllklapnfkndnshyighskntdikaleflekykiklyclpnsllanfielapnlnsnfiqeknytkdilipsseaygkfskni
nf

25

SEQ ID NO: 3

Plásmido de expresión que incluye CSTII con etiqueta (pNT1):

tcgctgttagcgggcccattaagtctgtctcggcgcgtctgctgtggctggctggcataaatatctcactcgcgaatcaaattcagcc
gatagcgggaacgggaaggcactggagtgccatgtccggtttcaacaaccatgcaaagtgaatgagggcatcgttcccactgc
gatgtggttccaacgatcagatggcgtgggcgcaalgcgcgccattaccgagtcgggctgcgcgttggcgggatctcgggt
agtgggatacgcgataccgaagacagctcatgttatatccgccgttaaccaccatcaaacaggattttcgcctgctggggcaaacc
agcgtggaccgcttctgcaactctcagggccaggcggggaaggcaatcagctgttcccgtctcactggtgaaaagaaaacc
accctggcgcccaatacgcgaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggttcccactgga
gctgggagtgagcgcgaacgcaattaatgtaagttagctcactcattaggcaccacaggctttacaattatgctccggctcgtatggc
ttcgggtgatgacgggtaaacctctgacacatgcagctcccggagacgggtcacagcttctctgtaaggatgcccgggagcagaca
agcccgtcagggcgcgtcagcgggtgttggcgggtgtcggggcgcagccatgaccagtcacgtagcgtatgcccggagtgtatact
ggcttaactatgcccgatcagagcagattgtactgagagtgcaccattatgcccgtgtaaataccgcacagatgctgaaggagaaaat
accgcatcagggcgtctctccgctctcctcgtcactgactcgcgtcggctgctcggctgcggcgagcggatcagctcactcaaa
gctgtaatacgggtatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgta
aaaaggccgctgtgctggcgttttccataggtccgccccctgacgagcatcaaaaaatcgacgctcaagtcagagggtgctgaa
accgacaggactataaagataaccaggcgtttccccctggaagctccctcgtcgcgtctcctgtccgacctgccgctaccggatac
ctgtccgctttctccctcgggaagcgtggcgtttctcatagctcagctgtaggatctcagttcgggtgtaggtcgttcccaagct
gggctgtgtgcaacccccgttcagcccgaccgctgccttatccggtaactatcgtcttgagtcacaaccgggtaagacacgac
ttatgccactggcagcagccactgtaacaggattagcagagcggaggtatgtagggcgtgctacagagttcttgaagtgtggccta
actacggctacactagaaggacagatattggatctcgcctctgctgaagccagttacctcggaaaaaggattgtagctcttgatccg
gcaacaaccaccgctggtagcgggtgtttttgttgaagcagcagattacgcgcagaaaaaaggatcacaagaagatcctttg
atctttctacgggctgacgctcagtggaacgaaaactcaggttaagggtttgtcatgagattatcaaaaaggatctcacctagat
ccttttaataaaaaatgaagtttaaatcaatctaaa gtatatatgagtaaaactggctgacagttaccaatgcttaatcagtgaggcacct
atctcagcgtatctgtctatttcgttcatccatagttgcctgactcccgtcgtgtagataactacgatacgggagggcttaccatctggccc
cagtgctgcaatgataccgcgagaccacgctcaccggctccagattatcagcaataaacagccagccggaaggccgagcgcga
gaagtggctcctgcaacttatccgctccatccagctctataattgttgcgggaagctagagtaagtagtccgagtaatagtttgcgc
aacgttgtgccattgctgcaaggcatcgtggtgtcacgctcgtcgttggtagcttaccagctccggttcccaacgatcaaggcgag
ttacatgatccccatgttgtgcaaaaaagcgggttagctcctcggctccgatcgggggggggggaaagccacgttgtgtctcaa
atctctgatgttaccatgcacaagataaaaatatacatcatgaacaataaaactgtctgcttacataaacagtaatacaagggtgttatga
gcatattcaacgggaaacgcttctccaggccgcgattaaattccaacatggatgctgatttatgggtataaatgggctcgcgata

atgtcgggcaatcaggtgacgaatctatcgactgtatgggaagccgatgcccagagttgttctgaaacatggcaaggtagcgtt
 gccaatgatgtacagatgagatgggtcagactaaactggctgacggaattatgacctctccgacctcaagcattttatccgtactcctga
 tgatgcatggttactcaccactgcatccccgggaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatattggtgat
 gcgctggcagtgctcctgcccgggtgcaatcattcctgttgaattgcttttaacagcgatcgcgtatttcgctcgcctcaggcgcaa
 tcacgaatgaataacggttgggtgatgagtgatttgatgacgagcgaatggctggcctgttgaacaagtctgaaagaaatgcat
 aagctattgccattcaccggattcagtcactcatggtgatttctactgataacctatgtttgacgaggggaaattaagggtgat
 tgatgttgacgagtcggaatcgcagaccgataccaggatcttccatcctatggaactgcctcgggtgagtttctccttaccagaaa
 cggcttttcaaaaataggattgataatcctgatataaataatgcaattcattgatgctcgaatgattttctaaagtactactctctt
 tcaatattattgaagcattatcagggttattgctcatgagcggatacatattgaatgtattgaaaaataaacaatagggttccgcg
 cacattccccgaaaagtgccacctgacgatgaaattgaaacgtaaatatttgttaaaattcgcgttaaatttgttaaatcagctcatttt
 taaccaataggccgaaatcggcaaaatccctataaatcaaaaagaatgcccagatagggtgaggtgtgtccagtttggacaaga
 gtccactataaagaactggactccaacgtcaaaaggcgcaaaaaccgtctatcaggggcgatggccactacgtgaacctacacca
 aatcaagtttttggggtcgaggtgccgtaaagctcaaatcggaaacctaaaggagccccgatttagagcttgacggggaaagcc
 ggcaactggtggcgaagaaagggaagaaagcgaagaggagggcgctaggcgctggcaagtgtagcgggtcacgctgcgc
 gtaaccaccacaccgcccgttaatgcccgtacagggcgctactatggttgccttgacgcatcgttaagaaaccattatataca
 tgacattaacctataaaaataggcgtatcacgagggcccttctgctcaagcagatctgaaaaaaaagcccgtcattaggcgggctca
 gatcgtcctatgttgacagcttatcagatgacggtaccgaattcctcagctatfaaaaataatattttgaaaatttccataagc
 ctactagaaggatgagatatactttgtagtatttttctgtatgataaaattgaattaaatttggcgtacttataaaattgctaa
 aagactgttaggacataaagaatatagttttatttgaatttttctgataaaattgaaatttataatctgtattttactatgtccaatatagtg
 tgaattatcattttaaaattaggagccaattttaaagatttttgttagatcaaaagcataaagatgaccattttgataaaaatcaattccc
 gaaagataaatttcttctatcctagggtatggctactgcacacataatagcccctgaggttaattcttattgaaataaattcgtgaaatt
 taaaataagcattaaaactttaagttgttgaaaaaatcatatcccaaatgagcatcaggaaaataatcgtaaaagttttcaaaaatttc
 atttctagatgagcttgggtgtaattagaacacataaattagtcggctcatattcttgattttggattaaatgtttaaagttagtattgttcaa
 aaaaaagactaggattgtaaaactgccttgcatttttaccagagatatttctcctcaaaaataaattgattacatcaaatatcaaaa
 tcatgttagtcttgaataatcaatttctttaaacttggctcattccagcaataataactttttcatggatccatgatgacctcctaagcatc
 gatagatcctgttctgtgtgaaattgttatccgctcacaattccacacattatacagccgatgattaattgtcaacagggggatgggga
 gtaagctgatcctgttctgtgtgaaattgttatccgctcacaattccacacattatacagccgatgattaattgtcaacagggggatgg
 ggagtaagctcatcagatgagatcctgttctgtgtgaaattgttatccgctcacaattccacacattatacagccggaagcataaa
 gtgtaaaacctggggtgcctaatgagtgagctaacctacattaattgcgttgcgctcactcccgttccagtcgggaaacctgtcgtg
 ccaggacaccatogaatggtgcaaaccttccggtatggcatgatacggccggaagagagtcattcagggtggtgaatgtgaa
 accagtaacgttatacagatgctgcagagatgccggtgctctatcagaccgttcccgtggtgaccaggccagccagtttctgc
 gaaaacgcgggaaaagtggaagcggcgatggcggagctgaattacattccaaccggtggcacaacaactggcgggcaaca
 gtcgttgcgtgattggcgttgcacctccagcttggccctgcacgcgctgcgcaattgtcgcggcgattaaatctgcgccgatcaac
 tgggtgccagcgtggtgctgatggtagaacgaagcggcgtgaagcgtgaaagcggcggtgcacaatcttctgcgcaacgc
 gtcagtggtgatcattaactatccgctggatgaccaggatgccattgctgtggaagctgcctgactaatgtccggcgttatttctga
 gtctctgaccagacaccatcaacagatattttctccatgaagacggtagcggactggcggtggagcatctgttcgattgggtca
 ccagcaaa

SEQ ID NO:4
 >Cst-II: Proteína de OH4384 (nativa)

5 MKKVI IAGNGPSLKEI DY SRLP NDFV FRCNQFY FEDKY YLGKKCKAVFY NPIL FFEQY YTLKHLIQNQEYETEL
 IMCSNY NQAHL ENENFVKTFY DYFP DAHLGYDFFKQLKDFNAYFKFHEI YFNQRITSGVYMCVAIAI ALGYKEIYL
 SGIDFYQNGSSYAFDTKQKNLLKLPNFKNDNSHY IGH SKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAP
 NLNSNFI IQEKNNYTKDILIP SSEAYGKFSKNI NFKKIKI KENIYYKLIKDLLRLPSDIKHYFKGK

SEQ ID NO: 5
 >Cst-II: Aminoácidos 1-260 de la proteína de OH4384 con la mutación lle53Ser

10 mkkviiagngpslkeidysrlpndfdvfrncqfyfedkyylgkkckavfynpslffeqqyylkhliqnqeyetelimsnynqahl
 enenfvtkydyfypdahlgdydfkqlkdfnayfkfheiyfnqrirtsqyvmcavaialgykeiylsgidfyqngssyafdkqknllk
 lpnfnkndnshyighskntdikaleflektykiklyclcpnsllanfielapnlnsnfiieqknytkdilipsseaygkfskninf

SEQ ID NO: 6
 >Cst-II: Proteína de OH4382

15 MKKVI IAGNGPSLKEI DY SRLP NDFV FRCNQFY FEDKY YLGKKCKAVFY NPIL FFEQY YTLKHLIQNQEYETEL
 IMCSNY NQAHL ENENFVKTFY DYFP DAHLGYDFFKQLKDFNAYFKFHEI YFNQRITSGVYMCVAIAI ALGYKEIYL
 SGIDFYQNGSSYAFDTKQKNLLKLPNFKNDNSHY IGH SKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAP
 NLNSNFI IQEKNNYTKDILIP SSEAYGKFSKNI NFKKIKI KENIYYKLIKDLLRLPSDIKHYFKGK

SEQ ID NO: 7
>Cst-II: Proteína de HB93-13

**MKKV I IAGNGPSLKEIDYSRLPNDFDVFR CNQFYFEDKY YLGKKCKAVFY NPILFFEQYYTLKHLIQNQEYETEL
IMCSNYNQAHL ENENFVKTFYDYFPDAHLGYDFFKQLKDFNAYFKFHEIYFNQRITSGVYMCVAIAIALGYKEIYL
SGIDFYQNGSSYAFDTKQKNLLKLAPNFKNDNSHYIGH SKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAP
NLNSNFI IQEKNNYTKDILIP SSEAYGKF SKNINFKKIKIKENIYYKLIKDLLRRLPSDIKH YFKGK**

5
SEQ ID NO: 8
>Cst-II: Proteína de O:10 ATCC 43438

**MKKV I ISGNGPSLKEIDYSRLPNDFDVFR CNQFYFEDKY YLGKKCKAVFY NPGLFFEQYYTLKHLIQNQEYETEL
IMCSNYNQAHL ENENFVKTFYDYFPDAHLGYDFFKQLKEFNAYFKFHEIYLNQRITSGVYMCVAIAIALGYKEIYL
SGIDFYQNGSSYAFDTKQENLLKLAPDFKNDRSHYIGH SKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAP
NLNSNFI IQEKNNYTKDILIP SSEAYGKF SKNINFKKIKIKENIYYKLIKDLLRRLPSDIKH YFKGK**

10
SEQ ID NO: 9
>Cst-II: Proteína de O:23 ATCC 43449

**MKKV I ISGNGPSLKEIDYSRLPNDFDVFR CNQFYFEDKY YLGKKCKAVFY NP SLFFEQYYTLKHLIQNQEY EIEI EL
IMCSNYNQAHL ENENFVKTFYDYFPDAHLGYDFFKQLKEFNAYFKFHEIYFNQRITSGVYMC T VAI AIALGYKEIYL
SGIDFYDNGGGYAFDTKQKNLLKLAPNFKNDNSHYIGH SKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAP
NLNSNFI IQEKNNYTKDILIP SSEAYGKF TKNINFKKIKIKENIYYKLIKDLLKLP S DIKH YFKGK**

15
SEQ ID NO: 10
>Cst-II: Proteína de O:41 ATCC 43460

**MKKV I ISGNGPSLKEIDYSRLPNDFDVFR CNQFYFEDKY YLGKKCKAVFY NP SLFFEQYYTLKHLIQNQEY ETEL
IMCSNFNQAHL ENQNFVKTFYDYFPDAHLGYDFFKQLKEFNAYFKFHEIYFNQRITSGVYMC T VAI AIALGYKEIYL
SGIDFYQNGSSYAFDTKQKNLLKLAPNFKNDNSHYIGH SKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAP
NLNSNFI IQEKNNYTKDILIP SSEAYGKF TKNINFKKIKIKENIYYKLIKDLLRRLPSDIKH YFKGK**

20
SEQ ID NO: 11
Motivo A de sialiltransferasa
DVFR CNQFYFED/E

25
SEQ ID NO: 12
Motivo B de sialiltransferasa
RITSGVYMC

30
SEQ ID NO: 13
DVFR CNQFYFED

SEQ ID NO: 14
DVFR CNQFYFEE

35
SEQ ID NO: 15
Especie principal de CST-II producida por el plásmido pNT1

**GSMKKV I IAGNGPSLKEIDYSRLPNDFDVFR CNQFYFEDKY YLGKKCKAVFY NP SLF
FEQYYTLKHLIQNQEYETELIMCSNYNQAHL ENENFVKTFYDYFPDAHLGYDFFKQL
KDFNAYFKFHEIYFNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQK
NLLKLAPNFKNDNSHYIGH SKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNL
NSNFIIQEKNNYTKDILIP SSEAYGKF SKNINF**

40

REIVINDICACIONES

- 5 1. Una proteína sialiltransferasa con etiqueta que comprende un polipéptido sialiltransferasa con al menos el 95% de identidad a SEQ ID NO:4 y una etiqueta MGS en el extremo amino de SEQ ID NO:4, en donde la proteína sialiltransferasa con etiqueta tiene actividad sialiltransferasa y la etiqueta MGS es el tripéptido Met-Gly-Ser.
- 10 2. La proteína sialiltransferasa con etiqueta de la reivindicación 1, en donde el polipéptido sialiltransferasa es SEQ ID NO:4.
- 15 3. La proteína sialiltransferasa con etiqueta según la reivindicación 1 que comprende un polipéptido sialiltransferasa y una etiqueta MGS en el extremo amino, en donde el polipéptido sialiltransferasa con etiqueta comprende el motivo A de sialiltransferasa, DVFRCNQF YFED/E (SEQ ID NO:11) y motivo B de sialiltransferasa, RITSGVYMC (SEQ ID NO: 12), y en donde la proteína sialiltransferasa con etiqueta tiene actividad sialiltransferasa.
- 20 4. Un ácido nucleico que codifica la proteína sialiltransferasa con etiqueta de la reivindicación 1, 2 o 3.
- 5 5. Un vector de expresión que comprende el ácido nucleico de la reivindicación 4.
6. El vector de expresión de la reivindicación 5 que tiene la secuencia de SEQ ID NO: 3.
- 25 7. Una célula huésped que comprende el vector de expresión de la reivindicación 5.
8. Un método de hacer una proteína sialiltransferasa con etiqueta, el método comprende cultivar la célula huésped de la reivindicación 7 en condiciones adecuadas para la expresión de la proteína sialiltransferasa con etiqueta.
- 30 9. El método de la reivindicación 8, que comprende además el paso de purificar la proteína sialiltransferasa con etiqueta.
- 35 10. Un método in vitro de producir un producto oligosacárido sialilado, el método comprende los pasos de a) poner en contacto un sustrato aceptor que comprende un oligosacárido con la sialiltransferasa con etiqueta de la reivindicación 1, 2 o 3 y con un sustrato donante que comprende un grupo ácido siálico; y b) permitir la transferencia del grupo ácido siálico del sustrato donante al sustrato aceptor, produciendo de esta manera el producto oligosacárido sialilado.
- 40 11. Un método in vitro de producir una proteína eritropoyetina sialilada, el método comprende los pasos de a) poner en contacto una proteína eritropoyetina que comprende un oligosacárido con la sialiltransferasa con etiqueta de la reivindicación 1, 2 o 3 y con un sustrato donante que comprende un grupo ácido siálico; y b) permitir la transferencia del grupo ácido siálico del sustrato donante a la proteína eritropoyetina, produciendo de esta manera la proteína eritropoyetina sialilada.
- 45 12. El método de la reivindicación 11, en donde el grupo ácido siálico es un ácido siálico PEGilado.

FIG 1

Vectores pcWin2/cst-II y II

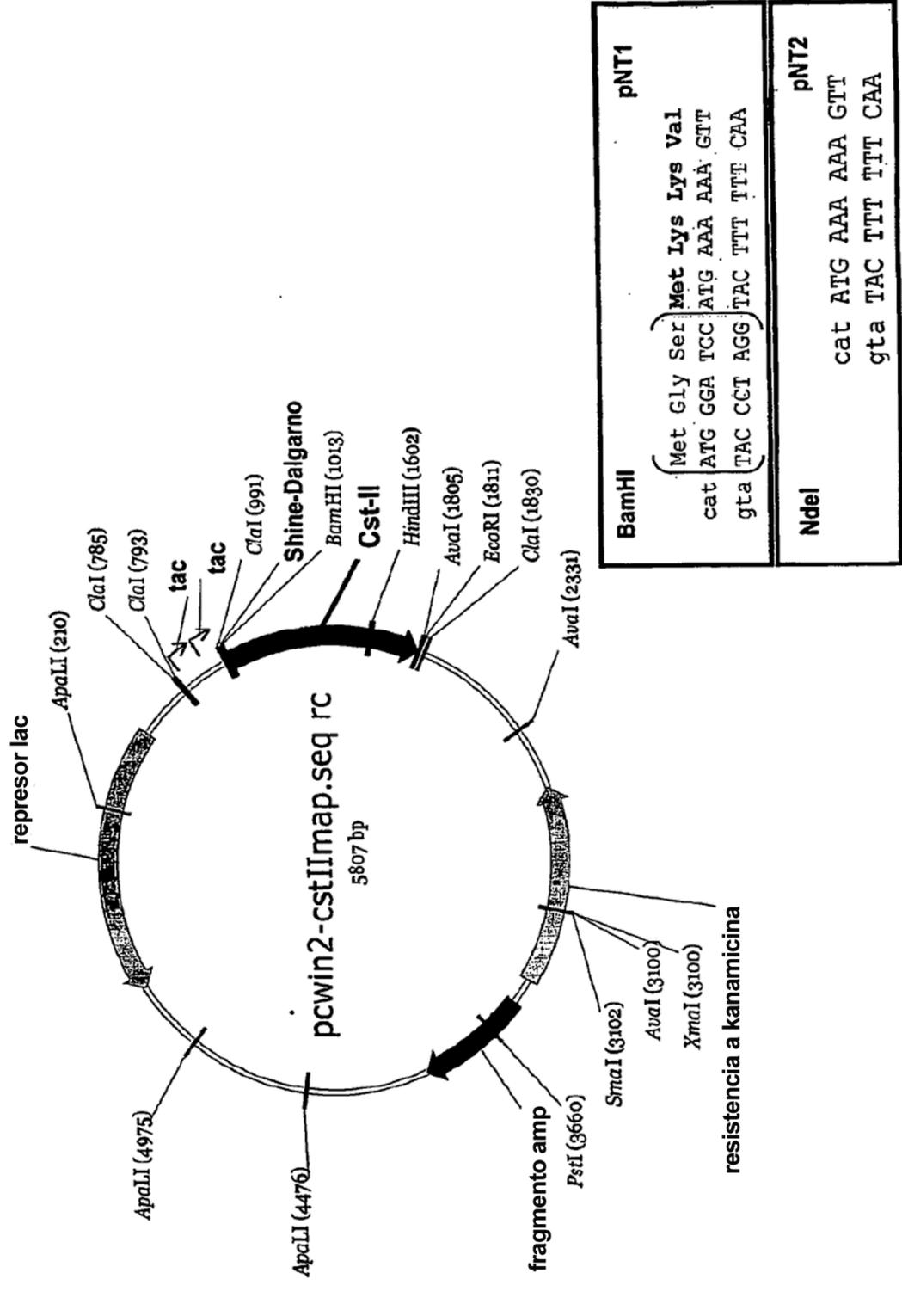


FIGURA 2

(MGS)MKKVIIAGNGPSLKEIDYSRLPNDFDVFR CNQFYFEDKY YLGKKCKAVFYNPS
LFFEQYYTLKHLIQNQEYETELIMCSNYNQA HLENENFVKTFYDYFPDAHLGYDFFK
QLKDFNAYFKFHEIYFNQRITSGVYMC AVAIALGYKEIYLSGIDFYQNGSSYAFDTKQ
KNLLKLAPNFKNDNSHYIGH SKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPN
LNSNFIIQEKNNYTKDILIP SSEAYGKFSKNINF

Componentes de los medios **FIG 3**

Medio E1

| Reactivo | Concentración |
|-----------------------------|--------------------------------------|
| Martone B-1 | 10g/L |
| Extracto de levadura Marcor | 5g/L |
| Cloruro de sodio | 10g/L |
| Agua de O.I. | Según se necesite |
| PPG 2000 | 1mL/L |
| Ácido fosfórico | Según se necesite para mantener pH=7 |
| Hidróxido de sodio | Según se necesite para mantener pH=7 |

Medio E6

| Reactivo | Concentración (g/L) |
|-----------------------------------|---------------------|
| Martone B-1 | 37.5 |
| Extracto de levadura Marcor | 19.75 |
| Dextrosa | 3.13 |
| Fosfato de amonio | 3.75 |
| PPG 2000 | 1 |
| Fosfato de potasio, monobásico | 3.75 |
| Fosfato de potasio, dibásico | 4.22 |
| Sulfato ferroso | 0.0205 |
| Sulfato de magnesio, heptahidrato | 1.01 |
| Agua de O.I. | Según se necesite |

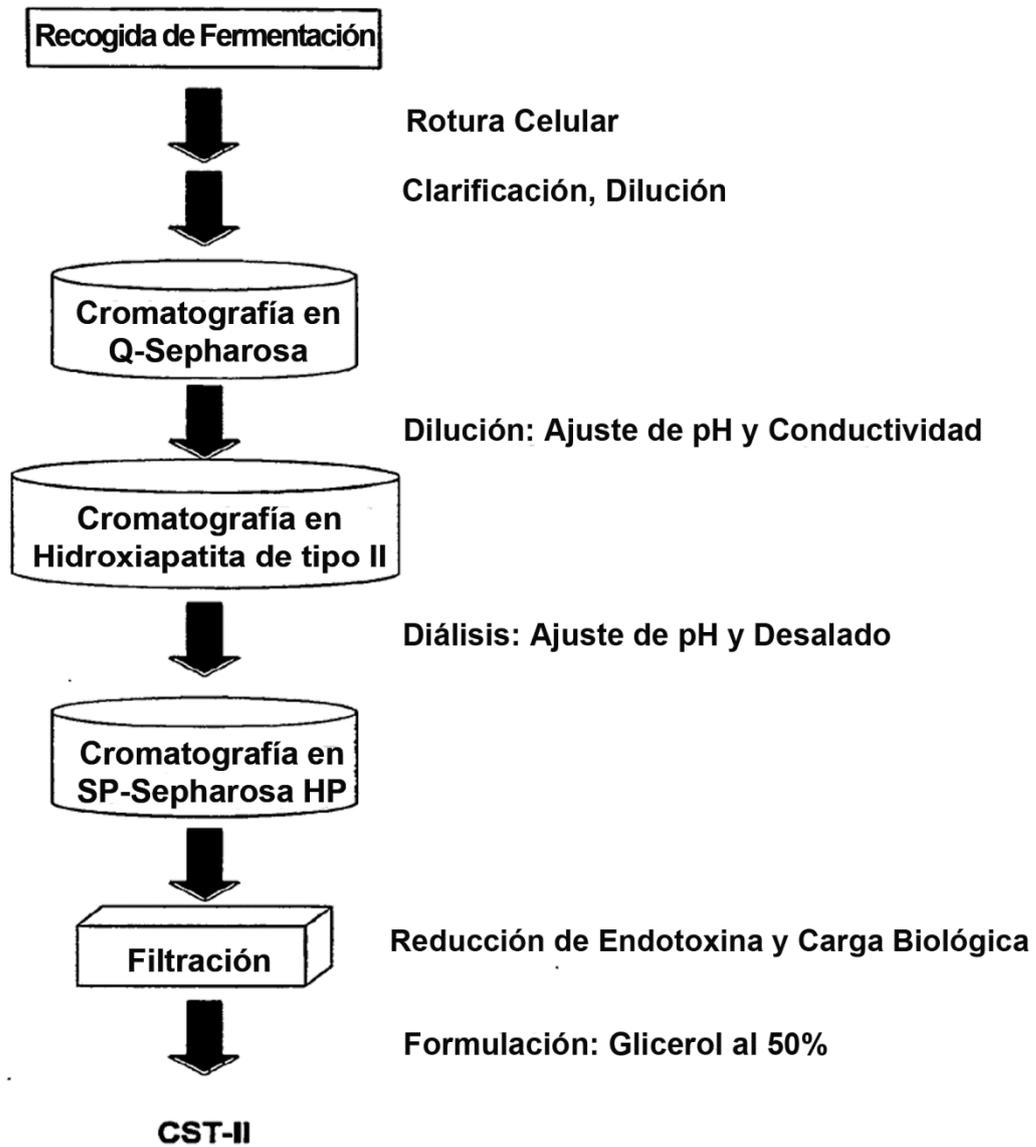


Figura 4

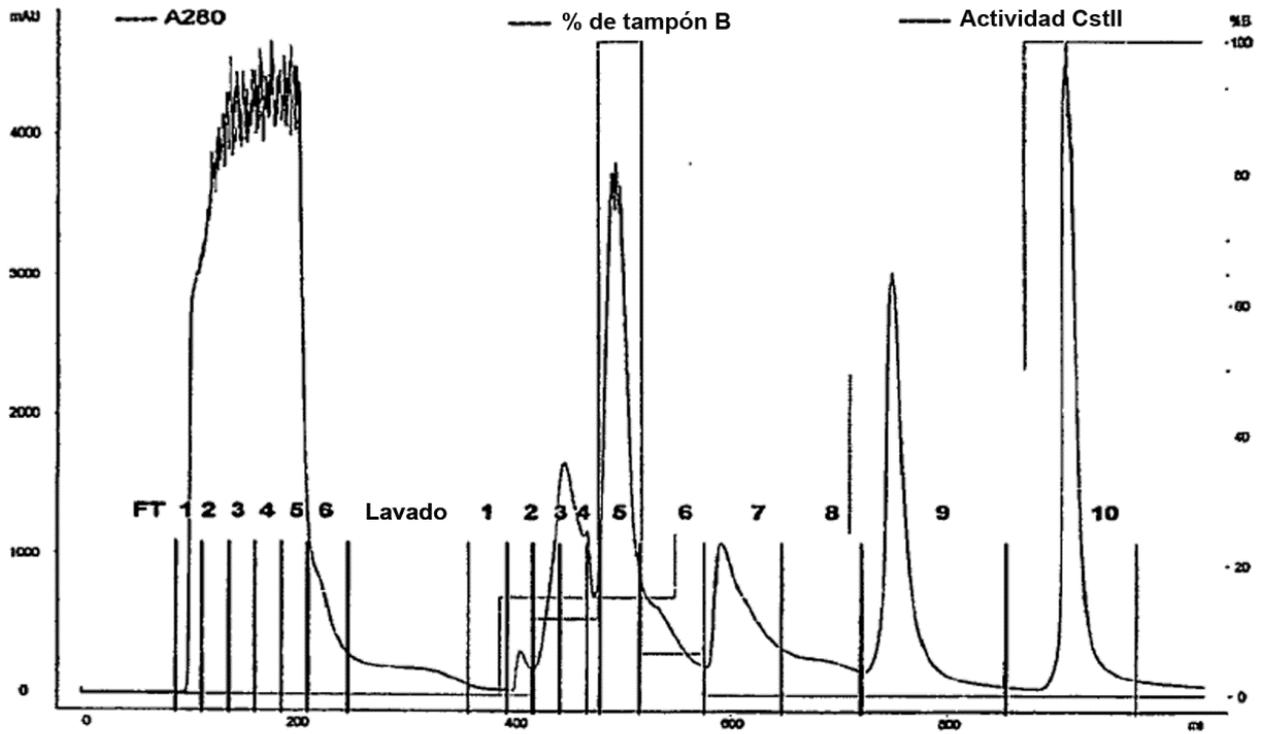


Figura 5

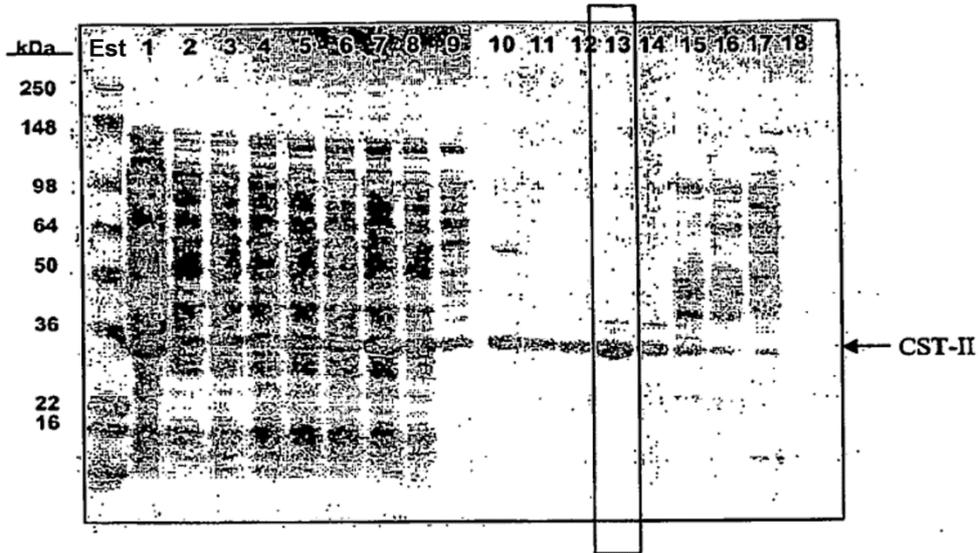


Figura 6

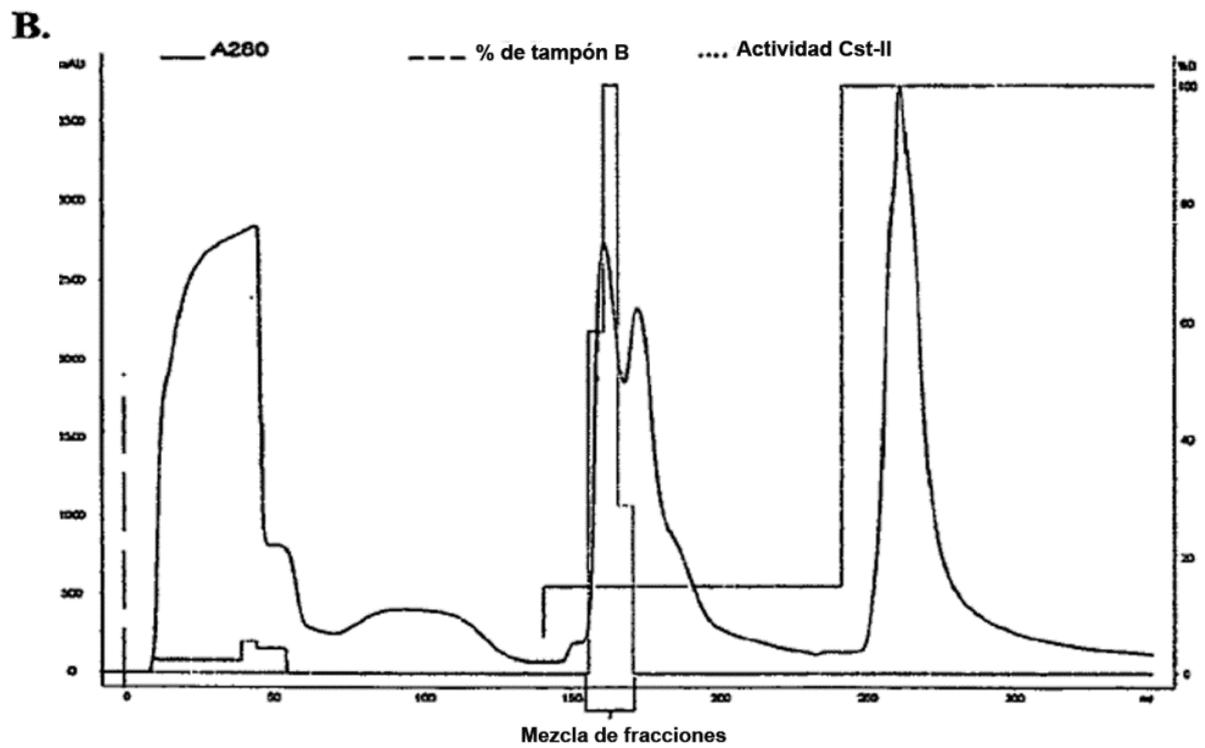
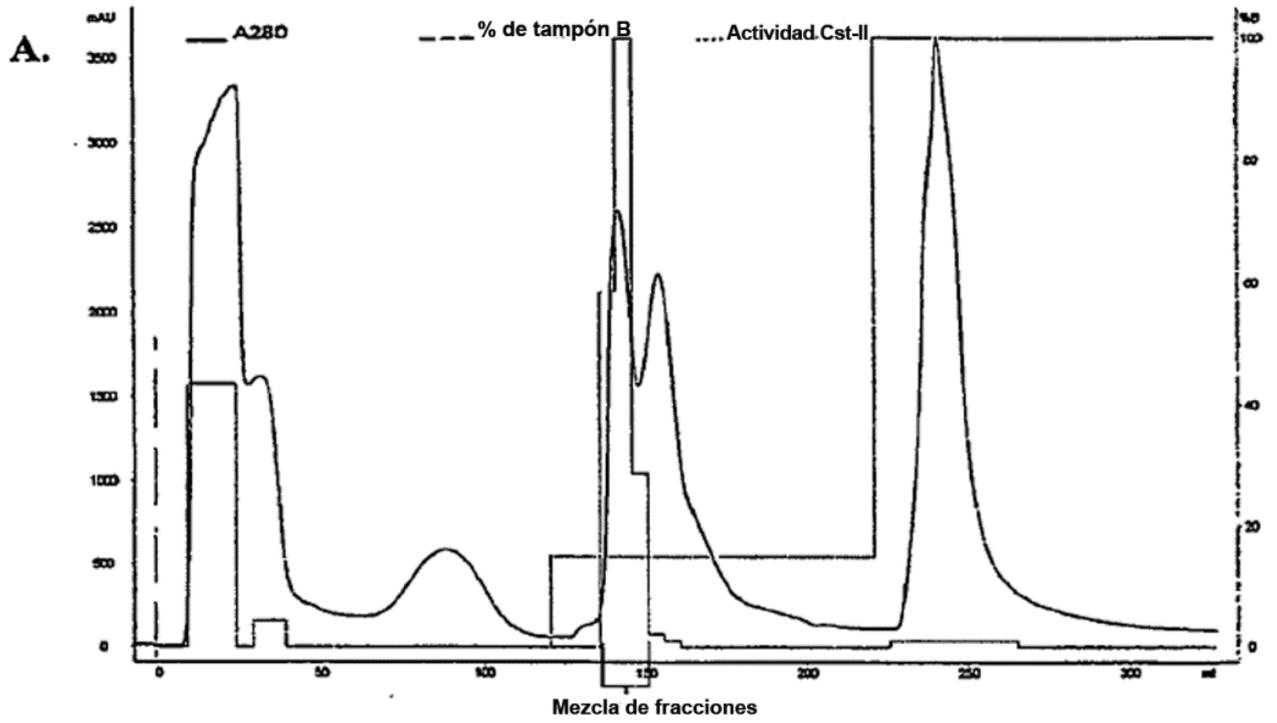


Figura 7

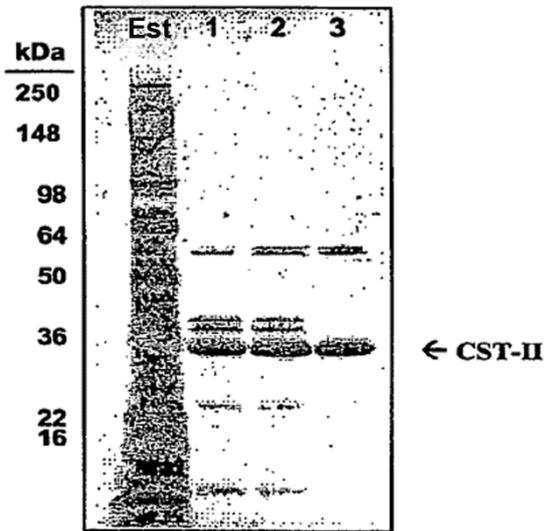


Figura 8

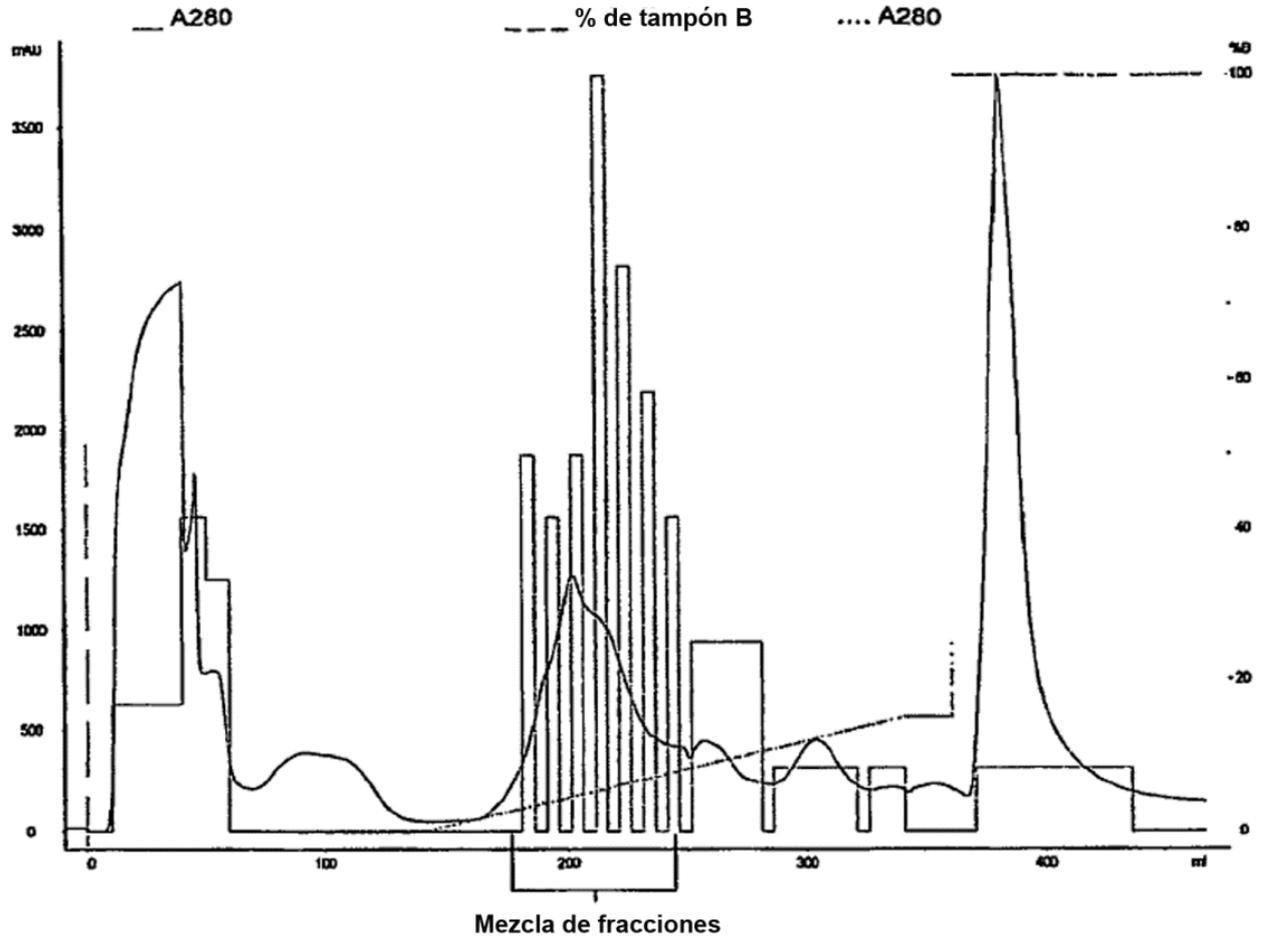


Figura 9

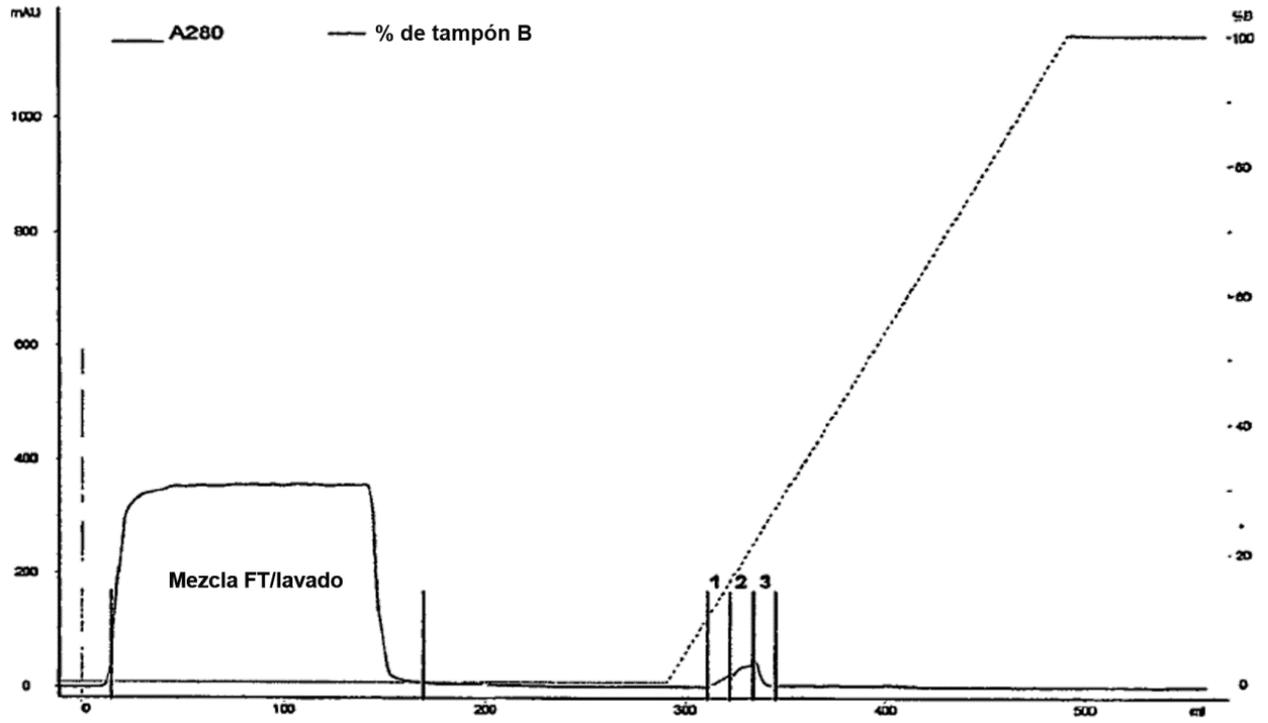


Figura 10

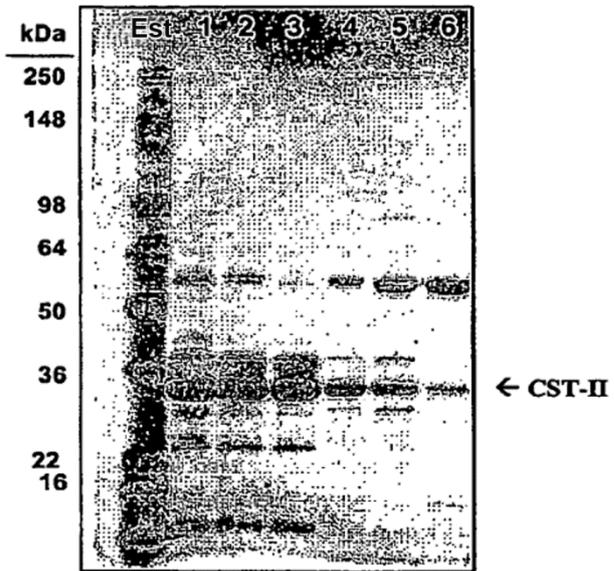


Figura 11

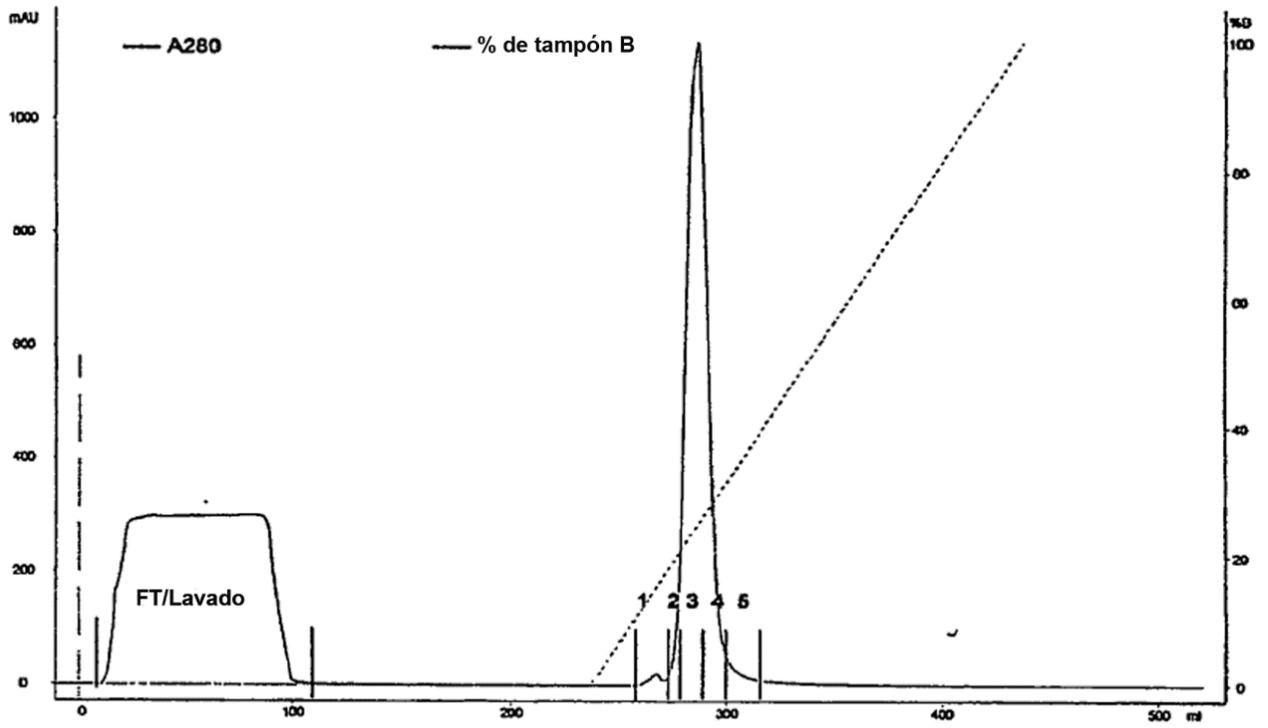


Figura 12

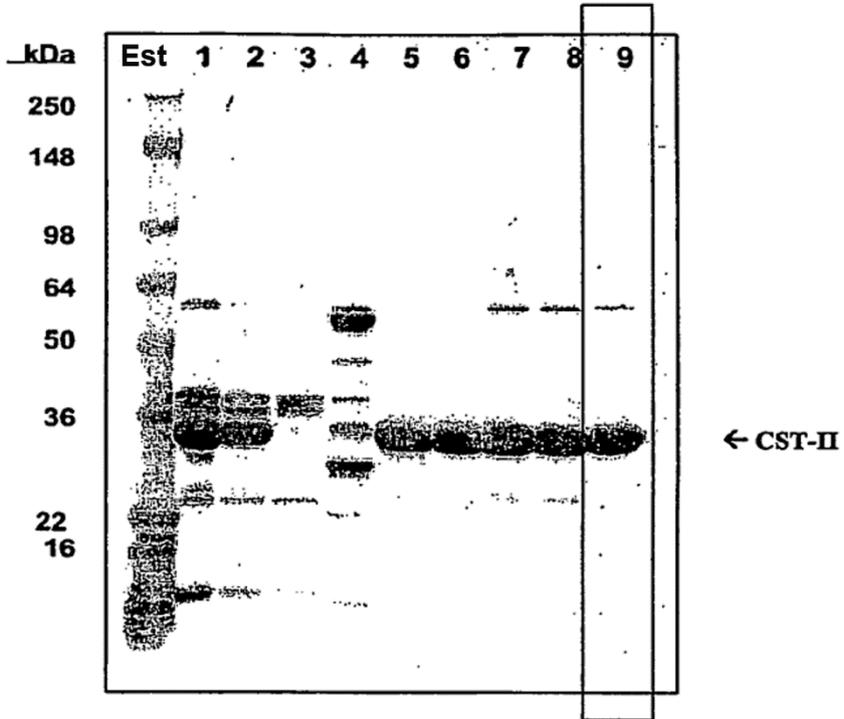


Figura 13

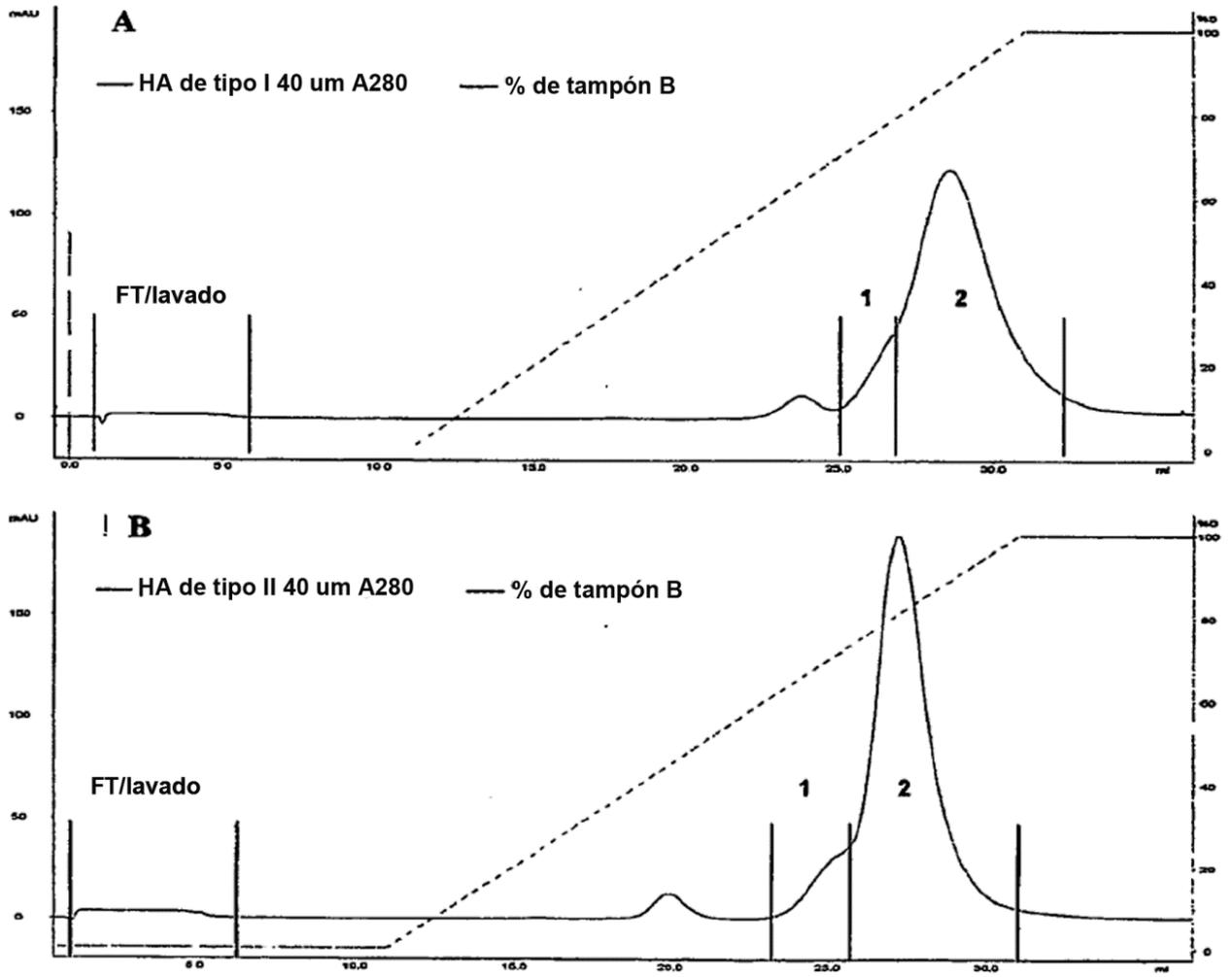


Figura 14

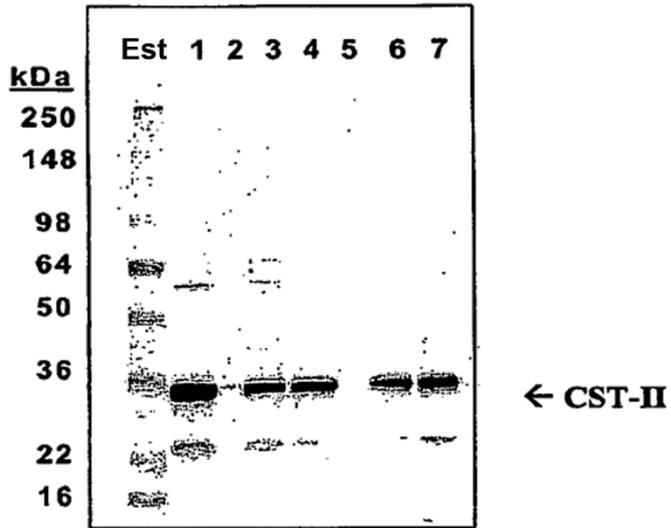


Figura 15

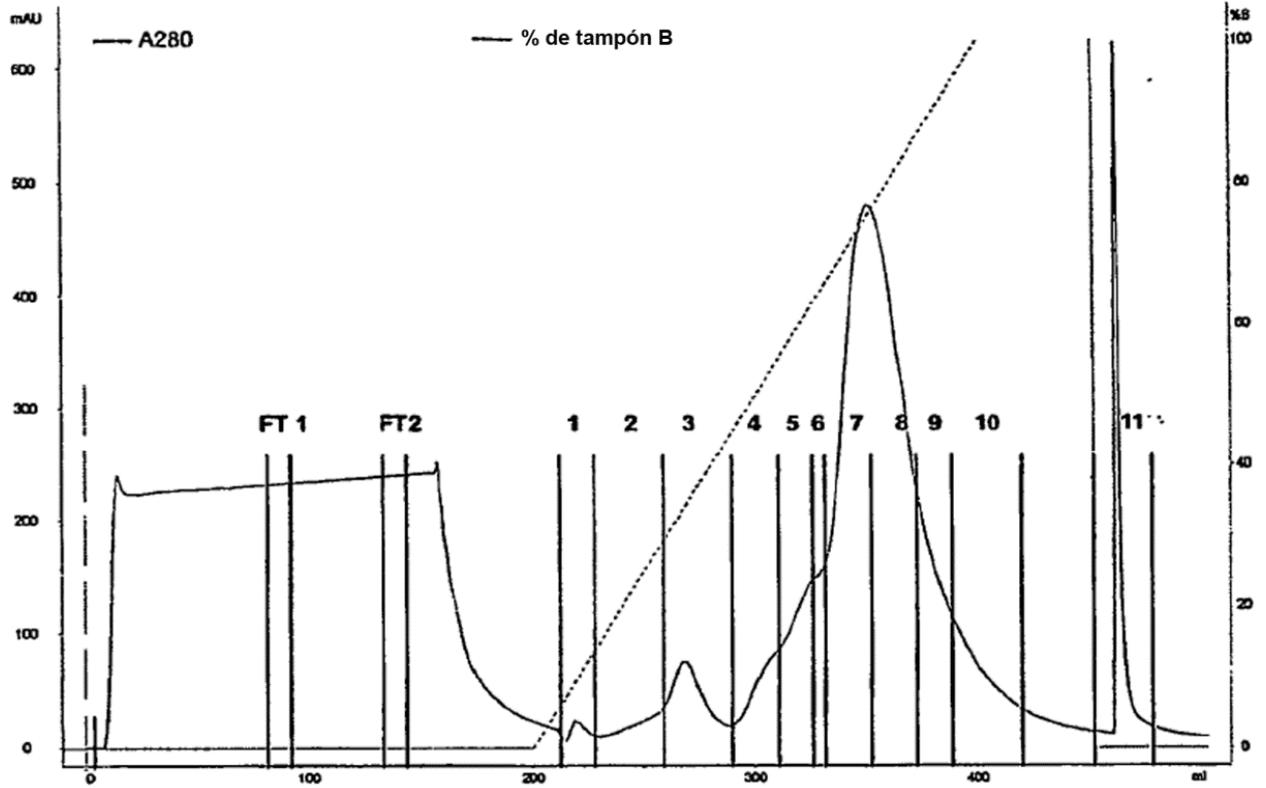


Figura 16

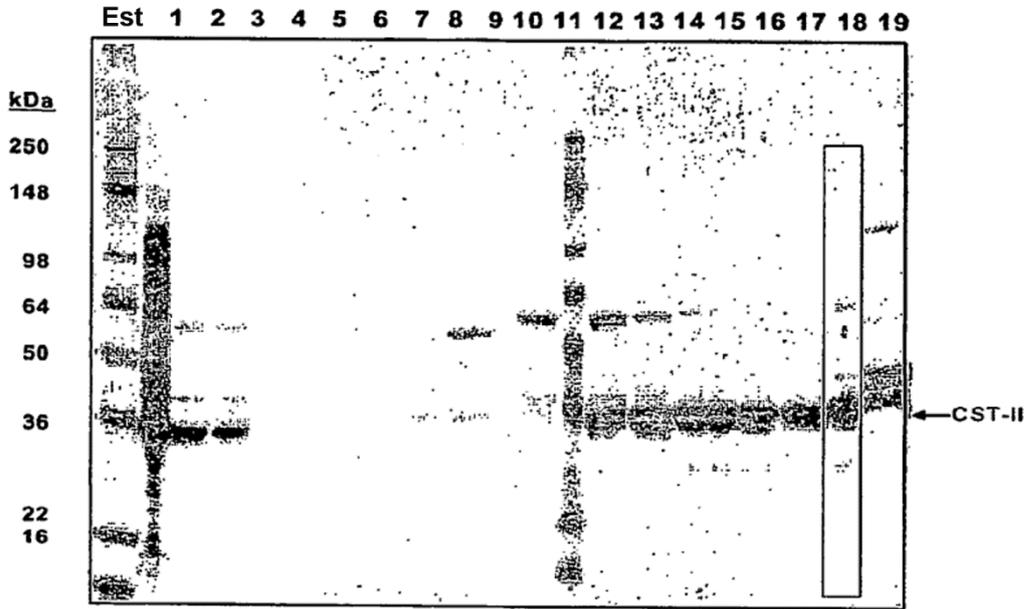


Figura 17

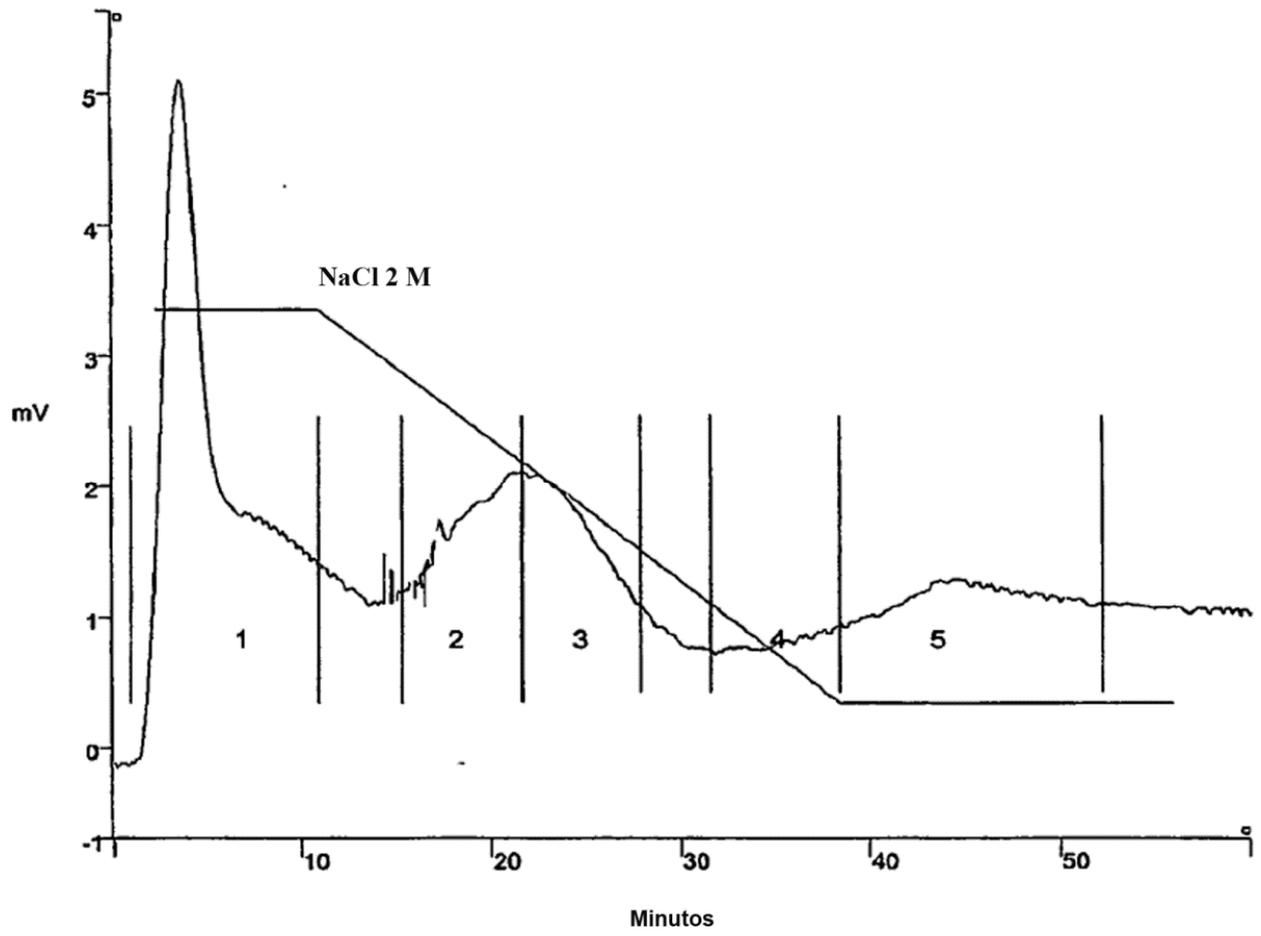


Figura 18

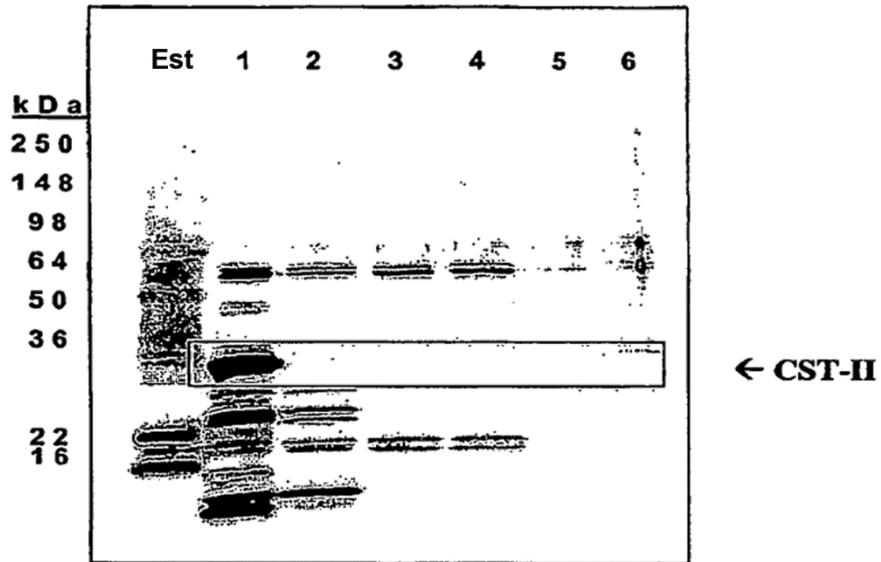


Figura 19

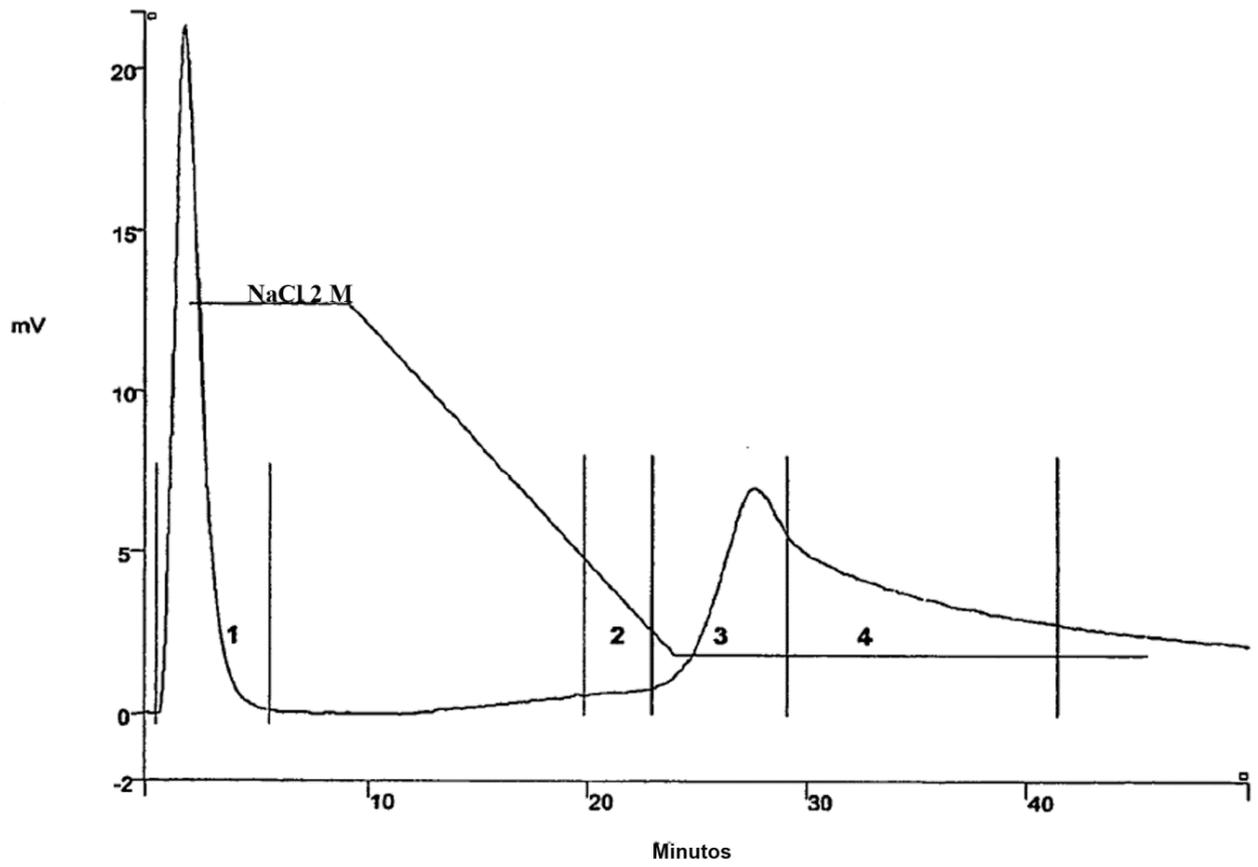


Figura 20

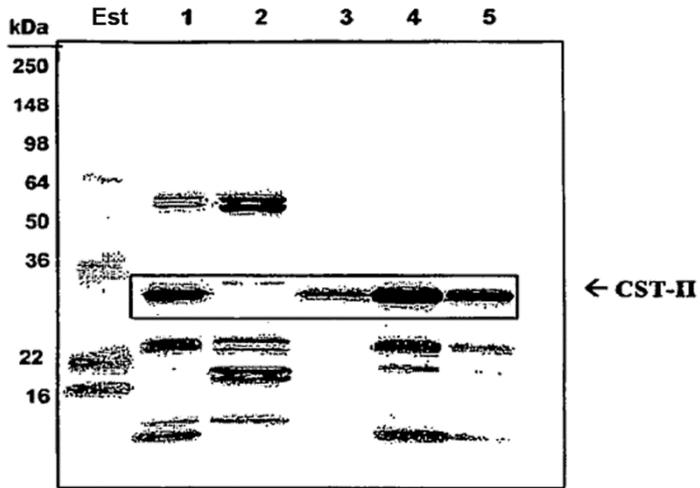


Figura 21

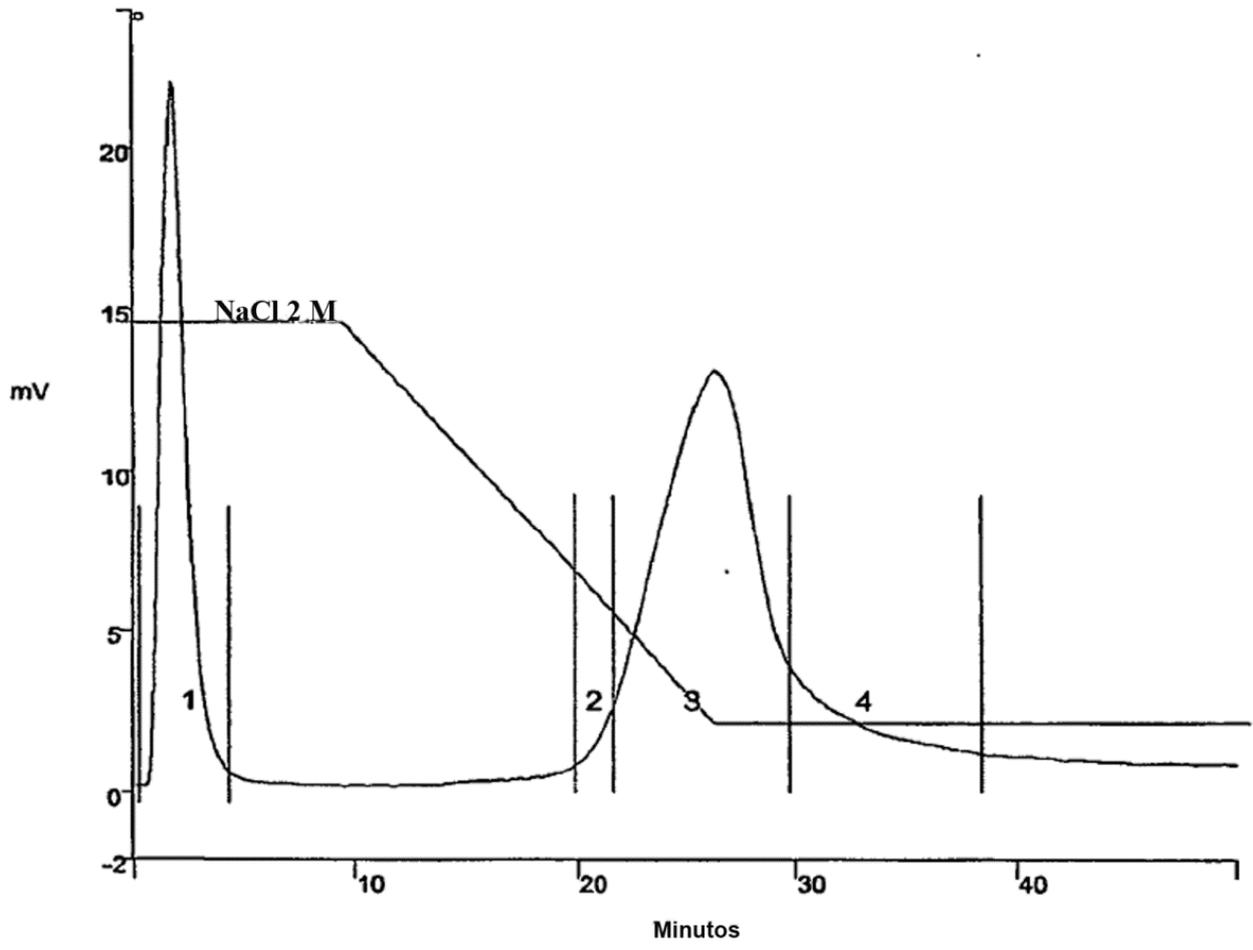


Figura 22

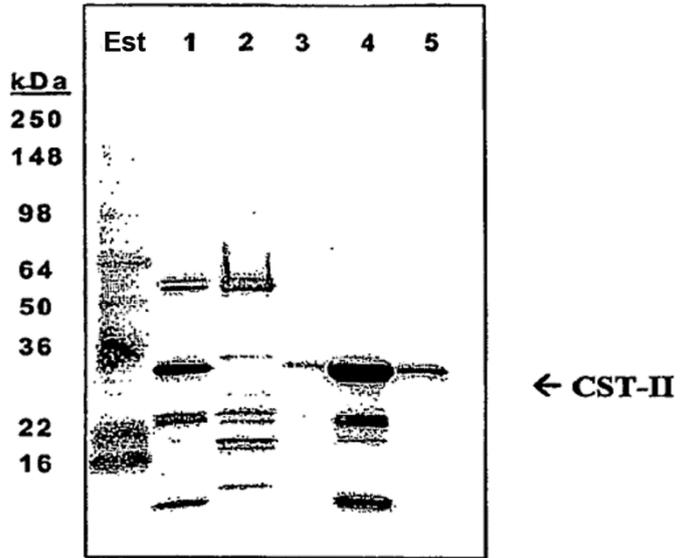


Figura 23

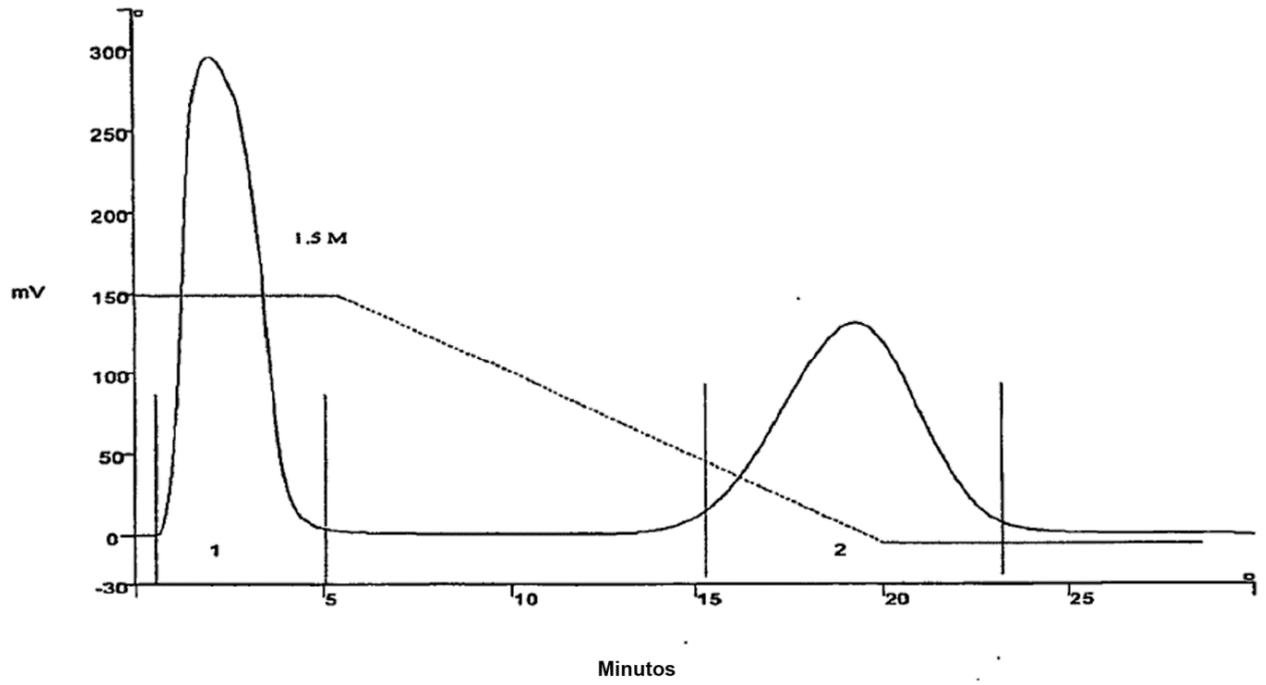


Figura 24

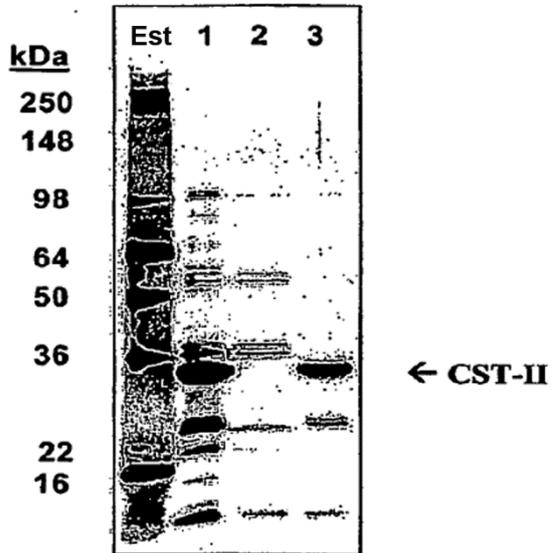


Figura 25

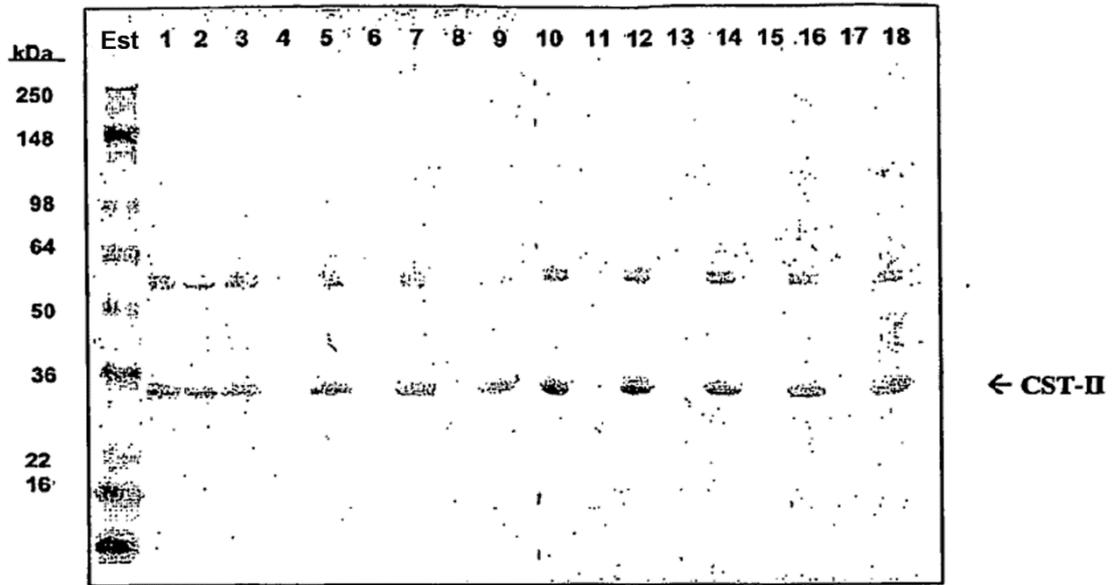


Figura 26

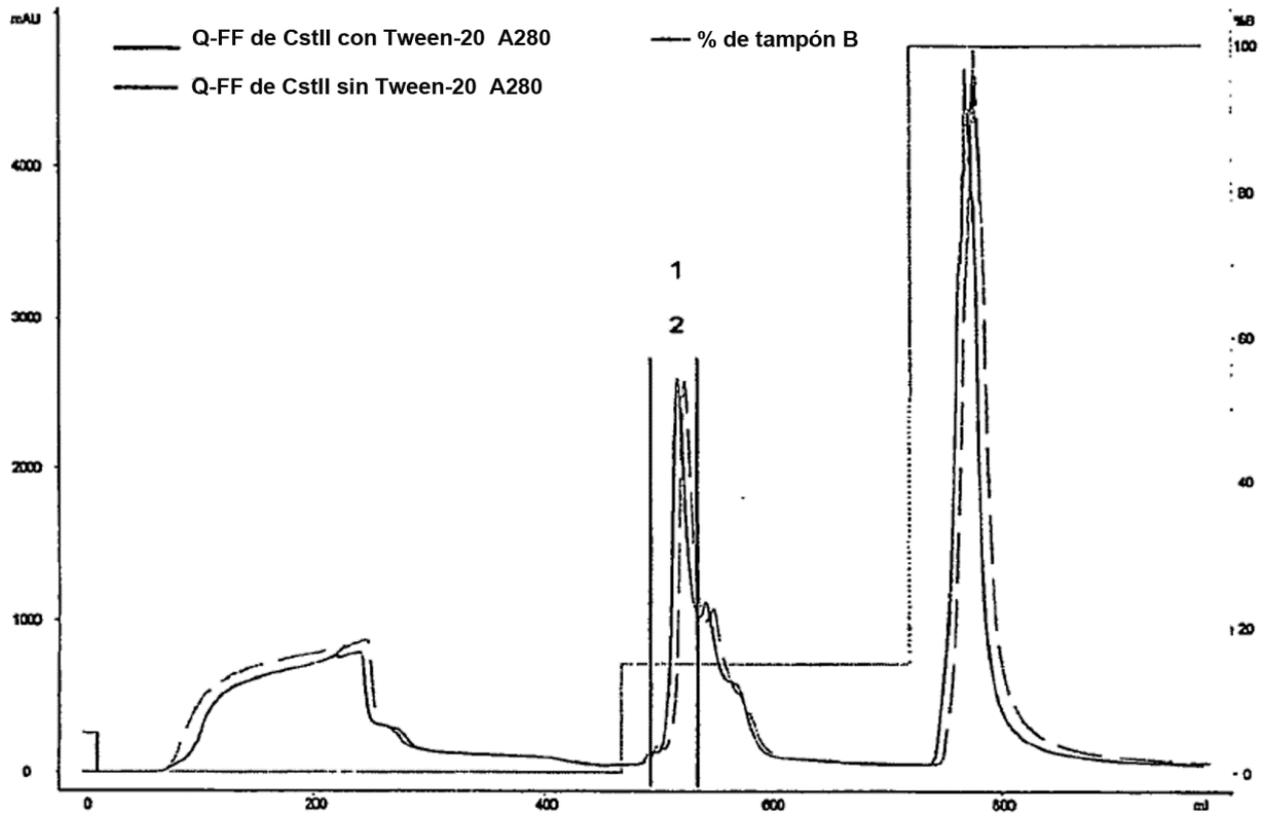


Figura 27

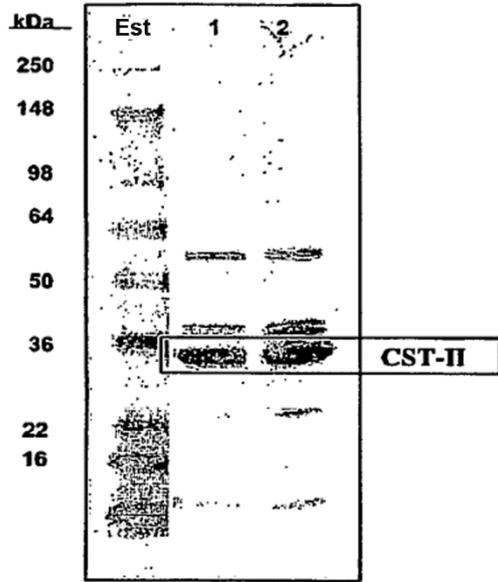


Figura 28

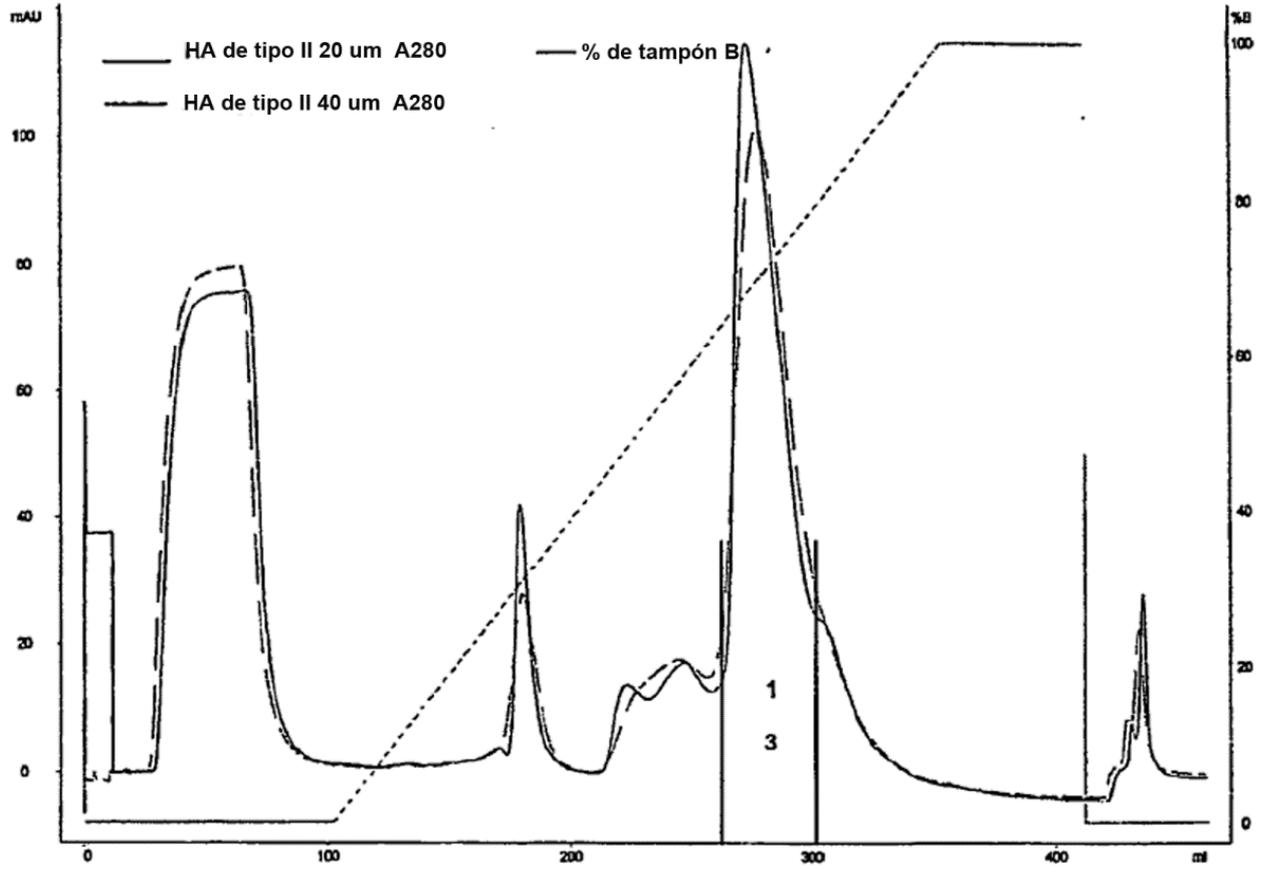


Figura 29

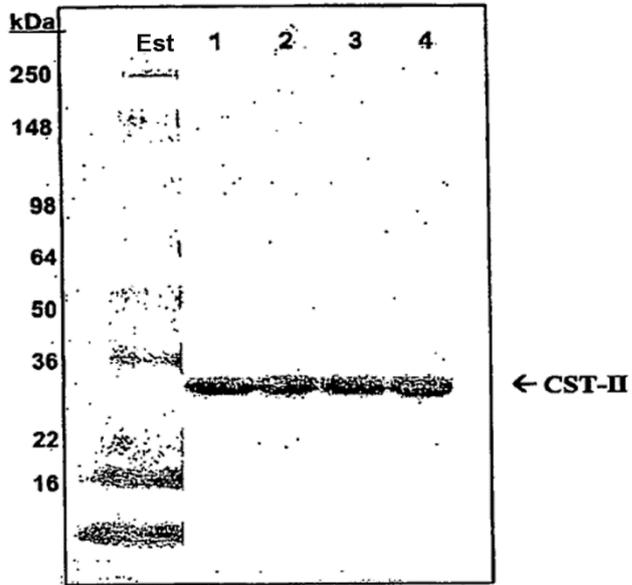


Figura 30

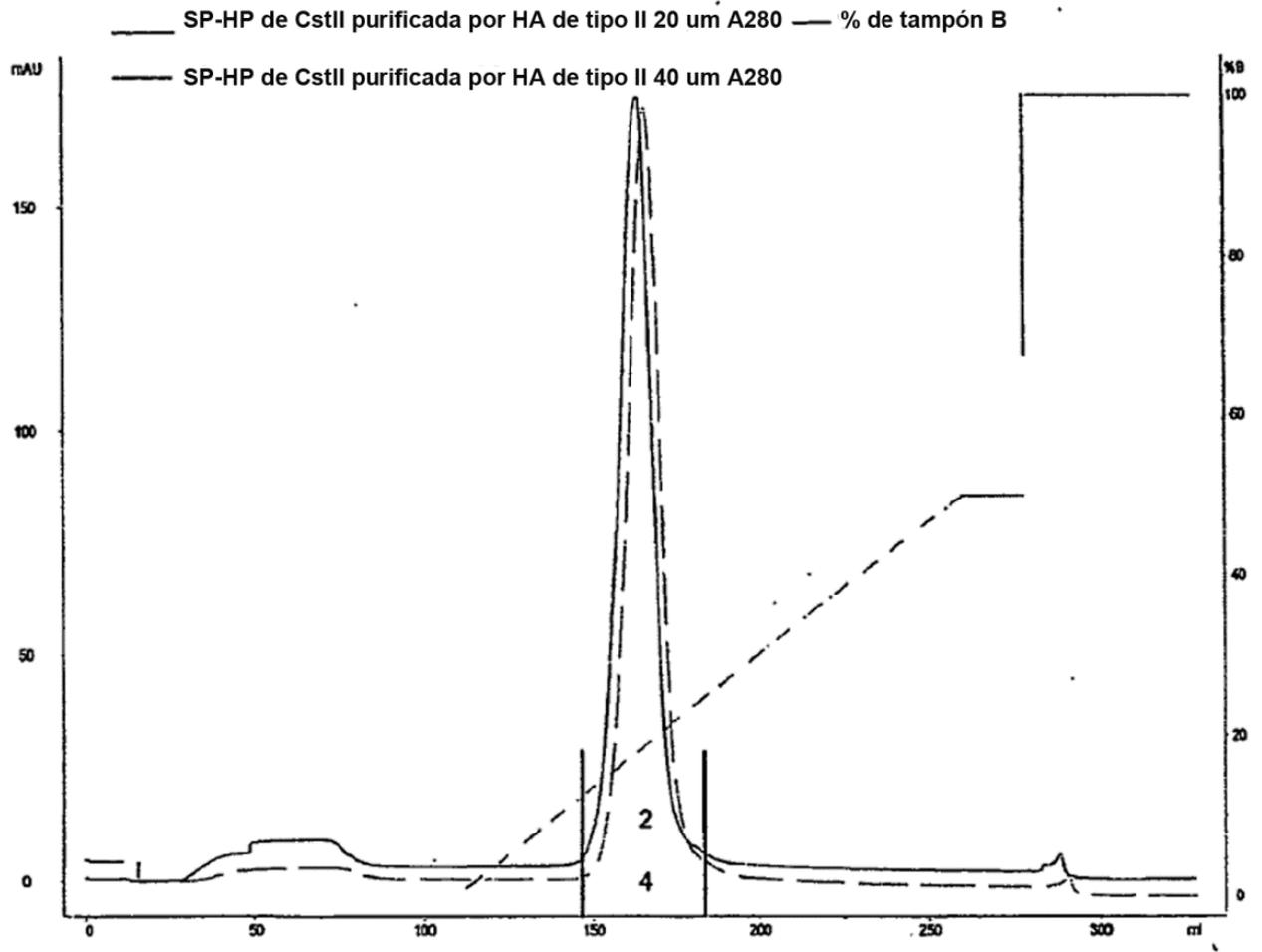


Figura 31

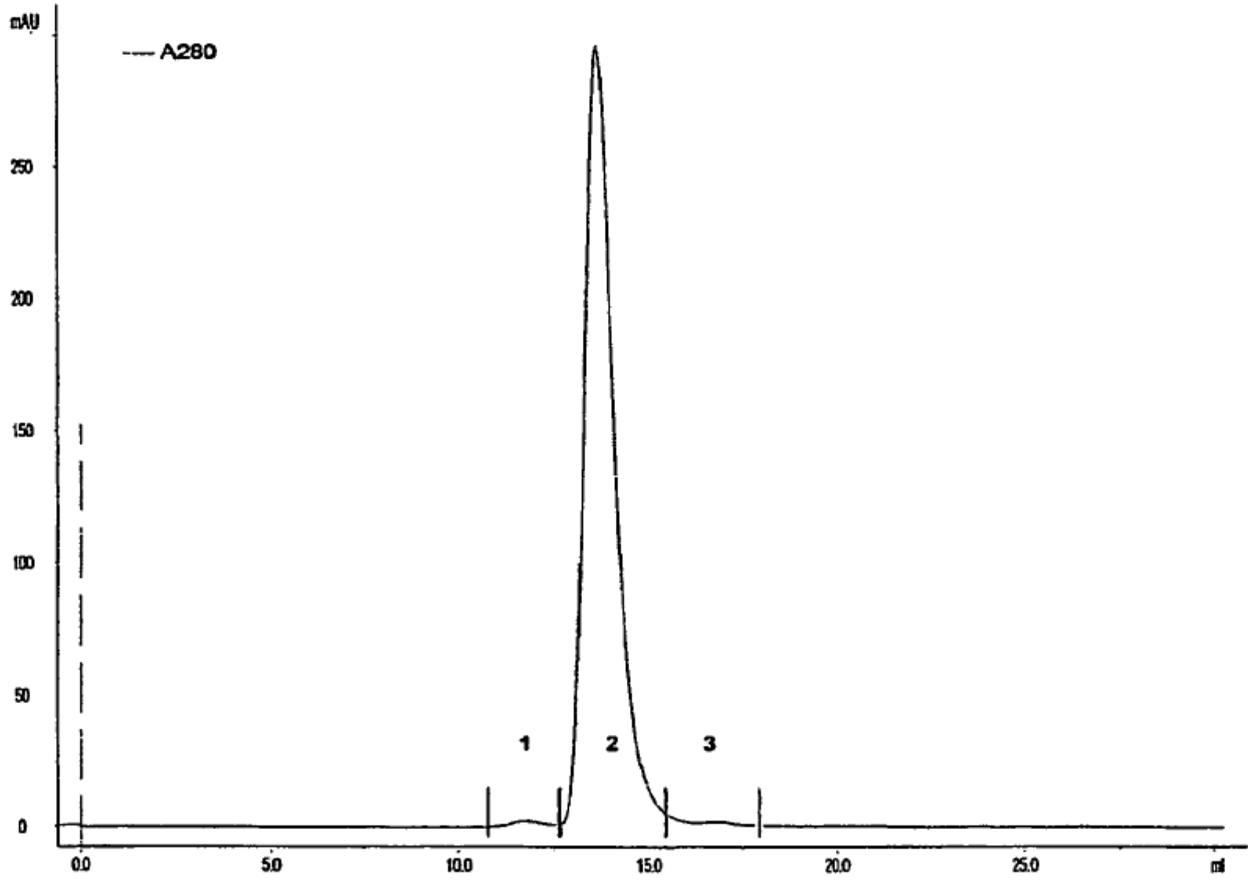


Figura 32

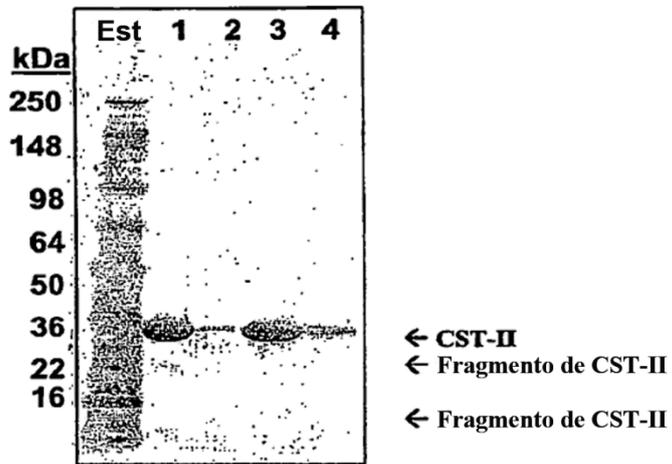


Figura 33

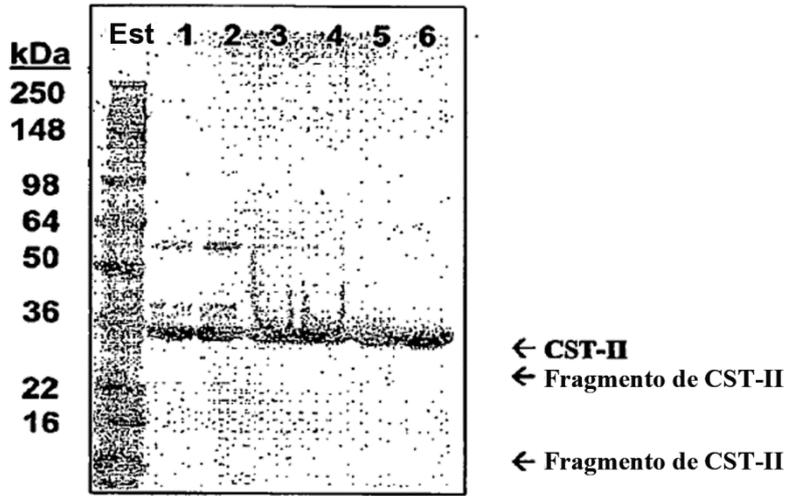


Figura 34

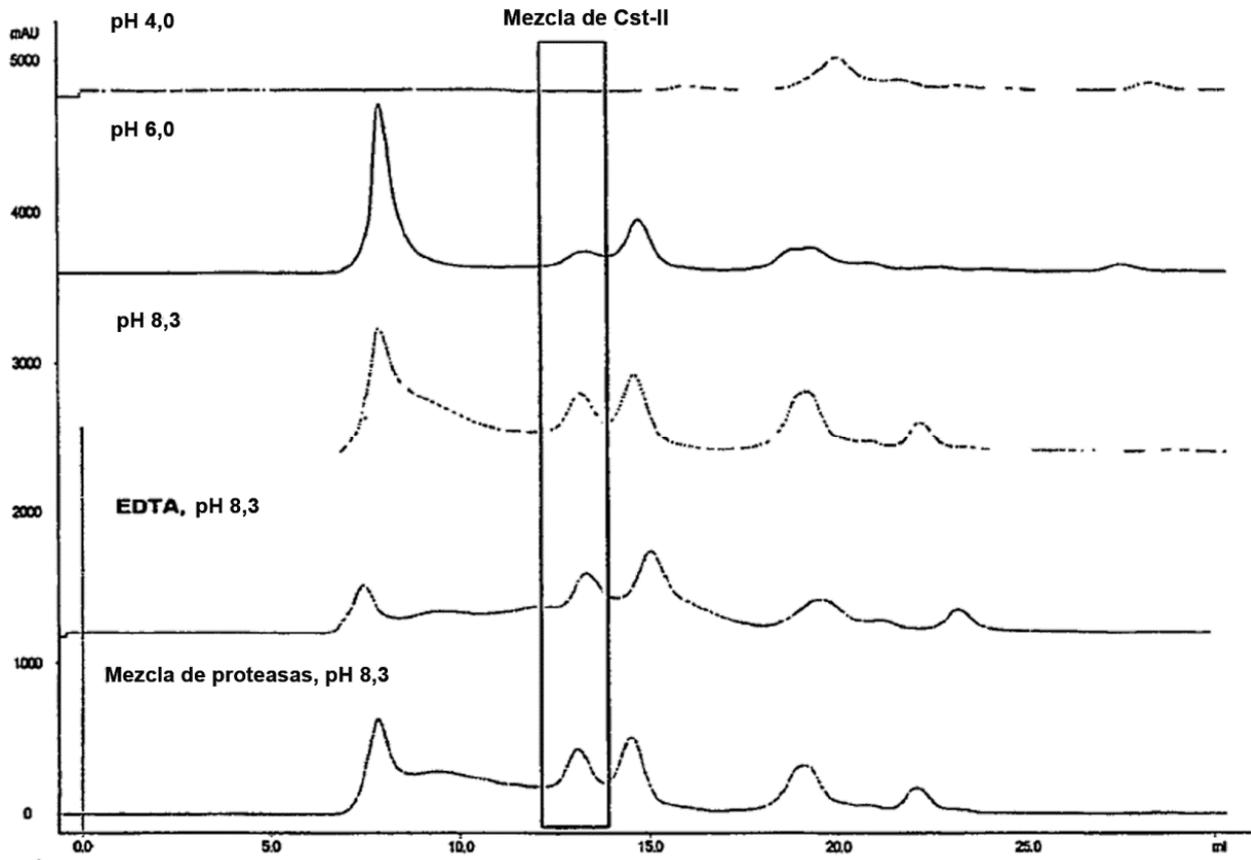


Figura 35

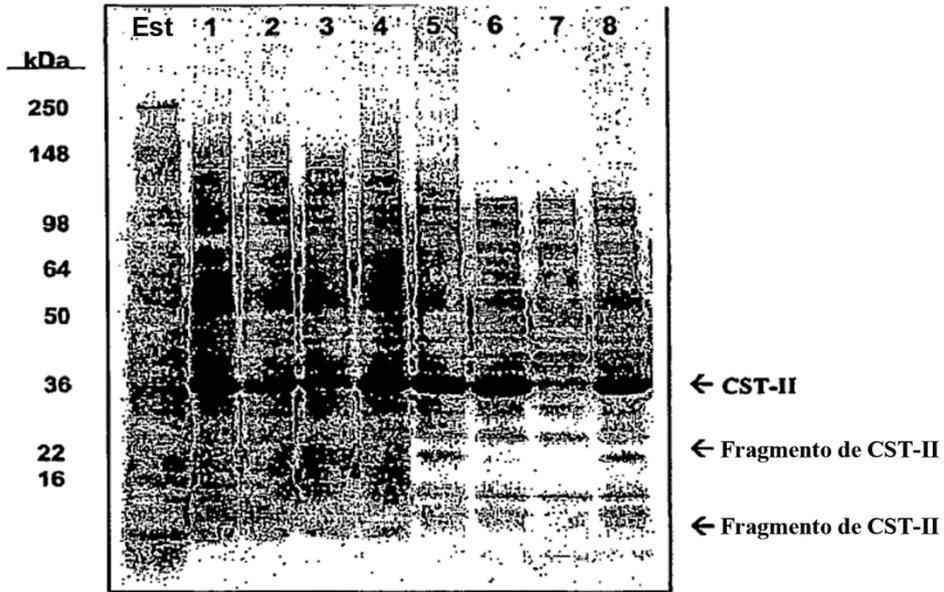


Figura 36

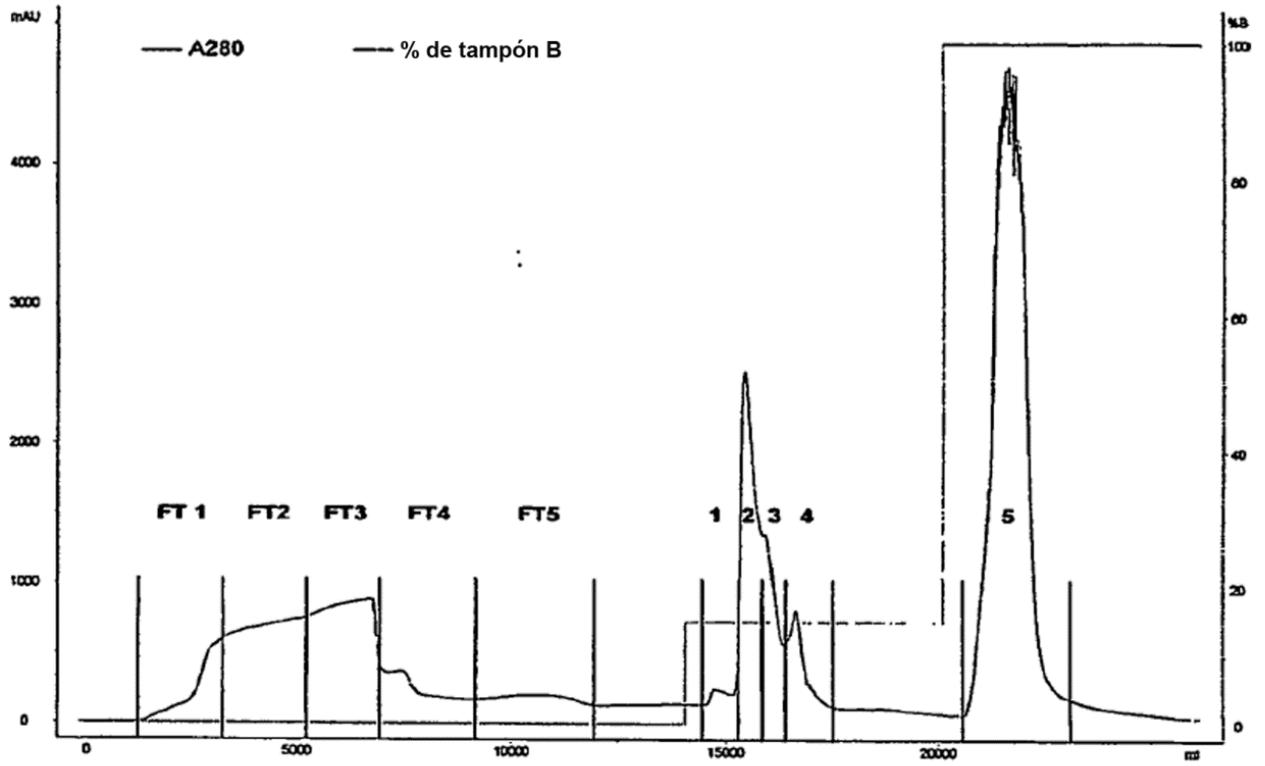


Figura 37

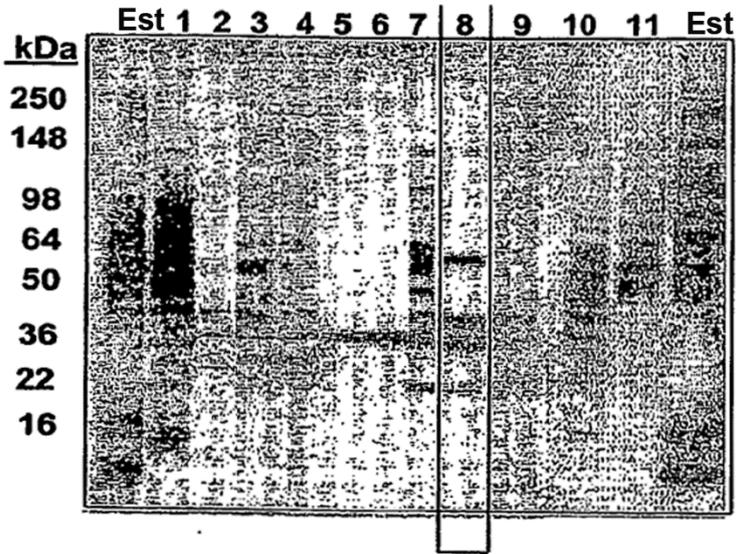


Figura 38

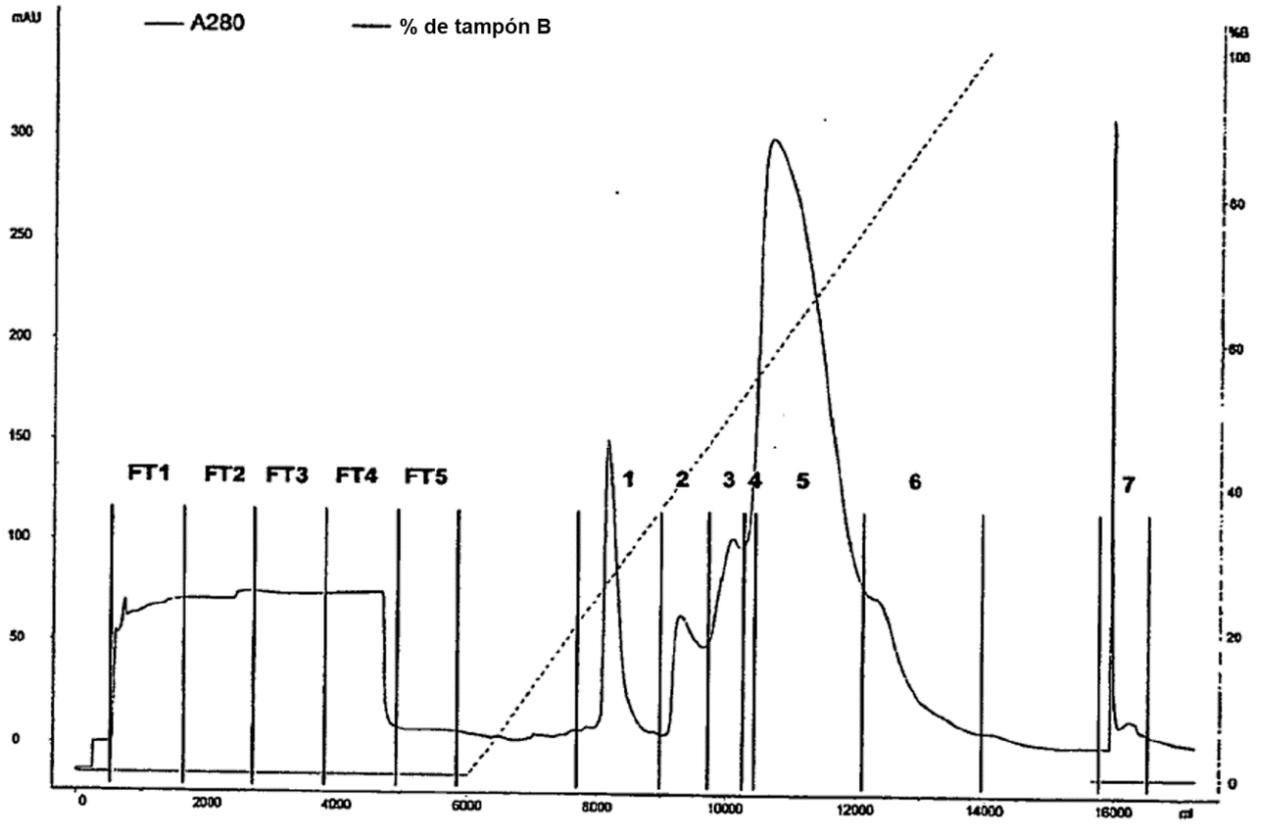


Figura 39

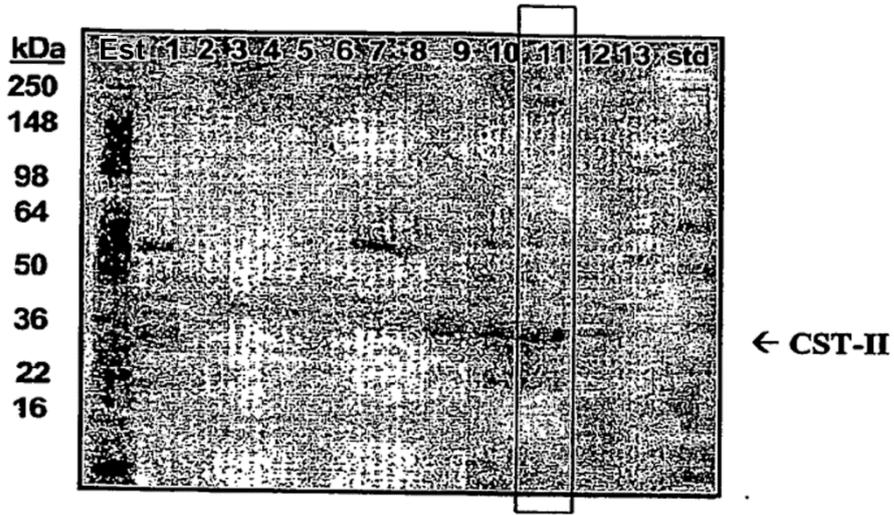


Figura 40

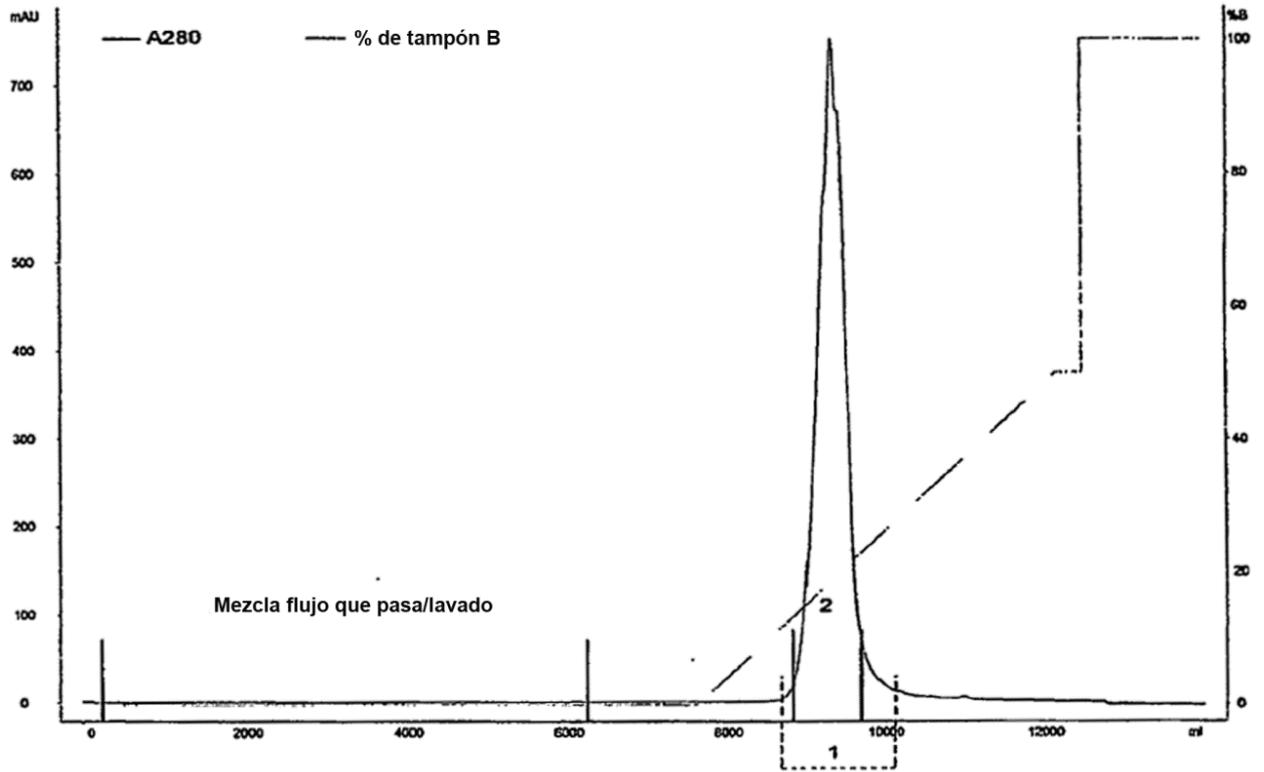


Figura 41

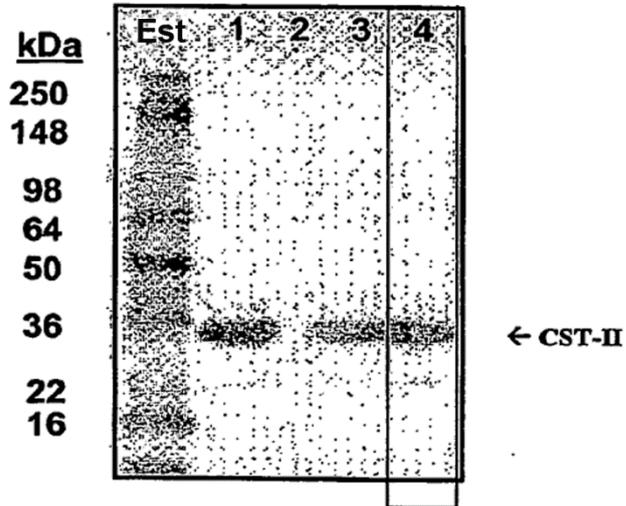


Figura 42

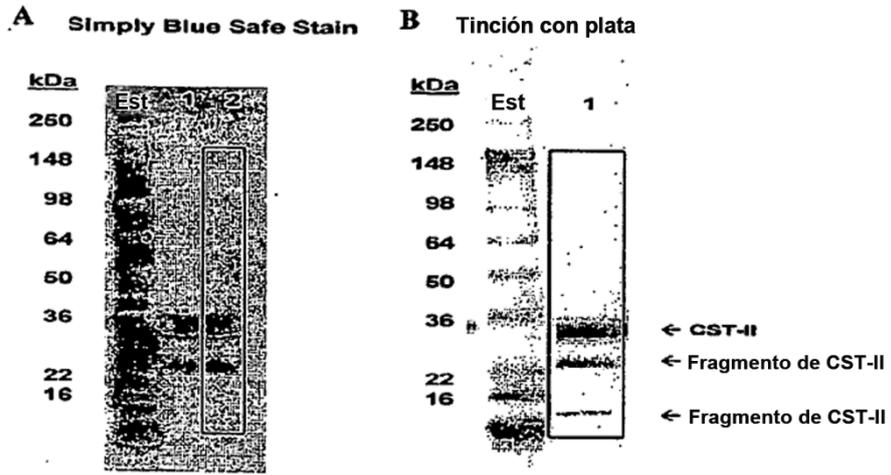


Figura 43

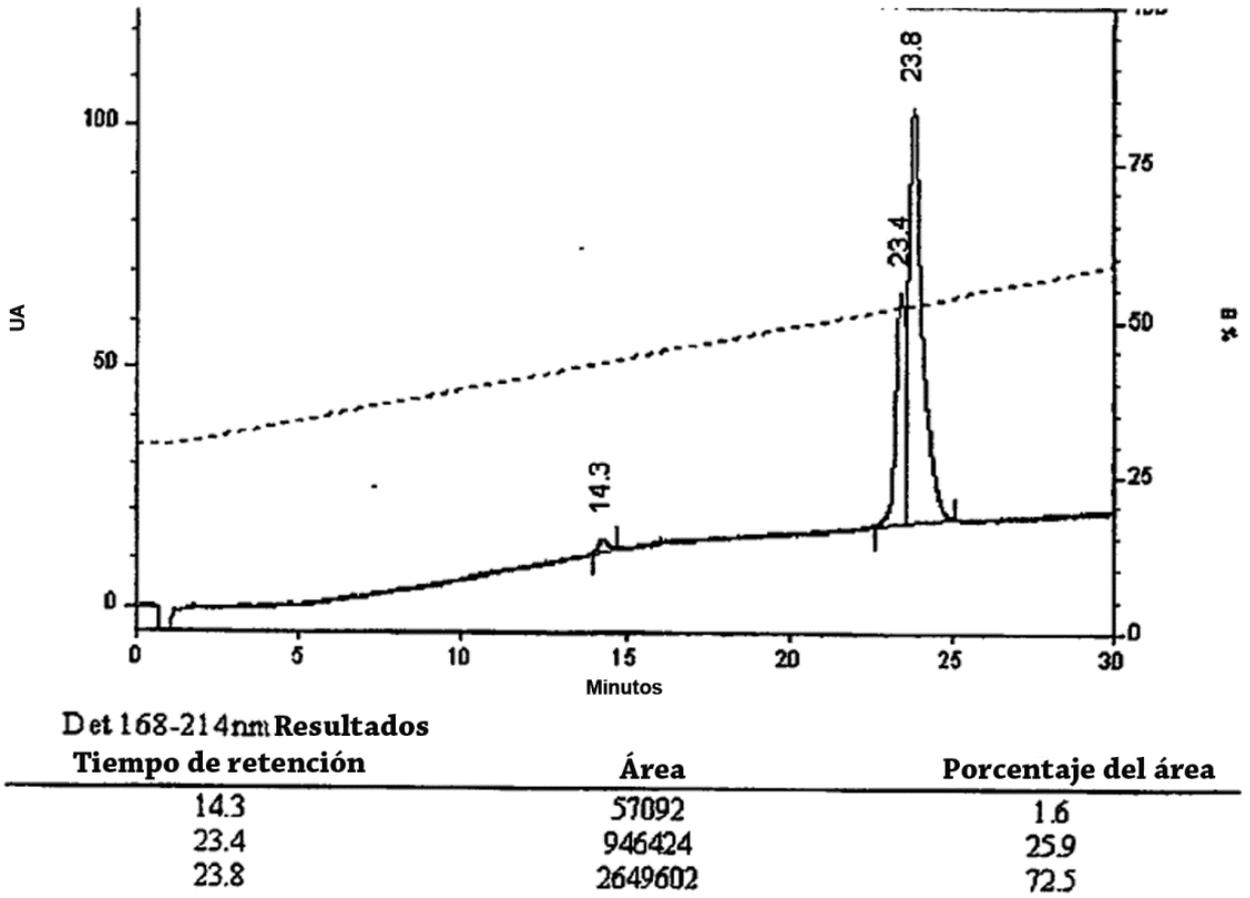


Figura 44