

(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



(11) Número de publicación: **2 460 672**

(51) Int. Cl.:

**C12N 15/82** (2006.01)  
**C12N 15/54** (2006.01)  
**C12N 9/10** (2006.01)  
**C12P 21/02** (2006.01)  
**A01H 5/00** (2006.01)  
**C12N 5/10** (2006.01)  
**C12N 15/11** (2006.01)

(12)

## TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **02.03.2001 E 05108014 (1)**

(97) Fecha y número de publicación de la concesión europea: **16.04.2014 EP 1621633**

---

(54) Título: **Glicoproteínas producidas en plantas con una expresión suprimida de la beta 1,2-xilosiltransferasa**

(30) Prioridad:

**03.03.2000 AT 3552000**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:  
**14.05.2014**

(73) Titular/es:

**GLOSSL, JOSEF (100.0%)**  
ZENTRUM FUR ANGEWANDTE GENETIK  
UNIVERSITAT FUR BODENKULTUR WIEN,  
MUTHGASSE 18  
1190 VIENNA, AT

(72) Inventor/es:

**GLÖSSL, JOSEF;**  
**STEINKELLNER, HERTA;**  
**WILSON, IAIN B;**  
**ALTMANN, FRIEDRICH;**  
**MACH, LUKAS;**  
**MUCHA, JAN y**  
**STRASSER, RICHARD**

(74) Agente/Representante:

**CURELL AGUILÁ, Mireia**

**ES 2 460 672 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

Glicoproteínas producidas en plantas con una expresión suprimida de la beta 1,2-xilosiltransferasa.

- 5 La presente invención se refiere a polinucleótidos que codifican una  $\beta$ 1,2-xilosil-transferasa. Además, la invención se refiere a vectores que comprenden estos polinucleótidos, a las células huésped recombinantes, a las plantas e insectos transfectados con los polinucleótidos o con ADN derivados de los mismos, respectivamente, así como a las glicoproteínas producidas en estos sistemas.
- 10 Las glicoproteínas presentan una variedad y complejidad de unidades de carbohidratos, siendo la composición y disposición de los carbohidratos característica de diferentes microorganismos. Las unidades de oligosacáridos de las glicoproteínas realizan numerosos cometidos, p. ej., son importantes en la regulación del metabolismo, están implicadas en la transmisión de las interacciones célula a célula, determinan los períodos de circulación de las proteínas en circulación y son decisivas en el reconocimiento de los epítopos en las reacciones antígeno-anticuerpo.
- 15 La glicosilación de las glicoproteínas comienza en el retículo endoplasmático (ER), donde los oligosacáridos están unidos a las cadenas laterales de asparagina por enlaces N-glicosídicos o a cadenas laterales de serina o treonina por enlaces O-glicosídicos. Los oligosacáridos unidos por N contienen un núcleo común procedente de una unidad de pentasacárido que consta de tres restos de manosa y de dos de N-acetilglucosamina. Para modificar más las unidades de carbohidratos, las proteínas se transforman desde el ER hasta el complejo de Golgi. La estructura de las unidades de oligosacárido unidas por N de las glicoproteínas está determinada por su conformación y por la composición de las glicosiltransferasas de los compartimentos de Golgi en los que se procesan.
- 20
- 25 Se ha demostrado que la unidad central de pentasacárido de los N-glucanos de algunas plantas está sustituida por xilosa unida a  $\beta$ 1,2 y fucosa unida a  $\alpha$ 1,3 (Lerouge *et al.*, 1998, *Plant Mol. Biol.* 38, 31-48; Rayon *et al.*, 1998, *J. Exp. Bot.* 49, 1463-1472). El heptasacárido "MMXF<sup>3</sup>" constituye el tipo de oligosacárido principal en las plantas (Kurosaka *et al.*, 1991, *J. Biol. Chem.*, 266, 4168-4172; Wilson y Altmann, 1998, *Glycoconj. J.* 15, 1055-1070). Estas estructuras se denominan también N-glucanos complejos o N-glucanos carentes de manosa o truncados, respectivamente. Los restos  $\alpha$ -manosil pueden estar además sustituidos por GlcNAc, a los que se unen la galactosa y la fucosa de modo que se prepara una estructura que corresponde al  $\alpha$ -epítopo humano de Lewis (Melo *et al.*, 1997, *FEBS Lett.* 415, 186-191; Fitchette-Laine *et al.*, 1997, *Plant J.* 12, 1411-1417).
- 30
- 35 En las glicoproteínas de los mamíferos no existe la  $\beta$ 1,2-xilosa ni la fucosa unida a  $\alpha$ 1,3. Se ha descubierto que la  $\beta$ 1,2-xilosa junto con la  $\alpha$ 1,3-fucosa desempeña una función importante en el reconocimiento por el epítopo de los anticuerpos que se dirigen contra los oligosacáridos unidos por N a la planta y por este motivo desencadenan reacciones inmunitarias en cuerpos humanos o de animales contra estos oligosacáridos (Faye *et al.*, 1993, *Anal. Biochem.* 209, 104-108). La  $\beta$ 1,2-xilosa y/o la  $\alpha$ 1,3-fucosa que contienen N-glucanos parecen ser además una de las principales causas de la reactividad alérgica cruzada muy extendida entre varios alérgenos vegetales y de insectos y se denomina también "determinante del carbohidrato reactivo en cruz" (CCD). Debido a la frecuente aparición de reacciones inmunológicas cruzadas, los CCD además enmascaran el diagnóstico de la alergia.
- 40
- 45 Las reacciones inmunológicas desencadenadas en el cuerpo humano por las proteínas vegetales son el principal problema en la utilización medicinal de las proteínas recombinantes humanas producidas en las plantas. Para soslayar este problema, debería evitarse la  $\beta$ 1,2-xilosilación junto con la  $\alpha$ 1,3-fucosilación. Según un estudio, se aisló un mutante de la planta *Arabidopsis thaliana* en el cual se pierde la actividad de la N-acetil-glicosaminil transferasa I, primera enzima en la biosíntesis de los glicanos complejos. Por esta razón se interrumpe la biosíntesis de glicoproteínas complejas. No obstante, estas plantas mutantes son capaces de desarrollarse normalmente en determinadas condiciones (A. Schaewen *et al.*, 1993, *Plant Physiol.* 102, 1109-1118).
- 50
- 55 Para bloquear específicamente la transferencia de la  $\beta$ 1,2-xilosa a un oligosacárido sin interferir también en otras etapas de la glicosilación, únicamente debería inactivarse esta enzima que es directamente responsable de esta glicosilación específica, es decir, la  $\beta$ 1,2-xilosiltransferasa. Esta transferasa que solamente aparece en las plantas y en algunas especies de animales invertebrados, p. ej., en *Schistosoma sp.* (Khoo *et al.*, 1997, *Glycobiology* 7, 663-677) y caracoles (p. Mulder *et al.*, 1995, *Eur. J. Biochem.* 232, 272-283), no estando presente todavía en el hombre o en otros vertebrados, debería ser inactivada a propósito o suprimida para que las proteínas humanas que se producen en las plantas o en las células vegetales, respectivamente, no contengan ya este epítopo desencadenante de la reacción inmunitaria, como es el caso desde hace tiempo.
- 60
- 65 La  $\beta$ 1,2-xilosiltransferasa transfiere la D-xilosa desde la UDP-xilosa a la manosa unida en beta de los oligosacáridos unidos por N a la planta.
- Esta enzima fue purificada a partir de los microsomas de la soja en 1997; Zeng *et al.*: *J. Biol. Chem.*, 272, 31340-31347, 1997). Según este artículo, el mejor receptor para la transferencia de la xilosa fue GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>-T, pero GlcNAc<sub>1</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> con el GlcNAc en el ramal 3, fue también un buen receptor. Además, otros numerosos oligosacáridos unidos por N son poco receptores, especialmente aquellos con unidades de galactosa en el terminal no reductor.

En el artículo de Rayon *et al.* (*Plant Physiology*, 1999, 119, 725-733) se menciona que las proteínas de *Arabidopsis* están N-glicosiladas por N-glucanos del tipo ricos en manosa y por oligosacáridos que contienen xilosa y fucosa. TEZUKA *et al.* (*Eur. J. Biochem.* 203, 401-413 (1992)) midió las actividades de diferentes enzimas, por ejemplo  $\beta$ 1,2-xilosil-transferasa en la fracción de Golgi de las células de sicomoro cultivadas en suspensión. Ellos demostraron que

5 la xilosa fue transferida a la manosa interna por  $\beta$ 1,2-xilosil-transferasa. Además, mencionaron que los oligosacáridos que contenían xilosa están ampliamente distribuidos en todo el reino vegetal si bien los oligosacáridos unidos por N que contenían xilosa se encontraban también en los gastrópodos y en las cloroficeas.

10 Para la supresión o inactivación específica de las proteínas es mejor realizar esto en el ámbito de las etapas de transcripción y traducción, respectivamente. Para esto es necesario aislar y secuenciar el nucleótido y la secuencia de nucleótidos que codifica la proteína activa.

15 Como se mencionó anteriormente, la  $\beta$ 1,2-xilosiltransferasa de soja se aisló y se purificó en 1997. Solamente se ha aislado una parte de la xilosiltransferasa (véase el documento WO 99/29835 A1, SEC. ID. nº: 6 y nº 7), sin embargo, el ADNc completo, que codifica la proteína activa no pudo aislarse y caracterizarse hasta ahora. La razón por la cual la secuencia de nucleótidos no ha sido identificada hasta ahora pudo ser debido a los grandes problemas en el procedimiento debido a la mucha escasez de ARNm que codifica la xilosiltransferasa en los organismos, como por ejemplo la soja. Hasta ahora, aunque varios grupos han tratado de identificar la secuencia de nucleótidos completa de este gen, normalmente de soja, no fue posible producir el ADNc completo que corresponde al ARNm de la xilosiltransferasa con ayuda de métodos convencionales conocidos por ser eficaces en los casos habituales, por ejemplo con ayuda de la ampliación RACE (ampliación rápida de los extremos del ADN). Con este método se amplían secuencias desconocidas con ayuda de cebadores de ampliación específicos. Las razones potenciales de los experimentos de RACE en 5' sin éxito pueden ser una selección inadecuada de los cebadores específicos de PCR así como la presencia de componentes que inhiben la transcriptasa inversa durante la síntesis del ADNc.

20 25 Un problema de la  $\beta$ 1,2-xilosiltransferasa es que su solubilidad y actividad depende de la presencia de detergentes.

30 Además del problema de la concentración sumamente baja del ARNm de la  $\beta$ 1,2-xilosiltransferasa existe asimismo el problema de que la estructura secundaria en el extremo 5' del ARN parece impedir la ampliación de esta zona. En estos casos la ampliación por RACE que es en si misma un método sensible no produce la secuencia correcta y completa del ADNc de xilosiltransferasa.

35 Desde luego es muy probable también que el ARNm y el ADNc derivado de éste, por el hecho de que está presente solamente en muy bajas concentraciones, se recomienda y muta muy fácilmente en el transcurso de varias manipulaciones. Por esta y otras razones potenciales la clonación y la expresión de este gen específico fue imposible hasta ahora.

40 Un objetivo de la presente invención consiste en clonar y secuenciar el gen completo que codifica una  $\beta$ 1,2-xilosiltransferasa y preparar los vectores que comprenden este gen o un ADN alterado o uno de los derivados del mismo ADN para transfectar plantas además de células de las mismas con uno de estos vectores, para producir glicoproteínas que no contienen la  $\beta$ 1,2-xilosa que se encuentra normalmente, además de proporcionar los métodos correspondientes.

45 Es divulgada además la producción de grandes cantidades de enzima recombinante purificada para permitir la síntesis *in vitro* de N-glucanos homogéneos o glicoconjungados que contienen  $\beta$ 1,2-xilosa. Esto facilitará la elucidación posterior del rol de la  $\beta$ 1,2-xilosa en la inmunogenia y alergenia de las glicoproteínas vegetales.

50 55 El objetivo según la invención se alcanza mediante una molécula de ADN que comprende una secuencia según la secuencia SEC. ID. nº: 8 con un marco de lectura abierto desde el par de bases 227 hasta el par de bases 1831 o que es por lo menos homóloga a la secuencia anterior o que se hibrida con la secuencia indicada anteriormente en condiciones severas, o que comprende una secuencia que ha degenerado en la secuencia de ADN anterior debido al código genético, codificando la secuencia una proteína vegetal que presenta actividad de  $\beta$ 1,2-xilosiltransferasa o que es complementaria de las mismas. Esta secuencia completa que nunca ha sido descrita es particularmente útil para experimentos, análisis y procesos productivos que se refieran a la actividad de  $\beta$ 1,2-xilosiltransferasa. Esta secuencia puede ser utilizada especialmente para la inactivación o supresión de la  $\beta$ 1,2-xilosiltransferasa así como para la sobreexpresión y producción de la enzima recombinante.

60 65 En la búsqueda en las bases de datos GenBank-EMBL+DDBJ+PDB que utilizan los péptidos derivados de la xilosiltransferasa de la soja mencionada anteriormente se recuperan varias secuencias de polipéptidos (de *Arabidopsis* y *Drosophila*) con homologías significativas. Sin embargo, estas secuencias no se relacionaron por otra parte entre sí además de con la secuencia de la  $\beta$ 1,2-xilosiltransferasa finalmente identificada. La recuperación lograda de la secuencia candidata para  $\beta$ 1,2-xilosiltransferasa fue posible solamente montando de manera apropiada las tres secuencias peptídicas derivadas de  $\beta$ 1,2-xilosiltransferasa. Todas las estrategias de investigación que utilizaron los solicitantes según el presente estado de la técnica (es decir, utilizando las secuencias peptídicas por separado o en combinación entre sí) no condujeron a una recuperación lograda de la secuencia para  $\beta$ 1,2-xilosiltransferasa.

El aislamiento y purificación de este gen se consiguió buscando en las bases de datos DDBJ+GenBank+EMBL+PBD correspondientes a las secuencias de tres péptidos conocidos (utilizados como péptidos ensamblados) de la xilosiltransferasa de la soja (Patente WO 99/29835 A1, SEC. ID. nº: 3 y nº 5). Se descubrió que una secuencia de ADN de *Arabidopsis thaliana* que no ha sido aún asignada a ninguna proteína anteriormente presentaba homología con dos de los tres péptidos. Con ayuda del programa buscador de genes se encontró la secuencia de la proteína prevista según la cual se diseñan los cebadores específicos de la secuencia para una RT-PCR. Fue posible producir una primera cadena de ADNc correspondiente al ARNm del gen de la xilosiltransferasa de *A. thaliana* tras lo cual la primera cadena de ADNc se sometió a PCR utilizando los cebadores diseñados específicamente. El motivo de producción exitosa del ADNc de la xilosiltransferasa de *A. thaliana* puede ser por una parte que el ARNm de xilosiltransferasa de *A. thaliana* es menos problemático en comparación con otras especies vegetales, por otra parte la PRC se realizó con cebadores específicos para el gen diseñados óptimamente.

El marco de lectura abierto de la SEC. ID. nº: 8 codifica una proteína con 534 aminoácidos y con un peso molecular teórico de 60,2 kDa, estando presente supuestamente un fragmento de transmembrana en la zona entre Ile11 y Phe29. El valor pl calculado de la proteína codificada de la secuencia según la SEC. ID. nº: 9 es 7,52.

La actividad de la  $\beta$ 1,2-xilosiltransferasa de la planta se detecta por un método y se mide, añadiéndose la xilosiltransferasa a una muestra que contiene UDP-xilosa y un receptor marcado (p. ej., un glicopéptido o un oligosacárido marcado). Después del tiempo de reacción, se mide el contenido de xilosa ligada. La actividad de la xilosiltransferasa en este caso se considera positiva si la medida de la actividad es mayor de por lo menos el 10 al 20%, en particular por lo menos del 30 al 50%, que la medida de la actividad de la referencia negativa. Además la estructura del oligosacárido puede verificarse mediante HPLC. Dichos protocolos son de la técnica anterior (Staudacher *et al.*, 1998, *Anal. Biochem.* 246, 96-101; Staudacher *et al.*, 1991, *Eur. J. Biochem.* 199, 745-751). Puede determinarse además si la xilosa está o no ligada al sustrato receptor midiendo la masa del producto mediante espectrometría de masas.

El emparejamiento de las dos moléculas de ADN puede cambiarse seleccionando la temperatura y la fuerza iónica de la muestra. Por condiciones severas, según las condiciones de la invención, se entienden las que permiten una unión exacta y severa. Por ejemplo, las moléculas de ADN se hibridan en dodecilsulfato de sodio (SDS) al 7%, NaPO<sub>4</sub> 0,5 M, pH 7,0, EDTA 1 mM a 50°C y se lavan con SDS al 1% a 42°C.

Si las secuencias presentan por lo menos el 70% de homología con la secuencia nº 8 pueden determinarse p. ej., mediante el programa FastDB de EMBL o el banco de datos SWISSPROT.

Existen numerosas diferencias relevantes entre la  $\beta$ 1,2-xilosiltransferasa codificada por la molécula de ADN de la presente invención y la enzima respectiva de la soja tal como se describe en el documento WO 99/29835 A1:

1. La enzima recombinante es soluble sin detergentes (p. ej., Triton X-100) mientras que la solubilidad de la enzima de la soja depende de la presencia de detergentes.
2. La enzima recombinante es totalmente activa en ausencia de detergentes (p. ej., Triton X-100).
3. La enzima recombinante está N-glicosilada, mientras que la enzima de la soja está descrito que no está glicosilada.
4. La enzima de *A. thaliana* presenta actividad enzimática completa también como forma truncada que carece de los 32 aminoácidos N-terminales.
5. En contraste con la enzima de la soja la enzima de *A. thaliana* tiene un pH óptimo amplio y presenta una pronunciada actividad en el intervalo de pH 6 a 8.
6. La secuencia de ADNc que codifica la enzima de soja corresponde solamente a los aminoácidos (aa) 199 a 469 de la proteína *A. thaliana*, véase la figura 11.
7. La secuencia de ADNc que codifica la xilosiltransferasa de *A. thaliana* contiene dos inserciones correspondientes a los aa 375 a 382 y a los aa 425 a 429 de la secuencia de la proteína prevista) en comparación con la secuencia parcial de la enzima de la soja, véase la figura 11.
8. Ninguno de los cinco péptidos (véase la figura 4 del documento WO 99/29835 A1) aislados de la enzima de la soja es idéntico a los de las correspondientes zonas de *A. thaliana*:

SEC. ID. nº: 1 del péptido: homóloga a la de los aa 411 a 422 de la enzima de *A. thaliana*

SEC. ID. nº: 2 del péptido: homóloga a la de los aa 192 a 205 de la enzima de *A. thaliana*

SEC. ID. nº: 3 del péptido: homóloga a la de los aa 451 a 477 de la enzima de *A. thaliana*

SEC. ID. nº: 4 del péptido: homóloga a la de los aa 191 a 205 de la enzima de *A. thaliana*

5 SEC. ID. nº: 5 del péptido: homóloga a la de los aa 503 a 512 de la enzima de *A. thaliana* (observación: la secuencia de ADNc relacionada en el documento WO 99/29835 A1 no contiene una secuencia de codificación para el péptido 5).

10 Por consiguiente la molécula de ADN según la presente divulgación es particularmente ventajosa ya que codifica una enzima activa recombinante que presenta características y efectos sorprendentemente ventajosos sobre la enzima purificada conocida.

15 Preferentemente, la secuencia de la molécula de ADN codifica una proteína con una actividad de  $\beta$ 1,2-xilosiltransferasa. Esta proteína específica es especialmente útil para el análisis, los experimentos y los métodos de producción que se relacionan con la  $\beta$ 1,2-xilosiltransferasa.

20 Preferentemente, la molécula de ADN es por lo menos el 70%, preferentemente por lo menos el 80%, particularmente preferida por lo menos el 95%, homóloga con la secuencia según la SEC. ID. nº: 8. Esta secuencia codifica una  $\beta$ 1,2-xilosiltransferasa particularmente activa.

25 Según una forma de realización ventajosa adicional, la molécula de ADN comprende 1750 a 1850 pares de bases, en particular 1831.

30 Al actuar de este modo, es particularmente ventajoso si una de las moléculas de ADN indicadas anteriormente está asociada covalentemente con un marcador detectable. Como marcador, puede utilizarse cualquier marcador normal, tales como, p. ej., marcadores fluorescentes, luminiscentes, radioactivos, biotina, etc. De esta manera, se proporcionan reactivos que son adecuados para la detección, selección y cuantificación de las correspondientes moléculas de ADN en muestras de tejido sólidas (p. ej., de plantas) o también en muestras líquidas, mediante métodos de hibridación.

35 Preferentemente, la molécula de ADN incluye una secuencia que comprende una mutación por delección, inserción y/o sustitución. El número de nucleótidos es variable y oscila entre uno solo a varios nucleótidos eliminados, insertados o sustituidos. También es posible que el marco de lectura esté desplazado por la mutación. En dicho "gen transgénico" es meramente importante que se altere la expresión de una  $\beta$ 1,2-xilosiltransferasa y se impida la formación de enzima activa y operativa. Al actuar de este modo, el sitio de la mutación es variable, con tal que se impida la expresión de una proteína enzimáticamente activa. Preferentemente, la mutación en la zona catalítica de la enzima está situada en la zona C-terminal. Los métodos de inserción de mutaciones en las secuencias de ADN son bien conocidos por los expertos en la materia, y por consiguiente se necesita tratar con detalle varias posibilidades de mutagénesis. En este caso pueden emplearse la mutagénesis fortuita así como, en particular, la mutagénesis dirigida, p. ej., la mutagénesis dirigida al sitio, la mutagénesis controlada por nucleótidos o la mutagénesis con ayuda de enzimas de restricción.

40 45 50 55 60 65 La solicitud divulga además una molécula de ADN que codifica una ribozima que comprende dos secciones de secuencias, cada una de las cuales tiene una longitud de por lo menos 15 pares de bases, cada una de las cuales son complementarias de las secciones de la secuencia de una molécula de ADN de la invención descrita anteriormente de modo que la ribozima forma un complejo y escinde el ARNm que es transcrita por una molécula de ADN de  $\beta$ 1,2-xilosiltransferasa natural. La publicación de John M. Burke "Clearing the way for ribozymes" (*Nature Biotechnology* 15:414-415; 1997) se refiere al modo general de funcionamiento de las ribozimas. La ribozima reconocerá el ARNm de la  $\beta$ 1,2-xilosiltransferasa mediante emparejamiento de las bases complementarias con el ARNm. Posteriormente, la ribozima se escindirá y destruirá el ARN de un modo específico para la secuencia, antes de traducir la enzima. Después de la disociación del sustrato escindido, la ribozima se hibrida de forma repetitiva con una molécula de ARN y actúa como endonucleasa específica. En general, las ribozimas pueden producirse específicamente por inactivación de un determinado ARNm, incluso si no se conoce la secuencia de ADN completa que codifica la proteína. Las ribozimas son particularmente eficaces si se desplazan lentamente a lo largo del ARNm. En este caso es más fácil para la ribozima encontrar un sitio libre de ribozima en el ARNm. Por esta razón, los mutantes de ribosomas lentos son también adecuados como sistemas para las ribozimas (J. Burke, 1997, *Nature Biotechnology* 15:414-415).

Una manera posible consiste también en utilizar una forma variada de una ribozima, es decir, una minizima. Las minizimas son eficaces particularmente para la escisión de las moléculas de ARNm mayores. Una minizima es una ribozima de cabeza de martillo que tiene un enlazador de oligonucleótico corto en lugar del tronco/bucle II. Las monozimas diméricas son particularmente eficaces (Kuwabara *et al.*, 1998, *Nature Biotechnology*, 16; 961-965).

60 65 Es divulgado además un vector biológicamente funcional que comprende una de las moléculas de ADN indicadas anteriormente. Para la transfección en las células huésped, es necesario un vector independiente ampliable, en el que, dependiendo de la célula huésped, del mecanismo de transfección, del cometido y del tamaño de la molécula de ADN, puede utilizarse un vector adecuado. Dado que se conoce un gran número de vectores diferentes, una enumeración de los mismos iría más allá de los límites de la presente solicitud y por consiguiente se efectúa sin ella, particularmente desde que los vectores son muy bien conocidos por los expertos en la materia (en cuanto a los vectores así como a

las técnicas y a los términos utilizados en la presente memoria que son conocidos por los expertos en la materia, véase también Maniatis). Teóricamente, el vector presenta una masa molecular pequeña y debería comprender genes seleccionables para que conduzcan a un fenotipo reconocible fácilmente en una célula de este modo capaz de una selección fácil de las células huésped que contienen el vector y de las exentas del vector. Para obtener un alto

5 rendimiento en ADN y en los productos génicos correspondientes, el vector debería comprender un activador potente, además de un potenciador, señales de ampliación génica y secuencias reguladoras. Para una replicación autónoma del vector, además es importante un origen de replicación. Los sitios de adenilación son responsables del tratamiento correcto del ARNm y de las señales de corte y empalme para los transcritos del ARN. Si se utilizan fagos, virus o 10 partículas de virus como vectores, las señales de empaquetamiento controlarán el empaquetamiento del ADN del vector. Por ejemplo, para la transcripción en plantas, son adecuados los plásmidos Ti y para la transcripción en células de insectos, baculovirus y en insectos, respectivamente, trasposones, tal como el elemento P.

Si se inserta el vector de la invención descrito anteriormente en una planta o en una célula vegetal, se consigue una supresión tras la transcripción de la expresión génica del gen endógeno de la  $\beta$ 1,2-xilosiltransferasa mediante la transcripción de un homólogo transgénico a éste o de partes de éste, en la orientación codificante (*sense*). Para esta técnica codificante, además, se hace referencia a las publicaciones de Baucombe 1996, *Plant. Mol. Biol.*, 9:373-382 y Brigneti *et al.*, 1998, *EMBO J.* 17:6739-6746. Esta estrategia de "censura génica" es un modo eficaz de suprimir la expresión del gen de  $\beta$ 1,2-xilosiltransferasa, véase también Waterhouse *et al.*, 1998, *Proc. Natl. Acad. Sci. USA*, 95:13959-13964.

20 Además, la solicitud divulga un vector biológicamente funcional que comprende una molécula de ADN según una de las formas de realización descritas anteriormente, que está orientado en sentido inverso con respecto al activador. Si este vector se transfecta en una célula huésped, se leerá un "ARNm no codificante (*antisense*)" que es complementario del ARNm de la  $\beta$ 1,2-xilosiltransferasa y forma un complejo con esta última. Este enlace impedirá el 25 tratamiento correcto, el transporte, la estabilidad o, impidiendo la hibridación del ribosoma, impedirá la traducción y de este modo la expresión génica normal de la  $\beta$ 1,2-xilosiltransferasa.

Aunque la secuencia completa de la molécula de ADN podría insertarse en el vector, las secuencias parciales de ésta 30 debido a su menor tamaño pueden presentar ventajas para determinados fines. En el aspecto no codificante, p. ej., es importante que la molécula de ADN sea lo bastante grande para formar un ARNm de cadena complementaria suficientemente grande que se una al ARNm de la transferasa. Una molécula de ARN no codificante comprende, p. ej., de 50 a 200 nucleótidos ya que muchas de las moléculas de ARNm no codificantes naturales conocidas comprende aproximadamente 100 nucleótidos.

35 Para una inhibición particularmente eficaz de la expresión de una  $\beta$ 1,2-xilosiltransferasa activa, es adecuada una combinación de la técnica codificante y de la técnica no codificante (Waterhouse *et al.*, 1998, *Proc. Natl. Acad. Sci. USA*, 95:13959-13964).

40 Se utilizan con ventaja las moléculas de ARN que se hibridan rápidamente. La eficacia de las moléculas de ARN no codificantes que tienen un tamaño de más de 50 nucleótidos dependerá de la cinética de hibridación *in vitro*. Por lo tanto, p. ej., las moléculas de ARN no codificante que se hibridan rápidamente presentan una inhibición mayor de la expresión de las proteínas que las moléculas de ARN que se hibridan lentamente (Wagner *et al.*, 1994, *Annu. Rev. Microbiol.*, 48:713-742; Rittner *et al.*, 1993, *Nucl. Acids Res.*, 21: 1381-1387). Dichas moléculas de ARN no codificante 45 que se hibridan rápidamente comprenden particularmente un gran número de bases externas (terminales libres y secuencias de conexión), un gran número de subdominios estructurales (componentes) así como un grado bajo de bucles (Patzel *et al.* 1998; *Nature Biotechnology*, 16; 64-68). Las estructuras secundarias hipotéticas de la molécula de ARN no codificante pueden determinarse, p. ej., con ayuda de un programa informático, según el cual se selecciona una secuencia adecuada de ADN del ARN no codificante.

50 En el vector pueden insertarse diferentes zonas de la secuencia de la molécula de ADN. Una posibilidad consiste, p. ej., en insertar en el vector solamente una parte que es responsable de la hibridación del ribosoma. El bloqueo en esta zona del ARNm será suficiente para interrumpir la traducción completa. Una eficacia particularmente grande de las moléculas no codificantes se produce también en las zonas 5' y 3' no traducidas del gen.

55 La solicitud divulga asimismo un vector biológicamente funcional que comprende una de las dos moléculas de ADN mencionadas anteriormente (mutación o molécula de ADN del ribozima). Lo que se ha dicho anteriormente respecto a los vectores es aplicable también en este caso.

60 Es divulgado asimismo un método de preparación de un ADNc que comprende la molécula de ADN de la invención, en la que el ARN se aísla de una célula vegetal, en particular de las células de la hoja, mediante el cual se realiza una transcripción inversa tras la adición de una transcriptasa inversa y cebadores. Cada una de las etapas de este método se realiza según los protocolos conocidos por sí mismos. Para la transcripción inversa, por una parte, es posible producir el ADNc del ARNm entero con ayuda de oligo(dT) cebadores y solamente entonces realizar una PCR mediante cebadores seleccionados a fin de preparar las moléculas de ADN que comprenden el gen de  $\beta$ 1,2-xilosiltransferasa. Por otra parte, los cebadores seleccionados pueden utilizarse directamente para la transcripción 65 inversa a fin de obtener el ADNc corto y específico. Pueden prepararse cebadores adecuados p. ej., por síntesis según

el modelo de secuencias de ADNc de la transferasa.

La solicitud divulga además un método de clonación de una  $\beta$ 1,2-xilosiltransferasa, caracterizado porque la molécula de ADN de la invención se clona en un vector que posteriormente se transfecta en una célula huésped o en un anfitrión, respectivamente, en el que, por selección y ampliación de las células huésped transfectadas, se obtienen 5 líneas celulares transfectadas que expresan la  $\beta$ 1,2-xilosiltransferasa activa. La molécula de ADN se inserta en el vector con ayuda, p. ej., de las endonucleasas de restricción. Para el vector, es aplicable lo dicho anteriormente. Lo que es importante en este método es que se selecciona un sistema eficaz anfitrión-vector. Para obtener una enzima 10 activa, son particularmente adecuadas las células huésped eucarióticas. Una manera posible consiste en transfectar el vector en las células de un insecto. Al actuar así, como vector debería utilizarse virus de insectos, tal como, p. ej., baculovirus.

Desde luego, pueden también transfectarse, plantas o células vegetales, células humanas o de otros vertebrados, en 15 cuyo caso estas últimas expresarían una enzima extraña a ellas.

Preferentemente, se proporciona un método de preparación de células huésped recombinantes, en particular células 20 vegetales o plantas, respectivamente, con una producción de  $\beta$ 1,2-xilosiltransferasa suprimida o completamente interrumpida, que se caracteriza porque, por lo menos, uno de los vectores según la invención, es decir, el que comprende la molécula de ADN de la invención, la molécula de ADN mutante o la molécula de ADN que codifica ribozimas o el que comprende la molécula de ADN en orientación inversa al activador, se inserta en la célula huésped o en la planta, respectivamente. Lo que se ha dicho anteriormente para la transfección también es aplicable a este caso.

Como células huésped pueden utilizarse, p. ej., células vegetales, en las que, p. ej., es elegible el plásmido Ti con el 25 sistema de *Agrobacterium*. Con el sistema de *Agrobacterium* es posible transfectar una planta directamente: las agrobacterias producen agallas en el tronco de la raíz en los vegetales. Si las agrobacterias infectan una planta dañada, las propias bacterias no penetran en la planta, pero insertan la fracción de ADN recombinante, el denominado T-ADN, del plásmido Ti que produce tumor anular y extracromosómico en las células vegetales. El T-ADN, y por lo tanto también la molécula de ADN insertada en éste, se instalan en el ADN cromosómico de la célula de manera 30 estable de modo que los genes del T-ADN se expresarán en la planta.

Existen numerosos mecanismos conocidos de transfección eficaz para sistemas huésped diferentes. Algunos ejemplos 35 son la electroporación, el método del fosfato de calcio, la microinyección y el método de los liposomas.

Posteriormente, se seleccionan las células transfectadas, p. ej., basándose en las resistencias a los antibióticos para 40 las cuales el vector comprende genes u otros genes marcadores. A continuación se amplían las líneas celulares transfectadas, en pequeñas cantidades, p. ej., en placas Petri, o en grandes cantidades, p. ej., en fermentadores. Además, las plantas presentan una característica particular, es decir, son capaces de volver a desarrollarse a partir de una célula (transfectada) o de un protoplasto, respectivamente, en una planta completa que puede desarrollarse.

Dependiendo del vector utilizado, tendrán lugar procesos en el anfitrión de modo que se suprimirá o se bloqueará 45 completamente la expresión enzimática:

Si se transfecta el vector que comprende la molécula de ADN con la mutación por delección, inserción o sustitución, se 50 producirá una recombinación homóloga: la molécula de ADN mutante reconocerá la secuencia idéntica en el genoma de la célula huésped a pesar de su mutación y se insertará exactamente en ese lugar de modo que se forme un "gen trangénico". De esta manera, se introduce una mutación en el gen para la  $\beta$ 1,2-xilosiltransferasa que es capaz de 55 inhibir la expresión de la  $\beta$ 1,2-xilosiltransferasa sin defectos. Como se ha explicado anteriormente, en esta técnica es importante que la mutación sea suficiente para bloquear la expresión de la proteína activa. Tras la selección y ampliación, el gen puede secuenciarse como comprobación adicional a fin de determinar el éxito de la recombinación homóloga o el grado de mutación, respectivamente.

Si se transfecta el vector que comprende la molécula de ADN que codifica una ribozima, la ribozima activa se 60 expresará en la célula huésped. La ribozima forma un complejo con la secuencia del ARNm complementario de la  $\beta$ 1,2-xilosiltransferasa por lo menos en un determinado sitio, escinde este sitio, y de este modo puede inhibir la traducción de la enzima. En esta célula huésped así como en las líneas celulares, u opcionalmente, en las plantas, respectivamente, derivadas de éstas, no se expresará la  $\beta$ 1,2-xilosiltransferasa.

En caso de que el vector comprenda la molécula de ADN de la invención en la dirección codificante o reversa al 65 activador, un ARNm codificante o no codificante se expresará en la célula transfectada (o en la planta, respectivamente). El ARNm no codificante es complementario por lo menos con una parte de la secuencia de ARNm de la  $\beta$ 1,2-xilosiltransferasa y puede inhibir asimismo la traducción de la enzima. Como ejemplo de un método de supresión de la expresión de un gen por la técnica no codificante, se hace referencia a la publicación de Smith *et al.*, 1990, *Mol. Gen. Genet.* 224:477-481, en la que en esta publicación se inhibe la expresión del gen implicado en el proceso de maduración de los tomates. Se ha demostrado recientemente que el ARN de doble cadena (ARNds) activa la censura del gen específico de la secuencia en una amplia variedad de organismos, incluyendo los nemátodos,

vegetales, tripanosomas, moscas de la fruta y planaria; se ha demostrado que un activador de ARN todavía no caracterizado produce la metilación del ADN en varios sistemas vegetales diferentes que conducen a la interferencia selectiva con la función génica (para reseña véase Fire A., 1999, *Trends Genet.* 15 (9): 358-363).

- 5 En todos los sistemas, la expresión de la  $\beta$ 1,2-xilosiltransferasa está por lo menos suprimida, preferentemente incluso completamente bloqueada. El grado de alteración de la expresión génica dependerá del grado de acomplejamiento de la recombinación homóloga, en las posibles mutaciones coincidentes ulteriores y en otros procedimientos en la zona del genoma. Se comprueba la actividad de  $\beta$ 1,2-xilosiltransferasa en las células transfectadas y se seleccionan.
- 10 Además, es posible aumentar todavía más la supresión de la expresión de la  $\beta$ 1,2-xilosiltransferasa descrita anteriormente introduciendo en el anfitrión un vector que comprende un gen que codifica una proteína de un mamífero, p. ej.,  $\beta$ 1,4-galactosiltransferasa, además de la inserción de un vector descrito anteriormente. La xilosilación puede reducirse mediante la actuación de otras enzimas de mamífero, siendo particularmente eficaz la combinación de la inhibición de la expresión de una  $\beta$ 1,2-xilosiltransferasa mediante el vector de la invención y por medio de un vector de la enzima de mamífero.

Para la transfección puede utilizarse cualquier tipo de planta, p. ej., judía mung, planta de tabaco, planta de tomate y/o patata.

- 20 Otro método ventajoso de producción de células huésped recombinantes, en particular de células vegetales o de plantas, respectivamente, consiste en que la molécula de ADN que comprende la mutación se inserta en el genoma de la célula, o planta, huésped respectivamente, en lugar de la secuencia homóloga no mutada (Schaefer *et al.*, 1997, *Plant J.*; 11(6):1195-1206). Este método por lo tanto no funciona con un vector, sino con una molécula de ADN pura. La molécula se inserta en el anfitrión, p. ej., por bombardeo de genes, microinyección o electroporación, para mencionar sólo tres ejemplos. Como ya se ha explicado, la molécula de ADN se une a la secuencia homóloga en el genoma del anfitrión de modo que se produzca en el genoma una recombinación homóloga y por lo tanto la recepción de la mutación por delección, inserción o sustitución, respectivamente: la expresión de la  $\beta$ 1,2-xilosiltransferasa puede suprimirse o bloquearse completamente.
- 30 Preferentemente, se proporcionan plantas o células vegetales recombinantes, respectivamente, que han sido preparadas por uno de los métodos descritos anteriormente, estando su producción de  $\beta$ 1,2-xilosiltransferasa suprimida o completamente bloqueada, respectivamente. Preferentemente, la actividad de su  $\beta$ 1,2-xilosiltransferasa es inferior al 50%, en particular inferior al 50%, particularmente preferida del 0%, de la actividad de la  $\beta$ 1,2-xilosiltransferasa que se produce en las plantas o en las células de las plantas naturales, respectivamente. La ventaja de estas plantas o células vegetales, respectivamente, es que las glicoproteínas producidas no comprenden ninguna o apenas comprenden alguna xilosa unida por  $\beta$ 1,2. Si los productos de estas plantas son absorbidos por cuerpos humanos o de vertebrados, no existirá reacción inmunitaria debida al epítopo de  $\beta$ 1,2-xilosa.

- 40 La solicitud divulga asimismo una molécula de APN que comprende una secuencia de bases complementaria con la secuencia de la molécula de ADN según la invención. El APN (ácido nucleico peptídico) es una secuencia similar a la del ADN, que está unida por nucleobases a un eje central pseudopeptídico. El APN generalmente se híbida con oligómeros de ADN, ARN o APN complementarios por emparejamiento Watson-Crick de bases y formación de la hélice. El eje central peptídico asegura una resistencia mayor a la degradación enzimática. Por lo tanto la molécula de APN es un agente no codificante mejorado.

- 45 Ni las nucleasas ni las proteasas pueden unirse a una molécula de APN. La estabilidad de la molécula de APN, si está unida a una secuencia complementaria, comprende un bloqueo estérico suficiente de las ADN y ARN polimerasas, transcriptasa inversa, telomerasa y ribosomas. La publicación de Pooga *et al.*, "Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission *in vivo*" (*Nature Biotechnology* 16:857-861; 1998) se refiere a las moléculas de APN en general y específicamente a una molécula de APN que es complementaria del ARNm tipo 1 del receptor de galanina humano.

- 55 Si la molécula de APN comprende la secuencia mencionada anteriormente, se unirá al ADN o a un sitio del ADN, respectivamente, que codifica la  $\beta$ 1,2-xilosiltransferasa y de este modo puede inhibir la transcripción de esta enzima. Ya que ni se transcribe ni se traduce, la molécula de APN se preparará por síntesis, p. ej., con ayuda de la técnica t-Boc. Ventajosamente, está prevista una molécula de APN que comprende una secuencia de bases que corresponde a la secuencia de la molécula de ADN de la invención. Esta molécula de APN formará un complejo con el ARNm o con un sitio del ARNm de la  $\beta$ 1,2-xilosiltransferasa de modo que se inhiba la traducción de la enzima. Argumentos similares a los publicados para el ARN no codificante son aplicables en este caso. Por este motivo, p. ej., una zona de acomplejamiento particularmente eficaz es la zona inicial de la traducción o también las zonas 5' no traducidas del ARNm.

- 65 Es divulgado además un método de preparación de plantas o de células vegetales, respectivamente, en particular de células vegetales que comprende una expresión bloqueada de la  $\beta$ 1,2-xilosiltransferasa y el nivel de transcripción o de traducción, respectivamente, que se caracteriza porque las moléculas de APN de la invención están insertadas en las células. Para insertar la molécula o las moléculas de APN, respectivamente, en la célula, se utilizan de nuevo métodos

convencionales, tales como, p. ej., electroporación o microinyección. La inserción es particularmente eficaz si los oligómeros de APN se unen a péptidos de penetración celular, p. ej., transportan o pAntp (Pooga *et al.*, 1998, *Nature Biotechnology* 16; 857-861).

- 5 La solicitud divulga un método de preparación de glicoproteínas recombinantes que se caracteriza porque las plantas o las células vegetales recombinantes de la invención, respectivamente, cuya producción de  $\beta$ 1,2-xilosiltransferasa está suprimida o completamente bloqueada, respectivamente, o las plantas o células, respectivamente, en las que se ha insertado la molécula de APN según el método, se transfecan con el gen que expresa la glicoproteína de manera que se expresan las glicoproteínas recombinantes. Al actuar así, como se ha descrito ya anteriormente, los genes que comprenden vectores para las proteínas deseadas se transfecan en el anfitrión o en las células huésped, respectivamente, como se ha descrito ya anteriormente. Las células vegetales transfecadas expresarán las proteínas deseadas y no tendrán ninguna o apenas xilosa unida por  $\beta$ 1,2. Por lo tanto, no activan las reacciones inmunitarias ya mencionadas anteriormente en el cuerpo humano o de vertebrados. En estos sistemas puede producirse cualquier proteína.
- 10 15 Ventajosamente, está previsto un método de preparación de glicoproteínas humanas recombinantes que está caracterizado porque las plantas o las células vegetales recombinantes, respectivamente, cuya producción de  $\beta$ 1,2-xilosiltransferasa está suprimida o completamente bloqueada, o las plantas o células, respectivamente, en las que se han insertado las moléculas de APN según el método de la invención, se transfecan con el gen que expresa la glicoproteína de manera que se expresan las glicoproteínas recombinantes. Mediante este método se hace posible producir proteínas humanas en plantas (células vegetales) que, si son absorbidas por el cuerpo humano, no activan ninguna reacción inmunitaria dirigida contra los restos de xilosa unida por  $\beta$ 1,2. Es posible utilizar tipos de plantas para producir glicoproteínas recombinantes que sirven como materiales de alimentación, p. ej., plátano, patata y/o tomate. Los tejidos de esta planta comprenden la glicoproteína recombinante de modo que, p. ej., por extracción de la glicoproteína recombinante del tejido y posterior administración o directamente calentando el tejido vegetal, respectivamente, la glicoproteína recombinante se absorbe en el cuerpo humano.
- 20 25 Preferentemente, está previsto un método de preparación de glicoproteínas recombinantes humanas para utilización medicinal, en el que las células vegetales o plantas recombinantes inventivas, respectivamente, cuya producción de  $\beta$ 1,2-xilosiltransferasa está suprimida o completamente bloqueada, respectivamente, o las plantas o células, respectivamente, en las que se han insertado moléculas de APN, se transfecan con el gen que expresa la glicoproteína de manera que se expresan las glicoproteínas recombinantes. Al actuar de este modo, puede utilizarse cualquier proteína que sea de interés médico.
- 30 35 40 Además, la presente solicitud divulga glicoproteínas recombinantes según un método descrito anteriormente, en el que se han preparado en sistemas vegetales, y en el que su secuencia peptídica comprende menos del 50%, en particular menos del 20%, particularmente preferido el 0%, de los restos de xilosa unida por  $\beta$ 1,2 que aparecen en las proteínas expresadas en los sistemas vegetales sin xilosiltransferasa reducida. Naturalmente, las glicoproteínas que no comprenden restos de xilosa unida por  $\beta$ 1,2 resultan preferidas. La cantidad de xilosa unida por  $\beta$ 1,2 dependerá del grado de supresión de la  $\beta$ 1,2-xilosiltransferasa mencionado anteriormente.
- 45 La solicitud divulga unas glicoproteínas humanas recombinantes que han sido producidas en sistemas vegetales según un método descrito anteriormente y cuya secuencia peptídica comprende menos del 50%, en particular menos del 20%, particularmente 0%, de los restos de xilosa unida por  $\beta$ 1,2 que aparecen en las proteínas expresadas en los sistemas vegetales de xilosiltransferasa no reducida.
- 50 55 Son divulgadas además unas glicoproteínas humanas recombinantes para utilización médica que han sido preparadas en sistemas vegetales según un método descrito anteriormente, cuya secuencia peptídica comprende menos del 50%, en particular menos del 20%, particularmente 0%, de los restos de xilosa unida por  $\beta$ 1,2 que aparecen en las proteínas expresadas en los sistemas vegetales de xilosiltransferasa no reducida.
- 60 Las glicoproteínas pueden incluir otras unidades de oligosacáridos unidas específicas de plantas, por lo que – en el caso de las glicoproteínas humanas – se diferencian de estas glicoproteínas naturales. No obstante, mediante las glicoproteínas según la invención se desencadena una reacción inmunitaria más ligera o una reacción no inmunitaria del todo, respectivamente, en el cuerpo humano, ya que, como se ha explicado ya en la parte introductoria de la memoria, los restos de xilosa unida por  $\beta$ 1,2, junto con los restos de  $\alpha$ 1,3-fructosa, son la causa principal de las reacciones inmunitarias o de las reacciones inmunitarias cruzadas, respectivamente, en las glicoproteínas vegetales.
- 65 Es divulgada además una composición farmacéutica que comprende las glicoproteínas según la invención. Además de las glicoproteínas de la invención, la composición farmacéutica comprende más adiciones normales en dichas composiciones. Existen, p. ej., agentes de dilución adecuados de contenidos de tampón variables (p. ej., Tris-HCl, acetato, fosfato, pH y fuerza iónica), aditivos, tales como tensioactivos y solubilizantes (p. ej., Tween 80, Polysorbate 80), conservantes (p. ej., Timersosal, alcohol bencílico) adyuvantes, antioxidantes (p. ej., ácido ascórbico, metabisulfito de sodio), emulsionantes, cargas (p. ej., lactosa y manitol), enlaces covalentes de polímeros, tal como polietilénglico a la proteína, incorporación del material en composiciones en partículas de compuestos poliméricos, tales como el ácido poliláctico, ácido poliglicólico, etc. o en liposomas, agentes auxiliares y/o sustancias portadoras que son adecuadas en

el tratamiento respectivo. Dichas composiciones influirán en el estado físico, estabilidad, velocidad de liberación *in vivo* y velocidad de excreción *in vivo* de las glicoproteínas.

5 La solicitud divulga asimismo un método de selección de moléculas de ADN que codifican una  $\beta$ 1,2-xilosiltransferasa, en una muestra, en el que las moléculas de ADN marcadas o secuencias parciales de las mismas, se añaden a la muestra, que se une a las moléculas de ADN que codifican una  $\beta$ 1,2-xilosiltransferasa. Las moléculas de ADN hibridadas pueden detectarse, cuantificarse y seleccionarse. En la muestra que contiene ADN de una sola cadena con el que las moléculas de ADN marcadas pueden hibridarse, la muestra se desnaturiza, p. ej., por calentamiento.

10 Una manera posible consiste en separar el ADN que debe analizarse, posiblemente tras la adición de endonucleasas, mediante electroforesis en gel o en gel de agarosa. Una vez han sido transferidas a una membrana de nitrocelulosa, se mezclan las moléculas de ADN marcadas las cuales se hibridan con la correspondiente molécula homóloga de ADN ("transferencia Southern").

15 Otra manera posible consiste en encontrar genes homólogos de otras especies por métodos dependientes de PCR utilizando cebadores específicos y/o degenerados, derivados de la secuencia de la molécula de ADN divulgada.

20 Las moléculas de ADN marcadas de la invención o las secuencias parciales de las mismas se inmovilizan de manera ventajosa en matrices del portador. La utilización de micromatrices de ADN ("chips génicos") es una manera posible más de hallar genes homólogos o de estudiar el nivel de expresión de los genes homólogos. Con esta finalidad, el ADN que representa la secuencia del gen genómica completa, la secuencia del ADNc completa, las partes de estas secuencias o cualquier combinación de las secuencias parciales se inmoviliza en las matrices del portador, para que los genes homólogos, tras añadir la muestra a las matrices del portador, se hibriden con las moléculas de ADN marcadas (para ejemplos véanse p. ej., Ferea T.L. y Brown, P. O., 1999, Current Opinion in Genetics & Development 9: 715-722 y las referencias citadas en la presente memoria).

25 Preferentemente, la muestra para el método de la invención identificado anteriormente comprende ADN genómico de un organismo vegetal. Mediante este método, se analiza la presencia del gen de la  $\beta$ 1,2-xilosiltransferasa en un gran número de plantas u otras especies de manera muy rápida y eficaz. De este modo, es posible seleccionar plantas o individuos respectivamente de otras especies que no comprendan este gen, o suprimir o bloquear completamente, respectivamente, la expresión de  $\beta$ 1,2-xilosiltransferasa en dichas plantas u otros organismos que comprendan este gen, mediante el método de la invención descrito anteriormente, de modo que pueda utilizarse posteriormente para la transfección y producción de proteínas (humanas).

30 35 La solicitud divulga asimismo las moléculas de ADN que codifican una  $\beta$ 1,2-xilosiltransferasa que pueden seleccionarse según los tres últimos métodos mencionados y aislarla posteriormente de la muestra. Estas moléculas pueden utilizarse para ensayos ulteriores. Pueden secuenciarse y a su vez pueden utilizarse como sondas de ADN para encontrar la  $\beta$ 1,2-xilosiltransferasa. Estas moléculas de ADN – marcadas – funcionarán para organismos, que están relacionados con los organismos de los que se han aislado, más eficazmente como sondas que las moléculas de ADN divulgadas.

40 45 Es divulgada además una preparación de  $\beta$ 1,2-xilosiltransferasa clonada según la invención que comprende isoformas con valores pI comprendidos entre 6,0 y 9,0, en particular entre 7,50 y 8,00. El valor pI de una proteína es el valor del pH en el cual su carga neta es cero y depende de la secuencia de aminoácidos, del tipo de glicosilación así como de la estructura espacial de la proteína. La  $\beta$ 1,2-xilosiltransferasa puede comprender varias isoformas que presentan un valor de pH en este intervalo. Las razones de las varias isoformas de la transferasa son, p. ej., las diferentes glicosilaciones así como la proteólisis limitada. El valor de pH de una proteína puede determinarse por enfoque isoelectrónico, que es conocido por el experto en la materia.

50 La principal isoforma de la enzima tiene un peso molecular aparente de 60,2 kDa.

En particular, la preparación de la invención comprende una isoforma que presenta un valor pI de 7,52.

55 La solicitud divulga asimismo a un método de preparación de unidades de carbohidrato "incorporadas a plantas" de glicoproteínas humanas y de otros vertebrados o de otros glicoconjungados, en el que la UDP-xilosa así como la  $\beta$ 1,2-xilosiltransferasa codificadas por una molécula de ADN descrita anteriormente se añaden a una muestra que comprende una unidad de carbohidrato o una glicoproteína, respectivamente, de modo que la xilosa en posición  $\beta$ 1,2 se une mediante la  $\beta$ 1,2-xilosiltransferasa a la unidad de carbohidrato o a la glicoproteína, respectivamente. Mediante el método para la clonación de la  $\beta$ 1,2-xilosiltransferasa es posible producir grandes cantidades de enzima purificada.

60 Para obtener una transferasa totalmente activa, se proporcionan condiciones de reacción adecuadas.

La invención se pondrá más claramente de manifiesto haciendo referencia a los dibujos a título de ejemplo no limitativo. Con detalle, en los dibujos,

65 la Fig. 1 representa la secuencia de aminoácidos de los péptidos 2 y 3 de soja (patente WO 99/29835; SEC. ID. n° 3 y n° 5), la homología entre estos péptidos y una secuencia de *A. thaliana* así como la secuencia de ADN de

cuatro cebadores 1 a 4;

la Fig.2 representa la secuencia de ADNc de  $\beta$ 1,2-xilosiltransferasa incluyendo el nt 226 de la zona 5' no traducida;

5 la Fig.3 representa la secuencia de aminoácidos de  $\beta$ 1,2-xilosiltransferasa derivada de la misma;

las Figs. 4a, 4b y 4c representan la alineación de los ADNc de  $\beta$ 1,2-xilosiltransferasa de *A. thaliana*, un ADN genómico y una secuencia EST;

10 la Fig. 5 representan la alineación de la secuencia de aminoácidos de  $\beta$ 1,2-xilosiltransferasa derivada de los ADN, de un ADN genómico y de una secuencia EST;

15 la Fig. 6 es una representación esquemática de la  $\beta$ 1,2-xilosiltransferasa así como de los productos de la PCR y de la hidrofobia de los restos de aminoácidos;

15 la Fig. 7 representa una comparación de la actividad de  $\beta$ 1,2-xilosiltransferasa de células de insectos transfectadas con el gen de  $\beta$ 1,2-xilosiltransferasa con la de una referencia negativa;

20 las Figs. 8a y 8b representan la estructura de un sustrato receptor y el producto de la  $\beta$ 1,2-xilosiltransferasa;

la Fig. 9 representa los espectros de masas;

la Fig. 10 representa el análisis del producto de  $\beta$ 1,2-xilosiltransferasa por HPLC en fase inversa, y

25 la Fig. 11 representa la alineación de la secuencia de aminoácidos prevista derivada del ADNc de la presente solicitud con la secuencia de aminoácidos de la  $\beta$ 1,2-xilosiltransferasa purificada en soja.

### Ejemplo 1

#### RT-PCR y clonación de ADNc

Se diseñaron cebadores para la ampliación del ADNc de  $\beta$ 1,2-xilosiltransferasa supuesto por RT-PCR de la forma siguiente: Una exploración por BLASTP de la base de datos utilizando dos péptidos de soja (SEC. ID. nº: 1 y nº 2;

35 correspondientes a SEC. ID. nº: 3 y nº 5 en la Fig. 4 en la patente WO 99/29835 A1; sin embargo, los aminoácidos LG C-terminales se omitieron en la SEC. ID. nº: 5) (véase la Fig. 1) presentó una secuencia de proteínas derivada de una secuencia de ADN genómico de *Arabidopsis thaliana* (nº de registro: AB015479) con mas del 80% de homología (SEC. ID. nº: 3).

35 Los cebadores 3 (SEC. ID. nº: 4) y 4 (SEC. ID. nº: 5) se basaban en la secuencia de *A. thaliana* homóloga a los péptidos 2 y 3 de soja. El análisis de la secuencia homóloga del ADN genómico que utiliza Gene-Finder en el BCM Search Launcher produjo una proteína prevista. El cebador 1 (SEC. ID. nº: 6) se diseñó para incluir el codón de iniciación de la proteína prevista, mientras que el cebador 2 (SEC. ID. nº: 7) contiene el codón de terminación de la proteína prevista.

40 Se aisló el ARN completo de las hojas nuevas de *Arabidopsis thaliana* var. Columbia utilizando el reactivo TRIzol (Life Technologies). Se trató ARN con ADNasa (Promega, RQ1 ARNasa exenta de ADNasa) para eliminar los vestigios de ADN genómico. Se sintetizó la primera cadena del ADNc a partir de 1  $\mu$ g de ARN total a 42°C utilizando oligo(dT) cebadores (Sigma) y transcriptasa inversa de AMV (Promega).

45 La primera cadena del ADNc se sometió a una PCR, en la que se utilizaron diferentes combinaciones de cebadores codificantes y no codificantes (ilustrados en la Fig. 6): El producto del cebador 3 y del cebador 4 fue un fragmento de ADN con una longitud de 174 bp (P1), el producto del cebador 1 y del cebador 2 fue un fragmento de ADN con una longitud de 1.605 bp (P2), el producto del cebador 1 y del cebador 4 fue un fragmento de ADN con una longitud de 1.525 bp (P3) y el cebador 3 y el cebador 4 produjeron un ADN de 254 bp (P4). Para la ampliación del marco de lectura abierto supuesto se utilizaron el cebador 1 y el cebador 2. Una reacción de PCR contenía en un volumen total de 50  $\mu$ l, 0,2  $\mu$ moles de cada cebador, dNTPs 0,05 mM, MgSO<sub>4</sub> 2 mM, Tris-HCl 20 mM (pH 8,2 a 25°C), KCl 10 mM, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 10 mM, Triton X-100 al 0,1%, 5  $\mu$ g de BSA exento de nucleasa y 2,5 unidades Pfu de ADN polimerasa de Promega. Despues de una primera etapa de desnaturación a 94°C durante 2 min, se realizaron 30 ciclos de 1 min a 92°C, 40 s a 53°C y 3 min y 30 s a 72°C. La última etapa de amplificación se realizó a 72°C durante 8 min. Los productos de la PCR se subclonaron en un vector puc19 linealizado y desfosforilado de SmaI, y se secuenciaron: Se obtuvieron las secuencias de los fragmentos subclonados mediante el método del didesoxinucleótido (Kit de reacción ABI PRISM

50 Dye Termination Cycle Sequencing Ready y el analizador ABI PRISM 310 Genetic de Perkin Elmer). Como resultado de la RT-PCR se obtuvieron tres secuencias de ADNc ligeramente diferentes. La secuencia de ADNc 6 que tiene un tamaño de 1.605 bp (xt-Ath6; SEC. ID. nº: 8) y codifica una proteína de 534 aminoácidos que tiene un peso molecular de 60,2 kDa y un valor pI teórico de 7,52 (véase la Fig. 3) es idéntica a la secuencia de nucleótidos derivada del clon genómico después de la eliminación de dos intrones (xt-Athgen.seq). El ADNc 9 presenta 4 cambios de pares de bases comparado con el ADNc 6, mientras que el ADNc 16 presenta 6 cambios de pares de bases comparado con el

55 ADNc 6 (ilustrados en las Figs. 4a, 4b y 4c). Por consiguiente la secuencia de aminoácidos derivada del ADNc 9

comprende dos cambios comparada con la secuencia de aminoácidos derivada del ADNc 6 (SEC. ID. nº: 8) y la secuencia de aminoácidos del ADN 16 presenta cuatro restos cambiados (ilustrados en la Fig. 5).

5 La Fig. 3 representa la secuencia (SEC. ID. nº: 9) de aminoácidos derivada del ADNc de la  $\beta$ 1,2-xilosiltransferasa (xt-Ath6; SEC. ID. nº: 8). Los sitios potenciales de glicosilación de la unión a la asparagina son Asn51, Asn301 y Asn479.

10 Las Figs. 4a, 4b y 4c representan la alineación de las secuencias de nucleótidos de la  $\beta$ 1,2-xilosiltransferasa del ADNc 6 de *A. thaliana* (xt-Ath6; SEC. ID. nº: 8), del ADNc 9 de *A. thaliana* (xt-Ath9), del ADNc 16 de *A. thaliana* (xt-Ath16), la secuencia del ADN genómico de *A. thaliana* tras la eliminación de los dos intrones (xt-Athgen) y la secuencia EST (xtAthEST). La línea de puntos representa la secuencia de consenso; la línea de rayas representa un hueco.

15 La secuencia genómica (xt-Athgen; nº de registro AB015479, codón de inicio en las posiciones 58.185 a 58.187, codón de terminación en las posiciones 60.214 a 60.216 del ADN genómico) procede de la eliminación de dos supuestos intrones (intrón 1: desde la posición 58.881 hasta la 59.116 del ADN genómico; intrón 2: desde la posición 59.268 hasta la 59.458 del ADN genómico) utilizando el servidor de predicción del sitio de corte y empalme NetPlantGene. La secuencia EST de *A. thaliana* (xt-AthEST; nº de registro AI994524) es el resultado de una exploración en la base de datos utilizando BLASTN.

20 La Fig. 5 representa la alineación de las secuencias de aminoácidos de la  $\beta$ 1,2-xilosiltransferasa procedente del ADNc 6 de *A. thaliana* (xt-Ath6; SEC. ID. nº: 9), del ADNc 9 de *A. thaliana* (xt-Ath9), del ADNc 16 de *A. thaliana* (xt-Ath16), de la secuencia genómica de *A. thaliana* (xt-Athgen) y procedente de la secuencia EST (xtAthEST) de *A. thaliana*. La línea de puntos representa la secuencia de consenso; la línea de rayas representa un hueco.

25 En la Fig. 6, se ilustra la representación esquemática prevista de la proteína  $\beta$ 1,2-xilosiltransferasa (arriba) y el índice de hidrofobia derivado utilizando ProtScale, de la proteína codificada (abajo), un índice de hidrofobia positivo significa un aumento de hidrofobia. Entre los tamaños de los cuatro productos de PCR indicados anteriormente (P1 a P4) se presentan en relación con el ADNc. Codificación "C" para la zona citoplasmática supuesta, "T" para la zona de la secuencia de la proteína por "TMpred" (de EMBnet) proporcionó una zona de transmembrana supuesta desde Ile11 hasta Phe29. La zona C-terminal de la enzima comprende probablemente la zona catalítica y por consiguiente debería dirigirse al lumen del aparato de Golgi. Según esto, esta transferasa parece ser una proteína de transmembrana de tipo II como todas las glicosiltransferasas analizadas hasta ahora que están involucradas en la biosíntesis de la glicoproteína (Joziasse, 1992, *Glycobiology*, 2, 271-277). Las zonas grises representan la posición de los dos péptidos, los hexágonos representan los sitios de N-glucosilación potencial. Una exploración BLAST en los bancos de datos accesible a través de NCBI presentó únicamente gran homología con otra secuencia de la planta (álamo temblón híbrido, nº de registro AI62640).

## Ejemplo 2

### Expresión de $\beta$ 1,2-xilosiltransferasa recombinante en células de insectos

40 La zona de codificación completa de la  $\beta$ 1,2-xilosiltransferasa supuesta incluyendo la zona citoplasmática y de transmembrana se separó del vector puc19 mediante digestión con BamHI y EcoRI y se subclonó en el vector pVL 1393 de transferencia del baculovirus digerido con BamHI/EcoRI y desfosforilado (PharMingen, San Diego, CA). Se confirmó la clonación correcta por secuenciado utilizando 5'-AACCATCTCGCAAATAAGTA-3' (SEC. ID. nº: 10) del cebador transcrto pVL1393 y 5'-GTCGGGTTAACATTACGGATTC-3' (SEC. ID. nº: 11) del cebador complementario pVL1393. Para asegurar una recombinación homóloga 1 $\mu$ g, se transfeció el vector de transferencia junto con 200 ng de ADN vírico Baculo-Gold lineal (PharMingen, San Diego, CA) en 1  $\times$  10<sup>6</sup> células Sf-9 en medio IPL-41 utilizando Lipofectina (Life Technologies) según el protocolo del fabricante. Tras 5 días de incubación a 27°C, se utilizaron varios volúmenes de sobrenadante con el virus recombinante para la infección de las células Sf-21 de insecto. Se incubaron las células en medio IPL-41 enriquecido con suero de ternero fetal inactivado por calor al 5% durante 4 días a 27°C, a continuación se recogieron y se lavaron 2 veces con solución salina tamponada con fosfato. Se volvieron a poner en suspensión las células y se homogeneizaron en el tampón siguiente (1 ml para 10<sup>7</sup> células): tampón MES 100 mM, pH 7,0, con Triton X-100 al 1%, DTT 1 mM, PMSF 1 mM, 5  $\mu$ g/ml de Leucopeptina (Sigma), 5  $\mu$ g/ml de E-64 (Serva) y se incubaron en hielo durante 30 min.

## Ejemplo 3

### Determinación de la actividad de $\beta$ 1,2-xilosiltransferasa

60 Se determinó la actividad de la  $\beta$ 1,2-xilosiltransferasa en los homogeneizados celulares. Se prepararon referencias negativas con el mismo número de células no infectadas. Las mezclas de ensayo contenían, en un volumen total de 20  $\mu$ l, 13  $\mu$ l de células homogeneizadas, 2 nmoles de hexapéptido GnGn dabsilado o de GnGn-piridilamina como sustrato receptor (Fig. 8a), UDP-xilosa 1 mM como sustrato donante, ATP 10 mM, MnCl<sub>2</sub> 20 mM y se incluyó 2-acetamido-1,2-didesoxi-nojirimicina para impedir la degradación del producto por N-acetilhesoxaminidasa. Se incubaron las muestras durante 1 hora a 37°C y se analizaron por espectrometría de masas MALDI-TOF.

La Fig. 7 presenta la actividad enzimática medida de la  $\beta$ 1,2-xilosiltransferasa así como de la referencia negativa. Las barras grises presentan la actividad cuando se utilizó el hexapéptido GnGn como sustrato, mientras que las barras negras indican la utilización de GnGn-piridilamina como sustrato. La actividad enzimática de las células cotransfectedas fue 30 veces mayor que la de las referencias negativas.

5 La estructura del sustrato receptor de  $\beta$ 1,2-xilosiltransferasa se representa en la Fig. 8a, y el producto supuesto en la Fig. 8b, en la que R representa una piridilamina o un resto de heptapéptido dabsilado.

#### Ejemplo 4

##### Espectrometría de masas del producto de xilosiltransferasa

10 Se realizó la espectrometría de masas en una DYNAMO (ByoAnalysis, Santa Fe, NM), un MALDITOF MS que permite la extracción dinámica (sinónimo de extracción tardía). Se utilizaron dos tipos de preparaciones de la matriz de muestra: se disolvieron glicopéptidos dabsilados en ácido fórmico al 5% y se aplicaron alícuotas a la diana, se secaron con aire y se cubrieron con ácido alfa-ciano-4-hidroxi cinnámico al 1%. Se diluyeron con agua los glucanos piridilaminados, se aplicaron a la diana y se secaron con aire. Tras la adición de ácido 2,5-dihidroxibenzoico al 2%, se secaron inmediatamente las muestras aplicando vacío.

15 20 La Fig. 9 representa el espectro de masas de estas muestras, (A) siendo la referencia negativa: El pico principal (S) muestra el sustrato dabsil-Val-Gly-Glu-(GlcNAc<sub>4</sub>Man<sub>3</sub>)Asn-Arg-Thr, siendo 2262,3 el valor  $[M+H]^+$  calculado. Este sustrato también aparece como producto de adición de sodio y es un ión más pequeño que se ha formado por fragmentación de la función azo del grupo dabsilo, en (S\*). El pico a  $m/z$  = 2424,4 presenta la desgalactosilación incompleta del sustrato. El espectro de masas (B) presenta la muestra con  $\beta$ 1,2-xilosiltransferasa recombinante tras 25 incubación durante 1 h a 37°C. El pico principal (P), que tiene un valor de  $[M+H]^+$  de 2393,4, representa el producto xilosilado.

#### Ejemplo 5

##### Ánálisis por HPLC del producto xilosiltransferasa

30 35 Se realizaron análisis de xilosiltransferasa como se describió anteriormente en el ejemplo 3, con la excepción de que se utilizaron 10 nmoles de GnGn-piridilamina como sustrato receptor. Tras 4 h de incubación se analizó la muestra tanto por espectrometría de masas MALDI-TOF y por HPLC en sed-fase inversa para comprobar la estructura del producto. Se recogió el supuesto pico del producto eluyendo ligeramente delante del sustrato GnGn-PA. Se determinó por MS MALDI-TOF que la masa del producto era 1550,9 lo que concuerda con GnGnX-PA. En la digestión con  $\beta$ -N-acetilglicosamidasa de riñón bovino (en tampón de citrato de sodio 50 mM de pH 5,0 durante 20 h a 37°C con 25 muestra de enzima), se eluyó el glucano con aproximadamente la retención de MM. Esto está de acuerdo con los datos publicados sobre la retención de MM-PA y MMX-PA (Wilson & Altmann, 1998, *Glycoconj. J.* 15, 1055-1070). La 40 digestión adicional con alfa-manosidasa de judías blancas en las condiciones seleccionadas (20 h a 37°C con 50 mU de enzima) produjo la aparición de dos nuevos picos. Como la manosa ligada por alfa-1,3 es considerablemente más sensible a la manosidasa que la manosa ligada por alfa-1,6, los picos se asignan a 00X y M0X (en orden de elución). De hecho, la M0X-piridilamina preparada a partir de bromelina por desfucosilación con ácido eluyó junto con la M0X 45 supuesta procedente del producto de xilosiltransferasa. El desplazamiento apenas importante del tiempo de elución debido a la eliminación del resto de manosa ligado por alfa-1,3 es una firme indicación de que la  $\beta$ -manosa se sustituye por la xilosa (Wilson & Altmann, 1998, *Glycoconj. J.* 15, 1055-1070; Wilson & Altmann, 1998, *Glycoconj. J.* 15, 79-82).

50 55 La Fig. 10 representa el análisis del producto de xilosiltransferasa por HPLC en fase inversa. (A) mezcla para incubación de transferasa; (B) producto aislado de la xilosiltransferasa; (C) producto aislado de la xilosiltransferasa tras la digestión con  $\beta$ -N-acetilglicosaminidasa; (D) producto aislado de la xilosiltransferasa tras digestión adicional con alfa-manosidasa. Las asignaciones de los picos son las siguientes: 1, GnGn-PA; 2, GnGnX-PA; 3, MMX-PA; 4, M0X-PA; 5, 00X-PAA; 6, M0-PA (procedente de vestigios de sustrato en el producto aislado). Para abreviaturas de las estructuras de N-glucano véase Wilson I.B.H. y Altmann, F., 1998, *Glycoconj. J.* 15, 1055-1077.

60 La Fig. 11 representa la alineación de la secuencia de aminoácidos prevista según el documento WO 99/29835 A1. Esta alineación demuestra que la secuencia de aminoácidos de la enzima de soja purificada corresponde solamente a los aminoácidos 199 a 469 de la secuencia procedente del ADNc según la presente invención. Además, la secuencia de aminoácidos prevista procedente del ADNc de la presente solicitud contiene dos inserciones (correspondientes a los aa 375 a 382 y a los aa 425 a 429 de la secuencia prevista) en comparación con la secuencia de la enzima de soja purificada.

**Listado de secuencias**

&lt;110&gt; Glössl Prof., Josef

5 &lt;120&gt; gen de xilosiltransferasa

&lt;130&gt; gen de xilosiltransferasa

&lt;140&gt;

10 &lt;141&gt;

&lt;160&gt; 11

&lt;170&gt; PatentIn Ver. 2.1

15

&lt;210&gt; 1

&lt;211&gt; 27

&lt;212&gt; PRT

&lt;213&gt; soja

20

&lt;400&gt; 1

**Ser Gln Val Gln Ala Ile His Asp Ala Ser Val Ile Ile Gly Ala His**  
 1 . . . . . 5 . . . . . 10 . . . . . 15

**Gly Ala Gly Leu Thr His Ile Val Ser Ala Leu**  
 . . . . . 20 . . . . . 25

25 &lt;210&gt; 2

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; soja

&lt;400&gt; 2

30

**Gly Leu Glu Tyr His Ala Ile Asn**  
 1 . . . . . 5

&lt;210&gt; 3

&lt;211&gt; 60

&lt;212&gt; PRT

35 <213> *arabidopsis thaliana*

&lt;400&gt; 3

**Asp Gln Val Arg Ala Ile Gln Asp Ala Ser Val Ile Ile Gly Ala His**  
 1 . . . . . 5 . . . . . 10 . . . . . 15

**Gly Ala Gly Leu Thr His Ile Val Ser Ala Thr Pro Asn Thr Thr Ile**

1

**Phe Glu Ile Ile Ser Val Glu Phe Gln Arg Pro His Phe Glu Leu Ile**  
 . . . . . 35 . . . . . 40 . . . . . 45

**Ala Lys Trp Lys Gly Leu Glu Tyr His Ala Met His**  
 . . . . . 50 . . . . . 55 . . . . . 60

40

&lt;210&gt; 4

<211> 21  
 <212> ADN  
 <213> Secuencia artificial

5 <220>  
 <223> Descripción de la secuencia artificial: cebador

<400> 4

10 **gatcaagtcc gagccattca a** 21

<210> 5  
 <211> 21  
 <212> ADN

15 <213> Secuencia artificial

<220>  
 <223> Descripción de la secuencia artificial: cebador

20 <400> 5

**cgcgtgatac tccaaatcctt t** 21

<210> 6  
 25 <211> 21  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 30 <223> Descripción de la secuencia artificial: cebador

<400> 6

**atgagtaaac ggaatccgaa g** 21

35 <210> 7  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial

40 <220>  
 <223> Descripción de la secuencia artificial: cebador

<400> 7

45 **ttagcagcca aggctttca t** 21

<210> 8  
 <211> 1831  
 <212> ADN

50 <213> *Arabidopsis thaliana*

<400> 8

aaatctgcag actctcaaaa ttccgattca tcttattgaa gaacaatttt ccggcgaaac 60  
 agccgatgaa gtctcgctg aatcttctgt acctttcacc ggcgattgac ttcacttcag 120  
 aatcgagaga gaagaaatcg atggaaaact aaaaatagaa agagttcaa attctcgctc 180  
 tctcttcaaa accgcaaattc aagggAACGA gagacgagag agagagatga gtaaacggaa 240  
 tccgaagatt ctgaagattt ttctgtatata gttacttctc aactctctct ttctcatcat 300  
 ctacttcgtt tttactcat cgctcggtttc accggagcag tcacagcctc ctcatatata 360  
 ccacgttca gtgaataacc aatcgccgat tcagaaaccg tggccgatct taccttctta 420  
 cctcccatgg acgcccggc agaggaatct accaactggc tcctgcgaag gttacttcgg 480  
 gaatggatt acaaagagag ttgacttcct taagccgagg attggaggag gaggagaagg 540  
 aagctggttc cgatgtttt acagtgagac attacagagt tcgatttgcg aaggaaggaa 600  
 tctgagaatg gttccggatc ggattgttat gtcgagagga ggtgagaagt tagaggaagt 660  
 tatggggagg aaagaggagg aggagcttcc tgcgttgcga caaggtgcgt ttgaggttagc 720  
 ggaagaggtt tcttcacggt taggtttaa gagacaccgt cggtttgggt gaggagaagg 780  
 aggttagtgcg gtttctcgcc ggctggtaa tgatgagatg ttgaatgaat atatgcaaga 840  
 aggtggaaatt gatagacata caatgagaga tttgggtgct tcgattcgtg ctgttgatac 900  
 caatgatttc gtttgtgaag agtgggtgga ggaaccgacc ttgcgttgcg ctagattcga 960  
 gtacgcaaat ctcttcata ctgtgacaga ttggtatagt gcctatgttt cgtctagagt 1020  
 caccgggtttc ctaatcgac ctcacgttgc ttgcgttgcg ggacactgca cgacgcagct 1080  
 agaagaaaca tggacagctt tggttccgg aatcagatac gcaaagaact tcaccaaacc 1140  
 ggtttgtttc cgccacgcga ttcttcacc attgggatac gaaaccgctc tttttaaagg 1200  
 cttgtccggg gaaatagact gcaaggggaga ttcaagtcac aatctgtggc aaaaccggg 1260  
 cgataaaagg actgcgagga tatcagagtt tggtgaaatg atcagagcag ctttcgggtt 1320  
 gcctgtcaat agacaccgct cattagaaaa gcccgttatca tcatcatcat catcagcctc 1380  
 agtttataat gttcttttg tccaccgtga agattactta gcccattcctc gtcatggcgg 1440  
 taaaagtccag tctcggtca tcaatgagga agaagtgttc gactcggtgc atcattgggt 1500  
 tgcaactggg tccaccgtc tgaccaaattg cgggattaaac cttgtgaatg gcttgcttgc 1560  
 acacatgtca atgaaagatc aagtccgagc cattcaagat gcttcagtga tcataggagc 1620  
 tcatggagca ggactgactc acattgtctc tgcaacacca aacacacaacga tatttgagat 1680  
 aataagcgtc gagtttcaga gacctcattt cgagcttata gctaagtggaa aaggattggaa 1740  
 gtatcacgctg atgcacatctgg cgaactcactg acggctgtga ttgagaagtt 1800  
 aacggagatc atgaagagcc ttggctgcta a 1831

<210> 9  
 <211> 533  
 <212> PRT

5 <213> *Arabidopsis thaliana*

<400> 9

Met Ser Lys Arg Asn Pro Lys Ile Leu Lys Ile Phe Leu Tyr Met Leu  
 1 5 10 15

Leu Leu Asn Ser Leu Phe Leu Ile Ile Tyr Phe Val Phe His Ser Ser  
 20 25 30

Ser Phe Ser Pro Glu Gln Ser Gln Pro Pro His Ile Tyr His Val Ser  
 35 40 45

Val Asn Asn Gln Ser Ala Ile Gln Lys Pro Trp Pro Ile Leu Pro Ser  
 50 55 60

Tyr Leu Pro Trp Thr Pro Pro Gln Arg Asn Leu Pro Thr Gly Ser Cys  
 65 70 75 80

Glu Gly Tyr Phe Gly Asn Gly Phe Thr Lys Arg Val Asp Phe Leu Lys  
 85 90 95

Pro Arg Ile Gly Gly Gly Glu Gly Ser Trp Phe Arg Cys Phe Tyr  
 100 105 110

Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Arg Asn Leu Arg Met  
 115 120 125

Val Pro Asp Arg Ile Val Met Ser Arg Gly Gly Glu Lys Leu Glu Glu  
 130 135 140

Val Met Gly Arg Lys Glu Glu Glu Glu Leu Pro Ala Phe Arg Gln Gly  
 145 150 155 160

Ala Phe Glu Val Ala Glu Glu Val Ser Ser Arg Leu Gly Phe Lys Arg  
 165 170 175

His Arg Arg Phe Gly Gly Glu Gly Ser Ala Val Ser Arg Arg  
 180 185 190

Leu Val Asn Asp Glu Met Leu Asn Glu Tyr Met Gln Glu Gly Gly Ile  
 195 200 205

Asp Arg His Thr Met Arg Asp Leu Val Ala Ser Ile Arg Ala Val Asp  
 210 215 220

Thr Asn Asp Phe Val Cys Glu Glu Trp Val Glu Glu Pro Thr Leu Leu  
 225 230 235 240

Val Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Val Thr Asp Trp  
 245 250 255

Tyr Ser Ala Tyr Val Ser Ser Arg Val Thr Gly Leu Pro Asn Arg Pro  
 260 265 270

His Val Val Phe Val Asp Gly His Cys Thr Thr Gln Leu Glu Glu Thr  
 275 280 285

Trp Thr Ala Leu Phe Ser Gly Ile Arg Tyr Ala Lys Asn Phe Thr Lys  
 290 295 300

Pro Val Cys Phe Arg His Ala Ile Leu Ser Pro Leu Gly Tyr Glu Thr  
 305 310 315 320

Ala Leu Phe Lys Gly Leu Ser Gly Glu Ile Asp Cys Lys Gly Asp Ser  
 325 330 335

Ala His Asn Leu Trp Gln Asn Pro Asp Asp Lys Arg Thr Ala Arg Ile  
 340 345 350

Ser Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Gly Leu Pro Val Asn  
 355 360 365

Arg His Arg Ser Leu Glu Lys Pro Leu Ser Ser Ser Ser Ala  
 370 375 380

Ser Val Tyr Asn Val Leu Phe Val Arg Arg Glu Asp Tyr Leu Ala His  
 385 390 395 400

Pro Arg His Gly Gly Lys Val Gln Ser Arg Leu Ile Asn Glu Glu Glu  
 405 410 415

Val Phe Asp Ser Leu His His Trp Val Ala Thr Gly Ser Thr Gly Leu  
 420 425 430

Thr Lys Cys Gly Ile Asn Leu Val Asn Gly Leu Leu Ala His Met Ser  
 435 440 445

Met Lys Asp Gln Val Arg Ala Ile Gln Asp Ala Ser Val Ile Ile Gly  
 450 455 460

Ala His Gly Ala Gly Leu Thr His Ile Val Ser Ala Thr Pro Asn Thr  
 465 470 475 480

Thr Ile Phe Glu Ile Ile Ser Val Glu Phe Gln Arg Pro His Phe Glu  
 485 490 495

Leu Ile Ala Lys Trp Lys Gly Leu Glu Tyr His Ala Met His Leu Ala  
 500 505 510

Asn Ser Arg Ala Glu Pro Thr Ala Val Ile Glu Lys Leu Thr Glu Ile  
 515 520 525

Met Lys Ser Leu Gly Cys  
 530

- 5 <210> 10  
 <211> 24  
 <212> ADN  
 <213> Secuencia artificial
- 10 <220>  
 <223> Descripción de la secuencia artificial: cebador  
 <400> 10

aaccatctcg caaataaata agta

- <210> 11  
<211> 24  
<212> ADN  
5 <213> Secuencia artificial  
  
<220>  
<223> Descripción de la secuencia artificial: cebador  
  
10 <400> 11

**gtcgggtta acattacgga ttcc**

**24**

## REIVINDICACIONES

- 5 1. Método para la expresión de glicoproteínas recombinantes, caracterizado porque las glicoproteínas recombinantes son expresadas en una planta o célula vegetal, en el que la expresión de la  $\beta$ 1,2-xilosiltransferasa es suprimida o bloqueada completamente de manera que la secuencia peptídica de la glicoproteína presente menos de 50%, particularmente menos de 20%, preferentemente particularmente 0% de restos de xilosa unida por  $\beta$ 1,2 presentes en las proteínas expresadas en los sistemas vegetales sin xilosiltransferasa reducida, mediante la transfección con un ARN no codificante que es complementario al ARNm de una  $\beta$ 1,2-xilosiltransferasa endógena, que se codifica mediante una secuencia según la SEC ID nº: 8 con un marco de lectura abierto desde el par de bases 227 hasta el par de bases 1831, una secuencia que es idéntica por lo menos 70% a la SEC ID nº: 8, o una secuencia que hibrida con la SEC ID nº: 8 en condiciones severas, o una secuencia que ha degenerado a SEC ID nº: 8 debido al código genético, o con un vector que codifica dicho ARN no codificante.
- 10 2. Método para la expresión de glicoproteínas recombinantes, caracterizado porque las glicoproteínas recombinantes son expresadas en una planta o célula vegetal, en el que la expresión de la  $\beta$ 1,2-xilosiltransferasa es suprimida o bloqueada completamente de manera que la secuencia peptídica de la glicoproteína presenta menos de 50%, particularmente menos de 20%, preferentemente particularmente 0% de restos de xilosa unida por  $\beta$ 1,2 presentes en las proteínas expresadas en los sistemas vegetales sin xilosiltransferasa reducida, en el que la reducción es mediante mutación por delección, inserción y/o sustitución de un gen que codifica una proteína que presenta una actividad  $\beta$ 1,2-xilosiltransferasa codificada por una secuencia según la SEC ID nº: 8 con un marco de lectura abierto desde el par de bases 227 al par de bases 1831, una secuencia que es idéntica por lo menos 70% a la SEC ID nº: 8, o una secuencia que hibrida con la SEC ID nº: 8 en condiciones severas, o una secuencia que ha degenerado a la SEC ID nº: 8 debido al código genético mediante la introducción de un vector o una molécula de ADN que comprende dicha secuencia con dicha mutación por delección, inserción y/o sustitución del gen de  $\beta$ 1,2-xilosiltransferasa en dichas plantas o células de manera que se produce la recombinación homóloga con el genoma de la planta o célula vegetal.
- 15 3. Método según la reivindicación 2, caracterizado porque la secuencia codificante es una secuencia con una identidad de por lo menos 70% con la secuencia según la SEC ID nº: 8.
- 20 4. Método según cualquiera de las reivindicaciones 1 a 3, caracterizado porque la secuencia codificante es una secuencia con una identidad de por lo menos 80% con la secuencia según la SEC ID nº: 8.
- 25 5. Método según cualquiera de las reivindicaciones 1 a 4, caracterizado porque la secuencia codificante es una secuencia con una identidad de por lo menos 95% con la secuencia según la SEC ID nº: 8.
- 30 6. Método según cualquiera de las reivindicaciones 1 a 5, caracterizado porque la glicoproteína es una proteína humana, especialmente una proteína humana para uso médico.
- 35

Péptido 3 de la patente WO 99/29835 A1      **SQVQAIHDASVIIGAHGAGLTHIVSAL**

Péptido 2 de la patente WO 99/29835 A1<sup>1</sup>      **GLEYHAIN**

Soja:	<b>SQVQAIHDASVIIGAHGAGLTHIVSAL</b>	<b>GLEYHAIN</b>
	<b>QV AI DASVIIGAHGAGLTHIVSA</b>	<b>GLEYHA</b>
A.thaliana:	<b>DQVRAIQDASVIIGAHGAGLTHIVSATPNTTIFELISVEFQRPHFELIAKWKGLEYHAMH</b>	

Cebador 1	<b>5'-ATGAGTAAACGGAATCCGAAG-3'</b>
Cebador 2	<b>5'-TTAGCAGCCAAGGCTTTCAT-3'</b>
Cebador 3	<b>5'-GATCAAGTCCGAGCCATTCAA-3'</b>
Cebador 4	<b>5'-CGCGTGATACTCCAATCCCTT-3'</b>

<sup>1</sup>Se omitieron los aminoácidos LG C-terminales

Fig. 1

AAATCTGCAGACTCTCAAAATTCCGATTCACTCTTATTGAAGAACAA	46
TTTCCGGCGAAACAGCCGATGAAGTCTCGCTGAATCTTCTGTACCTTCACCGGCGAT	106
TGACTTCACCTCAGAATCGAGAGAGAAGAAATCGATGGAAAACAAAAATAGAAAGAGTT	166
TCAAATTCTCGCTCTCTCAAAACCGCAAATCAAGGGAACGAGAGACGAGAGAGAG	226
ATGAGTAAACGGAATCCGAAGATTCTGAAGATTTCTGTATATGTTACTCTCAACTCT	286
CTCTTCTCATCATCTACTCGTTTCACTCATCGTCGTTTACCGGAGCAGTCACAG	346
CCTCCTCATATATACCACGTTCACTGAATAACCAATCGGCATTCAAGAAACCGTGGCCG	406
ATCTTACCTTCTTACCTCCCATTGGACGCCGCCAGAGGAATCTACCAACTGGCTCCTGC	466
GAAGGTACTTCGGAATGGATTACAAAGAGAGTTGACTTCCTTAAGCCGAGGATTGGA	526
GGAGGAGGAGAAGGAAGCTGGTTCCGATGTTTACAGTGAGACATTACAGAGTTCGATT	586
TGTGAAGGAAGGAATCTGAGAATGGTCCGGATCGGATTGTTATGTCGAGAGGAGGTGAG	646
AAGTTAGAGGAAGTTATGGGGAGGAAAGAGGAGGAGGAGCTTCTGCCTCGACAAGGT	706
CGCTTGAGGTACCGGAAGAGGTTTCTCACGGTTAGGTTTAAGAGACACCGTCGTTT	766
GGTGGAGGAGAAGGAGGTAGTGCGGTTCTGGCGGCTGGTGAATGATGAGATGTTGAAT	826
GAATATATGCAAGAAGGTGGAATTGATAGACATACAATGAGAGATTGGTGCTTCGATT	886
CGTGCTGTTGATACCAATGATTCTGTTGTGAAGAGTGGTGGAGGAACCGACCTGCTT	946
GTCACTAGATTCCGAGTACGCAAATCTCTCCATACTGTGACAGATTGGTATAGTGCCTAT	1006
GTTCGCTAGAGTCACCGGTTGCCTAATCGACCTCACGTTGTTTGTGACGGACAC	1066
TGCACGACGCAGCTAGAAGAACATGGACAGCTTGGTCCCGGAATCAGATACGCAAAG	1126
AACTTCACCAAACCGGTTGTTCCGCCACCGATTCTTCAACCATTGGGATACGAAACC	1186
GCTTTAAAGGCTTGTCCGGAGAAATAGACTGCAAGGGAGATTCAAGCTCACAACTG	1246
TGGCAAAACCGGACGATAAAAGGACTGCGAGGATATCAGAGTTGGTGAATGATCAGA	1306
GCAGCTTCGGGTTGCCTGTCAATAGACACCGCTCATTAGAAAAGCCGCTATCATCATCA	1366
TCATCATCAGCCTCAGTTATAATGTTCTTTGTCCGCCGTGAAGATTACTTAGCCCAT	1426
CCTCGTCATGGCGGTAAAGTCCAGTCTCGGCTCATCAATGAGGAAGAAGTGGTCACTCG	1486
TTGCATCATTGGGTTGCAACTGGTCCACCGGTCTGACCAATGCGGGATTAACTTGTG	1546
AATGGCTTGCTTGCACACATGTCAATGAAAGATCAAGTCCGAGCCATTCAAGATGCTTCA	1606
GTGATCATAGGAGCTCATGGAGCAGGACTGACTCACATTGTCTTGCAACACCAAACACA	1666
ACGATATTGAGATAATAAGCGTCGAGTTCAAGAGACCTCATTCGAGCTTATAGCTAAG	1726
TGGAAAGGATTGGAGTATCACGCGATGCATCTGGCGACTCACGAGCGGAACCAACGGCT	1786
GTGATTGAGAAGTTAACGGAGATCATGAAGAGCCTTGGCTGCTAA	1831

Fig. 2

10	20	30	40	50	60
MSKRNPKILKIFLYMLLNSLFLIIYFVFHSSFSPEQSQPPHIYHVSVNNQSAIQKPWP					
70	80	90	100	110	120
ILPSYLPWTPPQRNLPTGSCEGYFGNGFTKRVDFLKPRIGGGEGSWFRCFYSETLOSSI					
130	140	150	160	170	180
CEGRNLRMVPDRIVMSRGGEKLEEVAGRKEEEELPAFRQGAFEVAEEVSSRLGFKRHRRF					
190	200	210	220	230	240
GGGEGGSAVSRLVNDMLNEYMQEGGIDRHTMRDLVASIRAVDTNDFVCEEWVEEPTLL					
250	260	270	280	290	300
VTRFEYANLFHTVTDWYSAYVSSRVTGLPNRPHVVFDGHCTTQLEETWTALFSGIRYAK					
310	320	330	340	350	360
NFTKPVCFRHAILSPGYETALFKGLSGEIDCKGDSAHNLWQNPDDKRTARISEFGEMIR					
370	380	390	400	410	420
AAFGLPVNRHRSLEKPSSSSSASVYNVLFVRREDYLAHPRHGGKVQSRLINEEEVFDS					
430	440	450	460	470	480
LHHWVATGSTGLTKCGINLVNGLLAHMSMKDQVRAIQDASVIIGAHGAGLTHIVSATPNT					
490	500	510	520	530	
TIFEIIISVEFQRPHFELIAWKGLEYHAMHLANSRAEPTAVIEKLTEIMKSLGC					

Fig. 3

xt-Ath6.seq	:	ATGAGTAAACGGAATCCGAAGATTCTGAAGATTTCTGTATATGTTACTCTCACTCTCTCTCATCATCTACTT	:	80
xt-Ath9.seq	:	.....	:	80
xt-Ath16.seq	:	.....	:	80
xt-Athgen.seq	:	.....	:	80
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	CGTTTTCACTCATCGTCGTTTCACGGAGCAGTCACAGCCTCTCATATATACCACTGTTCACTGAATAACCAATCGG	:	160
xt-Ath9.seq	:	.....A.....	:	160
xt-Ath16.seq	:	.....	:	160
xt-Athgen.seq	:	.....	:	160
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	CGATTCAAGAAACGGTGGCGATCTTACCTCTTACCTCCATGGACGCCGCCAGAGGAATCTACCAACTGGCTCTGC	:	240
xt-Ath9.seq	:	.....	:	240
xt-Ath16.seq	:	.....	:	240
xt-Athgen.seq	:	.....	:	240
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	GAAGGTTACTTCGGGAATGGATTACAAAGAGAGTTGACTTCCTTAAGCCGAGGATTGGAGGGAGGAGAGGARGCTG	:	320
xt-Ath9.seq	:	.....	:	320
xt-Ath16.seq	:	.....T.....	:	320
xt-Athgen.seq	:	.....	:	320
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	GTTCGGATGTTTACAGTGAGACATTACAGAGTTGATTTGTGAGGAAGGAATCTGAGAATGGTCCCGATGGATTC	:	400
xt-Ath9.seq	:	.....	:	400
xt-Ath16.seq	:	.....	:	400
xt-Athgen.seq	:	.....	:	400
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	TTATGCGAGAGGGGTGAGAAGTTAGAGGAGTTATGGGGAGGAAAGAGGGAGGAGCTTCTGGCTTTCGACAAAGT	:	480
xt-Ath9.seq	:	.....	:	480
xt-Ath16.seq	:	.....A.....	:	480
xt-Athgen.seq	:	.....	:	480
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	GGGTGGAGGTACCGGAAGAGGTTCTTACGGTTAGGTAAAGAGACACCCGTCGTTTGGTGGAGGAAGGGAGTAG	:	560
xt-Ath9.seq	:	.....	:	560
xt-Ath16.seq	:	.....	:	560
xt-Athgen.seq	:	.....	:	560
xt-AthEST.seq	:	.....	:	-

Fig. 4a

xt-Ath6.seq	:	TCGGCTTCTCGGCGGCTGGTGAATGATGAGATGTTGAATGAATATGCAAGAAGGTGAAATTGATAGACATACAAATGA	:	640
xt-Ath9.seq	:	.....	:	640
xt-Ath16.seq	:	.....C.....	:	640
xt-Athgen.seq	:	.....	:	640
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	GAGATTTGGTTGCTTCGATTGGCTGTTGATACCAATGATTTGGTTGCAAGAGACTGGGTGGAGGAACCGACCTTGCTT	:	720
xt-Ath9.seq	:	.....	:	720
xt-Ath16.seq	:	.....	:	720
xt-Athgen.seq	:	.....	:	720
xt-AthEST.seq	:	.....	:	72
xt-Ath6.seq	:	GTCACTAGATTGGAGTACGCCAATCTCTTCCATACTGTGACAGATGGTATAGTGCCTATGTTGGCTAGAGTCACCGG	:	800
xt-Ath9.seq	:	.....	:	800
xt-Ath16.seq	:	.....A.....	:	800
xt-Athgen.seq	:	.....	:	800
xt-AthEST.seq	:	.....	:	152
xt-Ath6.seq	:	TTTGCCTAATCGACCTCACGGTTGTTTGGTACGGACACTGCACGACGGCAGCTAGAAGAAACATGGACAGCTTGT	:	880
xt-Ath9.seq	:	.....	:	880
xt-Ath16.seq	:	.....	:	880
xt-Athgen.seq	:	.....	:	880
xt-AthEST.seq	:	.....	:	232
xt-Ath6.seq	:	CCGGAAATCAGATACGCCAAGAACTTCACCAACCCGTTTGTGTTCCGCACGGATTCTTCACCAATTGGATACCAAAACC	:	960
xt-Ath9.seq	:	.....A.....	:	960
xt-Ath16.seq	:	.....G.....	:	960
xt-Athgen.seq	:	.....	:	960
xt-AthEST.seq	:	.....	:	312
xt-Ath6.seq	:	GCTCTTTAAAGGCTTGCCGGAGAAATAGACTGCAAGGGAGATTGAGCTACAATCTGTGCCAAAACCCGGACGATAA	:	1040
xt-Ath9.seq	:	.....G.....	:	1040
xt-Ath16.seq	:	.....	:	1040
xt-Athgen.seq	:	.....C.....	:	1040
xt-AthEST.seq	:	.....	:	392
xt-Ath6.seq	:	AAGGACTCCGAGGGATATCAGAGTTGGTGAATGATCAGAGCAGCTTGGGTTGCCGTCAATAGACACCGCTCATTAG	:	1120
xt-Ath9.seq	:	.....	:	1120
xt-Ath16.seq	:	.....	:	1120
xt-Athgen.seq	:	.....C....G.....GTC....TA.....C	:	1120
xt-AthEST.seq	:	.....	:	472

Fig. 4b

```

xt-Ath6.seq : AAAAGCCGCTATCATCATCATCATCAGCCTCAGTTATAATGTTCTTTGTCCGCCGTAAAGATTACTTAGCCCAT : 1200
xt-Ath9.seq : .....G..... : 1200
xt-Ath16.seq : ..... : 1200
xt-Athgen.seq : ..... : 1200
xt-AthEST.seq : G..... : 482

xt-Ath6.seq : CCTCGTCATGGGGTAAAGTCCAGTCGGTCATCAATGAGGAAGAAGTGTTCGACTCGTTGCATCATGGGTTGCAAC : 1280
xt-Ath9.seq : ..... : 1280
xt-Ath16.seq : ..... : 1280
xt-Athgen.seq : ..... : 1280
xt-AthEST.seq : ..... : -

xt-Ath6.seq : TGGGTCCACCGGTCTGACCAAATGGGGATTAACCTTGTAATGGCTTGCAACACATGTCAATGAAAGATCAAGTCC : 1360
xt-Ath9.seq : ..... : 1360
xt-Ath16.seq : ..... : 1360
xt-Athgen.seq : ..... : 1360
xt-AthEST.seq : ..... : -

xt-Ath6.seq : GAGCCATTCAAGATGCTTCAGTGTACAGGAGCTATGGAGCAGGACTGACTCACATTGCTCTGCAACACCAACACA : 1440
xt-Ath9.seq : ..... : 1440
xt-Ath16.seq : ..... : 1440
xt-Athgen.seq : ..... : 1440
xt-AthEST.seq : ..... : -

xt-Ath6.seq : ACCGATTTGAGATAATAAGCGTCAGTTCAAGAGACCTCATTGAGCTTATAGCTAAGTGGAAAGGATTGGAGTATCA : 1520
xt-Ath9.seq : ..... : 1520
xt-Ath16.seq : ..... : 1520
xt-Athgen.seq : ..... : 1520
xt-AthEST.seq : ..... : -

xt-Ath6.seq : CGCGATGCATCTGGCGAACTCACGAGCGAACCAACGGCTGTGATTGAGAAGTTACGGAGATCATGAAGAGCCTTGGCT : 1600
xt-Ath9.seq : ..... : 1600
xt-Ath16.seq : .....G..... : 1600
xt-Athgen.seq : ..... : 1600
xt-AthEST.seq : ..... : -

xt-Ath6.seq : GCTAA : 1605
xt-Ath9.seq : ..... : 1605
xt-Ath16.seq : ..... : 1605
xt-Athgen.seq : ..... : 1605
xt-AthEST.seq : ----- : -

```

Fig. 4c

xt-Ath6.seq	:	M3KRNPKILKIFLYHLLLN3LFLIIYFVFHSSFSPEQSOPPHIYHVSVNQSAIQKPUPILPSYLPSTPPQRNLPTGSC	:	80
xt-Ath9.seq	:	.....	:	80
xt-Ath16.seq	:	.....	:	80
xt-Athgen.seq	:	.....	:	80
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	EGYFGNGFTKRVDFLKPRIGGGEGSWFRCFYSETLOSSICECRNLRREVPRIVMSRGGEKLEEVNGRKEEEELPAFQG	:	160
xt-Ath9.seq	:	.....	:	160
xt-Ath16.seq	:	.....L.....	:	160
xt-Athgen.seq	:	.....	:	160
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	APEVAEEVSSRLGFRRRRFGGEGGSAUSRRLVNDENLNEYHQEGGIDRHTRDLVASIRAVDTNDFVKEEVVEPTLL	:	240
xt-Ath9.seq	:	.....	:	240
xt-Ath16.seq	:	.....	:	240
xt-Athgen.seq	:	.....	:	240
xt-AthEST.seq	:	.....	:	24
xt-Ath6.seq	:	VTRFETANLFHTVTDUTSAYVSSRVTGLPNRPHVVFDGHCTTQLEETUTALFSGIRYAKNFTKPVCFRHAXLSPGYET	:	320
xt-Ath9.seq	:	.....	:	320
xt-Ath16.seq	:	.....K.....	:	320
xt-Athgen.seq	:	.....	:	320
xt-AthEST.seq	:	.....	:	104
xt-Ath6.seq	:	ALFKGLSGEIDCKGDSANHLWQNPDDKRTARISEFGEMIRAAFLPVRHRSLEKPLSSSSSSASVYNVLFVRREDYLAH	:	400
xt-Ath9.seq	:	.....R.....	:	400
xt-Ath16.seq	:	.....	:	400
xt-Athgen.seq	:	.....	:	400
xt-AthEST.seq	:	.....R.....R.....GHSN...R.....	:	160
xt-Ath6.seq	:	PRHGGKVQSLRINEEVYDSLHHWATGCTGLTKCGINLVNGLLAHNSNMDQVRAIQDASVIIGAAGLTHIVSATPNT	:	480
xt-Ath9.seq	:	.....	:	480
xt-Ath16.seq	:	.....	:	480
xt-Athgen.seq	:	.....	:	480
xt-AthEST.seq	:	.....	:	-
xt-Ath6.seq	:	TIFETISVEFQRPHFELIAKUKGLEYHAMHLANSRAEPTAVIERLTEIMKSLGC	:	534
xt-Ath9.seq	:	.....	:	534
xt-Ath16.seq	:	.....	:	534
xt-Athgen.seq	:	.....	:	534
xt-AthEST.seq	:	.....	:	-

Fig. 5

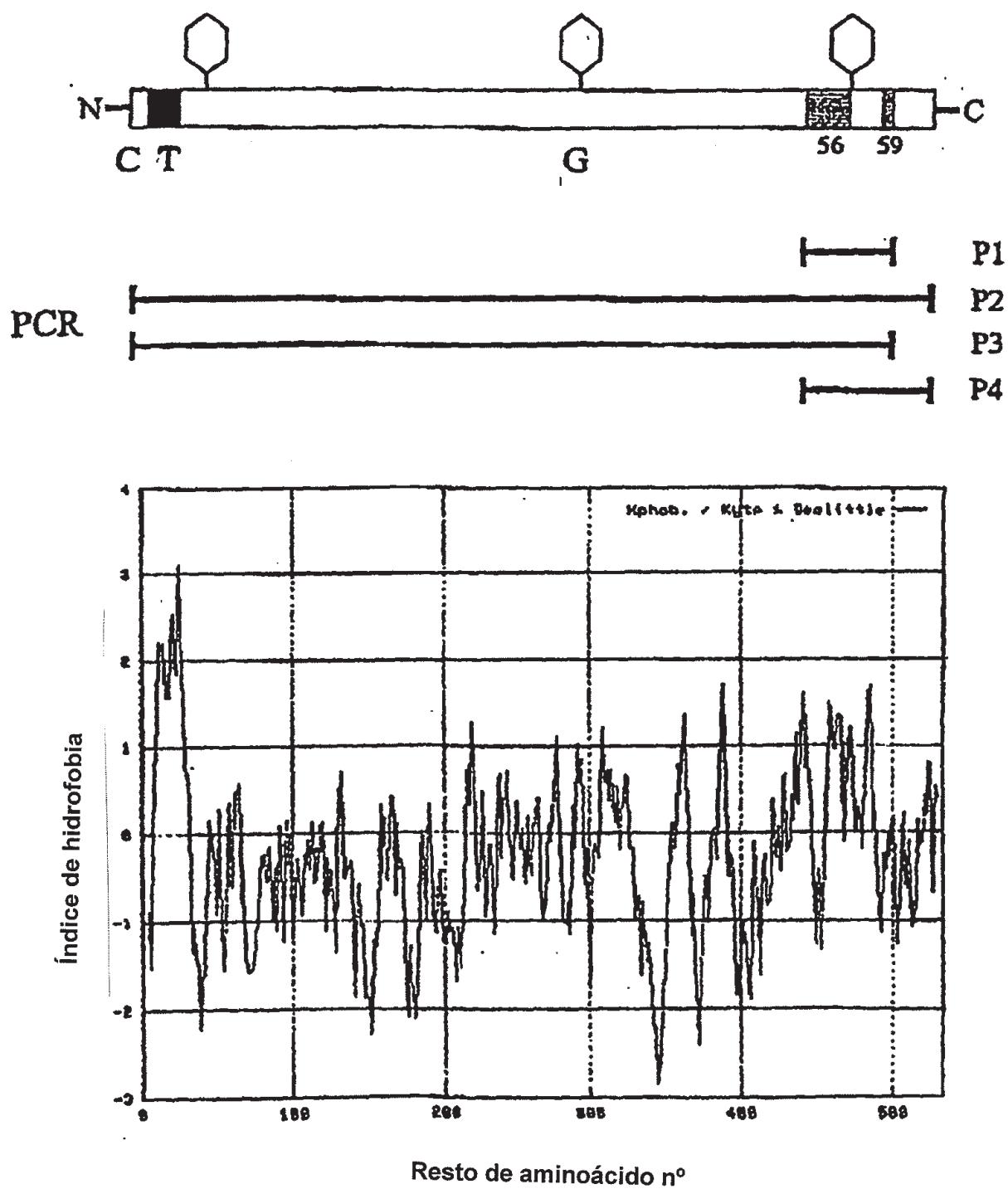


Fig. 6

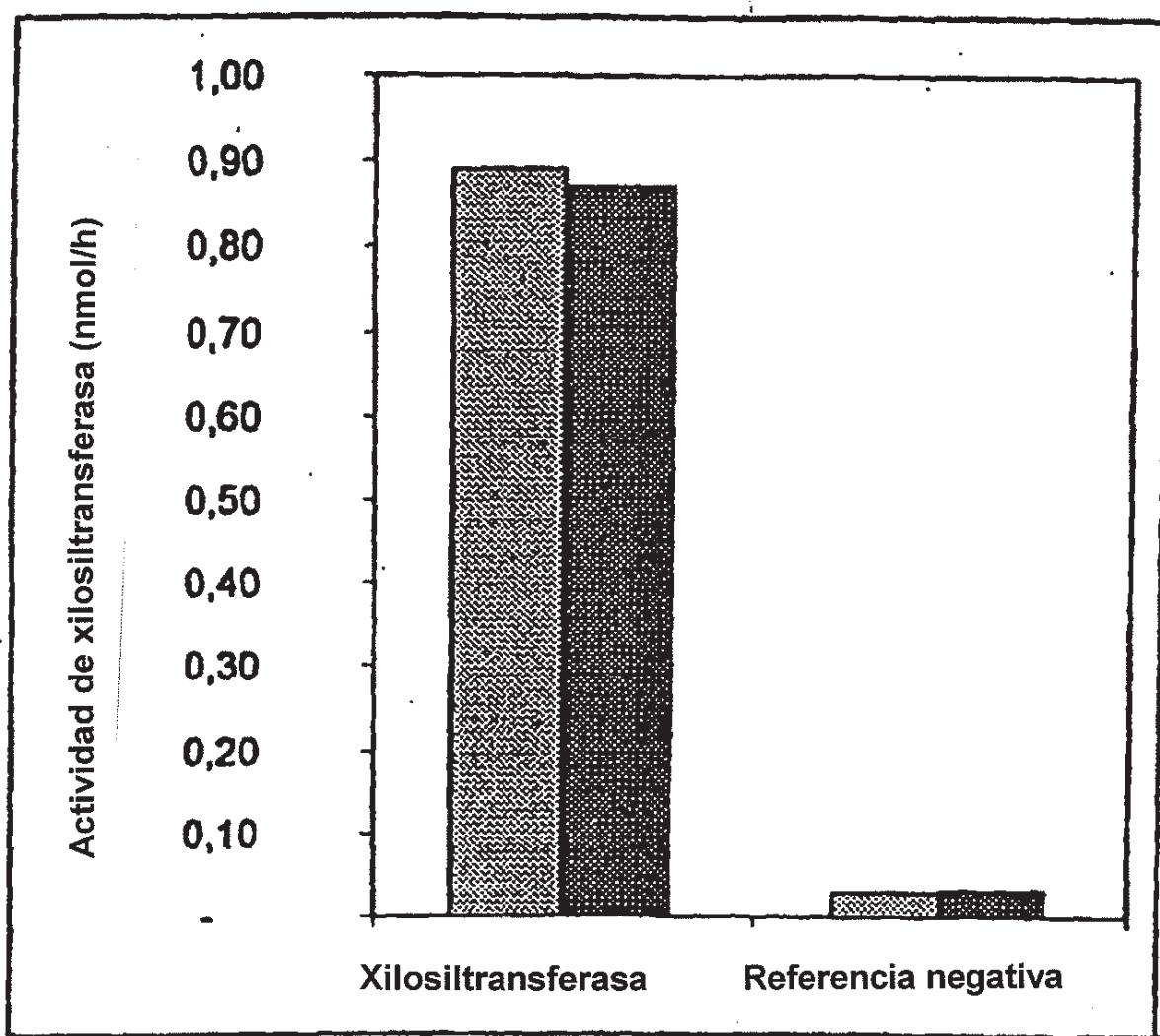


Fig. 7

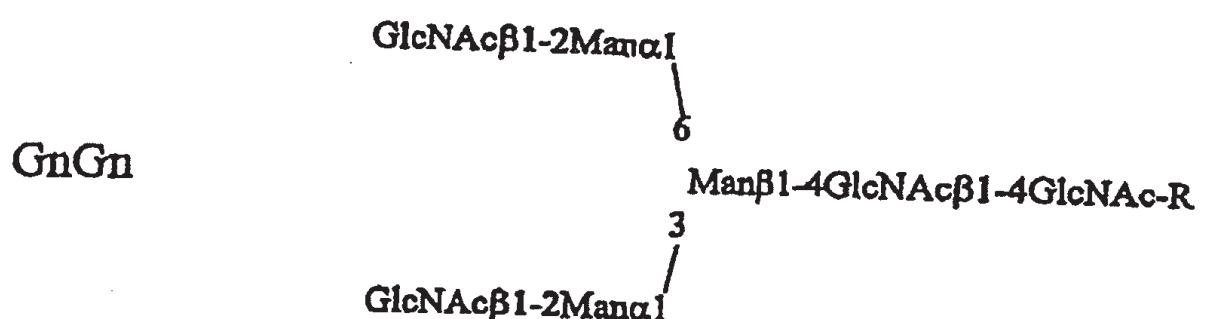


Fig. 8a

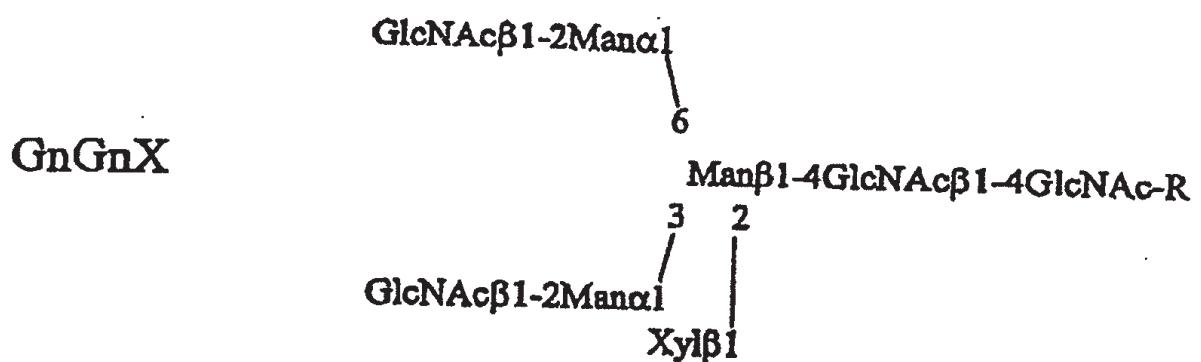
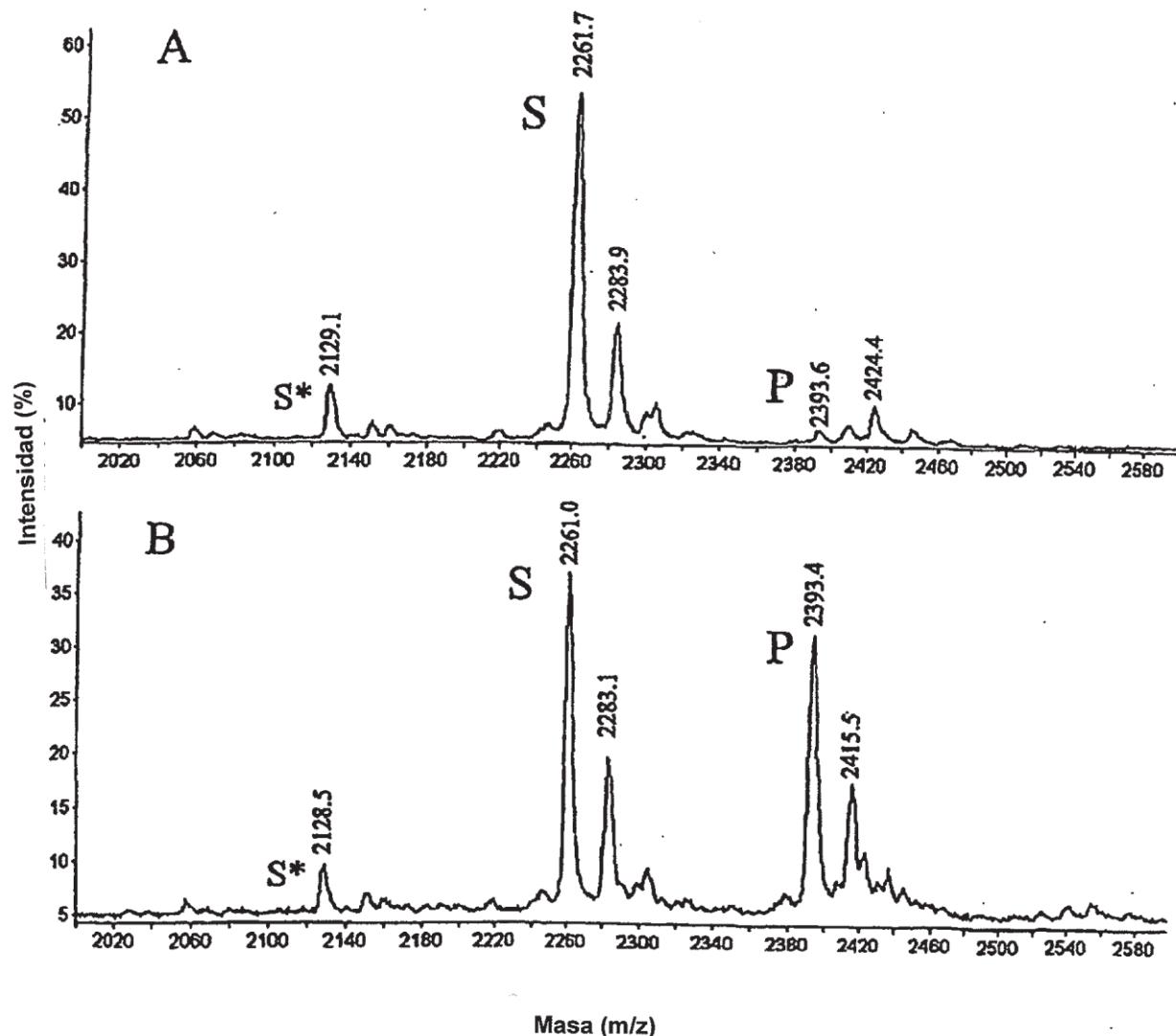


Fig. 8b



Masa (m/z)

Fig. 9

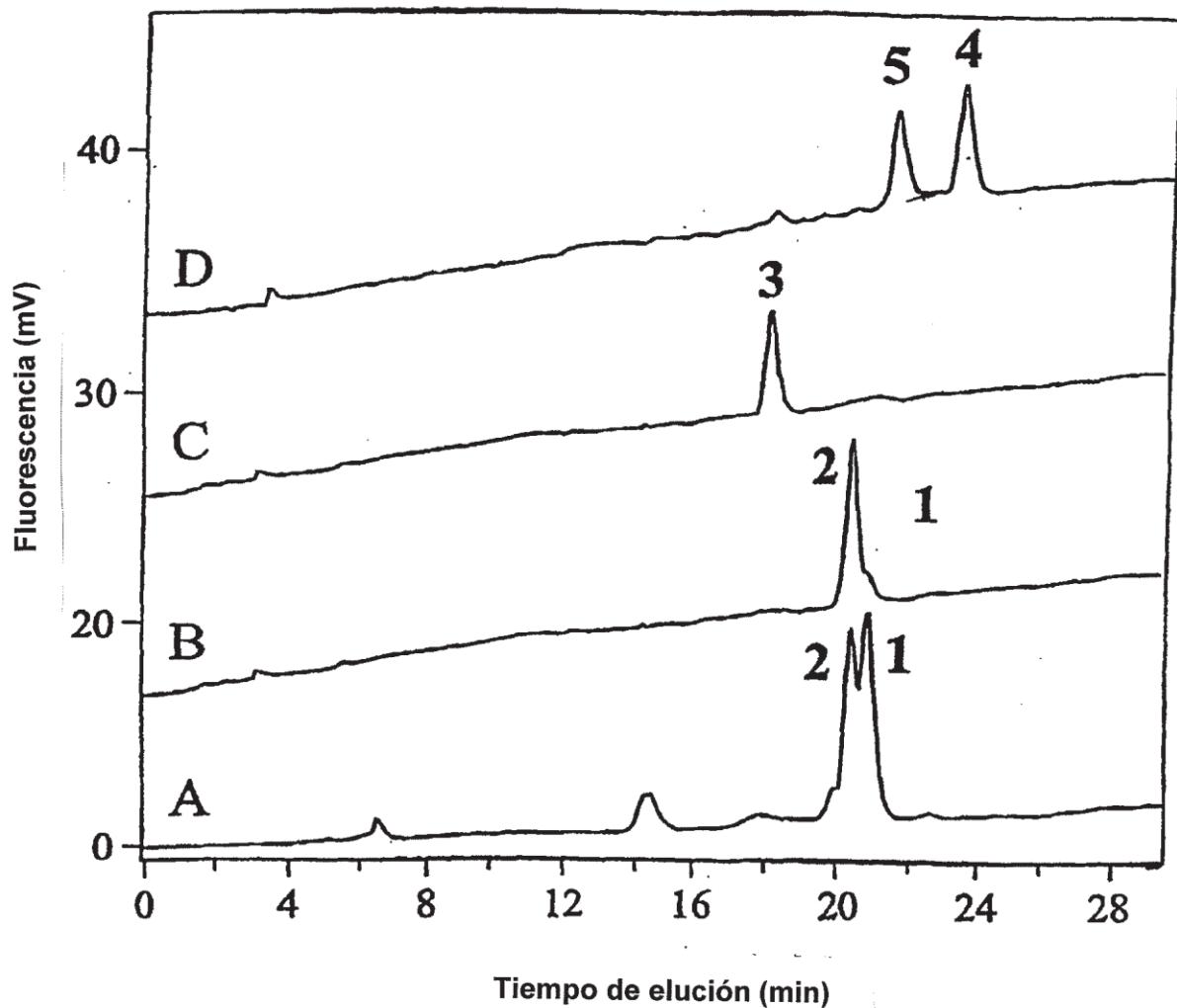


Fig. 10

A. thaliana de soja	: RXXXXXXXXXXXXXXXXXXXXXXSFLIIYVHSSSFSPSEQSQQPPHIYHVSNNQSAIQKPWPLIPSYLPWTTPQRLPTGSCE	: 81
A. thaliana de soja	: XXXXXXXXXXXXXXXXXXXXXXXXXSFLIIYVHSSSFSPSEQSQQPPHIYHVSNNQSAIQKPWPLIPSYLPWTTPQRLPTGSCE	: 162
A. thaliana de soja	: GFNGNGFTKRVDFLKPRIGGGEGSWFRCFYSETLQSSICEGRNLRMVPDRIVMSRGKKEKLEEVMGKREEEEELPAFRQGAF	: 162
A. thaliana de soja	: XXXXXXXXXXXXXXXXXXXXXXXXXSFLIIYVHSSSFSPSEQSQQPPHIYHVSNNQSAIQKPWPLIPSYLPWTTPQRLPTGSCE	: 242
A. thaliana de soja	: EVAEEVSSRLGFKRRREFGGEGGSAVSRRLVNDEMNLNEYMQEGGIDRHTMRDLVASIRAVDTNDFEVCEENVEEPTLLVTR	: 243
A. thaliana de soja	: FEYANLFHTVTDWYSAYVSSRVTALPNRPHVIFVDGHCKAPLEETWKALFSSVGYAKSFSGSVCFHHAILSPLGYETAMFR	: 323
A. thaliana de soja	: FEYANLFHTVTDWYSAYVSSRVTGLPNRPHVVFVDGHCTTQLEETWTALFGIRYAKNFTKPVCRHAILSPLGYETALFK	: 324
A. thaliana de soja	: GISEHIDCYGAPAQELLQNLDHKTARLSEEGEMVRAAFGLPLNVNHNXMEXXXXPLAGHNVLFVRRREDYLAHPRHSG	: 404
A. thaliana de soja	: GISEHIDCKGDSAHNLWQNPDDKRTARISEEGEMIRAAFGQPLPNRHRSLEKPLSSSSSSASVYNNLTVRRREDYLAHPRHGG	: 405
A. thaliana de soja	: KLESRTISNEQEVFNSLKSXXXXSNSYKGGKINLNGLFAHNSMKDQVQAIHDA SVLIGAHGAG-----	: 468
A. thaliana de soja	: KVQSRLINEEEVEDSILHHWVATGSTGLTKGGINLYNGLAHNSMKDQVRAIQDASVIGAHLTHIVSATPNTTIEII	: 486
A. thaliana de soja	: SVEFQRPHFELIAKWKGLLEYHAMHLANSRAEPTAVIEKLTEIMKSLGC-	: 534

Fig. 11