

(11) Número de Publicação: **PT 2062977 E**

(51) Classificação Internacional:

**C12P 21/00** (2013.01) **C12N 15/54** (2013.01)  
**C12N 9/10** (2013.01)

**(12) FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **2000.02.17**

(30) Prioridade(s): **1999.02.18 AT 27099**

(43) Data de publicação do pedido: **2009.05.27**

(45) Data e BPI da concessão: **2014.01.08**  
**068/2014**

(73) Titular(es):

**FRIEDRICH ALTMANN**  
**INSTITUT FUR CHEMIE, UNIVERSITAT FUR**  
**BODENKULTUR, MUTHGASSE 18 1190 WIENAT**

(72) Inventor(es):

**FRIEDRICH ALTMANN** AT

(74) Mandatário:

**ALBERTO HERMÍNIO MANIQUE CANELAS**  
**RUA VÍCTOR CORDON, 14 1249-103 LISBOA** PT

(54) Epígrafe: **GENE DA FUCOSILTRANSFERASE**

(57) Resumo:

É COLOCADA À DISPOSIÇÃO UMA MOLÉCULA DE DNA, QUE ABRANGE UMA SEQUÊNCIA DE ACORDO COM A SEQ ID Nº: 1 COM UMA GRELHA DE LEITURA ABERTA DO PAR DE BASES 211 AO PAR DE BASES 1740 OU APRESENTA PELO MENOS 50% DE HOMOLOGIA COM A SEQUÊNCIA SUPRA CITADA OU HIBRIDIZA SOB CONDIÇÕES RIGOROSAS COM A SEQUÊNCIA SUPRA CITADA OU ABRANGE UMA SEQUÊNCIA, QUE É DEGENERADA COMO RESULTADO DO CÓDIGO GENÉTICO PARA A SEQUÊNCIA DE DNA SUPRA CITADA, EM QUE A SEQUÊNCIA CODIFICA UMA PROTEÍNA DE PLANTA COM UMA ACTIVIDADE FUCOSILTRANSFERASE OU É COMPLEMENTAR A ELA.

**RESUMO****"GENE DA FUCOSILTRANSFERASE"**

É colocada à disposição uma molécula de DNA, que abrange uma sequência de acordo com a SEQ ID N°: 1 com uma grelha de leitura aberta do par de bases 211 ao par de bases 1740 ou apresenta pelo menos 50% de homologia com a sequência supra citada ou hibridiza sob condições rigorosas com a sequência supra citada ou abrange uma sequência, que é degenerada como resultado do código genético para a sequência de DNA supra citada, em que a sequência codifica uma proteína de planta com uma actividade fucosiltransferase ou é complementar a ela.

**DESCRIÇÃO****"GENE DA FUCOSILTRANSFERASE"**

A invenção refere-se a polinucleótidos, que codificam uma fucosiltransferase. Além disso a invenção refere-se a sequências parciais destes polinucleótidos assim como vectores com estes polinucleótidos, células hospedeiras, plantas e insectos recombinantes, que estão transfetados com os polinucleótidos ou o DNA derivado destes, assim como as glicoproteínas produzidas nestes sistemas.

As glicoproteínas apresentam uma variedade e complexidade de unidades de hidratos de carbono, em que a composição e disposição dos hidratos de carbono é característica para diferentes organismos. As unidades de oligossacáridos das glicoproteínas têm uma série de funções, assim elas são p.ex. importantes na regulação do metabolismo, estão envolvidas na mediação de interações célula-célula, elas determinam os tempos de circulação de proteínas na circulação sanguínea e elas são cruciais para o reconhecimento de epítopos nas reações antigénio-anticorpo.

A glicosilação de glicoproteínas começa no retículo endoplasmático (ER), onde os oligossacáridos são

ligados ou através de ligações N-glicosídicas às cadeias laterais da asparagina ou através de ligações O-glicosídicas às cadeias laterais da serina ou treonina. Os oligossacáridos N-ligados contêm um núcleo comum de uma unidade pentassacárido, que é constituída por três resíduos de manose e dois de N-acetilglucosamina. Para a modificação ulterior das unidades de hidratos de carbono, as proteínas são transportadas do ER para o complexo de Golgi. A estrutura das unidades de oligossacáridos N-ligados de glicoproteínas é determinada pela sua conformação e a composição das glicosiltransferases dos compartimentos de Golgi, nos quais elas são processadas.

Foi mostrado, que no complexo de Golgi de algumas células de plantas e de insectos a unidade núcleo-pentassacárido é substituída por xilose e fucose  $\alpha$ 1,3-ligada (P. Lerouge et al. 1998 Plant Mol. Biol. 38, 31-48; Rayon et al. 1998 L. Exp. Bot. 49, 1463-1472). O heptassacárido resultante "MMXF<sup>3</sup>" representa o tipo de oligossacárido principal em plantas (Kurosaka et al. 1991 J. Biol. Chem., 266, 4168-4172). Assim p.ex. a peroxidase de rábano,  $\beta$ -fructosidase da cenoura e lectina da *Erythrina cristagalli* assim como a fosfolipase A2 do veneno das abelhas produtoras de mel ou as glicoproteínas da membrana neuronal de embriões de insectos apresentam resíduos  $\alpha$ 1,3-fucose, que estão ligados ao núcleo de glicanos. Estas estruturas são também designadas de N-glicanos complexos ou N-glicanos deficientes em manose ou truncados. Os resíduos  $\alpha$ -manosil podem ser além disso ser substituídos por GlcNac,

ao qual estão ligadas galactose e fucose, de forma que é produzida uma estrutura que corresponde ao epítopo humano Lewis a (Melo et al. 1997, FEBS Lett. 415, 186-191; Fitchette-Laine et al. 1997 Plant J. 12, 1411-1417).

Nem a xilose nem a fucose  $\alpha$ 1,3-ligada ocorrem em glicoproteínas de mamíferos. Verificou-se que a núcleo- $\alpha$ 1,3-fucose tem um papel importante no reconhecimento do epítopo por anticorpos, que são dirigidos contra oligossacáridos N-ligados de plantas e de insectos (I.B.H. Wilson et al., Glycobiology vol. 8, nº 7, pág. 651-661, 1998) e desta forma desencadeia reações imunitárias no corpo humano ou animal contra estes oligossacáridos. O resíduo  $\alpha$ 1,3-Fucose parece além disso ser uma das causas principais da reactividade alérgica cruzada generalizada entre diferentes alergénios de plantas e de insectos (Tretter et al., Int. Arch. Allergy Immunol. 1993; 102:259-266) e é designado também de "determinante de reatividade cruzada de hidratos de carbono" (CCD). Num estudo de epítopos de tomates e pólens de gramíneas foi também verificado resíduos de fucose  $\alpha$ 1,3-ligado como determinante comum, o que parece ser a razão para o frequente surgimento comum de alergias em doentes aos tomates e a pólens de gramíneas (Peterson et al. 1996, J. Allergy Clin. Immunol., vol. 98, 4; 805-814). Os CCDs dissimulam além disso os diagnósticos de alergias através do aparecimento frequente de reatividade imunitária cruzada.

Estas reacções imunitárias, que são desencadeadas

por proteínas de plantas no corpo humano, são o problema principal na utilização médica de proteínas humanas recombinantes, produzidas em plantas. Para evitar este problema, teria de ser impedida a  $\alpha$ 1,3 núcleo-fucosilação. Pôde ser mostrado num estudo, que oligossacáridos que em vez de uma L-fucose (6-deoxi-L-galactose) apresentam uma L-galactose, apesar disso são totalmente biologicamente activos (E. Zabrackis et al. 1996, *Science*, vol 272). Segundo um outro estudo foi isolado um mutante da planta *Arabidopsis thaliana*, ao qual falta a N-acetyl-glucosaminil-transferase I, a primeira enzima na biossíntese de glicanos complexos. A biossíntese de glicoproteínas complexas nestes mutantes é assim perturbada. Contudo estas plantas mutantes têm a capacidade de se desenvolverem normalmente sob determinadas condições (A. Schaewen et al. 1993, *Plant Physiol.* 102; 1109-1118).

De forma a impedir especificamente a ligação do núcleo- $\alpha$ 1,3 fucose num oligossacárido, sem também intervir em outros passos de glicosilação, teria de ser desactivada apenas a enzima que é responsável por esta glicosilação específica, nomeadamente a núcleo- $\alpha$ 1,3, fucosiltransferase. Ela foi pela primeira vez isolada e caracterizada a partir do feijão mungo, em que foi verificado, que a actividade desta enzima é dependente da presença de GlcNAc-terminais não redutores (Staudacher et al., 1995, *Glycoconjugate J.* 12, 780-786). Esta transferase, que apenas existe em plantas e insectos, mas não em humanos ou outros vertebrados, teria de ser especificamente inactivada ou

reprimida, de forma que proteínas humanas, que são produzidas em plantas ou células de plantas ou também insectos ou células de insectos, não apresentem mais como até agora este epítopo que desencadeia a reacção imunitária.

A publicação de John M. Burke "Clearing the way for ribozymes" (Nature Biotechnology 15: 414-415; 1997) diz respeito ao funcionamento geral de ribozimas.

A publicação de Pooga et al "Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo" (Nature Biotechnology 16: 857-861; 1998) diz respeito a moléculas PNA em geral e especificamente a uma molécula PNA, que é complementar ao mRNA do receptor da galanina do tipo 1 humano.

A US 5 272 066 A diz respeito a um processo para alteração de proteínas eucarióticas e procarióticas, para aumentar a sua circulação *in vivo*. Neste caso os oligossacáridos ligados são alterados com a ajuda de diferentes enzimas, entre elas também a GlcNAc- $\alpha 1 \rightarrow 3(4)$ -fucosiltransferase.

A EP 0 643 132 A1 diz respeito à clonagem de uma  $\alpha 1,3$ -fucosiltransferase, que é isolada de células humanas (THP-1). As cadeias de hidrocarbonetos descritas nesta publicação correspondem aos sialil Lewis x- e sialil Lewis a-oligossacáridos. A especificidade da enzima de células

humanas é uma muito diferente da fucosiltransferase de células de plantas.

Outras fucosiltransferases animais estão descritas em Costache et al (J. Biol. Chem. 272(47), (1997) : 29721-29728).

É um objectivo da presente invenção, clonar e sequenciar o gene que codifica a fucosiltransferase de plantas, assim como colocar à disposição processos para produzir glicoproteínas, que não abrangem a  $\alpha$ 1,3-núcleo fucose que ocorre naturalmente.

O objectivo de acordo com a invenção é resolvido através de um processo para produção de glicoproteínas recombinantes, que abrangem a produção de uma glicoproteína recombinante em plantas ou células de plantas, cuja produção de GlcNAc- $\alpha$ 1,3-fucosiltransferase está reprimida ou completamente impedida e cuja actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase endógena é inferior a 50% da actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase que ocorre naturalmente em plantas ou células de plantas, e a GlcNAc- $\alpha$ 1,3-fucosiltransferase que é codificada através de uma molécula de DNA, que abrange uma sequência de acordo com a SEQ ID n°.1 (nesta publicação foi utilizado o código IUPAC, em que "N" representa inosina) com uma grelha de leitura aberta do par de bases 211 até ao par de bases 1740 ou apresenta pelo menos 50 % de homologia com a sequência supra citada ou hibridiza sob condições rigorosas com a sequência supra

citada, ou abrange uma sequência como resultado da degeneração do código genético para a sequência de DNA supra citada.

Esta sequência, que antes nunca tinha sido descrita, adequa-se perfeitamente para quaisquer experiências, análises e processos de produção, etc, que se relacionam com a actividade da fucosiltransferase de plantas. Neste caso são de interesse tanto a sequência de DNA como também a proteína codificada por esta sequência, em que no entanto a sequência de DNA é utilizada especialmente para impedir a actividade da fucosiltransferase.

A grelha de leitura aberta da sequência com a SEQ ID n°:1 codifica uma proteína com 510 aminoácidos e com um peso molecular teórico de 56,8 kDa, em que no domínio entre Asn36 e Gly54 existe presumivelmente uma parte transmembranar. O valor pI calculado da proteína codificada da sequência de acordo com a SEQ ID N°:1 compreende 7,51.

A actividade da fucosiltransferase de plantas é comprovada e medida através de um processo, em que a fucosiltransferase é adicionada a uma amostra abrangendo fucose marcada e um aceitador (p.ex. uma glicoproteína) ligado a um suporte, p.ex. Sepharose. Após o tempo de reacção a amostra é lavada e o conteúdo de fucose ligado é medido. A actividade da fucosiltransferase é neste caso considerada como positivo, quando a medição da actividade é pelo menos cerca de 10 a 20%, especialmente pelo menos

cerca de 30 a 50% mais elevado que a actividade do controlo negativo. A estrutura da glicoproteína pode adicionalmente ser verificada através de HPLC. Tais protocolos são do estado da técnica (Staudacher et al. 1998, Anal. Biochem. 246, 96-101; Staudacher et al. 1991, Eur.J.Biochem. 199, 745-751).

Por exemplo a fucosiltransferase é adicionada a uma amostra abrangendo uma fucose marcada radioactivamente e um aceitador, p.ex. GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-Asn. Após o tempo de reação a amostra é purificada através de cromatografia de permuta aniónica e é medido o conteúdo de fucose ligado. Da diferença entre a radioactividade medida da amostra com o aceitador e a da do controlo negativo sem aceitador pode ser calculada a actividade. A actividade da fucosiltransferase é já avaliada como positiva, quando a radioactividade medida se encontra pelo menos 30-40% mais elevada que a radioactividade medida na amostra negativa.

O emparelhamento de duas moléculas de DNA pode ser modificado através da escolha da temperatura e força iónica da amostra. Sob condições rigorosas, são entendidas condições de acordo com a invenção, que possibilitam uma ligação exacta, rigorosa. P. ex. as moléculas de DNA são hibridizadas em dodecil sulfato de sódio a 7% (SDS), 0,5 M NaPO<sub>4</sub> pH 7,0; 1mM EDTA a 50°C e lavado com 1% SDS a 42°C.

Se as sequências apresentam uma homologia de pelo

menos 50% com a SEQ ID N°:1, pode p.ex. ser determinado com o programa FastDB da base de dados EMBL ou SWISSPROT.

A sequência da molécula de DNA codifica uma proteína com uma actividade GlcNAc- $\alpha$ 1,3-fucosiltransferase, especialmente com uma actividade núcleo- $\alpha$ 1,3-fucosiltransferase. Como já anteriormente descrito a núcleo- $\alpha$ 1,3-fucosiltransferase está presente em plantas e insectos, mas não no organismo humano, de forma que esta sequência de DNA é especialmente apropriada para ser utilizada em análises e experiências específicas da fucosiltransferase assim como processos de produção. Sob uma núcleo- $\alpha$ 1,3-fucosiltransferase é especialmente entendido a GlcNAc- $\alpha$ 1,3-fucosiltransferase GDP-L-Fuc:Asn-ligada. No âmbito da presente invenção com a designação de  $\alpha$ 1,3-fucosiltransferase é em geral considerado especialmente a núcleo- $\alpha$ 1,3-fucosiltransferase. Para as medições da actividade anteriormente descritas são úteis neste caso especialmente aceitadores com um GlcNAc-terminal não redutor. Tais aceitadores são p.ex. GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-Asn, GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4(Fuc $\alpha$ 1-6GlcNAc $\beta$ 1-Asn e GlcNAc $\beta$ 1-2Man $\alpha$ 1-3[Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man $\alpha$ 1-6]Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-Asn. Se a fucose está ligada, pode ser verificado para além disso através da medição da insensibilidade relativamente à N-Glicosidase F, o que pode ser comprovado com a ajuda de espetrometria de massa.

Preferencialmente a molécula de DNA apresenta

pelo menos 70-80%, especialmente preferido pelo menos 95%, de homologia com a sequência de acordo com a SEQ ID N°:1. Esta sequência codifica uma GlcNAc- $\alpha$ 1,3-fucosiltransferase especialmente activa. Uma vez que a sequência de DNA dependendo se de planta ou de insecto pode estar mais ou menos modificada, uma sequência, que apresenta p.ex. 70% de homologia com a sequência de acordo com a SEQ ID N°:1, apresenta igualmente uma actividade da fucosiltransferase que é suficiente para ser utilizada no processo como supra citado.

Descrito é uma molécula de DNA que abrange 2150 a 2250, em especial 2198, pares de bases. Esta molécula de DNA apresenta 100 a 300, preferencialmente 210, pares de bases a montante antes do codão de iniciação assim como 350 a 440, especialmente 458, pares de bases a jusante após o codão de terminação da grelha de leitura aberta, em que o final da molécula de DNA abrange preferencialmente uma 3'-poli(A)-cauda. Desta forma é assegurada uma regulação plenamente funcional ao nível da tradução, e é descrita uma molécula de DNA que é especialmente eficiente e codifica sem problemas uma GlcNAc- $\alpha$ 1,3-fucosiltransferase activa.

A presente descrição divulga além disso uma molécula de DNA, que abrange uma sequência de acordo com a SEQ ID N°:3 ou que abrange uma sequência, que apresenta pelo menos 85%, especialmente preferido pelo menos 95%, especialmente pelo menos 99%, de homologia com a sequência supra citada ou hibridiza sob condições rigorosas com a

sequência supra citada ou que degenerou como resultado do código genético para a sequência de DNA supra citada. A homologia é preferencialmente determinada com um programa, que reconhece inserções e deleções e não as tem em conta no cálculo de homologia. Esta sequência de nucleótidos codifica um motivo peptídico conservado, ou seja que a maioria das GlcNAc- $\alpha$ 1,3-fucosiltransferases activas e funcionais abrangem a sequência de aminoácidos assim codificada. A sequência pode aqui apresentar tanto o mesmo comprimento que a sequência de acordo com a SEQ ID N°:3, ou também evidentemente pode ser mais comprida. Esta sequência apresenta um comprimento mais pequeno que a sequência, que codifica a proteína total, e é assim menos propensa a recombinações, deleções ou outras mutações. Devido ao motivo conservado e à sua maior estabilidade, esta sequência é especialmente apropriada para os testes de reconhecimento da sequência.

A sequência SEQ ID N°: 3 apresenta a seguinte sequência: 5'-gaagccctgaagcactacaaatttagcttagcgttgaaaattcgaat-gaggaagattatgtaactgaaaaattcttccaatcccttgtgctgaaactgtccct-3'

Divulgado é uma molécula de DNA, que abrange uma sequência parcial de uma molécula de DNA supra citada e apresenta um tamanho de 20 a 200, preferencialmente 30 a 50 pares de bases. A molécula de DNA pode ser por exemplo utilizada, para se ligar como sonda a sequências complementares de GlcNAc- $\alpha$ 1,3-fucosiltransferase, de forma que estas podem ser selecionadas de uma amostra. Desta forma

podem ser selecionadas, isoladas e caracterizadas outras GlcNAc- $\alpha$ 1,3-fucosiltransferases das mais diferentes plantas e insectos. Podem ser utilizadas quaisquer ou também várias diferentes sequências parciais, especialmente uma parte do motivo conservado já supra citado.

Aqui é especialmente vantajoso, quando uma das moléculas de DNA supra citadas é associada covalentemente com uma substância marcadora comprovável. Como substância marcadora é de interesse qualquer marcador usual, assim p.ex. marcadores fluorescentes, luminescentes, radioactivos, marcações não-isótopas, como biotina, etc. Desta forma são colocados à disposição reagentes, que são úteis para a comprovação, a seleção e quantificação de moléculas de DNA correspondentes em amostras de tecido sólidas (p.ex. de plantas) ou também em amostras líquidas através de processos de hibridação.

É divulgado além disso um vector biologicamente funcional, que abrange uma das moléculas de DNA supra citadas ou partes de diferentes comprimentos, com pelo menos 20 pares de bases. Para a transfecção em células hospedeiras é necessário um vector independente, com capacidade de se multiplicar, em que dependendo da célula hospedeira, o mecanismo de transfecção, função e tamanho da molécula de DNA, pode ser utilizado um vector adequado. Uma vez que é conhecido um grande número de diferentes vectores, uma enumeração iria ultrapassar o âmbito do pedido, de forma que tal é aqui renunciado, especialmente,

por os vectores serem bem conhecidos do especialista (relativamente aos vectores assim como todas as técnicas e terminologias utilizadas nesta publicação, que são do conhecimento do especialista, ver também Sambrook Maniatis). Idealmente o vector apresenta uma pequena massa molecular e deveria apresentar genes selecionáveis, de forma a levar numa célula a um fenótipo facilmente reconhecível, de forma que é possível uma seleção simples de células hospedeiras contendo o vector e isentas de vector. De forma a obter um elevado rendimento de DNA e dos correspondentes produtos de genes, o vector deverá apresentar um promotor forte, assim como um enhancer, sinais de amplificação de gene e sequências reguladoras. Para uma replicação autónoma do vector é além disso importante uma origem de replicação. Locais de poliadenilação são responsáveis por um processamento correcto do mRNA e sinais de excisão pelos RNA-transcritos. No caso de serem utilizados fagos, vírus ou partículas virais como vectores, os sinais de empacotamento controlam o empacotamento do DNA do vector. P.ex. são apropriados para a transcrição em plantas plasmídeos Ti e para a transcrição em células de insectos baculovírus, ou transposões em insectos, como o elemento P.

Se o vector supra descrito for inserido numa planta ou célula de planta, assim é alcançada uma repressão pós transcrição da expressão génica do gene da  $\alpha$ 1,3-fucosiltransferase endógena através da transcrição de um transgene ou partes deste, homólogo deste, na orientação

sense. Para esta técnica sense é referido para além disso às publicações Baulcombe 1996, *Plant.Mol.Biol.* 9:373-382 e Brigneti et al. 1998, *EMBO J.* 17: 6739-6746. Esta estratégia de "Gen-Silencing" representa uma possibilidade eficaz, de reprimir a expressão do gene da  $\alpha$ 1,3-fucosiltransferase, ver também Waterhouse et al. 1998, *Proc.Natl.Acad.Sci.USA*, 95:13959-13964.

Além disso a divulgação refere-se a um vector biologicamente funcional, que abrange uma molécula de DNA de acordo com uma das formas de execução supra citadas ou partes desta de diferentes comprimentos com orientação inversa ao promotor. Se este vector for transfectado numa célula hospedeira, é lido um "antisense-mRNA", que é complementar ao mRNA da GlcNAc  $\alpha$ 1,3-fucosiltransferase e a complexa. Esta ligação impede ou o processamento correcto, o transporte, a estabilidade, ou através da prevenção do emparelhamento dos ribossomas a tradução e assim a expressão génica normal da GlcNAc- $\alpha$ 1,3-fucosiltransferase.

Apesar de toda a sequência da molécula de DNA poder ser introduzida no vector, sequências parciais desta podem ser vantajosas para determinados fins devido ao seu pequeno tamanho. É importante p.ex. no aspecto antisense, que a molécula de DNA seja suficientemente grande, para formar um antisense-mRNA suficientemente grande para se ligar à mRNA-transferase. Por exemplo uma molécula antisense-RNA apropriada abrange 50 a 200 nucleótidos, pois muitas das moléculas de antisense-RNA conhecidas que

ocorrem naturalmente abrangem aproximadamente 100 nucleótidos.

Para uma inibição especialmente eficaz da expressão de uma  $\alpha$ 1,3-fucosiltransferase activa é apropriada uma combinação da técnica sense e técnica antisense (Waterhouse et al. 1998, Proc.Natl.Acad.Sci.USA, 95:13959-13964).

Vantajosamente são utilizadas moléculas de RNA que hibridizam rapidamente. A eficácia das moléculas antisense-RNA, que apresentam um tamanho acima de 50 nucleótidos, depende da cinética de emparelhamento *in vitro*. Assim p.ex. moléculas antisense-RNA que se emparelham rapidamente mostram uma forte inibição da expressão das proteínas como molécula de RNA que hibridiza lentamente (Wagner et al. 1994, Annu. Rev. Microbiol., 48:713-742; Rittner et al. 1993, Nucl. Acids Res, 21:1381-1387). Tais moléculas de antisense-RNA que hibridizam rapidamente apresentam especialmente um grande número de bases externas (terminais livres e sequências de ligação), um grande número de sub-domínios estruturais (componentes), assim com um baixo grau de ansas (Patzel et al. 1998; Nature Biotechnology, 16; 64-68). As estruturas secundárias hipotéticas da molécula de antisense-RNA, podem p.ex. ser determinadas com a ajuda de um programa de computador, de acordo com o qual é selecionada uma sequência de DNA antisense-RNA apropriada.

Podem ser introduzidas diferentes regiões da

sequência da molécula de DNA no vector. Uma possibilidade consiste p.ex. em introduzir apenas a parte que é responsável pelo emparelhamento dos ribossomas, no vector. Um bloqueio nesta região do mRNA é suficiente para parar toda a tradução. Uma eficiência especialmente elevada da molécula antisense resulta também para as regiões 5' e 3' que não são traduzidas do gene.

Uma molécula de DNA mencionada pode apresentar uma sequência, que abrange uma mutação por deleção, por inserção e/ou por substituição. O número dos nucleótidos mutados é neste caso variável e vai desde um único a vários nucleótidos deletados, inseridos ou substituídos. Também é possível, que através da mutação a grelha de leitura seja deslocada. Importante num tal "Knockout-Gene" é apenas, que a expressão de uma GlcNAc- $\alpha$ 1,3-fucosiltransferase esteja perturbada e seja impedida a formação de uma enzima funcional, activa. Neste caso o local da mutação é variável, desde que a expressão de uma proteína enzimaticamente activa seja impedida. Preferencialmente a mutação é no local catalítico da enzima, que se encontra no local C-terminal. Os processos para introdução de mutações nas sequências de DNA são muito bem conhecidos do especialista, de forma que aqui é renunciado, aprofundar as diferentes possibilidades de mutagénese. Podem ser utilizadas tanto mutagéneses accidentais, mas também especialmente mutagéneses orientadas, p.ex. a site-directed-mutagenesis, a mutagénese dirigida com oligonucleótidos ou mutagéneses com a ajuda de enzimas de restrição.

Além disso é descrito uma molécula de DNA, que codifica uma ribozima, a qual apresenta duas secções da sequência de respectivamente pelo menos 10 a 15 pares de bases, as secções da sequência são complementares a uma molécula de DNA como anteriormente descrito, de forma que a ribozima complexa e corta o mRNA, que é transcrito de uma molécula de DNA GlcNAc- $\alpha$ 1,3-fucosiltransferase natural. A ribozima reconhece o mRNA da GlcNAc- $\alpha$ 1,3-fucosiltransferase através de emparelhamento complementar de bases com o mRNA. De seguida a ribozima corta e destrói o RNA numa forma específica da sequência, antes da enzima ser traduzida. Após dissociação do substrato cortado a ribozima hibridiza repetidamente as moléculas de RNA e actua como endonuclease específica. De um modo geral as ribozimas podem ser produzidas especificamente para inactivação de um mRNA determinado, mesmo quando não é conhecida a sequência de DNA total, que codifica a proteína. As ribozimas são especialmente então eficazes, quando os ribossomas deslizam lentamente ao longo do mRNA. Neste caso, é mais fácil para a ribozima, encontrar um local livre de ribossomas no mRNA. Por este motivo os mutantes de ribossomas lentos são também apropriados como sistema para ribozimas (J. Burke, 1997, *Nature Biotechnology*; 15, 414-415). Esta molécula de DNA é especialmente apropriada para a repressão ou impedimento da expressão de uma GlcNAc- $\alpha$ 1,3-fucosiltransferase de planta.

Uma possibilidade consiste também em utilizar uma forma modificada de um ribozima, nomeadamente uma minizima.

As minizimas são especialmente eficazes para o corte de moléculas de mRNA maiores. Uma minizima é uma ribozima-de cabeça martelo, que em vez do Stem/Loop II apresenta um nucleótido-linker curto. Especialmente eficazes são dímeros-minizimas (Kuwabara et al. 1998, *Nature Biotechnology*, 16; 961-965).

Por conseguinte é descrito um vector biologicamente funcional, que abrange uma das duas ultimas moléculas de DNA mencionadas (mutações ou molécula de DNA-ribozima). Neste caso é válido o já descrito relativamente a vectores. Um tal vector pode p.ex. ser inserido num microrganismo e ser utilizado para a produção de elevadas concentrações das moléculas de DNA anteriormente descritas. Além disso um tal vector é especialmente apropriado para a inserção da molécula de DNA específica num organismo planta ou insecto, para limitar ou impedir completamente a produção da GlcNAc- $\alpha$ 1,3-fucosiltransferase nesse organismo.

Divulgado é um processo para produção de um cDNA abrangendo a molécula de DNA divulgada, em que o RNA é isolado de células de insectos ou de plantas, especialmente de células de hipocótilo, com o qual é realizada uma transcrição reversa após adição de uma transcriptase reversa e iniciadores. Cada um dos passos deste processo é realizado de acordo com protocolos em si conhecidos. Para a transcrição reversa existe por um lado a possibilidade, com a ajuda de iniciadores oligo(dT) produzir o cDNA do mRNA total e só subsequentemente realizar uma PCR com

iniciadores escolhidos, para produzir moléculas de DNA que abrangem o gene da GlcNAc- $\alpha$ 1,3-fucosiltransferase. Por outro lado, os iniciadores escolhidos podem ser utilizados directamente para a transcrição reversa, para se obter um cDNA curto, específico. Os iniciadores apropriados podem ser produzidos p.ex. sinteticamente através da apresentação de sequências de cDNA da transferase. Com a ajuda deste processo podem ser produzidas muito rapidamente e simplesmente, com pequena quota de erros, grandes quantidades de moléculas cDNA.

Além disso é divulgado um processo para clonagem de uma GlcNAc- $\alpha$ 1,3-fucosiltransferase, caracterizado por, a molécula de DNA ser clonada num vector, o qual é subsequentemente transfectado numa célula hospedeira ou num hospedeiro, em que através de seleção e amplificação de células hospedeiras transfetadas são obtidas linhagens celulares, que exprimem a GlcNAc- $\alpha$ 1,3-fucosiltransferase activa. A molécula de DNA é por exemplo inserida com a ajuda de endonucleases de restrição no vector. Para o vector é válido novamente o já anteriormente mencionado. Importante neste processo, é que se escolha um sistema hospedeiro-vector eficiente. Para se obter uma enzima activa, são especialmente apropriadas células hospedeiras eucarióticas. Uma possibilidade consiste em transfetar o vector em células de insectos. Neste caso seria de utilizar especialmente um vírus de insecto como vector, tal como p.ex. Baculovírus.

Evidentemente podem ser transfectadas também células humanas ou de outros vertebrados, em que estas iriam exprimir uma enzima estranha para elas.

Divulgado é um processo para produção de células hospedeiras recombinantes, especialmente células de plantas ou de insectos, ou plantas ou insectos com uma produção de GlcNAc- $\alpha$ 1,3-fucosiltransferase reprimida ou completamente impedida, que é caracterizado por, pelo menos um dos vectores, nomeadamente aquele abrangendo a de acordo com a invenção molécula de DNA, a molécula de DNA mutada ou a molécula de DNA que codifica o ribozima, ou aquele abrangendo a molécula de DNA em orientação inversa para o promotor, ser inserido numa célula hospedeira ou planta ou no insecto. Aqui também é válido o já anteriormente descrito para a transfecção.

Como células hospedeiras podem ser utilizadas p.ex. células de plantas, em que aqui por exemplo interessa o plasmídeo Ti com o sistema Agrobacterium. É possível com o sistema Agrobacterium, transfectar directamente uma planta: As agrobactérias causam em plantas nódulos na zona do colo. Quando as agrobactérias infectam uma planta ferida, as bactérias não vão elas próprias para a planta, mas sim elas inserem a secção de DNA recombinante, o chamado T-DNA do plasmídeo Ti em forma de anel, extracromossómico, indutor de tumores, nas células de plantas. O T-DNA, e assim também a molécula de DNA aí inserida, é integrado estavelmente no DNA cromossómico das células, de forma que são expressos os genes do T-DNA na planta.

Existem vários mecanismos de transfecção conhecidos, eficientes, para diferentes sistemas de hospedeiros. Alguns exemplos são a electroporação, o método de fosfato de cálcio, microinjeção, método de lipossomas.

As células transfectadas são subsequentemente selecionadas, p.ex. com base em resistências a antibióticos, para o vector que apresenta o gene, ou outros genes marcadores. Depois as linhagens celulares transfectadas são amplificadas, ou em pequenas quantidades, p.ex. em placas de Petri, ou em grandes quantidades, tal como em fermentadores. Além disso as plantas apresentam uma característica especial, elas são nomeadamente capazes de se desenvolverem outra vez a partir de uma célula (transfectada) ou um protoplasma numa planta completa, que pode ser cultivada.

Dependendo do vector utilizado, ocorrem processos no hospedeiro, de forma que a expressão da enzima é reprimida ou completamente impedida:

Se o vector abrangendo a molécula de DNA com a mutação por deleção, inserção ou substituição for transfectado, então ocorre uma recombinação homóloga: a molécula de DNA mutada reconhece apesar da mutação a sequência idêntica no genoma da célula hospedeira e é integrada exactamente no local, de forma que é originado um "knockout-gene". Desta forma é introduzida uma mutação no

gene da GlcNAc- $\alpha$ 1,3-fucosiltransferase, o que pode inibir a expressão plenamente operacional da GlcNAc- $\alpha$ 1,3-fucosiltransferase. Como já anteriormente explicado, é importante nesta técnica, que a mutação seja suficiente, para impedir a expressão da proteína activa. Após seleção e amplificação, o gene pode ser sequenciado como verificação adicional, de forma a verificar o sucesso da recombinação homóloga ou o grau da mutação.

Se o vector abrangendo a molécula de DNA, que codifica um ribozima, for transfetado, assim é exprimido na célula hospedeira o ribozima activo. O ribozima complexa a sequência de mRNA complementar da GlcNAc- $\alpha$ 1,3-fucosiltransferase pelo menos num local determinado, corta esse local e pode dessa forma inibir a tradução da enzima. Nesta célula hospedeira, assim como na linhagem celular dela descendente ou eventualmente na planta, não é exprimido nenhuma GlcNAc- $\alpha$ 1,3-fucosiltransferase.

No caso do vector abrangendo a molécula de DNA divulgada na direcção sense ou inversa do promotor, é exprimido na célula transfetada (ou planta) um mRNA sense ou antisense. O antisense-mRNA é complementar para pelo menos uma parte da sequência de mRNA da GlcNAc- $\alpha$ 1,3-fucosiltransferase e pode eventualmente inibir a tradução da enzima. Como exemplo para um processo para reprimir a expressão de um gene através da técnica de antisense é remetido para a publicação de Smith et al. 1990, Mol.Gen.Genet. 224:477-481, em que nesta publicação é

inibida a expressão de um gene que está envolvido no processo de maturação de tomates.

Em todos os sistemas a expressão da GlcNAc- $\alpha$ 1,3-fucosiltransferase é pelo menos reprimida, preferencialmente até completamente impedida. O grau da perturbação da expressão génica depende do grau da complexação, recombinação homóloga, de mutações accidentais eventualmente subsequentes, e outros processos no domínio do genoma. As células transfectadas são verificadas e selecionadas quanto à actividade de GlcNAc- $\alpha$ 1,3-fucosiltransferase.

Existe além disso a possibilidade, de reforçar ainda mais a repressão acima descrita da expressão da  $\alpha$ 1,3-fucosiltransferase, ao adicionalmente à inserção de um vector anteriormente descrito, ser inserido no hospedeiro um vector abrangendo um gene, que codifica uma proteína de um mamífero, p.ex.  $\beta$ 1,4-galactosiltransferase. A fucosilação pode ser reduzida através da acção de outras enzimas de mamíferos, em que a combinação da inibição da expressão de uma  $\alpha$ 1,3-fucosiltransferase activa com a ajuda de um vector segundo a invenção e com a ajuda de um vector de enzima de mamífero é especialmente eficaz.

Para a transfecção pode ser utilizada qualquer tipo de planta, por exemplo o feijão mungo, planta do tabaco, planta do tomate e/ou da batata.

Um outro processo vantajoso para a produção de

células hospedeiras recombinantes, especialmente células de plantas ou de insectos ou plantas ou insectos, consiste em, que a molécula de DNA abrangendo a mutação é inserida no genoma da célula hospedeira ou planta ou do insecto no local da sequência não mutada, homóloga (Schaefer et al. 1997, Plant J.; 11(6):1195-1206). Este processo funciona por conseguinte não com um vector, mas sim com uma molécula de DNA pura. A molécula de DNA é inserida no hospedeiro p.ex. através de bombardeamento de genes, microinjeção ou eletroporação, apenas para mencionar três exemplos. Como já anteriormente explicado, a molécula de DNA encosta-se à sequência homóloga no genoma do hospedeiro, de forma que ocorre uma recombinação homóloga e assim a incorporação da mutação por deleção, inserção ou substituição no genoma: A expressão da GlcNAc- $\alpha$ 1,3-fucosiltransferase pode ser reprimida ou completamente impedida.

Descrito são plantas ou células de plantas assim como insectos ou células de insectos, em que a sua actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase é inferior a 50%, especialmente inferior a 20%, especialmente preferido 0%, da actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase que ocorre naturalmente em plantas ou células de plantas e insectos ou células de insectos. A vantagem destas plantas ou células de plantas consiste em que, as glicoproteínas, que são produzidas por elas, não apresentam nenhuma ou quase nenhuma fucose  $\alpha$ 1,3-ligada. Se então os produtos destas plantas ou insectos forem absorvidos pelo corpo humano ou de um mamífero, não ocorre nenhuma reacção imunitária contra o epítopo  $\alpha$ 1,3-fucose.

Descritas são plantas ou células de plantas recombinantes, que são produzidas de acordo com um dos processos anteriormente descritos e em que a sua produção de GlcNAc- $\alpha$ 1,3-fucosiltransferase está reprimida ou completamente impedida.

Divulgado são também insectos ou células de insectos recombinantes, que são produzidos de acordo com um dos processos anteriormente descritos e em que a sua produção de GlcNAc- $\alpha$ 1,3-fucosiltransferase está reprimida ou completamente impedida. Também aqui não são produzidas nenhuma glicoproteína com resíduos de fucose  $\alpha$ 1,3-ligado, de forma que também não ocorre nenhuma reacção imunitária contra o epítopo  $\alpha$ 1,3-fucose.

Referido é também uma molécula PNA que abrange uma sequência de bases, que é complementar à sequência da molécula de DNA assim como sequências parciais desta. PNA (péptido ácido nucleico) é uma sequência semelhante ao DNA, em que as bases nucleótidas estão ligadas a uma estrutura central pseudopeptídica. PNA hibridiza em geral com oligómeros de DNA, RNA ou PNA complementares através de emparelhamento de bases de Watson-Crick e formação de hélice. A estrutura central peptídica garante uma maior resistência contra a degradação enzimática. A molécula PNA representa por isso um agente antisense melhorado. Nem nucleases nem proteases têm a capacidade de atacar uma molécula PNA. A estabilidade da molécula PNA, quando está

ligada a uma sequência complementar, mostra um bloqueio estérico suficiente de polimerases DNA e RNA, da transcriptase reversa, telomerase e dos ribossomos.

Caso a molécula PNA apresente a sequência anteriormente mencionada, assim ela liga-se ao DNA ou a um local do DNA, que codifica a GlcNAc- $\alpha$ 1,3-fucosiltransferase, e pode desta forma inibir a transcrição desta enzima. A molécula PNA, uma vez que não é nem transcrita nem traduzida, é produzida sinteticamente, p.ex. com a ajuda da técnica t-Boc.

Divulgado é uma molécula PNA, que abrange uma sequência de bases, que corresponde à sequência da molécula de DNA assim como sequências parciais desta. Esta molécula PNA complexa o mRNA ou um local do mRNA da GlcNAc- $\alpha$ 1,3-fucosiltransferase, de forma que a tradução da enzima é inibida. Aqui é válido semelhante, que para o antisense-RNA. Assim p.ex. um local de complexação especialmente eficiente é a região de iniciação da tradução ou também as regiões do mRNA 5'- que não são traduzidas.

Divulgado é um processo para produção de plantas ou insectos ou células, especialmente células de plantas ou de insectos, que apresentam uma expressão bloqueada da GlcNAc- $\alpha$ 1,3-fucosiltransferase ao nível da transcrição ou da tradução, que é caracterizado por, serem inseridas moléculas PNA nas células. Para inserir a molécula PNA ou as moléculas PNA na célula, são utilizados novamente os

métodos usuais, como p.ex. electroporação ou microinjeção. Especialmente eficiente é a inserção, quando os oligómeros PNA estão ligados a péptidos de penetração nas células, p.ex. Transportan ou pAntp (Pooga et al. 1998, *Nature Biotechnology*, 16; 857-861).

A invenção põe à disposição um processo para produção de glicoproteínas recombinantes, que é caracterizado por, as plantas ou células de plantas recombinantes cuja produção de GlcNAc- $\alpha$ 1,3-fucosiltransferase está reprimida ou completamente impedida, serem transfectadas com o gene, que codifica a glicoproteína, de forma que são expressas as glicoproteínas recombinantes. Neste caso são transfectados como anteriormente já descrito vectores abrangendo genes para a proteína desejada no hospedeiro ou células do hospedeiro como também já anteriormente descrito. As células de plantas ou de insectos transfectadas expressam a proteína desejada, em que elas não apresentam nenhuma ou quase nenhuma fucose  $\alpha$ 1,3-ligada. Desta forma elas não desencadeiam as reacções imunitárias anteriormente já mencionadas no corpo humano ou no de mamífero. Podem ser produzidas quaisquer proteínas nestes sistemas.

Preferencialmente é posto à disposição um processo para produção de glicoproteínas humanas recombinantes, que é caracterizado por, as plantas ou células de plantas recombinantes, cuja produção de GlcNAc- $\alpha$ 1,3-fucosiltransferase está reprimida ou completamente impedida, serem transfectadas com o gene, que codifica a glico-

proteína, de forma que são expressas glicoproteínas recombinantes. Através deste processo torna-se possível, produzir proteínas humanas em (células de) plantas, que, quando elas são absorvidas pelo corpo humano, não desencadeiam nenhuma reacção imunitária dirigida aos resíduos de fucose  $\alpha 1,3$ -ligado. Neste caso existe a possibilidade, de utilizar tipos de plantas para a produção das glicoproteínas recombinantes, que servem como alimento, p.ex. bananas, batatas e/ou tomates. Os tecidos destas plantas abrangem a glicoproteína recombinante, de forma que p.ex. através da extração da glicoproteína recombinante do tecido e subsequente administração ou directamente através do consumo do tecido vegetal, a glicoproteína recombinante é absorvida no corpo humano.

Preferencialmente é aqui um processo para produção de glicoproteínas humanas recombinantes para utilização terapêutica, em que as plantas ou células de plantas recombinantes, de acordo com a invenção, cuja produção da GlcNAc- $\alpha 1,3$ -fucosiltransferase está reprimida ou completamente impedida, são transfetadas com o gene, que codifica a glicoproteína, de forma que são expressas as glicoproteínas recombinantes. Aqui interessa qualquer proteína, que tenha interesse terapêutico.

Além disso a presente invenção refere-se a glicoproteínas recombinantes, em que elas foram produzidas de acordo com um dos processos anteriormente descritos em sistemas de plantas ou de insectos e em que a sua sequência de péptidos apresenta abaixo de 50%, especialmente abaixo

de 20%, especialmente preferido 0%, de resíduos de fucose  $\alpha$ 1,3-ligada presentes nas proteínas expressas nos sistemas de plantas ou de insetos não reduzidos pela fucosiltransferase. Preferíveis são naturalmente glicoproteínas que não apresentam nenhum resíduo de fucose  $\alpha$ 1,3-ligado. A quantidade de fucose  $\alpha$ 1,3-ligada é dependente do grau da repressão anteriormente descrita da GlcNAc- $\alpha$ 1,3-fucosiltransferase.

Preferencialmente a invenção refere-se a glicoproteínas humanas recombinantes, que foram produzidas de acordo com um dos processos anteriormente descritos em sistemas de plantas, e que apresentam na sua sequência peptídica abaixo de 50%, especialmente abaixo de 20%, especialmente preferido 0%, de resíduos de fucose  $\alpha$ 1,3-ligada presentes nas proteínas expressas nos sistemas de plantas não reduzidos pela fucosiltransferase.

Descritas são glicoproteínas humanas recombinantes para utilização terapêutica, que foram produzidas de acordo com um dos processos anteriormente descritos em sistemas de plantas ou de insetos, e que apresentam na sua sequência peptídica abaixo de 50%, especialmente abaixo de 20%, especialmente preferido 0%, de resíduos de fucose  $\alpha$ 1,3-ligada presentes nas proteínas expressas nos sistemas de plantas ou de insetos não reduzidos pela fucosiltransferase.

Estas glicoproteínas podem apresentar outras unidades de oligossacáridos ligadas, que são específicas para

plantas ou insectos, pelo qual elas se diferenciam - no caso de glicoproteínas humanas - destas glicoproteínas naturais. Não obstante é desencadeado pelas glicoproteínas uma reacção imunológica menor ou até nenhuma no corpo humano, pois, como já descrito na introdução da descrição, os resíduos de fucose  $\alpha$ 1,3-ligados são a principal causa das reacções imunitárias ou reacções imunitárias cruzadas contra glicoproteínas de plantas e de insectos.

Divulgado é uma composição farmacêutica, que abrange as glicoproteínas. Adicionalmente às glicoproteínas a composição farmacêutica abrange outros aditivos usuais para tais composições. Estes são p.ex. diluentes apropriados de diferentes conteúdos de tampão (p.ex. Tris-HCl, acetato, fosfato), pH e força iônica, aditivos, tal como agentes tensioactivos e agentes solubilizantes (Tween 80, polissorbato 80), conservantes (p.ex. timerosal, álcool benzílico), adjuvantes, antioxidantes (p.ex. ácido ascórbico, metabissulfito de sódio), emulsionantes, substâncias inertes (p.ex. lactose, manitol), ligação covalente de polímeros, tal como polietilenoglicol, à proteína, integração do material em preparações em forma de partículas de compostos químicos poliméricos, tal como ácido poliláctico, ácido poliglicólico, etc. ou em lipossomas, excipientes e/ou agentes de transporte, que são úteis no respectivo tratamento. Tais composições irão influenciar o estado físico, a estabilidade, a velocidade da libertação in vivo e a velocidade da eliminação in vivo das glicoproteínas.

Descrito é também um processo para selecção de

moléculas de DNA, que codificam uma GlcNAc- $\alpha$ 1,3-fucosiltransferase, numa amostra, em que são adicionadas as moléculas de DNA marcadas, divulgadas, à amostra, que se ligam às moléculas de DNA que codificam uma GlcNAc- $\alpha$ 1,3-fucosiltransferase. As moléculas de DNA hibridizadas podem ser detectadas, quantificadas e selecionadas. Para que a amostra apresente DNA monocatenário, com o qual as moléculas de DNA marcadas podem hibridizar, a amostra é desnaturada, p.ex. através de aquecimento.

Uma possibilidade consiste em separar o DNA a investigar, eventualmente após adição de endonucleases, através de electroforese em gel num gel de agarose. Após passagem sobre uma membrana de nitrocelulose, são adicionadas as moléculas de DNA marcadas, que hibridizam com a molécula de DNA homóloga correspondente ("Southern Blotting").

Uma outra possibilidade consiste em localizar genes homólogos de outras espécies através de processos dependentes de PCR com a utilização de iniciadores específicos e/ou degenerados, derivados da sequência da molécula de DNA divulgada.

Preferencialmente a amostra abrange para os processos supra citados DNA genómico de um organismo de planta ou de insecto. Através deste processo é investigado de uma forma muito rápida e eficiente um grande número de plantas e insectos relativamente à presença do gene da GlcNAc- $\alpha$ 1,3-fucosiltransferase. Desta forma plantas e

insectos, que não apresentam este gene, podem ser selecionados ou em tais plantas e insectos, que apresentam este gene, pode através de um processo supra citado ser reprimida ou completamente impedida a expressão da GlcNAc- $\alpha$ 1,3-fucosiltransferase, de forma que eles podem ser subsequentemente utilizados para a transfecção e produção de glicoproteínas (humanas).

Referido são também moléculas de DNA, que codificam uma GlcNAc- $\alpha$ 1,3-fucosiltransferase, que foram selecionadas de acordo com os dois últimos processos mencionados e subsequentemente foram isoladas da amostra. Estas moléculas podem ser utilizadas para outras investigações. Elas podem ser sequenciadas e serem utilizadas por seu lado também como sondas de DNA para localizar GlcNAc- $\alpha$ 1,3-fucosiltransferases. Estas moléculas de DNA - marcadas - irão actuar mais eficientemente como sondas em organismos, que são parentes dos organismos a partir dos quais foram isoladas, do que as moléculas de DNA de acordo com a invenção.

Divulgada é uma preparação de GlcNAc- $\alpha$ 1,3-fucosiltransferase clonada, que apresenta isoformas com valores pI entre 6,0 e 9,0, especialmente entre 6,8 e 8,2. O valor pI de uma proteína é aquele valor de pH, no qual a sua carga líquida é zero e é dependente da sequência de aminoácidos, do padrão de glicosilação assim como também da estrutura espacial da proteína. A GlcNAc- $\alpha$ 1,3-fucosiltransferase abrange pelo menos 7 isoformas, que apresentam um valor pI neste intervalo. A razão para as diferentes

isofomas da transferase é p.ex. diferentes glicosilações assim como proteólise limitada. Experiências mostraram, que plântulas de feijão mungo de diferentes plantas apresentam diferentes proporções das isozimas. O valor pI de uma proteína pode ser determinado através da focagem isoelétrica, que é do conhecimento do especialista.

A isofoma principal da enzima apresenta um peso molecular aparente de 54 kDa.

Especialmente a preparação apresenta isofomas com valores de pI de 6,8, 7,1 e 7,6.

Descrito é um processo para produção de unidades de hidrato de carbono "tornadas planta" de glicoproteínas humanas e de outros mamíferos, em que é adicionado a uma amostra, que abrange uma unidade de hidrato de carbono ou uma glicoproteína, unidades de fucose assim como uma GlcNAc- $\alpha$ 1,3-fucosiltransferase codificada por uma molécula de DNA supra citada, de forma que a fucose é ligada na localização  $\alpha$ 1,3 à unidade de hidrato de carbono ou à glicoproteína através da GlcNAc- $\alpha$ 1,3-fucosiltransferase. Através do processo para clonagem de GlcNAc- $\alpha$ 1,3-fucosiltransferase é possível, produzir grandes quantidades da enzima purificada. Para se obter uma transferase completamente activa, são produzidas condições de reacção apropriadas. Verificou-se que a transferase apresenta uma actividade especialmente elevada num valor de pH de aproximadamente 7, quando é utilizado como tampão ácido 2-(N-morfolino)etanossulfónico-HCl. Na presença de catiões

bivalentes, especialmente  $Mn^{2+}$ , a actividade da transferase recombinante é aumentada. A unidade de hidrato de carbono é ou adicionada à amostra na forma não ligada ou ligada a uma proteína. A transferase recombinante é activa para ambas as formas.

A invenção é adicionalmente elucidada com base nos seguintes exemplos e nas figuras, às quais ela evidentemente não está limitada. Em particular a ilustração Fig. 1a e 1b mostra a quantidade medida de proteína e a actividade enzimática medida nas várias frações do eluato como curvas; a Fig. 2 a ilustração de uma electroforese em gel da GlcNAc- $\alpha 1,3$ -fucosiltransferase; a Fig. 3 o resultado da focagem isoeléctrica e a actividade da transferase medida das várias isoformas; a Fig. 4 as sequências N-terminais de 4 péptidos tripticos 1 - 4 assim como a sequência de DNA de três iniciadores S1, A2 e A3; a Fig. 5a e 5b a sequência de cDNA da  $\alpha 1,3$ -fucosiltransferase; a Fig. 6a e 6b a sequência de aminoácidos derivada da  $\alpha 1,3$ -fucosiltransferase; a Fig. 7 uma representação esquemática da  $\alpha 1,3$ -fucosiltransferase assim como a hidrofobicidade dos resíduos de aminoácidos; a Fig. 8 uma comparação dos motivos conservados de diferentes fucosiltransferases; a Fig. 9 uma comparação da actividade da fucosiltransferase de células de insectos tranfectadas com o gene da  $\alpha 1,3$ -fucosiltransferase com um controlo negativo; a Fig. 10a e 10b estruturas de aceitadores diferentes da  $\alpha 1,3$ -fucosiltransferase; a Fig. 11 e 12 espectros de massa; e a Fig. 13 o resultado de uma HPLC.

Exemplo 1:

Isolamento da núcleo- $\alpha$ 1,3-fucosiltransferase

Todos os passos foram realizados a 4°C. Plântulas de feijão mungo foram homogeneizadas numa misturadora, em que foi utilizado 0,75 volumes de tampão de extração por kg de feijões. O homogeneizado foi de seguida filtrado através de duas camadas de tecido de algodão e o filtrado foi centrifugado a 30000xg 40 min. O sobrenadante foi descartado e o pelete foi extraído com tampão de solução durante a noite com agitação contínua. Centrifugação subsequente a 30000xg durante 40 min formou o extrato-Triton.

O extrato-Triton foi como de seguinte purificado:

Passo 1: O extracto-Triton foi aplicado numa coluna de celulose de um permutador de anões-dietilaminoetil-celulose DE52 (5x28 cm, firma Whatman), que antes foi calibrado com tampão A. A fração não ligada foi processada posteriormente no passo 2.

Passo 2: A amostra foi aplicada numa coluna Affi-Gel Blue coluna (2,5x32) calibrada com tampão A. Após lavagem da coluna com este tampão foi feita a eluição da proteína adsorvida com tampão A abrangendo 0,5 M NaCl.

Passo 3: Após diálise do eluato do passo 2 contra

tampão B, foi aplicado numa coluna S-Sepharose calibrada com o mesmo tampão. Foi feita a eluição da proteína ligada com um gradiente linear de 0 a 0,5 M NaCl no tampão B. As fracções com GlcNAc- $\alpha$ 1,3-fucosiltransferase foram agrupadas e dialisadas contra tampão C.

Passo 4: A amostra dialisada foi aplicada numa coluna GnGn-Sepharose calibrada com tampão C. Foi feita a eluição da proteína ligada com tampão C abrangendo 1 M NaCl em vez de MnCl<sub>2</sub>.

Passo 5: A enzima foi de seguida dialisado contra tampão D e aplicado numa coluna GDP-hexanolamina-Sepharose. Após lavagem da coluna com tampão D a transferase foi eluida através da substituição de MgCl<sub>2</sub> e NaCl por 0,5 mM GDP. As fracções activas foram agrupadas, dialisadas contra 20 mM tampão Tris-HCl, pH 7,3 e liofilizadas.

A actividade enzimática da GlcNAc  $\alpha$ 1,3 fucosil-transferase foi determinada através da utilização de Gn-Gn-péptido e GDP-L-[U-<sup>14</sup>C]fucose a concentrações do substrato de respetivamente 0,5 e 0,25, na presença de tampão ácido 2-(N-morfolino)etanossulfónico-HCl, Triton X-100, MnCl<sub>2</sub>, GlcNAc e AMP (de acordo com Staudacher et al. 1998 Glycoconjugate J. 15, 355-360; Staudacher et al. 1991 Eur.J.Biochem. 199, 745-751).

As concentrações da proteína foram determinadas com a ajuda do método "Ácido bichinchonínico" (Pierce) ou,

nos últimos passos da purificação enzimática, através da análise dos aminoácidos (Altmann 1992 Anal.Biochem. 204, 215-219).

Na Fig. 1a e 1b estão representadas as quantidades medidas de proteína e a actividade enzimática medida nas várias fracções do eluato como curvas. A Fig. 1a mostra a separação supra citada na coluna S-Sepharose, a Fig. 1b a separação na coluna GnGn-Sepharose, em que o círculo representa a proteína, o círculo preto, cheio a GlcNAc- $\alpha$ 1,3-fucosiltransferase e o quadrado a N-acetil- $\beta$ -glucosaminidase. Um U é definido como a quantidade de enzima, que transfere 1  $\mu$ mol fucose para um aceitador por minuto.

A tabela 1 mostra os passos individuais da purificação da transferase.

Tabela 1

Passo de purificação	Proteína total	Actividade total	Actividade específica	Factor de purificação	Rendimento
	mg	mU	mU/mg	-vezes	%
Extracto Triton X-100	91500	4846	0,05	1	100
DE52	43700	4750	0,10	2	98,0
Affigel Blue	180,5	4134	23	460	85,3
S-Sepharose	8,4	3251	390	7800	67,1
GnGn-Sepharose	0,13 <sup>1</sup>	1044	8030	160000	21,5
GDP-hexanolamin-Sepharose	0,02 <sup>1</sup>	867	43350	867000	17,9

<sup>1</sup> foi determinado através de análise de aminoácidos

Tampão de extracção:

0,5 mM ditiotreit

1 mM EDTA

0,5 % polivinilpolipirrolidona

0,25 M sacarose

50 mM tampão Tris-HCl, pH 7,3

Tampão de solução:

0,5 mM ditiotreit

1 mM EDTA

1,5 % Triton X-100

50 mM Tris-HCl, pH 7,3

Tampão A:

25 mM tampão Tris-HCl, pH 7,3 com:

0,1 % Triton X-100 e

0,02 % NaN<sub>3</sub>

Tampão B:

25 mM tampão Na-citrato, pH 5,3 com:

0,1 % Triton X-100 e

0,02 % NaN<sub>3</sub>

Tampão C:

25 mM tampão Tris-HCl, pH 7,3 com:

5 mM MnCl<sub>2</sub> e

0,02 % NaN<sub>3</sub>

Tampão D:

25 mM Tris-HCl, pH 7,3 com:

10 mM MgCl<sub>2</sub>,

0,1 M NaCl e

0,02 % NaN<sub>3</sub>

Exemplo 2:

SDS-Page e focagem isoeléctrica

Foi realizada uma SDS-page numa Biorad-Mini-Protean-célula em géis com 12,5% acrilamida e 1% bisacrilamida. Os géis foram corados ou com Comassie Brilliant Blau R-250 ou prata. A focagem isoeléctrica da fucosiltransferase foi realizada em géis pré-preparados com um intervalo pI entre 6-9 (Servalyt precotes 6-9, Serva). Os géis foram corados com prata de acordo com o protocolo do fabricante. Para a electroforese bidimensional foram cortadas bandas do gel de focagem, tratadas com reagentes de S-alquilação e SDS, e submetido a uma SDS-page, com anteriormente descrito.

A Fig. 2 representa a figura de um gel de electroforese da GlcNAc- $\alpha$ 1,3-fucosiltransferase, em que à esquerda está representada a electroforese bidimensional e à direita a SDS-Page unidimensional. Aqui a banda designada de A é um padrão, a banda designada de B a GlcNAc  $\alpha$ 1,3-fucosiltransferase da coluna GnGn-Sepharose e a banda designada de C a GlcNAc- $\alpha$ 1,3-fucosiltransferase "purificada", ou seja a fracção da coluna GDP-hexanolamina-

Sepharose. As duas bandas a 54 e 56 kDa representam isoformas da transferase.

A Fig. 3 mostra o resultado de uma focagem isoeléctrica. A banda A foi corada com prata, na banda B foi testada a actividade das isoformas da transferase. A actividade é aqui indicada como % fucose, a GDP-fucose que foi transferida para o substrato.

Exemplo 3:

#### Sequenciação dos péptidos

Para a sequenciação da proteína foram cortadas bandas do SDS-gel de poliacrilamida corado com Coomassie, carboxi amidometilado e dissociado com tripsina de acordo com Görg et al. 1988, Electrophoresis, 9, 681-692. Os péptidos tripticos foram separados com HPLC de fase reversa num 1,0x250 mm Vydac C18 a 40°C com uma taxa do fluxo de 0,05 ml/min, em que foi utilizado um aparelho HP 1100 (Hewlett-Packard). Os péptidos isolados foram sequenciados com um sistema de sequenciação de proteína Hewlett-Packard G1005A de acordo com o protocolo do fabricante. Para além disso a mistura peptídica foi analisada através de digestão com Ingel com MALDI-TOF MS (ver abaixo).

A Fig. 4 mostra as sequências N-terminais de 4 péptidos tripticos 1 - 4 (SEQ ID N°: 5-8). A partir dos três primeiros péptidos foram produzidos os iniciadores S1, A2 e A3 (números de identificação da sequência 9-11).

Exemplo 4:

RT-PCR e clonagem de cDNA

Todo o RNA foi isolado de hipocótilos de feijão mungo com 3 dias de idade, em que foi utilizado o sistema de isolamento de RNA total SV de Promega. Para produção de cDNA de uma cadeia, todo o RNA foi incubado 1 h a 48°C com AMV transcriptase reversa e iniciadores oligo(dT), em que foi utilizado o sistema de transcrição reversa de Promega.

O cDNA de uma cadeia foi submetido a uma PCR, em que foi utilizado uma combinação de iniciadores sense e antisense:

A 10 µl da reação de transcrição reversa foi adicionado o seguinte:

50 µl com 0,1 µmmol de cada um dos iniciadores, 0,1 mM dNTPs, 2 mM MgCl<sub>2</sub>, 10 mM tampão Tris-HCl, pH 9,0, 50 mM KCl e 0,1% Triton X-100.

Após um primeiro passo de desnaturação a 95°C durante 2 min, foram feitos 40 ciclos de 1 min a 95°C, 1 min a 49°C e 2 min a 72°C. O último passo de extensão foi realizado a 72°C durante 8 min. Os produtos da PCR foram subclonados no vector pCR2.1, em que foi utilizado o TA Cloning Kit de Invitrogen, e sequenciado. Os produtos desta PCR foram dois fragmentos de DNA com o comprimento de 744bp

e 780 bp, em que ambos os fragmentos de DNA apresentam o mesmo 5'-terminal (ver também Fig. 7).

A partir destes dois fragmentos de DNA, as regiões 5' e 3' em falta do cDNA foram obtidas através de 5'- e 3'- Rapid Amplifikation dos terminais-cDNA (RACE), em que foi utilizado o RACE Kit de Gibco-BRL. Como iniciador antisense foi utilizado o iniciador de amplificação universal do Kit e como iniciador sense ou 5'-CTGGAACTGT-CCCTGTGGTT-3' (SEQ ID N°: 12) ou 5'-AGTGCACTAGAGGCCAGAA-3' (SEQ ID N°: 13). Também foi utilizado como iniciador sense o Anker-Primer encurtado do Kit e como iniciador antisense 5'-TTCGAGCACCACAATTGGAAAT-3' (SEQ ID N°: 14) ou 5'-GAATGCAAAGACGGCACGATGAAT-3' (SEQ ID N°: 15).

A PCR foi realizada com uma temperatura de emparelhamento de 55°C e as condições supra citadas. Os produtos 5' e 3' RACE foram subclonados no vector pCR2.1 e sequenciados: as sequências dos fragmentos subclonados foram sequenciadas com o método didesoxinucleótido (ABI PRISM Dye Terminator Cycle Sequencing Ready reaction Kit e ABI PRISM 310 Genetic analyser (Perkin Elmer)). Os iniciadores-para a frente T7 e M13 foram utilizados para a sequenciação dos produtos clonados no vector pCR2.1. Ambas as cadeias da região de codificação foram sequenciadas pelo Vienna VBC Genomics-Sequencing Service, em que foi utilizado o iniciador marcado com infravermelhos (IRD700 e IRD800) e um sequenciador LI-COR Long Read IR 4200 (Lincoln, NE).

As Fig. 5a e 5b mostram o cDNA total, que apresenta um tamanho de 2198 bp e uma grelha de leitura aberta de 1530 bp (SEQ ID N°: 1). A grelha de leitura aberta (codão de iniciação nos pares de bases 211-213, codão de terminação nos pares de bases 1740-1743) codifica uma proteína de 510 aminoácidos com um peso molecular de 56,8 kDa e um valor pI teórico de 7,51.

As Fig. 6a e 6b mostram a sequência de aminoácidos derivada do cDNA da GlcNAc- $\alpha$ 1,3-fucosiltransferase (SEQ ID N°: 2). Os locais para a glicosilação ligada à asparagina são no Asn346 e ASN 429.

Na Fig. 7 está representado o cDNA da GlcNAc- $\alpha$ 1,3-fucosiltransferase esquematicamente e o índice de hidrofobicidade derivado da proteína codificada (abaixo), em que um índice de hidrofobicidade positivo significa uma hidrofobicidade aumentada. Entre eles estão mostrados os tamanhos de ambos os produtos de PCR supra citados em relação com o cDNA completo. O domínio da codificação é representado através das barras, em que "C" codifica o domínio citoplasmático postulado, T o domínio transmembranar postulado e G o domínio catalítico no lúmen do Golgi postulado da transferase. A análise da sequência de DNA através de "TMpred" (de EMBnet, Suíça) deu uma região transmembranar provável entre Asn36 e Gly54. O domínio C-terminal da enzima abrange provavelmente a região catalítica e deverá consequentemente estar voltada para o lúmen do aparelho de Golgi. Por conseguinte esta transferase parece ser uma proteína transmembranar do tipo II, como

todas as glicosiltransferases analisadas até à data, que têm um papel na biossíntese de glicoproteínas (Joziasse, 1992 Glycobiology 2, 271-277). Os domínios cinzentos representam os quatro péptidos tripticos, os hexágonos os locais de N-glicosilação potenciais. Uma procura BLASTP acessível através de NCBI de todas as bases de dados mostrou uma semelhança entre a GlcNAc- $\alpha$ 1,3 fucosiltransferase e outras  $\alpha$ 1,3/4-fucosiltransferases, p.ex. a fucosiltransferase-VI humana. Com 18-21 % (investigado por SIM-LALVIEW, Expase, Suiça) a semelhança total era além de qualquer significância. No entanto um domínio da sequência de 35 aminoácidos (SEQ ID N°: 4) mostra uma homologia visivelmente elevada para outras  $\alpha$ 1,3/4 fucosiltransferases (Fig. 8). Este domínio da sequência encontra-se entre Glu267 e Pro301 da SEQ ID N°: 2.

Exemplo 5:

Expressão da GlcNAc- $\alpha$ 1,3 fucosiltransferase recombinante em células de insecto

A região que codifica a GlcNAc- $\alpha$ 1,3 fucosiltransferase provável incluindo a região citoplasmática e transmembranar foi amplificada com o iniciador-para a frente 5'-CGGCGGATCCGCAATTGAATGATG-3' (SEQ ID N°: 16) e o iniciador-para trás 5'-CCGGCTGCAGTACCATTTAGCGCAT-3' (SEQ ID N°: 17) com o sistema Expand High Fidelity PCR de Boehringer Mannheim. O produto da PCR foi duplamente digerido com PstI e BamHI e subclonado no vector de transferência pVL1393 do baculovírus tratado com alcalina-fosfatase, o qual foi

previamente digerido com PstI e BamHI. Para garantir uma recombinação homóloga, o vector de transferência foi cotransfetado com DNA viral Baculo Gold (PharMingen, San Diego, CA) em células de insecto Sf9 em meio IPL-41 com lipofectina. Após uma incubação de 5 dias a 27°C foram utilizados diferentes volumes do sobrenadante com o vírus recombinante para infecção das células de insecto Sf21. Após uma incubação de 4 dias a 27°C em meio IPL-41 com 5% FCS as células Sf1 foram colhidas e lavadas 2x com solução salina tamponada com fosfato. As células foram ressuspensas em 25 mM tampão Tris-HCl, pH 7,4, com 2% Triton X-100 e fracionadas através de sonificação em gelo.

Exemplo 6:

---

Ensaio para a actividade da GlcNAc  $\alpha$ 1,3 fucosiltransferase

O homogeneizado e o sobrenadante celular foram testados relativamente à GlcNAc  $\alpha$ 1,3 fucosiltransferase. Foram realizadas amostras cegas com baculovírus recombinante, que codifica a tabaco-GlcNAc-transferase I (Strasser et al. 1999, *Glycobiology*, em impressão).

A Fig. 9 mostra a actividade enzimática medida da GlcNAc- $\alpha$ 1,3-fucosiltransferase recombinante assim como do controlo negativo. No melhor dos casos a actividade enzimática das células cotransfetadas e do seu sobrenadante era 30x mais elevada que os controlos negativos. Esta actividade endógena, que é mensurável na ausência da transferase recombinante, deve-se essencialmente à insecto- $\alpha$ 1,6-

fucosiltransferase e apenas em pequena percentagem à GlcNAc- $\alpha$ 1,3-fucosiltransferase. Por conseguinte a elevação da GlcNAc- $\alpha$ 1,3-fucosiltransferase, que deriva dos baculovírus recombinantes, compreende muito acima das 100 vezes.

A enzima mostrou uma larga actividade máxima à volta de um valor de pH de 7,0, quando foi medida a actividade em tampão ácido 2-(N-morfolino)etanossulfónico-HCl. Como é visível na tabela 2, a adição de catiões bivalentes promove, especialmente Mn<sup>2+</sup>, a actividade da transferase recombinante.

Tabela 2

Aditivo (conc. 10 mM)	Actividade relativa (Aceitador : GnGn-péptido)
	%
Nenhum	21
EDTA	18
MnCl <sub>2</sub>	100
CaCl <sub>2</sub>	82
MgCl <sub>2</sub>	52
CdCl <sub>2</sub>	44
CoCl <sub>2</sub>	35
CuCl <sub>2</sub>	3
NiCl <sub>2</sub>	24
ZnCl <sub>2</sub>	0,6

A Tabela 3 mostra, que entre os aceitadores

utilizados o GnGn-péptido apresenta as taxas de incorporação mais elevadas com condições teste padrão, seguido de perto pelo GnGnF<sup>6</sup>-péptido e M5GnAsn. Não pôde ser verificado nenhuma transferência para o MM-péptido, o qual não apresenta o GlcNAc-terminal redutor na manose 3-ligada. Esta estrutura parece ser necessária para a núcleo-fucosil-transferase. A transferase recombinante era além disso inactiva relativamente aos aceitadores usuais, as  $\alpha$ 1,3/4-fucosiltransferases para a determinação dos grupos sanguíneos, que transferem a fucose de GlcNAc para os terminais não redutores de oligossacáridos. Os valores  $K_m$  aparentes para os substratos dos aceitadores GnGn-péptido, GnGnF<sup>6</sup>-péptido, M5Gn-Asn e para o substrato doador GDP-fucose foram estimados em 0,19, 0,13, 0,23 ou 0,11. As estruturas das moléculas estão representadas na Fig. 10a e 10b.

Tabela 3

Aceitador-Substrato	Actividade relativa	Valor $K_m$
	%	mM
GnGn-péptido	100	0,19
GnGnF <sup>6</sup> -péptido	87	0,13
M5Gn-Asn	71	0,23
MM-péptido	0	
Gal $\beta$ 1-4GlcNAc	0	
Gal $\beta$ 1-3GlcNAc	0	
Gal $\beta$ 1-3GlcNAc $\beta$ 1-	0	
3Gal $\beta$ 1-4Glc		

Exemplo 7:

Espectrometria de massa do produto da fucosil-transferase

O GnGn-hexaméptido dabsilado (2nmol) foi incubado com o homogeneizado de células de insecto abrangendo a GlcNAc- $\alpha$ 1,3 fucosiltransferase recombinante (0,08 mU) na presença de GDP-L-fucose não radioactiva (10 nmol), tampão ácido 2-(N-morfolino)etanossulfónico-HCl, Triton X-100, MnCl<sub>2</sub>, GlcNAc e AMP. Foi realizado um controlo negativo com um homogeneizado das células de insecto infectadas para as amostras cegas. As amostras foram incubadas durante 16h a 37°C e analisadas através de espectrometria de massa MALDI TOF.

A espectrometria de massa foi realizada num DYNAMO (Therrmo BioAnalysis, Santa Fe, NM), um MALDI-TOF MS, que tem a capacidade de extração dinâmica (sinónimo para extração tardia). Foram utilizados dois tipos de preparações de matriz da amostra: péptidos e glicopéptidos dabsilados foram dissolvidos em 5% ácido fórmico e foram aplicadas alíquotas no alvo, seco ao ar e coberto com 1% ácido  $\alpha$ -ciano-4hidroxicinâmico. Glicanos piridilaminados, oligossacáridos redutores e glicopéptidos não derivados foram diluídos com água, aplicados no alvo e secos ao ar. Após adição de 2% ácido 2,5-dihidroxibenzóico as amostras foram imediatamente secas através da colocação de um vácuo.

A Fig. 11 mostra o espetro de massa destas

amostras, em que A representa o controlo negativo: o pico principal (S) mostra o substrato dabsil-Val-Gly-Glu-(GlcNAc<sub>4</sub>Man<sub>3</sub>)Asn-Arg-Thr, em que o valor  $[M+H]^+$  calculado comprehende 2262,3. Este substrato aparece também como produto de adição sódio e como ião mais pequeno, que surgiu através da fragmentação da função azo do grupo dabsil, em (S\*). Uma pequena quantidade de produto (P,  $[M+H]^+ = 2408,4$ ) é atribuída à  $\alpha 1,6$ -fucosiltransferase endógena. O pico a  $m/z = 2424,0$  mostra a degalactosilação incompleta do substrato. O espectro de massa B mostra a amostra com  $\alpha 1,3$ -fucosiltransferase recombinante. O pico principal (P) representa o produto fucosilado, (P\*) o seu ião fragmentado.

Adicionalmente foram misturadas alíquotas de ambas as amostras, para se obter concentrações de substrato e produto semelhantes (amostra A). Esta mistura foi diluída com 0,1 M acetato de amónio, pH 4,0, abrangendo 10  $\mu$ U N-glicosidase A (amostra B) ou com 50 mM Tris/HCl, pH 8,5, abrangendo 100  $\mu$ U (1 U hidroliza 1  $\mu$ mol substrato por min) N-glicosidase F (amostra C). Após 2 e 20 h foram retiradas pequenas alíquotas destas misturas e analisadas através de MALDI-TOF MS.

Na Fig. 12 estão representados os três espectros de massa das amostras A, B e C. A amostra A não digerida mostra dois picos principais: o substrato a 2261,4  $m/z$  e o produto fucosilado a 2407,7  $m/z$ . A curva no meio mostra o espectro de massa da amostra B, tratada com N-glicosidase A, que hidrolisa ambos os glicopéptidos. O pico a 963,32

representa o produto deglicosilado. A curva inferior mostra o espectro de massa da amostra C. A N-glicosidase F não tem a capacidade de hidrolisar o substrato  $\alpha$ 1,3 fucosilado, de forma que o espectro apresenta o pico a 2406,7 m/z do produto fucosilado, enquanto o pico do substrato hidrolisado aparece a 963,08 m/z.

Exemplo 8:

Análise HPLC do produto fucosiltransferase piridilaminado

Ambas as amostras acima descritas (produto fucosilado e controlo negativo) foram digeridas com N-glicosidase A. Os oligossacáridos obtidos foram piridilaminados e analisados através de HPLC de fase reversa (Wilson et al. 1998 Glycobiology 8, 651-661; Kubelka et al. 1994 Arch.Biochem.Biophys. 308, 148-157; Hase et al. 1984 J.Biochem. 95, 197-203).

Na Fig. 13 o diagrama superior B representa o controlo negativo, em que é visível adicionalmente ao substrato residual (GnGn-péptido) o produto  $\alpha$ 1,6 fucosilado. A mostra um pico com um tempo de retenção substancialmente menor, o que é específico para a fucose GlcNAc- $\alpha$ 1,3-ligada redutora.

No diagrama inferior foi comparado o produto da transferase isolado antes (curva A) e depois da digestão através de N-acetil- $\beta$ -glucosaminidase (curva B) com MMF<sup>3</sup> de fosfolipase A<sub>2</sub> das abelhas produtoras de mel (curva C).

## LISTAGEM DE SEQUÊNCIAS

&lt;110&gt; Altmann Dr., Friedrich

&lt;120&gt; Gene da fucosiltransferase

&lt;130&gt; R 53926 (divisão de R 46233)

&lt;150&gt; A 270/99

&lt;151&gt; 1999-02-18

&lt;150&gt; EP 05108012.5

&lt;151&gt; 2000-02-17

&lt;160&gt; 17

&lt;170&gt; Versão da patente 2.1

&lt;210&gt; 1

&lt;211&gt; 2198

&lt;212&gt; DNA

&lt;213&gt; vigna radiata

&lt;400&gt; 1

actaactcaa	acgctgcatt	ttctttttc	tttcagggaa	ccatccaccc	ataacaacaa	60
aaaaaaacaac	agcaagctgt	gttttttta	tcgttcttt	tcttaaaca	agcaccccca	120
tcatgaaatc	gtgctcataa	cgccaaaatt	ttccattttc	cttgatttt	tagtttattt	180
tgcgaaatg	gcagttgggg	gchgcaattt	atgatgggtc	tgttgacgaa	tcttcgaggc	240
tcgagaacag	atggtgccca	acaagacagc	ttaccgggtt	ttggctccggg	aggcaaccca	300
aagagggaaat	ggagcaatct	aatgcctctt	tttgcgttgc	tttgtggtcat	cgcggagatc	360
gcgtttctgg	gttaggttgg	tatggccaaa	aacgcggca	ttgggtgactc	cctcgctgac	420
ttcttcttacc	gctctcgagc	ggtcgttgg	ggtagcgtt	ttgggttggg	tttgggtggct	480
tctgatcgga	attctgttata	gtatagttgt	gagaatgtt	ttggagagggg	ggatgtgtc	540
acgttattcg	ggggcttttc	caaagagctt	tttttttttt	ttggagctgt	tcaggagtgg	600
aagtctgttt	cggttggatg	taaattttgg	tttagtgggg	atagaaagcc	agatgcgc	660
tttgggttac	ctcaaccaag	tggAACAGCT	agcattctgc	gatcaatgg	atcagcagaa	720
tactatgtct	agaacaatata	tgccatggc	agacggaggg	gatataacat	cgtaatgaca	780
accagtctat	tttcggatgt	ttctgttgg	tatittttat	gggtctgatgt	tgatatgtat	840
gcaccagtgc	agccggaaaac	tgaagctgt	tttgctgt	ttttccattt	caattgtgtt	900
gctcgaaatt	tccgggttgc	agctcttgag	gcccgttgc	atcaaaacat	caaaattgtat	960
tcttatgttgc	gttgtcacag	gaaccgtgat	ggaagagtga	acaaagtgg	agccctgt	1020
cactacaaat	ttagctttagc	gtttggaaaat	tgcgttgc	agatgtatgt	aactgaaaaaa	1080
ttcttccaaat	cccttgggtgc	tggactgtc	cctgtgggt	tttgtgtct	aaatatttcag	1140
gacttgc	tttctccctgg	ttaatttttca	catattaaag	agatagggaa	tgttggatgt	1200
gttgcggaa	ccatgagata	tcttagcagaa	aatcccgaag	catataatca	atcattgggg	1260
tggaaagtatg	agggtccatc	tgactccctc	aggcccctt	tggatgtggc	agctgtgtc	1320
tcacgtgtcc	gtctttgtcat	tcacttggc	acatgtgatg	gagagaaggaa	agaaaataat	1380
ccaaagctta	agagacgttc	tttgcgttgc	actagagggc	cagaaggccgt	atatcatatc	1440
tatgtcagag	aaaggggaaag	gtttggatgt	gagtccattt	acctggggc	tagcaattt	1500
actctgaatg	ctgtgaaggc	tgctgtgtt	tttgcgttca	catccctgaa	tcttgcct	1560
gtatggaaaga	ctgaaaggcc	tgaagtata	agagggggga	gtgtttttaaa	actctacaaa	1620
atatacccaa	ttggcttgcac	acagagacaa	gctttttata	ccttcgttgc	caaagggtat	1680
gctgtttca	ggagtactt	ggagaacaaat	ccttgcgttgc	agtttggatgt	cattttgtgt	1740
tagcatgc	taaatggta	ctctgtctt	cctgttgcatt	tttttttttt	ctgagacta	1800
gcttaggttt	taggaatgt	tatggcgtt	aatatggcat	ggctttttttt	atgcctagtt	1860
tcttggccaa	ctcattgtat	ttttgtataa	gacatcacac	ttaatttttta	aacttggttc	1920
tgttagaagt	caaatccata	tttaatgttt	agtttttagt	tttttttttt	atcatcttgc	1980
agtcacagg	tttgcgttata	gttgcgttgc	actgaaatct	aatagaaggg	tcagatgttt	2040
cactcaagac	acattattac	tttgcgttgc	tttgcgttgc	tttttttttt	tttgtgtct	2100
gaactgtccc	tgtgggttgc	gcacgttta	tttgcgttgc	gttactgtcc	agtgggttac	2160
gtttttgacc	tctaaaaaaa	aaaaaaa	aaaaaaa	aaaaaaa	2198	

&lt;210&gt; 2

&lt;211&gt; 510

&lt;212&gt; PRT

&lt;213&gt; vigna radiata

&lt;400&gt; 2

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

385 390 395 400  
Thr Arg Gly Pro Glu Thr Val Tyr His Ile Tyr Val Arg Glu Arg Gly  
405 410 415  
Arg Phe Glu Met Glu Ser Ile Tyr Leu Arg Ser Ser Asn Leu Thr Leu  
420 425 430  
Asn Ala Val Lys Ala Ala Val Val Leu Lys Phe Thr Ser Leu Asn Leu  
435 440 445  
Val Pro Val Trp Lys Thr Glu Arg Pro Glu Val Ile Arg Gly Gly Ser  
450 455 460  
Ala Leu Lys Leu Tyr Lys Ile Tyr Pro Ile Gly Leu Thr Gln Arg Gln  
465 470 475 480  
Ala Leu Tyr Thr Phe Ser Phe Lys Gly Asp Ala Asp Phe Arg Ser His  
485 490 495  
Leu Glu Asn Asn Pro Cys Ala Lys Phe Glu Val Ile Phe Val  
500 505 510

&lt;210&gt; 3

&lt;211&gt; 105

&lt;212&gt; DNA

&lt;213&gt; Sequência artificial

&lt;220&gt;

&lt;223&gt; Descrição da sequência artificial:cDNA

&lt;400&gt; 3

gaagccctga agcactacaa **atttagctta** gcgttgaaa **attcgaatga** ggaagattat 60  
gtaactgaaa aattcttcca atccctgtt gctggaactg tccct 105

&lt;210&gt; 4

&lt;211&gt; 35

&lt;212&gt; PRT

&lt;213&gt; Sequência artificial

&lt;220&gt;

&lt;223&gt; Descrição da sequência artificial:fragmento de péptido

&lt;400&gt; 4

Glu Ala Leu Lys His Tyr Lys Phe Ser Leu Ala Phe Glu Asn Ser Asn  
1 5 10 15Glu Glu Asp Tyr Val Thr Glu Lys Phe Phe Gln Ser Leu Val Ala Gly  
20 25 30Thr Val Pro  
35

&lt;210&gt; 5

&lt;211&gt; 15

&lt;212&gt; PRT

&lt;213&gt; Sequência artificial

&lt;220&gt;

&lt;223&gt; Descrição da sequência artificial:fragmento de péptido

&lt;220&gt;

&lt;221&gt; característica\_misc

&lt;222&gt; (5)

&lt;223&gt; incerto

<400> 5  
Lys Pro Asp Ala Xaa Phe Gly Leu Pro Gln Pro Ser Thr Ala Ser

1 5 10 15

<210> 6  
<211> 10  
<212> PRT  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:fragmento de péptido

<400> 6  
Pro Glu Thr Val Tyr His Ile Tyr Val Arg  
1 5 10

<210> 7  
<211> 13  
<212> PRT  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:fragmento de péptido

<400> 7  
Met Glu Ser Ala Glu Tyr Tyr Ala Glu Asn Asn Ile Ala  
1 5 10

<210> 8  
<211> 10  
<212> PRT  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:fragmento de péptido

<400> 8  
Gly Arg Phe Glu Met Glu Ser Ile Tyr Leu  
1 5 10

<210> 9  
<211> 29  
<212> DNA  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:iniciador universal

<220>  
<221> característica\_misc  
<222> (3), (15)  
<223> n é a, c, g ou t

<400> 9  
gcngartayt aygcngaraa yaayathgc 29

<210> 10  
<211> 22  
<212> DNA  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:iniciador universal

<220>  
<221> característica\_misc  
<222> (14), (17)  
<223> n é a, c, g ou t

<400> 10  
crtadatrtg rtanacngty tc 22

<210> 11  
<211> 20  
<212> DNA  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:iniciador universal

<220>  
<221> característica\_misc  
<222> (6)  
<223> n é a, c, g ou t

<400> 11  
tadatnswyt ccatytcraa 20

<210> 12  
<211> 20  
<212> DNA  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:iniciador

<400> 12  
ctggaaactgt ccctgtggtt 20

<210> 13  
<211> 20  
<212> DNA  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:iniciador

<400> 13  
agtgcactag agggccagaa 20

<210> 14  
<211> 22  
<212> DNA  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:iniciador

<400> 14  
ttcgagcacc acaattggaa at 22

<210> 15  
<211> 24  
<212> DNA  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:iniciador

<400> 15  
gaatgcaaag acggcacgat gaat 24

<210> 16  
<211> 24  
<212> DNA  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:iniciador

<400> 16  
cggcggatcc gcaattgaat gatg 24

<210> 17  
<211> 25  
<212> DNA  
<213> Sequência artificial

<220>  
<223> Descrição de sequência artificial:iniciador

<400> 17  
ccggctgcag taccatttag cgcat 25

**REIVINDICAÇÕES**

1. Processo para produção de glicoproteínas recombinantes, abrangendo a produção de uma glicoproteína recombinante em plantas ou células de plantas, cuja produção de GlcNAc- $\alpha$ 1,3-fucosiltransferase está reprimida ou completamente impedida e cuja actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase endógena compreende menos de 50% da actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase que ocorre naturalmente em plantas ou células de plantas, e cuja GlcNAc- $\alpha$ 1,3-fucosiltransferase é codificada por uma molécula de DNA, que abrange uma sequência de acordo com a SEQ ID N°. 1 com uma grelha de leitura aberta do par de bases 211 ao par de bases 1740 ou apresenta pelo menos 50% de homologia com a sequência supra citada ou hibridiza sob condições rigorosas com a sequência supra citada ou abrange uma sequência, que é degenerada como resultado do código genético para a sequência de DNA supra citada.

2. Processo de acordo com a reivindicação 1, caracterizado por, as glicoproteínas recombinantes serem proteínas humanas, especialmente proteínas para utilização terapêutica.

3. Processo de acordo com a reivindicação 1 ou 2, caracterizado por, a actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase endógena compreender menos de 20% da actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase que ocorre naturalmente em plantas e células de plantas.

4. Processo de acordo com uma das reivindicações 1 a 3, caracterizado por, a actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase endógena compreender 0% da actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase que ocorre naturalmente em plantas e células de plantas.

5. Processo de acordo com uma das reivindicações 1 a 4, caracterizado por, a redução da actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase ocorrer através de inibição antisense, em que é utilizado um polinucleótido, que é pelo menos parcialmente complementar à sequência da molécula de DNA.

6. Processo de acordo com uma das reivindicações 1 a 4, caracterizado por, a redução da actividade da GlcNAc- $\alpha$ 1,3-fucosiltransferase ocorrer através de mutação knock-out da sequência da molécula de DNA.

Lisboa, 31 de março de 2014

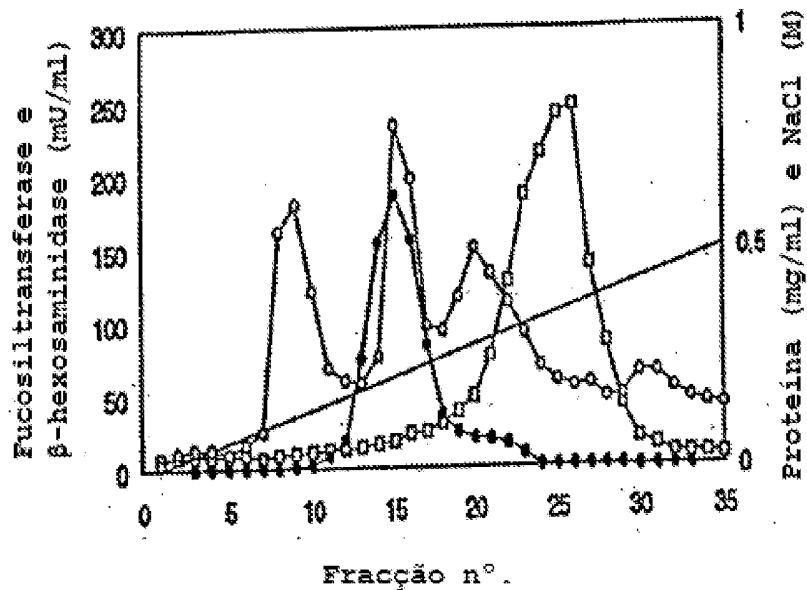


FIG.1 a

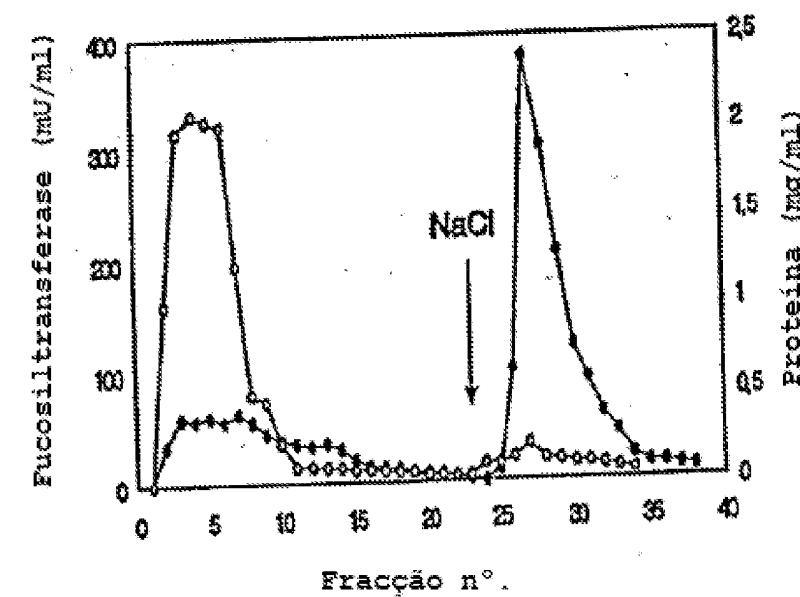


FIG.1 b

PE2062977

2/16

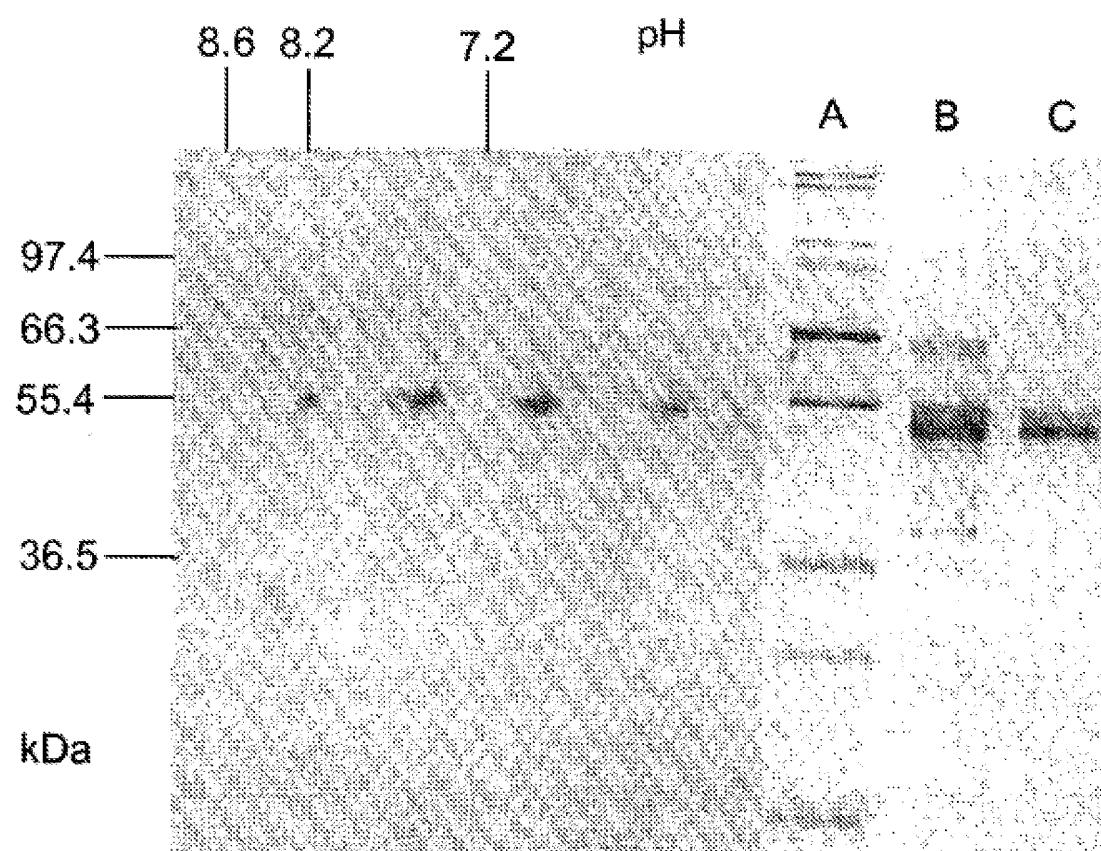


FIG. 2

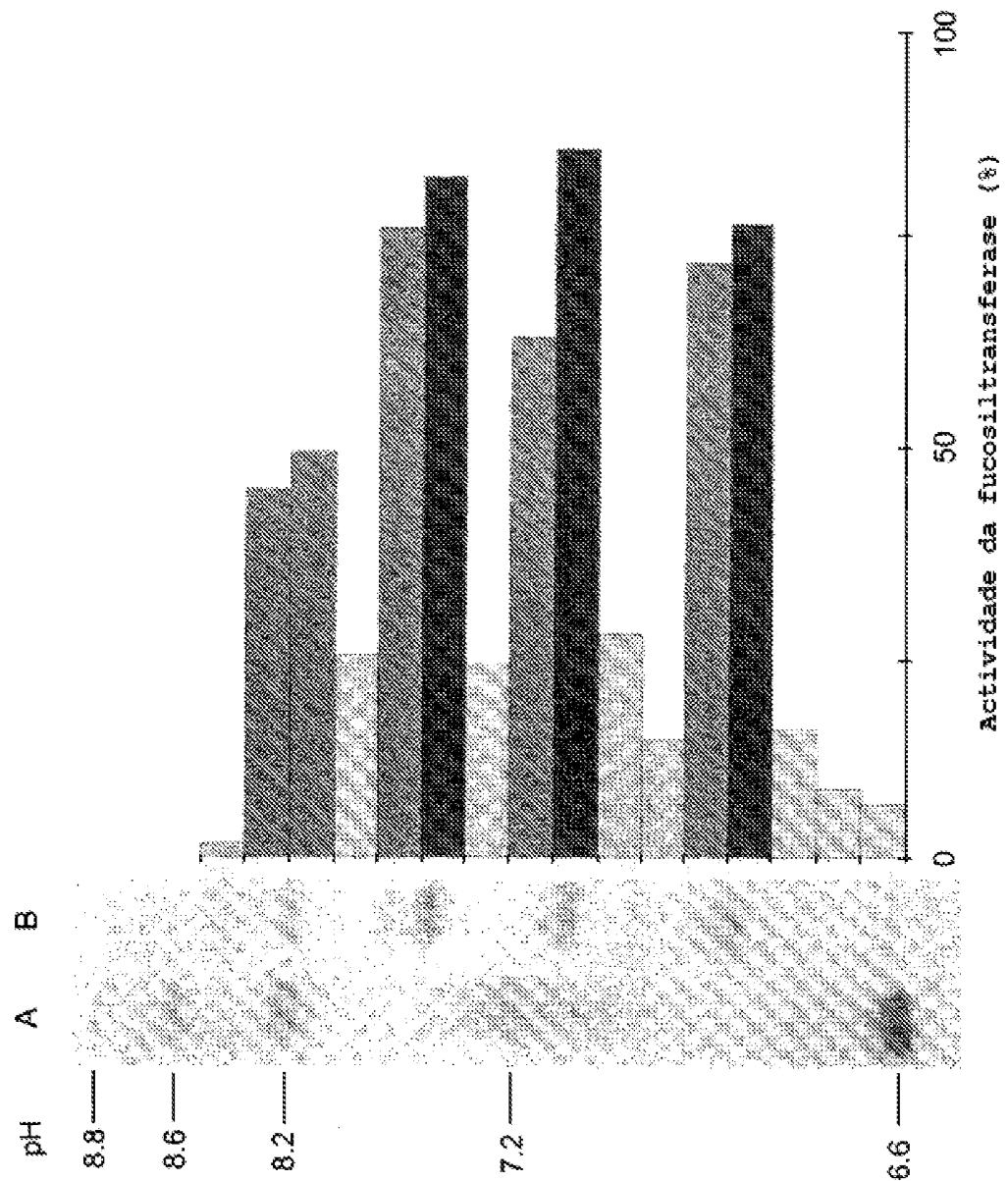


FIG. 3

1 KPDAxFGLPQPSTAS  
2 PETVYHIVVR  
3 MESAELYAENNIA  
4 GRFEMESIYL

S1 5'- GCIGAATACTACGCGIGAAAACAACATCGC -3'  
      G T T G T T T T

A2 5'- CATAGATATGATAIACIGTCTC -3'  
      G T G G T

A3 5'- TAGATICACTCCATCTCAAA -3'  
      T G T G

FIG.4

ACTAACTCAA	ACGCTGCATT	TTCTTTTTTC	TTTCAGGGAA	CCATCCACCC	ATAACAAACAA	60
AAAAAAACAAC	AGCAAGCTGT	GTTCCTTTTA	TCTTCTTTT	TCTTAAACA	AGCACCCCCA	120
TCATGGAATC	GTGCTCATAA	CGCCAAAATT	TTCCATTCC	CTTGATTTT	TAGTTTATT	180
TGCGGAATTG	GCAGTTGGGG	GGGCAATTCA	ATGATGGGTC	TGTTGACGAA	TCTTCGAGGC	240
TCGAGAACAG	ATGGTCCCCA	ACAAGACAGC	TTACCCGTTT	TGGCTCCGGG	AGGCAACCCA	300
AAGAGGAAAT	GGAGCAATCT	AATGCTCTT	GTGTTGCCC	TTGTTGGTCAT	GGGGAGATC	360
GGCTTTCTGG	GTAGGTTGGA	TATGGCCAAA	AACGGGGCCA	TGGTTGACTC	CTCTGCTGAC	420
TTCTTCTACC	GCTCTGGAGC	GGTCGTTGAA	GGTGACGATT	TGGGTTGGG	TTGGTGGCT	480
TCTGATCGGA	ATTCTGAATC	GTATAGTTGT	GAGGAATGGT	TGGAGAGGGG	GGATGCTGTC	540
ACGTATTGCA	GGGGCTTTTC	CNAAGAGCCT	ATTTTTGTTT	CTGGAGCTGA	TCAGGAGTGC	600
AAAGTGTGTT	CGGTTGGATG	TAATTTGGG	TTTATGGGG	ATAGAAAGCC	AGATGGGGCA	660
TTTGGGTAC	CTCAACCAAG	TGGAACAGCT	ACCATCTSC	GATCAATGGA	ATCAGCAGAA	720
TACTATGCTG	AGAACAAAT	TGCCATGGCA	AGACGGAGGG	GATATAACAT	CGTAATGACA	780
ACCAAGTCTAT	CTTCGGATGT	TCCGTTGGA	TATTTTCTAT	GGGCTGAGTA	TGATATGATG	840
GCACCAAGTGC	AGCCGAAAC	TGAAAGCTCT	CTTGCAGCTG	CTTTCATTTC	CAATTGTGGT	900
GCTCGAAATT	TCCGGTTGCA	AGCTCTTGAG	GCCCTTGAAA	AATCAAACAT	CAAATTGAT	960
TCTTATGGTG	GTTGTCACAG	GAACCGTGAT	GGAAAGAGTGA	ACAAAGTGGA	AGCCCTGAAG	1020
CACTACAAT	TTAGCTTAGC	GTTCGAAAAT	TCGAATGAGG	AAGATTATGT	AACTGAAAAA	1080
TTCTTCCAAAT	CCCTTGTGTC	TGGAACGTGTC	CCTGTGGTTG	TTGGTGCTCC	AAATATTCAAG	1140

FIG.5 a

GACTTTGCTC CTTCTCTGG TTCAATTITA CATATTAAAG AGATAGAGGA TGTTGAGTCT	1200
GTTGCAAAGA CCATGAGATA TCTAGCAGAA AATCCCGAAG CATATAATCA ATCATTGAGG	1260
TGGAAGTATG AGGGTCCATC TGACTCCCTTC AAGGCCCTTG TGGATATGGC AGCTGTGCAT	1320
TCATCGTGCC GTCCTTGCA TCACCTGGCC ACAGTGAGTA GAGAGAAGGA AGAAAATAAT	1380
CCAAAGCCTTA AGAGACGTCC TTGCAAGTGC ACTAGAGGGC CAGAAACCGT ATATCATATC	1440
TATGTCAGAG AAAGGGGAAG GTTTGAGATG GAGTCATTT ACCTGAGGTC TAGCAATTIA	1500
ACTCTGAATG CTGTGAASSGC TGCTGTTGTT TTGAAAGTCA CATCCCTGAA TCTTGTGCCT	1560
GTATGGAAGA CTGAAAGGCC TGAAAGTTATA AGAGGGGGGA GTGCTTTAAA ACTCTACAAA	1620
ATATACCCAA TTGGCTTGAC ACAGAGACAA CCTCTTTATA CCTTCACCTT CAAAGGTGAT	1680
GCTGATTTCAGGACTT GGAGAACAAAT CCTTGTGCCA AGTTGAACT CATTTCGTG	1740
TAGCCTGCGC TAAATGGTAC CTCTGCTCTA CCTGAATTAG CTTCACTTAG CTGAGCACTA	1800
GCTAGAGTTT TAGGAATGAG TATGGCAGTG AATATGGCAT GGCTTTATTT ATGCCTAGTT	1860
TCTTGGCCAA CTCATTGATG TTTGTATAA GACATCACAC TTTAATTATA AACTTGTTC	1920
TGTAGAAGTG CAAATCCATA TTTAATGCTT AGTTTGTG CTCTTATCTG ATCATCTAGA	1980
ACTCACAGTT CTTGTATATT GTGAGTGAAA ACTGAAACTT AATAGAAGGA TCAGATGTTT	2040
CACTCAAGAC ACATTATTAC TTCAATGTTGT TTTGATGATC TCGAGCTTTT TTAGTGTCTG	2100
GAACCTGCCC TGTGGTTTGA GCACCTGTTA TTGCTTCAGT GTTACTGTCC AGTGGTTATC	2160
GTTTTGACCC TCTAAAAAAA AAAAAAAA AAAAAAAA	2198

FIG.5 b

Met Met Gly Leu Leu Thr Asn Leu Arg Gly Ser Arg Thr Asp Gly Ala  
 1 — 5 10 15

Gin Gin Asp Ser Leu Pro Val Leu Ala Pro Gly Gly Asn Pro Lys Arg  
 20 25 30

Lys Trp Ser Asn Leu Met Pro Leu Val Val Ala Leu Val Val Ile Ala  
 35 40 45

Glu Ile Ala Phe Leu Gly Arg Leu Asp Met Ala Lys Asn Ala Ala Met  
 50 55 60

Val Asp Ser Leu Ala Asp Phe Phe Tyr Arg Ser Arg Ala Val Val Glu  
 65 70 75 80

Gly Asp Asp Leu Gly Leu Gly Leu Val Ala Ser Asp Arg Asn Ser Glu  
 85 90 95

Ser Tyr Ser Cys Glu Glu Trp Leu Glu Arg Glu Asp Ala Val Thr Tyr  
 100 105 110

Ser Arg Gly Phe Ser Lys Glu Pro Ile Phe Val Ser Gly Ala Asp Gln  
 115 120 125

Glu Trp Lys Ser Cys Ser Val Gly Cys Lys Phe Gly Phe Ser Gly Asp  
 130 135 140

Arg Lys Pro Asp Ala Ala Phe Gly Leu Pro Gln Pro Ser Gly Thr Ala  
 145 150 155 160

Ser Ile Leu Arg Ser Met Glu Ser Ala Glu Tyr Tyr Ala Glu Asn Asn  
 165 170 175

Ile Ala Met Ala Arg Arg Arg Gly Tyr Asn Ile Val Met Thr Thr Ser  
 180 185 190

Leu Ser Ser Asp Val Pro Val Gly Tyr Phe Ser Trp Ala Glu Tyr Asp  
 195 200 205

Met Met Ala Pro Val Gln Pro Lys Thr Glu Ala Ala Leu Ala Ala Ala  
 210 215 220

Phe Ile Ser Asn Cys Gly Ala Arg Asn Phe Arg Leu Gln Ala Leu Glu  
 225 230 235 240

Ala Leu Glu Lys Ser Asn Ile Lys Ile Asp Ser Tyr Gly Gly Cys His

FIG.6 a

Arg Asn Asn Asp Gly Arg Val Asn Lys Val Glu Ala Leu Lys His Tyr  
 260 265 270  
 Lys Phe Ser Leu Ala Phe Glu Asn Ser Asn Glu Glu Asp Tyr Val Thr  
 275 280 285  
 Glu Lys Phe Phe Gln Ser Leu Val Ala Gly Thr Val Pro Val Val Val  
 290 295 300  
 Gly Ala Pro Asn Ile Gln Asp Phe Ala Pro Ser Pro Gly Ser Ile Leu  
 305 310 315 320  
 His Ile Lys Glu Ile Glu Asp Val Glu Ser Val Ala Lys Thr Met Arg  
 325 330 335  
 Tyr Leu Ala Glu Asn Pro Glu Ala Tyr Asn Gln Ser Leu Arg Trp Lys  
 340 345 350  
 Tyr Glu Gly Pro Ser Asp Ser Phe Lys Ala Leu Val Asp Met Ala Ala  
 355 360 365  
 Val His Ser Ser Cys Arg Leu Cys Ile His Leu Ala Thr Val Ser Arg.  
 370 375 380  
 Glu Lys Glu Glu Asn Asn Pro Ser Leu Lys Arg Arg Pro Cys Lys Cys  
 385 390 395 400  
 Thr Arg Gly Pro Glu Thr Val Tyr His Ile Tyr Val Arg Glu Arg Gly  
 405 410 415  
 Arg Phe Glu Met Glu Ser Ile Tyr Leu Arg Ser Ser Asn Leu Thr Leu  
 420 425 430  
 Asn Ala Val Lys Ala Ala Val Val Leu Lys Phe Thr Ser Leu Asn Leu  
 435 440 445  
 Val Pro Val Trp Lys Thr Glu Arg Pro Glu Val Ile Arg Gly Gly Ser  
 450 455 460  
 Ala Leu Lys Leu Tyr Lys Ile Tyr Pro Ile Gly Leu Thr Gln Arg Gln  
 465 470 475 480  
 Ala Leu Tyr Thr Phe Ser Phe Lys Gly Asp Ala Asp Phe Arg Ser His  
 485 490 495  
 Leu Glu Asn Asn Pro Cys Ala Lys Phe Glu Val Ile Phe Val  
 500 505 510

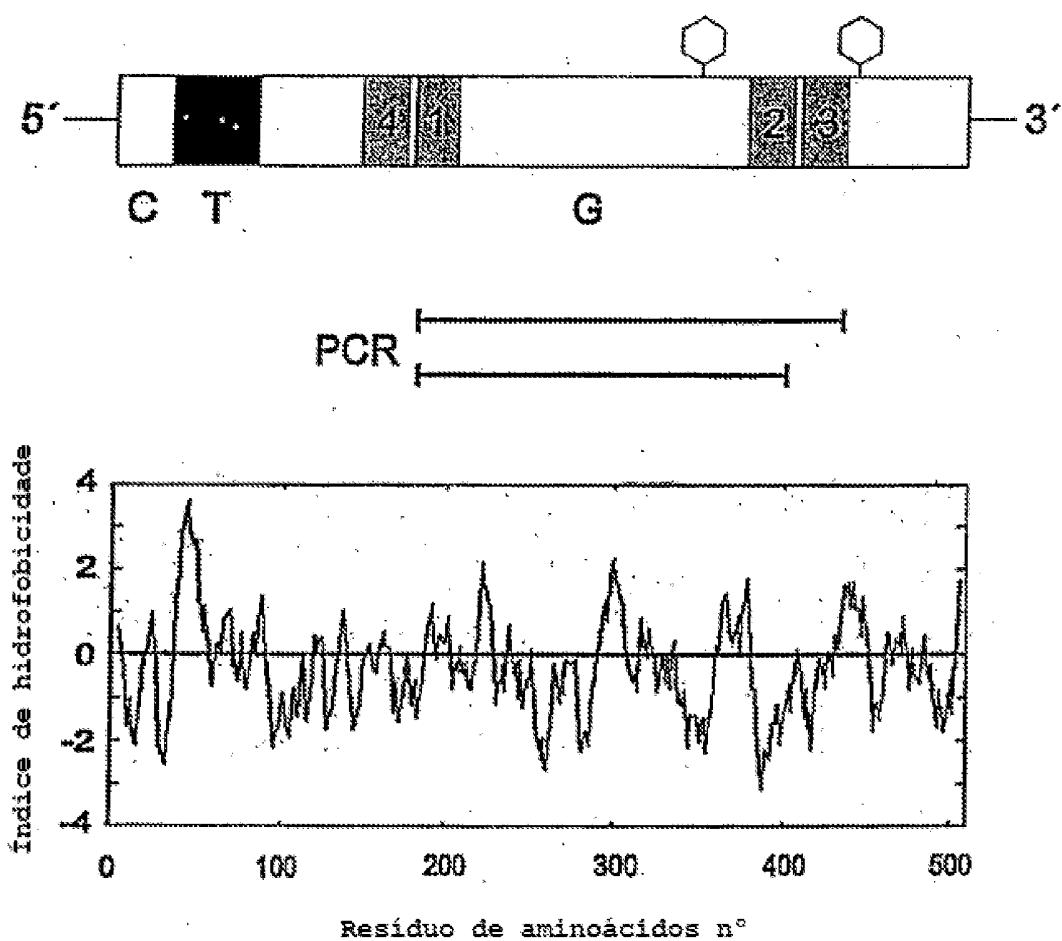


FIG.7

Transferase Tipo \* Motivo conservado

Fuct-C3	Mungo Bohne	P21217	267 - EALKKXKESTLAEENSNEEDYTEKEFQ-SILVACTVP
Fuct-III	Mensch	Q16938	236-ELTISRRIKEYTAAEENSLEPDVYIIEKLMRNLLERWAVP
Fuct-III	Schimpanse	Q11128	247 - ETLISRRIKEYTAAEENSLEPDVYIIEKLMRNLLERWAVP
Fuct-V	Mensch	P56433	249 - ETLISRRIKEYTAAEENSLEPDVYIIEKLMRNLLERWAVP
Fuct-V	Schimpanse	P51993	249 - ETLISRRIKEYTAAEENSLEPDVYIIEKLMRNLLERWAVP
Fuct-VI	Mensch	P56434	235 - ETLISRRIKEYTAAEENSLEPDVYIIEKLMRNLLERWAVP
Fuct-VI	Schimpanse	Q11126	235 - ETLISRRIKEYTAAEENSLEPDVYIIEKLMRNLLERWAVP
Fuct-VII	Rind	Q35886	249 - KQLSQQKEYTAAEENSLEPDVYIIEKLMRNLLQAVAVP
Fuct-VII	chin. Hamster	Q11130	237 - GHLARKKAYTAAEENSLEPDVYIIEKLMRNLLERWAVP
Fuct-VII	Mensch	Q11131	217 - PTVAQTYRIVTSENSQIRDQXTEKEFWRNALVACTVP
Fuct-VII	Maus	076204	264 - PTIAKRPVYIATENSQIRDQXTEKEFWRNALAAGAVP
Fuct-VII	Schistosoma mansoni	P22083	226 - PTIAKRPVYIATENSQIRDQXTEKEFWRNALAAGAVP
Fuct-IV	Mensch	Q11127	277 - HTVVARXKKEYTAAEENSQIRDQXTEKEFWRNALLAGAVP
Fuct-IV	Maus	Q62894	305 - HTVVARXKKEYTAAEENSQIRDQXTEKEFWRNALLAGAVP
Fuct-IV	Ratte	Q98952	305 - HTVVARXKKEYTAAEENSQIRDQXTEKEFWRNALLAGAVP
Fuct-IV	Huhn	088819	228 - KTVSAKXKKEYTAAEENSQIRDQXTEKEFWRNALLAGAVP
Fuct-IX	Maus	076544	233 - PTISRCKTLYLSTENSTIKDIXTEKLYNAPF-AGSTVP
FutA	Dicyostelium discoideum	032631	247 - DYLKRNFAATAEENSQIRDQXTEKEFWRNALVSETIP
IpFact1	Dictyostelium discoideum	036511	106 - EFLISQXKRNLAATCENSQGCGXVTEKLL-AYFSETIP
IpFact2	Helicobacter pylori	Q21362	227 - EFLISQXKRNLAATCENSQGCGXVTEKLL-AYFSETIP
CEFT-1	Caenorhabditis elegans		303 - MLDDTDXHEVYVTFENSICEDYVTERLINKSGYQNTLIP

\* Número de acesso para a Base de Dados de Proteínas SwissProt

FIG.8

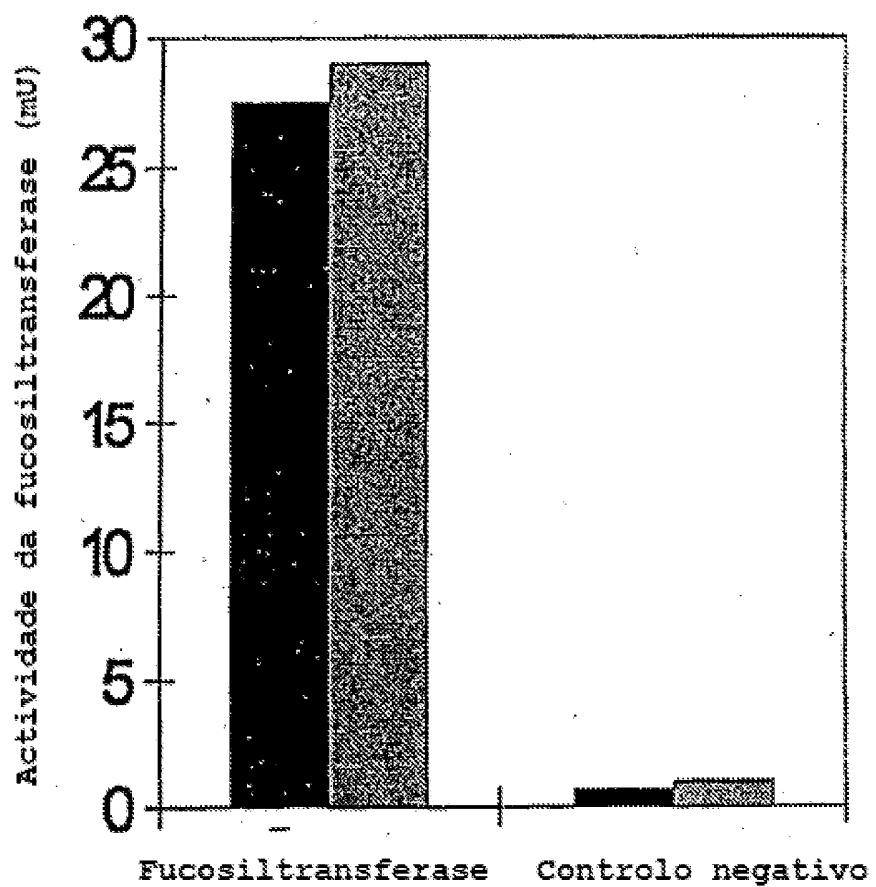


FIG.9

PE2062977

12/16

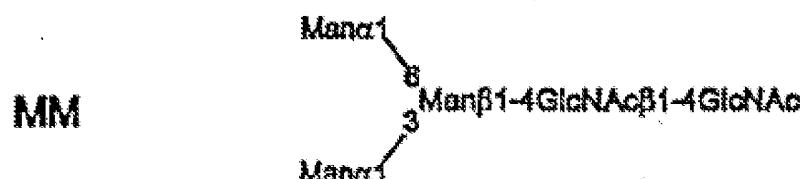
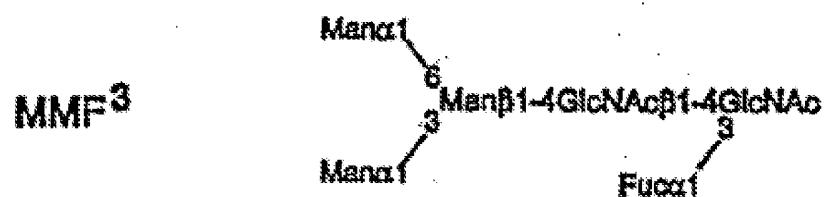
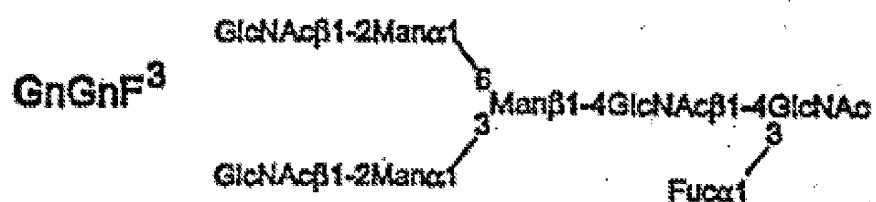
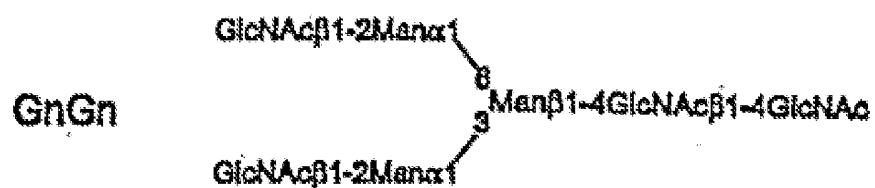
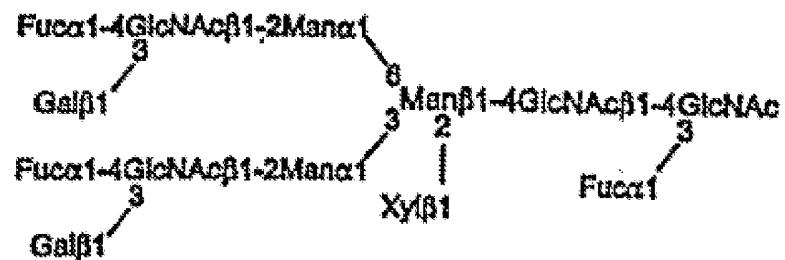


FIG.10 a

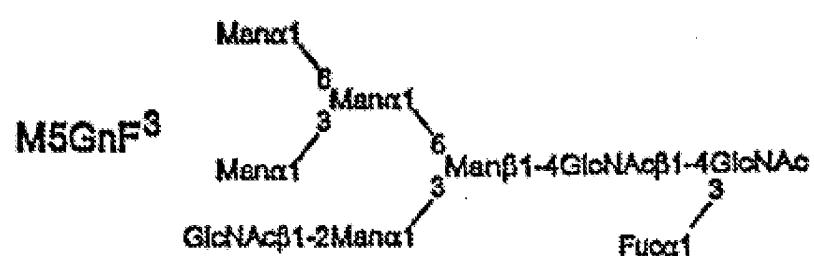
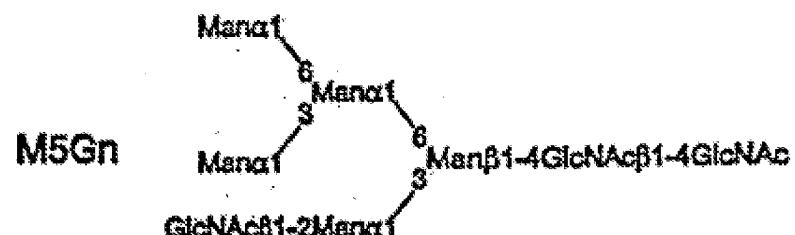
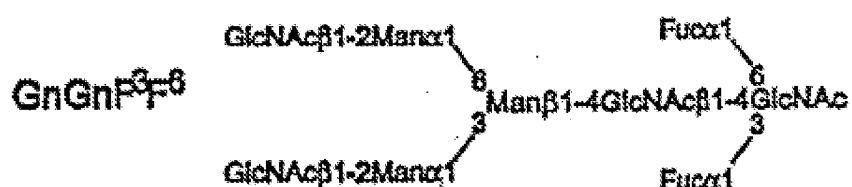
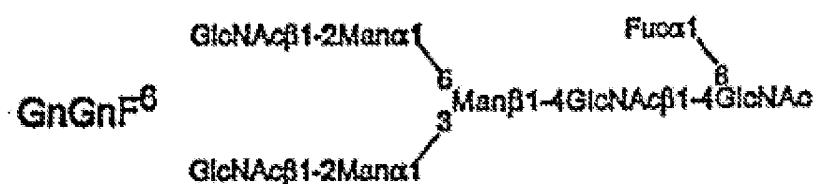
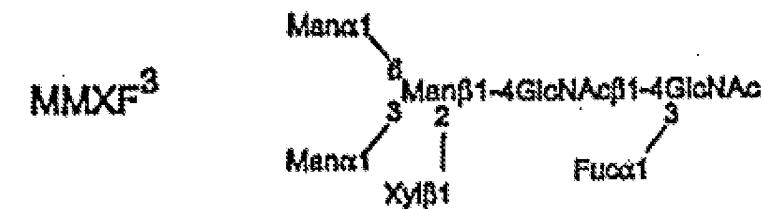


FIG.10 b

PE2062977

14/16

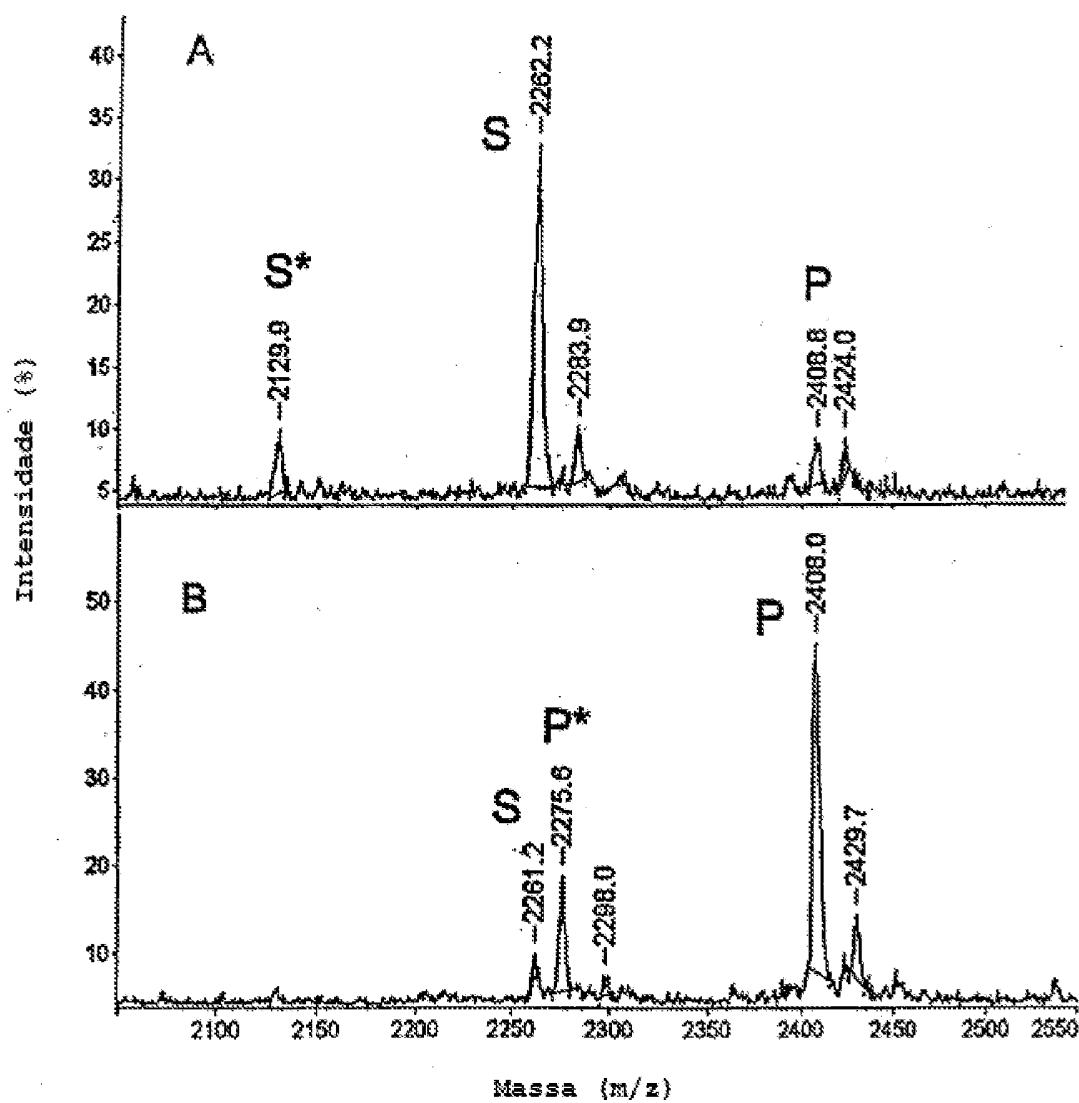
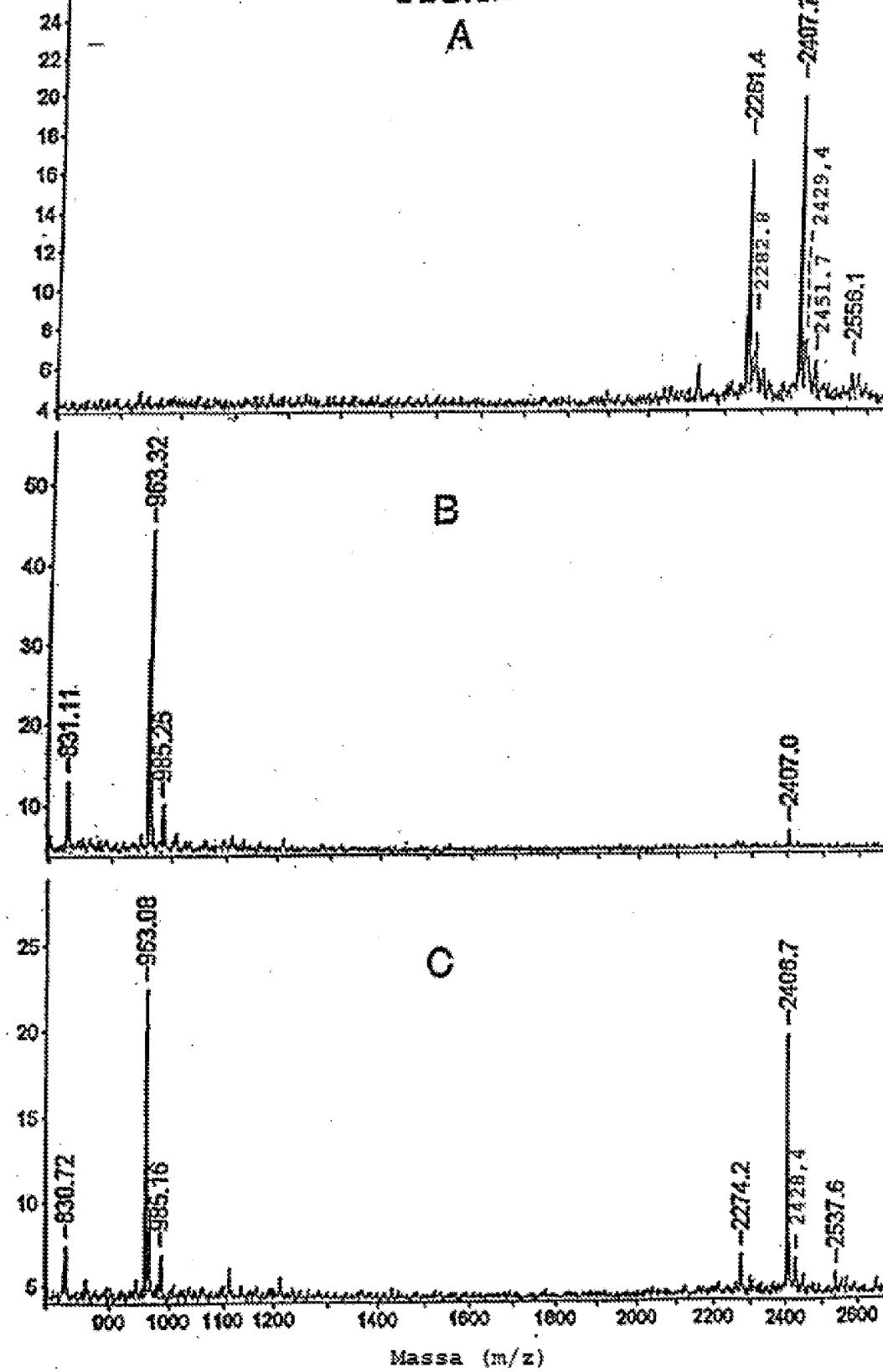


FIG.11

FIG.12



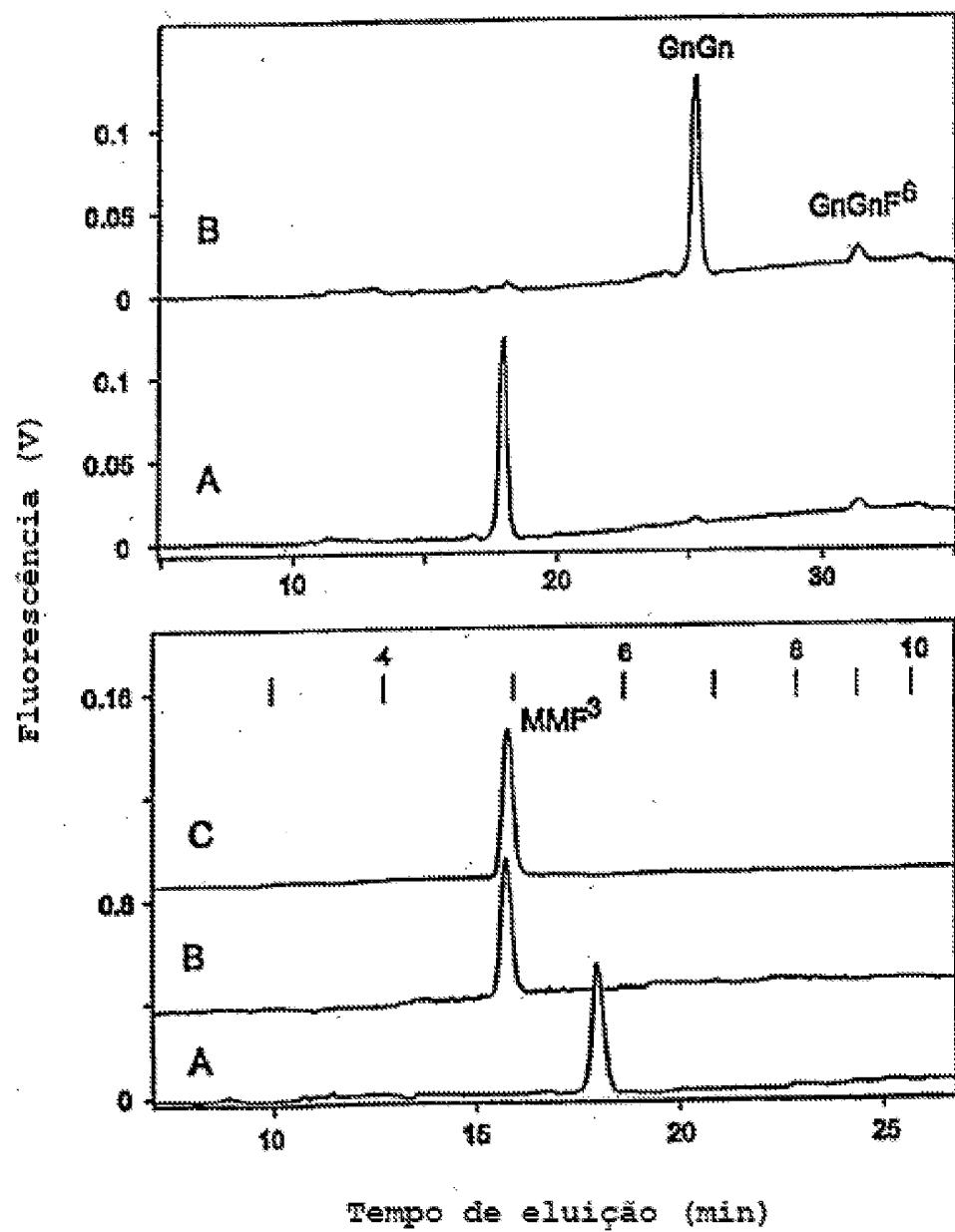


FIG. 13

## REFERÊNCIAS CITADAS NA DESCRIÇÃO

Esta lista de referências citadas pelo requerente é apenas para conveniência do leitor. A mesma não faz parte do documento da patente Europeia. Ainda que tenha sido tomado o devido cuidado ao compilar as referências, podem não estar excluídos erros ou omissões e o IEP declina quaisquer responsabilidades a esse respeito.

### Documentos de patentes citadas na Descrição

- US 5272966 A
- EP 0643132 A1
- EP 05108012 A

### Literatura que não é de patentes citada na Descrição

- P. LEROUGE et al. *Plant Mol. Biol.*, 1996, vol. 38, 31-48
- RAYON et al. *L. Exp. Bot.*, 1998, vol. 49, 1463-1472
- KUROSAKA et al. *J.Biol.Chem.*, 1991, vol. 266, 4168-4172
- MELO et al. *FEBS Lett.*, 1997, vol. 415, 186-191
- FITCHETTE-LAINE et al. *Plant J.*, 1997, vol. 12, 1411-1417
- I.B.H. WILSON et al. *Glycobiology*, 1998, vol. 8 (7), 651-661
- TRETTER et al. *Int. Arch. Allergy Immunol.*, 1993, vol. 102, 269-266
- PETERSEN et al. *J. Allergy Clin. Immunol.*, 1996, vol. 98 (4), 805-814
- E. ZABLACKIS et al. *Science*, 1996, vol. 272
- A. SCHAEWEN et al. *Plant Physiol.*, 1993, vol. 102, 1109-1118
- STAUDACHER et al. *Glycoconjugate J.*, 1995, vol. 12, 780-786
- JOHN M. BURKE. *Cleaning the way for ribozymes*. *Nature Biotechnology*, 1997, vol. 15, 414-415
- POOGA et al. *Cell-penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo*. *Nature Biotechnology*, 1998, vol. 16, 857-861
- COSTACHE et al. *J. Biol. Chem.*, 1997, vol. 272 (47), 29721-29728
- STAUDACHER et al. *Anal. Biochem.*, 1998, vol. 246, 98-101
- STAUDACHER et al. *Eur.J. Biochem.*, 1991, vol. 199, 745-751
- SCHRIFTEN BAULCOMBE. *Plant Mol. Biol.*, 1996, vol. 9, 373-382
- BRIGNET et al. *EMBO J.*, vol. 17, 6739-6746
- WATERHOUSE et al. *Proc.Natl.Acad.Sci.USA*, 1998, vol. 95, 13959-13964
- WATERHOUSE et al. *Proc.Natl.Acad.Sci.USA*, 1998, vol. 95, 13959-13964
- WAGNER et al. *Annu. Rev. Microbiol.*, 1994, vol. 48, 713-742
- RITTNER et al. *Nucl. Acids Res.*, 1993, vol. 21, 1361-1387
- PATZEL et al. *Nature Biotechnology*, 1998, vol. 16, 64-68
- J.BURKE. *Nature Biotechnology*, 1997, vol. 15, 414-415
- KUWABARA et al. *Nature Biotechnology*, 1998, vol. 16, 961-965
- SMITH et al. *Mol.Gen.Genet.*, 1990, vol. 224, 477-481
- SCHAEFER et al. *Plant J.*, 1997, vol. 11 (6), 1195-1208
- POOGA et al. *Nature Biotechnology*, 1998, vol. 16, 857-861
- STAUDACHER et al. *Glycoconjugate J.*, 1998, vol. 15, 355-360
- STAUDACHER et al. *Eur.J.Biochem.*, 1991, vol. 199, 745-751
- ALTMANN. *Anal.Biochem.*, 1992, vol. 204, 215-219
- WILSON et al. *Glycobiology*, 1998, vol. 8, 651-661
- KUBELKA et al. *Arch.Biochem.Biophys.*, 1994, vol. 308, 148-157
- HASE et al. *J.Biochem.*, 1984, vol. 95, 197-203