



República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(21) BR 112013013045-8 A2



(22) Data do Depósito: 24/11/2011

(43) Data da Publicação Nacional: 26/10/2021

(54) **Título:** PROTEÍNAS RECOMBINANTES COM ATIVIDADE N-ACETILGLUCOSAMINILTRANSFERASE, POLINUCLEOTÍDEO, VETOR DE EXPRESSÃO, CÉLULAS E MÉTODOS DE PRODUÇÃO DE N-GLICANO COMPLEXO

(51) **Int. Cl.:** C12N 9/10; C12P 21/00; C12N 5/10; C12N 1/14; C12N 15/80.

(30) **Prioridade Unionista:** 24/11/2010 US 61/417144.

(71) **Depositante(es):** NOVARTIS INTERNATIONAL PHARMACEUTICAL LTD.; GLYKOS FINLAND OY.

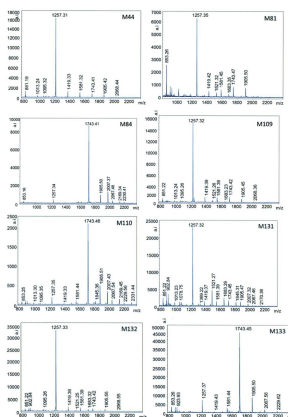
(72) **Inventor(es):** JARI NATUNEN; GLYKOS FINLAND OY; ANNE KANERVA; JUKKA HILTUNEN; MARKKU SALOHEIMO; HELI VISKARI; ANNE HUUSKONEN.

(86) **Pedido PCT:** PCT EP2011070956 de 24/11/2011

(87) **Publicação PCT:** WO 2012/069593 de 31/05/2012

(85) **Data da Fase Nacional:** 24/05/2013

(57) **Resumo:** PROTEÍNAS RECOMBINANTES COM ATIVIDADE N-ACETILGLUCOSAMINILTRANSFERASE, POLINUCLEOTÍDEO, VETOR DE EXPRESSÃO, CÉLULAS E MÉTODOS DE PRODUÇÃO DE N-GLICANO COMPLEXO. A presente descrição refere-se a proteínas recombinantes tendo atividade de N-acetilglucosaminiltransferase. A presente descrição também se refere a métodos para a produção de N-glicanos complexos, incluindo as etapas de fornecimento de células hospedeiras que contêm tais proteínas recombinantes e a cultura das células hospedeiras de modo que as proteínas recombinantes sejam expressas.



Relatório Descritivo da Patente de Invenção para: **"PROTEÍNAS RECOMBINANTES COM ATIVIDADE N-ACETILGLUCOSAMINILTRANSFERASE, POLINUCLEOTÍDEO, VETOR DE EXPRESSÃO, CÉLULAS E MÉTODOS DE PRODUÇÃO DE N-GLICANO COMPLEXO".**

5 REFERÊNCIA CRUZADA AO PEDIDO DE PATENTE RELACIONADO

Este pedido de patente reivindica o benefício do Pedido de Patente Provisório U.S. N ° 61/417.144, depositado em 24 de novembro de 2010, o qual é incorporado na presente invenção por meio da referência na sua totalidade.

10 APRESENTAÇÃO DE LISTAGEM DE SEQUÊNCIA NO ARQUIVO DE TEXTO ASCII

O conteúdo da seguinte apresentação em arquivo de texto ASCII é incorporado na presente invenção por meio da referência em sua totalidade: em suporte informático (CRF) da Listagem de Sequência (nome do arquivo : 619672001040SEQLIST.txt, data de gravação : 22 de novembro de 2011, o tamanho : 305 KB).

CAMPO DA INVENÇÃO

A presente invenção refere-se a composições e métodos úteis para a produção de N-glicanos.

20 ANTECEDENTES

A modificação pós-traducional das proteínas é muitas vezes necessária no que diz respeito ao dobramento apropriado da proteína e da função. Uma modificação de proteínas comum é a adição de oligossacarídeos (glicanos) a polipeptídeos nascentes no retículo endoplasmático com a finalidade de formar glicoproteínas, um processo conhecido por meio da glicosilação. A n-glicosilação é de particular importância no que diz respeito à produção de proteínas recombinantes utilizadas para fins terapêuticos. Uma vez que os sistemas de expressão procarióticos padrão faltam a maquinaria adequada necessária para tais modificações, as alternativas de sistemas de expressão têm de ser usadas na produção dessas proteínas terapêuticas. A levedura e fungos são opções atrativas no que diz respeito à expressão de proteínas em que podem ser facilmente cultivadas em grande escala em

meios simples, o que permite baixos custos de produção. Além disso, existem ferramentas para manipular a constituição genética relativamente sim-

ples de células de levedura e fungos, bem como as células eucarióticas mais complexas, tais como células de mamífero ou de inseto (De Pourcq et al. *Appl Microbiol Biotechnol*, 87 (5) : 1617 a 31) .

As células fúngicas e as células de mamíferos dividem etapas
 5 comuns nos estágios iniciais de glicosilação que resultam na formação de manose (8) N-acetilglucosamina (2) (Man8GlcNAc2). No entanto, existem diferenças significativas nas fases posteriores do processo. Por exemplo, em levedura, as subunidades de manose adicionais são adicionadas a Man8GlcNAc2 por meio de manosiltransferases e manano polimerases com
 10 a finalidade de se obter N-glicanos de tipo manose elevada. Em contraste, os açúcares de manose são removidos do Man8GlcNAc2 para originar Man5GlcNAc2 humano, seguidos de três reações sequenciais envolvendo as enzimas N-acetilglucosaminiltransferase I (GnTI), manosidase II (II, Mns), e N-acetilglucosaminiltransferase II (GnTII), com a finalidade de converter
 15 Man5GlcNAc2 em GlcNAc2Man3GlcNAc2.

As diferenças entre o processo de glicosilação em células de mamífero e fúngicas representam um desafio no que diz respeito à expressão de proteínas de mamífero glicosiladas em células fúngicas uma vez que as glicoproteínas com N-glicanos de tipo manose elevada não são adequa-
 20 das para utilização terapêutica em seres humanos (De Pourcq et al ., 2010 ; Wildt e Gerngross, *Nature Reviews Microbiology*, 3 : 119 a 128). Consequentemente, foram realizados estudos de re-engenharia das vias de glicosilação em leveduras e espécies de fungos que lhes permitam expressar as proteínas recombinantes humanas. A abordagem geral na glicoengenharia de le-
 25 vedura ou células fúngicas tem sido a de interromper os genes endógenos que estão envolvidos na formação de N-glicanos de tipo manose elevada. Estas perturbações do gene podem ser combinadas com a sobre-expressão de manosidases endógenas /e/ou de glicosiltransferases e glicosidases de diferentes espécies (Chiba et al, 1998, *J Biol Chem* 273 : 26298-304 ; Kainz et al, 2008, *Appl Environ Microbiol* 74 : 1076 a 1086 ; Maras et al, 1997, *Euro J Biochem* 249 : 701 a 07 ; Maras et al, 1999, *FEBS Letters* 452 : 365 a 70 ;
 30 Hamilton et al, 2003, *Science* 301 : 1244 a 6 ; De Pourcq et al., 2010). No

entanto, a produção de proteínas de mamífero glicosiladas nas células de mamífero ainda não requerem a engenharia genética complicada e demorada e pode ser ineficaz em produzir uma glicoproteína desejada.

5 Dessa maneira, permanece uma necessidade na técnica de um sistema mais simples e mais eficiente com a finalidade de expressar N-glicanos complexos em células não mamíferas.

SUMÁRIO

10 As composições são descritos na presente invenção, incluindo as proteínas recombinantes que possuem atividade de N-acetilglucosaminiltransferase. Além disso, são descritos na presente invenção os métodos de produção de complexos de N-glicanos e métodos de produção de Man3GlcNAc2 glicanos.

Dessa maneira, um aspecto inclui as proteínas recombinantes que possuem a atividade de N-acetilglucosaminiltransferase, em que as proteínas recombinantes catalisam a transferência de N-acetilglucosamina com um resíduo terminal e Man α 3 catalisam a transferência de N-acetilglucosamina com um terminal do resíduo Man α 6 de um glicano aceitador, e onde a proteína recombinante contém domínios catalíticos de pelo menos duas enzimas diferentes. Em certas modalidades, o glicano aceitador é ligado a uma molécula selecionada a partir de um aminoácido, um peptídeo, ou um polipeptídeo. Em certas modalidades, a molécula é um polipeptídeo heterólogo. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o glicano aceitador é Man3. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a proteína recombinante é uma proteína de fusão que contém um domínio catalítico de N-acetilglucosaminiltransferase I e II, um domínio catalítico de N-acetilglucosaminiltransferase. Em certas modalidades, um domínio catalítico de N-acetilglucosaminiltransferase I e Domínio catalítico de N-acetilglucosaminiltransferase II são as enzimas humanas. Em certas modalidades, o domínio catalítico N -acetilglucosaminiltransferase I inclui uma sequência que é de pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 % , pelo menos 97 %, pelo menos 98 %, pelo

menos 99 %, ou 100 % idêntica aos resíduos de aminoácidos 105 a 445 da SEQ ID NO : 1. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o domínio catalítico N -acetilglucosaminiltransferase II inclui uma sequência que é pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 % ou 100 % idêntica resíduos de aminoácidos 30 a 447 da SEQ ID NO : 21. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o domínio catalítico de N-acetilglucosaminiltransferase I é um Terminal N para o domínio catalítico de N-acetilglucosaminiltransferase II. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o domínio catalítico de N-acetilglucosaminiltransferase II é um Terminal N para o domínio catalítico de N-acetilglucosaminiltransferase I.

Em certas modalidades que podem ser combinadas com as modalidades anteriores, as proteínas recombinantes ainda contêm um espaçador entre o domínio catalítico de N-acetilglucosaminiltransferase I e o domínio catalítico de N-acetilglucosaminiltransferase II . Em certas modalidades, o espaçador contém uma sequência de domínio haste. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o espaçador é de pelo menos 5, pelo menos 10, pelo menos 15, pelo menos 20, pelo menos 30, pelo menos 40, ou pelo menos 50 aminoácidos de comprimento. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o espaçador contém uma sequência que é selecionada a partir da SEQ ID NO : 118, SEQ ID NO : 120, SEQ ID NO : 122 e SEQ ID NO : 124. Em certas modalidades, o espaçador contém uma sequência que é selecionada a partir da SEQ ID NO : 118, SEQ ID NO : 120 e SEQ ID NO : 124. Em certas modalidades, o espaçador contém a sequência de SEQ ID NO : 120 ou SEQ ID NO : 124. Em certas modalidades, o espaçador contém a sequência de SEQ ID NO : 124.

Em certas modalidades que podem ser combinadas com as modalidades anteriores, as proteínas recombinantes ainda contêm um peptídeo alvo ligado à extremidade Terminal N de um dos domínios catalíticos. Em

certas modalidades, o peptídeo alvo contém um domínio da haste. Em certas modalidades, o domínio de haste é de uma enzima N-acetilglucosaminiltransferase I ou uma enzima N-acetilglucosaminiltransferase II. Em certas modalidades, a enzima N-acetilglucosaminiltransferase I e a enzima N-acetilglucosaminiltransferase II são enzimas humanas. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o domínio da haste é de uma proteína selecionada a partir de uma manosidase, uma manosiltransferase, uma glicosiltransferase, um tipo de proteína 2 de Golgi, MNN2, MNN4, MNN6, MNN9, MNN10, MNS1, KRE2, VAN1, ou OCH1. Em certas modalidades, a proteína é a partir de um organismo selecionado a partir de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filibasidium*, *Fusarium*, *Gibberella*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, ou *Trichoderma*. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o peptídeo alvo é um peptídeo Kre2 alvo. Em certas modalidades, o peptídeo alvo contém um domínio transmembranar. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o peptídeo alvo contém ainda um domínio transmembranar ligado à extremidade do Terminal N do domínio da haste. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o domínio transmembranar é de uma enzima N-acetilglucosaminiltransferase I ou uma enzima N-acetilglucosaminiltransferase II. Em certas modalidades, a enzima N-acetilglucosaminiltransferase I e a enzima N-acetilglucosaminiltransferase II são enzimas humanas. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o domínio transmembranar é de uma proteína selecionada a partir de uma manosidase, uma manosiltransferase, uma glicosiltransferase, um tipo de proteína 2 de Golgi, MNN2, MNN4, MNN6, MNN9, MNN10, MNS1, KRE2, VAN1, ou OCH1. Em certas modalidades, a proteína é a partir de um organismo selecionado a partir de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filibasidium*,

Fusarium, Gibberella, Humicola, Magnaporthe, Mucor, Myceliophthora, Myrothecium, Neocallimastix, Neurospora, Paecilamyces, Penicillium, Piromyces, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, ou *Trichoderma*. Em certas modalidades, o peptídeo alvo contém um domínio

5 citoplasmático. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o peptídeo alvo contém ainda um domínio citoplasmático ligado à extremidade Terminal N do domínio da haste. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o peptídeo alvo contém ainda um domínio citoplasmático ligado à extremidade

10 Terminal N do domínio transmembranar. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o domínio citoplasmático é de uma enzima N-acetilglucosaminiltransferase I ou uma enzima N-acetilglucosaminiltransferase II. Em certas modalidades, a enzima N-acetilglucosaminiltransferase I e a enzima N-acetilglucosaminiltransferase II são enzimas

15 humanas. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o domínio citoplasmático é de uma proteína selecionada a partir de uma manosidase, uma manosiltransferase, uma glicosiltransferase, um tipo de proteína 2 de Golgi, MNN2, MNN4, MNN6, MNN9, MNN10, MNS1, KRE2, VAN1, ou OCH1. Em certas modalidades, a proteína é a partir

20 de um organismo selecionado a partir de *Acremonium, Aspergillus, Aureobasidium, Cryptococcus, Chrysosporium, Chrysosporium lucknowense, Filibasidium, Fusarium, Gibberella, Humicola, Magnaporthe, Mucor, Myceliophthora, Myrothecium, Neocallimastix, Neurospora, Paecilamyces, Penicillium, Piromyces, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium*, ou *Trichoderma*.

25

Um outro aspecto inclui as proteínas recombinantes contendo um domínio catalítico de N-acetilglucosaminiltransferase II humano e um domínio catalítico de N-acetilglucosaminiltransferase I humano onde o domínio catalítico de N-acetilglucosaminiltransferase II está localizado no terminal

30 N para o domínio catalítico de N-acetilglucosaminiltransferase I, um espaçador sequência contendo a sequência a partir de uma haste do Domínio catalítico de N-acetilglucosaminiltransferase I humano localizada entre os domí-

nios catalíticos, e de um peptídeo alvo localizado no terminal N ao domínio catalítico de N-acetilglucosaminiltransferase II na qual o peptídeo alvo contém um domínio citoplasmático, um domínio transmembranar, e uma haste do domínio de N-acetilglucosaminiltransferase II humana. Outro aspecto inclui uma proteína recombinante que contém uma sequência que é pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica à SEQ ID NO : 95.

Um outro aspecto inclui proteínas recombinantes contendo Domínio catalítico de N-acetilglucosaminiltransferase II e um Domínio catalítico de N-acetilglucosaminiltransferase I, em que o domínio catalítico de N-acetilglucosaminiltransferase II está localizado no terminal N para o Domínio catalítico de N-acetilglucosaminiltransferase I, um espaçador situado na entre os domínios catalíticos, onde o espaçador contém uma sequência selecionada a partir da SEQ ID NO : 118, SEQ ID NO : 120, SEQ ID NO : 122 e SEQ ID NO : 124, e um peptídeo alvo Terminal N localizado ao N- acetilglucosaminiltransferase domínio catalítico II na qual o peptídeo alvo contém um domínio citoplasmático, um domínio transmembranar, e um domínio de haste de N-acetilglucosaminiltransferase II humana. Em certas modalidades, o espaçador contém uma sequência que é selecionada a partir da SEQ ID NO : 118, SEQ ID NO : 120 e SEQ ID NO : 124. Em certas modalidades, o espaçador contém a sequência de SEQ ID NO : 120 ou SEQ ID NO : 124. Em certas modalidades, o espaçador contém a sequência de SEQ ID NO : 124.

Um outro aspecto inclui polinucleotídeos isolados que codificam a proteína recombinante de qualquer uma das modalidades anteriores. Outro aspecto inclui os vetores de expressão que contêm o polinucleotídeo isolado da modalidade anterior, operativamente ligada a um promotor. Em certas modalidades, o promotor é um promotor constitutivo. Em certas modalidades, o promotor é um promotor indutível. Em certas modalidades, o promotor é de um gene selecionado a partir de *gpdA*, *cbh1*, *TAKA* amilase de *Aspergillus oryzae*, *prbateladainase* aspártica de *Rhizomucor miehei*, *Aspergillus niger* neutro alfa-amilase, *Aspergillus niger* de ácido estável de alfa-amilase,

glucoamilase de *Aspergillus niger* (glaA), *Aspergillus awamori* glaA, lipase de *Rhizomucor miehei*, protease alcalina de *Aspergillus oryzae*, *Aspergillus oryzae*, triose fosfato isomerase, acetamidase de *Aspergillus nidulans*, acetamidase de *Aspergillus oryzae*, *Fusarium oxysporum*, protease do tipo tripsina, 5 fúngica endo α -L-arabinase (ABNA), α -L-arabinofuranosidase A fúngica (ABFA), α -L-arabinofuranosidase B fúngica (abfB), xilanase por fungos (xlnA), fitase fúngica, fungos ATP-sintetase, subunidade de fungos 9 (PBR), triose fosfato isomerase de fungos (IPT), álcool desidrogenase de fungos (Adha), α -amilase fúngica (amy), amilaglicosidase fúngica (glaA), acetamidase fúngica (amdS), desidrogenase de fungos gliceraldeído-3-fosfato 10 (GPD), álcool-desidrogenase de levedura, levedura de álcool oxidase, lactase de levedura, levedura de 3-fosfoglicerato-quinase, levedura triosefosfatossomerase, SpO2 α -amilase bacteriana bacteriana, ou SSO. Outro aspecto inclui as células hospedeiras que contêm a expressão do vetor de qualquer 15 uma das modalidades anteriores.

Um outro aspecto inclui os métodos capazes de produzir a proteína recombinante de qualquer uma das modalidades precedentes, incluindo as etapas de introdução de um polinucleotídeo isolado que codifica a proteína recombinanteem uma célula hospedeira e cultivar a célula hospedeira de 20 tal modo que a proteína recombinante é expressa. Em certas modalidades, os métodos incluem ainda uma etapa de purificação da proteína recombinante a partir da célula hospedeira. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira é uma célula fúngica. Em certas modalidades, a célula fúngica é selecionada a partir 25 de leveduras ou fungos filamentosos.

Um outro aspecto inclui os métodos de produção de um complexo de N-glicanos, incluindo as etapas de proporcionar uma célula hospedeira, onde a célula hospedeira contém uma proteína de fusão contendo um domínio catalítico de N-acetilglucosaminiltransferase I e II, um domínio cata- 30 lítico de N-acetilglucosaminiltransferase do polinucleotídeo que codifica e cultura da célula hospedeira de tal modo que a proteína de fusão é expressa, a proteína de fusão em que catalisa a transferência de N-acetilglucosa-

- mina com um resíduo $\text{Man}\alpha 3$ terminal e N-acetilglucosamina a um resíduo $\text{Man}\alpha 6$ terminal de um glicano aceitador com a finalidade de produzir um complexo de N-glicanos. Em certas modalidades, o complexo de N-glicanos é ligado a uma molécula selecionada a partir de um aminoácido, um peptídeo, ou um polipeptídeo. Em certas modalidades, a molécula é um polipeptídeo heterólogo. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o glicano aceitador é $\text{Man}3$. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o complexo de N-glicanos é $\text{GlcNAc}\beta 2\text{Man}\alpha 3$ ($\text{GlcNAc}\beta 2\text{Man}\alpha 6$) $\text{Man}\beta 4\text{GlcNAc}\beta 4\text{GlcNAc}$.
- 5 Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira é uma célula eucariótica. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira é uma célula fúngica. Em certas modalidades, a célula é uma célula fúngica levedura selecionada a partir de *S. cerevisiae*, *K. lactis*, *P. pastoris*, *H. polymorpha*, *C. albicans*, *Schizosaccharomyces*, ou *Yarrowia*. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula é uma célula fúngica filamentosa selecionada a partir de *Trichoderma sp.*, *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filibasidium*, *Fusarium*, *Gibberella*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, ou *Tolypocladium*. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira contém ainda um polinucleotídeo que codifica um veículo de
- 10 UDP-GlcNAc. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira tem um nível reduzido de atividade de um dolichil-P-Man : Man (5) GlcNAc (2)-PP-dolichyl manosiltransferase em relação ao nível de atividade em uma célula hospedeira de tipo selvagem. Em certas modalidades, a célula hospedeira tem um nível reduzido
- 15 de expressão de um gene *alg3* em comparação com o nível de expressão em uma célula hospedeira de tipo selvagem. Em certas modalidades, o gene *alg3* é eliminado da célula hospedeira. Em certas modalidades que podem
- 20
- 25
- 30

ser combinadas com as modalidades anteriores, a célula hospedeira tem um nível reduzido de atividade de um α -1, 6-manosiltransferase em relação ao nível da atividade de uma célula hospedeira de tipo selvagem. Em certas modalidades, a célula hospedeira tem um nível reduzido de expressão de um gene OCH1 em comparação com o nível de expressão em uma célula hospedeira de tipo selvagem. Em certas modalidades, o gene OCH1 é eliminado da célula hospedeira. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira contém ainda um polinucleotídeo que codifica uma α -1,2-manosidase. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira contém ainda um polinucleotídeo que codifica uma β -1,4-galactosiltransferase. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira contém ainda um polinucleotídeo que codifica uma sialiltransferase. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira é uma célula de Trichoderma que tenha um nível reduzido de atividade de um dolichil-P-Man : Man (5) GlcNAc (2)-PP- manosiltransferase em relação ao nível de atividade em uma célula Trichoderma do tipo selvagem. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira é uma célula de levedura ou fúngica que tem um nível reduzido de atividade de um dolichil-P-Man: Man (5) GlcNAc (2)-PP-manosiltransferase e dolichil em um nível reduzido de atividade de uma alfa-1,6-manosiltransferase em relação aos níveis de atividade de uma célula de levedura do tipo selvagem e outra contém um polinucleotídeo que codifica uma α -1,2-manosidase.

Um outro aspecto inclui os métodos de produção de um complexo de N-glicanos, incluindo as etapas de proporcionar uma célula hospedeira de Trichoderma, quando a célula hospedeira tem um nível reduzido de expressão de um gene alg3 em comparação com o nível de expressão em um hospedeiro de tipo selvagem célula e que contém um primeiro polinucleotídeo que codifica um Domínio catalítico de N-acetilglucosaminiltransferase I e um segundo polinucleotídeo que codifica um domínio catalítico II N-acetilglu-

cosaminiltransferase, e a cultura da célula hospedeira com a finalidade de produzir um complexo de N-glicanos.

Um outro aspecto inclui os métodos de produção de um complexo de N-glicanos, incluindo as etapas de incubação de uma proteína de fusão que contém um Domínio catalítico de N-acetilglucosaminiltransferase I e um domínio catalítico II N-acetilglucosaminiltransferase, um glicano aceitador, e um doador de N-acetilglucosamina em conjunto em um tampão, onde a proteína de fusão catalisa a transferência de N-acetilglucosamina com um Resíduo Man α 3 terminal e N-acetilglucosamina a um resíduo Man α 6 terminal de um glicano aceitador com a finalidade de produzir um complexo de N-glicanos. Em certas modalidades, o glicano aceitador é ligado a uma molécula selecionada a partir de um aminoácido, um peptídeo, ou um polipeptídeo. Em certas modalidades, a molécula é um polipeptídeo heterólogo. Em certas modalidades, o glicano aceitador é Man3. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o doador N-acetilglucosamina é um veículo de UDP-GlcNAc.

Um outro aspecto inclui células de fungos filamentosos que contém uma mutação de alg3 e Man3GlcNAc2, onde Man3GlcNAc2 inclui pelo menos 50 %, pelo menos 60 %, pelo menos 70 %, pelo menos 80 %, pelo menos 90 %, ou 100 % (% em mol) de N-glicanos neutros segregados por meio das células. Os N-glicanos neutros podem ser ligados a uma molécula selecionada a partir do grupo que consiste em um aminoácido, um peptídeo, e um polipeptídeo. Em certas modalidades, a mutação de alg3 é uma deleção de alg3. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula é uma célula de *Trichoderma reesei*. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula fúngica filamentosa contém ainda um primeiro polinucleotídeo que codifica um Domínio catalítico de N-acetilglucosaminiltransferase I e um segundo polinucleotídeo que codifica um domínio catalítico II N-acetilglucosaminiltransferase. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula fúngica filamentosa contém ainda um polinucleotídeo que codifica uma proteína de fusão que contém um Domínio

catalítico de N-acetilglucosaminiltransferase I e um domínio catalítico II N-acetilglucosaminiltransferase.

Um outro aspecto inclui os métodos de produzir um glicano Man3GlcNAc2 em uma célula hospedeira, incluindo as etapas de proporcionar uma célula hospedeira com um nível reduzido de atividade de manosiltransferase em relação ao nível da atividade de uma célula hospedeira de tipo selvagem e a cultura do hospedeiro célula com a finalidade de produzir um glicano Man3GlcNAc2, onde o glicano Man3GlcNAc2 inclui pelo menos 50 %, pelo menos 60 %, pelo menos 70 %, pelo menos 80 %, pelo menos 90 %, or 100 % (% em mol) de N-glicanos neutros segregados por meio da célula hospedeira. Os neutros N-glicanos podem ser ligados a uma molécula selecionada a partir de um aminoácido, um peptídeo, e um polipeptídeo. Em certas modalidades, o glicano Man3GlcNAc2 está ligado a um polipeptídeo heterólogo. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a manosiltransferase é um dolichil-P-Man : Man (5) GlcNAc (2)-PP-dolichil manosiltransferase. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira tem um nível reduzido de expressão de um gene *alg3* em comparação com o nível de expressão em uma célula hospedeira de tipo selvagem. Em certas modalidades, o gene *alg3* é eliminado da célula hospedeira. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira é uma célula de *Trichoderma*. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o nível de atividade de alfa-1, 6-manosiltransferase na célula hospedeira é reduzido em comparação com o nível de atividade em uma célula hospedeira de tipo selvagem. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula hospedeira contém um polinucleotídeo que codifica um α -1,2-manosidase.

Um outro aspecto inclui uma célula fúngica filamentosa ter uma redução do nível de expressão de um gene *alg3* em comparação com o nível de expressão em uma célula fúngicas filamentosos do tipo selvagem, em que a célula fúngica filamentosa contém uma proteína recombinante de

qualquer das modalidades precedentes . Em certas modalidades, o gene *alg3* contém uma mutação. Preferencialmente, a proteína recombinante possui uma atividade N-acetilglucosaminiltransferase, onde a proteína recombinante catalisa a transferência de N-acetilglucosamina com um Resíduo

5 *Man* α 3 terminal e catalisa a transferência de N-acetilglucosamina a um resíduo *Man* α 6 terminal de um glicano aceitador, e em que a proteína recombinante é uma proteína de fusão contendo um domínio catalítico de N-acetilglucosaminiltransferase I e II, um domínio catalítico de N-acetilglucosaminiltransferase. Em certas modalidades, a mutação do gene *alg3* é uma

10 deleção do gene *alg3*. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a proteína de fusão é codificada por um polinucleotídeo operativamente ligado a um promotor. Em certas modalidades, o promotor é um promotor indutível. Em certas modalidades, o promotor induzível é o promotor *cbh1*. Em certas modalidades que podem ser combinadas

15 com as modalidades anteriores, a célula fúngica filamentosa contém ainda um polinucleotídeo que codifica um veículo de UDP-GlcNAc. Em certas modalidades que podem ser combinadas com as modalidades anteriores, o fungo filamentoso tem um nível reduzido de atividade de um α -1, 6-manosiltransferase em relação ao nível da atividade de uma célula fúngica filamentosa de tipo selvagem. Em certas modalidades, o fungo filamentoso tem

20 um nível reduzido de expressão de um gene *OCH1* em comparação com o nível de expressão em uma célula fúngica filamentosa do tipo selvagem. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula fúngica filamentosa contém ainda um polinucleotídeo que

25 codifica uma α -1,2-manosidase. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula fúngica filamentosa contém ainda um polinucleotídeo que codifica uma β -1,4-galactosiltransferase. Em certas modalidades que podem ser combinadas com as modalidades anteriores, a célula fúngica filamentosa contém ainda um polinucleotídeo que codifica uma sialiltransferase. Em certas modalidades que podem

30 ser combinadas com as modalidades anteriores, a célula filamentosa fúngica é selecionada a partir de *Trichoderma sp.*, *Acremonium*, *Aspergillus*, *Aureo-*

basidium, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filibasidium*, *Fusarium*, *Gibberella*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paecilamyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, e *Tolypocladium*.

5 DESCRIÇÃO DAS FIGURAS

A Figura 1 mostra os perfis de espectrometria de massa N-glicanos neutros de glicosilação média de cepas de *T. reesei* M44, M81, M84, M109, M110, M131, M132, M133, M134 e M124.

10 A Figura 2 mostra a análise da fragmentação de Man7Gn2 monofosforilado. Apenas um exemplo de estrutura de Man7Gn2 monofosforilado é mostrado.

A Figura 3 mostra perfis de glicanos ácidos de espectrometria de massa de cepas *T. reesei* M44, M81, M84, M109, M110, M131, M132, M133, M134 e M124.

15 A Figura 4 mostra Os perfis de N-glicano neutros (a) e acídicos (b) da cepa de *T. reesei* M44 cultivadas em um fermentador durante 131,4 horas (batelada alimentada).

A Figura 5 mostra Os perfis de N-glicano neutros (a) e acídicos (b) de N-glicano em espectrometria de massa do meio de cultura de *T. reesei*.

A Figura 6 mostra uma mancha da membrana das proteínas secretadas *T. reesei* M44.

25 A Figura 7 mostra um exemplo das bandas de proteína *T. reesei* M44 analisadas cultivadas em um fermentador. A glicosilação de proteínas não diferiu de uma forma significativa no que diz respeito à glicosilação média em *T. reesei*. O espectro foi focado para o sinal da linha de base menor, e o sinal principal do espectro não foi quantitativo, em comparação com outros sinais.

30 A Figura 8 mostra uma mancha de Southern de DNA a partir da cepa progenitora e das cepas Alg3 knockout com uma sonda alg3.

A Figura 9A mostra um mapa de enzimas de restrição de uma seção do construto pTTv38 com tamanhos de produtos de restrição previs-

tos. A Figura 9B mostra uma mancha de Southern de DNA genômico a partir da cepa parental e as cepas Alg3 knockout digeridas com EcoRI + PvuI (E + P) ou KpnI NheI + (K + N). O DNA de controle foi pTTv38 DNA de plasmídeo digerido com NotI. A mancha foi sondada com uma sonda AMDS.

5 A Figura 10 mostra a análise MALDI de N-glicanos neutros. A parte A mostra a cepa parental M124. A Parte B mostra os Alg3 4A knockout. Os quadrados representam N-acetilglicosamina e os círculos representam manose, exceto para a glicose marcada.

10 A Figura 11 mostra a análise de fragmentação de Man3Gn2 da cepa Alg3 4A knockout.

A Figura 12 mostra a análise de fragmentação de Hex5Gn2 da cepa Alg3 4A knockout (parte A) e cepa de M124 parente (parte B). O sinal marcado com uma caixa existe somente como um isômero da cepa Alg3 knockout

15 A Figura 13 mostra N-glicanos neutros da cepa Alg3 4A knockout após digestão α -mannosidase.

A Figura 14 mostra a separação dos dois principais glicanos a partir da cepa Alg3 knockout por meio da cromatografia líquida.

20 A Figura 15 mostra os espectros de RMN de prótons das frações Hex3HexNAc2 (parte A) e Hex6HexNAc2 (parte B). Os espectros foram recolhidos a 40°C usando um Varian Unity Inova de 600 MHz equipado com um espectrômetro de criosonda.

25 A Figura 16 mostra a fração ácida da cepa parental M124 (parte A) e da cepa Alg3 4A knockout (B). N-glicanos com duas unidades de fosfato são marcados com um asterisco.

A Figura 17 apresenta N-glicanos neutros a partir do sobrenadante da cepa *T. reesei* Alg3 4A knockout que foi cultivada em um frasco durante 5 dias.

30 A Figura 18 apresenta N-glicanos neutros a partir do sobrenadante da cepa *T. reesei* Alg3 4A knockout que foi cultivada em um fermentador de 10 dias.

A Figura 19 mostra um espectro de MALDI de mistura de reação

GnTI. GnTI converteu 54 % do aceitador para o produto com um HexNAc adicional.

5 A Figura 20 mostra a análise de Western blot da expressão GnTII. As amostras foram executadas em 12 % de SDS-PAGE gel e apagou em membrana de nitrocelulose. Histidina marcada com GnTII foi detectada na membrana utilizando os anticorpos monoclonais α -HIS. Os números mostrados no lado esquerdo são os tamanhos das proteínas marcadoras de peso molecular (kDa).

10 A Figura 21 mostra um espectro de MALDI de mistura de reação GnTII. 83 % do aceitador (m / z 913,340) que foi convertido no produto (m / z 1.136,433).

A Figura 22 mostra a atividade de GnTI observada para a proteína de fusão GnTI / GnTII.

15 A Figura 23 mostra os N-glicanos presentes em GnTI / GnTII *T. reesei* transformantes obtidos por meio do direcionamento para o local de alg3.

A Figura 24 mostra um espectro de MALDI da mistura de reação purificada a partir do ensaio da proteína de fusão GnTII / GnTI de atividade da enzima.

20 A Figura 25 mostra um espectro da mistura de reação β 1-2,3,4,6-N-acetilglucosaminidase.

A Figura 26 mostra um espectro da mistura de reação de MALDI- β 1 4GalT.

25 A Figura 27 mostra diagramas de N-glicanos observados a partir de proteínas do sobrenadante de *T. reesei* M127 pTTv110 transformantes (gnt II / I, em locais alg3) nos dias 3 (A), 5 (B) e 7 (C e D). O clone 17A produziu G0 adicional no dia 7. (E) O espectro de massa de N-glicanos de proteínas a partir do sobrenadante da cepa *T. reesei* M127 GnT II / I transformantes 17A dos clones cultivados durante 7 dias em balões de agitação
30 neutro. Sinais marcados com asteriscos originados a partir do meio de cultura.

A Figura 28 mostra N-glicanos neutros de rituximab a partir de *T.*

reesei M202 GnT II / I clones transformantes (A) 9A-1 e (B) 31A-1, tanto cultivados com inibidor de tripsina de soja, e (C) O espectro de massa de N-glicanos neutro do rituximab purificada a partir da cepa T. reesei M202 GnT II / I transformante clone 9A-1 cultivadas durante 5 dias em balões de agitação, na presença de inibidor de tripsina de soja.

Figura 29 mostra espectros de MALDI espaçador modificado misturas de reação de fusão GnTII / GnTI. A parte (A) mostra uma mistura de reação de GnTII / GnTI com 3xG4S modificação espaçador. 36 % do aceitador foi convertido para o produto com duas HexNAcs adicionais. A parte (B) apresenta uma mistura de reação de GnTII / GnTI com 2xG4S modificação espaçador. 38 % do aceitador foi convertido para o produto com duas HexNAcs adicionais. Calculado m / z valores para [M + Na] +-sinais de GnTI produto, Hex3HexNAc2 (cale. m / z 933,318), não foi detectado em qualquer dos espectros porque todo o produto GnTI foi convertido diretamente Hex3HexNAc3, (cale. m / z 1.136,318).

A Figura 30 apresenta Western blots de péletes celulares variantes do espaçador GnTII / I (A), e os sobrenadantes (B). Linhas 1.GnTII controle positivo, cepa simulada 2 GY3, 3.GY7-2 do tipo selvagem GnTII / I, 4.GY32-5 espaçador 3xG4S, 5. GY32-9 espaçador 3xG4S, 6.GY33-7 espaçador 2xG4S, 7.GY33-8 espaçador 2xG4S, 8.GY49-3 espaçador cbhl e 9.GY50-10 espaçador EGIV.

A Figura 31 mostra as atividades de GnT GnTII / I de tipo selvagem e espaçador de variantes a partir de sobrenadantes e expressas na presença de inibidores de protease após o dia 3 (a) nas fases de expressão e o dia 4 (B) nas fases de expressão. O eixo x representa a identidade da amostra (= em peso de tipo selvagem, _1, _2 = clones paralelas do espaçador da variante), e o eixo y descreve a percentagem de produtos formados (produtos de reação GnTI e GnTII somados).

A Figura 32 mostra as atividades da proteína de fusão GNT GnTII / I (com o tipo selvagem espaçador) em sobrenadante, células e lisado. Os produtos GnTI e GnTII foram adicionados juntos

A Figura 33 mostra as atividades de GnT GnTII / I de tipo selva-

gem e o espaçador da variante em sobrenadante (A), (B) células, e (c) os lisados.

A Figura 34 mostra exemplo espectros de N-glicanos da cepa parental M124 e GnT1 transformantes neutras no dia 5. Os sinais com adição Gn (m / z 1460) estão marcados com uma seta. (PTTv11 com promotor cbh1, pTTv13 com promotor gpdA).

A Figura 35 mostra as quantidades de man5 e Gn1Man5 em quatro transformantes GNT1 positivos nos dias 3 e 5. A quantificação foi realizada contra calibrador interno (Hex2HexNAc4, 2 pmol).

A Figura 36 mostra exemplo espectros de N-glicanos da cepa parental M124 e GnT1 transformantes com calibrador interno (NeuAcHex4HexNAc2, 0,5 pmol). Fosforilada. Os produtos GnT1 são marcados com uma seta.

A Figura 37 mostra os diagramas de N-glicanos neutros de diferentes cepas GnTII / clones a partir do dia 5. Parte (A) mostram o clone pTTv140. A parte (B) mostra o clone pTTv142. A parte (C) mostra o clone pTTv143. A parte (D) mostra o clone pTTv141.

A Figura 38 mostra um exemplo de N-glicanos neutros de diferentes cepas de GnTII / clones e a cepa parental M198 dos dias 3, 5 e 7. A parte (A) mostra o clone 1-117A. A parte (B) mostra o clone 3-11A. A parte (C) mostra o clone 30A. A parte (D) mostra a mancha parental M198.

A Figura 39 mostra a membrana das proteínas separadas da cepa de *T. reesei* M198 e GnTII do clone 3-17A. A proteína de 50 kDa é marcada com uma seta.

A Figura 40 mostra que os diagramas de coluna de proteínas totais segregados contra a proteína (s) secretada (s) individual (s) da cepa parental M198 (A) e do clone 3-17A GnTII (B).

A Figura 41 mostra um diagrama de coluna de fermentador da cepa GnTII M329 do dia 3 ao dia 7, e agitar a cultura do frasco de cepa M329 do dia 5.

A Figura 42 mostra um alinhamento da sequência de aminoácido múltiplo de *T. reesei* ALG3 e ALG3 homólogos.

DESCRIÇÃO DETALHADA

A presente invenção refere-se a proteínas recombinantes que possuem atividade de N-acetilglucosaminiltransferase onde a proteína recombinante catalisa a transferência de N-acetilglucosamina (GlcNAc) a um

5 Resíduo Man α 3 terminal e catalisa a transferência de N-acetilglucosamina com um resíduo Man α 6 terminal de um aceitador glicano, e em que a proteína recombinante contém os domínios catalíticos de pelo menos duas enzimas diferentes.

Em algumas modalidades, as proteínas recombinantes da presente invenção, incluem dois domínios catalíticos, em que um domínio catalítico possui uma atividade N-acetilglucosaminiltransferase I (GnTI) (por exemplo, reage com um resíduo Man α 3 terminal), e o outro tem domínio catalítico N - acetilglucosaminiltransferase II (GnTII) (por exemplo, reage com um resíduo terminal de Man α 6).

15 Em algumas modalidades, as proteínas recombinantes da presente invenção catalisam as reações que ocorrem essencial e sequencialmente. Por exemplo, as proteínas recombinantes da presente invenção podem catalisar a transferência de GlcNAc a um resíduo terminal de Man α 3, em primeiro lugar, e, em seguida, catalisar a transferência de GlcNAc a um

20 resíduo Man α 6 terminal um glicano aceitador. Em uma modalidade, as reações sequenciais são essencialmente de pelo menos 10 vezes, pelo menos 20 vezes, pelo menos 30 vezes, pelo menos 40 vezes, pelo menos 50 vezes, pelo menos 60 vezes, pelo menos 70 vezes, pelo menos 80 vezes, pelo menos 90 vezes, ou pelo menos 100 vezes mais eficazes do que as

25 duas reações em ordem inversa. Em certas modalidades, a reação sequencial significa que, essencialmente, ou absolutamente nenhuma GlcNAc podem ser transferidas para o resíduo Man α 6 terminal -se GlcNAc ainda não tenha sido transferido para o resíduo Man α 3 terminal. Em uma modalidade específica, o glicano aceitador contém um GlcNAc β 2Man α 3 ramificado .

30 Em algumas modalidades, as proteínas recombinantes reagem, de uma maneira específica, com ambos os resíduos Man α 3 Man α 6 e, opcionalmente, em glicanos aceptadores ramificados, mas não substancialmente

ou totalmente com outras estruturas Man α , por exemplo, Conjugados Man α -monossacarídeos, com Man α benzil /e/ou Man α Ser / Thr-peptídeo. A reatividade não substancial é de preferência inferior a 10 %, abaixo de 8 %, inferior a 6 %, inferior a 4 %, inferior a 2 %, inferior a 1 %, ou inferior a 0,1 % da Vmax, com concentrações de 0,1 mM de glicano aceitador de reações com os resíduos Man α 3 e Man α 6 terminais. Em uma modalidade específica, as proteínas recombinantes têm reatividades substancialmente similares com o terminal Man α 3 (preferencialmente como reação GnTI) e o resíduo Man α 6 terminal (preferencialmente como reação GnTII) do glicano aceitador. De preferência, nem a atividade catalítica tem mais do que uma diferença de 5 vezes, 10 vezes, 3 vezes ou 2 vezes na eficiência de reação em comparação com o outro a atividade catalítica sob as mesmas condições.

Em uma modalidade específica, a transferência de GlcNAc ao terminal Man α 3 e Man α 6 causa uma conversão de pelo menos 10 %, pelo menos 25 %, pelo menos 50 %, pelo menos 70 %, pelo menos 90 %, ou pelo menos 95 % de Man3 glicano com um glicano dois GlcNAcs terminais. A eficácia da reação pode ser medido pelo método in vitro ou em ensaios in vivo como descrito nos exemplos aqui revelados. A eficácia das reações de transferência de GlcNAc pode ser medida essencialmente tal como descrito nos exemplos ou como velocidade de reação máxima Vmax com concentrações de 0,1 mM e as concentrações de aceitador doador saturantes. Em uma modalidade específica, a eficácia da reação é medida por meio de um glicano aceitador de Man3 ligado a um aminoácido, um peptídeo, ou um polipeptídeo.

A presente descrição também se refere a métodos de produção de um complexo de N-glicanos, incluindo as etapas de proporcionar uma célula hospedeira, onde a célula hospedeira contém um ácido nucleico que codifica para uma proteína de fusão que contém um domínio catalítico N -acetilglucosaminiltransferase I e um domínio catalítico N -acetilglucosaminiltransferase II, e cultura da célula hospedeira de tal modo que a proteína de fusão é expressa, a proteína de fusão em que catalisa a transferência de N-acetilglucosamina com um Resíduo Man α 3 terminal e N-acetilglucosamina a

um resíduo Man α 6 terminal de um glicano aceitador com a finalidade de produzir um complexo de N-glicanos.

5 A presente invenção refere-se também a uma célula fúngica filamentosa tendo uma redução do nível de expressão de um gene *alg3* em comparação com o nível de expressão em uma célula fúngicas filamentosos do tipo selvagem, em que a célula fúngica filamentosa contém uma proteína recombinante da presente invenção .

Definições

10 Tal como usado na presente invenção, o termo "proteína recombinante" refere-se a qualquer proteína que tenha sido produzida a partir de um ácido nucleico recombinante. O termo "ácido nucleico recombinante", tal como usado na presente invenção refere-se a um polímero de ácidos nucleicos em que pelo menos um dos seguintes é verdadeiro: (a) a sequência de ácidos nucleicos que é estranha a (isto é, não é de uma forma natural en-

15 contrada em) uma determinada célula hospedeira, (b) a sequência pode ser encontrada de uma forma natural em uma determinada célula hospedeira, mas está presente em uma (por exemplo, maior do que o esperada) quantidade anormal ou expressa a um nível que é mais ou menos do que o nível de expressão natural, ou (c) a sequência de ácidos nucleicos que inclui duas

20 ou mais subsequências que não se encontram na mesma relação entre si na natureza. Por exemplo, em relação exemplo (c), uma sequência de ácido nucleico recombinante que tem duas ou mais sequências de genes não relacionados arranjadas com a finalidade de fazer um novo ácido nucleico funcional. Em outro exemplo, uma sequência de ácido nucleico recombinante

25 irá conter uma sequência promotora e uma sequência codificadora do gene que não é encontrada de uma forma natural adjacente um ao outro.

Tal como usado na presente invenção, o termo "atividade de N-acetilglucosaminiltransferase" refere-se à atividade de uma enzima que transfere um resíduo de N-acetilglucosaminil (GlcNAc) para um glicano acei-

30 tador. Normalmente, as enzimas que têm esta atividade são N-acetilglucosaminiltransferases (GlcNAc transferase). Em certas modalidades, GlcNAc transferases são eucariotas. Em certas modalidades, as transferases de

GlcNAc são enzimas de mamíferos que formam uma ligação- β da posição 1 de um resíduo GlcNAc para os resíduos de manose terminais. Em certas modalidades, as transferases de GlcNAc são β 2-N-acetilglucosaminiltransferases transferência β 2 ligada GlcNAc-resíduo (s) na posição-2 resíduos de manose terminais de glicanos, em particular a um glicano N-ligado. Em certas modalidades, as transferases β 2-GlcNAc são enzimas com atividade GnTI e GnTII atividade. Atividade GnTI transfere um resíduo de GlcNAc a um ramo Man α 3. O ramo Man α 3 pode ser um Man α 3 (R-Man α 6) Man β -ramo da estrutura em N-glicanos ligados núcleo, tais como Man3GlcNAc2 ou Man5GlcNAc2 ou Man3 ou man5. GnTI enzimas podem ser enzimas de mamíferos, enzimas de plantas, enzimas ou eucariotas inferiores. Atividade GnTII transfere um resíduo de GlcNAc a um Man α 6-galho, como um Man α 6 (GlcNAc β 2Man α 3) Man β -galho de uma estrutura central glicanos N-linked. Um exemplo de um tal Man α 6 ramos é GlcNAc1Man3GlcNAc2.

Tal como usado na presente invenção, "N-acetilglucosamina" refere-se a um resíduo de N-acetilglucosamina (GlcNAc). GlcNAc pode ser parte de uma estrutura de glicano. O grupo amina está na posição 2, tem uma configuração D, e tem uma estrutura de piranose como um resíduo. Ele pode ser de uma maneira alternativa designado 2-acetamido-2-desoxi-D-glucopiranosose (D-GlcpNAc). GlcNAc pode também ser um monossacarídeo livre redutor (isto é, não faz parte do glicano).

Tal como usado na presente invenção, "Man", refere-se a um resíduo de manose. Um "terminal Man α 3" ou um "terminal Man α 6" refere-se a um de manose não substituída para o resíduo terminal da extremidade não redutora por um outro resíduo de monossacarídeo ou resíduos.

Tal como usado na presente invenção, "glicano" refere-se a uma cadeia de oligossacarídeo, que pode ser ligado a um veículo, tal como um aminoácido, peptídeo, polipeptídeo, lípido ou um conjugado extremidade redutora. Em certas modalidades, a presente invenção relaciona-se com glicanos N-ligados conjugados a um local de N-glicosilação de polipeptídeo, tais como N-Asn-Xxx-Ser/Thr- por ligação a cadeia lateral de amida de nitrogênio resíduo de asparagina (Asn), onde XXX é qualquer resíduo de ami-

noácido excepto Pro. A presente invenção pode ainda relacionar glicanos, como parte de dolicol-fosfo-oligossacarídeo (PP-Dol-OS), estruturas lipídicas precursores, que são precursores dos glicanos N-ligados no retículo endoplasmático de células eucariotas. Os oligossacarídeos precursores são ligados a partir de sua extremidade redutora a dois resíduos de fosfato sobre o lípido dolicol. Por exemplo, Alg3 α 3-manosiltransferase modifica o precursor PP-Dol-oligossacarídeo de N-glicanos. Geralmente, as estruturas de glicano aqui descritas são as estruturas de glicano de terminais, onde os resíduos não redutores não são modificadas por outro resíduo de monossacarídeo ou resíduos.

Tal como usado na presente invenção, "glicoproteína" refere-se a um peptídeo ou polipeptídeo ligado a um glicano. O glicano pode ser ligado ao peptídeo ou polipeptídeo em uma cotranslacional ou modificação pós-translacional.

Tal como usado na presente invenção, "glicolípido" refere-se a um lípido ligado a um glicano e inclui gliceroglicolipídios, glicoesfingolipídios e Glicosilfosfatidilinositóis.

Como utilizado em toda a divulgação, glicolipid e carboidratos nomenclatura atual é essencialmente de acordo com as recomendações da Comissão IUPAC-IUB sobre Nomenclatura Bioquímica (por exemplo, carboidratos Res 1998, 312, 167 ; . Res carboidrato 1997, 297, 1 ; . Eur. Eur. J. Biochem. 1.998, 257, 29). Supõe-se que a Gal (galactose), Glc (glicose), GlcNAc (N-acetilglucosamina), GalNAc (N-acetilgalactosamina), Man (manose) e Neu5Ac são da configuração D, Fuc da configuração L, e todas as unidades de monossacarídeo sob a forma de piranose (D-Galp D-Glcp, D-GlcpNAc, D-GalpNAc, MANP-D, L-Fucp, D-Neup5Ac). O grupo amina é como definido para a galactose natural e glucosaminas na posição 2 de GlcNAc ou GalNAc. Glicosídicas são mostradas parcialmente em mais curtos e, em ptécnica, já nomenclatura, as ligações dos SA/Neu5X-residues ácido siálico α 3 α 6 e significa o mesmo que α 2-3 e α 2-6, respectivamente, e para os resíduos monossacarídeos de hexoses α 1-3, α 1-6, β 1-2, β 1-3, 4- β 1 e β 1-6 pode ser encurtada como α 3, α 6, β 2, β 3, β 4 e β 6, respectivamente. Lacto-

samina tipo II refere-se a N-acetilactos, Gal β 4GlcNAc, /e/ou de tipo I, N-acetilactosamina. Gal β 3GlcNAc e ácido siálico (SA) referem-se ao ácido N-acetilneuramínico (Neu5Ac), ácido N-glicolilneuramínico (Neu5Gc), ou qualquer outro ácido siálico natural, incluindo derivados de Neu5X. Ácido siálico

5 é referido como NeuNX ou Neu5X, onde a preferência X é Ac ou Gc. Ocasionalmente Neu5Ac/Gc/X pode ser referido como NeuNAc / NeuNGc / NeuNX.

As proteínas recombinantes da presente invenção

A presente invenção refere-se a proteínas recombinantes que

10 possuem atividade de N-acetilglucosaminiltransferase, onde as proteínas recombinantes catalisam a transferência de N-acetilglucosamina com um resíduo Man α 3terminal e catalisam a transferência de N-acetilglucosamina com um resíduo Man α 6 terminal de um glicano aceitador. As proteínas re-

combinantes da presente invenção podem incluir, sem limitação, as proteí-

15 nas de comprimento completo possuindo atividade N-acetilglucosaminiltransferase, fragmentos de proteínas que possuem atividade de N-acetilglucosaminiltransferase, domínios catalíticos com atividade N-acetilglucosaminiltransferase, e proteínas de fusão que possuem atividade de N-acetilglucosaminiltransferase. Uma única proteína recombinante da presente in-

20 venção tem a capacidade de catalisar a ambas as transferências de N-acetilglucosaminas. A transferência de N-acetilglucosamina com um Resíduo Man α 3 terminal pode ocorrer antes ou após a transferência de N-acetilglucosamina com um resíduo Man α 6 terminal. De uma maneira alternativa, as transferências podem ocorrer simultaneamente.

25 O glicano aceitador pode ser ligado a uma molécula tal como um aminoácido, um peptídeo, ou um polipeptídeo. Em certas modalidades, o aminoácido é um resíduo de asparagina. O resíduo de asparagina pode ser em aminoglicosídico ligado a partir da amida da cadeia lateral (uma estrutura de ligação de polipeptídeo de N-glicanos de mamíferos biológico) e podem

30 ser parte de uma cadeia peptídica, tais como um dipeptídeo, um oligopeptídeo ou um polipeptídeo. O glicano pode ser um derivado final de redução, tal como uma N-, O-, ou C-ligado, de preferência glicosídico, derivado da Glc-

NAC reduzindo ou Man, tal como um separador ou resíduo orgânico terminal com uma determinada estrutura de glicano de ligação selecionada a partir de entre o grupo de um aminoácido, um grupo alquila, heteroalquila, acila, alquilóxi, arila, arilalquila, ou heteroarilalquila. O espaçador pode ser ainda ligado a um suporte polivalente ou uma fase sólida. Em certas modalidades, as estruturas contendo alquila incluem metila, etila, propila, e C4-C26 alquilas, tais como lípidos, fosfolípidos, glicerolípídios dolicol-fosfolípidos e as ceramidas e derivados. A extremidade redutora pode ser também derivatizada por meio da aminação redutiva de uma ligação amina secundária ou uma estrutura derivada. Certos veículos incluem biopoli- ou oligômeros, tais como (poli) peptídeos, poli (sacarídeos), tais como dextrano, celulose, amilase, ou glicosaminoglicanos, e outros oligômeros ou polímeros orgânicos, tais como plásticos, incluindo o polietileno, polipropileno, poliamidas (por exemplo, de náilon ou de poliestireno), poliacrilamida, e os ácidos polilácticos, tais como os dendrímeros PAMAM, Starburst ou dendrímeros do mar, ou a polilisina, e polialquilglicóis tais como polietilenoglicol (PEG). As fases sólidas podem incluir cavidades de microtitulação, as partículas de sílica, vidro, metal (incluindo o aço, ouro e prata), grânulos de polímeros, tais como esferas de poliestireno ou resina de ácido poliláctico, grânulos, grânulos de polissacarídeos ou espaçadores orgânicos contendo contas magnéticas.

Em certas modalidades, o glicano aceitador está ligado a um polipeptídeo heterólogo. Tal como usado na presente invenção, o termo um "peptídeo" e um "polipeptídeo" são sequências de aminoácidos incluindo uma pluralidade de resíduos de aminoácidos polimerizados consecutivos. Para efeitos da presente invenção, tipicamente, os peptídeos são aquelas moléculas, incluindo um máximo de 50 resíduos de aminoácidos e polipeptídeos incluem mais de 50 resíduos de aminoácidos. O peptídeo ou polipeptídeo pode incluir resíduos de aminoácidos modificados, que ocorrem de uma forma natural, os resíduos de aminoácidos não codificados por meio de um códon, e que não ocorrem de uma forma natural, dos resíduos de aminoácidos. Tal como usado na presente invenção, o termo "proteína" pode referir-se a um peptídeo ou um polipeptídeo de qualquer tamanho. A expressão

"polipeptídio heterólogo" refere-se a um polipeptídeo que não é encontrado de uma forma natural em uma determinada célula hospedeira, ou não é endógeno para uma determinada célula hospedeira. Em certas modalidades, o polipeptídeo heterólogo é uma proteína terapêutica. As proteínas terapêuticas, por exemplo, podem incluir os anticorpos monoclonais, interferonas, eritropoetinas, hormonas de crescimento, enzimas ou fatores de coagulação do sangue. Por exemplo, o glicano aceitador s pode ser ligado a uma proteína terapêutica, tais como rituximab.

Glicano Aceitador

Em certas modalidades, a estrutura do glicano aceitador tem a seguinte fórmula $[R1] y\text{Man}\alpha3 ([R2] z\text{Man}\alpha6) \text{Man} \{ \beta4\text{GlcNAc} (\text{Fuc}\alpha x) n [\beta4\text{GlcNAc}] m \} q$, onde q, y, z, n e m são 0 ou 1, x é a posição de ligação 3 ou 6, dos resíduos de fucose opcional ; R1 é GlcNAc, preferencialmente GlcNAc β 2, e R2 é uma estrutura ramificada $\text{Man}\alpha3$ ($\text{Man}\alpha6$), com a condição de que, quando z é 1, y é então 0, e quando z for 0, então Y é 0 ou 1. () define uma filial na estrutura do núcleo de N-glicano regular, presente ou ausente. [] e { } definem uma parte da estrutura do glicano que pode estar presente ou ausente em uma sequência linear. Quando z é 0 e y é 0, então a estrutura é um glicano $\text{Man}3$ e, quando z é 0 e y é 1, a estrutura é uma GlcNAc $\text{Man}3$ glicano. Quando Y é 0 e z é 1, o glicano é um $\text{man}5$ glicano. O glicano aceitador pode ser beta-glicosidicamente ligado a um resíduo de Asn e de preferência entre a redução GlcNAc final. Em uma modalidade, o glicano aceitador é um polipeptídeo ligada a N-glicano, em que m e q são 1, e a estrutura aceitadora contém um derivado de $[R1] y\text{Man}\alpha3 ([R2] z\text{Man}\alpha6) \text{Man}\beta4\text{GlcNAc} (\text{Fuc}\alpha x) n\beta4\text{GlcNAc}$. Os derivados opcionais incluem as substituições de resíduos monossacarídeos tais como GlcNAc ou xilase.

O glicano aceitador pode ser $\text{Man}3$, GlcNAc $\text{Man}3$ ou $\text{man}5$. Em certas modalidades, o glicano aceitador é $\text{Man}3$ ou GlcNAc $\text{Man}3$. $\text{Man}3$ é um glicano trimannosila que compreende pelo menos um dos $\text{Man}\alpha3$ ou $\text{Man}\alpha6$ resíduos e é de preferência um Man oligossacarídeo, como $\text{Man}\alpha3$ ($\text{Man}\alpha6$) ramificado. Outros certos $\text{Man}3$ oligossacarídeos são $\text{Man}\alpha3$ ($\text{Man}\alpha6$) $\text{Man}\beta$, $\text{Man}\alpha3$ ($\text{Man}\alpha6$) $\text{Man}\beta4\text{GlcNAc}$ e $\text{Man}\alpha3$ ligado ao polipeptídeo ($\text{Man}\alpha6$)

Man β 4GlcNAc β 4GlcNAc. Além disso, dependendo da célula hospedeira, o glicano pode conter Fuc, Xil ou em Man β GlcNAc /e/ou os resíduos de GlcNAc, como Man α 3 (Man α 6) Man β 4GlcNAc β 4 (Fuc α x) nGlcNAc, onde x é 3 ou 6 e n é 0 ou 1, também descrita por uma fórmula que indica a composição de monossacarídeos estrutura de manose terminais e reduzindo composição final como Man3GlcNAc2 (n é 0) e Man3GlcNAc2Fuc (n é 1). Em certas modalidades, especialmente aquelas com uma estrutura de polipeptídeo ligado, a estrutura é uma Man α 3 (Man α 6) Man β 4GlcNAc β 4 (Fuc α 6) Man3nGlcNAc. Em certas modalidades, a estrutura de polipeptídeo ligada a GlcNAcMan3 é GlcNAc β 2Man α 3 (Man α 6) Man β 4GlcNAc β 4 (Fuc α 6) nGlcNAc, também descrita por meio de uma composição de monossacarídeos de fórmula GlcNAcMan3GlcNAc2 (n é 0) e GlcNAcMan3GlcNAc2Fuc (n é 1). Em certas modalidades, a estrutura de man5 ligada ao polipeptídeo é Man α 3 { Man α 3 (Man α 6) Man α 6 } Man β 4GlcNAc β 4 (Fuc α 6) nGlcNAc, onde { } e () indicam uma ramificação e n é 0 ou 1, também descrita por meio de uma composição de monossacarídeos de fórmula Man5GlcNAc2 (n é 0) e Man5GlcNAc2Fuc (n é 1).

Por conseguinte, certos Man3 glicanos possuem estruturas de acordo com a seguinte fórmula, Man α 3 (Man α 6) Man β 4GlcNAc (Fuc α x) n β 4GlcNAc, onde n é 0 ou 1, indicando a presença ou ausência de uma parte da molécula, em que x é 3 ou 6, e onde () define uma filial na estrutura. Em modalidades da presente invenção em que o glicano aceitador é Man3, a proteína recombinante catalisa a transferência de N-acetilglucosamina do terminal Man α 3 e Man α 6 de Man3, resultando assim em GlcNAc2Man3, GlcNAc β 2Man α 3 (GlcNAc β 2Man α 6) Man β 4GlcNAc β 4 (Fuc α x) nGlcNAc, onde n é 0 ou 1, também descrita por uma composição de monossacarídeos fórmula GlcNAc2Man3GlcNAc2 (n é 0) e GlcNAc2Man3GlcNAc2Fuc (n é 1).

Em modalidades da presente invenção em que o glicano aceitador é man5, a proteína recombinante, catalisa a transferência de N-acetilglucosamina do terminal de Man α 3 man5. Depois de duas manoses terem sido removidas GlcNAcMan5 (por exemplo, por manosidase II) para formar GlcNAcMan3, a proteína recombinante, catalisa a transferência de N-acetil-

glucosamina do terminal Man α 6, resultando assim em GlcNAc β 2Man α 3 (que tem a estrutura GlcNAc β 2Man α 3 (GlcNAc β 2Man α 6) Man β 4GlcNAc β 4 (Fuc α x) nGlcNAc, onde n é 0 ou 1, também referido como G0 se ligado a um anticorpo).

5 Proteínas de fusão contendo os domínios catalíticos de N-acetilglucosaminiltransferase

Em certas modalidades, as proteínas recombinantes da presente invenção são proteínas de fusão contendo um domínio catalítico de N-acetilglucosaminiltransferase I e II, um domínio catalítico de N-acetilglucosaminiltransferase. O termo "proteína de fusão" refere-se a qualquer proteína ou polipeptídeo contendo uma proteína ou um polipeptídeo ligado aos aminoácidos heterólogos.

N-acetilglucosaminiltransferase I (GlcNAc-TI; GnTI; CE 2.4.1.101) catalisa a reação de UDP-N-acetil-D-glucosamina + 3 - (alfa-D-manosila)-beta-D-manosil-R \rightarrow UDP + 3 - (2 - (N-acetil-beta-D-glucosaminila)-alfa-D-manosila)-beta-D-manosil-R, onde R representa o restante do oligossacarídeo ligado a N no glicano aceitador. Um Domínio catalítico de N-acetilglucosaminiltransferase I é qualquer porção de uma enzima N-acetilglucosaminiltransferase I que é capaz de catalisar esta reação. As sequências de aminoácidos para as enzimas N-acetilglucosaminiltransferase I a partir de vários organismos estão listadas nas SEQ ID NOs : 1 a 19. As enzimas GnTI adicionais estão listadas na base de dados CAZy na família glicosiltransferase 13 (cazy.org/GT13_all). As espécies enzimaticamente caracterizadas incluem *A. thaliana* AAR78757.1 (US6 653 a 459), *C. elegans* AAD03023.1 (Chen S. et al J. Biol 1999, 274 (1) : 288 a 97), *D. melanogaster* AAF57454.1 (Sarkar e Schachter Biol Chem, Fevereiro de 2001, 382 (2) : 209 a 17) ; *C. griseus* AAC52872.1 (Puthalakath H. et al J. Biol 1996 271 (44) : 27818 a 22) ; *H. sapiens* AAA52563.1 (Kumar R. et al Proc Natl Acad Sci EUA. dezembro de 1990, 87 (24) : 9948 a 52) ; *M. auratus* AAD04130.1 (Opat Como et al Biochem J. dezembro de 1998 15 ; 336 (Pt 3) : 593 a 8), (incluindo por exemplo, desativar as mutante), coelho, *O. cuniculus* AAA31493.1 (Sarkar M. et al Proc Natl Acad Sci EUA 01 de janeiro 1991 A., 88 (1.) : 234 a 8). Exemplos

adicionais de enzimas ativas caracterizadas podem ser encontrados em cazy.org/GT13_characterized. A estrutura 3D do domínio catalítico de GnTI de coelho foi definida por meio da cristalografia de raios X em Unligil UM et al. EMBO J. 2000 16 de outubro, 19 (20) : 5269 a 80. As estruturas dos bancos de dado da proteína (APO) para GnTI são 1FO8, 1FO9, 1FOA, 2AM3, 2AM4, 2AM5 e 2APC. Em certas modalidades, o domínio catalítico de N-acetilglucosaminiltransferase I é da enzima humana N-acetilglucosaminiltransferase I (SEQ ID NO : 1), ou as suas variantes. Em certas modalidades, o domínio catalítico de N-acetilglucosaminiltransferase I contém uma sequência que é pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica aos resíduos de aminoácidos 84 a 445 da SEQ ID NO : 1. Em algumas modalidades, uma sequência mais curta pode ser usada como um domínio catalítico (por exemplo, os resíduos de 105 a 445 aminoácidos da enzima humana, ou os resíduos de aminoácidos 107 a 447 da enzima de coelho, Sarkar et al (1998) Glicoconjugate J 15 : 193 a 197). As sequências adicionais que podem ser usadas como o domínio catalítico de GnTI incluem os resíduos de aminoácidos desde cerca de 30 a 445 aminoácidos da enzima humana ou de qualquer domínio da haste do terminal C começando entre o resíduo de 30 a 105 aminoácidos e continuando até cerca dos 445 aminoácidos da enzima humana, ou a correspondente sequência homóloga de outro GnTI ou uma variante cataliticamente ativa ou mutante. O domínio catalítico pode incluir as partes do Terminal N da enzima, tal como a totalidade ou parte do domínio da haste, o domínio transmembranar, ou o domínio citoplasmático.

Tal como usado na presente invenção, o termo "citoplasmático" é utilizado para se referir a uma parte de uma proteína que interage com o citoplasma de uma célula.

N-acetilglucosaminiltransferase II (GlcNAc-TII; GnTII; CE 2.4.1.143) catalisa a reação de UDP-N-acetil-D-glucosamina + 6 - (alfa-D-manosila)-beta-D-manosil-R \rightleftharpoons UDP + 6 - (2 - (N-acetil-beta-D-glucosaminila)-alfa-D-manosila)-beta-D-manosil-R, onde R representa o restante do oligossacari-

deo ligado a N no glicano aceitador. Um domínio catalítico de N-acetilglucosaminiltransferase II é qualquer porção de uma enzima N-acetilglucosaminiltransferase II, que é capaz de catalisar esta reação. As sequências de aminoácidos para as enzimas N-acetilglucosaminiltransferase II a partir de vários organismos está listada em SEQ ID NOs : 20-33. Em certas modalidades, o domínio catalítico de N-acetilglucosaminiltransferase II da enzima é de N-acetilglucosaminiltransferase II humana (SEQ ID NO : 20). As espécies GnTII adicionais estão listadas na base de dados CAZy na família glicosiltransferase 16 (cazy.org/GT16_all). As espécies enzimaticamente caracterizados incluem GnTII de *C. elegans*, *D. melanogaster*, o *Homo sapiens*, *Rattus norvegicus*, *Sus scrofa* (cazy.org/GT16_characterized). Em certas modalidades, o domínio catalítico de N-acetilglucosaminiltransferase II contém uma sequência que é pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica aos resíduos de aminoácidos a partir de cerca de 30 a 447 da SEQ ID NO : 21. O domínio catalítico podem incluir partes Terminal N da enzima, tal como a totalidade ou parte do domínio da haste, o domínio transmembranar, ou o domínio citoplasmático.

Em certas modalidades, o domínio catalítico de N-acetilglucosaminiltransferase I é o Terminal N para o domínio catalítico de N-acetilglucosaminiltransferase II. Em outras modalidades, o domínio catalítico de N-acetilglucosaminiltransferase II do Terminal N é o domínio catalítico de N-acetilglucosaminiltransferase I. O termo "Terminal N" refere-se ao posicionamento de um conjunto de resíduos de aminoácidos mais perto da extremidade de um polipeptídeo que é terminado por um aminoácido com um grupo amina livre (-NH₂), em comparação com um conjunto de referência de resíduos de aminoácidos .

Espaçadores

Em certas modalidades da presente invenção, a proteína recombinante contém um espaçador entre a N-acetilglucosaminiltransferase I e domínio catalítico Domínio catalítico de N-acetilglucosaminiltransferase II.

O termo "espaçador" refere-se a qualquer número de aminoácidos consecutivos de qualquer sequência de separação a N-acetilglucosaminiltransferase I e domínio catalítico da II domínio catalítico de N-acetilglucosaminiltransferase de tal forma que o espaçador não tem efeito sobre a função enzimática dos domínios catalíticos. Tipicamente, o espaçador é de pelo menos 5, pelo menos 10, pelo menos 15, pelo menos 20, pelo menos 30, pelo menos 40, ou pelo menos 50 aminoácidos de comprimento. Em certas modalidades, o espaçador contém uma sequência de domínio haste. "Domínio haste" refere-se a um domínio de proteína ou um seu fragmento, que está localizada adjacente ao domínio de transmembrana de uma enzima nativa, tal como uma glicosiltransferase ou uma hidrolase de glicosila, e, opcionalmente, tem como alvo o enzima ou auxilia na retenção da enzima na ER/Golgi. Domínios Estaminais geralmente começam com a primeira sequência de aminoácidos do domínio transmembrantar hidrófobo e termina no domínio catalítico.

Domínios estaminais exemplares incluem, mas não estão limitados a, o domínio da haste de GnTI humano, resíduos de aminoácidos a partir de cerca de 30 até cerca de 83 ou desde cerca de 30 a cerca de 105 para o GnTII humana, ou os resíduos de aminoácidos a partir de cerca de 26 a cerca de 106 ou desde cerca de 26 a cerca de 83 para o *T. reesei* KRE2. Em certas modalidades onde o espaçador contém uma sequência de domínio de haste, o espaçador inclui os 30 a 83 aminoácidos da sequência de GnTI humana (SEQ ID NO : 34). Em outras modalidades, o espaçador pode incluir qualquer uma das sequências listadas nas SEQ ID NOs : 35 a 38.

Outros exemplos de espaçadores adequados incluem, sem limitação, os 3XG4S flexíveis separadores (SEQ ID NO : 118), o espaçador 2XG4S flexível (SEQ ID NO : 120), o espaçador para o *T. reesei* cbhl (SEQ ID NO : 122), e o espaçador para a celulase de *T. reesei* EGIV (SEQ ID NO : 124).

Em certas modalidades, o comprimento do espaçador é o mesmo que o comprimento de uma haste de domínio GnT1. Em certas modalidades, o comprimento é de cerca de 74 resíduos de aminoácidos, mais ou menos cerca de 37 aminoácidos. Por exemplo, o comprimento do espaçador

é de cerca de 30 aminoácidos e cerca de 110 aminoácidos, ou a partir de cerca de 35 aminoácidos e cerca de 100 aminoácidos, ou como exemplificado nos exemplos descritos na presente invenção, mais ou menos 2, 3, 4, ou 5-aminoácidos. Em uma modalidade, o comprimento do espaçador corresponde a um domínio de GnT1 haste truncado, por exemplo, começar desde o aminoácido 25 até ao aminoácido 104, ou entre o aminoácido 30 até ao aminoácido 101, até ao final do domínio haste GnT1. Em certas modalidades, o espaçador pode incluir uma parte do caule do domínio GnT1 humano, que pode começar a partir de um aminoácido posicionado entre aminoácido 70 ao aminoácido 87 (de acordo com a numeração da SEQ ID NO : 34), ou entre o aminoácido e aminoácido 76 104, ou no início do aminoácido 30, 35, 40, 45, 50, 60, 70, 73, 74, 75, 76, 80, 83, 84, 85, 86, 87, 100, 101, 102, 103, ou 104, para o fim do domínio haste GnT1 humano. Em outras modalidades, o espaçador pode incluir um peptídeo espaçador heterólogo, que pode incluir um peptídeo espaçador de fungos /e/ou de um peptídeo espaçador oligômero repetitivo.

Tipicamente, o espaçador é um peptídeo alongado sem conformação específica e que contém resíduos de aminoácidos permitindo uma elevada flexibilidade (por exemplo, Gli e Ala), hydrophobicity (por exemplo, Ser e Thr) e opcionalmente Pro para evitar conformação. O espaçador pode ser glicosilado. Em certas modalidades do espaçador é O-glicosilado incluindo O-manosilação fúngica. Em certas modalidades, o espaçador é um peptídeo endógeno fúngico filamentosos ou Trichoderma espaçador, tal como um espaçador que separa de uma forma natural os domínios protéicos. O espaçador pode ser derivado a partir de uma enzima de um fungo, tais como fungos filamentosos (por exemplo, de *T. reesei*), um seu fragmento, ou um múltiplo do espaçador /e/ou o seu fragmento ou mutante seu análogo ou equivalente segregada ou celulolítica. O espaçador fúngico singular pode conter prolina /e/ou glicina /e/ou serina /e/ou treonina dimérica ou oligomérica, /e/ou múltiplos resíduos de aminoácidos selecionados a partir de Ser, Thr, Gli, Pro ou Ala ou quaisquer combinações dos mesmos. Em certas modalidades, o espaçador é um oligômero que contém uma repetição de monômero com 1 a

10 ou 1 a 5 resíduos de aminoácidos selecionados a partir de Ser, Thr, Gli, Pro ou Ala e, opcionalmente, um resíduo de aminoácido carregado de uma forma negativa selecionado a partir dos resíduos de Glu ou Asp ou os resíduos carregados de uma forma positiva Lis ou Arg. Em certas modalidades, o resíduo de carga é carregado de uma forma negativa. Em certas modalidades, o monômero contém os resíduos diméricos ou oligoméricos de aminoácidos, /e/ou múltiplos resíduos de aminoácidos individuais selecionados a partir de Ser, Thr, Gli, Pro e Ala. Em certas modalidades do oligômero contém um monômero de um dímero ou oligômero de glicina e um resíduo único selecionado a partir do Ser, Tre, Gli, Pro e Ala. Em certas modalidades, o único resíduo é Ser ou Thr. Em certas modalidades, o resíduo é Ser. Em certas modalidades, a sequência de repetição é o espaçador { (YYY) Nxxx } m, onde n é 2 a 10, m é de 2 a 10, e Xxx e Yyy são selecionados a partir de entre Ser, Thr, Gli, Pro e Ala, com a condição de que Xxx Yyy e não são o mesmo resíduo de aminoácido. Em certas modalidades o espaçador de repetição é {(Gli) Nxxx} m, onde n é 2 a 10, m é de 2 a 10, e XXX é selecionado a partir de Ser, Thr, Gli, Pro Ala e em certas modalidades Xxx é Ser ou Thr. Em certas modalidades Ser é Xxx.

Peptídeos de Direcionamento

20 Em certas modalidades, as proteínas recombinantes da presente invenção incluem um peptídeo alvo ligado aos domínios catalíticos. O termo "ligado", tal como usado na presente invenção, significa que dois polímeros de resíduos de aminoácidos, no caso de um ou dois polímeros polipeptídicos de nucleotídeos, no caso de um polinucleotídeo ou são acoplados de uma maneira direta ao lado uns dos outros ou estão dentro do mesmo polipeptídeo ou polinucleotídeo mas são separados por meio dos resíduos de aminoácidos ou de nucleotídeos. Um "peptídeo alvo", tal como é usado na presente invenção, refere-se a qualquer número de resíduos de aminoácidos consecutivos da proteína recombinante, que são capazes de localizar a proteína recombinante para o retículo endoplasmático (ER) ou aparelho de Golgi (Golgi) dentro da célula hospedeira. O peptídeo alvo pode ser de Terminal N ou de terminal C para os domínios catalíticos. Em certas modalidades, o

peptídeo alvo é um Terminal N para os domínios catalíticos. Em certas modalidades, o peptídeo de ligação ao alvo proporciona um ER ou Golgi componente, tal como uma enzima de manosidase II. Em outras modalidades, o peptídeo de ligação de direcionamento proporciona directamente a membrana do RE ou Golgi.

Os componentes do peptídeo alvo podem provir de qualquer enzima que normalmente reside no ER ou aparelho de Golgi. Essas enzimas incluem manosidases, manosiltransferases, Glicosiltransferases, Tipo 2 proteínas de Golgi, e MNN2, MNN4, MNN6, MNN9, MNN10, MNS1, KRE2, VAN1 e OCH1 enzimas. Tais enzimas podem ser provenientes de uma espécie de levedura ou fúngicas, tais como aquelas de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filabasidium*, *Fusarium*, *Gibberella*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, e *Trichoderma*. As sequências de tais enzimas podem ser encontradas na base de dados de sequência do GenBank.

Em certas modalidades, o peptídeo alvo provém da mesma enzima e do organismo como um dos domínios catalíticos da proteína recombinante. Por exemplo, se a proteína recombinante compreende um domínio catalítico GnTII humano, o peptídeo alvo da proteína recombinante é da enzima GnTII humano. Em outras modalidades, o peptídeo alvo pode provir de uma enzima diferente /e/ou organismo como os domínios catalíticos da proteína recombinante.

Exemplos de vários peptídeos de direcionamento para utilização na detecção de proteínas para o ER ou Golgi que podem ser utilizadas para alvejar as proteínas recombinantes da presente invenção, incluem : peptídeo Terminal N fundido com galactosiltransferase Kre2/Mnt1 (Schwientek, JBC 1996, 3398), HDEL para localização de manosidase de ER de células de levedura com a finalidade de produzir man5 (Chiba, JBC 1998, 26.298-304 ; Callewaert, FEBS Lett 2001, 173 a 178), peptídeo alvo OCH1 fundido com o domínio catalítico GnTI (Yoshida et al, 1999 Glicobiology, 53 a 8), o peptídeo

Terminal N de levedura fundido com MNS1 α 2-manosidase (Martinet et al, Bibateladach Lett 1998, 1171), a porção Terminal N da Kre2 ligada ao domínio catalítico de GnTI ou β 4GalT (Vervecken, Appl. Environ Microb 2004, 2639 a 46), várias abordagens de avaliação em Wildt e Gerngross (Nature Rev Bibateladach 2005, 119), GnTI full-length em *Aspergillus nidulans* (Kalsner et al, Glicocon. J 1995, 360-370), GnTI de comprimento total em *Aspergillus oryzae* (Kasajima et al, Biosci Bibateladach Biochem 2006, 2662 a 8), a porção da estrutura de localização de levedura Sec12 fundido com GnTI de *C. elegans* em *Aspergillus* (Kainz et al 2008), a porção Terminal N de levedura MNN9 fundido com GnTI humano *Aspergillus* (Kainz et al 2008), a porção Terminal N de *Aspergillus* Mnn10 fundido com GnTI humano (Kainz et al, Appl. Environ Microb 2008, 1076-1086), e de comprimento total GnTI humano em *T. reesei* (Maras et al, FEBS Lett 1999, 365 a 70).

Em certas modalidades, o peptídeo alvo é o Kre2/Mnt1 (isto é, Kre2) peptídeo de polipeptídeo possuindo a sequência de aminoácidos da SEQ ID NO : 115 ou SEQ ID NO : 116.

Outros exemplos de sequências que podem ser utilizados para direcionar os peptídeos incluem as sequências listadas na Tabela 1 abaixo.

Tabela 1: Peptídeos de Direcionamento. Domínios transmembranares putativos estão sublinhados. Em KRE2, o domínio da haste permitindo a localização de Golgi é sublinhado e duplo sublinhado. Outro1 e Outro02 são proteínas relacionadas com a manosiilação putativa.

Homólogo a	Citoplásmico	Transmembrana	Luminal
KRE2 estExt_fgenes1_pm.C_30039	MASTNARYVR SEQ ID NO : 39	<u>YLLIAFFTILVFYFVSN</u> SEQ ID NO : 40	<u>SKYEGVDLNKGTFTAPDSTKTTTPKPPATGDAKDF-</u> <u>PLALTPNDP</u> <u>GFNDLVGIAPGPRMNATFVTLARNSDVWDIARSIRQVEDRFN</u> RRNYDWVFLNDKPFNDTFKVVTTSLVSGKTHYGEIAPEHW SFPDWIDQDKAKKVRDMAERKIY- DSVSYRHMCRFESGFFRQPLMMNYEYYWRV EPSIELICDIHYDPFRLMVEQKKYSFVISLIEYPATIALWEST KKFMKNHPEHIAPDNSMRFLSDDGGGETYNNCHFWSNFEIGSL EWLRSKQYIDFFESLDKDGFFYERWGDAPVHSIAAGLMLNR SEIHFFNDIAYWHVPFTHCPTGEKTRLDLKCHCDPKENFDWK GYSCTSRFFEMNGMDKPEGWENQQD SEQ ID NO : 41
KRE2 alternativa e_gw1.28.231.1	MAIARPVR SEQ ID NO : 42	<u>ALGGLAAILWCFFLI</u> SEQ ID NO : 43	QLLRPSSSYNSPGDRYINFERDPNL- PTGEPEGILVRTSDRYAPDAK DTDRASATLLALVRNEEVDDMVASMVDLERTWNSKFNYPW

Homólogo a	Citoplásmico	Transmembrana	Luminal
			<p>TFNDKPFSEEFKKKTSAVTNATCNYPKEHWDAPSWIDPA</p> <p>IFEESA AVLKKNQVQYANMMSYHQMCRWNSGMFYKHPALK</p> <p>DVRYYYWRVEPKVHFFCDVDYDFRYMQDNNKTYGFTINLID</p> <p>DPHLTPLWPQTAKFLADHPNYLHEHSAIKWVIDDARRPQHN</p> <p>REAQGFSTCHFWSNFEVADMEFWRSKVYEDYFEHLDRAGGF</p> <p>FYERWGDAPVHSIALGLFEDSSKIHWFRDIGYQHIPPFCNPSP</p> <p>KCKGCVTGRLTDGEPFLHREDCRPNWFKYAGMG</p> <p>SEQ ID NO : 44</p>
<p>OCH1</p> <p>e_gwl.16.371.1</p>	<p>MLNPRR</p> <p>SEQ ID NO : 45</p>	<p><u>ALIAAAFILTVFFLI</u></p> <p>SEQ ID NO : 46</p>	<p>SRSNSESASTSEPKDAEAEALSAANAQQRAAPPPPPQKPMID</p> <p>MSGMSTYDKLAYAYEYDIESKFPAYIWQTWKTPSEGDFFR</p> <p>EQEASWSIEHPGFIHEVITDSVADTLLQLLIGSIPEVLEAYHALP</p> <p>LPVLKADLFRYLILIAAGGIYSDIDTYAIRSALEWIPPQIPKETV</p> <p>GLVIGIEADPRPDWADWYSRRIQFCQWTTQSKPGHPVLRDII</p> <p>SRITNQTLKEMKSKGLSAFQGNRVVDLTGPAVWTDTIMDYF</p> <p>NDERYFDMENSKGRIDYRNFTGMTSKRVGDVVVLPTSFSFSP</p> <p>GVGQMGAKDYDDPMAFVKHDFEGTWKPE-</p> <p>ERHIGEIVQELGEGQGEAPKEQ</p> <p>SEQ ID NO : 47</p>

Homólogo a	Citoplásmico	Transmembrana	Luminal
OCH1 alternativa fgenes1_pm.C_scaffold_13000080	MGMGQCQWSPFRNKVPTQMRRRC SEQ ID NO : 48	<u>LPLIITVVCVFLVIV</u> SEQ ID NO : 49	NFDWILAIPNPASVLRREPAPPLPGSTFPQKIWQTWKVDPLN FDERDLVTARTWTINPGMRYEVVTDANEMAYIEDRYGPN FDRPDIVEFYKMINLPIIKADLLRYMIMYAEAGGIYADIDVETM KPFHRFIPDRYDEKDIDIIGVEIDQPDFKDHPILGKKSMSFCQ WTFVARPQQPVMMRLIENIMKWFKTVARDQGVPLGEVQLDF DQVISGTGPSAFTKAMLEEMNRKTKGPKVTWDAFHNLDSEK LVGGVLVLTVEAFACAGQHSDSGNHNARNALVKHHFHASN WPSRHPRYKHPAYGQVEDCNWVPECVRKWDEDTSNWDKYS ENEQKKILQDIENARLERERQQAALP SEQ ID NO : 50
MNN9 e_gw1.5.262.1	MARPMGSVRLKKANPST SEQ ID NO : 51	<u>LILGAVLCIFIIIFLV</u> SEQ ID NO : 52	SPSSPASASRLSIVSAQHHLSPPTSPYQSPRSGAVQGPPPVTRY NLNKVTVTSDPVRNQEHLILTPMARFYQEYWDNLLRLNYPH ELITLGFILPKTEGNQATSMLOKQIQKTQNYGPEKDRFKSIII LRQDFDPAVVSQDESERHKLANKQKARREVMKARNSSLFTT LGPSTSWVLWLDADITETAPTLIQDLASHDKPIIVANCFQKYY DPESKKMAERPYPDFNSWQDSETALKMAEQMGPDILLLEGYA EMATYRTLAYMSTPGGSKDLVVPDGVGGTALLVKADVH RDGAMFPFPFAFYHLIESEGEFAKMAKRLGWQPYGLPNYKVYH

Homólogo a	Citoplásmico	Transmembrana	Luminal
			YNE SEQ ID NO : 53
MNN9 Alternativa1 estExt_GeneWisePlus.C_230146	MLLPKGGLDWRSAQAQIPPTRAL WNAVTRTR SEQ ID NO : 54	<u>FILLVGITGLILLW</u> SEQ ID NO : 55	RGVSTASSEMQSYCWGPAKPPMEMSPNEHNRWNGHLQTPV IFNHHAPVEVNSSTIEHVLDLPINSTKQAVTKEERILITPLKD AAPYLSKYFELLAELTYPHRLIDLAFLVSDSTDDTLAVLASEL DRIQKRPDQIPFHSATVIEKDFGFKLSQNVVEERHSFEAQGPRR KAMGRARNYLLITALKPEHSWVYWRDVIDVDSPGTGILEDFIA HNRDILVPNIWFHRYRDGVIEGRFDYNSWVESDKGRKLANS LDKDVVLAEGYKQYDTGRTYMAKMGDWRENKDVLELELDGI GGVNILVKADVHRSGINFPCYAFEN- AETEGFAKMAKRAGYEVYGLPNYVWWHIDTEKGGNA SEQ ID NO : 56
MNN9 alternativa2 estExt_GeneWisePlus.C_400029	MMPRHHSSGFSNGYPADTTFEI SPHRFQPRATLPPHRRKRKRTAIR SEQ ID NO : 57	<u>VGIADVVLVLVLF</u> SEQ ID NO : 58	QPRSVASLISLGLSGYDDLKLETVRYDLSNVQGTARGWER EERILLCVPLRDAEQHLPMTFFSHLKNFTYPHNLDLAFVSDS KDHTLESLTEHLEAIQADDPKQPYGEISIEKDFGQKVNQDV ESRHGFAAQASRRKLMAQARNWLLSAALRPYHSWVYWRDV DVETAPFTILEDLMRHNKDVIVPNVWRPLPDWLGGEQPYDL NSWQSEETALALADTLDEDAVIVEGYAEYATWRPHLAYLRD

Homólogo a	Citoplásmico	Transmembrana	Luminal
			PYGDPDMEIDGVGGVSILAKAKVFRAGVHFPFSFEKHA TEGFGKMAKRMHFSVVGLPHYTIWHLIEPSVDDIKHMEEME RERIAREKEEEEEERKKKEAQIKEEFGDANSQWEQDKQQMQDL KLQDRGGDKKEAAAAGVNGGAAAKAAGAMEGQKN SEQ ID NO : 59
MNN10 fgenes5_pg.C_scaffold_5000342	MSLSRSPSPVPGGGWSSPGLNINS GRSSPSNAAGSSVSWESAKMRKQG ANGYPSTQNQGFFTRHMRRI SSSLPRFAAGPGNTYAEREKYERG GHSPHAGGGRLRAFLARIGRRLKWR SEQ ID NO : 60	<u>ILLPLIIICTIVAYY</u> SEQ ID NO : 61	GTHEAPGFVHWRRISMGGGGEKVFILGANVGGGVMEWK GAREWAIERDSVRNKRKYATRWGYDLEIVDMKTKKRYAHE WRSEWEKVDIFIRAAMRKYPKAEWFWWLDLNTYVMEPSYL QRHLFNHLDRHVYRDINVFNPLNITHPPTEEYLDARESPVGD GNINSVNLMLTQDCSGFNLSFFIRSAWTEQLLDIWWDPVLI EQKHMEWEHKEQDALEQLIRTQPWIRQHTGFLPQRLINSFPP AACADESGLNNTRIHYNEKDRDFVNNMAGCEWGRDCWGE MYHYREFSYWLNRRNPWELFKEEIVAVIWKLTGQRVKL SEQ ID NO : 62
MNN10 alternativa estExt_GeneWisePlus.C_150339	MHFAYPSRKSSNPPFRPRSTRLP LRRSRIKT SEQ ID NO : 63	<u>IGLVFLVLATLWFF</u> SEQ ID NO : 64	SNPRVPRPDPERVPSGRPPVVLVTVIDPTQYPNAYLKTIKENR EQYAAKHGYEAFIVKAYDYDTQGAPQSWSKLMAMRHALTK FPECRFVWYLDQDIAIMDMMSKSLLEEQLNRRQKLESLEMIKNYP VVPPDSIIKTFSHLRPDEVLDIVSQDSSGLVAGSVVVRNSQWS

Homólogo a	Citoplásmico	Transmembrana	Luminal
			KFLLETWMDPLIRSYNFQKAERHALEHIVQWHPHTILSKLALVP QRTLGPYTRTDQGDIAQDGDVVMF- GCTKSGEQSCETVSASYQKWSSSL SEQ ID NO : 65
MNS1 fgenes1_pm.C_scaffold_3000175	MIRDPFGIHSKNFAKATALRAARDIK EAATQAGANALEMSFSLPKHVPDFG DPSRALEDRAWAALLPMYKDKP- YAYAPSMRLRPWWRRRK SEQ ID NO : 66	<u>VLGMIAAAVVMF-</u> <u>VLIVTGFF</u> SEQ ID NO : 67	SSGQTEEAKKKASGSAFSLGLSQERGGVDWDERRKSVVEA FEVWDIAERYAWGKDEFHPISKNGRNMAPKGLGWIIDS LDT MMLMNQTTRLQHAREWISTSLTWDQDQDVNTFETTIRMLGG LLSAHYLSTEFPELAPLTEDDEGAPGEDLILEKAKDLADRLLS AFESGIPYASVNIGEYKGPSSHSDNGASSTAEATTLQLEFKY LAKLTGEKNFWDKVEKVMVDDNQPEDGLVPIYIYATTGE FRGNIRLGSRGDSYYEYLIKQYLQTNKQEPYIYEEWDEALA GVRKHLVTYTEPSEFTIAERPDLGLEHPMSPKMDHLVCFMPTGT IALAAATGGLTEAEARKLSTWNKKKDDDDMLARELMHTCWG MYKYMKTGLAPEIMYFNIPNPPPESSAPHQAPAAAFDEDPHAE WRKDFVVHSNDVHNLQRPETVESLFYMWWRITGDDVKYREWVG WDMFKSFVNYTAVEDQGGFTSLLDANSIPPTPKDNMESFWL AETLKYMVYLLFSPNDVPLPLHKIVLNTEAHPFPRFDMGMLFSTG WKRKPRDGSAKKATTAATTDAE

Homólogo a	Citoplásmico	Transmembrana	Luminal
MNS1 alternativa estExt_fgenes1_pm.C_80182	MARRRYR SEQ ID NO : 69	<u>LFMICA</u> AIVILFLLR SEQ ID NO : 70	SEQ ID NO : 68 VSQNTWDDSAHYATLRHPPASNPAAAGGESPLKPAAKPEHEH EHENGYAPESKPKPQSEPKESKPAPEHAAGGQKSQKGKPSYE DDEETGKNPPKSAVIPSDTRLPPDNKVHWRPVKEHFPVPSES ISLPTGKPLKVPRVQHEFGVESPEAKSRRVARQERVGKEIERA WSGYKKFAWMHDELSPVSAKHRDPFCGWAATLVDSLDTLW IAGLKEQFDEAARAVEQIDFTTTPRNNIPVFETIRYLGGLGA FDVSGGHGGYPMLLTKAVELAEILMGIFDTPNRMPIIYQW QPEYASQPHRAGSVGIAELGTLSMFEFTRLAQLTSQYKYDDAV DRITDALIELQKQGTIPGLFPENLDASGCNHTATALRSSLSEA AQKQMEDLSDSNKPENYRPGKNSKADPQTVEKQPAKKQNEPV EKAKQVPTQQTAKRGKPPFGANGFTANWDCVPQGLVVGGY GFQQY HMGGGQDSAYEYFPKEYLLLGLESKYQKLIVDAVEAINEW LLIRPMTDGDWDILFPKAVSTAGNPSQDLVATFEVTHLTCTFIG GMYGLGGKIFGREKDDLETAKRLTDG- VWAYQSTVSGIMPEGSQVLACPTLEKCDFN ETLWWEKLDPKDWDRDKQVADDDKDKATVGEALKETANSH

Homólogo a	Citoplásmico	Transmembrana	Luminal
			DAAGGSAVHKRAAVPLPKGADDDVGSSELPQSLKDKIGFK NGEQKKTGSSVGIQRDPDAPVDSV - EAHRLPPQPEEEQVILPDKPQTHEEFVKQRIAE MGFAPGVVHIQSRQYILRPEAIESVWYMYRITGDPWMEKGW KMFEATRATRTEIANSADDVNSEEPGLKDEMESFWLAETLK YYYLLFSEPSVISLDEWVNLNTEAHPFKRPGGSGVIGHSI SEQ ID NO : 71
MNS1 alternativa2 estExt_GeneWisePlus.C_120298	MLNQLQGRVPRRY SEQ ID NO : 72	<u>IALVAFAEFVAELLW</u> SEQ ID NO : 73	SGYDFVPRATVGRFKYVPSSYDWSKAKVYYPVKDMKTLPO GTPVTFPRLQLRNQSEAQDDTTKARKQAVKDAFVKSWEAYK TYAWTKDQLQLSLSGKETFSGWSAQLVDAALDTLWIMDLKD DFFLAVKEVAVIDWSKTKDNKVINLFEVTIRYLGGLIAAYDLS QEPVLRAKAIELGDTLIAFTDTPNRLPSHWLDYSKAKKGTQR ADDMSGAAAGGTLCMEFTRLSQITGDPKYYDATERIKQFFYR FQNETTLPGMWPMVMNYREETMVESRYSMGGSADSLIEYLV KMPALLGGLDPQYPEMAIRALDTARDNLLFRPMTEKGDNILA LGNALVDHGNVQRITEMQHLTCFAGGMYAMAGKLFKRDDY VDLGSRISGGCVWAYDSFPSPGIMPESADMAACAKLDGPCPYD EVKAPVDPDGRRRPHGFHVKSRHYLLRPEAIESVFYMWIRITG

Homólogo a	Citoplásmico	Transmembrana	Luminal
			DQVWRDTAWRMWENIVREATEHFAIVEDVTRTASKLTNN YLLQTFWLAETLKYFYLIFFDDESAIDLKWWFNTEAHPFKRP AV SEQ ID NO : 74
MNS1 alternativa3 estExt_GeneWisePlus.C_160228	MLVVGRPRLVRNS SEQ ID NO : 75	<u>IILTLAILSIWHLGLL</u> SEQ ID NO : 76	SRTPTSASALVSASVSASSEWSRLRLMNRGAPLTPYPDSNSS FDWSAIPFRYPHNTHLPPRHKQPPLPRIQHRFGPESPAAAKE RIKRLKAVKQVFLRAWQAYKGYAWKQDALLPISGGGREQFS GWAA TLVDALDTL WIMGLREEFDEAVAAVAEIDFGSSTSRV NIFETNIRYLGLLAAAYDLSGREVLLKKAVELGDLIYAGFNTE NGMPVDFLNFYSAKSGEGLVVESSVVSASPGTLSLELAHLSQ VTGDDKYYSAVSQVMDVIFYQGQNKTRLPGVWPIDVNMRAK DVVSGSRFTLGGCADSLIEYLPKMHQLLGGGEPKYETMSRTF LQAADRHVFRPMLPGAEEEDVLMPGNVNVNVEDSDGEAVLDPE TEHLACFVGGMFGLAGRLFSRPDDVETGVRLTNGCVVYAYRA FPTGMMPERLDLAPCRDRSSRCPWDEEHWLEERAKRPEWEP HLPRGFTSAKDPRYLLRPEAIESVFYSYRITGRQEFQTAAWDM FTAVEKGTRTQFANAALVDVTRAADEL PQEDYMESFWLAET LKYFYLMFTTPDIISLDDYYVLNTEAHPFKLVG

Homólogo a	Citoplásmico	Transmembrana	Luminal
MNS1 alternativa4 e_gw1.13.279.1	-	<u>MVMLVAIALA-</u> <u>WLGCSLL</u> SEQ ID NO : 78	SEQ ID NO : 77 RPVDAMRADYLAQLRQETVDMFYHGYSNYMEHAFPEDEL PISCTPLTRDRDNPGISLNDALGNYSLLTLDLSLTLAILAGGPQ NGPYTGPQALSDFQDGVAEFVRHYGDGRSGPSGAGIRARGFD LDSKVQVFETVIRGVGGLLSAHLFAIGELPITGYVPRPEGVAG DDPLELAPIWPNGFRYDGGQLRLALDLSERLLPAFYTPTGIP YPRVNLRSIGIPFYVNSPLHQNLGEAVEEQSGRPEITETCSAGA GSLVLEFTVLSRLTGDARFEQA AKRAFWEVWHRRSEIGLIGN GIDAERGLWIGHAGHAGMDSFFEYALKSHILLS GLGMPNASTSRRQSTTSWLDPNSLHPPLPEMHTSDAFLQAW HQAHASVKRYLITDRSHFPYYSNNHRATGQPYAMWIDSLGA FYPGLLALAGEVEEAIEANLVYTAL- TRYSA LPERWSVREGNVEAGIGWWPGRPEFI ESTYHIYRATRDPWYLHV GEMVLRDIRRRCYAECCGWAGLQD VQTGEKQDRMESEFFLGETAKYMYLLFPDHPHPLNKLDAA YVF TTEGHPLIIPKSKRGSGSHNRQDRA- KAKKSRDVAVYTY YDEFTNSCPAPRPPSEHH LIGSATAARPD LFSVSRFTDLIRTPNVHGLEKVMRDKKKGR

Homólogo a	Citoplásmico	Transmembrana	Luminal
			<p>VVRYRATSNHTIPWTLPPAMLPENGTCAAPPERIISLIEFPAN</p> <p>DITSGITSRFGNHLNLSWQTHLGPTVN-</p> <p>LEGLRLQLEQVSDPATGEDKWRITHIG</p> <p>NTQLGRHETVFFHAEHVRHLKDEVFSCRRRRDAVEIELLVDK</p> <p>PSDTNNNTLASSDDDDVVVDAKAEEDGMLADDDGDTLNA</p> <p>ETLSSNSLFQSLLRASVSVFEPVYT-</p> <p>IPESDPSAGTAKVYSFDIATSTGPGAYPMPSI</p> <p>SDTPIPGNPFYFNFRNPASNFPWSTVFLAGQACEGPLPASAPRE</p> <p>HQVIVMLRGCSFSRKLNDNIPFSPHDRALQLVVVLDPEPPPP</p> <p>PPPPANDRRRDVTRPLLDTEQTTPKG-</p> <p>KRLHGIPMVLVRAARGDYELFGHAIGVG</p> <p>MRRKYRVESQGLVVENAVVL</p> <p>SEQ ID NO : 79</p>
<p>VAN1</p> <p>estExt_GeneWisePlus.C_400029</p>	<p>MMPRHHSSGFSNGYPRADTTFEISPFR</p> <p>FQPRATLPPHRKRKRRTAIR</p> <p>SEQ ID NO : 80</p>	<p><u>VGI</u>AVVVVLVLVLF</p> <p>SEQ ID NO : 81</p>	<p>QPRSVASLISLILSGYDDLKLETVRYDLSNVQGTARGWER</p> <p>EERILLCVPLRDAEQHLPMFFSHLKNFTYPHNLDLAFVSDS</p> <p>KDHTLESLEHLEAIQADPPKQPYGEISIIIEKDFGQKVNQDV</p> <p>ESRHGFAAQASRRKLMQAQARNWLLSAALRPYHSHVWVYWRDV</p> <p>DVETAPFTILEDLMRHNKDVIVPNVWRPLPDWLGGGEQPYDL</p>

Homólogo a	Citoplásmico	Transmembrana	Luminal
			NSWQESETALALADTLDEDAVIVEGYAEYATWRPHLAYLRD PYGDPDMEIDGVGGVSILAKAKVFRAGVHFPAFSFEKHAE TEGFGKMAKRMHFSVVGLPHYTTIWHLIEPSVDDIKHMEEME RERIAREKEEEEEERKKKEAQIKEEFGDANSQWEQDKQQMQDDL KLQDRGGDKAAAAAGVNGQAAAAKAAAGAMEGQKN SEQ ID NO : 82
VAN1 alternativa1 estExt_GeneWisePlus.C_230146	MLLPKGGLDWRSARAQIPPTR ALWNAVTRTR SEQ ID NO : 83	<u>FILLVGITGLILLW</u> SEQ ID NO : 84	RGVSTASSEMQSYFCWGPAPPMEMSPNEHNRWNGHLQTPV IFNHHAPVEVNSSTIEHVLDLPINS- KQAVTKEERILILTPLKDAAPYLSKYF ELLAELTYPHRLIDLAFLVSDSTDDTLAVLASELDRIQKRPDQI PFHSATVIEKDFGFKLSQNVEERHSFEAQGPRRKAMGRARNY LLITALKPEHSWVYWRDVIDVDSTGILEDFIAHDRDILVPNI WFHRYRDGVDIEGRFDYNSWVESDKGRKLANSLDKDVVLA EGYKQYDTGRITYMAKMGDWRENKDVELELDGIGGVNILVK ADVHRSGINFPCYAFENQAETEGFAKMAKR- AGYEVYGLPNYVVVWHIDTEEKGGNA SEQ ID NO : 85
VAN1 alternativa2	MARPMGSVRLKKANPST	<u>LILGAVLCIFIIIFLV</u>	SPSSPASASRLSIVSAQHHLSPPTSPYQSPRSGAVQGPVPVTRY

Homólogo a	Citoplásmico	Transmembrana	Luminal
e_gw1.5.262.1	SEQ ID NO : 86	SEQ ID NO : 87	NLNKVVTVTSDPVRNQEHLILTPMARFYQEYWDNLLRLNYPH ELITLGFILPKTKEGNQATSMQKQIKTQNYGPEKDRFKSIII LRQDFDPAVVSDSESERHKLANKQARREVMAKARNSLFTT LGPSTSWVLWLDADITETAPTLIQDLASHDKPIIVANCFQKYY DPESKKMAERPYPDFNSWQDSETALKMAEQMGPDILLLEGYA EMATYRTLAYMSTPGGSKDLVVPLDGVGGTALLVKADVH RDGAMFPFPFAFYHLIESEGFAMAKRLGWQPYGLPNYKVYH YNE SEQ ID NO : 88
Outro01 estExt_GeneWisePlus.C_150339	MHFAYPSRKSSNPPFPRSTRLP- GLRRSRIKT SEQ ID NO : 89	<u>IGIVLFLVLATLWFF</u> SEQ ID NO : 90	SNPRVPRDPERVPSPGRPPVVLVTVIDPTQYPNAYLKTIKENR EQYAAKHGYEAFIVKAYDYDTQGAPQSWSKLMAMRHALTK FPECRFVWYLDQDIAIMDMSKSLEEQLLNQRQKLESMLIKNYP VVPPDSIIKTFSHLRPDEVDLIVSQDSSGLVAGSVVVRNSQWS KFLLETWMDPLIRSYNFQKAERHALEHIVQWHPHTILSKLALVP QRTLGPYTRTDQGDIAQDGDFFVVMF- GCTKSGEQSCETVSASYQKWSSSL SEQ ID NO : 91
Outro02	MSLSRSPSPVPGGGWSSPGLNINSGR	<u>ILLPLIIICITIVAYYG</u>	THEAPGFVHWRRRISMGGGGEKFKVILGANVGGGVMEWKG

Homólogo a	Citoplásmico	Transmembrana	Luminal
fgenes5_pg.C_scaffold_5000342	SSPSNAAGSSVSWEAKMRKQGAN GYPSFSTQNQGFFTRHMRRISSSLPRF AAGPGNTYAEREKYERGGHS- PHAGGGRLRAFLARIGRRLKWR SEQ ID NO : 92	SEQ ID NO : 93	AREWAIERDSVRNKRKYATRWGYDLEIVDMKTKKRYAHEW RESWEKVDFIRAAAMRKYPKAEWFWWLDLNTYVMEPSYSLQ RHLFNHLDRHVYRDINVFNPLNTHPTTEEYLDAAEARSPVGDG NINSVNLMLTQDCSGFNLGSFFIRRSAWTEQLLDIWWDPVLIE QKHMEWEHKEQDALEQLIRTQPWIRQHTGFLPQRLINSFPPA ACADESGLNNTRIHYNEKDRDFVNVNMAGCEWGRDCWGEM YHYREFSYWLNLRNPWELFKEEIVAVIWKLTGQRVKL SEQ ID NO : 94

As sequências descaracterizadas podem ser testadas para o uso como peptídeos de direcionamento expressando proteínas na via de glicosilação em uma célula hospedeira, em que uma das proteínas contém a sequência não caracterizada como o único peptídeo alvo, e medição dos glicanos produzidos tendo em vista a localização citoplasmática da biossíntese de glicanos (por exemplo, como no Schwientek JBC 1996 3398), ou por meio da expressão de uma proteína fluorescente repórter fundido com o peptídeo alvo, e analisando a localização da proteína no Golgi por imunofluorescência ou por meio do fracionamento das membranas citoplasmáticas do Golgi e medindo a localização da proteína.

O peptídeo de direcionamento pode incluir um domínio da haste. Em certas modalidades, o domínio de haste é de uma enzima N-acetilglucosaminiltransferase I ou uma enzima N-acetilglucosaminiltransferase II. Em certas modalidades, especialmente, no domínio da haste é de uma enzima N-acetilglucosaminiltransferase I humana ou uma enzima humana N-acetilglucosaminiltransferase II. A sequência correspondente ao domínio da haste de enzima N-acetilglucosaminiltransferase I humana é a SEQ ID NO : 34. A sequência correspondendo ao domínio humano resultam da enzima N-acetilglucosaminiltransferase II é 30 a 85 resíduos de SEQ ID NO : 20.

O peptídeo de direcionamento pode incluir um domínio transmembranar. Um "domínio transmembranar" refere-se a qualquer sequência de resíduos de aminoácidos que é termodinamicamente estável em uma membrana, como uma estrutura tridimensional. Em modalidades em que o peptídeo alvo também inclui um domínio de haste, o domínio transmembranar é Terminal N ao domínio da haste. Em certas modalidades, o domínio transmembranar é de uma enzima N-acetilglucosaminiltransferase I ou uma enzima N-acetilglucosaminiltransferase II. Em certas modalidades especialmente, o domínio transmembranar é de uma enzima N-acetilglucosaminiltransferase I humana ou uma enzima humana N-acetilglucosaminiltransferase II. A sequência correspondente ao domínio de transmembrana de enzima N-acetilglucosaminiltransferase I humana é resíduos 7-29 de SEQ ID NO : 1. A sequência correspondente ao domínio de transmembrana de en-

zima humana N-acetilglucosaminiltransferase II é 10 a 29 resíduos de SEQ ID NO : 20.

O peptídeo de direcionamento pode incluir um domínio citoplasmático. O termo "domínio citoplasmático" refere-se a uma sequência de aminoácidos que é termodinamicamente estável em um ambiente citoplasmático como uma estrutura tridimensional. Em modalidades em que o peptídeo alvo também inclui um domínio de haste, é o domínio citoplasmático do Terminal N ao domínio da haste. Em modalidades em que o peptídeo alvo inclui também um domínio transmembranar, o domínio citoplasmático é o Terminal N com o domínio transmembranar. Em certas modalidades, o domínio citoplasmático é de uma enzima N-acetilglucosaminiltransferase I ou uma enzima N-acetilglucosaminiltransferase II. Em certas modalidades especialmente, o domínio citoplasmático é de uma enzima N-acetilglucosaminiltransferase I humana ou uma enzima humana N-acetilglucosaminiltransferase II. A sequência correspondendo ao domínio citoplasmático de enzima N-acetilglucosaminiltransferase I humana é os 1 a 6 resíduos da SEQ ID NO : 1. A sequência correspondendo ao domínio citoplasmático de enzima humana N-acetilglucosaminiltransferase II é os 1 a 9 resíduos da SEQ ID NO : 20.

Em certas modalidades, a proteína recombinante contém um domínio catalítico de GnTII humano do Terminal N a um domínio catalítico de GnTI humano com uma sequência separadora contendo a sequência da haste do domínio GnTI humano entre os domínios catalíticos. Nesta modalidade, a proteína recombinante também inclui um peptídeo alvo do Terminal N com o domínio catalítico com GnTII citoplasmático, transmembranar, e as hastes dos domínios de GnTII humano. A sequência da proteína recombinante na presente modalidade é de pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica à SEQ ID NO : 95, e a sequência de um cDNA que codifica o possível proteína recombinante desta modalidade é a SEQ ID NO : 96.

Em outras modalidades, a proteína recombinante contém um domínio catalítico de GnTII humano do Terminal N a um domínio catalítico

de GnTI humano com uma sequência separadora. A sequência espaçadora pode incluir, sem limitação, uma sequência que é pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica à SEQ ID NOs : 118, 120, 122, ou 124. Nesta modalidade, a proteína recombinante também inclui um peptídeo alvo com o domínio catalítico com GnTII citoplasmático do Terminal N, transmembrar, e as hastes dos domínios de GnTII humano. Dessa maneira, em certas modalidades, a sequência da proteína recombinante é de pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica a uma sequência selecionada a partir das SEQ ID NOs : 119, 121, 123, e 125. Em certas modalidades, a sequência de uma possível DNAC que codifica a proteína recombinante de SEQ ID NO : 119 é SEQ ID NO : 141. Em outras modalidades, a sequência de uma possível DNAC que codifica a proteína recombinante de SEQ ID NO : 121 é SEQ ID NO : 139. Em ainda outras modalidades, a sequência de uma possível DNAC que codifica a proteína recombinante de SEQ ID NO : 123 é SEQ ID NO : 143. Em outras modalidades, a sequência de um cDNA que codifica a possível proteína recombinante de SEQ ID NO : 125 é SEQ ID NO : 145.

Produção de Proteínas Recombinantes da Invenção

Um outro aspecto da presente invenção inclui os polinucleotídeos isolados que codificam para as proteínas recombinantes da presente invenção. Tal como usado na presente invenção, os termos "polinucleotídeo", "sequência de ácido nucleico", "sequência de ácidos nucleicos", e as variações das mesmas devem ser genéricas para polidesoxirribonucleotídeos (contendo 2-desóxi-D-ribose), polirribonucleotídeos (contendo a D-ribose), a qualquer outro tipo de polinucleotídeo que seja um N-glicosídeo de uma base purina ou pirimidina e a outros polímeros contendo as estruturas principais não nucleotídicas, desde que os polímeros contêm nucleobases em uma configuração que permite o emparelhamento de bases e empilhamento de bases, tal como encontradas no DNA e RNA. Dessa maneira, es-

tes termos incluem os tipos conhecidos de modificações de sequências de ácido nucleico, por exemplo, a substituição de um ou mais dos nucleotídeos que ocorrem de uma forma natural com um análogo, modificações internucleotídicas, tais como, por exemplo, aqueles com ligações não carregadas (por exemplo, fosfonatos de metila, fosfotriésteres, fosforamidatos, carbamatos, etc), ligações com carga negativa (por exemplo, fosforotioatos, fosforoditioatos, etc), e com ligações com carga positiva (por exemplo, aminoalquilfosforamidatos, aminoalquilfosfotriésteres), aqueles contendo as porções pendentes, tais como, por exemplo, proteínas (incluindo nucleases, toxinas, anticorpos, peptídeos de sinal, poli-L-lisina, etc), aqueles com intercaladores (por exemplo, acridina, psoraleno, etc), e os que contêm quelantes (por exemplo, metais, metais radioativos, boro, metais oxidativos, etc.) Tal como usado na presente invenção, os símbolos para nucleotídeos e polinucleotídeos são os recomendados pela Comissão IUPAC-IUB da Nomenclatura Bioquímica (Biochem 9 : 4022, 1970).

As sequências dos polinucleotídeos isolados são preparadas por meio de qualquer método adequado conhecido por aqueles que são versados na técnica, incluindo, por exemplo, síntese química direta ou clonagem. Para a síntese química direta, a formação de um polímero de ácidos nucleicos envolve tipicamente a adição sequencial de monômeros de nucleotídeos 3'- e 5'-bloqueados no terminal do grupo 5'-hidroxila de uma cadeia nucleotídica em crescimento, onde cada adição é efetuada por meio do ataque nucleofílico do terminal do grupo 5'-hidroxila da cadeia crescente na posição 3' do monômero adicionado, o que é tipicamente um derivado de fósforo, tal como um fosfotriéster, fosforamidita, ou semelhantes. Tal metodologia é conhecida por aqueles que são versados na técnica e está descrito nos textos pertinentes e na literatura [por exemplo, em Matteucci et al, (1980) Tetrahedron Lett 21 : 719 a 722 ; Pat. U.S. N ° s 4.500.707 ; 5.436.327 e 5.700.637]. Além disso, as sequências desejadas podem ser isoladas a partir de fontes naturais, por meio da divisão usando as enzimas de restrição de DNA apropriadas, separando os fragmentos, utilizando electroforese em gel, e em seguida, recuperando a sequência de ácido nucleico desejado a partir do gel

por meio de técnicas conhecidas por meio daqueles que são versados na técnica, tais como a utilização de reações em cadeia de polimerase (PCR, por exemplo, Patente U.S. No. 4.683.195.).

5 Cada polinucleotídeo da presente invenção pode ser incorporado em um vetor de expressão. O termo "vetor de expressão" ou "vetor" refere-se a um composto /e/ou composição que transduz, transforma, ou infecta uma célula hospedeira, fazendo assim com que a célula com a finalidade de expressar os ácidos nucleicos /e/ou proteínas não sejam os nativos à célula, ou em uma forma não nativa para a célula. Um "vetor de expressão" contém
10 uma sequência de ácidos nucleicos (DNA ou RNA, normalmente) a ser expressa por meio da célula hospedeira. Opcionalmente, o vetor de expressão compreende ainda os materiais para auxiliar na obtenção de entrada do ácido nucleico em uma a célula hospedeira, tal como um vírus, lipossomas, revestimentos de proteínas, ou semelhantes. Os vetores de expressão con-
15 templados para utilização na presente invenção incluem aqueles em que uma sequência de ácido nucleico pode ser inserida, juntamente com todos os elementos necessários determinados ou operacionais. Além disso, o vetor de expressão deve ser um que possa ser transferido para uma célula hospedeira e aí replicado. Certos vetores de expressão são plasmídeos, particularmente aqueles com locais de restrição que tenham sido bem docu-
20 mentados e que contêm os elementos operacionais determinados ou necessário para a transcrição do ácido nucleico. Tais plasmídeos, bem como outros vetores de expressão são bem conhecidos por meio daqueles que são versados na técnica.

25 A incorporação dos polinucleotídeos individuais pode ser realizada através de métodos conhecidos que incluem, por exemplo, a utilização de enzimas de restrição (por exemplo, BamHI, EcoRI, HhaI, XhoI, XmaI, e assim por diante) que clivam em locais específicos do vetor de expressão, por exemplo, o plasmídeo. A enzima de restrição produz as extremidades de
30 cadeia simples que podem ser recozidas a um polinucleotídeo tendo ou sintetizado para ter um terminal com uma sequência complementar às extremidades do vetor de expressão clivado. O recozimento é realizado usando

uma enzima adequada, por exemplo, DNA-ligase. Como será apreciado por meio daqueles que são versados na técnica, tanto o vetor de expressão e o polinucleotídeo pretendido são muitas vezes clivado com a mesma enzima de restrição, assegurando assim que as extremidades do vetor de expressão e as extremidades do polinucleotídeo sejam complementares umas as outras. Além disso, podem ser utilizados ligantes de DNA para facilitar a ligação de sequências de ácidos nucleicos em um vetor de expressão.

Uma série de polinucleotídeos individuais também podem ser combinada, utilizando métodos que são conhecidos por meio daqueles que são versados na técnica (por exemplo, Patente U.S. No. 4.683.195).

Por exemplo, cada um dos polinucleotídeos desejados podem ser gerados, inicialmente, em uma PCR separada. Depois disso, os iniciadores específicos foram concebidos de tal modo que as extremidades dos produtos de PCR contêm sequências complementares. Quando os produtos de PCR são misturados, desnaturados, e reanelados, as cepas possuindo as sequências correspondentes nas suas extremidades 3' se sobrepõem e podem atuar como iniciadores para o outro. A extensão desta sobreposição por meio da polimerase do DNA produz uma molécula em que as sequências originais são "divididas" em conjunto. Desta maneira, uma série de polinucleotídeos individuais pode ser "dividida" em conjunto e subsequentemente transduzida para uma célula hospedeira em simultâneo. Dessa maneira, a expressão de cada uma da pluralidade de polinucleotídeos é afetada.

Os polinucleotídeos individuais, ou polinucleotídeos "divididos", em seguida, são incorporados em um vetor de expressão. A presente invenção não é limitada no que diz respeito ao processo pelo qual o polinucleotídeo é incorporado no vetor de expressão. Aquelas pessoas que são versadas na técnica estão familiarizados com as etapas necessárias para a incorporação de um polinucleotídeo em um vetor de expressão. Um vetor de expressão típico contém o polinucleotídeo pretendido precedido por meio de uma ou mais regiões reguladoras, juntamente com um local de ligação ao ribossoma, por exemplo, uma sequência de nucleotídeos que é de 3 a 9 nucleotídeos de comprimento, 3 a 11 nucleotídeos localizados a montante do

códon de iniciação na *E. coli*. Vide Shine e Dalgarno (1975) *Nature* 254 (5495) : 34 a 38 e Steitz (1979) *Regulamento Biológica e Desenvolvimento* (ed. Goldberger, RF), 1 : 349 a 399 (Plenum, New York).

O termo como usado na presente invenção "operativamente ligado" refere-se a uma configuração na qual uma sequência de controle é colocada em uma posição adequada em relação à sequência de codificação de uma sequência de DNA ou polinucleotídeo de tal modo que a sequência de controle dirige a expressão de um polipeptídeo.

As regiões reguladoras incluem, por exemplo, as regiões que contêm um promotor e um operador. Um promotor está operativamente ligado ao polinucleotídeo desejado ou parte de um polinucleotídeo que codifica um polipeptídeo, iniciando desse modo a transcrição do polinucleotídeo, ou uma porção do polinucleotídeo que codifica para um polipeptídeo, por meio de uma enzima RNA polimerase. O operador é uma sequência de ácidos nucleicos adjacentes ao promotor, que contém um domínio da proteína de ligação de uma proteína repressora se pode ligar. Na ausência de uma proteína repressora, a transcrição é iniciada por meio do promotor. Quando presente, a proteína do repressor específico para o domínio da proteína de ligação do operador liga-se ao operador, inibindo desse modo a transcrição. Deste modo, o controle da transcrição é realizado, com base nas regiões reguladoras particulares utilizadas e a presença ou ausência da proteína do repressor correspondente. Exemplos incluem os promotores de lactose (alterações LAD da conformação da proteína do repressor, quando contatado com lactose, evitando assim a proteína repressora Lad de se ligar ao operador) e promotores do triptofano (quando complexado com triptofano, trpR proteína repressora tem uma conformação que se liga ao operador, na ausência de triptofano, a proteína do repressor trpR tem uma conformação que não se liga ao operador). Outro exemplo é o promotor de tac (vide Bóer et al., (1983) *Proc Natl Acad Sci EUA* 80 (1) : 21 a 25). Como será apreciado por meio daqueles que são versados na técnica, estas e outras regiões reguladoras podem ser utilizadas na presente invenção, e a presente invenção não está limitada a este respeito.

Os exemplos de determinados promotores para a ligação com os polinucleotídeos isolados que codificam para as proteínas recombinantes da presente invenção, incluem os promotores dos seguintes genes : GPDA, cbh1, TAKA amilase de *Aspergillus oryzae*, prbateladainase aspártica de *Rhizomucor miehei*, *Aspergillus niger* neutro alfa-amilase de ácido *Aspergillus niger* estável de alfa-amilase, glucoamilase de *Aspergillus niger* (glaA), glaA *Aspergillus awamori*, lipase de *Rhizomucor miehei*, protease alcalina de *Aspergillus oryzae*, *Aspergillus oryzae*, triose fosfato isomerase, acetamidase de *Aspergillus nidulans*, acetamidase de *Aspergillus oryzae*, *Fusarium oxysporum*, a protease do tipo tripsina, endo α -L-arabinase fúngica (ABNA), α -L-arabinofuranosidase A fúngica (ABFA), α -L-arabinofuranosidase B fúngica (abfB), xilanase fúngica (xlnA), fitase fúngica, fúngica ATP-sintetase, subunidade 9 fúngica (PBR), isomerase de fosfato de triose fúngica (tpi), desidrogenase de álcool de fungos (ADHA), fúngica α -amilase (amy), amilaglicosidase fúngica (glaA), acetamidase fúngica (amdS), desidrogenase de gliceraldeído-3-fosfato-fúngica (gpd), álcool desidrogenase de levedura, lactase de fermento, levedura de 3-fosfoglicerato quinase, triosefosfato-isomerase de leveduras, SpO2 α -amilase bacteriana, e ASN bacteriana. Em certas modalidades, os polinucleotídeos isolados que codificam para as proteínas recombinantes da presente invenção são operativamente ligados a um promotor constitutivo. Em outras modalidades, os polinucleotídeos isolados que codificam para as proteínas recombinantes da presente invenção são operativamente ligados a um promotor indutível. Em certas modalidades preferidas, o promotor indutível é de um gene cbh1.

Embora qualquer vetor de expressão adequado pode ser utilizado com a finalidade de incorporar as sequências desejadas, os vetores de expressão prontamente disponíveis incluem, sem limitação : os plasmídeos, tais como pSCIOI, pBR322, pBBRIMCS-3, pur, pEX, pMRIOO, pCR4, pBAD24, pUC19 ; bacteriófagos, tais como M1 3 e fagos λ fago. Claro que, tais vetores de expressão podem ser apenas adequados para células hospedeiras particulares. Aqueles que são versados na técnica, no entanto, podem determinar facilmente através de experimentação de rotina, se qualquer

vetor de expressão particular é adequado para qualquer célula hospedeira dada. Por exemplo, o vetor de expressão pode ser introduzido na célula hospedeira, a qual é, em seguida, monitorizada para a viabilidade e expressão das sequências contidas no vetor. Além disso, pode ser feita referência aos textos e publicações relevantes que descrevem os vetores de expressão e a sua adequação a qualquer célula hospedeira particular.

Um outro aspecto da presente invenção inclui as células hospedeiras contendo vetores de expressão contendo os polinucleotídeos isolados que codificam para as proteínas recombinantes da presente invenção. O termo "célula hospedeira", tal como usado na presente invenção refere-se a uma célula biológica viva que podem ser transformada através da inserção de DNA recombinante ou RNA. Tal DNA ou RNA recombinante pode estar em um vetor de expressão. Dessa maneira, uma célula hospedeira, tal como descrita na presente invenção pode ser um organismo procariótico (por exemplo, um organismo do reino eubactérias) ou uma célula eucariótica. Como será apreciado por meio daquele que é versado na técnica, uma célula procariótica carece de um núcleo ligado à membrana, enquanto que uma célula eucariótica tem um núcleo ligado à membrana. Em certas modalidades, as células hospedeiras utilizadas para a produção das proteínas recombinantes da presente invenção são células de fungos, tais como leveduras ou fungos filamentosas. Em outras modalidades, as células hospedeiras são células de mamíferos. Tais células podem humanas ou não humanas.

Um outro aspecto da presente invenção inclui os métodos para a produção das proteínas recombinantes da presente invenção. O método inclui as etapas de introdução de um polinucleotídeo isolado que codifica a proteína recombinante em uma célula hospedeira e cultivar a célula hospedeira de tal modo que a proteína recombinante é expressa. O método pode também incluir uma etapa de purificação da proteína recombinante a partir da célula hospedeira.

Os métodos de produção das proteínas recombinantes da presente invenção podem incluir a introdução ou transferência de vetores de expressão contendo os polinucleotídeos recombinantes da presente inven-

ção dentro da célula hospedeira. Tais métodos para transferir os vetores de expressão em células hospedeiras são bem conhecidos por meio daqueles que são versados na técnica. Por exemplo, um método para a transformação de *E. coli* com um vetor de expressão envolve um tratamento com cloreto de cálcio, onde o vetor de expressão é introduzido, através de um precipitado de cálcio. Outros sais, por exemplo, fosfato de cálcio, podem também ser utilizados na sequência de um procedimento semelhante. Além disso, a electroporação (isto é, a aplicação de corrente para aumentar a permeabilidade das células com sequências de ácido nucleico) pode ser usada para transfectar a célula hospedeira. Além disso, a microinjeção de sequências de ácido nucleico proporciona a capacidade de transfectar as células hospedeiras. Outros meios, tais como os complexos de lípidos, lipossomas, dendrímeros, e podem também ser empregues. Aqueles que são versados na técnica podem transfectar uma célula hospedeira com uma sequência desejada com estes ou outros métodos.

O vetor pode ser um vetor de replicação autônoma, ou seja, um vetor que existe como uma entidade extracromossômica, cuja replicação é independente da replicação cromossômica, por exemplo, um plasmídeo, um elemento extracromossômico, um minicromossoma, ou um cromossoma artificial. O vetor pode conter qualquer meio para assegurar a autorreplicação. De uma maneira alternativa, o vetor pode ser um que, quando introduzido no hospedeiro, é integrado no genoma e replicado juntamente com o (s) cromossoma (s) em que foi integrado. Além disso, pode ser utilizado um único vetor ou plasmídeo ou dois ou mais vetores ou plasmídeos que em conjunto contêm o DNA total a ser introduzido no genoma do hospedeiro, ou uma transposona pode ser usada.

Os vetores podem conter um ou mais marcadores selecionáveis que permitam a seleção fácil dos hospedeiros transformados. Um marcador de seleção é um gene, o produto do qual fornece, por exemplo, resistência biocida ou viral, resistência a metais pesados, prototrofia a auxotrofos, e similares. A seleção de células bacterianas pode ser baseada na resistência antimicrobiana que tenha sido conferida pelos genes tais como os genes

amp, gpt, neo e hyg.

Os marcadores adequados para hospedeiros de levedura são, por exemplo, ADE2, HIS3, LEU2, LIS2, MET3, TRP1, e URA3. Os marcadores selecionáveis para utilização em um hospedeiro fúngico filamentoso incluem, mas não se limitam a, amdS (acetamidase), argB (ornitina carbamoyl-transferase), bar (fosfotricina-acetiltransferase), hph (higromicina fosfo-transferase), niaD (nitrato reductase), pyrG (orotidina 5 descarboxilase 5'-fosfato), SC (sulfato adeniltransferase), trpC (antranilato e sintase), assim como os seus equivalentes. Determinada para o uso em *Aspergillus* são os genes amdS e pyrG de *Aspergillus nidulans* ou *Aspergillus oryzae* e o gene bar de *Streptomyces hygroscopicus*. Certa para o uso em *Trichoderma* são bar, pyr4 e amdS.

Os vetores podem conter um elemento (s) que permite a integração do vetor no genoma do hospedeiro ou a replicação autônoma do vetor na célula independente do genoma.

Para a integração no genoma do hospedeiro, o vetor pode contar com a sequência do gene ou por meio de qualquer outro elemento do vetor para a integração do vetor no genoma por meio da recombinação homóloga ou não homóloga. De uma maneira alternativa, o vetor pode conter sequências de nucleotídeos adicionais para dirigir a integração por recombinação homóloga no genoma do hospedeiro. As sequências de nucleotídeos adicionais permitem que o vetor seja integrado no genoma do hospedeiro a uma localização precisa no (s) cromossoma (s). Com a finalidade de aumentar a probabilidade de integração em um local preciso, os elementos integradores podem conter um número suficiente de ácidos nucleicos, tais como 100 a 10,000 pares de bases, de preferência de 400 a 10,000 pares de bases, e mais preferencialmente de 800 a 10,000 pares de bases, que são altamente homólogos com a sequência alvo correspondente com a finalidade de aumentar a probabilidade de recombinação homóloga. Os elementos integrador pode ser qualquer sequência que é homóloga com a sequência alvo no genoma do hospedeiro. Além disso, os elementos integradores podem ser não codificados ou que codificam para as sequências de nucleotídeos.

Por outro lado, o vetor pode ser integrado no genoma do hospedeiro por meio da recombinação não homóloga.

Para a replicação autônoma, o vetor pode ainda compreender uma origem de replicação que permite ao vetor replicar de forma autônoma no hospedeiro em causa. A origem de replicação pode ser qualquer plasmídeo replicador de mediar a replicação autônoma que funciona em uma célula. O termo "origem de replicação" ou "replicador plasmídeo" é definido na presente invenção como uma sequência que permite que um plasmídeo ou vetor replicar-se in vivo. Exemplos de origens de replicação para uso em um hospedeiro de levedura são a origem de 2 micra de replicação, ARS1, ARS4, a combinação de ARS1 e CEN3, e a combinação de ARS4 e CEN6. Exemplos de origens de replicação úteis em uma célula fúngica filamentosa são AMA1 e ANS1 (Gems et al, 1991, Cullen et al, 1987, WO 00/24883). O isolamento do gene de AMA1 e o construto de plasmídeos ou vetores que compreendem o gene pode ser realizada de acordo com os métodos descritos no documento WO 00/24883.

Para outros hospedeiros, os procedimentos de transformação pode ser encontrada, por exemplo, em Jeremias D. Read, et al., Applied and Environmental Microbiology, agosto de 2007, p. 5088-5096, por Kluyveromyces, em Osvaldo Delgado, et al., FEMS Microbiology Letters 132, 1995, 23-26, para Zymomonas, nos EUA 7.501.275 de Pichia stipitis, e em WO 2008/040387 para o Clostridium.

Mais do que uma cópia de um gene pode ser inserido dentro do hospedeiro com a finalidade de aumentar a produção do produto do gene. Um aumento no número de cópias do gene pode ser obtida mediante a integração de pelo menos uma cópia adicional do gene no genoma do hospedeiro ou incluindo um gene marcador selecionável amplificável, com a sequência de nucleotídeos em que as células contendo as cópias amplificadas do gene marcador selecionável, e assim as cópias adicionais do gene podem ser selecionadas cultivando as células na presença do agente de seleção adequado.

Os procedimentos empregados para enlaçar os elementos aci-

ma descritos com a finalidade de construir os vetores de expressão recombinantes da presente invenção são bem conhecidos por meio daqueles que são versados na técnica (vide, por exemplo, Sambrook et al., 1989, supra).

5 A célula hospedeira é transformada com pelo menos um vetor de expressão. Quando apenas um único vetor de expressão é utilizado (sem a adição de um produto intermédio), o vetor vai conter todas as sequências de ácidos nucleicos necessárias.

Uma vez que a célula hospedeira foi transformada com o vetor de expressão, a célula hospedeira é deixada crescer. Os métodos da presente invenção podem incluir a cultura da célula hospedeira de tal modo que os ácidos nucleicos recombinantes na célula são expressos. Para hospedeiros microbianos, este processo envolve a cultura das células em um meio apropriado. Tipicamente, as células são cultivadas a 35°C em meio apropriado. Certos meios de crescimento da presente invenção incluem, por exemplo, 15 os meios comercialmente preparados comuns, tais como o caldo Luria-Bertani (LB), o caldo Sabouraud Dextrose (SD) ou o caldo do meio de levedura (YM). Outros meios de crescimento definidos ou sintéticos podem também ser utilizados e o meio apropriado para o crescimento da célula hospedeira em particular será conhecido por alguém que seja versado na técnica da microbiologia ou ciência da fermentação. As variações de temperatura e 20 as outras condições adequadas para o crescimento são conhecidos na técnica (vide, por exemplo, Bailey e Ollis 1986).

Os métodos para a purificação de proteínas recombinantes da presente invenção a partir da célula hospedeira são bem conhecidos na técnica (vide ELV Harris e S. Angel eds (1989), Métodos de Purificação de Proteína : Uma abordagem prática, IRL Press, Oxford, Inglaterra) . Tais métodos 25 incluem, sem limitação, a eletroforese em disco de gel preparativa, focagem, cromatografia líquida de alta desempenho isoeletrico (HPLC), HPLC de fase reversa, filtração em gel, cromatografia de troca iônica e de partição e distribuição em contracorrente e as suas combinações. Em certas modalidades, as proteínas recombinantes transportam os marcadores de sequência adicionais com a finalidade de facilitar a purificação. Tais marcadores incluem 30

as etiquetas de epítomos e as etiquetas de proteínas. Exemplos não limitativos de marcadores de epítopo incluem c-myc, hemaglutinina (HA), polihistidina (6x-HIS), GLU-GLU e DYKDDDDK (FLAG) (SEQ ID NO : 117) marcadores de epítopo. Os marcadores de epítopo podem ser adicionados aos

5 peptídeos por meio de uma série de métodos reconhecidos. As sequências de DNA de marcadores de epítopo podem ser inseridas em sequências de codificação de proteínas recombinantes, como oligonucleotídeos ou através de iniciadores utilizados na amplificação por meio de PCR. Como alternativa, as sequências codificadoras de peptídeos podem ser clonados em vetores

10 específicos que geram fusões com marcadores de epítopo, por exemplo, vetores pRSET (Invitrogen Corp, San Diego, Calif.) Exemplos não limitantes de marcas de proteínas incluem glutathione-S-transferase (GST), proteína fluorescente verde (GFP) e proteína de ligação à maltose (MBP). Proteína etiquetas estão ligados a peptídeos ou polipeptídeos através de vários métodos bem conhecidos. Em uma abordagem, a sequência de codificação de

15 um polipeptídeo ou peptídeo pode ser clonado em um vetor que gera uma fusão entre o polipeptídeo ou peptídeo e a marcação de proteínas de interesse. Os vetores adequados incluem, sem limitação, os plasmídeos exemplares, pGEX (Amersham Pharmacia Biotech, Inc., Piscataway, NJ), pEGFP (Clontech Laboratories, Inc., Palo Alto, Califórnia), e pMAL.TM. (New

20 England Biolabs, Inc., Beverly, Mass.) Após a expressão, o epítopo ou uma proteína marcada com polipeptídeo ou peptídeo pode ser purificado a partir de um lisado em bruto da célula hospedeira por meio da cromatografia em uma matriz de fase sólida apropriada. Em alguns casos, pode ser preferível

25 remover o epítopo ou marcação da proteína (isto é, através de clivagem de protease) seguinte de purificação.

Métodos de Produção de Complexos de Glicanos

Um outro aspecto da presente invenção inclui os métodos para a produção de um complexo de N-glicanos, incluindo as etapas de proporcionar uma célula hospedeira, onde a célula hospedeira contém um polinucleotídeo que codifica uma proteína de fusão que compreende uma N-acetilglucosaminiltransferase I e um domínio catalítico N- acetilglucosaminiltransfera-

30

se II e domínio catalítico da cultura da célula hospedeira de tal modo que a proteína de fusão é expressa, a proteína de fusão em que catalisa a transferência de N-acetilglucosamina com um Resíduo $\text{Man}\alpha 3$ terminal e N-acetilglucosamina a um resíduo $\text{Man}\alpha 6$ terminal de um glicano aceitador com a finalidade de produzir um complexo N-glicano. Em certas modalidades, este aspecto inclui os métodos de produzir N-glicanos do tipo humano em uma célula de *Trichoderma*.

Conforme usado na presente invenção, o termo "complexo de N-glicanos" refere-se a um N-glicano que compreende uma estrutura $\text{GlcNAc}_2\text{Man}_3$ terminal.

O complexo de N-glicano inclui qualquer glicano possuindo a fórmula $[\text{GlcNAc}\beta 2]_z \text{Man}\alpha 3 ([\text{GlcNAc}\beta 2]_w \text{Man}\alpha 6) \text{Man} \{\beta 4 \text{GlcNAc}\beta (\text{Fuc}\alpha)_x\}_n [4\text{GlcNAc}]_m \text{p}$, onde n, m e p são 0 ou 1, indicando a presença ou ausência de uma parte da molécula, com a condição de que quando m é 0, então n é 0 (fucose é um ramo ligado a GlcNAc), onde x é 3 ou 6, onde () define uma ramificação na estrutura, onde [] define uma parte da estrutura do glicano pode estar presente ou ausente em uma sequência linear, e onde z e w são 0 ou 1. Preferivelmente, W e Z são : 1. Em certas modalidades, o complexo de N-glicanos inclui $\text{GlcNAc}\beta 2\text{Man}\alpha 3$ ($\text{GlcNAc}\beta 2\text{Man}\alpha 6$) $\text{Man}\beta 4\text{GlcNAc}\beta 4\text{GlcNAc}$, $\text{GlcNAc}\beta 2\text{Man}\alpha 3$ ($\text{Man}\alpha 6$) $\text{Man}\beta 4\text{GlcNAc}\beta 4\text{GlcNAc}$, $\text{GlcNAc}\beta 2\text{Man}\alpha 3$ ($\text{GlcNAc}\beta 2\text{Man}\alpha 6$) $\text{Man}\beta 4\text{GlcNAc}\beta 4$ ($\text{Fuc}\alpha 6$) GlcNAc , $\text{GlcNAc}\beta 2\text{Man}\alpha 3$ ($\text{Man}\alpha 6$) $\text{Man}\beta 4\text{GlcNAc}\beta 4$ ($\text{Fuc}\alpha 6$) GlcNAc e $\text{Man}\alpha 3$ ($\text{Man}\alpha 6$) $\text{Man}\beta 4\text{GlcNAc}\beta 4\text{GlcNAc}$. Em certas modalidades, os complexos de N-glicanos são GlcNAcMan_3 , $\text{GlcNAc}_2\text{Man}_3$, e ou Man_3 fúngicos não fucosilados

Em certas modalidades, o método de produção de um complexo de N-glicanos irá gerar uma mistura de diferentes glicanos. O complexo de N-glicanos, pode constituir pelo menos 1 %, pelo menos 3 %, pelo menos 5 %, pelo menos 10 %, pelo menos 15 %, pelo menos 20 %, pelo menos 25 %, pelo menos 50 %, ou pelo menos 75 % ou mais de uma mistura de tais glicano.

O glicano aceitador e, portanto, o complexo de N-glicanos, pode ser ligado a uma molécula tal como um aminoácido, um peptídeo, ou um

polipeptídeo. Em certas modalidades, o derivado de aminoácido é um resíduo de asparagina. O resíduo de asparagina pode ser uma ligação aminoglicosídica a partir da amida da cadeia lateral (uma estrutura de ligação de N-glicanos polipeptídeo de mamífero biológico) e podem ser parte de uma cadeia peptídica, tais como um dipeptídeo, um oligopeptídeo ou um polipeptídeo. O glicano pode ser um derivado final de redução, tal como uma N-, O-, ou C-ligado, de preferência glicosídica, derivado da redução de GlcNAc ou Man, tal como um separador ou resíduo orgânico terminal com uma determinada estrutura de glicano ligado selecionada de entre o grupo de um aminoácido, um grupo alquila, heteroalquila, acila, alquilóxi, arila, arilalquila, e heteroarilalquila. O espaçador pode ser ainda ligado a um suporte polivalente ou uma fase sólida. Em certas modalidades, as estruturas alquila contendo incluem metila, etila, propila, e C4-C26, alquilas, tais como lípidos, fosfolípidos, glicerolípids dolicol-fosfolípidos e as ceramidas e derivados. A extremidade redutora pode ser também derivatizada por meio da aminação redutiva de uma ligação amina secundária ou uma estrutura derivada. Certos veículos incluem biopoli- ou oligômeros, tais como (poli) peptídeos, poli (sacarídeos), tais como dextrano, celulose, amilase, ou glicosaminoglicanos, e outros oligômeros ou polímeros orgânicos, tais como plásticos, incluindo o polietileno, polipropileno, poliamidas (por exemplo, de náilon ou de poliestireno), poliacrilamida, e os ácidos polilácticos, tais como os dendrímeros PAMAM, Starburst ou dendrímeros do mar, ou a polilisina, e polialquilglicóis tais como polietilenoglicol (PEG). As fases sólidas podem incluir cavidades de microtitulação, as partículas de sílica, vidro, metal (incluindo aço, ouro e prata), grânulos de polímeros, tais como poliestireno ou resina de contas, contas de ácido poliláctico, contas de polissacarídeos ou espaçadores orgânicos contendo as contas magnéticas.

Em certas modalidades, o glicano aceitador está ligado a um polipeptídeo heterólogo. Em certas modalidades, o polipeptídeo heterólogo é uma proteína terapêutica. As proteínas terapêuticas podem incluir os anticorpos monoclonais, interferonas, eritropoetinas, hormonas de crescimento, enzimas ou fatores de coagulação do sangue e podem ser úteis no trata-

mento de seres humanos ou animais. Por exemplo, o glicano aceitador pode ser ligado a uma proteína terapêutica, tais como rituximab.

O glicano aceitador pode ser qualquer um dos glicanos aceitadores descritas na seção intitulada "proteínas recombinantes da presente invenção."

Em certas modalidades, o glicano aceitador pode ser man5. Nessas modalidades, a cepa de expressão *T. reesei* man5 é transformada com uma enzima de fusão GnTII / GnTI usando a integração aleatória ou por meio da integração direcionada para um local conhecido conhecido por não afetar a glicosilação man5. As cepas que produzem GlcNAcMan5 são selecionadas. As cepas selecionadas são ainda transformadas com um domínio catalítico de manosidase do tipo II, manosidase capaz de clivar as estruturas man5 com a finalidade de gerar GlcNAcMan3. Em certas modalidades, as enzimas do tipo manosidase II pertencem a família glicosídeo hidrolases 38 (cazy.org/GH38_all.html). As enzimas caracterizadas incluem as enzimas enumeradas no cazy.org/GH38_characterized.html. As enzimas especialmente úteis são as enzimas do tipo Golgi que clivam as glicoproteínas, tais como as da subfamília α -manosidase II (Man2A1 ; ManA2). Exemplos de tais enzimas incluem enzima humana AAC50302, enzima *D. melanogaster* (Van den Elsen, JM et al (2001) EMBO J. 20 : 3.008-3.017), aqueles com a estrutura 3D de acordo com APO-1HTY referência, e outros com o referido domínio catalítico no PDB. Para expressão citoplasmática o domínio catalítico da manosidase é tipicamente fundida com um peptídeo alvo do Terminal N ou expressos com animais endógenos ou as estruturas de direcionamento da planta Golgi das enzimas de plantas ou de animal de manosidase II. Após a transformação com o domínio catalítico de manosidase de uma manosidase do tipo II, uma cepa que produz de forma eficaz GlcNAc2Man3 está selecionada.

Células hospedeiras

Os métodos de produção de um complexo de N-glicanos incluem uma primeira etapa de proporcionar uma célula hospedeira. Qualquer célula hospedeira procariótica ou eucariótica pode ser utilizada na presente

invenção, desde que ela permaneça viável após ter sido transformada com uma sequência de ácidos nucleicos. De preferência, a célula hospedeira não é adversamente afetada por meio da transdução das sequências necessárias de ácido nucleico, a expressão subsequente de proteínas recombinantes, ou os intermediários resultantes. As células eucarióticas adequadas incluem, mas não estão limitadas a, fungos, plantas, insetos ou células de mamíferos.

Em certas modalidades, o hospedeiro é uma cepa fúngica. O termo "fungos" usado na presente invenção inclui as cepas *Ascomycota*, *Basidiomycota*, *Chytridiomycota* e *Zygomycota* (como definido por Hawksworth et al., In, de Ainsworth e Bisby Dicionário dos fungos, 8^a edição, 1995, CAB Internacional, University Press, Cambridge, Reino Unido), assim como o *Oomycota* (como citado em Hawksworth et al. de 1995, supra, página 171) e todos os fungos mitospóricos (Hawksworth et al., 1995, supra).

Em modalidades particulares, o hospedeiro é uma cepa de fungos de levedura. O termo "levedura" tal como é usado na presente invenção inclui a levedura ascospórogena (Endomycetales), levedura basidiospórogena, e levedura que pertence ao *Fungi Imperfecti* (Blastomycetes). Uma vez que a classificação da levedura pode mudar no futuro, para os fins da presente invenção, a levedura deve ser definida como descrito em Biologia e Atividades de levedura (Skinner, FA, Passmore, SM, e Davenport, RR, eds, Soc.. App. Bacteriol. Symposium Series No. 9, 1980).

Em certas modalidades, a cepa hospedeira é uma levedura *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, ou *Yarrowia*.

Em certas modalidades, a cepa hospedeira é uma levedura *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Pichia pastoris*, *Candida albicans*, *Hansenula polymorpha*, *Schizosaccharomyces*, ou *Yarrowia*.

Em outra modalidade particular, a célula hospedeira fúngica é uma cepa de fungos filamentosos. O termo "fungos filamentosos" incluem todas as formas filamentosas da subdivisão Eumycota e Oomycota (como definido por Hawksworth et al., 1995, supra). Os fungos filamentosos são

geralmente caracterizados por uma parede micelial composta por quitina, celulose, glucano, quitosano, manano, e outros polissacarídeos complexos. O crescimento vegetativo é por meio do alongamento de hifas e catabolismo de carbono é obrigatoriamente aeróbico. Em contraste, o crescimento vegetativo por leveduras tal como *Saccharomyces cerevisiae*, é por meio do brotamento de um talo unicelular e o catabolismo de carbono pode ser fermentativo.

A célula hospedeira fúngica filamentosa pode ser, por exemplo, uma cepa *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Scytalidium*, *Thielavia*, *Tolypocladium*, ou *Trichoderma*.

Em certas modalidades, a célula hospedeira de fungo filamentoso é uma cepa *Trichoderma sp.*, *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filibasidium*, *Fusarium*, *Gibberella*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, ou *Tolypocladium*.

Em certas modalidades, a célula hospedeira é uma célula de mamífero. Tais células podem ser humanas ou não humanas.

Em outras certas modalidades, a célula hospedeira é procariótica, e em certas modalidades, as procariotas são *E. coli*, *Bacillus subtilis*, *Zymomonas mobilis*, *Clostridium sp.* *Clostridium phytofermentans*, *Clostridium thermocellum*, *Clostridium beijerinckii*, *Clostridium acetobutlicum* (*Moorella thermoacetica*), *Thermoanaerobacterium saccharoliticum*, ou *Klebsiella oxytoca*. Em outras modalidades, as células hospedeiras procarióticas são *Carboxydocella sp.*, *Corynebacterium glutamicum*, *Enterobacteriaceae*, *Erwinia chrysanthemi* e *Lactobacillus sp.*, *Acidilactici*, *Rhodopseudomonas capsulata*, *Streptococcus lactis*, *Vibrio furnissii*, *Vibrio furnissii M1*, *Caldicellulosiruptor saccharoliticus* ou *Xanthomonas campestris*. Em outras modalidades, as células hospedeiras são cianobactérias. Exemplos adicionais de células hospedeiras bacterianas incluem, sem limitação, aquelas espécies atribuídas à *Escherichia*, *Enterobacter*, *Nitrogênobacter*, *Erwinia*, *Bacillus*,

Pseudomonas, *Klebsiella*, *Prbateladaus*, *Salmonella*, *Serratia*, *Shigella*, *Rhizobia*, *Vitreoscilla*, *Synechococcus*, *Synechocystis*, e as classes taxonômicas *Paracoccus*.

Em métodos da presente invenção para a produção de um complexo de N-glicanos, os métodos incluem uma etapa de cultura da célula hospedeira de tal modo que a proteína de fusão é expressa. Para hospedeiros microbianos, este processo envolve a cultura das células em um meio apropriado. Tipicamente, as células são cultivadas a 35°C em meio apropriado. Certos meios de crescimento da presente invenção incluem, por exemplo, os meios comercialmente preparados comuns, tais como caldo de Luria-Bertani (LB), caldo de Sabouraud Dextrose (SD) ou caldo do meio de levedura (YM). Outros meios de crescimento definidos ou sintéticos podem também ser utilizados e o meio apropriado para o crescimento da célula hospedeira em particular será conhecido por alguém que seja versado na técnica da microbiologia ou na ciência da fermentação. As variações de temperatura e as outras condições adequadas para o crescimento são conhecidas na técnica (ver, por exemplo, Bailey e Ollis 1986). Em certas modalidades, o pH da cultura de células situa-se entre 3,5 e 7,5, entre 4,0 e 7,0, entre 4,5 e 6,5, entre 5 e 5,5, ou 5,5.

As células hospedeiras utilizadas nos métodos de produção de um complexo de N-glicanos contêm um polinucleotídeo que codifica qualquer uma das proteínas recombinantes da presente invenção, como descrito na seção intitulada "Proteínas recombinantes da presente invenção." Em certas modalidades, a célula hospedeira contêm um polinucleotídeo que codifica uma proteína de fusão que compreende um domínio catalítico de N-acetilglucosaminiltransferase I e um domínio catalítico de N-acetilglucosaminiltransferase II, em que a proteína de fusão catalisa a transferência de N-acetilglucosamina com um Resíduo Man α 3 terminal e N-acetilglucosamina, a um terminal de Man α 6 resíduo um glicano aceitador com a finalidade de produzir um complexo de N-glicanos.

Em certas modalidades, a célula hospedeira contêm um polinucleotídeo que codifica um veículo de UDP-GlcNAc. O polinucleotídeo que

codifica o veículo de UDP-GlcNAc podem ser endógenos (isto é, de uma forma natural presente) na célula hospedeira, ou pode ser heterólogo para a célula hospedeira.

Em certas modalidades, a célula hospedeira contém um polinucleotídeo que codifica uma α -1,2-manosidase. O polinucleotídeo que codifica o α -1,2-manosidase pode ser endógeno na célula hospedeira, ou pode ser heterólogo para a célula hospedeira. Estes polinucleotídeos são especialmente úteis para uma célula hospedeira que expressa glicanos de manose elevada transferidos do Golgi para o ER, sem clivagem eficaz de exo- α -2-manosidase. O α -1,2-manosidase pode ser uma enzima manosidase do tipo I pertencente à família glicosídeo hidrolases 47 (cazy.org/GH47_all.html). Em certas modalidades, α -1,2-manosidase é uma enzima listada no cazy.org/GH47_characterized.html. Em particular, o α -1,2-manosidase pode ser uma enzima do tipo ER que cliva glicoproteínas, tais como as enzimas na subfamília de ER α -manosidase I EC 3.2.1.113 enzimas. Exemplos de tais enzimas incluem humano 1B α -2-manosidase (AAC26169), uma combinação de mamíferos manosidases ER, ou uma enzima, tais como fungos filamentosos α -1,2-manosidase (MDS1) (*T. reesei* AAF34579 ; Maras M et al J Bibateladach. 77, 2000, 255). Para a expressão citoplasmática do domínio catalítico da manosidase é tipicamente fundida com um peptídeo alvo, tais como HDEL, KDEL, ou parte de uma proteína ER ou Golgi inicial, ou expressa com um ER endógeno de direcionamento das estruturas de uma enzima de animal ou de planta manosidase I.

Em certas modalidades, a célula hospedeira contém um polinucleotídeo que codifica uma galactosiltransferase. Galactosiltransferases transferem os resíduos de galactose β ligados ao terminal de resíduo N-acetilglucosaminil. Em certas modalidades da galactosiltransferase é uma β -4-galactosiltransferase. Geralmente, β -4-galactosiltransferases pertencem à família CAZy glicosiltransferase 7 (cazy.org/GT7_all.html) e incluem β -N-acetilglucosaminil-glicopeptídeo β -1,4-galactosiltransferase (CE 2.4.1.38), que também é conhecido como a N-acetilactosamina sintase (EC 2.4.1.90). As subfamílias úteis incluem β 4-GalT1, β 4-GalT-II,-III-IV-V, e VI, tais como

β 4-GalTI ou β 4GalT-II,-III-IV, V-e - VI mamíferos ou humanos ou quaisquer combinações dos mesmos. β 4-GalT1, β 4-GalTII ou β 4-GalTIII são especialmente úteis para a galactosilação do terminal de GlcNAc β 2 estruturas de N-glicanos, tais como GlcNAcMan3, GlcNAc2Man3 ou GlcNAcMan5 (Guo S. et al. Glicobiology 2001, 11 : 813-20) . A estrutura tridimensional da região catalítica é conhecida (por exemplo, (2006) J. Mol. Biol 357 : . 1619 a 1633), e a estrutura foi representada na base de dados PDB com código 2FYD. O banco de dados CAZy inclui os exemplos de certas enzimas. As enzimas caracterizadas também estão listadas no banco de dados CAZy em cazy.org/GT7_characterized.html. Exemplos de enzimas úteis incluem β 4GalT β 4GalT1, por exemplo bovino *Bos taurus* enzima AAA30534.1 (Shaper NL et al Proc Natl Acad Sci EUA 83 (6), 1573 a 1577 (1986).), enzima humana (Guo S. et al 2001 Glicobiology, 11 : 813 a 20) e de enzima *Mus musculus* AAA37297 (Shaper, NL et al 1998, J. Biol Chem 263 (21), 10.420 a 10.428), enzimas, tais como β 4GalTII β 4GalTII BAA75819.1 humano, de hamster chinês *Cricetulus griseus* AAM77195, enzima *Mus musculus* BA-A34385 e japonês Medaka peixe *Oryzias latipes* BAH36754 e enzimas β 4GalTIII como humano β 4GalTIII BAA75820.1, hamster chinês *Cricetulus griseus* AAM77196 e enzima *Mus musculus* AAF22221.

A galactosiltransferase pode ser expressa no citoplasma da célula hospedeira. Um peptídeo alvo heterólogo, tal como um peptídeo Kre2 Schwientek descrito no J. Biol. Chem 1996 3398, pode ser utilizado. Os promotores que podem ser utilizados para a expressão da galactosiltransferase incluem os promotores constitutivos tais como promotores de *gpd*, enzimas de glicosilação endógenas e glicosiltransferases, tais como manosiltransferases que sintetizam N-glicanos no Golgi ou ER, e os promotores induzíveis de proteínas endógenas de elevado rendimento, tais como o promotor *cbh1*.

Em certas modalidades da presente invenção em que a célula hospedeira contém um polinucleotídeo que codifica uma galactosiltransferase, a célula hospedeira contém também um polinucleotídeo que codifica uma UDP-Gal A /e/ou o veículo de UDP-Gal. Em certas modalidades da presente invenção em que a célula hospedeira contém um polinucleotídeo que codifi-

ca uma galactosiltransferase, a lactose pode ser usada como fonte de carbono, em vez de glicose, quando a cultura da célula hospedeira. O meio de cultura pode estar compreendido entre pH 4,5 e 7,0 ou entre 5,0 e 6,5. Em certas modalidades da presente invenção em que a célula hospedeira contém um polinucleotídeo que codifica uma galactosiltransferase e um polinucleotídeo que codifica uma UDP-Gal /e/ou UDP-Gal veículo, um cátion bivalente tal como Mn^{2+} , Ca^{2+} ou Mg^{2+} pode ser adicionado ao meio de cultura de células.

Em certas modalidades, a célula hospedeira contém um polinucleotídeo que codifica uma sialiltransferase. A transferência sialiltransferases $\alpha 3$ ou ácido siálico $\alpha 6$ -linked, como Neu5Ac, a Gal terminal glicanos complexos galactosilação. Exemplos de sialiltransferases adequadas podem ser encontrados na família de proteínas de glicosilação 29 (cazy.org/GT29.html). A3-ou- $\alpha 6$ úteis incluem sialiltransferases β -galactósido α -2,6-sialiltransferase (EC 2.4.99.1), com uma certa subfamília ST6Gal-I, e N-acetilactosaminide α -2,3-sialiltransferase (EC 2.4.99.6) com uma possível reatividade cruzada com β -galactósido α -2,3-sialiltransferase (EC 2.4.99.4). Subtipos úteis de $\alpha 3$ -sialiltransferases incluem ST3Gal-III e IV-ST3Gal. Certas espécies enzimaticamente caracterizados destes estão listados como caracterizado por a base de dados cazy de enzimas de glicosilação (cazy.org/GT29_characterized.html). O polinucleotídeo que codifica a sialiltransferase $\alpha 3$ ou $\alpha 6$ -ligado pode ser endógeno à célula hospedeira, ou pode ser heterólogo para a célula hospedeira. Sialilação da célula do hospedeiro pode requerer a expressão de enzimas que sintetizam o dador de ácido CMP-siálico, tais como CMP-Neu5Ac, especialmente em células de fungos, plantas, nematóides / parasita, ou inseto.

A célula hospedeira pode ter níveis de atividade de várias enzimas endógenas aumentada ou reduzida. Um nível reduzido de atividade pode ser fornecido através da inibição da atividade da enzima endógena com um inibidor, um anticorpo, ou semelhantes. Em certas modalidades, a célula hospedeira é geneticamente modificada de forma a aumentar ou reduzir a atividade de várias enzimas endógenas. "Geneticamente modificado" refere-

se a qualquer DNA recombinante ou de RNA método usado para criar uma célula hospedeira procariótica ou eucariótica que expressa um polipeptídeo em níveis elevados, em níveis baixos, ou em uma forma mutada. Em outras palavras, a célula hospedeira foi transfectada, transformada ou transduzida com uma molécula de polinucleotídeo recombinante, e, dessa maneira, foi alterado de modo a fazer com que a célula para alterar a expressão de uma proteína desejada.

Modificações genéticas que conduzam a uma redução na expressão do gene, na função do gene, ou em função do produto do gene (ou seja, a proteína codificada pelo gene) pode ser referido como a inativação (completa ou parcial), deleção, interrupção, o bloqueio, o silenciamento, ou regulação, ou atenuação da expressão de um gene. Por exemplo, uma modificação genética em um gene que resulta em um decréscimo na função da proteína codificada por tal gene, pode ser o resultado de uma deleção completa do gene (ou seja, o gene não existe, e, portanto, a proteína não existe), uma mutação no gene que resulta na incompleta ou nenhuma tradução da proteína (por exemplo, a proteína não é expressa), ou uma mutação no gene que reduz ou suprime a função natural da proteína (por exemplo, uma proteína é expressa que diminuiu ou nenhuma atividade enzimática ou ação). Mais, de uma maneira específica,, a referência a redução da acção de proteínas aqui discutido, geralmente refere-se a qualquer modificação genética na célula hospedeira em questão, o que resulta na diminuição da expressão /e/ou a funcionalidade (a atividade biológica) das proteínas e inclui a diminuição da atividade das proteínas (por exemplo,, diminuição da catálise), o aumento da inibição ou da degradação das proteínas, bem como a redução ou a eliminação da expressão das proteínas. Por exemplo, a acção ou atividade de uma proteína da presente invenção pode ser diminuída através do bloqueio ou reduzindo a produção da proteína, reduzindo a acção da proteína, ou de inibir a acção da proteína. Combinações de algumas destas modificações também são possíveis. Bloqueando ou reduzindo a produção de uma proteína pode incluir a colocação do gene que codifica a proteína sob o controle de um promotor que requer a presença de um composto

indutor no meio de crescimento. Através do estabelecimento de condições de tal forma que o indutor se esgota do meio, a expressão do gene que codifica a proteína (e, portanto, a síntese de proteína) pode ser desligada. O bloqueio ou redução da acção de uma proteína pode também incluir o uso de

5 uma abordagem da tecnologia excisão semelhante ao descrito na Patente U.S.. No. 4743546. Para utilizar esta abordagem, o gene que codifica a proteína de interesse é clonado entre as sequências genéticas específicas que permitem a excisão controlada específica do gene a partir do genoma. Excisão pode ser solicitada por, por exemplo, uma mudança na temperatura de

10 cultivo da cultura, tal como na Patente U.S. N ° 4.743.546, ou por qualquer outro sinal físico ou nutricional.

Em geral, de acordo com a presente invenção, um aumento ou uma diminuição de uma dada característica de uma proteína mutante ou modificada (por exemplo, atividade enzimática) é feita com referência à

15 mesma característica de um tipo selvagem (isto é, normais, não modificado) proteína que é derivada do mesmo organismo (a partir da mesma fonte ou sequência pai), que é medido ou estabelecida sob as mesmas condições, ou equivalente. Do mesmo modo, um aumento ou uma diminuição na característica de uma célula hospedeira geneticamente modificada (por exemplo, a

20 expressão /e/ou a atividade biológica de uma proteína, ou a produção de um produto) é feita com referência à mesma característica de uma célula hospedeira de tipo selvagem da mesma espécie, e de preferência da mesma cepa, sob as mesmas condições, ou equivalente. Tais condições incluem as condições de ensaio ou de cultura (por exemplo, componentes do meio,

25 temperatura, pH, etc) em que a atividade da proteína (por exemplo, a expressão ou a atividade biológica) ou outras características da célula hospedeira é medido, bem como os tipo de ensaio utilizado, a célula hospedeira que é avaliada, etc Tal como discutido acima, as condições são condições equivalentes (por exemplo, condições de cultura), que são semelhantes,

30 mas não necessariamente idênticos (por exemplo, algumas alterações conservadoras em condições podem ser toleradas) e que não alteram substancialmente o efeito sobre o crescimento celular ou a expressão da enzima ou

a atividade biológica, em comparação com uma comparação feita sob as mesmas condições.

De preferência, uma célula hospedeira geneticamente modificada que tem uma modificação genética que aumenta ou diminui a atividade de uma dada proteína (por exemplo, uma enzima) tem um aumento ou uma diminuição, respectivamente, na atividade ou acção (por exemplo, a expressão, produção /e/ou a atividade biológica) da proteína, em comparação com a atividade da proteína de tipo selvagem em uma célula hospedeira de tipo selvagem, de pelo menos cerca de 5 %, e mais preferivelmente pelo menos cerca de 10 %, e mais preferencialmente pelo menos cerca de 15 %, e mais preferivelmente pelo menos cerca de 20 %, e mais preferivelmente pelo menos cerca de 25 %, e mais preferivelmente pelo menos cerca de 30 %, e mais preferivelmente pelo menos cerca de 35 %, e mais preferivelmente pelo menos cerca de 40 %, e mais preferencialmente pelo menos cerca de 45 %, e mais preferivelmente pelo menos cerca de 50 %, e mais preferivelmente pelo menos cerca de 55 %, e mais preferivelmente pelo menos cerca de 60 %, e mais preferivelmente pelo menos cerca de 65 %, e mais preferivelmente pelo menos cerca de 70 %, e mais preferivelmente pelo menos cerca de 75 %, e mais preferivelmente pelo menos cerca de 80 %, e mais preferivelmente pelo menos cerca de 85 %, e mais preferivelmente pelo menos cerca de 90 %, e mais preferivelmente pelo menos cerca de 95 %, ou de qualquer percentagem, em números inteiros entre 5 % e 100 % (por exemplo, 6 %, 7 %, 8 %, etc.) As mesmas diferenças são determinadas comparando uma molécula de ácido nucleico modificado ou isolado de proteína directamente para a molécula do tipo selvagem isolada de ácido nucleico ou proteína (por exemplo, se a comparação é realizada in vitro, em comparação com in vivo).

Em outro aspecto da presente invenção, uma célula hospedeira geneticamente modificada que tem uma modificação genética que aumenta ou diminui a atividade de uma dada proteína (por exemplo, uma enzima) tem um aumento ou uma diminuição, respectivamente, na atividade ou acção (por exemplo, a expressão, a produção /e/ou atividade biológica) da proteína, em comparação com a atividade da proteína de tipo selvagem em uma

célula hospedeira de tipo selvagem, de pelo menos cerca de 2 vezes, e mais preferivelmente pelo menos cerca de 5 - vezes, e mais preferivelmente pelo menos cerca de 10 vezes, e mais preferivelmente cerca de 20 vezes, e mais preferivelmente pelo menos cerca de 30 vezes, e mais preferivelmente pelo menos cerca de 40 vezes, e mais preferivelmente pelo menos cerca de 50 vezes, e mais preferivelmente pelo menos cerca de 75 vezes, e mais preferivelmente pelo menos cerca de 100 vezes, e mais preferivelmente pelo menos cerca de 125 vezes, e mais preferivelmente pelo menos cerca de 150 vezes, ou de todo qualquer incremento inteiro a partir de, pelo menos, cerca

5 2 vezes (por exemplo, 3 vezes, 4 vezes, 5 vezes, 6 vezes, etc.).

10

Em certas modalidades, a célula hospedeira tem um nível reduzido de atividade de um dolichil-P-Man : Man (5) GlcNAc (2)-PP-dolichil manosiltransferase em relação ao nível da atividade de uma célula hospedeira de tipo selvagem . Dolichil-P-Man : Man (5) GlcNAc (2)-PP-dolichil manosiltransferase (CE 2.4.1.130) transfere um resíduo de alfa-D-manosila dolichilfosfato a partir de D-manose em uma membrana lipídica ligada ao oligossacarídeo. Tipicamente, o dolichil-P-Man : Man (5) GlcNAc (2)-PP-manosiltransferase dolichil enzima é codificada por um gene alg3. Em certas modalidades, a célula hospedeira tem um nível reduzido de expressão de um gene

15 alg3 comparação com o nível de expressão em uma célula hospedeira de tipo selvagem. Em certas modalidades, o gene alg3 é eliminado da célula hospedeira.

20

Em certas modalidades, a célula hospedeira tem um nível reduzido de atividade de alfa-1, 6-manosiltransferase em relação ao nível da atividade de uma célula hospedeira de tipo selvagem. Alfa-1,6-manosiltransferase (CE 2.4.1.232) transfere um resíduo de alfa-D-manosila de GDP-manose em um oligossacarídeo da proteína ligada, formando um alongamento iniciando alfa-(1 -> 6)-D-manosil-D ligação-manose no Golgi apparatus .. Tipicamente, a enzima alfa-1,6-manosiltransferase é codificada por

25 um gene OCH1. Em certas modalidades, a célula hospedeira tem um nível reduzido de expressão de um gene OCH1 em comparação com o nível de expressão em uma célula hospedeira de tipo selvagem. Em certas modali-

30

dades, o gene OCH1 é eliminado da célula hospedeira.

Em certas modalidades, a célula hospedeira tem um nível reduzido de atividade das proteases. Em certas modalidades, os genes que codificam para várias proteases são suprimidos a partir da célula hospedeira.

- 5 Estes genes incluem, por exemplo, os genes que codificam as proteases, tais como pep1 (pepA em *Aspergillus*) e as enzimas celulolíticas, como cellobiohydrolase1 (cbh1).

- Em certas modalidades, a célula hospedeira pode ter um nível reduzido de atividade das proteínas envolvidas na extremidade não homóloga de união (NHEJ), a fim de melhorar a eficiência de recombinação homóloga. Em certas modalidades, os genes que codificam estas proteínas são suprimidos a partir da célula hospedeira. Os genes e os seus homólogos incluem, mas não estão limitados a, Ku70, Ku80, lig4, RAD50, Xrs2, Sir4, Lif1, Neil ou como descrito, por exemplo, de Ninomiya et al. 2004, Ishibashi et al. 2006, Villalba et al. 2008 e Mizutani et al. 2008.

- Em certas modalidades de métodos de produção de um complexo de N-glicanos, a célula hospedeira é uma célula de *Trichoderma* que tenha um nível reduzido de atividade de um dolichil-P-Man : Man (5) GlcNAc (2)-PP-dolichil manosiltransferase em relação ao nível da atividade de uma célula de tipo selvagem de *Trichoderma*.

- Em outras certas modalidades de métodos de produção de um complexo de N-glicanos, a célula hospedeira é uma célula de levedura que tem um nível reduzido de atividade de um dolichil-P-Man : Man (5) GlcNAc (2)-PP- dolichil manosiltransferase e um nível reduzido de atividade de alfa-1, 6-manosiltransferase em relação aos níveis de atividade de uma célula de levedura do tipo selvagem e compreende ainda um polinucleotídeo que codifica uma α -1,2-manosidase.

Métodos *in vitro* de produção de N-glicanos complexos

- Em outro aspecto, a presente invenção proporciona um método de produção de um complexo de N-glicanos, incluindo uma etapa de incubação de uma proteína de fusão que compreende um domínio catalítico de N-acetilglucosaminiltransferase I e um domínio catalítico de N-acetilglucosa-

miniltransferase II, de um glicano aceitador, e um doador N-acetilglucosamina juntos em um tampão, onde a proteína de fusão catalisa a transferência de N-acetilglucosamina com um Resíduo Man α 3 terminal e N-acetilglucosamina a um resíduo Man α 6 terminal de um glicano aceitador com a finalidade de produzir um complexo de N-glicanos. Em certas modalidades, o glicano aceitante está ligada a um aminoácido, um peptídeo, ou um polipeptídeo. Em certas modalidades, o glicano aceitante está ligada a um polipeptídeo heterólogo. Em certas modalidades, o glicano aceitador é Man3. Em certas modalidades do doador N-acetilglucosamina é um veículo de UDP-GlcNAc. Tipicamente, o tampão compreende um cátion divalente, tais como Mn²⁺, Ca²⁺ ou Mg²⁺ em concentrações de 1 mM a 100 mM, 100 μ M a 50 mM ou 0,1 mM e 25 mM. O doador N-acetilglucosamina é tipicamente utilizado em um excesso molar, tal como excesso de 1,1 a 100 vezes no que diz respeito aos locais aceptadores reativos no glicano aceitador. A concentração do glicano aceitador é tipicamente entre 1 pM a 100 mM, 100 μ M a 50 mM, ou de 1 a 25 mM. Sempre que o glicano aceitante está ligada a um polipeptídeo, as gamas de concentração são geralmente na extremidade inferior por causa de pesos moleculares mais elevados. As concentrações dos componentes da reação podem ser ajustadas com base nas suas solubilidades no tampão. A quantidade de atividade da enzima (unidades) pode ser ajustada com a finalidade de permitir uma reação eficaz dentro de um tempo de reação razoável. Um tempo razoável de reação é tipicamente desde alguns minutos a vários dias. Em certas modalidades o tempo de reação será de cerca de 0,5 horas a um dia, ou de 1 a 6 horas.

Os tampões úteis incluem tampões adequados para a proteína de fusão, tais como tris, HEPES, MOPS em intervalos de cerca de 5 a 8,5, pH 5,5 a 8,0, ou 6,0 e 7,5. Tipicamente, as concentrações de tampão TRIS, HEPES, MOPS ou tampões será entre 5 e 150 mM, entre 10-100 mM, 10-60 mM ou ajustada para manter o pH. A reação pode ser otimizada por adição de sal, tal como NaCl a 10-200 mM e/ou um estabilizador de enzima, mas a proteína não glicosylatable (por exemplo, um glicano não glicosilada ou não aceitador puro contendo albumina. Em uma certa modalidade da reação in

vitro é ajustado para ser realizado em meio de cultura celular. tampões de fosfato podem ser utilizados com a finalidade de reduzir a velocidade de reação.

Células e Métodos para a Produção de Glicanos $\text{Man}_3\text{GlcNAc}_2$

5 Em outro aspecto, a presente invenção proporciona as células de fungos filamentosos que contém uma mutação de *alg3* e $\text{Man}_3\text{GlcNAc}_2$, onde o $\text{Man}_3\text{GlcNAc}_2$ inclui pelo menos 50 %, pelo menos 60 %, pelo menos 70 %, pelo menos 80 %, pelo menos 90 %, ou 100 % (% em mol) de N-glicanos secretados pelas células do neutro. Os neutros N-glicanos podem
10 ser ligados a um aminoácido, um peptídeo, ou um polipeptídeo. O gene *alg3* pode ser mutado por meio de quaisquer meios conhecidos na técnica, tais como as mutações pontuais ou deleção de todo o gene *alg3*. De preferência, a função da proteína *alg3* é reduzida ou eliminada mediante a mutação de *alg3*. A célula fúngica filamentosa pode ser uma célula *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filibasidium*, *Fusarium*, *Gibberella*, *Humicola*, *Magnaporthe*, *Mu-*
15 *cor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, ou *Trichoderma*. Em certas modalidades, a célula filamentosa fúngica é uma célula *T. reesei*. Em certas modalidades, a célula fúngica filamentosa contém ainda um ou mais polinucleotídeos que codificam qualquer uma das proteínas recombinantes da presente invenção. Por exemplo, a célula fúngica filamentosa pode ainda conter um
20 primeiro polinucleotídeo que codifica um Domínio catalítico de N-acetilglucosaminiltransferase I e um segundo polinucleotídeo que codifica um domínio catalítico II N-acetilglucosaminiltransferase. De uma maneira alternativa, a célula fúngica filamentosa pode ainda conter um polinucleotídeo que codifica uma proteína de fusão incluindo um Domínio catalítico de N-acetilglucosaminiltransferase I e um domínio catalítico de N-acetilglucosaminiltrans-
25 ferase II.
30

Em ainda outro aspecto, a presente invenção proporciona os métodos de produzir um glicano $\text{Man}_3\text{GlcNAc}_2$ em uma célula hospedeira,

incluindo as etapas de proporcionar uma célula hospedeira com um nível reduzido de atividade de manosiltransferase em relação ao nível de atividade em um selvagem tipo de célula hospedeira e a cultura da célula hospedeira com a finalidade de produzir um glicano Man3GlcNAc2, onde o glicano

5 Man3GlcNAc2 faz pelo menos 50 %, pelo menos 60 %, pelo menos 70 %, pelo menos 80 %, pelo menos 90 %, ou 100 % (% em mol) de N-glicanos neutros segregados por meio da célula hospedeira.

Os glicanos Man3GlcNAc2 podem ser ligados a uma molécula tal como um aminoácido, um peptídeo, ou um polipeptídeo. Em certas modalidades, o aminoácido é um resíduo de asparagina. O resíduo de asparagina

10 pode ser em uma ligação aminoglicosídica a partir da amida da cadeia lateral (a estrutura biológica da proteína de ligação de mamífero de N-glicanos) e pode ser parte de uma cadeia peptídica, tais como um dipeptídeo, um oligopeptídeo ou um polipeptídeo. O glicano pode ser um derivado final de redução, tal como uma N-, O-, ou C-ligado, de preferência glicosídica, derivado

15 da redução de GlcNAc ou humano, tal como um separador ou resíduo orgânico terminal com uma determinada estrutura de glicano ligado selecionado a partir de entre o grupo de um aminoácido, um grupo alquila, heteroalquila, acila, alquilóxi, arila, arilalquila, e heteroarilalquila. O espaçador pode ser ainda

20 ligada a um suporte polivalente ou uma fase sólida. Em certas modalidades, as estruturas alquila contendo incluem metila, etila, propila, e C4-C26, alquilas, tais como lípidos, fosfolípidos, glicerolípids dolicol-fosfolípidos e as ceramidas e derivados. A extremidade redutora pode ser também derivatizada por meio da aminação redutiva de uma ligação amina secundária ou uma

25 estrutura derivada. Certos veículos incluem biopoli- ou oligômeros, tais como (poli) peptídeos, poli (sacarídeos), tais como dextrano, celulose, amilase, ou glicosaminoglicanos, e outros oligômeros ou polímeros orgânicos, tais como plásticos, incluindo o polietileno, polipropileno, poliamidas (por exemplo, de náilon ou de poliestireno), poliacrilamida, e os ácidos polilácticos, tais como

30 os dendrímeros PAMAM, Starburst ou dendrímeros do mar, ou a polilisina, e polialquilglicóis tais como polietilenoglicol (PEG). As fases sólidas podem incluir as cavidades de microtitulação, as partículas de sílica, de vidro, de

metal incluindo aço, de ouro e de prata, esferas de polímero, tais como esferas de poliestireno ou resina de ácido poliláctico, grânulos, grânulos de polissacarídeos ou espaçadores orgânicos contendo as contas magnéticas.

Em certas modalidades, o glicano Man3GlcNAc2 está ligado a um polipeptídeo heterólogo. Em certas modalidades, o polipeptídeo heterólogo é uma proteína terapêutica. As proteínas terapêuticas podem incluir anticorpos monoclonais, interferonas, eritropoetinas, hormonas de crescimento, enzimas ou fatores de coagulação do sangue e podem ser úteis no tratamento de seres humanos ou animais. Por exemplo, os glicanos Man3GlcNAc2 podem ser ligados a uma proteína terapêutica, tais como rituximab. Tipicamente, o glicano Man3GlcNAc2 será ainda modificado para se tornar um glicano complexo. Tal modificação pode ocorrer in vivo, na célula hospedeira, ou por meio dos métodos in vitro.

Em certas modalidades, o manosiltransferase é um dolichil-P-Man : Man (5) GlcNAc (2)-PP-dolichil manosiltransferase. Tipicamente, a enzima dolichil-P-Man : Man (5) GlcNAc (2)-PP-manosiltransferase dolichila é codificada por meio de um gene *alg3*. Em certas modalidades, a célula hospedeira tem um nível reduzido de expressão de um gene *alg3* em comparação com o nível de expressão em uma célula hospedeira de tipo selvagem. Em certas modalidades, o gene *alg3* é eliminado da célula hospedeira. As SEQ ID NOs : 97 e 98 fornecem os ácidos nucleicos e as sequências de aminoácidos do gene *alg3* em *T. reesei*, respectivamente.

Em certas modalidades, o nível de atividade de alfa-1, 6-manosiltransferase na célula hospedeira não é reduzido em comparação com o nível de atividade em uma célula hospedeira de tipo selvagem. Tipicamente, a enzima alfa-1,6-manosiltransferase é codificada por meio de um gene *OCH1*. Em certas modalidades, a célula hospedeira contém um polinucleotídeo que codifica um endógeno α -1,2-manosidase.

Em certas modalidades, a célula hospedeira é uma célula de *Trichoderma* e, em certas modalidades, a célula hospedeira é uma célula de *Trichoderma reesei*.

Células de Fungos Filamentosos da Presente Invenção

De acordo com outro aspecto, a presente invenção proporciona as células de fungos filamentosos, com uma diminuição do nível de expressão de um gene *alg3* da presente invenção, em comparação com o nível de expressão do gene *alg3* em uma célula fúngica filamentosa de tipo selvagem, em que as células de fungos filamentosos também contém qualquer uma das proteínas recombinantes da presente invenção, como descrito na seção intitulada "Proteínas recombinantes da presente invenção". Por exemplo, em certas modalidades da célula fúngica filamentosa contém adicionalmente uma proteína de fusão incluindo um domínio catalítico de N-acetilglucosaminiltransferase I e um domínio catalítico de N-acetilglucosaminiltransferase II do polinucleotídeo de codificação. A expressão da proteína de fusão pode ser controlada por meio de um promotor que está operativamente ligado ao polinucleotídeo. O promotor pode ser um promotor constitutivo ou um promotor indutível. Em certas modalidades preferidas, o promotor é um promotor indutível, tal como o promotor *cbh1* indutível.

Em outro aspecto, a presente invenção proporciona as células de fungos filamentosos, com uma diminuição do nível de expressão de um gene *alg3* da presente invenção, em comparação com o nível de expressão do gene *alg3* em uma célula fúngica filamentosa de tipo selvagem, em que as células fúngicas filamentosas também contém um polinucleotídeo que codifica um primeiro domínio catalítico de N-acetilglucosaminiltransferase I recombinante e um segundo polinucleotídeo que codifica um Domínio catalítico de N-acetilglucosaminiltransferase II recombinante. Em tais modalidades, a expressão do domínio catalítico de N-acetilglucosaminiltransferase I recombinante é controlada por meio de um promotor que está ligado operativamente ao primeiro polinucleotídeo e a expressão do Domínio catalítico de N-acetilglucosaminiltransferase II recombinante é controlada por meio de um promotor que está ligado operativamente ao segundo polinucleotídeo. O promotor pode ser um promotor constitutivo ou um promotor indutível. Em certas modalidades preferidas, o promotor é um promotor indutível, tal como o promotor *cbh1* indutível.

Em outras modalidades, um único polinucleotídeo pode codificar tanto o domínio catalítico de N-acetilglucosaminiltransferase I recombinante e o Domínio catalítico de N-acetilglucosaminiltransferase II recombinante de tal modo que elas são expressos como polipeptídeos separados. Em tais

5 modalidades, o polinucleotídeo pode conter um local de entrada de ribossoma interno que permite a tradução em separado de cada domínio catalítico do polinucleotídeo. Em tais modalidades, a expressão do domínio catalítico de N-acetilglucosaminiltransferase I recombinante é controlada por meio de um promotor que está operativamente ligado à porção do polinucleotídeo,

10 que codifica o domínio catalítico de N-acetilglucosaminiltransferase I e a expressão da proteína recombinante de domínio catalítico da N-acetilglucosaminiltransferase II é controlada por meio de um promotor que está operativamente ligado à porção do polinucleotídeo, que codifica o domínio catalítico de N-acetilglucosaminiltransferase II. O promotor pode ser um promotor

15 constitutivo ou um promotor indutível. Em certas modalidades preferidas, o promotor é um promotor indutível, tal como o promotor *cbh1* indutível.

Conforme descrito na presente invenção, N-acetilglucosaminiltransferase I (GlcNAc-TI ; GnTI ; CE 2.4.1.101) catalisa a reação de UDP-N-acetil-D-glucosamina + 3 - (alfa-D-manosila)-beta-D- manosil-R \rightleftharpoons UDP +

20 3 - (2 - (N-acetil-beta-D-glucosaminila)-alfa-D-manosila)-beta-D-manosil-R, onde R representa o restante do N-ligados oligossacarídeo no glicano aceitador. Um Domínio catalítico de N-acetilglucosaminiltransferase I é qualquer porção de uma enzima N-acetilglucosaminiltransferase I que é capaz de catalisar esta reação. As sequências de aminoácidos para as enzimas N-

25 acetilglucosaminiltransferase I a partir de vários organismos está listada em SEQ ID NOs : 1-19. GnTI enzimas adicionais estão listados na base de dados CAZy na família glicosiltransferase 13 (cazy.org/GT13_all). As espécies enzimaticamente caracterizado inclui *A. thaliana* AAR78757.1 (US6 653 459), de *C. elegans* AAD03023.1 (Chen S. et al J. Biol 1999, 274 (1) : 288-

30 97), *D. melanogaster* AAF57454. 1 (Sarkar e Schachter Biol Chem fevereiro 2001,. 382 (2) : 209 a 17) ; *C. griseus* AAC52872.1 (Puthalakath H. et al J. Biol 1996 271 (44) : 27818 a 22) ; *H. sapiens* AAA52563.1 (Kumar R. et al

Proc Natl Acad Sci EUA A. dezembro de 1990, 87 (24) : 9948 a 52) ; *M. auratus* AAD04130.1 (Opat Como et al Biochem J. dezembro de 1998 15 ; 336 (Pt 3) : 593 a 8), (incluindo por exemplo, desativar as mutante), coelho, *O. cuniculus* AAA31493.1 (Sarkar M. et al Proc Natl Acad Sci EUA 01 de janeiro 1991 A., 88 (1.) : 234 a 8). Exemplos adicionais de enzimas ativas caracterizadas podem ser encontrados em cazy.org/GT13_characterized. A estrutura 3D do domínio catalítico de GnTI de coelho foi definida por cristalografia de raios X em Unligil UM et al. EMBO J. 2000 16 de outubro, 19 (20): 5269-80. Os dados bancários (APO) estruturas de proteínas para GnTI são 1FO8, 1FO9, 1FOA, 2AM3, 2AM4, 2AM5 e 2APC. Em certas modalidades, o domínio catalítico de N-acetilglucosaminiltransferase I é da enzima humana N-acetilglucosaminiltransferase I (SEQ ID NO : 1), ou as suas variantes. Em certas modalidades, o domínio catalítico de N-acetilglucosaminiltransferase I contém uma sequência que é pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica aos resíduos de aminoácidos 84-445 da SEQ ID NO : 1. Em algumas modalidades, uma sequência mais curta pode ser usada como um domínio catalítico (por exemplo, resíduos de aminoácidos 105-445 da enzima humana, ou os resíduos de aminoácidos 107-447 da enzima de coelho, Sarkar et al (1998) Glicoconjugate J 15 : 193. -197). As sequências adicionais que podem ser usadas como o domínio catalítico de GnTI inclui os resíduos de aminoácidos desde cerca dos 30 a 445 aminoácido da enzima humana ou de qualquer domínio da haste do terminal C começando entre resíduo de 30 a 105 aminoácidos e continuando até cerca dos 445 aminoácidos do enzima humana, ou a correspondente sequência homóloga de outro GnTI ou uma variante cataliticamente ativo ou mutante. O domínio catalítico pode incluir as partes Terminal N da enzima, tal como a totalidade ou parte do domínio da haste, o domínio transmembranar, ou o domínio citoplasmático.

Conforme aqui descrito, a N-acetilglucosaminiltransferase II (GlcNAc-TII ; GnTII ; CE 2.4.1.143) catalisa a reação de UDP-N-acetil-D-glucosamina + 6 - (alfa-D-manosila)-beta-D- manosil-R < = > + UDP 6 - (2 -

(N-acetil-beta-D-glucosaminila)-alfa-D-manosila)-beta-D-manosil-R, onde R representa o restante do N-ligados oligossacarídeo no glicano aceitador. Um II domínio catalítico de N-acetilglucosaminiltransferase é qualquer porção de uma enzima N-acetilglucosaminiltransferase II, que é capaz de catalisar esta

5 reação. As sequências de aminoácidos para as enzimas N-acetilglucosaminiltransferase II a partir de vários organismos está listada em SEQ ID NOs : 20 a 33. Em certas modalidades, o domínio catalítico de N-acetilglucosaminiltransferase II é a enzima humana N-acetilglucosaminiltransferase II (SEQ ID NO : 20), ou as suas variantes. As espécies GnTII adicionais estão lista-

10 das na base de dados cazy na família glicosiltransferase 16 (cazy.org/GT16_all). As espécies enzimaticamente caracterizadas incluem GnTII de *C. elegans*, *D. melanogaster*, o *Homo sapiens*, *Rattus norvegicus*, *Sus scrofa* (cazy.org/GT16_characterized). Em certas modalidades, o domínio catalítico de N-acetilglucosaminiltransferase II contém uma sequência

15 que é pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 % , pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica aos resíduos de aminoácidos a partir de cerca de 30 a 447 da SEQ ID NO : 21. O domínio catalítico podem incluir partes Terminal N da enzima, tal como a totalidade

20 ou parte do domínio da haste, o domínio transmembranar, ou o domínio citoplasmático.

Em modalidades em que a célula fúngica filamentosa contém uma proteína de fusão da presente invenção, a proteína de fusão pode ainda conter um espaçador entre a N-acetilglucosaminiltransferase I e domínio ca-

25 talítico Domínio catalítico de N-acetilglucosaminiltransferase II. Qualquer um dos espaçadores da presente invenção, tal como descrito na seção intitulada "espaçadores" pode ser usado. Em certas modalidades preferidas, o espaçador é um espaçador EGIV, um espaçador 2xG4S, um espaçador 3xG4S, ou um espaçador de cbhl. Em outras modalidades, o espaçador contém uma

30 sequência de um domínio de haste.

Para a expressão dos RE / Golgi a N-acetilglucosaminiltransferase I /e/ou o domínio catalítico de N-acetilglucosaminiltransferase II é tipi-

camente fundida com um peptídeo alvo ou uma parte de uma proteína ER ou Golgi inicial, ou expressa com um ER endógeno alvo das estruturas de um animal ou vegetal enzima N-acetilglucosaminiltransferase. Em certas modalidades preferidas, a N-acetilglucosaminiltransferase I /e/ou o domínio catalítico de N-acetilglucosaminiltransferase II contém qualquer dos peptídeos objetivos da presente invenção, tal como descrito na seção intitulada "peptídeos de direcionamento". De preferência, o peptídeo alvo é ligado à extremidade do terminal N do domínio catalítico. Em algumas modalidades, o peptídeo alvo contém qualquer um dos domínios estaminais da presente invenção, tal como descrito na seção intitulada "peptídeos de direcionamento". Em certas modalidades preferidas, o peptídeo alvo é um peptídeo alvo Kre2. Em outras modalidades, o peptídeo alvo contém ainda um domínio transmembranar ligada à extremidade Terminal N do domínio de haste ou de um domínio citoplasmático ligado à extremidade Terminal N do domínio da haste. Em modalidades em que o peptídeo alvo contém ainda um domínio transmembranar, o peptídeo alvo pode ainda conter um domínio citoplasmático ligado à extremidade Terminal N do domínio transmembranar.

O nível de expressão de um gene da presente invenção alg3 pode ser reduzido através de qualquer método adequado conhecido na técnica, incluindo, sem limitação, a mutação do gene alg3. O alg3 pode ser mutado através de, por exemplo, mutações pontuais ou deleção de todo o gene alg3. De preferência, a função da proteína alg3 é reduzida ou eliminada mediante a mutação de alg3. O gene codifica uma dolichil alg3-P-Man : Man (5) GlcNAc (2)-PP-dolichil alfa-1,3-manosiltransferase . Tal como descrito na presente invenção, um dolichil-P-Man : Man (5) GlcNAc (2)-PP-dolichil manosiltransferase da presente invenção transfere um resíduo de alfa-D-manosila dolichil-fosfato a partir de D-manose em uma membrana lipídica ligada ao oligosacarídeo.

Em certas modalidades, a célula fúngica filamentosa pode conter um polinucleotídeo que codifica um veículo de UDP-GlcNAc. O polinucleotídeo que codifica o veículo de UDP-GlcNAc podem ser endógenos (isto é, de uma forma natural presente) na célula fúngica filamentosa, ou pode ser hete-

rólogo para a célula filamentosa fúngica.

Em outras modalidades, a célula fúngica filamentosa pode conter também um polinucleotídeo que codifica uma α -1,2-manosidase da presente invenção, tal como descrito na seção intitulada "células hospedeiras."

- 5 O polinucleotídeo que codifica o α -1,2-manosidase pode ser endógeno na célula fúngica filamentosa, ou pode ser heterólogo para a célula filamentosa fúngica. Estes polinucleotídeos são especialmente úteis para uma célula filamentosa fúngica expressando glicanos de manose elevada transferidos do Golgi para o ER, sem clivagem eficaz de α -2-manosidase. Para expressão citoplasmática do domínio catalítico da manosidase é tipicamente fundida com um peptídeo alvo, tais como HDEL, KDEL, ou parte de uma proteína ER ou Golgi inicial, ou expressa com um ER endógeno direcionamento das estruturas de uma enzima animal ou planta manosidase I.

- Em outras modalidades, a célula fúngica filamentosa também
- 15 pode conter um polinucleotídeo que codifica para uma galactosiltransferase da presente invenção, conforme descrito na seção intitulada "células hospedeiras." Galactosiltransferases transferem os resíduos de galactose- β ligados ao Terminal resíduo N-acetilglucosaminil. Em certas modalidades, galactosiltransferase é uma β -4-galactosiltransferase. A galactosiltransferase pode ser expressa no citoplasma do fungo filamentoso. Um peptídeo alvo heterólogo, tal como um peptídeo Kre2 Schwientek descrito no J. Biol. Chem 1996 3398, pode ser utilizado. Os promotores que podem ser utilizados para a expressão da galactosiltransferase incluem os promotores constitutivos tais como promotores de *gpd*, enzimas de glicosilação endógenas e glicosiltransferases, tais como manosiltransferases que sintetizam N-glicanos no Golgi ou ER, e promotores induzíveis de proteínas endógenas de elevado rendimento, tais como o promotor *cbh1*. Em modalidades da presente invenção em que a célula hospedeira contém um polinucleotídeo que codifica uma galactosiltransferase, a célula hospedeira contém também um polinucleotídeo que codifica uma UDP-Gal A /e/ou o veículo de UDP-Gal. Em certas modalidades da presente invenção em que a célula fúngica filamentosa contém um polinucleotídeo que codifica uma galactosiltransferase, a lactose po-
- 20
- 25
- 30

de ser usada como fonte de carbono, em vez de glicose, quando a cultura da célula filamentosa fúngica. O meio de cultura pode estar compreendido entre pH 4,5 e 7,0 ou entre 5,0 e 6,5. Em certas modalidades da presente invenção em que a célula fúngica filamentosa contém um polinucleotídeo que codifica uma galactosiltransferase e um polinucleotídeo que codifica uma UDP-Gal /e/ou UDP-Gal veículo, um cátion bivalente tal como Mn^{2+} , Ca^{2+} ou Mg^{2+} podem ser adicionados ao meio de cultura de células.

Em outras modalidades, a célula fúngica filamentosa pode conter também um polinucleotídeo que codifica uma sialiltransferase da presente invenção, como descrito na seção intitulada "células hospedeiras". A transferência sialiltransferases ao ácido siálico $\alpha 3$ ou $\alpha 6$ ligados, como Neu5Ac, a Gal terminal de complexos glicanos de galactosilação. O polinucleotídeo que codifica a sialiltransferase ligada $\alpha 3$ ou $\alpha 6$ pode ser endógeno à célula fúngica filamentosa ou pode ser heterólogo para a célula filamentosa fúngica. A sialilação na célula fúngica filamentosa pode requerer a expressão de enzimas que sintetizam o dador de ácido CMP-siálico, tais como CMP-Neu5Ac, especialmente em células de fungos, plantas, nematóides / parasita, ou inseto.

Além disso, a célula fúngica filamentosa pode ter os níveis de atividade de várias enzimas endógenas adicionais aumentados ou reduzidos. Um nível reduzido de atividade pode ser fornecido através da inibição da atividade da enzima endógena com um inibidor, um anticorpo, ou semelhantes. Em certas modalidades, a célula fúngica filamentosa é geneticamente modificada de forma a aumentar ou reduzir a atividade de uma ou mais enzimas endógenas. Os métodos para modificar geneticamente uma célula filamentosa fúngica para aumentar ou reduzir a atividade de uma ou mais enzimas endógenas são bem conhecidos na técnica e incluem, sem limitação, aqueles descritos na seção intitulada "células hospedeiras". Em certas modalidades, a célula fúngica filamentosa tem um nível reduzido de atividade de de alfa-1, 6-manosiltransferase em relação ao nível da atividade de uma célula fúngica filamentosa de tipo selvagem. Alfa-1,6-manosiltransferase (CE 2.4.1.232) no aparelho de Golgi transfere um alongamento iniciando resíduo

de alfa-D-manosila de GDP-manose em uma proteína ligada ao N-glicano oligosacarídeo, formando uma alfa-(1 -> 6)-D-manosil-D-manose de ligação. Tipicamente, a enzima alfa-1,6-manosiltransferase é codificada por um gene OCH1. Em certas modalidades, a célula filamentosa fúngica tem um nível

5 reduzido de expressão de um gene OCH1 em comparação com o nível de expressão em uma célula fúngica filamentosa de tipo selvagem. Em certas modalidades, o gene OCH1 é eliminado a partir da célula filamentosa fúngica.

A célula fúngica filamentosa pode ser, por exemplo, uma célula *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*,
 10 *Chrysosporium lucknowense*, *Filibasidium*, *Fusarium*, *Gibberella*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium* ou *Trichoderma*. Em certas modalidades, a célula filamentosa fúngica é uma célula *T. reesei*.

15 Composições farmacêuticas contendo o complexo de N-glicanos produzidos por meio dos métodos da invenção

Em outro aspecto, a presente invenção fornece uma composição, por exemplo, uma composição farmacêutica, contendo um ou mais N-glicanos complexos ligados a uma molécula heteróloga produzida or meio

20 dos métodos da presente invenção, formulados em conjunto com um veículo farmacêuticamente aceitável. As composições farmacêuticas da presente invenção também podem ser administradas em terapia de combinação, ou seja, combinadas com outros agentes. Por exemplo, a terapia de combinação pode incluir um complexo de N-glicanos ligado a uma molécula heteró-

25 loga de acordo com a presente invenção, combinado com pelo menos um outro agente terapêutico.

Tal como usado na presente invenção, o termo "veículo farmacêuticamente aceitável" inclui qualquer e todos os solventes, meios de dispersão, revestimentos, agentes antibacterianos e antifúngicos, agentes iso-

30 tônicos e retardadores da absorção, e similares que sejam fisiologicamente compatíveis. De preferência, o veículo é adequado para administração intravenosa, intramuscular, subcutânea, administração espinal ou epidérmica

parentérica (por exemplo, por injeção ou infusão). Dependendo da via de administração, o composto ativo, isto é, o complexo de N-glicanos ligado a uma molécula heteróloga de acordo com a presente invenção, pode ser revestido com um material para proteger o composto da ação de ácidos e outras condições naturais que podem inativar o composto.

As composições farmacêuticas da presente invenção podem incluir um ou mais sais farmaceuticamente aceitáveis. Um "sal farmaceuticamente aceitável" refere-se a um sal que retém a atividade biológica desejada do composto parental e não confere quaisquer efeitos toxicológicos indesejáveis (vide, por exemplo, Berge, SM, et al (1977) J. Pharm Sci. 66 : 1 a 19). Exemplos de tais sais incluem sais de adição de ácidos e sais de adição de base. Os sais de adição de ácido incluem aqueles derivados de ácidos inorgânicos não tóxicos, tais como clorídrico, nítrico, fosfórico, sulfúrico, bromídrico, iodídrico, fosforoso e similares, bem como a partir de ácidos orgânicos não tóxicos tais como ácidos mono e dicarboxílicos alifáticos, fenil-alcanóico substituído ácidos, ácidos hidroxil alcanóicos, ácidos aromáticos, alifáticos e ácidos sulfônicos aromáticos e similares. Os sais de adição de base incluem aqueles derivados de metais alcalino-terrosos, tais como sódio, potássio, magnésio, cálcio e similares, bem como a partir de aminas orgânicas não tóxicas, tais como N, N'-dibenziletlenodiamina, N-metilglucamina, cloropropil, colina, dietanolamina, etilenodiamina, procaína e similares.

Uma composição farmacêutica da presente invenção também pode incluir um anti-oxidante farmaceuticamente aceitável. Exemplos de antioxidantes farmaceuticamente aceitáveis incluem: (1) antioxidantes solúveis em água, tais como ácido ascórbico, cloridrato de cisteína, bissulfato de sódio, metabissulfato de sódio, sulfato de sódio e similares, (2) antioxidantes solúveis em óleo, tais como palmitato de ascorbilo, hidroxianisole butilado (BHA), hidroxitolueno butilado (BHT), lecitina, gálico de propila, alfa-tocoferol, e similares, e (3) agentes quelantes de metais, tais como ácido cítrico, ácido etilenodiaminotetracético (EDTA), o ácido tartárico, o sorbitol, o ácido fosfórico, e similares.

Exemplos de veículos aquosos e não aquosos adequados que

podem ser empregues nas composições farmacêuticas da presente invenção incluem água, etanol, polióis (tais como glicerol, propileno glicol, polietileno glicol, e similares), e suas misturas adequadas, óleos vegetais, tal como o azeite, e ésteres orgânicos injetáveis, tais como oleato de etila. A fluidez
5 adequada pode ser mantida, por exemplo, pelo uso de materiais de revestimento, tais como lecitina, pela manutenção do tamanho de partícula requerido no caso das dispersões, e pelo u.

Estas composições podem também conter adjuvantes tais como conservantes, agentes molhantes, agentes emulsionantes e agentes disper-
10 santes. Prevenção da presença de microrganismos pode ser assegurada tanto pelos procedimentos de esterilização, e pela inclusão de vários agentes antibacterianos e antifúngicos, por exemplo, parabeno, clorobutanol, ácido fenolsórbico, e outros semelhantes. Pode também ser desejável incluir agentes isotônicos, tais como açúcares, cloreto de sódio, e outros semelhan-
15 tes nas composições. Além disso, a absorção prolongada da forma farmacêutica injetável pode ser provocada pela inclusão de agentes que retardam a absorção, tais como monoestearato de alumínio e gelatina.

Veículos farmacêuticamente aceitáveis incluem soluções aquo-
sas estéreis ou dispersões e pós estéreis para a preparação extemporânea
20 de soluções ou dispersões injetáveis estéreis. A utilização de tais meios e agentes para substâncias farmacêuticamente ativas é conhecida na arte. Exceto na medida em que qualquer meio ou agente convencional é incompatível com o composto ativo, a sua utilização nas composições farmacêuticas da presente invenção está contemplada. Os compostos ativos suple-
25 mentares podem também ser incorporados nas composições.

As composições terapêuticas devem ser tipicamente estéreis e estáveis sob as condições de fabrico e armazenamento. A composição pode ser formulada como uma solução, microemulsão, lipossoma, ou outra estrutura ordenada adequada para concentração de droga elevada. O veículo
30 pode ser um solvente ou meio de dispersão contendo, por exemplo, água, etanol, poliol (por exemplo, glicerol, propileno glicol, e polietileno glicol líquido, e similares), e suas misturas adequadas. A fluidez adequada pode ser

mantida, por exemplo, pelo uso de um revestimento tal como lecitina, pela manutenção do tamanho de partícula requerido no caso de dispersão e pela utilização de tensoativos. Em muitos casos, será preferível incluir agentes isotônicos, por exemplo, açúcares, polialcoóis, tais como, sorbitol, manitol ou cloreto de sódio na composição. A absorção prolongada das composições injetáveis pode ser provocada pela inclusão na composição de um agente que atrasa a absorção, por exemplo, sais de monoestearato e gelatina.

As soluções injetáveis estéreis podem ser preparadas por incorporação do composto ativo na quantidade requerida em um solvente apropriado com um ou uma combinação dos ingredientes enumerados acima, como requerido, seguido de esterilização por microfiltração. Geralmente, as dispersões são preparadas por incorporação do composto ativo em um veículo estéril que contém um meio de dispersão básico e os outros ingredientes necessários entre os enumerados acima. No caso de pós estéreis para a preparação de soluções injetáveis estéreis, os métodos de preparação de certos são a secagem por vácuo e liofilização (liofilização), que produzem um pó do ingrediente ativo mais qualquer ingrediente adicional desejado a partir de uma solução previamente esterilizada por filtração.

A quantidade de ingrediente ativo que pode ser combinada com um material veículo com a finalidade de produzir uma forma de dosagem única irá variar dependendo do indivíduo a ser tratado e do modo particular de administração. A quantidade de ingrediente ativo que pode ser combinada com um material veículo com a finalidade de produzir uma forma de dosagem individual será geralmente aquela quantidade da composição que produz um efeito terapêutico. Geralmente, fora dos cem por cento, esta quantidade irá variar desde cerca de 0,01 por cento a cerca de noventa e nove por cento de ingrediente ativo, de preferência entre cerca de 0,1 por cento a cerca de 70 por cento, mais preferencialmente desde cerca de 1 por cento a cerca de 30 por cento de ingrediente ativo em combinação com um veículo farmacêuticamente aceitável.

Os regimes de dosagem são ajustados para proporcionar a resposta desejada, ótima (por exemplo, uma resposta terapêutica). Por exem-

plo, um único bolus pode ser administrado, várias doses divididas podem ser administradas ao longo do tempo ou a dose pode ser proporcionalmente reduzida ou aumentada como indicado por meio das exigências da situação terapêutica. É especialmente vantajoso formular as composições parenterais em forma de dosagem unitária para facilidade de administração e uniformidade de dosagem. A forma de dosagem unitária tal como usado na presente invenção refere-se a unidades fisicamente discretas adequadas como dosagens unitárias para os sujeitos a serem tratados : cada unidade contém uma quantidade predeterminada de composto ativo calculada com a finalidade de produzir o efeito terapêutico desejado em associação com o veículo farmacêutico requerido. A especificação para as formas de dosagem unitárias da presente invenção são ditadas por e directamente dependentes (a) das características únicas do composto ativo e do efeito terapêutico particular a ser alcançado, e (b) as limitações inerentes na técnica compõem um tal composto ativo para o tratamento da sensibilidade nos indivíduos.

Para a administração do complexo de N-glicanos ligados a uma molécula heteróloga, em particular, onde a molécula heteróloga é um anticorpo, a dosagem varia de cerca de 0,0001 a 100 mg / kg, e mais usualmente de 0,01 a 5 mg / kg, de o peso corporal do hospedeiro. Por exemplo, as dosagens podem ser de 0,3 mg / kg de peso corporal, de 1 mg / kg de peso corporal, de 3 mg / kg de peso corporal, de 5 mg / kg de peso corporal ou 10 mg / kg de peso corporal ou no intervalo de 1 a 10 mg / kg. Um regime de tratamento exemplar envolve administração uma vez por semana, uma vez a cada duas semanas, uma vez a cada três semanas, uma vez a cada quatro semanas, uma vez por mês, uma vez a cada 3 meses ou uma vez a cada três a seis meses. Certos esquemas de dosagem para um complexo de N-glicanos ligados a um anticorpo heterólogo incluem 1 mg / kg de peso corporal ou de 3 mg / kg de peso corporal por administração intravenosa, com o anticorpo a ser administrado utilizando um dos seguintes programas de dosagem : (i) a cada quatro seis semanas de dosagens, em seguida, a cada três meses, (ii) a cada três semanas, (iii) 3 mg / kg de peso corporal, uma vez seguido por 1 mg / kg de peso corporal de três em três semanas.

Em alternativa, um complexo de N-glicanos ligados a uma molécula heteróloga de acordo com a presente invenção pode ser administrada como uma formulação de liberação controlada, caso em que é necessária uma administração menos frequente. A dosagem e frequência variam dependendo da semi-vida da substância administrada no paciente. Em geral, os anticorpos humanos mostra a meia-vida mais longa, seguido por meio dos anticorpos humanizados, anticorpos quiméricos, e anticorpos não humanos. A dosagem e frequência de administração pode variar dependendo de se o tratamento é profilático ou terapêutico. Em aplicações profiláticas, uma dosagem relativamente baixa é administrada em intervalos relativamente pouco frequentes ao longo de um longo período de tempo. Alguns pacientes continuam a receber tratamento para o resto de suas vidas. Em aplicações terapêuticas, uma dosagem relativamente elevada em intervalos relativamente curtos por vezes é necessário, até que a progressão da doença seja reduzida ou terminada, e de preferência até que o paciente apresenta melhoria parcial ou completa dos sintomas da doença. Depois disso, o paciente pode ser administrado em um regime profilático.

Os níveis de dosagem reais dos ingredientes ativos nas composições farmacêuticas da presente invenção podem ser variados de modo a obter uma quantidade do ingrediente ativo que é eficaz para alcançar a resposta terapêutica desejada para um paciente em particular, composição e modo de a administração, sem ser tóxico para o paciente. O nível de dosagem selecionado a partir dependerá de uma variedade de fatores, incluindo a atividade de farmacocinética das composições particulares da presente invenção utilizado, ou do éster, sal ou amida, a via de administração, o tempo de administração, a taxa de excreção do composto particular a ser utilizado, a duração do tratamento, outros fármacos, compostos /e/ou materiais utilizados em combinação com as composições particulares empregues, a idade, sexo, peso, condição, saúde geral e história médica anterior do paciente a ser tratado, e fatores semelhantes bem conhecidos nas artes médicas.

A "dose terapeuticamente eficaz" de imunoglobulina da presente

invenção de preferência resulta em uma diminuição na gravidade dos sintomas da doença, um aumento na frequência e duração dos períodos de doença sem sintomas, ou a prevenção da deficiência ou incapacidade devido à aflição doença . Por exemplo, para o tratamento de tumores, uma "dose terapêuticamente eficaz" de preferência inibe o crescimento celular ou o crescimento do tumor em, pelo menos, cerca de 20 %, mais preferencialmente em pelo menos cerca de 40 %, ainda mais preferencialmente em pelo menos cerca de 60 %, e ainda mais preferencialmente em pelo menos cerca de 80 % relativamente a sujeitos não tratados. A capacidade de um composto para inibir o crescimento do tumor pode ser avaliada em um sistema de modelo animal preditivo de eficácia em tumores humanos. De uma maneira alternativa, esta propriedade de um composto pode ser avaliada examinando a capacidade do composto para inibir essa inibição in vitro por meio dos ensaios conhecidos por aqueles que são versados na técnica. Uma quantidade terapêuticamente eficaz de um composto terapêutico pode diminuir o tamanho do tumor, ou de outra forma melhorar os sintomas em um sujeito. Um versado na técnica seria capaz de determinar tais quantidades com base em fatores tais como o tamanho do sujeito, a gravidade dos sintomas do sujeito, e da composição particular ou da via de administração selecionada.

A composição da presente invenção pode ser administrada através de uma ou mais vias de administração, utilizando uma ou mais de uma variedade de métodos conhecidos na técnica. Como será apreciado por uma pessoa que é versada na técnica, a via /e/ou modo de administração irão variar dependendo dos resultados desejados. Certas vias de administração para porções de ligação da presente invenção incluem injeção intravenosa, intramuscular, intradérmica, intraperitoneal, subcutânea, espinal ou outras vias parentéricas de administração, como por exemplo por meio da injeção ou por meio da infusão. A frase "administração parentérica", tal como usado na presente invenção, significa os diferentes modos de administração da administração entérica e tópica, usualmente por meio da injeção, e inclui, sem limitação, injeção intravenosa, intramuscular, intra-arterial, intratecal, intracapsular, intraorbital, intracardiaca, intradérmica, intraperitoneal, trans-

traqueal, subcutânea, subcuticular, intra-articular, subcapsular, subaracnóide, a injeção intra-espinhal, epidural e intrasternal e infusão.

Em alternativa, um complexo de N-glicanos ligados a uma molécula heteróloga de acordo com a presente invenção podem ser administrados através de uma via nonparenteral, tal como uma via tópica, epidérmica ou mucosa da administração, por exemplo, por via intranasal, oral, vaginal, retal, sublingual ou tópica.

Os compostos ativos podem ser preparados com veículos que irão prbateladager o composto contra a liberação rápida, tal como uma formulação de liberação controlada, incluindo implantes, adesivos transdérmicos e sistemas de entrega microencapsulados. Os polímeros biodegradáveis, biocompatíveis, podem ser utilizados, tais como acetato de etileno vinila, polianidridos, ácido poliglicólico, colagénio, poliortoésteres e ácido poliláctico. Muitos métodos para a preparação de tais formulações são patenteados ou geralmente conhecidos por meio daqueles que são versados na técnica. Vide, por exemplo, Sistemas de Entrega de Fármaco de liberação sustentada e controlada, JR Robinson, ed., Marcel Dekker, Inc., New York, 1978.

As composições terapêuticas podem ser administradas com dispositivos médicos conhecidos na técnica. Por exemplo, em uma determinada modalidade, uma composição terapêutica da presente invenção pode ser administrada com um dispositivo de injeção sem agulha hipodérmica, tal como os dispositivos descritos na Patente U.S. N ° s 5.399.163, 5.383.851, 5.312.335, 5.064.413, 4.941.880, 4.790.824, ou 4.596.556. Exemplos de implantes bem conhecidos e módulos úteis na presente invenção incluem: Patente U.S. No. 4.487.603, a qual descreve uma bomba de micro-infusão implantável para a distribuição de medicação a uma taxa controlada, Patente U.S. N ° 4.486.194, que descreve um dispositivo terapêutico para administrar os medicamentos através da pele ; Patente U.S. No. 4.447.233, que descreve uma bomba de infusão de medicação para o fornecimento de medicamentos, a uma taxa de infusão precisa ; Patente U.S. No. 4.447.224, que descreve um aparelho de infusão de caudal variável para fornecimento de fármacos implantáveis contínuos, Patente U.S. N ° 4.439.196, que descreve

um sistema de entrega osmótica de fármaco tendo compartimentos de câmaras múltiplas, e Patente U.S. No. 4.475.196, que descreve um sistema osmótico de entrega de fármacos.

Em certas modalidades, a utilização do complexo de N-glicanos ligados a uma molécula heteróloga de acordo com a presente invenção é para o tratamento de qualquer doença que pode ser tratada com os anticorpos terapêuticos.

É para ser compreendido que, embora a presente invenção tenha sido descrita em conjunção com as certas modalidades específicas do mesmo, a descrição anterior destina-se a ilustrar e não limitar o âmbito da presente invenção. Outros aspectos, vantagens e modificações dentro do âmbito da presente invenção serão evidentes para aqueles que são versados na técnica à qual pertence a presente invenção.

A presente invenção tem sido descrita, são oferecidos os seguintes exemplos para ilustrar a presente invenção por meio de ilustração, não como forma de limitação.

EXEMPLOS

Exemplo 1 - Seleção de Cepas Hospedeiras para Glicoengenharia

O objetivo deste exemplo foi de identificar as cepas de *T. reesei* para glicoengenharia. Uma cepa ideal produz a quantidades elevadas de man5 N-glicanos e as baixas quantidades de glicanos ácidos.

Amostras

As diferentes cepas de *T. reesei*, incluindo M44 (VTT-D-00775 ; Selinheimo et al, FEBS J. 2006, 273 (18) : . 4322-35), M81, M84, M109, M110, M131, M132, M133, M134 e M124 (uma cepa mus53-excluído de M44) foram analisadas. Cada uma das dez cepas foi cultivada em culturas em frasco de agitação. As amostras foram colhidas em três pontos diferentes de tempo: 3 dias, 5 dias, e 7 dias. Ambos os sobrenadantes (proteínas secretadas) e péletes de células foram coletados e armazenados congelados a -20 ° C até a análise de glicanos ter sido conduzida.

N-glicanos foram isolados de proteínas segregadas a partir dos pontos de tempo indicados, seguido por meio da dessorção de laser assisti-

da por meio da matriz / ionização-tempo-de-voo (MALDI-TOF) de perfis de glicano. Os péletes de células do ponto de tempo de 5 dias foram submetidos a caracterização de N-glicanos. Um total de 80 amostras (30 de cada uma das frações de sobrenadante neutras e ácidas, e 10 de cada uma das

5 frações de péletes neutros e ácidos) foi submetido à análise.

A cepa M44 também foi submetida à batelada e cultivo de fermentação em batelada alimentada a fim de avaliar a diferença no perfil de glicanos entre frasco agitado e a cultura do fermentador. Para a análise de glicano, as amostras dos três pontos de tempo diferentes foram analisadas

10 para um total de 12 amostras (6 frações neutras e 6 frações ácidas). Como um controle, meio de cultura foi analisado.

Métodos de Espectrometria de Massa

A espectrometria de massa MALDI-TOF foi realizada com instrumento Ultraflex TOF / TOF Bruker (Bruker Daltonics, Alemanha). Os N-

15 glicanos neutros foram detectados no modo de refletor de íons positivos como íons $[M + Na]^+$, e os N-glicanos ácido foram detectados em modo linear como os íons negativos $[MH]^-$. A abundância relativa de componentes molar de N-glicanos neutros foi atribuída com base nas suas intensidades relativas de sinais no espectro. Os sinais resultantes glicano nos perfis glicanos

20 apresentados foram normalizadas para 100 % tendo em vista permitir a comparação entre amostras.

Métodos de Glicosilação de Proteínas Específicas

As proteínas a partir de uma amostra fermentador-cultivada foram separadas por meio de SDS-PAGE e riscadas para uma membrana de

25 PVDF. As bandas de proteínas de interesse foram excisadas e N-glicanos foram liberados por meio da liberação enzimática com PNGase F.

Perfil de N-glicanos Neutros de Cepas *T. reesei*

A estrutura man5 desejada pode ser observada como um sinal $[M + Na]$ para m/z 1257,4 valor de no espectro de massa apresentada na

30 Figura 1. O glicoma neutro das cepas de *T. reesei* analisadas foram encontradas para ter ou man5 ou Man8 como as principais espécies de glicano neutras (H5N2 e H8N2 na Tabela 2).

Tabela 2: A percentagem de diferentes sinais de N-glicanos neutros de cepas de *T. reesei* analisadas.

Composição	m/z	Cepa									
		M44	M81	M84	M109	M110	M131	M132	M133	M134	M124
		%	%	%	%	%	%	%	%	%	%
H3N2	933	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H4N2	1095	1.2	0.0	0.0	1.8	0.9	0.0	2.3	0.0	2.3	4.1
H5N2	1257	81.0	70.8	4.0	78.9	5.8	78.8	84.1	10.7	73.2	77.9
H6N2	1419	5.8	5.3	0.0	5.3	0.9	4.8	4.6	0.9	6.0	7.3
H7N2	1581	4.8	7.3	1.5	4.7	3.0	4.8	3.9	3.8	5.8	4.8
H8N2	1743	3.7	8.6	81.5	5.1	68.2	5.9	2.6	68.1	6.3	3.3
H9N2	1905	2.9	8.0	9.0	3.4	16.0	4.6	2.0	12.8	5.7	2.3
H10N2	2067	0.5	0.0	2.5	0.8	3.7	1.1	0.4	2.5	0.7	0.4
H11N2	2229	0.0	0.0	1.5	0.0	1.4	0.0	0.0	1.2	0.0	0.0
H12N2	2391	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Alguns N-glicanos ácidos foram observados nas frações N-glicanos neutros. Isto pode ter sido devido às propriedades específicas dos glicanos fosforilados, por exemplo, a presença de estruturas fosfodiéster, ou outras propriedades do fosfoglicanos o que poderia levar a fugas de espécies ácidas à fração neutra por meio das condições experimentais usadas neste estudo. Para verificar a estrutura correspondente, o sinal de interesse foi submetido a análise por meio de MS / MS. A fragmentação por meio da espectrometria de massa de glicanos foi realizada utilizando Bruker Ultraflex TOF / TOF no modo de análise por MS / MS (Figura 2). Uma vez que os glicanos não foram permetilados, não poderia ser obtida a atribuição estrutural definitiva com base nos dados MS/MS.

15 Perfis de N-glicanos Ácidos de Cepas *T. reesei*

Para fins de glicoengenharia foi útil ter cepas com uma quantidade mínima de N-glicanos ácido. Portanto, os perfis de ácidos N-glicanos foram analisados a partir das cepas utilizadas para o rastreio. O espectro de N-glicanos ácido das cepas analisadas é mostrado na Figura 3 e na Tabela 3 abaixo.

Tabela 3 : A percentagem de diferentes sinais de N-glicano ácidos de cepas de *T. reesei* analisadas.

		M44	M81	M84	M109	M110	M131	M132	M133	M134	M124
	m/z	%	%	%	%	%	%	%	%	%	%
Hex3HexNAc2SP	989	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex4HexNAc2SP	1151	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex5HexNAc2SP	1313	4.0	5.2	0.0	3.7	0.0	2.8	7.4	0.0	5.2	2.8
Hex5HexNAc2SP2	1393	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.9	0.0
Hex6HexNAc2SP	1475	23.7	27.3	2.2	18.1	2.1	22.4	21.0	3.9	24.9	26.3
Hex6HexNAc2SP2	1555	0.0	2.8	0.0	2.4	0.0	3.2	1.1	0.0	3.6	1.7
Hex7HexNAc2SP	1637	30.3	18.8	1.1	16.2	2.0	14.9	24.7	0.0	17.2	23.3
Hex7HexNAc2SP2	1717	0.0	7.7	0.0	8.6	0.0	10.7	2.5	0.0	10.4	7.0
Hex8HexNAc2SP	1799	18.4	11.8	17.9	12.8	9.7	9.1	19.7	14.5	8.8	11.2
Hex8HexNAc2SP2	1879	5.1	8.8	0.0	11.0	0.0	14.8	4.0	0.0	12.4	10.0
Hex9HexNAc2SP	1961	7.3	6.4	49.1	9.5	37.9	5.9	6.1	53.9	4.1	3.5
Hex9HexNAc2SP2	2041	4.2	5.0	0.0	5.7	0.0	7.3	5.1	0.0	5.9	7.2
Hex10HexNAc2SP	2123	2.8	2.9	19.7	4.5	28.1	2.6	2.3	19.3	2.1	1.6
Hex10HexNAc2SP2	2203	2.8	2.1	0.0	2.2	0.0	2.7	3.6	0.0	1.9	3.3
Hex11HexNAc2SP	2285	1.5	1.3	3.7	2.1	9.5	1.2	0.9	5.0	1.0	0.8
Hex11HexNAc2SP2	2365	0.0	0.0	0.0	0.9	0.0	1.3	1.5	0.0	0.8	1.3
Hex12HexNAc2SP	2447	0.0	0.0	1.3	1.0	1.6	1.0	0.0	0.0	0.5	0.0
Hex12HexNAc2SP2	2527	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.3	0.0
Hex13HexNAc2SP	2609	0.0	0.0	1.2	0.4	1.1	0.0	0.0	0.0	0.0	0.0
Hex14HexNAc2SP	2771	0.0	0.0	0.6	0.0	0.9	0.0	0.0	0.0	0.0	0.0

Perfil de N-glicanos a partir da cepa M44 Cultivada em fermentador

A cepa M44 foi cultivada em um fermentador, a fim de descobrir se diferentes condições de cultivo podem causar mudanças no seu perfil de glicanos. A análise N-glicano foi realizada para as amostras cultivadas em um fermentador (Batelada ; 41 : 10, 88 : 45 e 112 : 50 horas, e batelada de alimentação ; 45 : 50, 131 : 40 e 217 : 20 horas) e comparada com a de agitar frascos cultivados. N-glicanos neutros e ácidos de proteínas secretadas de cepa *T. reesei* M44 em cultura em fermentador é apresentada na Figura 4. A comparação entre as percentagens de N-glicanos a partir de culturas de frascos e fermentador é apresentada abaixo na Tabela 4.

Tabela 4: A percentagem de sinais de N-glicano da cepa de *T. reesei* M44 cultivada em frasco e no fermentador.

Composição	m/z	Frasco %	Fermentador %
H3N2	933	0,0	0,0
H4N2	1095	1,2	0,0
H5N2	1257	81,0	91,3
H6N2	1419	5,8	4,5
H7N2	1581	4,8	4,2
H8N2	1743	3,7	0,0
H9N2	1905	2,9	0,0
H10N2	2067	0,5	0,0
H11N2	2229	0,0	0,0
H12N2	2391	0,0	0,0

Análise de N-glicanos de Vibração do Meio de Cultura do Frasco

Conforme uma experiência de controle, meio de cultura (sem contato com o fungo) de *T. reesei* foi analisado. A Figura 5a mostra a análise de N-glicanos neutros em que não foram observadas N-glicanos. Apenas
5 sinais menores de oligômeros de hexoses, provavelmente derivadas do material de planta utilizado no meio, eram visíveis acima da linha de base. Na Figura 5b (glicanos ácidos), não foram observados sinais correspondentes aos N-glicanos.

N-glicosilação de proteínas secretadas

10 Para verificar se existe uma variação de glicosilação entre as proteínas segregadas individuais e as amostras de sobrenadantes de cultura de fermentação foram separadas por meio de SDS-PAGE e riscadas para membrana de PVDF. Os N-glicanos de bandas selecionadas foram então destacados com a membrana de liberação enzimática. Os resultados são
15 mostrados nas Figuras 6 e 7.

Conclusões: glicanos neutros

O objetivo deste estudo foi identificar as cepas de *T. reesei* para glicoengenharia com a maior quantidade de N-glicanos man5 e menor quantidade de glicanos ácidos. As cepas que possuem man5 como um pico principal em análise por meio da espectrometria de massa pode ter maior atividade de α -1,2-manosidase endógeno. Com base na informação de base
20 sobre N-glicosilação de *T. reesei*, a estrutura provável para man5 é Man α 3 [Man α 3 (Man α 6) Man α 6] Man β 4GlcNAc β 4GlcNAc (Salovuori et al, 1987 ; Stals et al Glicobiologia 14, 2004, página 725).

25 Algumas cepas continham H8N2 como um importante glicoforma neutro. Com base na literatura, esta glicoforma é mais provável que seja uma estrutura Glc α 3Man α 2Man α 2Man5 (Stals et al. Glicobiology 14, 2004, página 725). É possível que a deficiência de glicosidase nestas cepas impeça o corte dos glicanos com as glicoformas menores.

30 Em algumas cepas, N-glicanos ácido foram observados nos espectros neutros. Esta situação pode ter sido devido a uma maior proporção de N-glicanos ácido ou a fugas de estruturas específicas para a fração neu-

tra durante a separação de glicanos neutros de glicanos ácidos.

O perfil de glicanos de cepas foi um pouco mais favorável para glicoengenharia quando cultivada em um fermentador em comparação com os frascos de agitação. A glicosilação de proteínas a partir de amostras individuais do fermentador de cultura não diferiu de uma forma significativa da glicosilação média. Todas as proteínas analisadas continham man5 como glicoforma principal. Esta observação sugeriu que todas as proteínas segregadas passariam pelo processamento glicano semelhante. Dessa maneira, verificou-se que a maioria das proteínas segregadas foram similarmente glicosiladas por meio das células hospedeiras de *T. reesei*, que nem sempre é o caso com as células de mamíferos.

Glicanos ácidos

A fosforilação de N-glicanos não é geralmente desejada por causa do resíduo de glicoengenharia do fosfato terminal não estar presente em proteínas terapêuticas regulares, incluindo os anticorpos. Algumas exceções a essa regra são algumas proteínas especializadas utilizadas para doenças de depósito lisossômico da glicosilação. A fosforilação de N-glicanos pode ser específica para a proteína em fungos. Em animais, a fosforilação de manose é um sinal de direcionamento lisossomal conservado.

Até o momento não houve relatos de sulfatação de *T. reesei* de N-glicanos. Portanto, as estruturas de ácidos referidas neste relatório eram susceptíveis de ser glicanos fosforiladas.

A fosforilação é mais comum quando *T. reesei* é cultivada em valores baixos de pH, como é o caso de culturas em frascos, que podem estar relacionadas com o stress do pH baixo e quebra de micélio (Stals et al, 2004, Glicobiology 14 : 713. - 724). Neste estudo observou-se uma clara diferença entre frasco e as amostras cultivadas em fermentador. N-glicanos ácido, todos fosforilados, não foram observados em amostras de cultura em frasco de agitação. A quantidade de N-glicanos ácido nas amostras do fermentador pode ter sido inferior ao limite de detecção, ou, por causa de um pH mais elevado que pode ter havido nenhuma fosforilação significativa de glicanos. A proporção de N-glicanos ácidos para a quantidade total de N-

glicanos não pôde ser verificada com o método utilizado no presente estudo, devido às diferentes eficiências de ionização entre as espécies de glicanos neutros e ácidos.

A fim de determinar os níveis de fosforilação, os N-glicanos foram liberados por meio de N-glicanase de 10 µg da batelada secretada cultivada de *T. reesei* em batelada e fermentador em batelada de alimentação. A concentração de bateladas foi medida utilizando um método baseado em Bradford com BSA como padrão. Um pmol de molécula NeuAcHex4HexNAc2 padrão foi adicionado às amostras de N-glicanos ácidos antes da análise de MALDI-TOF. Os valores das principais glicoformas (Hex7HexNAc2P para fermentador e Hex6-8HexNAc2P para a cultura frasco) foram 0,9 pmol/10 mg de proteína secretada da cultura em batelada de 0,6 pmol/10 mg de proteína secretada da cultura em batelada alimentada, e 160 mg de pmol/10 secretada proteína da cultura em frasco, quando o pH da cultura foi deixado cair. A quantidade de N-glicanos neutros foi medida utilizando 10 pmol de padrão glicano Hex2HexNAc4 adicionado em amostras de N-glicanos neutros, antes da análise de MALDI-TOF. A quantidade de grandes glicoformas Hex5HexNAc2 pmol/10 foi de 87 mg de proteína secretada em batelada e culturas em batelada de alimentação e 145 pmol/10 mg de proteína secretada em cultura frasco. Dessa maneira, a proporção de ácidos N-glicanos a quantidade total de N-glicanos foi de 1 %, em cultura descontínua, 0,7 % em culturas em batelada de alimentação e 52 % no frasco de cultura. Quantificação foi baseada apenas na comparação intensidade do sinal usando dados MALDI-TOF.

N-glicanos também foram maiores na fração ácida. Isto pode ter sido devido a reações de fosfo-manosilação em que o fósforo com uma unidade de hexose é ligado a uma estrutura principal de glicano. Algumas estruturas difosforiladas de espectros foram observadas em ácido. Esta explicação está de acordo com os dados publicados anteriormente sobre glicanos fosforilados encontrados em *T. reesei* (Stals et al. 2004, Glicobiologia 14 : 725 a 737). Quando cultivada em um fermentador, a proporção de N-glicanos ácido foi muito baixa, abaixo do limite de detecção.

Os espectros de N-glicanos dos meios de cultura de *T. reesei*

não revelaram contaminação dos N-glicanos glicoma de *T. reesei* com derivados de material vegetal contendo meio.

Em conclusão, a análise de N-glicanos de diferentes cepas de *T. reesei* revelou que a maior glicoforma em cepas M44, M109, M131, M132 e M124 é man5 ou Man α 3 [Man α 3 (Man α 6) Man α 6] Man β 4GlcNAc β 4GlcNAc. A possível presença de glicose, incluindo H8N2 como um componente minoritário em cepas produtoras de man5 foi considerada. Duas cepas (M109 e M131) continham uma quantidade maior de H8N2 de H7N2. O enriquecimento de H8N2 poderia indicar deficiência parcial glicosidase.

A cepa M44 continha quase nenhum glicano fosforilado. Os glicanos ácidos observados na fração de glicanos neutros como os sinais a m / z 1521 e m / z 1683 foram observados em amostras de cepas M131, M109, M132 e M124, que indicaram os níveis mais elevados de fosforilação e a presença de estruturas fosfodiéster parciais.

O objetivo deste estudo foi o de encontrar uma cepa com a produção máxima de Man5Gn2 estrutura e produção de baixo nível de N-glicanos ácidos (fosforilados). As melhores linhagens teve mais de 80 % de man5 sob apertar as condições de cultura do frasco pH controlados. As melhores linhagens também haviam reduzido a produção de glicanos difosforilados /e/ou maiores estruturas fosforilada (vide a Tabela 3).

Exemplo 2 - Geração de uma Cepa de *Trichoderma Alg3* deficiente

Construção de Vetor e Geração de Cepa

O gene que codifica o ALG3 manosiltransferase foi identificado na sequência do genoma de *Trichoderma reesei*. Um construto de ruptura foi concebida com a finalidade de inserir o marcador de seleção acetamidase entre 1000 bp 5' e 3' que flanqueiam os fragmentos da região do gene *alg3*. Os fragmentos da região flangeadora foram amplificados por meio de PCR, e o construto foi feito por meio da clonagem de recombinação homóloga em *Saccharomyces cerevisiae*. O cassete de interrupção foi liberado da sua estrutura principal do vetor por meio da digestão e transformado na cepa *T. reesei* M124. Os transformantes foram selecionados em meio de acetamidase e rastreados por meio de PCR com um iniciador para a frente fora da 5' do

fragmento da região flanqueando o construto e com o iniciador inverso dentro do marcador de seleção amdS.

Triagem de transformantes

Cinquenta e oito de 62 transformantes selecionados deram um
5 produto de PCR de tamanho esperado para a integração do construto ao lugar alg3. Nove transformantes positivas para a PCR foram purificados através de clones individuais mononuclear culturas de esporos, e de suspensões de esporos foram feitas a partir deles. Estes nove clones foram analisados para a integração correcta do cassete de interrupção por hibridação de
10 Southern. DNA genômico digerido com EcoRI a partir da cepa parental a partir de nove clones e foi hibridado com uma sonda alg3 sob condições de hibridação convencionais. A sonda hibridou com o DNA a partir da cepa parental, mas não com o DNA a partir de qualquer um dos clones, indicando supressão bem sucedida de alg3 (Figura 8).

15 Uma análise mais aprofundada foi feita por meio da hibridização Southern com uma sonda AMDS. O gene amdS foi incluído no cassete de supressão e foi previsto para ser detectável no DNA a partir dos transformantes, mas não no DNA a partir da cepa parental. O DNA genômico da cepa parental M124 e nove transformantes foram digeridos com EcoRI +
20 PvuI (E + P) + NheI e KpnI (K + N). NotI digeriu o plasmídeo transportando o cassete de deleção alg3-amdS que foi utilizado como um controle positivo. A sonda reconheceu o esperado fragmento de ~ 2,7 kb (AMD) a partir do controle positivo, mas não hibridizam com a cepa parental. Todos os transformantes deram os sinais esperados (1,6 + 2,8 kb para E + P e 1,7 + 3,4 kb
25 para K + N, mostradas com setas na Figura 9B), indicando a integração correcta do cassete de eliminação. Clones 11A e 15A mostraram também de alguns fragmentos de hibridação adicionais, sugerindo a integração inespecífica do cassete de eliminação para o genoma (Figura 9B).

N-glicanos Analíticos

30 As culturas misturadas em frasco de cinco linhagens knockout diferentes Alg3 (4A, 5A, 6A, 10A e 16A) e cepa dos pais M124 foram analisadas para N-glicanos. As amostras foram coletadas a partir de pontos de

tempo de 3, 5, 7 e 9 dias. Todas as culturas foram cultivadas em duplicado.

A concentração de proteína das proteínas segregadas a partir de uma cepa selecionada aleatoriamente knockout (4A) de todos os pontos de tempo foi medida usando um ensaio baseado em Bradford contra uma curva padrão de BSA. A concentração de proteína mais elevada foi detectada no dia 5. Dessa maneira, as amostras de 5 dias foram utilizadas para análise de N-glicanos para todas as cinco cepas knockout. Todas as amostras, incluindo as culturas em duplicado, foram analisadas em triplicado. Dez µg foram usados para análise de N-glicanos. Ambos os N-glicanos ácidos e neutros foram analisados por meio de MALDI-TOF.

A principal glicoforma em M124 cepa parental foi Man5Gn2. Em todas as cepas knockout Alg3 o principal glicoforma foi Man3 (Figura 10). Man3 não foi encontrada na cepa parental M124. Em diferentes linhagens knockout Alg3 a quantidade de Man3 variou entre 49,7 % - 55,2 % nas culturas do frasco agitado permitindo queda de pH. Hex6Gn2 foi aumentado na cepa parental. Intensidades de sinal, como percentagens dos sinais observados neutros N-glicanos são apresentados na Tabela 5 abaixo.

Tabela 5: Teor de N-glicanos Neutros de Cepas Alg3 knockout.

Composição	m/z	Cepa											
		M124 parental		4A		5A		6A		10A		16A	
		Média	STDEV	Média	STDEV	Média	STDEV	Média	STDEV	Média	STDEV	Média	STDEV
Hex3HexNAc2	933.31	0.0	0.0	53.6	0.2	55.2	4.2	49.7	0.5	53.3	0.9	53.4	0.9
Hex4HexNAc2	1095.37	1.6	0.1	2.7	0.0	2.9	0.7	3.4	0.1	3.2	0.4	3.4	0.4
Hex5HexNAc2	1257.42	70.2	3.3	8.5	0.2	7.3	1.1	10.4	0.5	8.6	0.9	9.7	0.9
Hex6HexNAc2	1419.48	7.9	1.1	35.0	0.3	34.4	1.9	36.1	0.6	34.9	0.5	33.2	0.7
Hex7HexNAc2	1581.53	7.8	0.6	0.3	0.4	0.3	0.4	0.3	0.4	0.0	0.0	0.3	0.4
Hex8HexNAc2	1743.58	5.9	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex9HexNAc2	1905.63	6.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex10HexNAc2	2067.69	0.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

A presença de diferentes isômeros de cada glicoforma não pode ser observada através de análise de MALDI MS, estudos posteriores de espectrometria de massa em tandem foram realizados. Em primeiro lugar, foram investigadas as estruturas e Hex5Gn2 Man3. Para Man3 foi perguntado se a estrutura Man3 é ramificada ou linear. Para esta análise, uma amostra contendo estas duas estruturas foi permitilada e analisada por meio da espectrometria de massa com a fragmentação usando o Bruker Ultraflex III TOF/TOF instrumento de acordo com as instruções do fabricante (Figuras 11

e 12).

Em seguida, determinou-se se a unidade de hexose na extremidade não redutora da estrutura Hex6Gn2 é uma manose ou glicose. Alfa-manosidase digestão foi realizada em todas as cepas knockout ea cepa parental (Figura 13). Manosidase Jack bean, que cliva α -manoses e deixa o β -manose a partir da estrutura principal intacta, foi usado. A estrutura resultante foi prevista para ser Man1Gn2.

Devido a efeitos de gama de baixo peso molecular em MALDI, a intensidade relativa do glicano Man1GlcNAc2 pode ter sido um pouco reduzida, que explica um pequeno aumento na quantidade relativa de Hex6. Após a digestão α -mannosidase, as glicoformas Man3 e man4 desapareceram. Não se observou qualquer estrutura Man2. No entanto, Hex6 (m/z 1419) não foi digerido (Tabela 6), indicando que houve uma unidade de glicose na extremidade não redutora da estrutura. Alguns Hex5 não digeríveis também estavam presentes, provavelmente produzidos por meio de uma reação fraca de remoção estericamente impedida de Man6-ramificado da Hex6.

Tabela 6: N-glicanos neutros da cepa Alg3 4A knockout antes (nativa) e após a digestão de α -manosidase.

		4A	
		Nativa	α -Man'ase
Composição	m/z	Média	%
Hex1HexNAc2	609.21	0.0	53.2
Hex2HexNAc2	771.26	0.0	0.0
Hex3HexNAc2	933.31	47.5	0.0
Hex4HexNAc2	1095.37	3.8	0.0
Hex5HexNAc2	1257.42	11.7	5.0
Hex6HexNAc2	1419.48	36.8	41.0
Hex7HexNAc2	1581.53	0.2	0.8
Hex8HexNAc2	1743.58	0.0	0.0
Hex9HexNAc2	1905.63	0.0	0.0
Hex10HexNAc2	2067.69	0.0	0.0

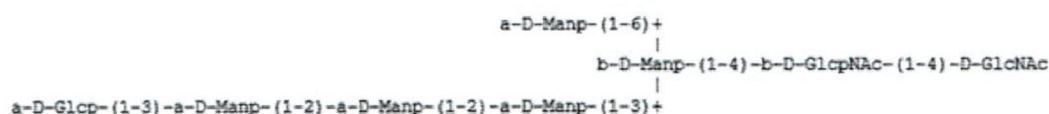
Para a análise de diferentes estruturas encontradas nas linha-

gens Alg3 knockout, uma digestão PNGase F em larga escala foi realizada para cepa Alg3 4A knockout. Dois grandes glicanos foram purificados por meio de HPLC (Figura 14) e analisados por meio de RMN (Figura 15).

Com base nos dados apresentados na Figura 15A, as espécies Hex3HexNAc2 foram inequivocamente identificadas como Man α 1-3 (Man α 1-6)-Man β 1-4GlcNAc β 1-4GlcNAc. As unidades H-1 Man α 3 e Man α 6 ressoaram em 5,105 e 4,914 ppm, respectivamente. A unidade Man β 4 H-2 foi observada em 4,245 ppm. Este sinal foi muito característico, devido à substituição vizinha de Man α 3-OH. Os sinais do grupo CH₃ de N-acetila das unidades de núcleo GlcNAc foram observados em 2,038 e 2,075. Estes valores concordaram bem com os relatados para este pentassacarídeo no Sugabase-database ([www.boc.chem.uu.nl / sugabase / sugabase.html](http://www.boc.chem.uu.nl/sugabase/sugabase.html)). Além disso, o espectro de RMN do próton foi medido por um Man α 1-3 (Man α 1-6)-Man β 1-4GlcNAc β 1-4GlcNAc produzido comercialmente (Glicoseparations, Inc.) em condições experimentais idênticas, e os deslocamentos químicos quase idênticos foram obtidos.

O espectro de RMN do componente Hex6HexNAc2 é mostrado na Figura 15B. Os dados mostraram que este componente representa o octassacarídeo Glc α 1-3Man α 1-2Man α 1-2Man α 1-3 (Man α 1-6)-Man β 1-4GlcNAc β 1-4GlcNAc. A presença de uma unidade de glucose era evidente a partir do sinal de 5,255 mostra um α Glc acoplamento típico 2,4 Hz. Todos os sinais Man mostram tipicamente acoplamento <1 Hz devido à configuração equatorial de H-2. Observaram-se pequenas diferenças em comparação com os dados Sugabase (Tabela 7), a qual pode ser atribuída à temperatura diferente usada na presente medição de RMN (40°C vs 26°C).

Tabela 7: Dados de RMN Publicados de Glc α 1-3Man α 1-2Man α 1-2Man α 1-3 (Man α 1-6)-Man β 1-4GlcNAc β 1-4GlcNAc. Os dados foram obtidos a partir de Sugabase (encontrado em [www.boc.chem.uu.nl / sugabase / sugabase](http://www.boc.chem.uu.nl/sugabase/sugabase)).



Resíduo	Ligação	Próton	PPM	J	Hz
D-GlcNAc		H-1a	5.189		
		H-1b	4.694		
		H-2a	3.867		
		H-2b	3.692		
		NAC	2.038		
b-D-GlcpNAc	4	H-1	4.606		
		H-2	3.792		
		NAC	2.077		
b-D-Manp	4, 4	H-1	4.773		
		H-2	4.237		
a-D-Manp	6, 4, 4	H-1	4.913		
		H-2	3.964		
a-D-Manp	3, 4, 4	H-1	5.346		
		H-2	4.080		
a-D-Manp	2, 3, 4, 4	H-1	5.304		
		H-2	4.103		
a-D-Manp	2, 2, 3, 4, 4	H-1	5.038		
		H-2	4.224		
a-D-Glcp	3, 2, 2, 3, 4, 4	H-1	5.247		
		H-2	3.544		

Por fim, os perfis de N-glicanos de cepa 4A knockout seleciona-
 dAs aleatoriamente foram analisados em diferentes momentos (dias 3, 5, 7 e
 9). O pH de cultura do frasco agitado foi de 4,8 no ponto de hora de início e
 5 de 2,6 no ponto de tempo final. As amostras triplicatas de todos os pontos de
 tempo de culturas duplicadas foram analisadas. Observou-se que em ambos
 os duplicados, a quantidade relativa de sinal de Man3Gn2 diminuiu como
 uma função do tempo de crescimento devido à redução do pH. No entanto, a
 quantidade de sinal Hex6Gn2 aumentou como uma função do tempo de
 10 crescimento (Tabela 8).

Tabela 8: As porcentagens de intensidade de sinal a partir dos sinais de glicanos neutros observados de cepa 4A Alg3 knockout. As culturas duplicadas (3A e 4A) de quatro diferentes momentos (dias 3, 5, 7 e 9) foram analisadas.

cepa Alg3 4A knock out (frasco 3A)

Composição	m/z	Dia 3, 3A		Dia 5, 3A		Dia 7, 3A		Dia 9, 3A	
		média	stdev	média	stdev	média	stdev	média	stdev
Hex3HexNAc	730,24	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Hex2HexNAc2	771,26	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Hex3HexNAc2	933,31	61,7	3,7	61,3	0,8	61,1	1,9	52,7	7,7
Hex4HexNAc2	1095,37	2,6	0,2	2,5	0,1	2,1	0,4	3,7	1,0
Hex5HexNAc2	1257,42	4,3	0,6	6,5	0,4	5,7	0,6	6,4	1,0
Hex6HexNAc2	1419,48	31,4	3,5	29,8	0,4	31,1	1,6	37,2	5,7

cepa Alg3 4A knock out (frasco 4A)

Composição	m/z	Dia 3, 4A		Dia 5, 4A		Dia 7, 4A		Dia 9, 4A	
		média	stdev	média	stdev	média	stdev	média	stdev
Hex3HexNAc	730,24	0,0	0,0	0,0	0,0	0,0	0,0	0,7	1,2
Hex2HexNAc2	771,26	0,0	0,0	0,0	0,0	0,0	0,0	0,3	0,5
Hex3HexNAc2	933,31	61,7	3,2	58,6	1,1	55,6	1,9	54,8	5,9
Hex4HexNAc2	1095,37	3,4	1,0	2,6	0,2	3,1	0,2	2,6	0,5
Hex5HexNAc2	1257,42	5,2	1,5	6,7	0,4	7,1	0,4	7,6	3,7
Hex6HexNAc2	1419,48	29,7	0,9	32,1	0,8	34,3	1,5	34,0	3,6

A diferença entre essas duas análises (Tabelas 4 e 7) sobre o percentual de Man3 em clone 4A (dia 5) foi anotada. Esta diferença pode ser devido a diferenças nos processos de análises. Alguma labilidade das preparações de proteína do meio de cultivo heterogênea foi observada após o (s) ciclo (s) de congelação-descongelação, provavelmente devido ao glicano /e/ou degradação de proteínas, resultando em quantidades reduzidas de glicanos maiores. A geração dos dados na Tabela 5 incluíram os ciclos de congelamento-descongelamento adicionais.

As frações de N-glicanos ácidos foram também analisadas por meio de MALDI (Figura 16). A abundância de diferentes compostos ácidos na cepa M124parental difere de todas as cepas Alg3 knockout, entre os quais a fração ácida parece ser muito semelhante.

Três glicanos importantes na linhagem parental foram H6N2P1, H7N2P1 e H8N2P1. No Alg3 knockout o tamanho deslocado para glicanos menores : H5N2P1, H6N2P1 e H4N2P1. Além disso, os glicanos difosforilados foram mais abundantes na linhagem paterna. Isto pode ter sido devido a uma falta de um substrato apropriado para a enzima em particular que se liga fosforilados de manose a um glicano. A manose pode ser fosforilada mais alongada por outros resíduos de manose. A fosforilação não estava substancialmente presente nos glicanos da cepa M124 produzido sob condições de fermentação.

Comparação de fermentador e amostras crescidas do Frasco agitado

Uma Cepa Alg3 knockout (transformante 4A) foi cultivada em fermentação em batelada em lactose e passou meio extrato de grãos. O meio foi de 60 g / l de lactose com 20 g / l de extrato de grão passado com um volume de 7 litros (fermentador limitada bio01616) após a inoculação. Outros componentes do meio foram KH₂PO₄ e (NH₄)₂SO₄. PH Cultura foi controlado entre 5,5 e 5,8. Biomassa da cultura e amostras de sobrenadante foram tomadas durante o curso da execução e armazenadas a -20 ° C. Micélio também amostras foram recolhidas para análise do RNA possível e foram imediatamente congelados em nitrogênio líquido e transferidos para a temperatura de -70°C. As amostras recolhidas a partir de todo o curso destas

fermentações foram analisadas para a composição de N-glicanos. A análise de N-glicanos foi realizada para fermentador prazo bio01616) e para o dia tempo da amostra de 5 pontos a partir da cultura frasco agitado de transformante 4A (Figuras 17 e 18). O sinal principal na cultura em frasco de agitação foi Man3 (59 %). Na cultura do fermentador, o sinal principal foi Man3 (85 %), e a proporção de Hex6 foi reduzida para 8 %.

Conclusões

O knockout Alg3 foi bem sucedido na produção de 50 % ou mais do esperado Man3 glicoforma. A estrutura ramificada desejado de Man α 3 (Man α 6)-Man β foi verificada por meio da espectrometria de massa e espectroscopia de RMN de fragmentação.

Os outros produtos resultantes do knockout Alg3 incluíram man4 (produto minoritário contendo manose), Hex5 (um produto de degradação de Hex6 como indicado na Figura 13) e Hex6, que foi o segundo maior componente. O componente Hex6 foi caracterizada para conter terminais Glc pela resistência manosidase e sinais de RMN específicos, incluindo Glc α 3Maternal N. Considerou-se que a estrutura de glicano podem ser otimizadas através de métodos com a finalidade de reduzir a quantidade do terminal Glc, que foi provavelmente causando eficácia sub-ótima de glucosidase II com o glicano desprovido de manoses no Man α 6-braço da molécula. A otimização das condições de fermentação pode reduzir a quantidade de terminais Glc.

Estes dados indicaram glicosilação melhores resultados no T. reesei Alg3 knockout em comparação com dados anteriores para Alg3 knockouts em Aspergillus (Kainz et al., Appl Environ Microb. 2008 1076-1086) e P. pastoris (Davidson et al., Glicobiology 2004, 399-407). Nas obras de Kainz et al. e Davidson et al., os níveis Hex6 de produto correspondentes semelhantes ou superiores foram relatados. Estes estudos relataram também problemas adicionais com α 2-manose, produtos OCH1 e tamanho maior e glicanos de células específicas do tipo produzido por P. pastoris. Em conclusão, a análise de N-glicanos de T. reesei Alg3 knockout revelou que a maior glicoforma nas cepas knockout é Man3Gn2, um ponto de partida para a geração eficiente desejado do tipo de mamífero de N-glicanos.

Exemplo 3 - Purificação e Atividade de Enzimas GnTI e GnTII Individuais

GnTI e GnTII Humana (N-acetilglucosaminiltransferase I e N-acetilglucosaminiltransferase II) foram expressos como proteínas segregadas, solúveis em *Pichia pastoris*, a fim de estudar a sua especificidade e atividade de aceitador.

Geração de Construto GnTI para a Produção em *P. pastoris*

A sequência de GnTI humano (P26572) foi obtida como uma sequência de comprimento completo e subclonada em vetores de sobre-expressão de *Trichoderma reesei*. As sequências de codificação da proteína (CDS) que codificam a parte solúvel de GnTI humano, foram clonados para o vetor de expressão pBLARG-SX, a fim de produzir uma forma segregada da proteína na levedura *Pichia pastoris* para estudos enzimáticos. Durante o procedimento de clonagem, uma sequência de codificação Sua marcação foi adicionado ao extremo 5' do quadro para obter uma marcação na extremidade Terminal N da proteína truncada. A sequência foi verificada por análise de sequenciação. Resultando pTTg5 vetor foi linearizado e transformada por electroporação para células de *P. pastoris* com a finalidade de produzir GY190 cepa GY4. Arg + transformantes foram escolhidos e selecionados por meio de PCR. GY4 clones contendo o plasmídeo integrado foram testados quanto à expressão da proteína.

Expressão e purificação de GnTI Solúvel

A cepa de *P. pastoris* que expressa GY4 GnTI solúvel foi primeiramente cultivada durante a noite com agitação a 30°C em meio BMGY (extrato de levedura 1 %, 2 % de peptona, 100 mM de fosfato de potássio pH 6,0, 1,34 % de base azotada de levedura, 4 x 10⁻⁵ % de biotina, 1 % de glicerol) até DO600 2-6. As células foram então colhidas por centrifugação e ressuspensas para um OD600 de em meio BMMY (como BMGY, mas com 0,5 % de metanol em vez de 1 % de glicerol). A cultura foi colocada em um frasco e voltou iludido a uma incubadora com agitação a 16°C. 100 % de metanol foi adicionado a uma concentração final de 0,5 % a cada 24 h para manter a indução. Amostras de 1 ml da cultura de expressão foi feita 0, 24, 48 e 72 horas após a indução, e ambos os péletes de células e os sobrena-

dantes foram armazenados para análise. Após 3 dias de indução, as células de toda a cultura foram colhidas por meio da centrifugação, e o sobrenadante foi recolhido para posterior purificação de GnTI.

Preparação de amostra de GnTI bruto para Atividade de Ensaio

5 A cultura de células de *Pichia pastoris*, o qual continha solúvel marcada com His GnTI foi processada para ensaio de atividade por meio da concentração e troca de tampão. Em suma, 40 ml de sobrenadante a partir da cultura de *P. pastoris* em frasco de agitação foi colhido no dia 3 após a indução com MeOH por sedimentação das células no tubo de 50 ml Falcon
10 (Eppendorf 5810R, 3220 rcf, 5 min a 4°C) e a recolha sobrenadante. O sobrenadante foi, em seguida, concentrou-se para <2,5 ml, por centrifugações sequenciais (Eppendorf 5810R ou comparável, 3220 rcf, 10 min a 4°C) com um concentrador Millipore Ultracel Amicon 30K. O volume do concentrado foi ajustado para 2,5 ml com 100 mM de MES, pH 6,1. Concentrado foi submetido a troca de tampão, com uma coluna de filtração em gel PD-10 (GE Healthcare 17-0851-01). A coluna foi primeiro equilibrada com MES 100 mM, pH 6,1 e, em seguida, a amostra (2,5 ml) foi adicionado, por escoamento foi descartado e eluição com 2,25 ml de tampão MES foi recolhido. Finalmente, 500 µl do eluato foram concentrados para 100 µl com concentrador Millipore
15 Biomax 30K (Eppendorf 5417, 12 000 rcf, 5 min 4°C) e utilizado diretamente em ensaios de atividade.

Ensaio da Atividade da Enzima GnTI

Man α 1-6 (Man α 1-3) Man β 1-4GlcNAc (Man3Gn) foi utilizado como um aceitador para GnTI no ensaio de atividade de GnTI. A reação GnTI
25 foi realizado através da incubação da mistura de reação, que continha 0,1 mM de aceitador Man3GlcNAc, 20 mM de UDP-GlcNAc, de GlcNAc a 50 mM, 100 mM de MnCl₂, 0,5 % de BSA e 8 µl GnTI em MES 100 mM, pH 6,1, em um volume total de 10 µl à temperatura ambiente durante a noite. A reação foi parada por meio da incubação da reação a 100 °C durante 5 min.

30 Em paralelo com o ensaio de atividade de GnTI, a possível atividade HexNAc'ase na preparação de enzima em bruto foi controlada. GlcNAc β 1-2Man α 1-6 (GlcNAc β 1-2Man α 1-3)-Man β 1 4GlcNAc β 1-4GlcNAc-Asn

(= Gn2Man3Gn2-Asn) foi utilizado como um substrato para HexNAc'ase. A reação foi realizada de forma semelhante como para GnTI, excepto 100 pmol de Gn2Man3Gn2-Asn foi adicionado em vez de Man3Gn e UDP-GlcNAc. Nenhuma atividade HexNAc'ase foi detectada.

5 A mistura da reação foi purificada para análise MALDI por Hypersep sequencial C18 (100 mg, Thermo Scientific, gato não : 60300-428) e Hypercarb (10 mg/96 cavidade / placa 1 PKG, cat no. 60302-606) cromatografia sobre HyperSep 96 cavidades Manifold Vacuum, Thermo Scientific. Hypersep C18 foi preparado com 300 mL EtOH e 300 mL de água MQ, o
10 prato de coleta foi então colocado sob, e as amostras foram carregadas e eluída com 150 mL de água MQ. Hypercarb foi preparado com 300 mL MeOH e 300 ml de água MQ. Eluatos de Hypersep C18 foram carregados, os sais foram removidos com 150 µl de 0,5 M de NH₄Ac, e os cavidades foram lavados com 2 x 300 ml de água MQ. Os produtos de reação de GnTI foram
15 eluídos com 150 µl de 25 % de ACN, e os produtos de reação HexNAc'ase foram eluídas com 25 % de ACN e 0,05 % de TFA. As amostras foram secas em um Speedvac.

A matriz assistida por laser de ionização de dessorção de tempo de luz (MALDI-TOF), espectrometria de massa (MS) foi realizada com um
20 Ultraflex TOF / TOF instrumento Bruker (Bruker Daltonics, Alemanha). Sacarídeo aceitador e produto foram detectados no modo de reflector de íons positivos como íons [M + Na] ⁺. Calculado m / z valores para [M + Na] ⁺-sinais de Hex3HexNAc1 e Hex3 HexNAc2 foram 733,238 e 933,318, respectivamente. A proporção percentual do aceitador e o produto foi calculada a partir
25 dos sinais correspondentes a Hex3HexNAc1 e Hex3 HexNAc2 (Figura 19).

Geração de Construto GnTII para a Produção em *P. pastoris*

A sequência nucleotídica que codifica GnTII humana foi amplificada por meio de PCR com iniciadores GP3 e GP13, que continha os locais de restrição KpnI e EcoRI, respectivamente. O fragmento de PCR com EcoRI /
30 KpnI digerido foi ligado a um pBLARG-SX vetor de clonagem digerido de forma semelhante. Após verificação da sequência, o construto final foi transformada em cepa de *P. pastoris* GS190 para originar GY22 cepa. Os transformantes

de levedura positiva foram testados por meio de PCR. Dois clones (dos quais apenas uma é mostrado na Figura 20) foram estudados quanto à expressão de GnTII, sob o controle do metanol AOX1 induzível em 16°C e a 30°C.

Expressão de GnTII Solúvel

5 De acordo com a análise de Western blot (Figura 20), a cepa *P. pastoris* GY22 produziu a enzima recombinante solúvel GnTII. GnTII tem uma massa molecular calculada de 49049,0 Da e dois locais previstos de N-glicosilação. O GnTII recombinante foi secretado para o meio de cultura a 16°C (pista 9). Quando cultivadas a 30°C, GnTII recombinante foi presa no
10 interior das células (pista 4).

Ensaio de atividade de GnTII Solúvel

P. pastoris de cultura de células contendo GnTII solúvel marcada com His foi processado para um ensaio de atividade, tal como descrito acima para o GnTI. A cultura celular foi centrifugado, o sobrenadante foi colhido e
15 concentrado, a troca de tampão MES 100 mM, pH 6,1 foi realizado, e a amostra resultante foi adicionalmente concentrado antes do teste de atividade.

O ensaio de atividade foi realizado de forma semelhante como para GnTI. GnMan3Gn foi utilizado como um aceitador de GnTII.

GnTII A reação foi realizada na presença de 0,1 mM de aceitador GnMan3Gn, 20 mM de UDP-GlcNAc, de GlcNAc a 50 mM, 100 mM de
20 MnCl₂, 0,5 % de BSA, e GnTII em MES 100 mM, pH 6,1. A purificação da mistura de reação para a análise de MALDI-TOF MS foi realizada por cromatografia em C18 e Hypersep Hypercarb sequenciais em uma placa de 96 cavidades no colector de vácuo tal como descrito acima para o GnTI.

25 MALDI-TOF MS foi realizado com uma Ultraflex TOF / TOF instrumento Bruker (Bruker Daltonics, Alemanha). Sacarídeo aceitador e produto foram detectados no modo de reflector de íons positivos como íons [M + Na] ⁺. Relação entre o produto e aceitador no final da reação foi calculada a partir de suas intensidades de sinal (calculado m / z valores para [M + Na] ⁺
30 sinais de GnMan3Gn aceitador e do produto com uma adição de GlcNAc são 933,318 e 1.136,397, respectivamente).

A cultura de *P. pastoris* produzindo GnTII foi repetida, e GnTII

concentrado (60x) foi preparado a partir do sobrenadante e a sua atividade medida de acordo com os métodos descritos acima. MALDI espectro de amostras no ponto de tempo 2,5 h, 5 h, e durante a noite mostraram que 80 %, 83 % e 82 % do aceitador foi convertido no produto, respectivamente. A

5 reação próxima do máximo foi atingida em 2,5 horas.

Além disso, uma amostra GnTII em bruto foi preparada, e o ensaio de atividade foi realizado como descrito acima para a amostra de GnTI bruto. A mistura de reação foi incubada durante a noite, purificado e sujeito a análise de MALDI. MALDI espectros revelou atividade GnTII (Figura 21).

10 HexNAc'ase atividade não foi detectada na amostra GnTII bruto.

Os métodos utilizados para sintetizar um aceitador de GnTII para utilização nos ensaios de atividade GnTII descritos acima foram como segue. Uma amostra GnTI foi preparada a partir de um meio de cultura de *P. pastoris*, tal como descrito acima. Este exemplo mostrou GnTI elevada atividade GnTI e, portanto, pode ser usado na conversão de cerca de 40 nmols de Man3Gn para GnMan3Gn. A reação foi levada a cabo na presença de Man3Gn 0,5 mM, 20 mM de UDP-GlcNAc, de GlcNAc a 50 mM, 100 mM de MnCl₂, 0,5 % de BSA, e amostra GnTI. A mistura de reação foi incubada três dias à temperatura ambiente. Uma amostra de cerca de 1 % foi submetida a purificação por

15 meio de cromatografia Hypercarb e análise MALDI. A reação GnTI convertido quase todos Man3Gn aceitador GnMan3Gn ao produto de acordo com espectro de MALDI. Apenas 2,8 % do aceitante não foi convertida.

20

Exemplo 4 - Proteína de Fusão GnTI / GnTII

Geração do construto de expressão de GnTI / GnTII

25 Uma proteína de fusão GnTI/II recombinante foi construída por meio da amplificação de um fragmento de 1313 pb GnTII fusão com um iniciador 65-mero na extremidade 5', que continha um local de fusão em-estrutura(uma pequena sequência de GnTI contendo um de uma forma natural Alei do local de restrição que ocorre com o códon de parada removido e

30 sobreposto com a sequência de GnTII) e os iniciadores da extremidade 3' homóloga GntII contendo quer SpeI ou locais de restrição NdeI. Este local de fusão permitiu a clonagem de um fragmento de fusão diretamente a um vetor

super-expressão de *T. reesei* com GnTI tipo selvagem sob o controle do promotor de *cbh1* (clonagem com AleI / NdeI) ou com GnTI tipo selvagem sob o controle do promotor de *gpdA* (clonagem com AleI / SpeI). Foram utilizados de alta fidelidade da polimerase Phusion (Finnzymes) e procedimentos de amplificação e clonagem padrão. A sequência foi verificada por sequenciação directa de vetores de expressão. O vetor resultante foi utilizado com a finalidade de expressar a fusão de uma proteína transmembranar em *T. reesei*.

Para obter mais informação sobre a funcionalidade das proteínas de fusão, GnTI / II também as proteínas de fusão foram expressas como proteínas solúveis em *P. pastoris*. CDS de GnTI / II de fusão que codifica a parte solúvel da proteína foi clonado para o vetor de expressão pBLARG-SX, a fim de produzir a proteína para estudos enzimáticos. Durante o procedimento de clonagem, a sua sequência de codificação de marcação foi adicionada à extremidade 5 'do quadro para obter uma marcação na extremidade Terminal N da proteína truncada. A sequência foi verificada por meio da análise de sequenciação. O vetor resultante foi linearizado e transformado por electroporação a cepa de *P. pastoris* GS190 para originar cepa GY6. Arg + transformantes foram escolhidos e selecionados por meio de PCR. Os clones *P. pastoris* contendo o plasmídeo integrado foram testados quanto à expressão da proteína.

Purificação do Solúvel GnTI / II Produzido em *P. pastoris*

A expressão em *P. pastoris* e os procedimentos de purificação foram efetuados tal como descrito acima com a proteína recombinante de GnTI.

Testes de Atividade de Enzima da Proteína de Fusão de GnTI/II

Os ensaios de atividade foram realizados como descrito acima para os ensaios de GnTI utilizando Man3Gn oligossacarídeo como um aceitador e dador de UDP-GlcNAc. Os produtos da reação foram analisados por meio da espectrometria de massa MALDI-TOF. Apenas a atividade de GnTI foi observado para a proteína de fusão de GnTI / GnTII (Figura 22).

30 Transformação de *T. reesei* com Construto GnTI/GnTII por meio da Integração Aleatória

Um plasmídeo quimico humano GnTI / GnTII com um promotor

de *gpdA* foi co-transformado na cepa de *T. reesei* com integração aleatória de M124. A seleção foi obtido por cotransformação de um gene contendo um marcador acetamidase plasmídeo. Vinte PCR transformantes positivos foram purificados para clones mononucleares e cultivados em culturas frasco de 5 agitação para análise de glicanos. Todos os transformantes e a cepa parental M124 foram cultivadas em meio mínimo de *Trichoderma* (TRMM), pH 4,8, suplementado com 4 % de lactose e 2 % de extrato de grão gasto. Sobrenadante e micélio amostras foram coletadas nos dias 3, 5 e 7, e foram armazenadas congeladas até a análise. Além disso, como controle, *T. reesei* foram transformadas com um construto de *GnTI* por meio da integração aleatória. 10

Análise de Glicano de Cepas T. reesei GnTI/GnTII obtidas por meio da integração Aleatória

Amostras de 20 clones diferentes em três momentos diferentes (dias 3, 5 e 7) de *T. reesei* M124 *GnTI* / *GnTII* transformantes foram analisadas. 15 Amostras de duas cepas progenitoras M124 foram analisadas para os controles. As reações com N-glicanase sem desnaturação SDS foram realizados em placas de 96 cavidades, em triplicado, para 5 µg de proteína sobrenadante. A concentração de proteína dos sobrenadantes foi medida por ensaio à base de Bradford (Bio-Rad Quick Start Bradford Protein Assay) utilizando 20 BSA como um padrão. Ambos neutro e N-glicanos ácido foram analisados por MALDI-TOF MS. Nenhum produto foi detectada utilizando o *GnTI* / *GnTII* construto em qualquer um dos clones em qualquer ponto de tempo, bem como nos clones dos transformantes com *GnTI gpdA* promotor.

Transformação de T. reesei com Construtos GnTI/GnTII por meio da Integração alvo 25

Uma sequência de *GnTI/GnTII* quimérico foi subclonada em uma estrutura principal pTTv38, um vetor que contém um gene marcador de acetamidase e locais de sequência 5'- e 3'- flanqueadora para o local de integração *alg3*. O vetor foi transformado em cepa *T. reesei* M124 como um 30 fragmento digerido. A partir desta transformação, 18 transformantes positivos de PCR, produzindo fragmentos de PCR, indicando a integração correta para o local *alg3*, foram detectados. Estes transformantes foram cultivados

em frascos de agitação após uma única etapa de purificação e os esporos foram analisados conforme descrito abaixo.

Análise de Glicano de *T. reesei* de Cepas GnTI/GnTII obtido visando o local alg3

As amostras de sobrenadante de 10 clones diferentes em três
 5 momentos diferentes (dias 3, 5 e 7) de Δ alg3 *T. reesei* GnTI / GnTII transformantes foram obtidos. Os clones foram cultivados em frascos agitados com duas composições diferentes meios de comunicação. TRMM, pH 5,5, com 2 % de extrato de grão gasto, 4 % de lactose, e K-ftalato de tampão foi utilizado para todos os clones e, em paralelo, TRMM, pH 5,5, com 2 % de
 10 extrato de grão gasto, 4 % de lactose, 1 % de casaminoácidos, e K-ftalato tamponantes foi usado para cinco dos clones. O cultivo foi mantido por 7 dias : 5 dias a 28°C e nos dias 6 e 7 a 24°C.

As análises N-glicanos foram feitas em triplicado em placas de 96 cavidades durante 5 µg de proteína sobrenadante. As amostras foram
 15 analisadas a partir de dia 3, 5 e 7. A concentração de proteína dos sobrenadantes foi medida por ensaio à base de Bradford (Bio-Rad Quick Start Bradford Protein Assay) utilizando BSA como um padrão. Ambos N-glicanos neutro e ácido foram analisados por meio de MALDI-TOF MS.

As quantidades detectáveis de glicofórm G0 foram encontrados em
 20 cada clone. O clone continha o 201A com mais de 1,2 % de Gn2Man3 (Figura 23 e Tabela 9). Além disso, a quantidade de Hex6 foi menor neste clone particular. O segundo meio com 1 % casaminoácidos não deu qualquer produção extra de G0/GlcNAc β 2Man α 3 (GlcNAc β 2Man α 6) Man β 4GlcNAc β 4GlcNAc β . Os resultados dos dias 3 e 7 amostras foram essencialmente os mesmos
 25 como para a amostra de 5 dias.

Tabela 9: As percentagens de intensidade de sinal de N-glicanos observados a partir de proteínas secretadas de transformantes de *T. reesei* GnTI / II (GnTI/II integrados no local alg3). Os clones com a letra A no seu nome foram cultivados em meio A) e os clones com B em meio B), que tinha um extra de 1 % de casaminoácidos em comparação com meio A).

 30

composição	mV _e	clone 201A, dia 5				clone 202A, dia 5				clone 208A, dia 5				clone 210A, dia 5							
		Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX
Man2	771.3	0.6	0.5	86.8	0.0	1.0	0.4	0.7	173.2	0.0	1.1	0.6	0.6	92.3	0.0	1.1	0.1	0.2	173.2	0.0	0.4
Man3	933.3	47.9	14.5	30.2	39.0	64.6	41.3	0.2	0.4	41.1	41.5	38.2	1.1	2.8	37.0	38.9	38.2	1.1	3.0	37.5	39.5
Man4	1095.4	7.9	2.9	36.5	5.9	11.3	6.4	0.6	8.7	6.0	7.0	5.3	0.2	4.0	5.0	5.5	6.0	0.4	6.6	5.5	6.2
GnMan3	1136.4	1.4	0.7	46.9	1.0	2.2	1.1	0.3	23.5	0.8	1.3	1.0	0.2	17.0	0.9	1.2	1.1	0.1	8.9	1.0	1.2
Man5	1257.4	10.5	2.5	23.5	8.7	13.3	8.6	0.8	9.7	7.7	9.4	8.2	0.3	4.0	7.8	8.5	8.9	0.3	3.7	8.6	9.3
Gn2Man3	1339.5	1.2	0.8	69.1	0.6	2.2	0.6	0.1	21.0	0.5	0.8	0.6	0.1	21.5	0.5	0.7	0.6	0.1	17.5	0.6	0.8
Hex6	1419.5	27.3	23.7	86.7	0.0	42.0	40.5	0.6	1.5	39.9	41.1	44.7	0.7	1.6	43.9	45.2	43.2	0.7	1.6	42.7	44.0
Hex7	1581.5	2.9	3.0	103.3	1.1	6.4	1.0	0.1	11.0	1.0	1.2	1.1	0.1	11.7	1.0	1.2	1.2	0.0	3.7	1.2	1.2
Hex8	1743.6	0.1	0.2	173.2	0.0	0.4	0.2	0.3	173.2	0.0	0.5	0.3	0.2	87.0	0.0	0.4	0.6	0.3	57.0	0.3	1.0
composição	mV _e	clone 212A, dia 5				clone 213A, dia 5				clone 215A, dia 5				clone 216A, dia 5							
		Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX
Man2	771.3	0.4	0.4	86.8	0.0	0.7	0.6	0.6	94.4	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Man3	933.3	45.6	1.3	2.8	44.2	46.8	40.0	2.8	7.0	37.3	42.9	43.4	1.9	4.4	41.3	45.1	42.6	2.0	4.6	40.5	44.4
Man4	1095.4	5.6	0.3	5.1	5.4	5.9	6.5	0.6	8.8	6.0	7.1	6.3	0.5	8.5	5.7	6.8	6.1	0.6	10.3	5.4	6.7
GnMan3	1136.4	0.9	0.2	22.4	0.7	1.1	0.9	0.1	8.5	0.8	1.0	1.1	0.1	6.9	1.0	1.2	1.1	0.2	14.1	0.9	1.2
Man5	1257.4	7.2	0.5	7.0	6.8	7.7	9.5	0.4	3.8	9.1	9.8	8.5	0.4	4.2	8.2	8.9	7.7	0.6	8.4	7.0	8.3
Gn2Man3	1339.5	0.5	0.1	11.9	0.5	0.6	0.6	0.1	18.3	0.5	0.7	0.7	0.2	29.3	0.6	1.0	0.7	0.2	26.4	0.5	0.9
Hex6	1419.5	38.6	1.2	3.0	37.4	39.7	40.7	2.5	6.1	38.2	43.2	38.5	1.8	4.6	37.4	40.5	40.5	1.7	4.2	39.0	42.4
Hex7	1581.5	0.8	0.0	4.1	0.8	0.8	1.0	0.1	10.8	0.9	1.2	1.1	0.1	4.5	1.1	1.2	1.0	0.1	6.4	0.9	1.0
Hex8	1743.6	0.4	0.1	34.8	0.3	0.5	0.1	0.2	173.2	0.0	0.3	0.4	0.3	88.5	0.0	0.6	0.4	0.3	87.6	0.0	0.6
composição	mV _e	clone 217A, dia 5				clone 219A, dia 5				clone 201B, dia 5				clone 202B, dia 5							
		Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX
Man2	771.3	0.9	0.0	5.0	0.9	1.0	0.5	0.4	96.7	0.0	0.9	0.4	0.7	173.2	0.0	1.1	0.6	1.1	173.2	0.0	1.8
Man3	933.3	54.1	1.1	1.9	53.0	55.0	44.0	1.8	4.1	42.4	45.9	46.9	0.2	0.5	46.6	47.1	40.6	1.7	4.3	38.6	41.8
Man4	1095.4	5.2	0.3	6.5	4.9	5.5	5.7	0.1	1.5	5.6	5.8	6.9	0.9	12.7	6.0	7.8	8.5	0.9	10.0	7.7	9.4
GnMan3	1136.4	0.9	0.2	17.4	0.7	1.0	1.0	0.2	16.6	0.9	1.2	1.2	0.4	32.1	0.9	1.6	1.3	0.4	0.0	0.9	1.8
Man5	1257.4	5.8	0.1	2.6	5.6	5.9	8.0	1.2	15.6	6.7	9.2	8.1	0.5	5.7	7.8	8.6	10.0	0.6	6.2	9.5	10.6
Gn2Man3	1339.5	0.7	0.1	14.7	0.6	0.7	0.9	0.1	14.2	0.8	1.0	0.8	0.1	7.1	0.8	0.9	0.7	0.5	70.8	0.3	1.3
Hex6	1419.5	31.5	1.5	4.7	30.5	33.3	38.5	1.1	2.8	37.3	39.2	34.2	0.7	2.1	33.8	35.1	37.5	1.1	2.8	36.7	38.7
Hex7	1581.5	0.9	0.1	12.9	0.8	1.0	1.0	0.2	15.4	0.8	1.1	1.1	0.1	5.2	1.0	1.2	0.8	0.7	86.9	0.0	1.2
Hex8	1743.6	0.0	0.0	0.0	0.0	0.0	0.4	0.1	17.9	0.3	0.5	0.4	0.3	90.7	0.0	0.7	0.0	0.0	0.0	0.0	0.0
composição	mV _e	clone 208B, dia 5				clone 210B, dia 5				clone 219B, dia 5											
		Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX	Medida	SD	RSD	MIN	MAX
Man2	771.3	0.9	0.8	87.1	0.0	1.5	0.8	0.7	86.7	0.0	1.2	1.2	0.1	10.3	1.0	1.3					
Man3	933.3	48.4	1.2	2.4	47.3	49.6	39.6	1.1	2.7	38.6	40.8	34.9	1.8	5.2	33.2	36.8					
Man4	1095.4	7.2	0.2	2.2	7.0	7.3	7.9	0.6	8.0	7.3	8.5	8.1	0.3	4.1	7.8	8.4					
GnMan3	1136.4	0.6	0.6	92.1	0.0	1.1	1.0	0.1	12.7	0.9	1.1	1.1	0.1	12.1	1.0	1.2					
Man5	1257.4	8.7	0.7	7.6	7.9	9.1	9.6	0.2	2.0	9.4	9.8	11.3	0.8	7.5	10.7	12.3					
Gn2Man3	1339.5	0.4	0.2	44.3	0.2	0.6	0.6	0.2	32.4	0.4	0.8	0.6	0.1	13.9	0.5	0.6					
Hex6	1419.5	32.4	0.4	1.4	32.1	32.9	38.5	0.3	0.8	38.3	38.9	40.6	0.7	1.8	39.8	41.1					
Hex7	1581.5	1.0	0.2	15.5	0.8	1.1	1.5	0.1	8.2	1.4	1.6	1.4	0.2	13.5	1.2	1.5					
Hex8	1743.6	0.4	0.4	87.7	0.0	0.7	0.5	0.5	92.4	0.0	0.9	0.8	0.1	16.3	0.7	0.9					

Exemplo 5 - Proteína de Fusão GnTII/GnTI

Geração de construto de expressão GnTII / GnTI

Um construto de expressão de fusão GnTII/GnTI foi gerado através da aplicação de técnicas de PCR de sobreposição. Os fragmentos de fusão foram amplificados a partir de modelos de GnTI e GnTII separadamente com iniciadores contendo 50 pb se sobrepõe ao local de fusão em- estrutura. Os fragmentos foram purificados a partir de um gel de agarose e utilizado como modelo de PCR para a amplificação do construto de fusão de acordo com os procedimentos padrão. O construto de fusão foi clonado para um vetor com locais de restrição Apal/Spel. O construto resultante foi verificado por meio da análise de sequenciação. Um vetor foi gerado para a expressão da forma solúvel de GnTII / GnTI de *P. pastoris* com a sua marcação na extremidade Terminal N da proteína alvo. Este vetor foi gerado de uma maneira semelhante à descrita acima para o construto de fusão GnTI / II.


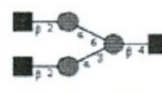
Purificação de GnTII / GnTI Solúvel Produzido em *P. pastoris*

A expressão em *P. pastoris* e os procedimentos de purificação foram efetuados tal como descrito acima para a proteína recombinante de GnTI.

Testes de Atividade de Enzima de Proteína de Fusão de GnTII / GnTI

Os ensaios de atividade foram realizados como descrito acima para GnTI usando Man3Gn oligossacarídeo como aceitador. Um espectro de MALDI da mistura de reação a partir da reação GnTII / GnTI purificada mostrou que dois resíduos GlcNAc β foram transferidos para o receptor (figura 24).

Tabela 10: Sumário das atividades da proteína de fusão GnTII / GnTI.

transformante GnTII/GnTI	Concentração do Aceitante	Produtos Formados	
			
Transformante 1	0.5 mM	47%	5%
Transformante 1	0.1 mM	-	11%
Transformante 2	0.5 mM	3%	2.4%

Caracterização por β -N-acetilglucosaminidase

A mistura de reação formada na atividade GnTII / GnTI foi tratada com β 1-2,3,4,6-N-acetilglucosaminidase de *Streptococcus pneumoniae*. A análise MALDI MS foi utilizada para determinar que ambos os resíduos transferidos ligados a β GlcNAc foram clivados (figura 25).

Galactosilação por β 1-4GalT

A mistura de reação formada na atividade GnTII / GnTI foi tratada com β 1-4GalT a partir de leite bovino. β 1-4GalT era esperado que os resíduos de GlcNAc galactosylate terminais na mistura de produto. De acordo com o espectro de MALDI-se a mistura de reação β 1-4GalT, ambos os produtos foram galactosilado. Dois galactoses foram transferidas para o produto Gn2Man3Gn, o que indicou que os resíduos de GlcNAc foram ligados para separar as β ramificações de manose (Figura 26).

Transformação de *T. reesei* com o Construto GnTII / GnTI por meio da Integração Aleatória

Uma sequência GnTII / GnTI quimérica foi concebida e clonada em um vetor contendo o promotor de *gpdA*. Após verificação da sequência do plasmídeo, que foi cotransformada na cepa M124 *T. reesei* com o gene marcador da higromicina. Treze PCR transformantes positivos foram identificados. Todos os transformantes positivos e à cepa parental M124 foram cultivadas em TRMM, pH 4,8, suplementado com 4 % de lactose e 2 % de extrato de grão gasto. Além disso, sete transformantes e a cepa parental foram cultivados em TRMM, pH 5,5, com 4 % de lactose, 2 % de extrato de grão gasto, e 1 % de casaminoácidos tamponados com 100 mM (Pipps piperazine1, ácido 4bis2propanossulfônico). As medições de pH foram utilizadas para monitorar a taxa de crescimento das cepas. Sobrenadante e micélio amostras foram coletadas nos dias 3, 5 e 7, armazenado congelado, e analisadas para estruturas de glicano. A sequência de GnTII / GnTI também foi clonado em um plasmídeo que contém o promotor *cbh1*. Além disso, como controle, de *T. reesei* foram transformadas com um construto de GnTI por integração aleatória.

Análise de Glicano de Cepas de *T. reesei* GnTII/GnTI obtidas por meio da Integração Aleatória

156 amostras de sobrenadante de *T. reesei* da cepa M124 GnTII / GnTI transformantes e cepa M124 parental cultivada em dois meios diferentes foram analisadas. O primeiro meio foi TRMM, pH 4,8, suplementado com 2 % de extrato de grão gasto e 4 % de lactose, e o segundo meio foi TRMM, pH 5,5, suplementado com 2 % de extrato de grão gasto, 4 % de lactose, 100 Píppos mM, e 1 % casaminoácidos. As células foram cultivadas em ambos os tipos de mídia para 3, 5 e 7 dias.

10 As reações de N-glicanase sem desnaturação SDS foram realizadas em placas de 96 cavidades, em triplicado, para 5 µg de proteína de amostras de sobrenadante para pontos de tempo de 3 e 5 dias. A concentração de proteína dos sobrenadantes foi medida por ensaio à base de Bradford (Bio Rad Quick Start Bradford Protein Assay) utilizando BSA como
15 um padrão. Ambos neutro e N-glicanos ácido foram analisados por MALDI-TOF MS.

Nenhum sinal do produto GnTII / GnTI esperado era visível em nenhum dos clones a partir dos pontos de tempo de 3 e 5 dias. Além disso, não foi observado a partir de GnTI e GnTI / II transformantes com promotores GPDA que foram gerados por meio da integração aleatória.

Transformação de *T. reesei* com Construto GnTII / GnTI por meio da Integração alvo

Um vetor possuindo a sequência de GnTII / GnTI quimérico sob o controle do promotor de *cbh1* foi construído com um marcador LOOPOUT gene *pyr4* e subclonado em um vetor de espinha dorsal entre *alg3* flanqueando fragmentos da região alvo para a integração. Um cassete de expressão *PmeI* digerido foi transformado em *T. reesei* cepa M127 (*pyr4*-cepa de M124). Após a seleção da placa, os clones foram PCR-selecionados e purificados através de esporos individuais. Com a finalidade de se obter material
25 para análises de glicano, agitar culturas em frasco foram realizadas como descrito. Cinco transformantes positivos de PCR, indicando a integração correta para o local de *alg3* na transformação M127 foram cultivadas em um
30

volume de 300 ml durante sete dias a 28°C em um meio contendo TRMM, pH 5,5, suplementado com 40 g / l de lactose, 20g / l de extrato de grãos e 100 Pipps mm. Para evitar a contaminação bacteriana, com 100 mg / l de ampicilina foi adicionado aos balões no momento da inoculação. As

5 amostras para análises glicano foram coletados nos dias 3, 5 e 7.

Análise de Glicano de *T. reesei* de Cepas GnTII / GnTI obtidas visando o Local alg3

As amostras de sobrenadante de *T. reesei* da cepa M124 (controle), cinco clones diferentes de GnTII M127 / GnTI transformantes, e amostras de meio de controle foram preparadas em triplicado, em placas de 96 cavidades durante 5 µg de proteína sobrenadante. As concentrações de proteína dos sobrenadantes foram medidas por ensaio à base de Bradford (Bio-Rad Quick Start Bradford Protein Assay) utilizando BSA como um padrão. Reações PNGase F foram realizadas como descrito, mas sem a desnaturação SDS. Os N-glicanos libertados foram purificados em primeiro lugar com Hypersep C-18 e, em seguida, com Hypersep Hypercarb (ambos a partir de Thermo Scientific) onde glicanos neutros e ácidos foram separados. Ambas as purificações foram realizadas no formato de 96 cavidades. N-glicanos neutros foram analisados por MALDI-TOF MS.

20 As proporções de N-glicanos a partir de *T. reesei* GnTII M127 / GnTI neutro transformantes foram comparados com as proporções da cepa M124, que foi de outro modo o mesmo como cepa M127 mas pyr4 positivo. Quatro dos cinco GnTII / GnTI transformantes G0 produzido como uma glicofoma principal em todos os pontos de tempo (3, 5 e 7 dias). Apenas os clones 46A foi G0 negativa (Figura 27). A proporção de Man3Gn era pequeno em cada clone em todos os pontos de tempo, mas a proporção de Hex6 era ainda bastante grande. No dia 7, clone 17A produziu o mais e o menos G0 Hex6 em comparação com outros clones (Figura 27). Quatro clones dos transformantes GnTII / GnTI produzidos cerca de 40 % glicofoma G0 no dia 5 em condições de frasco de agitação (Figura 27). Condições de fermentação com um pH controlado pode aumentar a quantidade de produto G0 e reduzir a quantidade de Hex6 em alg3 knock-out.

Em um meio de amostra, foram observados uma série de N-glicanos do tipo plantas, mas não foram observados sinais correspondentes a G0.

Transformação de Rituximab Produzindo *T. reesei* com Construto GnTII /

5 GnTI por meio da Integração alvo

O cassete de expressão descrito na seção intitulada "Transformação de *T. reesei* com Construto GnTII / GnTI por meio da Integração alvo" foi transformado em *T. reesei* de cepa M279 (pyr4-deformação da cepa M202). M202 se obteve por meio da eliminação pep1 protease em M124 e
10 introduzindo o rituximab cadeia pesada e leve (com o sítio de clivagem de Kex2). Após a seleção da placa, os clones foram PCR-selecionados e purificados através de esporos individuais. Com a finalidade de se obter material para análises de glicano, as culturas agitadas em frasco foram realizadas conforme descrito na seção intitulada "Transformação de *T. reesei* com
15 Construto GnTII / GnTI por meio da Integração alvo" e, além disso, alguns meios de cultura foram suplementados com 0,3 mg / ml de tripsina de soja inibidor (SBTI) e 1 % de casaminoácidos. SBTI foi introduzido pela primeira vez na inoculação e, em seguida, por dia nos dias 3-6. PMSF e Pepstatina A foram adicionados a todas as amostras antes da congelação.

20 Análise de Glicano de Cepas GnTII / GnTI de *T. reesei* Produzindo Rituximab obtidas visando o local alg3

Rituximab foi purificado com Proteína G cromatografia de afinidade a partir de dia 5 amostras de sobrenadante com SBTI e desde o dia 5 e 7 amostras sem SBTI. Reações PNGase F foram realizadas para ~ 10 µg de
25 proteína desnaturada. Os N-glicanos libertados foram purificados em primeiro lugar com Hypersep C-18 e, em seguida, com Hypersep Hypercarb (ambos a partir de Thermo Scientific) onde glicanos neutros e ácidos foram separadas. As etapas de purificação foram realizadas no formato de 96 cavidades. Os N-glicanos Neutro e ácidos foram analisados por meio de MAL-
30 DI-TOF MS. Dois dos clones de transformantes GnTII / GnTI, 9A e 31A-1-1, produzido G0 glicoforma de ~ 30 % e ~ 24 %, respectivamente. No entanto, quantidades razoáveis de Hex6 e GnMan3 ainda foram observadas (Figura

28). Rituximab dos outros clones continham pouca ou nenhuma G0.

Otimização de espaçadores

Uma série de modificações de espaçadores para proteínas de fusão GnTII / GnTI foram construídos. Estas variantes foram produzidas em Pichia e estudados in vitro para a estabilidade e a atividade da enzima.

Os materiais e métodos para a clonagem de proteínas de fusão de GnTII / GnTI são descritos na presente invenção. A sequência T45 foi amplificada em duas partes, utilizando a estratégia PCR de sobreposição. Primeiro, um fragmento foi amplificado com GP13 iniciador 5' e GP93 iniciador 3', e um segundo fragmento foi amplificado com GP92 iniciador 5' e GP2 iniciador 3'. A amplificação foi realizada com Phusion alta fidelidade da polimerase PCR (Finnzymes) sob as condições padrão fornecidos pelo fornecedor. As condições de amplificação foram como se segue: desnaturação inicial a 98°C durante 30 segundos e desnaturação a 98°C para extensão de 5 segundos, emparelhamento a 65°C durante 30 segundos e extensão a 72°C durante 45 segundos, repetido 20 vezes, e finais a 72°C durante 20 minutos. Os produtos resultantes da PCR foram purificados a partir do gel de agarose com um kit de extração de gel GeneJET Fermentas. Estes fragmentos sobrepostos, com sequências modificadas foram combinados na mesma mistura de reação com condições convencionais sem iniciadores. Dez ciclos de recozimento/extensão foram efetuados como se segue: desnaturação inicial a 98°C durante 30 segundos de desnaturação a 98°C durante 5 segundos, emparelhamento a 65°C durante 30 segundos de extensão a 72°C durante 45 segundos, 10 de repetição vezes, e de extensão final a 72°C durante 20 minutos. Os iniciadores GP13 (5') e GP2 (3') foram adicionadas, e foi continuada ciclisto, tal como descrito acima para 20 ciclos de amplificação. T45 O fragmento amplificado foi purificado com um kit de purificação de PCR GeneJET Fermentas, digerido com EcoRI / KpnI (New England Biolabs) de acordo com protocolos convencionais, e clonado em EcoRI / KpnI vetor de expressão de levedura digerido pBLARG-SX. O vetor resultante foi sequenciado com iniciadores 3'AOX, 5'AOX, GP9, GP37, GP38 e GP122. A sequência foi encontrada para ser correta.

Este plasmídeo resultante foi utilizado como um molde para a modificação do espaçador 3xG4S. A clonagem da sequência T46 foi feita como descrito acima com T45. GP13 5'-iniciador e GP95 3'-iniciador foram usados para a síntese do primeiro fragmento, e GP94 5'-iniciador e GP2 3'-iniciador foram usadas para a síntese da segunda fragmento. Os fragmentos foram combinados, e os iniciadores GP13 (5 ') e GP2 (3') foram adicionados para a amplificação. T46 fragmento amplificado foi então digerido com EcoRI / KpnI e clonado no vetor de expressão de levedura pBLARG-SX. O vetor resultante foi sequenciado com os iniciadores descritos acima, e a sequência foi encontrada para ser correta.

Os espaçadores celulase naturais relacionados foram construídos com um método de sobreposição de PCR similares. Com o espaçador cbhl-relacionada, o primeiro fragmento foi amplificado com o iniciador 5'-GP13 GP107 e 3'-escorva. O segundo fragmento foi amplificado com o iniciador 5'-GP108 e GP2 3'-iniciador (Tabela 11). Com o espaçador EGIV-relacionada, o primeiro fragmento foi amplificado com GP13 5'iniciador e GP109 3'iniciador. O segundo fragmento foi amplificado com o iniciador 5'-GP110 e GP2 3'-iniciador (Tabela 11). Em ambos os casos, os produtos de PCR foram purificados a partir de gel de agarose, combinadas e usadas como um molde para a próxima reação de PCR para amplificar as sequências de T50 e T51. T50 e T51 produtos de PCR foram, em seguida, digerido com EcoRI / KpnI e clonado no vetor de expressão de levedura pBLARG-SX.

Todas as amplificações por meio de PCR foram realizadas com polimerase de alta fidelidade Phusion (Finnzymes). Os iniciadores (Tabela 11) foram encomendados à MWG Operon. A sequenciação foi realizada pelo Laboratório de Sequenciação de DNA do Instituto de Biotecnologia, Universidade de Helsínquia, como um serviço comercial.

Tabela 11: sequências de iniciadores.

Iniciador	Sequência 5'-3'
3'AOX	GCAAATGGCATTCTGACATCC (SEQ ID NO : 99)
5'AOX	GACTGGTTCCAATTGACAAGC (SEQ ID NO : 100)
GP2	CAGTGGTACCCCTAATTCAGCTAGGATCATAGCCCTCCCACG (SEQ ID NO : 101)
GP9	CGGACCACCGCAAGTTCC (SEQ ID NO : 102)
GP13	ATGCGGAAATTCTGCAATCATCATCATCATCGCCAGCGTAAGAACGAGGCCCT (6 x HIS) (SEQ ID NO : 103)
GP37	CCTTTCTCTATCCAACTCTACC (SEQ ID NO : 104)
GP38	GGAACCTGCGGTGGTCCG (SEQ ID NO : 105)
GP92	CCGCCGGCTCCAGGAGGTGGGGCAGTGGAGGTGGCGGCAGTGGAGGTGGCGGCGGTGGAAGTGGGAGGTGCCACCGCGCCCC (SEQ ID NO : 106)
GP93	GCGGTGGGCACCCCTCCCACTGCCGCCACCTCCACTGCCCCCCACCTCCCTGGAGCCGGCGGTAAGAC (SEQ ID NO : 107)
GP94	AGGTGGGGCAGTGGAGGTGGCGGCAGTGGCGGCGGTGGAAGTGGGAGGTGCCACCGCGCCCC (SEQ ID NO : 108)
GP95	CGGTGGGCACCCCTCCCACTTCCACCGCGCCCACTGCCGCCACCTCCCACTGCCCCCCACCTCCCTG (SEQ ID NO : 109)
GP107	GTTTCGCGGGAGGGTTGCCGCCCGCTAGGGTTGCCGGTGCTCTGGAGCCGGCGGTAAGACTTGC (SEQ ID NO : 110)
GP108	GCAACCCCTCCCGCGGAACCCGCCTGGCAGCACCGGGAGGTGCCACCGCGCCCTCCCGCCC (SEQ ID NO : 111)
GP109	CCGCCTCCAGGAACAGTGGCGCTGGCGGTGCCCGTCCGCGGAGCTCTGGAGCCGGCGGTAAGACTTGC (SEQ ID NO : 112)
GP110	CGCCACTGTTCTTGAGGCGGTAGCGGGCCCCACACAGCGGGAGGGTGCCACCGCGCCCTCCCGCCCCAGC (SEQ ID NO : 113)
GP122	CATTAGCGAGAAGTTTACGG (SEQ ID NO : 114)

Os espaçadores modificados (3xG4S e 2xG4S) das enzimas de fusão GnTII / GnTI foram transformados para um ensaio de atividade por meio da concentração e troca de tampão de uma maneira semelhante ao descrito para GnTI no Exemplo 3. Os ensaios de atividade foram efetuados com Man3Gn aceitador, e as misturas da reação foram purificadas como descrito no ensaio de atividade de GnTI. A análise MALDI também foi realizada conforme descrito com a mistura de reação GnTI, mas, além disso, a formação do produto GnTII, Hex3HexNAc3, foi seguido. O m/z valores calculados para o sinal de Hex3HexNAc3 + [M + Na] foi 1136,318 (Figura 29).

10 Variantes de Espaçadores

As variantes de espaçador GnTII / I foram modificadas a partir de sequência espaçadora da GnTII / I da proteína de fusão do tipo selvagem. Os espaçadores modificados são listados na Tabela 12. Todas as cepas de quatro espaçadores variantes (GY32, GY33, GY49 e GY50), do tipo selvagem GnTII / cepa fusão I (GY7-2), e cepa simulada (GY3) foram expressos em 16°C com inibidores da protease. As cepas foram inoculadas em 60 mL de meio BMGY-a +30°C, 220 rpm, durante a noite (o / n). As culturas durante a noite e as células foram sedimentadas foram ressuspensas em 60 ml de-meio BMMY. Os inibidores de protease, EDTA a 1 mM, Pepstatina A 1,5 uM (Sigma) e 1 de EDTA completo livre do comprimido de cquetel do inibidor de protease (Roche) foram adicionados em culturas, ao mesmo tempo, quando a indução MeOH foi iniciado e, depois de que uma vez por dia. Amostras de 25 ml foram retiradas de culturas no dia 3 e no dia 4, e as amostras de sobrenadante foram concentrados utilizando tubos de concentração (Millipore), tampão foi trocado em colunas PD-10 em MES 100 mM pH 6,1 e concentrou-se em 50x finais. Os péletes de células foram ressuspensos em 500 µl de 1xPBS, excepto sedimento celular de tipo selvagem (3), que foi novamente suspenso em 500 mL de MES a 100 mM pH 6,1 e coquetel inibidor completo (livre de EDTA).

30 A sequência de aminoácidos da proteína de fusão GnTII / GnTI contendo o espaçador 3xG4S está estabelecida na SEQ ID NO : 119. A sequência de nucleotídeos da proteína de fusão GnTII / GnTI contendo o es-

- paçador 3xG4S está estabelecido na SEQ ID NO : 141. A sequência de aminoácidos da proteína de fusão GnTII / GnTI contendo o espaçador 2xG4S está estabelecido na SEQ ID NO : 121. A sequência de nucleotídeos da proteína de fusão GnTII / GnTI contendo o espaçador 2xG4S está estabelecido
- 5 na SEQ ID NO : 139. A sequência de aminoácidos da proteína de fusão GnTII / GnTI contendo o espaçador cbhl está estabelecido na SEQ ID NO : 123. A sequência de nucleotídeos da proteína de fusão GnTII / GnTI contendo o espaçador cbhl está estabelecido na SEQ ID NO : 143. A sequência de aminoácidos da proteína de fusão GnTII / GnTI contendo o espaçador EGIV está estabelecido na SEQ ID NO : 125. A sequência de nucleotídeos da proteína de fusão GnTII / GnTI contendo o espaçador EGIV está estabelecido na
- 10 SEQ ID NO : 145.

- Uma amostra de 200 µl de suspensão de células foi lavada por meio da ressuspensão e centrifugação repetindo células em MES a 100 mM
- 15 pH 6,1 com coquetel inibidor completo (livre de EDTA). Um ligante de células foi preparado tomando 200 µl de amostra de células lavadas, adicionando 50 µl esferas de vidro e 2 µl de Triton X-100 e colocando em batedor durante 6 min. Ensaios de 50x concentradas sobrenadantes de cultura *P. pastoris*, a amostra de células e atividade GnTI do ligante celular foram realizadas como
- 20 acima.

Tabela 12 : Descrição de cepas de levedura.

Cepas de Levedura	Descrição	Sequência da variante de espaçador
GY3	cepa Mock	
GY7-2	Fusão de GnTII/I do tipo selvagem	
GY32-5	Fusão de GnTII/I da variante de espaçador 3xG4S	SEQ ID NO : 118
GY32-9		
GY33-7	Fusão de GnTII/I da variante de espaçador 2xG4S	SEQ ID NO : 120
GY33-8		
GY49-3	Fusão de GnTII/I da variante de espaçador cbhl	SEQ ID NO : 122
GY50-7	Fusão de GnTII/I da variante de espaçador EGIV	SEQ ID NO : 124
GY50-10		

A análise Western blot de sedimentos celulares e sobrenadantes de cultura de 50x concentradas a partir do dia 3 estão apresentados na Figura 30. O espaçador da variante cbhl (GY49) deu um sinal forte a partir da célula amostra de sedimento, mas não a partir do sobrenadante. O espaçador da variante EGIV (GY50) foi detectado a partir do sobrenadante, mas apenas o sinal fraco foi obtido. Sinais fracos de amostras de sobrenadante foram também obtidos com o tipo selvagem GnTII / cepa de fusão I (GY7-2), e as cepa espaçador da variante 2xG4S GY33-7 e GY33-8 (Figura 30).

As atividades da proteína de fusão GnTII / I contendo o espaçador da variantes foram então comparadas com a atividade da proteína de fusão GnTII / I contendo o espaçador do tipo selvagem.

Atividade da Fusão GnTII / I em sobrenadantes. O substrato Man3Gn GnTI foi fornecido, e o produto da reação, GnMan3Gn, atuou como o aceitador de GnTII para a atividade da proteína de fusão. As amostras para ensaios de atividade foram tomadas após o dia 3 e dia 4 fases de expressão. A Figura 31 mostra os resultados do ensaio de atividade de culturas de GnTII / eu proteínas de fusão contendo tanto o tipo selvagem ou o espaçador da variante espaçador. Exemplos de culturas foram feitas na presença de inibidores (a 1,5 uM de pepstatina A, 1 mM de EDTA, 1 tablet/50 ml de EDTA completo livre comprimido coquetel inibidor de protease). Para simplificar, o GnTI e GnTII produtos de reação foram adicionados em conjunto. Todas as amostras de ensaio de atividade continha apenas pequenas quantidades (<5 %) de produto GnMan3Gn GnTI, GnTII, indicando que transformado ativamente a GnMan3Gn para Gn2Man3Gn.

Todas as quatro variantes de espaçador apresentaram atividades GNT, embora houvesse alguma variabilidade entre clones e dias de cultivo. As proteínas de fusão GnTII / I (contendo os 2xG4S clone_1), 3xG4S (clone_1 e clone_2), ou variantes EGIV espaçador ativy apresentaram maior do que a enzima com o espaçador do tipo selvagem (Figura 31). A proteína de fusão GnTII / I contendo o espaçador da variante cbhl mostrou uma atividade comparável com a enzima, com o espaçador do tipo selvagem (Figura 31). O GnTII / I da proteína de fusão variante de contining 2xG4S (clone_2)

tinham menor atividade do que a enzima com o espaçador do tipo selvagem (Figura 31). Dia 4 amostras apresentaram maior atividade do que o dia 3 amostras, com exceção da GnTII / I proteína de fusão contendo o espaçador da variante 3xG4S (clone_1 e clone_2), que mostraram maior atividade no dia 3 (Figura 31). A proteína de fusão GnTII / I do espaçador da variante E-GIV teve a maior atividade no dia 4 (Figura 31).

Atividade de fusão GnTII / I nas células e lisados celulares. Os ensaios de atividade das células, ligante celular, e as amostras de sobrenadante de células contendo o GnTII / I da proteína de fusão tendo o espaçador do tipo selvagem indicaram que as amostras de lisados continham a maior atividade (Figura 32). A segunda mais elevada foi de atividade na superfície da célula e, em menor atividade foi observado nas amostras de sobrenadante (Figura 32). Por conseguinte, parece que a maior parte do GnTII / I da proteína de fusão foi localizada em células ou na superfície da célula, com apenas uma pequena quantidade a ser segregada.

As atividades GnT de células contendo GnTII / I e proteínas de fusão possuindo o espaçador ou de tipo selvagem ou as variantes de espaçadores são mostrados na Figura 33. As células foram ressuspensas em 500 µl de 100 mM de MES, pH 6,1 com EDTA de coquetel inibidor completo livre e espaçador da variante em 500 µl de PBS e as células e os lisados para testes de atividade foram preparados como acima.

Conforme mostrado na Figura 33, as proteínas de fusão GnTII / I contendo o espaçador da variantes eram muito mais GnTII / I da atividade nas células do que nos sobrenadantes. Em lisados, as enzimas parecem estar inativas. Acredita-se que esta falta de atividade é devida à ação de proteases libertadas. A proteína de fusão GnTII / I contendo o espaçador da variante cbhl mostrou uma atividade elevada em células e os lisados (Figura 33), que se correlaciona com a análise de Western blot mostrando maior sinal na célula de amostra de sedimento (Figura 30).

Discussão. Em sobrenadantes, as proteínas de fusão GnTII / I que contêm os espaçadores da variante 2xG4S e 3xG4S tiveram maior atividade que a GnTII / I da proteína de fusão contendo o espaçador do tipo sel-

vagem, enquanto que o espaçador da variante cbhl tinha uma atividade comparável à da proteína de fusão GnTII / I contendo o do tipo selvagem espaçador. Além disso, a proteína de fusão GnTII / I contendo o espaçador da variante EGIV apresentou a maior atividade de GnT. Análise de Western blot de dia 3 amostras apresentaram alguma correlação com os resultados do dia quatro atividades. Análise de Western blot mostraram bandas fracas com amostras de sobrenadante de tipo selvagem, os dois clones de 2xG4S e EGIV. As atividades foram detectados na seguinte ordem: EGIV > 2xG4S (clone_1) > 3xG4S (clone_2) > 3xG4S (clone_1) ≥ cbhl = wild-type = 2xG4S (clone_2).

A determinação da atividade da proteína de fusão GnTII / I no sobrenadante de células, e as amostras de ligado celular da proteína de fusão GnTII / I contendo o espaçador do tipo selvagem mostrou que a maioria da atividade está associada no interior das células e menor quantidade é secretada. Acredita-se que isto explica o porquê muito melhores sinais de His-tag GnTII / I foram vistas em frações de células, em vez de nas frações de sobrenadante em análise de Western blot.

A inibição de serina e proteases de cisteína por inibidor comprimido completo livre de EDTA, metaloproteases por EDTA e proteases aspárticas por pepstatina A, a melhoria do rendimento do GnTII / I da proteína de fusão. Esta observação sobre o uso de inibidor de serina protease é, de acordo com o trabalho de Salamin et al. (Appl. Environ. Microbiol., 76 (2010) 4269-4276), o que mostrou que a atividade de protease tipo serina nos meios de *P. pastoris* foi completada inibida com PMSF. Além disso, Vad et al. (J. Biotechnol. 116 (2005) 251-260) relatou a produção elevada, superior a 300 mg / l, da hormona paratiróide humana intacta em *P. pastoris*, na presença de EDTA a 10 mM juntamente com a co-expressão de *Saccharomyces cerevisiae*, a proteína isomerase de dissulfureto.

Todas as proteínas de fusão de GnTII / I que contêm cada uma dos quatro espaçadores da variante possuía atividade GnTII / I, bem como a atividade das enzimas que tenham as 2xG4S e EGIV espaçador da variante tinha atividades mais elevadas que a proteína de fusão GnTII / I contendo o

espaçador do tipo selvagem.

Exemplo 6 - Uso de proteínas de fusão com Man5 como o Glicano aceitante
Construção Cepa T. reesei Expressando Rituximab com N-glicosilação do
Tipo Man5

5 A sequência de rituximab nativa é códon harmonizada. Os plasmídeos de origem, contendo a síntese rituximab cadeia leve e cadeia pesada são gerados. As cadeias de anticorpos e proteínas de fusão de cbhl foram concebidos com sequências sobrepostas de 40 nucleotídeos como estão os vetores de expressão pHHO1 (marcador de seleção acetamidase, cbh1 flancos para a integração no local de cbh1) para a cadeia pesada ou
 10 pHHO2 (marcador de seleção higromicina, egl1 flancos para a integração no local egl1) para a cadeia leve, com a finalidade de permitir a clonagem de levedura usando recombinação homóloga.

Os plasmídeos de genes obtidos são transformados em E. coli.
 15 O DNA é preparado e os genes sintéticos são digeridos e isolados a partir das espinhas dorsais do plasmídeo. Os vetores de expressão de levedura são construídos por recombinação homóloga nos vetores de expressão de T. reesei com a proteína de fusão de cbhl e qualquer cadeia pesada ou leve. Os plasmídeos recombinantes são resgatados da levedura e transformados
 20 em E. coli. Após a triagem PCR, os clones corretos são isolados e sequenciados. As cassetes de expressão de fragmentos são digeridos e isolado a partir da espinha dorsal do plasmídeo, resultando em fragmentos de cerca de 10,2 kb para as construções de cadeias pesadas e fragmentos de 10,8 kb para as construções de cadeias leves. Os fragmentos de cadeia pesada e leve são
 25 cotransformados para a cepa de T. reesei M124. Os transformantes são selecionados por meio da resistência a higromicina e capacidade para crescer em acetamida como única fonte de nitrogênio. Os transformantes são semeados em um meio seletivo duplo para duas voltas sucessivas e testados por meio de PCR, para a integração de construtos de expressão no genoma.

30 Introdução de Enzima GnTII / I tandem e Manosidase II para cepa de T. reesei que expressa anticorpo rituximab

Em adição à introdução de uma GnTII recombinante / I em uma

cepa produtora de man5 como M124, uma atividade de manosidase II, é ainda necessário para remover duas manoses a partir da estrutura de modo a que GlcNAcMan5 glicano GnTII / I pode usar GlcNAcMan3 como uma molécula receptora .

5 O cassete de expressão GnTII / I descrito nos exemplos anteriores podem ser direcionados para, por exemplo, o local de cbh2 de *T. reesei*, utilizando métodos essencialmente como descrito acima. Para gerar uma molécula receptora GlcNAcMan3 para GnTII / I da proteína de fusão, a atividade de manosidase II é, em seguida, introduzida na cepa de transformação
10 utilizando os métodos descritos acima.

A atividade de Manosidase II é apresentada ao anticorpo rituximab-expressando cepa M124 criando um cassete de expressão contendo manosidase desejada com um promotor para dirigir a expressão de manosidase. Os promotores úteis são aqueles de *gpdA* ou *cbh1*. A atividade Manosidase II pode ser transformada através da integração aleatória, seguida por
15 meio da triagem de linhagens com níveis de expressão mais apropriados. O cassete de expressão está ligado a um gene marcador de seleção de propriedade, ou um marcador de seleção é cotransformado como um cassete de expressão separado. A transformação é realizada de acordo com métodos
20 descritos acima.

Um construto de fusão de manosidase II pode ser derivado de um *T. reesei* citoplasmático, transmembranar e domínio da haste, ou peptídeo alvo, de KRE2 e ligado em grelha a uma deleção de aminoácidos Terminal N de uma manosidase II humana. A proteína de fusão codificada localiza no RE / Golgi por meio do direcionamento KRE2 sequência peptídica,
25 mantendo a sua atividade de domínio catalítico de manosidase e capaz de hidrolisar GlcNAcMan5GlcNAc2 em GlcNAcMan3GlcNAc2. Em certas modalidades, uma manosidase humana de comprimento completo II pode ser expressa em uma cepa M124.

30 O peptídeo alvo KRE2 compreende os aminoácidos de cerca de 1 a cerca de 106 ou desde cerca de 1 a cerca de 83 KRE2.

Kre2 aa 1-106

MASTNARYVRYLLIAFFTILVFYFVSNSKYEGVDLNKGTFTAPDSTKTTPKPPATG
DAKDFPLALTPNDPGFNDLVGIAPGPRMNATFVTLARNSDVWDIARSIRQ (SEQ
ID NO: 115)

Kre2 aa 1-83

MASTNARYVRYLLIAFFTILVFYFVSNSKYEGVDLNKGTFTAPDSTKTTPKPPATG
DAKDFPLALTPNDPGFNDLVGIAPGPR (SEQ ID NO: 116)

Após a transformação de *Trichoderma* com o construto de manosidase II acima descrito, as cepas de *Trichoderma* são selecionadas, se-
meadas em meio selectivo para dois ciclos sucessivos, e testados por meio
de PCR, para a integração dos construtos de expressão no genoma. Os
transformantes selecionados a partir de cepas de *Trichoderma* que produzem
man5 e expressando a GnTII / proteína de fusão I, manosidase II, e o anticorpo
rituximab, são então cultivadas em frascos de agitação ou condições fermen-
tador e analisadas quanto ao teor de glicano como descrito acima.

Exemplo 7 - Expressão de GnTI e GnTII em *T. reesei*

Transformação de *T. reesei* M124 com Construto GnTI por meio da Integra- ção Aleatória

O códon otimizado humano Gntl foi transformado na cepa de *T.*
reesei M124. O gene Gntl foi clonado em um vetor sob o controle de dois
promotores diferentes : (1) o promotor induzível do gene *cbh1*, e (2) o pro-
motor constitutivamente expresso do gene *gpdA*. Os vetores contendo Gntl
sob qualquer um dos dois promotores foram cada um dos co-transformado
na cepa de *T. reesei* M124 com um plasmídeo contendo um ou acetamidase
ou um gene marcador de resistência à higromicina.

Trinta e quatro transformantes com Gntl sob o promotor de *gpdA*
e sob seleção acetamida foram rastreadas por meio de PCR, e todos foram
positivos para Gntl. Para transformantes com Gntl sob a *cbh1* promotor e
sob seleção acetamida, 19 dos 26 foram PCR-positivos para o construto Gn-
tl. Além disso, a extracção inicial do DNA foi realizada durante cinco cepas
com Gntl sob o promotor *cbh1* e sob seleção higromicina. Todas estas a-
mostras foram PCR-positivos. Vinte e cinco transformantes promotor *gpdA* e
todos os transformantes promotor *cbh1* (14 5) foram purificados a clones

mononuclear e suspensões de esporos foram preparadas.

Para fins de análise inicial, 23 transformantes de promotor *gpdA* e 19 transformantes de promotor *cbh1* (14 crescido de acetamida e cinco da seleção de higromicina), bem como a cepa parental M124 foram cultivadas em frascos de agitação de 250 ml com 50 ml de meio mínimo de *Trichoderma* fornecido com 2 % de extrato de grão gasto e 4 % de lactose. O crescimento das cepas foi monitorizado por medições de pH. Amostras de sobrenadante (e micélios) foram recolhidas nos dias 3, 5 e 7, armazenadas, congeladas até serem utilizadas para a análise da estrutura de glicano.

10 Análise de Glicano de de Cepas *GnT1* de *T. reesei* obtidas por meio da Integração Aleatória

A concentração de proteína de todas as amostras de sobrenadante foi medida por meio do ensaio à base de Bradford (BioRad Prbateladain Assay início rápido de Bradford), utilizando BSA como padrão. Teor de proteína segregada de amostras sujeitas a análise N-glicano foi ajustado para 5 µg ou 10 µg. A análise N-glicano foi realizada quer em pratos de 96 cavidades durante 5 µg de proteína sobrenadante, ou em tubos de 1,5 ml para 10 µg de proteína sobrenadante. Todas as análises de N-glicanos foram realizadas em triplicado. Ambos os N-glicanos neutros e ácidos foram analisados com MALDI-TOF MS.

Para obter medidas mais exatas da quantidade do produto produzido no *GnT1* *Gn1Hex5* quatro *GnT1* transformantes (dos dias 3 e 5) e também da quantidade de ácidos produzidos N-glicanos, os espectros de MALDI foi enriquecida com um glicano conhecido. Para os N-glicanos neutros e ácidos N-glicanos, um calibrador interno de 2 pmol / espectro Hex2HexNAc4 no valor de massa de 1,177 Da e 0,5 pmol de Hex4HexNAc2 monosialilated no valor de massa de 1,362 Da foram utilizados, respectivamente. As análises foram realizadas em triplicado.

Nenhum produto *GnT1* foi observado em qualquer dos transformantes promotor *gpdA*. No entanto, oito *cbh1* transformantes promotor produzido o produto *GnT1* *Gn1Man5* (Figuras 34 e 35 e Tabela 13), cinco com a seleção higromicina, três com a seleção acetamida.

Tabela 13: As percentagens de intensidades de sinal em comparação com Man5 e Gn1Man5 com calibrante interno Hex2HexNAc4 em quatro transformantes GnT1 positivos e da cepa parental M124 nos dias 3 e 5. Man5 é a principal glicoforma na cepa parental M124.

		M124 1, dia 3					M124 1, dia 5					M124/GnT1, clone HM1, dia 3					M124/GnT1, clone HM1, dia 5				
Composição	m/z	Média	SD	RSD	MIN	MAX	Média	SD	RSD	MIN	MAX	Média	SD	RSD	MIN	MAX	Média	SD	RSD	MIN	MAX
Hex2HexNAc4	1177.42	97.7	0.5	0.5	97.1	98.0	36.5	0.8	2.3	35.9	37.1	78.5	14.5	18.4	68.2	88.7	50.1	10.6	21.2	42.6	57.6
Hex5HexNAc2	1257.42	2.3	0.5	22.5	2.0	2.9	63.5	0.8	1.3	62.9	64.1	14.5	9.9	68.0	7.5	21.5	44.0	9.6	21.9	37.2	50.8
Hex5HexNAc3	1460.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.1	4.6	65.6	3.8	10.3	5.9	1.0	16.7	5.2	6.6
		M124/GnT1, clone 8, dia 3					M124/GnT1, clone 8, dia 5					M124/GnT1, clone 39, dia 3					M124/GnT1, clone 39, dia 5				
Composição	m/z	Média	SD	RSD	MIN	MAX	Média	SD	RSD	MIN	MAX	Média	SD	RSD	MIN	MAX	Média	SD	RSD	MIN	MAX
Hex2HexNAc4	1177.42	77.3	7.6	9.8	72.0	82.7	67.3	10.0	14.9	56.5	76.3	83.7	1.5	1.8	82.7	84.8	40.0	1.9	4.6	37.9	41.1
Hex5HexNAc2	1257.42	15.0	5.2	34.4	11.4	18.7	18.9	6.2	32.5	12.8	25.1	8.3	1.0	11.7	7.6	8.9	46.9	1.8	3.8	45.6	49.0
Hex5HexNAc3	1460.5	7.6	2.4	31.6	5.9	9.3	13.8	4.0	29.1	10.8	18.3	8.0	0.6	6.9	7.6	8.4	13.1	0.3	2.1	12.7	13.3
		M124/GnT1, clone 90, dia 3					M124/GnT1, clone 90, dia 5														
Composição	m/z	Média	SD	RSD	MIN	MAX	Média	SD	RSD	MIN	MAX										
Hex2HexNAc4	1177.42	93.8	1.6	1.7	92.4	95.6	92.6	2.7	2.9	89.8	95.3										
Hex5HexNAc2	1257.42	3.7	1.0	25.9	2.6	4.5	4.7	1.4	30.9	3.2	6.0										
Hex5HexNAc3	1460.5	2.5	0.7	26.2	1.8	3.1	2.7	1.3	47.8	1.5	4.1										

5

Os produtos GnT1 Gn1Man6P1, Gn1Man7P1 e Gn1Man8P1 também foram encontrados em N-glicanos fosforilados de todos os transformantes positivos. A quantidade de N-glicanos fosforilados tinha aumentado em GnT1 transformantes, e o perfil foi inclinado para N-glicanos maiores, com Man7P1 ou Man8P1 tendo o sinal mais forte (em Man6P1 parental M124) (Figura 36).

10

Oito GnT1 transformantes produziram a estrutura Gn1Man5. Gn1Man5 era mais abundante no clone 39. No entanto, o melhor clone pareceu ser o clone 8, que produziu o segundo nível mais elevado de Gn1Man5, mas tinha uma elevada proporção de man5 e Gn1Man5 (Figura 35). O clone 8, que contém GnT1 sob o controle do promotor de cbhl, foi denominado cepa M198, e foi selecionado para a análise continuada.

15

Transformação da Cepa M198 de *T. reesei* com Construto GnTII por meio da Integração Alvo

20

Cinco vetores abrigando GnTII foram criados (Tabela 14). Dois dos vetores continha os peptídeos alvo Golgi em mamíferos nativos em GN-TII. Nos outros três vetores, o peptídeo alvo de mamífero foi substituído por um de *T. reesei* MNT1 (α -1,2-manosiltransferase) dirigido ao peptídeo. Todos os cinco vetores continham um promotor cbh1 ou um promotor gpdA e um marcador de pyr4 loop-out. Além disso, todos os cinco vetores foram direcionados para integrar o local alg3, apagando assim o gene alg3. Nos

25

construtos MNT1/GnTII sob a *cbh1* promotor, duas diferentes exclusões da sequência GnTII foram testadas.

Tabela 14 : vetores GNT2 Construídos.

Nome de plasmídeo	Promotor	Peptídeo alvo	Deleção do Terminal N (GnTII)
pTTv140	<i>cbh1</i>	Mamífero	N/A
pTTv141	<i>gpdA</i>	Mamífero	N/A
pTTv142	<i>cbh1</i>	<i>Trichoderma</i> MNT1	74 aminoácidos
pTTv143	<i>cbh1</i>	<i>Trichoderma</i> MNT1	104 aminoácidos
pTTv144	<i>gpdA</i>	<i>Trichoderma</i> MNT1	74 aminoácidos

Esses vetores, exceto para o vetor pTTv144, foram transformados na melhor cepa GnTI produtora py4 negativa M198 (M319) como fragmentos Pmel. Os transformantes foram purificados a clones mononucleares e PCR triados. Os clones mostrando a integração correta em ambas as extremidades foram então selecionados para a análise continuada.

Para estudar as características de crescimento das tensões geradas expressando GNTII, grandes culturas frasco de agitação foram preparadas. A cultura em frasco agitado foi preparada em duas bateladas separadas. A primeira batelada continha pTTv140, pTTv142 e pTTv143. O segundo batelada continha pTTv141. A cepa parental M198 foi utilizada como uma cepa de controle. As células foram cultivadas em meio TRMM suplementado com 40 g / l de lactose, 20 g / l de extrato de grão gasto, e PIPPS 100 mM, pH 5,5. Cinco transformantes por construto foram cultivadas. O pTTv140, pTTv142 e pTTv143 culturas foram amostrados nos dias 3, 5, 7 e 9. Os pTTv141 culturas foram amostrados nos dias 3, 5, 7 e 10. O pH e a célula de peso seco de cada amostra foram medidos e as amostras de sobrenadante de cultura foram usados para a análise da estrutura de glicano.

Análise de Glicano de cepas de *T. reesei* obtidas visando GnTII para o local *alg3* de Cepa *T. reesei* M198

Cinco clones diferentes contendo o vetor pTTv140 (contendo o peptídeo alvo nativo e o promotor de *cbh1*), o vetor pTTv142 (contendo o peptídeo alvo MNT1, a 74 aa GNTII deleção Terminal N, e o promotor de

cbhl), o vetor pTTv143 (contendo o peptídeo MNT1 direcionamento, o aa 110 deleção Terminal N GNTII, e o promotor de cbhl) e a pTTv141 vetor (contendo o peptídeo alvo e o promotor gpdA) foram analisados.

As análises de N-glicanos foram preparadas em triplicado, para as amostras 5 dias e, em duplicado, para o dia 3 e 7 amostras em placas de 96 cavidades durante 5 µg de proteína sobrenadante. As concentrações de proteína dos sobrenadantes foram medidos por ensaio à base de Bradford (BioRad Prbateladain Assay início rápido de Bradford), utilizando BSA como padrão. Reações PNGase F foram realizadas como descrito. Os N-glicanos libertados foram purificados em primeiro lugar com Hypersep C-18, 100 mg e, em seguida, com 10 mg Hypersep Hypercarb (ambos a partir de Thermo Scientific) onde glicanos neutros e ácidos foram separadas. Ambas as purificações foram realizadas no formato de 96 cavidades. N-glicanos neutros foram analisados por meio de MALDI-TOF MS.

N-glicanos de quatro diferentes cepas transformadas com GnTII foram analisadas. O clone 117A-1, que foi transformado com o vetor pTTv140, e, portanto, continha o peptídeo alvo nativo e o promotor de cbhl, produziu cerca de 40 % de G0 e cerca de 13 % de Hex6 (Figura 37A). Os clones pTTv143 transformados com o vetor, contendo o peptídeo, dessa maneira, MNT1 de direcionamento, a deleção 110 aa do Terminal N de GnTII, e o promotor de cbhl, produziu cerca de 10 % de G0 (Figura 37C). Clone 3B, que continha o promotor gbdA produziu cerca de 28 % de G0 e cerca de 19 % de Hex6 (Figura 37D).

Os padrões de glicosilação de clones representativos contendo o pTTv140, pTTv141 e pTTv142 vetores também mostraram ser estável em função do tempo (Figura 38).

Glicosilação de Proteína específica

Para analisar as mudanças específicas na glicosilação das proteínas, amostras da pTTv142 clone contendo o vetor e 3-17A a partir da cepa parental M198 foram separados por SDS-PAGE e riscados para uma membrana de PVDF. As bandas de proteína de interesse (quatro bandas de M198 e quatro do clone 3-17A) foram excisadas e os N-glicanos foram liber-

tados com a membrana de liberação enzimática com PNGase F (Figura 39).

Os N-glicanos neutros isolados e purificados foram analisados utilizando MALDI-TOF MS. O padrão de glicosilação do total das proteínas secretadas foi semelhante a uma proteína de 50 kDa separado da cepa M198 parental (Figura 40). A banda da proteína de menor tamanho não foi glicosilada.

Em GnTII do clone 3-17A, a maioria dos sinais atípicos tinham desaparecido, o que confirma a sua origem a partir do meio. Além disso, o padrão de glicosilação do clone 3-17A difere dos padrões de glicano de proteínas segregadas totais (Figura 40B). A quantidade de G0 a partir do clone 3-17A foi de cerca de 35 a 36 % (Figura 40B).

Cultivo do fermentador da cepa GnTII

O cultivo do fermentador da cepa GnTII 1-117A M329 (que contém o pTTv140vetor) foi fermentado em TRMM pH 5,5 + 2 % de extrato de grãos Passado + 6 % de lactose + 0,5 % de KH₂PO₄ + 0,5 % de (NH₄)₂SO₄ a 28°C (pH 5,5). A análise N-glicano foi realizada em triplicado, a 5 µg de proteínas secretadas descritos na seção "proteína específica de glicosilação" acima em amostras colhidas no dia 3. A quantidade de G0 era cerca de 48 % e a quantidade de Hex6 foi de cerca de 19 % no dia 3 (Figura 41).

20 Exemplo 8 - homólogos T. reesei ALG3

Transformação de T. reesei M124 com Construto GnTI por meio da Integração Aleatória

Os homólogos T. reesei ALG3 foram identificados a partir de outros organismos. Esses homólogos podem ser utilizados para conceber ALG3 para outros construtos de deleção de células T. reesei fúngicas filamentosas. Os homólogos ALG3 estão listados na Tabela 15. Um alinhamento da sequência de aminoácido múltiplo de T. reesei ALG3 e homólogos ALG3 são mostrados na Figura 42.

Tabela 15: Homólogos ALG3.

seqüência de referência	Organismo	SEQ ID NO :
Trire2 104121 fgenes5_pg.C_scaffold_3000076	<i>Trichoderma reesei</i>	126
Triat2 270085 fgenes1_pg.contig_14_#_149	<i>Trichoderma atroviride</i>	127
TriviGv29_8_2 194462 fgenes1_pm.87_#_115	<i>Trichoderma virens</i>	128
EGU81920.1	<i>Fusarium oxysporum Fo5176</i>	129
XP_389829.1	<i>Gibberella zeae PH-1</i>	130
AEO60805.1	<i>Myceliophthora thermophila</i>	131
XP_962259.1	<i>Neurospora crassa OR74A</i>	132
XP_001824044.1	<i>Aspergillus oryzae RIB40</i>	133
XP_001259497.1	<i>Neosartorya fischeri NRRL 181</i>	134
XP_001398696.2	<i>Aspergillus niger CBS 513.88</i>	135
XP_362427.2	<i>Magnaporthe oryzae 70-15</i>	136
NP_593853.1	<i>Schizosaccharomyces pombe 972h</i>	137

Exemplo 9 - Variantes de Proteína de Fusão GnTII/GnTIGeração do construto de expressão GnTII / GnTI

Uma proteína de fusão GnTI / II recombinante, sob o controle do promotor de cbh1 indutível e contendo um de quatro espaçador da variante é

5 construída como descrito nos Exemplos 4 e 5. As quatro variantes de espaçador são 2xG4S o espaçador, o espaçador 3xG4S, o espaçador de cbhl e o espaçador EGIV.

Em suma, os fragmentos de fusão são amplificados a partir de modelos GnTII e GnTI separadamente com iniciadores contendo 50 pb sobreposições in-estrutura no local de fusão. Os fragmentos são purificados a

10 partir de um gel de agarose e utilizados como modelo de PCR para a amplificação do construto de fusão de acordo com os procedimentos padrão. O construto de fusão foi clonado em um vetor com locais de restrição Apal / Spel, sob o controle do promotor de cbh1 indutível. Além disso, o peptídeo

15 alvo do mamífero nativo Golgi no domínio GNTII foi substituído por um de T. reesei MNT1 (α -1,2-manosiltransferase) dirigido ao peptídeo.

Para introduzir o espaçador da variante 2xG4S na proteína de fusão, T45 sequência é amplificado em duas partes, utilizando a estratégia PCR de sobreposição. Primeiro, um fragmento é amplificado com AKT1-maio 6-1

20 'iniciador
(GGTACCGGGCCCACTGCGCATCATGCGCTTCCGAATCTACAAGCG
(SEQ ID NO : 146)) e GP93 iniciador 3', e um segundo fragmento foi amplificado com GP92 iniciador 5 'e AKT1 -6 4 3' iniciador (GGCGCGCCAC-TAGTCTAATTCCAGCTGGGATCATAGCC (SEQ ID NO : 147)). A amplificação

25 ção é realizada com Fusão de alta fidelidade da polimerase PCR (Finnzymes) sob as condições padrão fornecidos pelo fornecedor. As condições de amplificação são como descritos no Exemplo 5. O produto de PCR resultante foi purificado a partir do gel de agarose, e os fragmentos que se sobrepõem, com sequências modificadas são combinados na mesma mistura de

30 reação com condições convencionais sem iniciadores. Dez ciclos de recozimento / extensão são realizados como descrito no Exemplo 5. Os iniciadores AKT1-6-1 (5 ') e AKT1-6-4 (3') são adicionados, e a bicicleta é continuado tal

como descrito no Exemplo 5 para 20 ciclos de amplificação. T45 O fragmento amplificado é então purificado, digerido com Apal / SpeI (New England Biolabs) de acordo com protocolos convencionais, e clonado no vetor de expressão de *Trichoderma reesei*. O fragmento clonado é então verificado por meio do sequenciamento com o conjunto apropriado de iniciadores e a sequência gerada é usada para o construto de *T. reesei* vetor de expressão com o promotor 2xG4S e *alg3* alvo.

O plasmídeo resultante é utilizado como molde para a modificação do espaçador 3xG4S. A clonagem da sequência T46 é realizada como descrito acima com o T45. O iniciador AKT1-6-1 5'- e o iniciador GP95 3'- são utilizados para a síntese do primeiro fragmento, e GP94 5'-iniciador e AKT1-6-4 3'-iniciador são utilizados para a síntese da segunda fragmento. Os fragmentos são combinados, e os iniciadores AKT1-6-1 (5') e AKT1-6-4 (3') são adicionados para a amplificação. O fragmento T46 amplificado é então digerido com Apal / SpeI e clonado no vetor de expressão de *Trichoderma reesei*. O fragmento clonado é então verificado por meio do sequenciamento com um conjunto adequado de iniciadores e a sequência gerada é usada para o construto de *T. reesei* vetor de expressão com promotor 3xG4S e *alg3* alvo.

O CBHI e espaçadores EGIV são construídos com um método de sobreposição de PCR similares. Para o espaçador de CBHI, o primeiro fragmento é amplificado com AKT1-6-1 e iniciador 5'-GP107 3'. O segundo fragmento foi amplificado com o iniciador 5'-GP108 e AKT1-6-4 3'-iniciador (Tabela 11). Para o espaçador EGIV, o primeiro fragmento é amplificado com AKT1-6-1 e 5'iniciador 3'iniciador GP109. O segundo fragmento foi amplificado com o iniciador 5'-GP110 e AKT1-6-4 3'-iniciador (Tabela 11). Em ambos os casos, os produtos de PCR são purificados a partir de gel de agarose, combinados e usadas como um molde para a próxima reação de PCR para amplificar as sequências de T50 e T51. T50 e T51 produtos de PCR são então digerido com Apal / SpeI e clonado no vetor de expressão de *Trichoderma reesei*. Os fragmentos clonados são, então, verificados por sequenciação com conjuntos apropriados de iniciadores e as sequências gera-

das são utilizados para construto de vetores de expressão de *T. reesei* com qualquer cbhl ou promotor EGIV e alg3 alvo.

5 Todas as amplificações por meio de PCR são feitas com a Fusão de alta fidelidade da polimerase (Finnzymes). Os iniciadores (Tabela 11) são ordenados de MWG Operon. A sequenciação é realizada pelo Laboratório de Sequenciação de DNA do Instituto de Biotecnologia, Universidade de Helsínquia, como um serviço comercial.

Os vetores de expressão de *Trichoderma reesei* com as sequências GnTII / GnTI quiméricas descritas com variações separadoras (2xG4S, 3xG4S, cbhl e EGIV) são subclonados sob o controle do promotor de cbh1, com um marcador LOOPOUT gene pyr4 e alg3 flanqueando os fragmentos da região para integração alvo no esqueleto são, então, construídos. As cassetes de expressão são transformados em *T. reesei* cepa M279 (pyr4-cepa de M202). Após a seleção da placa, os clones PCR são selecionados e purificados através de esporos individuais. Com a finalidade de se obter material para análises de glicano, as culturas agitadas em frasco são executadas como descrito.

Introdução de variantes de Proteína de Fusão GnTII / I na Cepa *T. reesei* Expressando o anticorpo rituximab

20 As variantes de proteínas de fusão GnTII / I recombinantes são introduzidas em rituximab expressando a cepa *T. reesei* M279 descrito no Exemplo 5.

Em suma, os vetores contendo a proteína de fusão GnTII / GnTI sob o controle do promotor de cbh1, o peptídeo MNTI alvo, o marcador de quadro de saída pyr4, e cada uma das quatro variantes de espaçador são, cada um subclonados em um vetor da estrutura principal entre os fragmentos alg3 da região para a integração alvo, apagando assim o gene alg3. Um cassete de expressão Pmel digerido é transformado em *T. reesei* cepa M279 (pyr4 uma exigência). Após a seleção da placa, os clones são PCR-selecionados e purificados através de esporos individuais.

Análise de Glicano de cepas variantes GnTII / GnTI de *T. reesei* Produzindo Rituximab obtidas visando o Local *alg3*

5 Com a finalidade de se obter material para análise de glicano, as culturas agitadas em frasco são realizadas como descrito no Exemplo 5 e, além disso, alguns meios de cultura são suplementadas com 0,3 mg / ml de inibidor de tripsina de soja (SBTI) e 1 % de casaminoácidos. SBTI é adicionado pela primeira vez na inoculação e, em seguida, por dia nos dias 3-6. PMSF e Pepstatina A é adicionada a todas as amostras antes de congelamento.

10 O rituximab é purificado com a proteína G da cromatografia de afinidade a partir de dia 5 amostras de sobrenadante com SBTI e desde o dia 5 e 7 amostras sem SBTI. As reações são realizadas com PNGase F para ~ 10 µg de proteína desnaturada. Os N-glicanos libertados são purificados primeiro com Hypersep C-18 e, em seguida, com Hypersep Hypercarb
15 (ambos a partir de Thermo Scientific) onde os glicanos neutros e ácidos são separados. As etapas de purificação são realizadas no formato de 96 cavidades. Os N-glicanos Neutros e ácidos são analisadas por meio de MALDI-TOF MS para testar a presença da glicoforma G0 no anticorpo rituximab.

REIVINDICAÇÕES

1. Proteína recombinante tendo atividade N-acetilglucosaminiltransferase, caracterizada pelo fato de que a proteína recombinante catalisa a transferência de N-acetilglucosamina a um resíduo Man α 3 terminal e catalisa a transferência de N-acetilglucosamina a um resíduo Man α 6 terminal de um glicano aceitador, e em que a proteína recombinante compreende um domínio catalítico a partir de pelo menos duas enzimas diferentes.

2. Proteína recombinante de acordo com a reivindicação 1, caracterizada pelo fato de que a proteína recombinante é uma proteína de fusão que compreende um domínio catalítico de N-acetilglucosaminiltransferase I e um domínio catalítico de N-acetilglucosaminiltransferase II.

3. Proteína recombinante de acordo com a reivindicação 2, caracterizada pelo fato de que o domínio catalítico de N-acetilglucosaminiltransferase I e o domínio catalítico de N-acetilglucosaminiltransferase II são de enzimas humanas.

4. Proteína recombinante de acordo com a reivindicação 3, caracterizada pelo fato de que o domínio catalítico de N-acetilglucosaminiltransferase I compreende uma sequência que é pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica aos resíduos de aminoácidos 105 a 445 da SEQ ID NO : 1.

5. Proteína recombinante de acordo com a reivindicação 3 ou 4, caracterizada pelo fato de que o domínio catalítico de N-acetilglucosaminiltransferase II compreende uma sequência que é pelo menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica aos resíduos de aminoácidos 30 a 447 da SEQ ID NO : 21.

6. Proteína recombinante de acordo com qualquer uma das reivindicações 2 a 5, caracterizada pelo fato de que o domínio catalítico de N-acetilglucosaminiltransferase II é um terminal N para o domínio catalítico de

N-acetilglucosaminiltransferase I.

7. Proteína recombinante de acordo com qualquer uma das reivindicações 2 a 6, caracterizada pelo fato de que compreende ainda um espaçador entre o domínio catalítico de N-acetilglucosaminiltransferase I e o
5 domínio catalítico de N-acetilglucosaminiltransferase II.

8. Proteína recombinante de acordo com a reivindicação 7, caracterizada pelo fato de que o espaçador compreende uma sequência do domínio de haste.

9. Proteína recombinante de acordo com a reivindicação 7 ou 8,
10 caracterizada pelo fato de que o espaçador compreende uma sequência selecionada a partir do grupo que consiste em SEQ ID NO: 118, SEQ ID NO : 120, SEQ ID NO : 122 e SEQ ID NO : 124.

10. Proteína recombinante de acordo com qualquer uma das reivindicações 2 a 9, caracterizada pelo fato de que compreende ainda um peptídeo alvo ligado à extremidade do terminal N de um dos domínios catalíticos.
15

11. Proteína recombinante de acordo com a reivindicação 10, caracterizada pelo fato de que o peptídeo alvo compreende um domínio de haste, um domínio transmembranar ligado à extremidade do terminal N do
20 domínio da haste e, opcionalmente, um domínio citoplasmático ligado à extremidade do terminal N do domínio transmembranar.

12. Proteína recombinante de acordo com a reivindicação 8 ou 11, caracterizada pelo fato de que o peptídeo alvo, o domínio de haste, o domínio transmembranar e/ou o domínio citoplasmático é de uma proteína selecionada do grupo que consiste em uma manosidase, uma manosiltransferase, uma glicosiltransferase, uma proteína de Golgi do tipo 2, MNN2, MNN4, MNN6, MNN9, MNN10, MNS1, KRE2, VAN1 e OCH1 ou de uma proteína da Tabela 1.
25

13. Proteína recombinante de acordo com a reivindicação 12,
30 caracterizada pelo fato de que a proteína é de um organismo selecionado a partir do grupo que consiste em *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filibasidium*,

Fusarium, Gibberella, Humicola, Magnaporthe, Mucor, Myceliophthora, Myrothecium, Neocallimastix, Neurospora, Paecilamyces, Penicillium, Piromyces, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, e Trichoderma.

5 14. Proteína recombinante de acordo com a reivindicação 1 ou
2, caracterizada pelo fato de que compreende um domínio catalítico de N-acetilglucosaminiltransferase II humana e um domínio catalítico de N-acetilglucosaminiltransferase I humana, em que o domínio catalítico de N-acetilglucosaminiltransferase II está localizado no terminal N para o domínio cata-
10 lítico de N-acetilglucosaminiltransferase I, uma sequência de espaçador que compreende a sequência de um domínio de haste de N-acetilglucosaminiltransferase I humana localizada entre os domínios catalíticos, e um peptídeo alvo localizado no terminal N para o domínio catalítico de N-acetilglucosaminiltransferase II, em que o peptídeo alvo compreende um domínio
15 citoplasmático, um domínio transmembranar e um domínio de haste a partir de N -acetilglucosaminiltransferase II humana.

 15. Proteína recombinante de acordo com a reivindicação 1 ou
2, caracterizada pelo fato de que compreende uma sequência que é pelo
20 menos 70 %, pelo menos 75 %, pelo menos 80 %, pelo menos 85 %, pelo menos 90 %, pelo menos 95 %, pelo menos 96 %, pelo menos 97 %, pelo menos 98 %, pelo menos 99 %, ou 100 % idêntica à SEQ ID NO: 95.

 16. Proteína recombinante de acordo com a reivindicação 1 ou
2, caracterizada pelo fato de que compreende um domínio catalítico de N-acetilglucosaminiltransferase II humana e um domínio catalítico de N-acetilglucosaminiltransferase I, em que o domínio catalítico de N-acetilglucosaminiltransferase II está localizado no terminal N para o domínio catalítico de
25 N-acetilglucosaminiltransferase I, um espaçador localizado entre os domínios catalíticos, em que o espaçador compreende a sequência selecionada do grupo que consiste em SEQ ID NO : 118, SEQ ID NO : 120, SEQ ID NO :
30 122 e SEQ ID NO : 124, e um peptídeo alvo localizado no terminal N para o domínio catalítico N- acetilglucosaminiltransferase II em que o peptídeo alvo compreende um domínio citoplasmático, um domínio transmembranar, e um

domínio de haste de N-acetilglucosaminiltransferase II humana.

17. Proteína recombinante de acordo com a reivindicação 16, caracterizada pelo fato de que o espaçador compreende uma sequência selecionada do grupo que consiste em SEQ ID NO: 118, SEQ ID NO: 120 e
5 SEQ ID NO: 124.

18. Polinucleotídeo isolado, caracterizado pelo fato de que codifica a proteína recombinante, como definida em qualquer uma das reivindicações 1 a 17.

19. Vetor de expressão, caracterizado pelo fato de que compreende o polinucleotídeo isolado, como definido na reivindicação 18, operacionalmente ligado a um promotor.
10

20. Célula hospedeira, caracterizada pelo fato de que compreende o vetor de expressão, como definido na reivindicação 19.

21. Método de produção de um N-glicano complexo, caracterizado pelo fato de que compreende:
15

(1) proporcionar uma célula hospedeira, em que a célula hospedeira compreende um polinucleotídeo que codifica uma proteína de fusão que compreende um domínio catalítico de N-acetilglucosaminiltransferase I e um domínio catalítico de N-acetilglucosaminiltransferase II; e

(2) cultura da célula hospedeira de tal modo que a proteína de fusão é expressa, em que a proteína de fusão catalisa a transferência de N-acetilglucosamina a um resíduo Man α 3 terminal e N-acetilglucosamina a um resíduo Man α 6 terminal de um glicano aceitador com a finalidade de produzir um N- glicano complexo.
20

22. Método de acordo com a reivindicação 21, caracterizado pelo fato de que o glicano aceitor está ligado a um polipeptídeo heterólogo.
25

23. Método de acordo com a reivindicação 21 ou 22, caracterizado pelo fato de que o N-glicano complexo é GlcNAc β 2Man β 3 (GlcNAc β 2 Man β 6) Man β 4GlcNAc β 4GlcNAc.

24. Método de acordo com qualquer uma das reivindicações 21 a 23, caracterizado pelo fato de que a célula hospedeira é uma célula fúngica filamentosa selecionada do grupo que consiste em *Trichoderma sp.*,
30

Acremonium, Aspergillus, Aureobasidium, Cryptococcus, Chrysosporium, Chrysosporium lucknowense, Filibasidium, Fusarium, Gibberella, Magnaporthe, Mucor, Myceliophthora, Myrothecium, Neocallimastix, Neurospora, Paecilamyces, Penicillium, Piromyces, Schizophyllum, Talaromyces, Thermoascus, Thielavia, e Tolypocladium.

25. Método de acordo com qualquer uma das reivindicações 21 a 24, caracterizado pelo fato de que a célula hospedeira compreende adicionalmente um polinucleotídeo que codifica um transportador de UDP-GlcNAc.

26. Método de acordo com qualquer uma das reivindicações 21 a 25, caracterizado pelo fato de que a célula hospedeira tem um nível reduzido de atividade de um dolichil-P-Man: Man (5) GlcNAc (2)-PP-dolichil manosiltransferase em comparação ao nível de atividade em uma célula hospedeira do tipo selvagem ou o gene *alg3* é deletado da célula hospedeira.

27. Método de acordo com qualquer uma das reivindicações 21 a 26, caracterizado pelo fato de que a célula hospedeira tem um nível reduzido de atividade de uma α -1,6- manosiltransferase ou um nível reduzido de expressão de um gene *och1* em comparação ao nível da atividade ou expressão de uma célula hospedeira do tipo selvagem.

28. Método de acordo com qualquer uma das reivindicações 21 a 27, caracterizado pelo fato de que a célula hospedeira compreende adicionalmente:

- i) um polinucleotídeo que codifica uma α -1,2-manosidase;
- ii) um polinucleotídeo que codifica uma β -1,4-galactosiltransferase; e/ou
- iii) um polinucleotídeo que codifica uma sialiltransferase.

29. Método de produção de um N-glicano complexo, caracterizado pelo fato de que compreende a incubação de uma proteína de fusão que compreende um domínio catalítico de N-acetilglucosaminiltransferase I e um domínio catalítico de N-acetilglucosaminiltransferase II, um glicano aceitador e um doador de N-acetilglucosamina em conjunto em um tampão, em que a proteína de fusão catalisa a transferência de N-acetilglucosamina a um resíduo Man α 3 terminal e N-acetilglucosamina a um resíduo Man α 6 ter-

minal de um glicano aceitador com a finalidade de produzir um N-glicano complexo.

30. Célula fúngica filamentosa tendo uma redução do nível de expressão de um gene *alg3* em comparação com o nível de expressão em uma célula fúngica filamentosa do tipo selvagem ou em que o gene *alg3* é deletado, caracterizada pelo fato de que compreende uma proteína recombinante como definida em qualquer uma das reivindicações 1 a 17.

31. Célula fúngica filamentosa de acordo com a reivindicação 30, caracterizada pelo fato de que a célula fúngica filamentosa compreende ainda um polinucleotídeo que codifica um transportador de UDP-GlcNAc.

32. Célula fúngica filamentosa de acordo com a reivindicação 30 ou 31, caracterizada pelo fato de que a célula fúngica filamentosa tem um nível reduzido de atividade de α -1,6-manosiltransferase em comparação ao nível da atividade de uma célula fúngica filamentosa do tipo selvagem.

33. Célula fúngica filamentosa de acordo com qualquer uma das reivindicações 30 a 32, caracterizada pelo fato de que a célula fúngica filamentosa tem um nível reduzido de expressão de um gene *och1* em comparação com o nível de expressão em uma célula fúngica filamentosa do tipo selvagem.

34. Célula fúngica filamentosa de acordo com qualquer uma das reivindicações 30 a 33, caracterizada pelo fato de que a célula fúngica filamentosa compreende ainda

i) um polinucleotídeo que codifica uma α -1,2-manosidase;

ii) um polinucleotídeo que codifica uma β -1,4-galactosiltransferase; e/ou

iii) um polinucleotídeo que codifica uma sialiltransferase.

35. Célula fúngica filamentosa de acordo com qualquer uma das reivindicações 30 a 34, caracterizada pelo fato de que a célula fúngica filamentosa é selecionada do grupo que consiste em *Trichoderma sp.*, *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Chrysosporium*, *Chrysosporium lucknowense*, *Filibasidium*, *Fusarium*, *Gibberella*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Myrothecium*, *Neocallimastix*, *Neurospora*, *Paeci-*

lamyces, Penicillium, Piromyces, Schizophyllum, Talaromyces, Thermoascus, Thielavia, e Tolypocladium .

- 5 36. Invenção, caracterizada por qualquer forma de suas concretizações ou por qualquer categoria aplicável de reivindicação, por exemplo, de produto ou de processo ou de uso ou de uso para preparar uma composição/medicamento englobados pela matéria inicialmente descrita, revelada ou ilustrada no pedido de patente.

FIG. 1

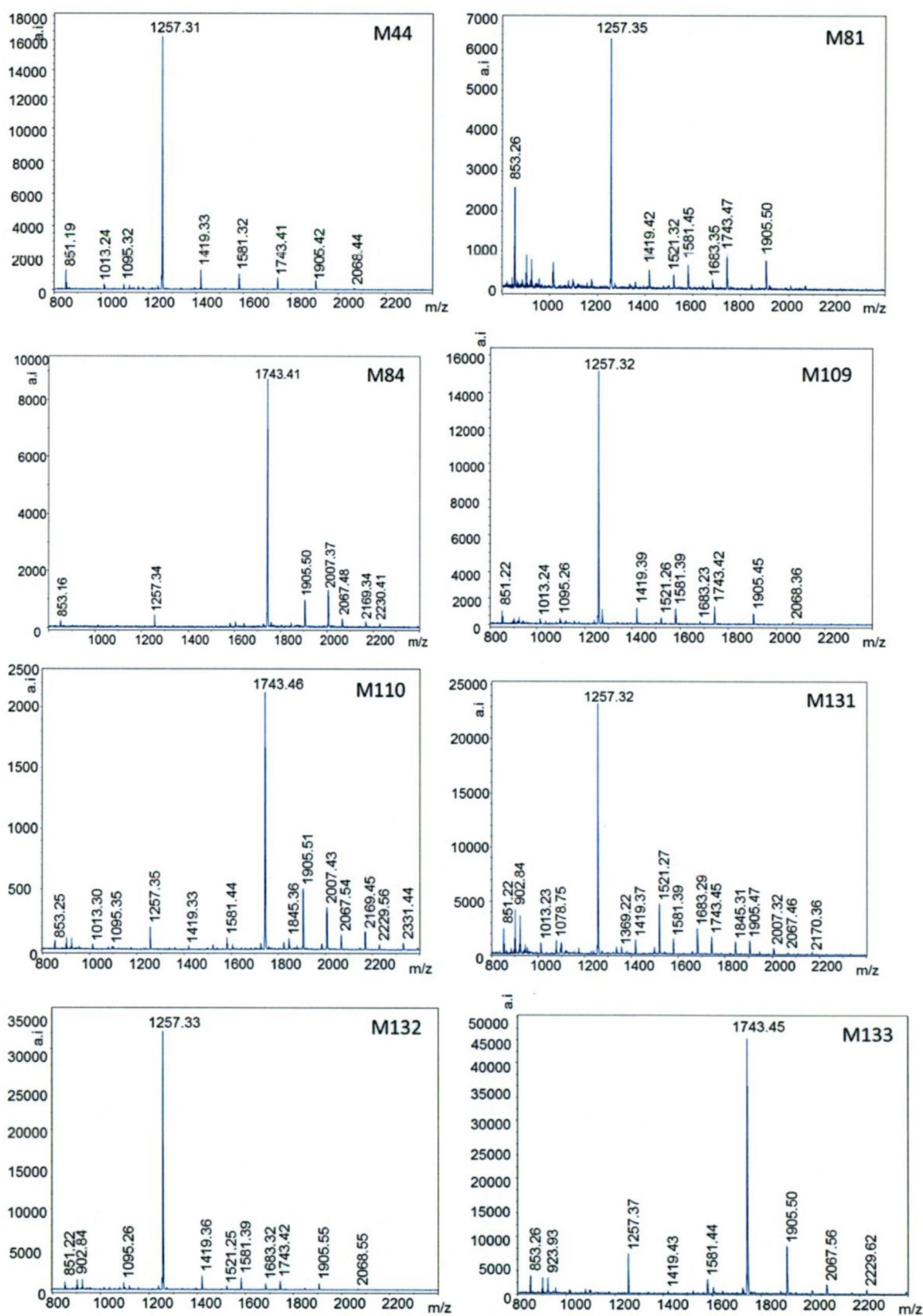


FIG. 1 (Continuação)

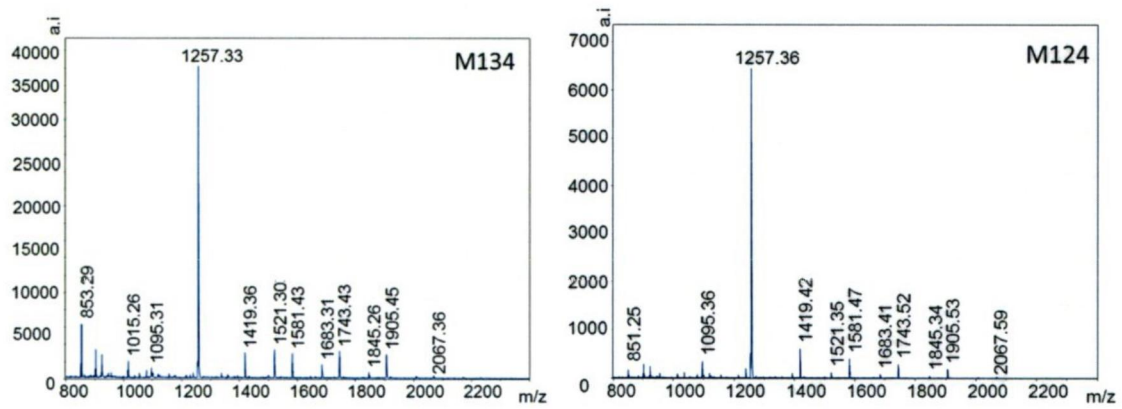


FIG. 2

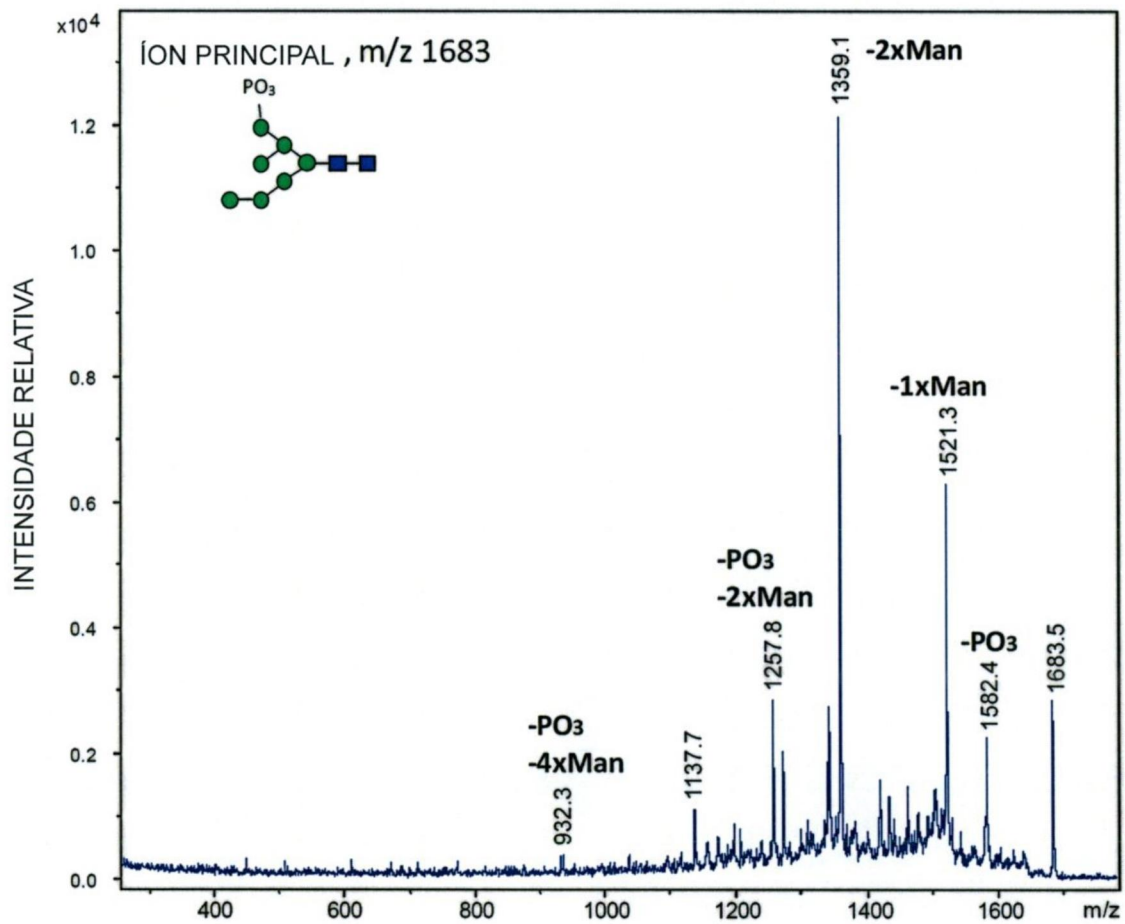


FIG. 3

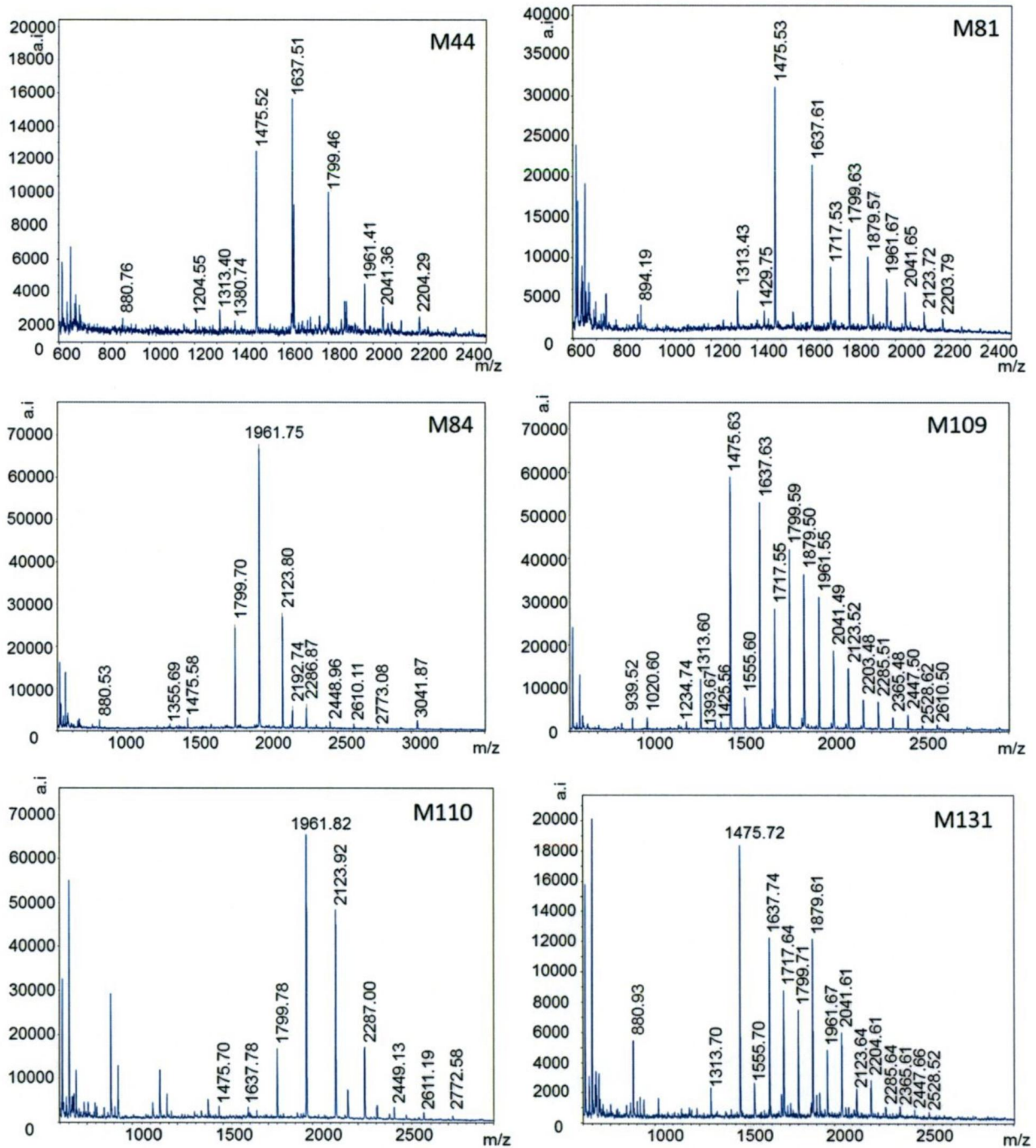


FIG. 3 (Continuação)

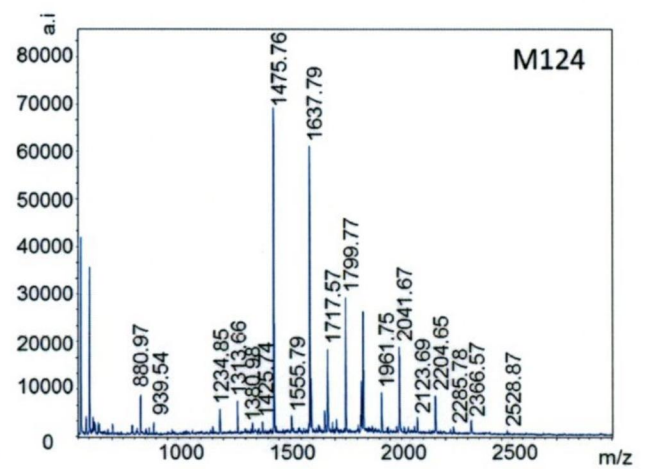
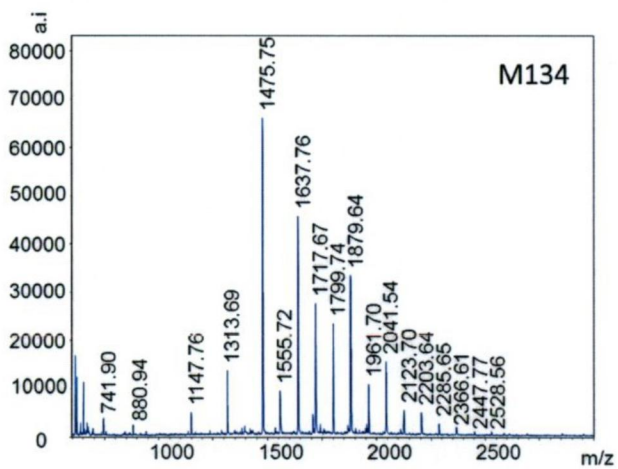
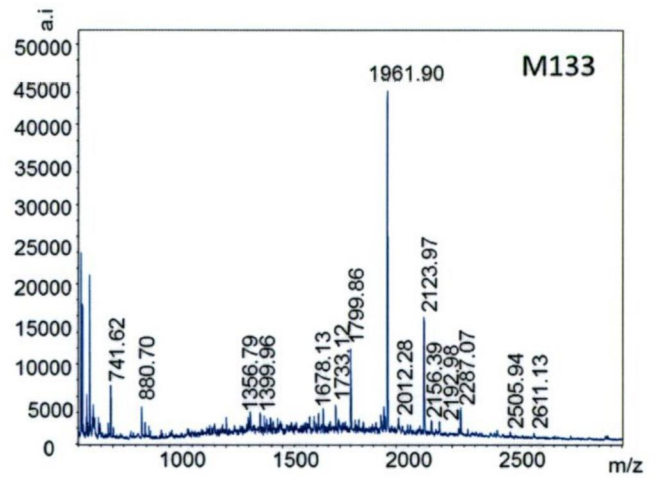
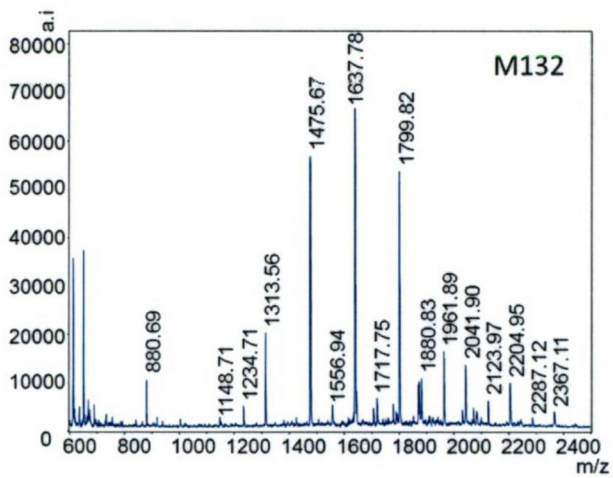
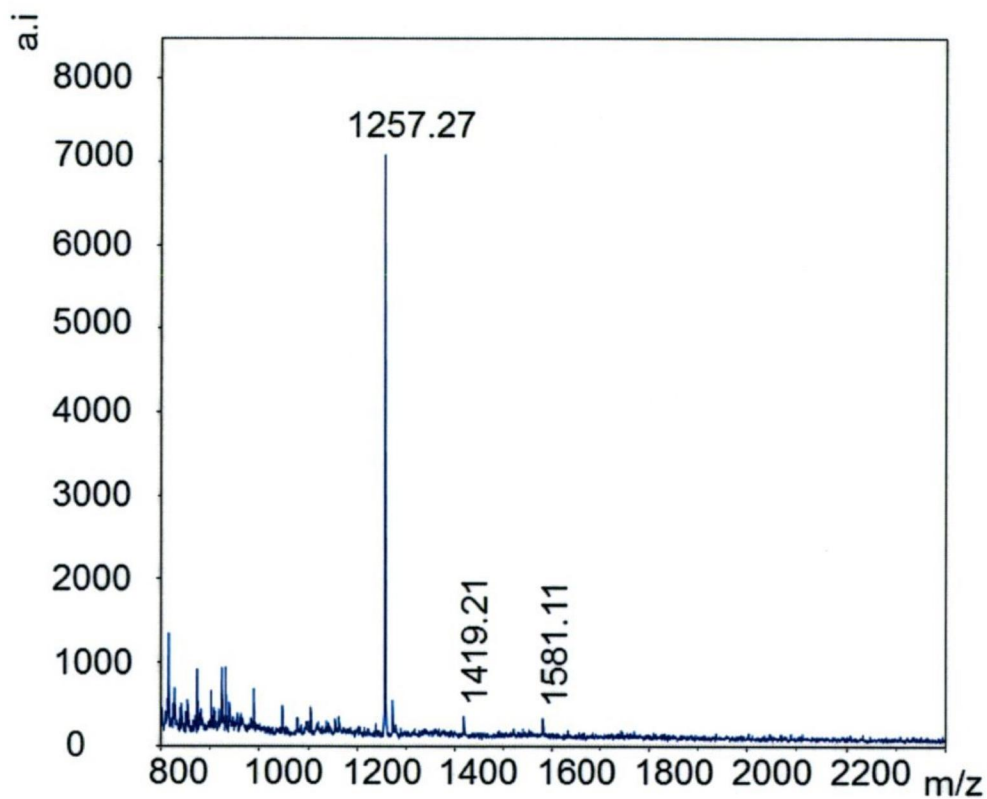


FIG. 4

A.



B.

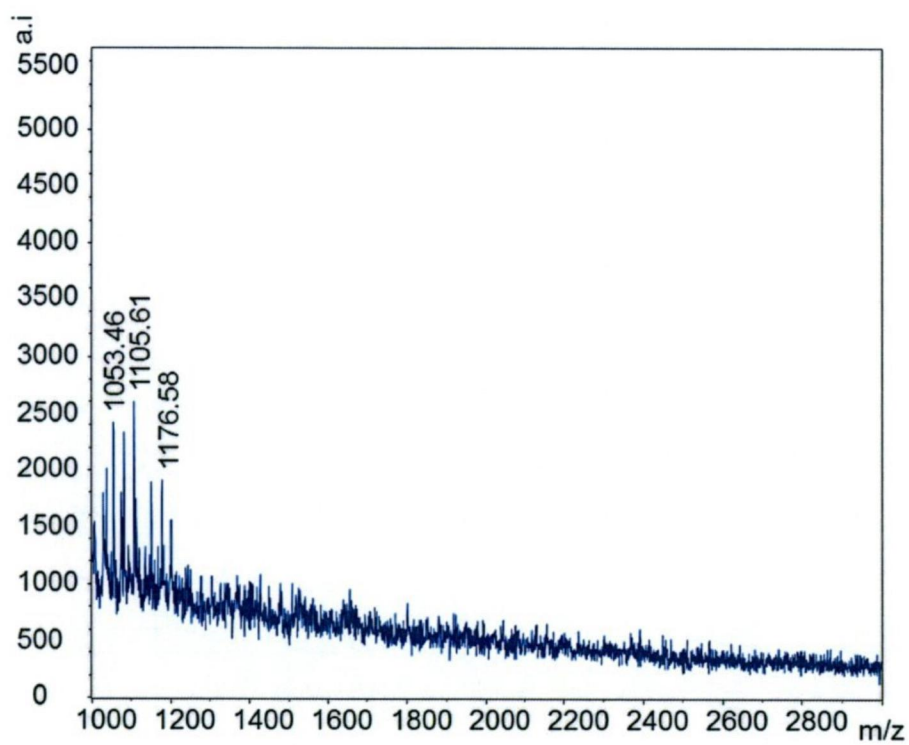
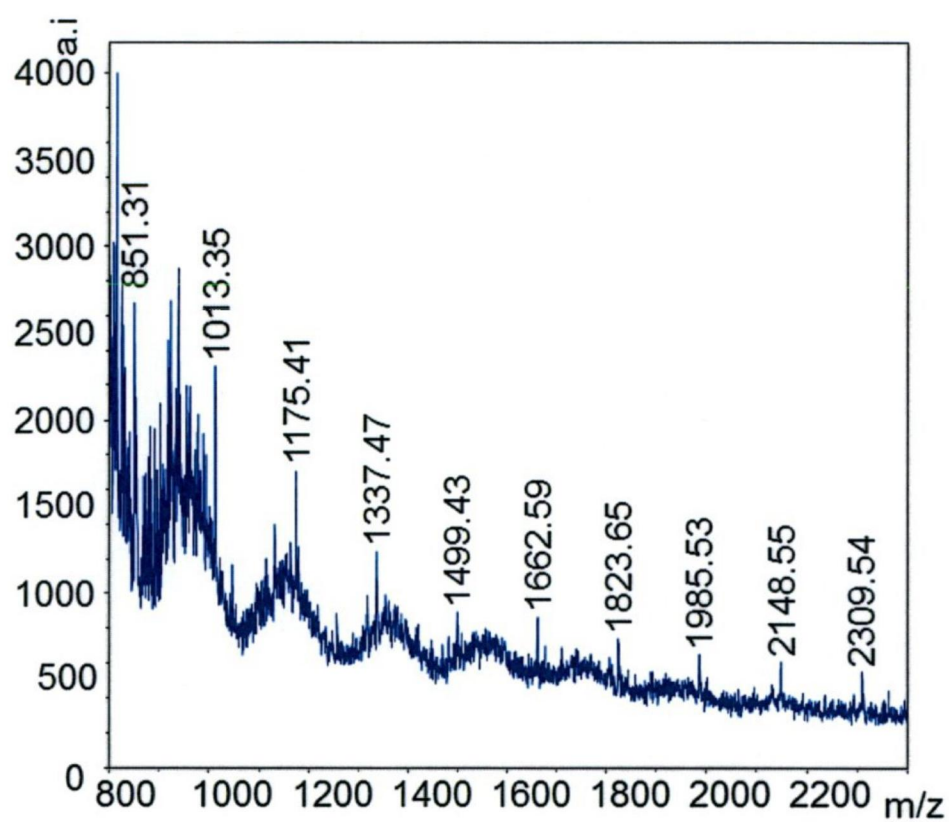


FIG. 5

A.



B.

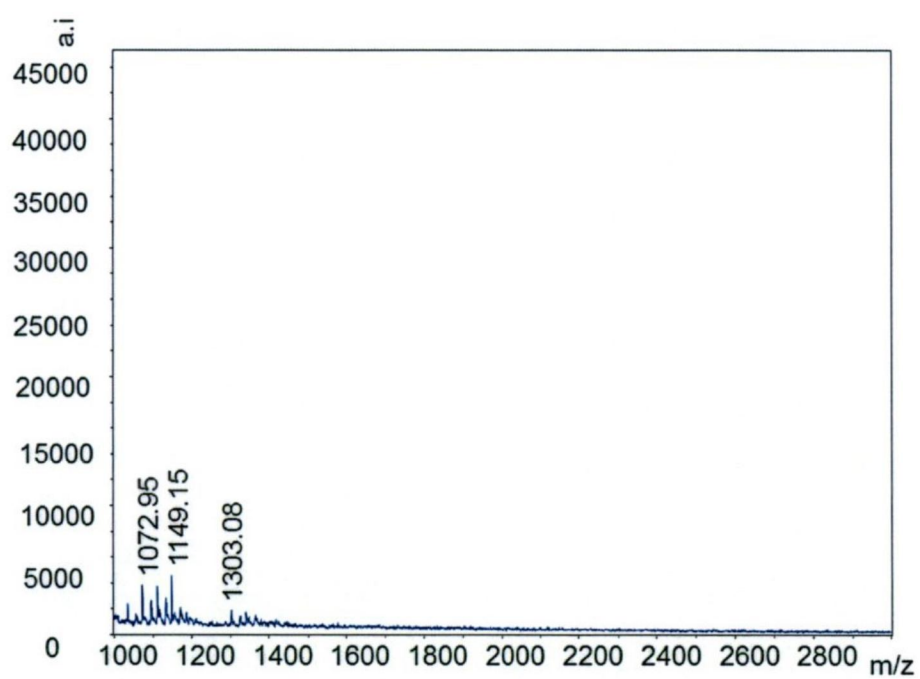


FIG. 6

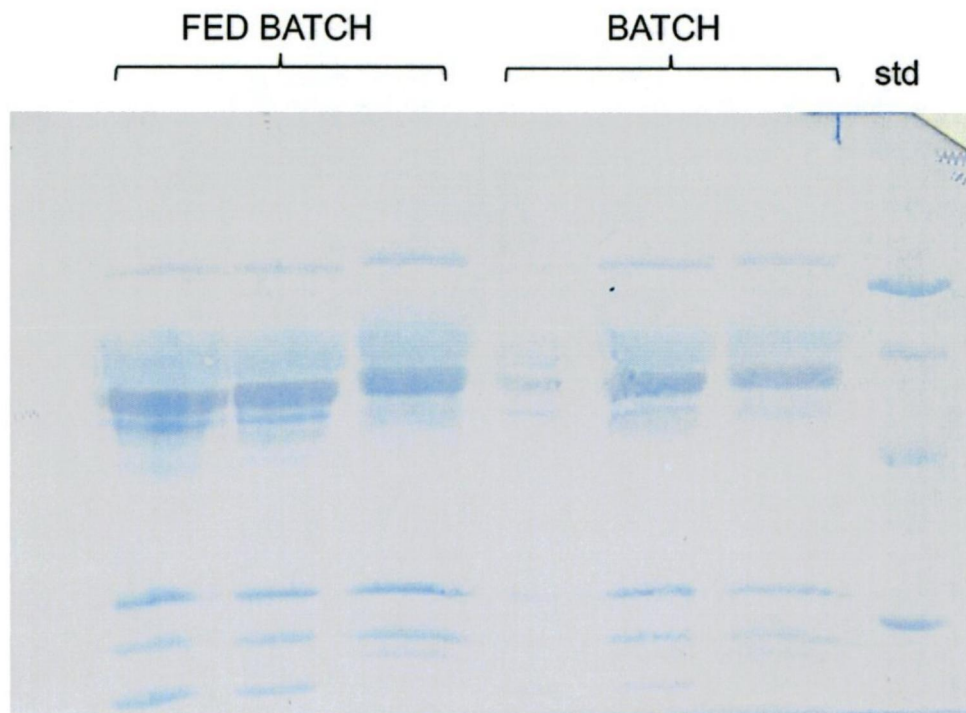


FIG. 7

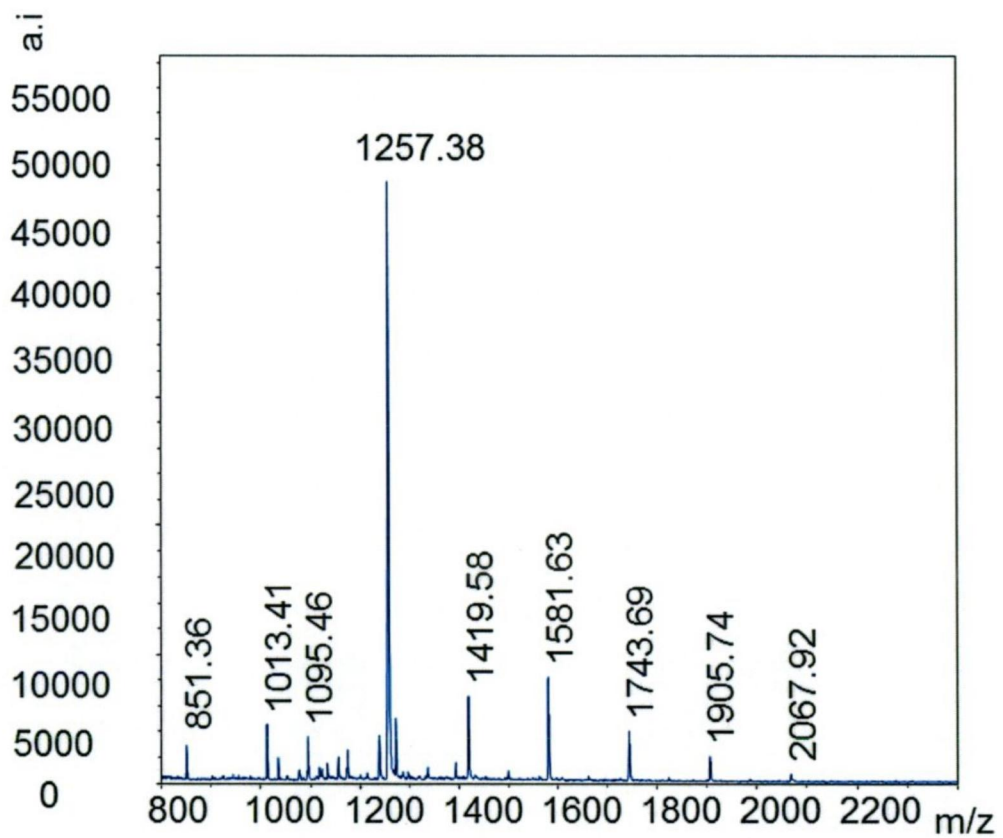


FIG. 8

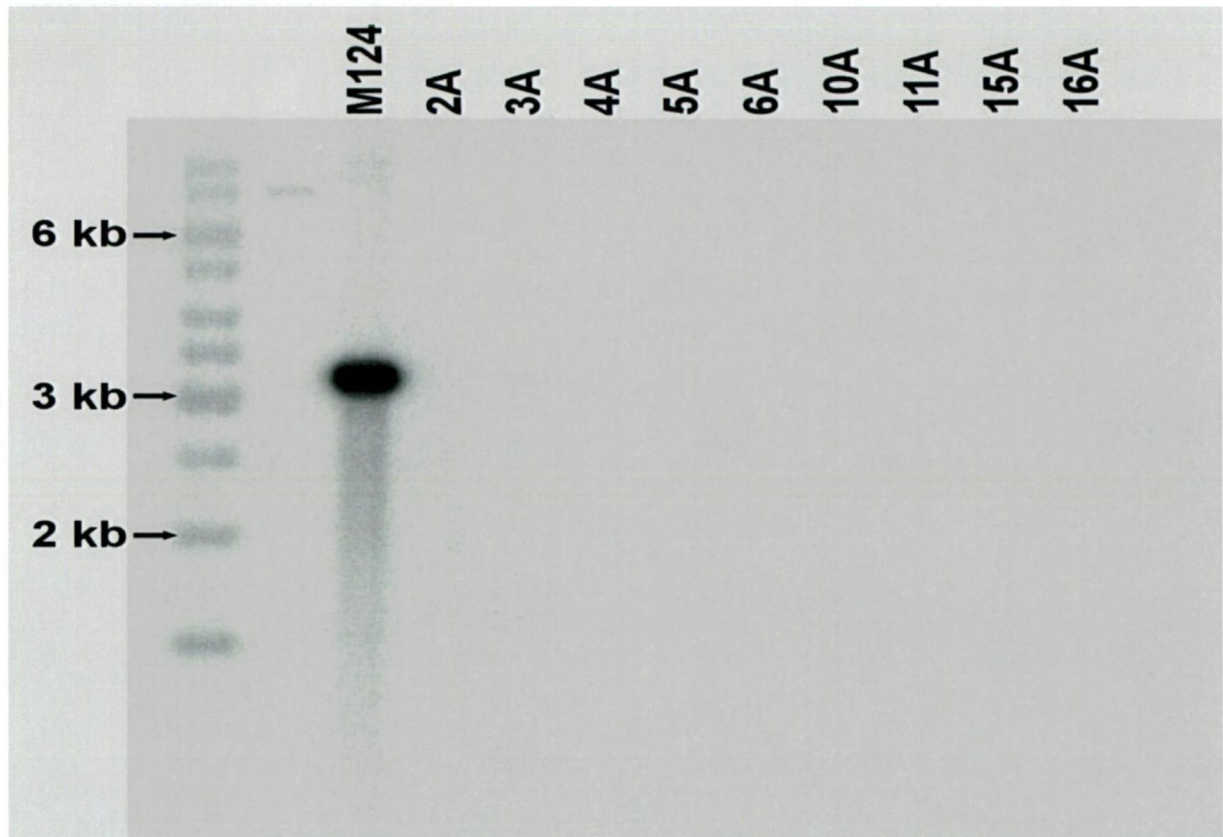
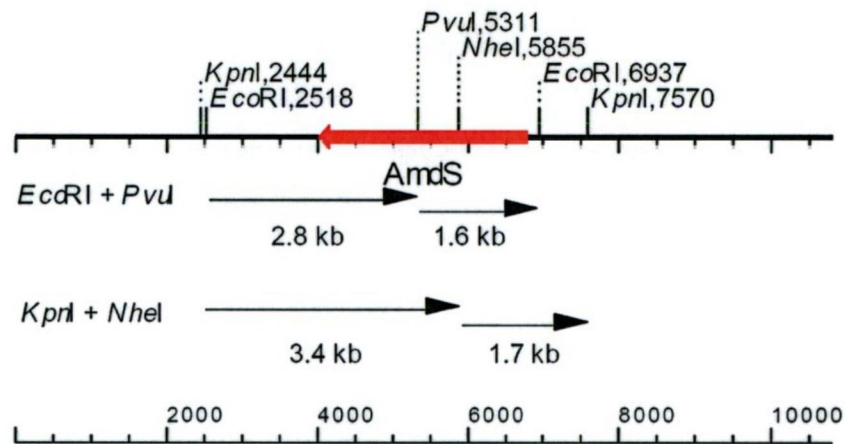


FIG. 9

A.



B.

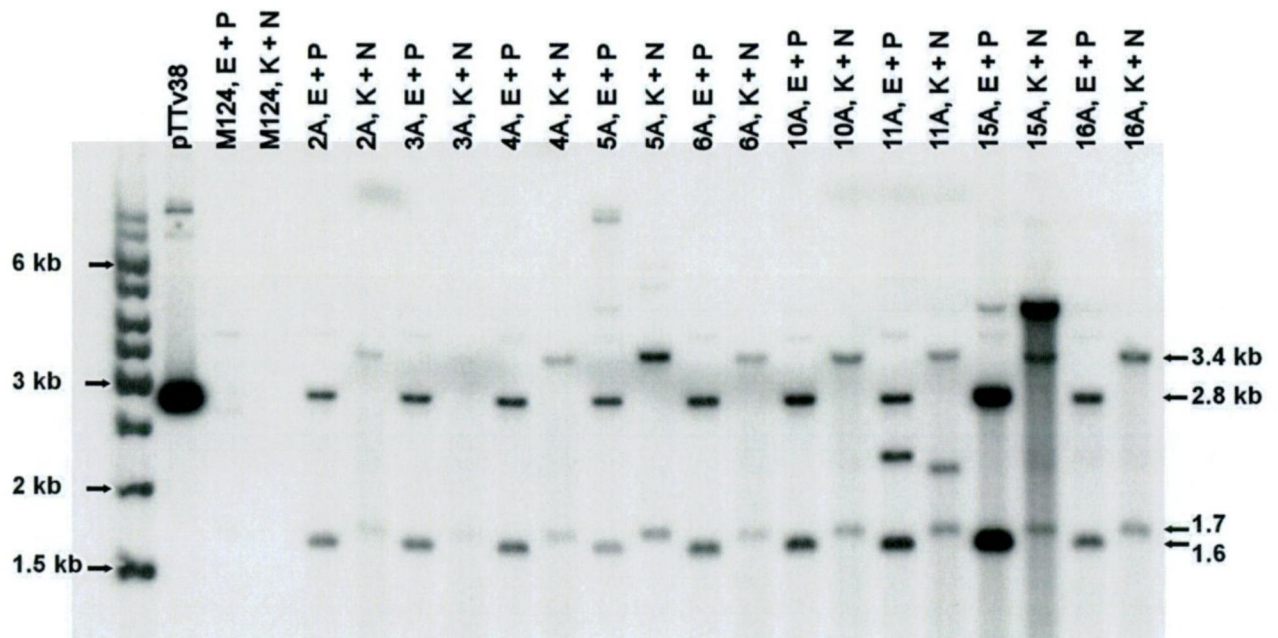


FIG. 10

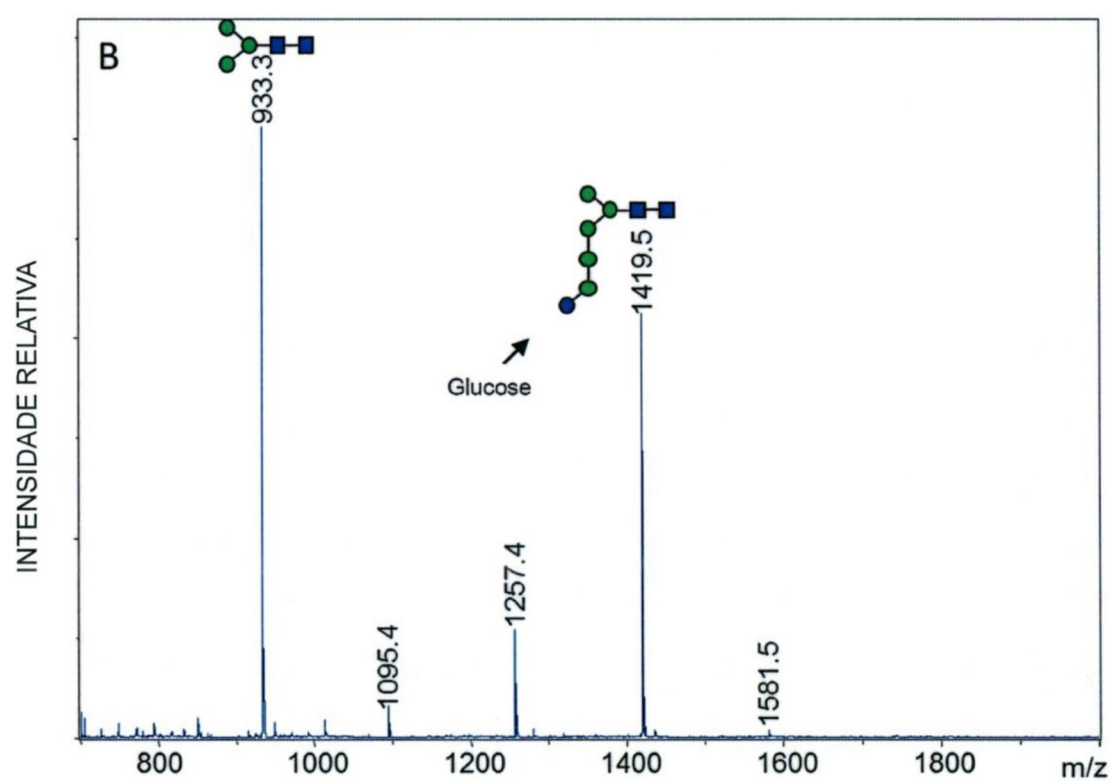
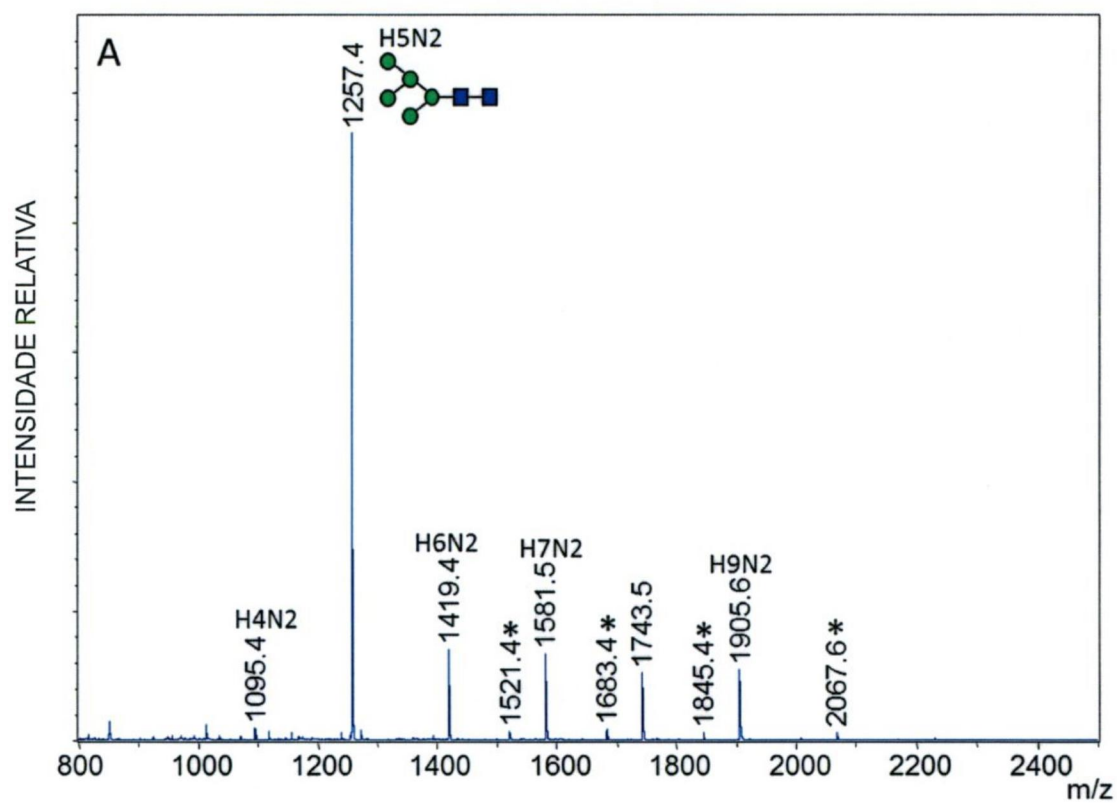


FIG. 11

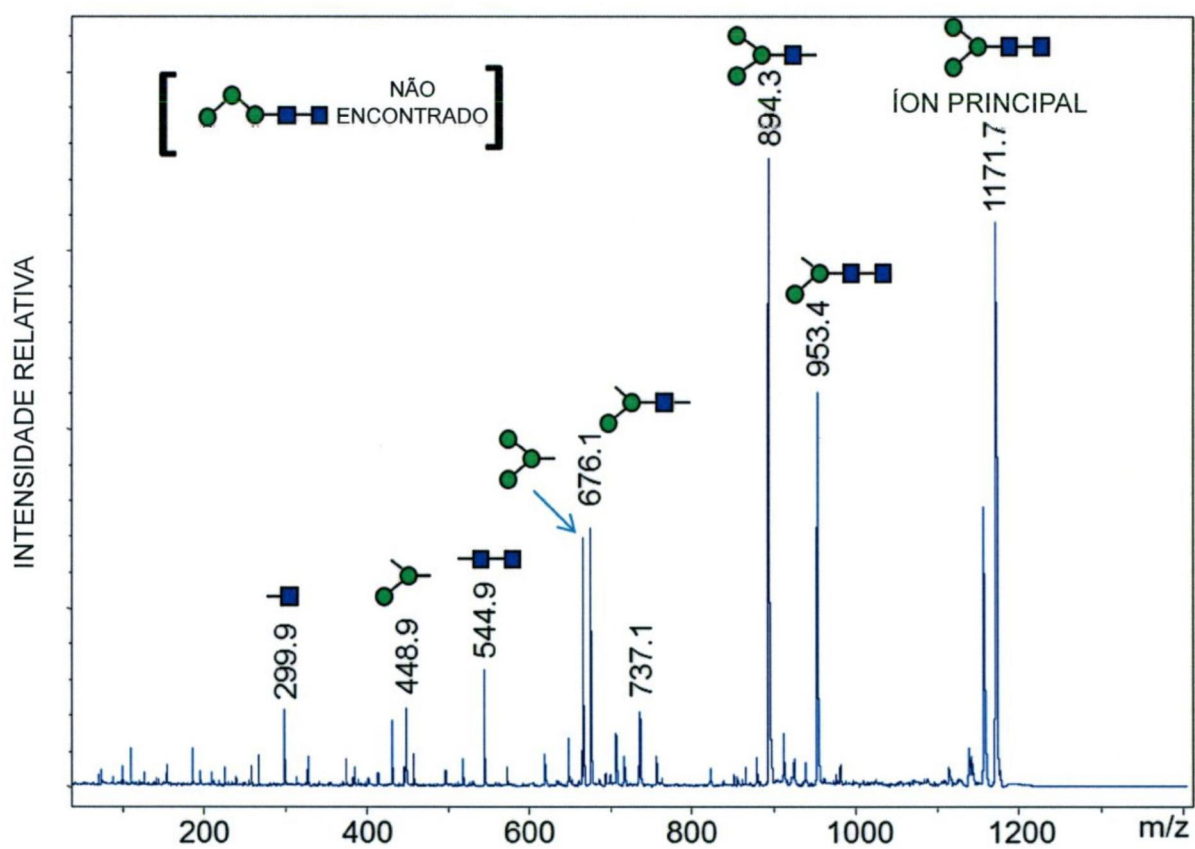


FIG. 12

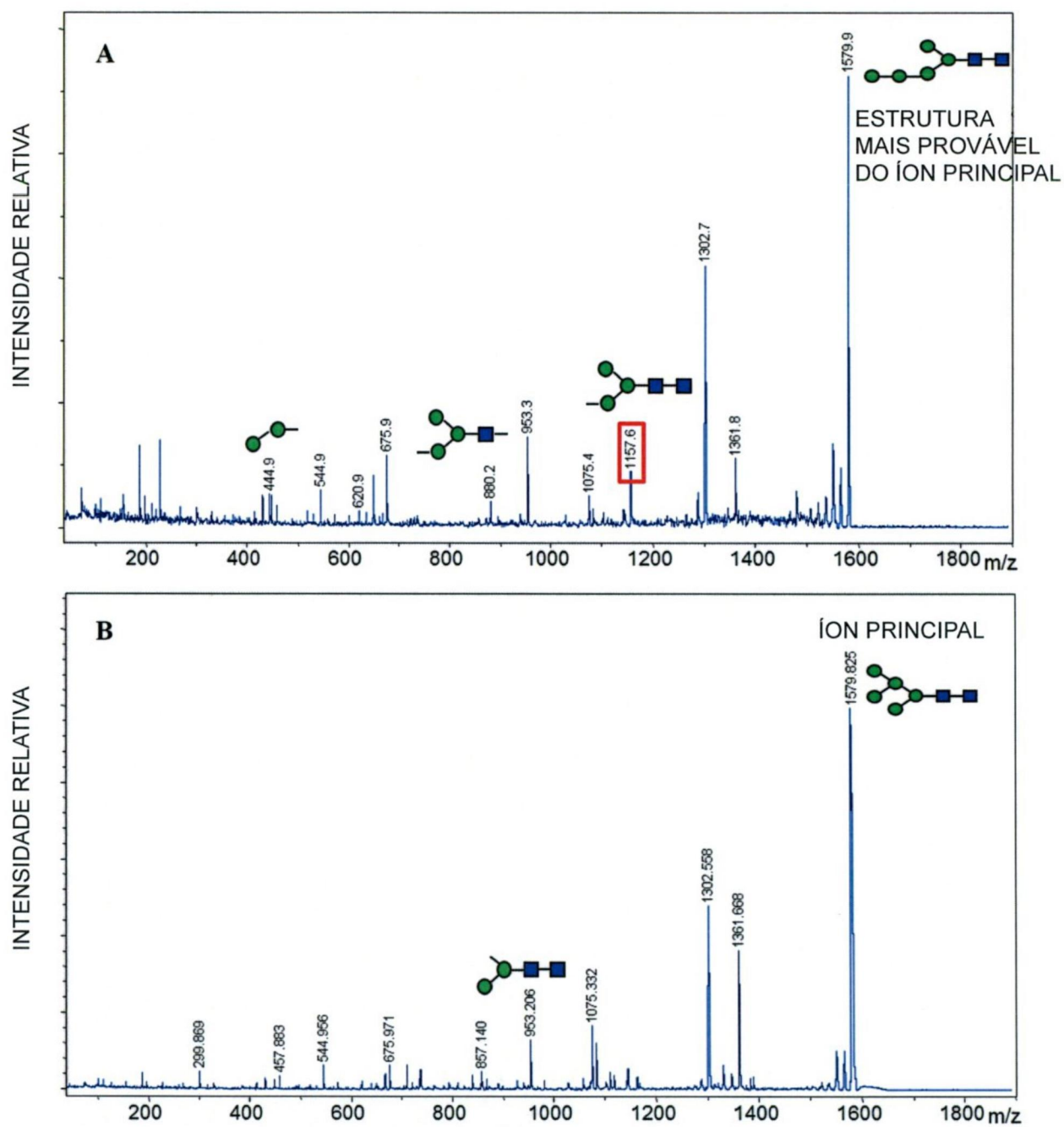


FIG. 13

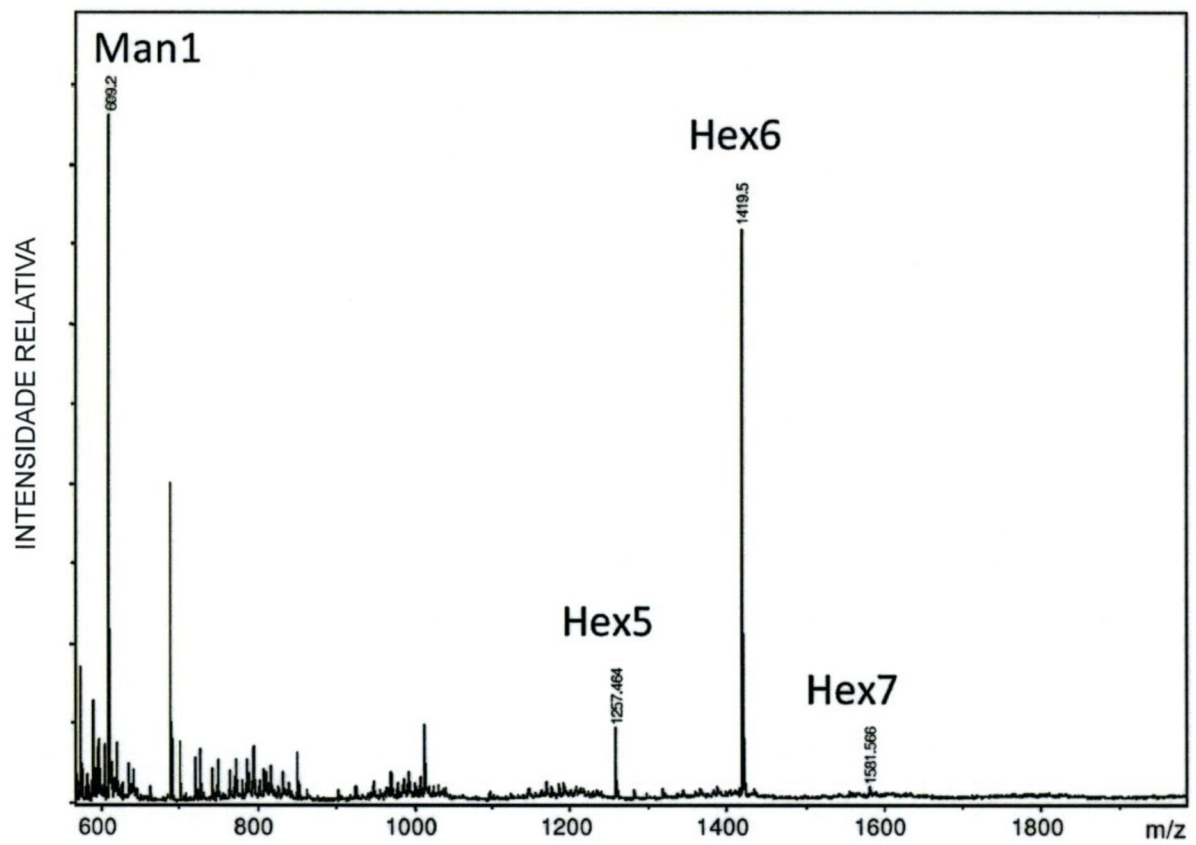


FIG. 14

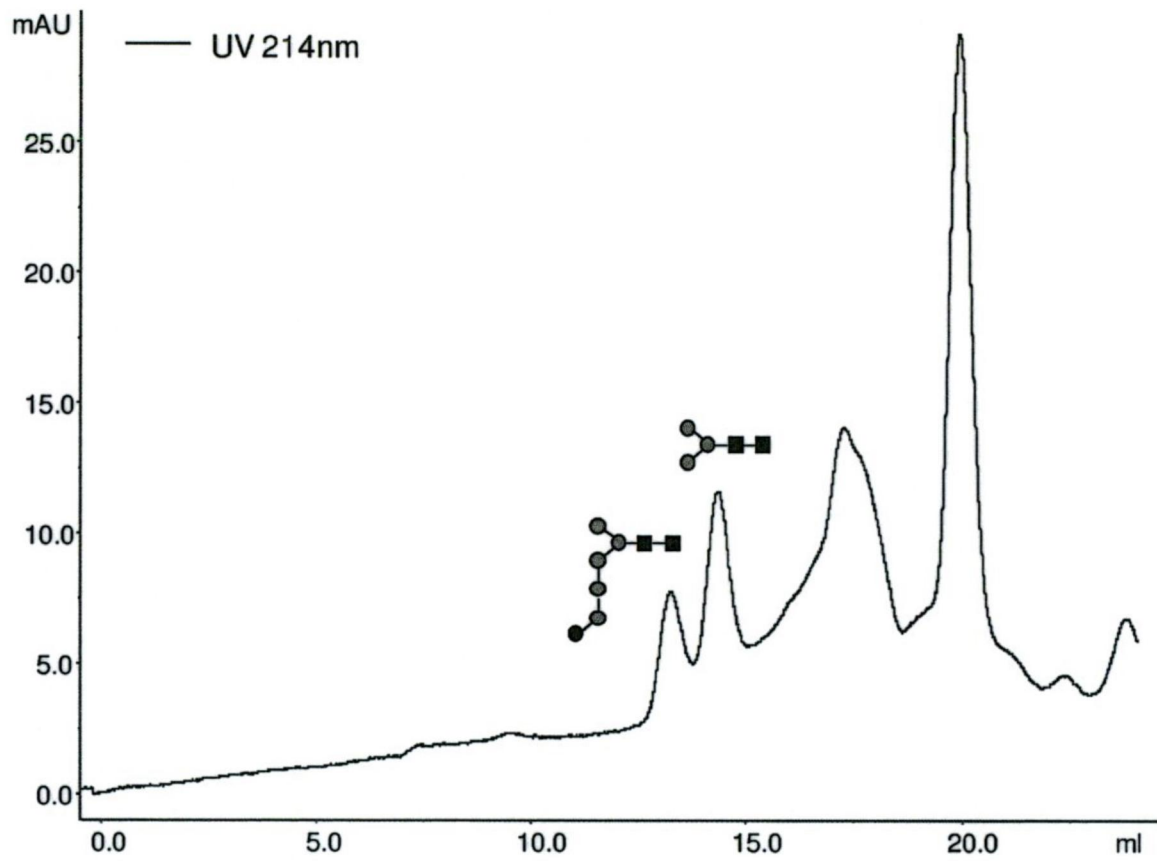


FIG. 15

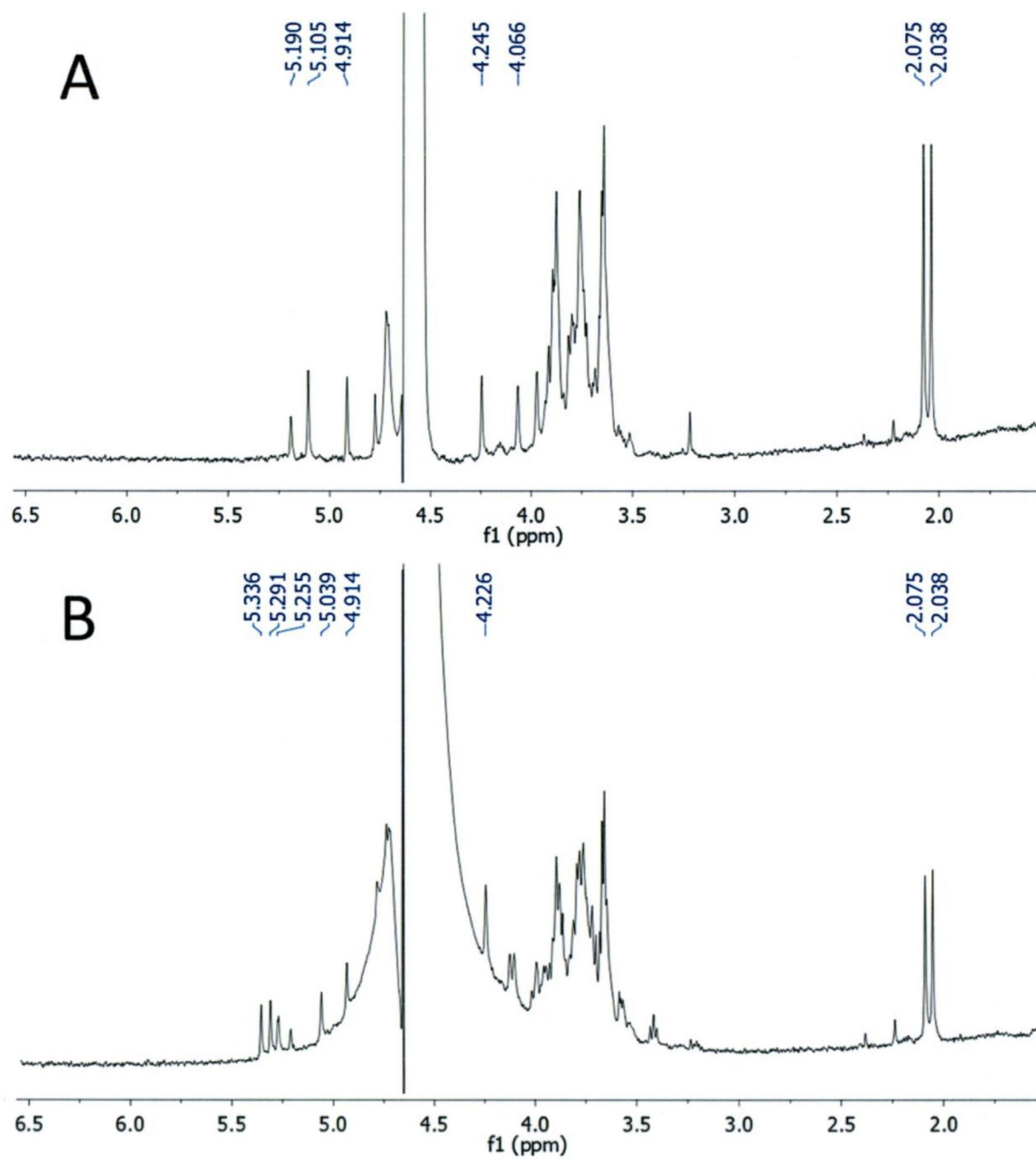


FIG. 16

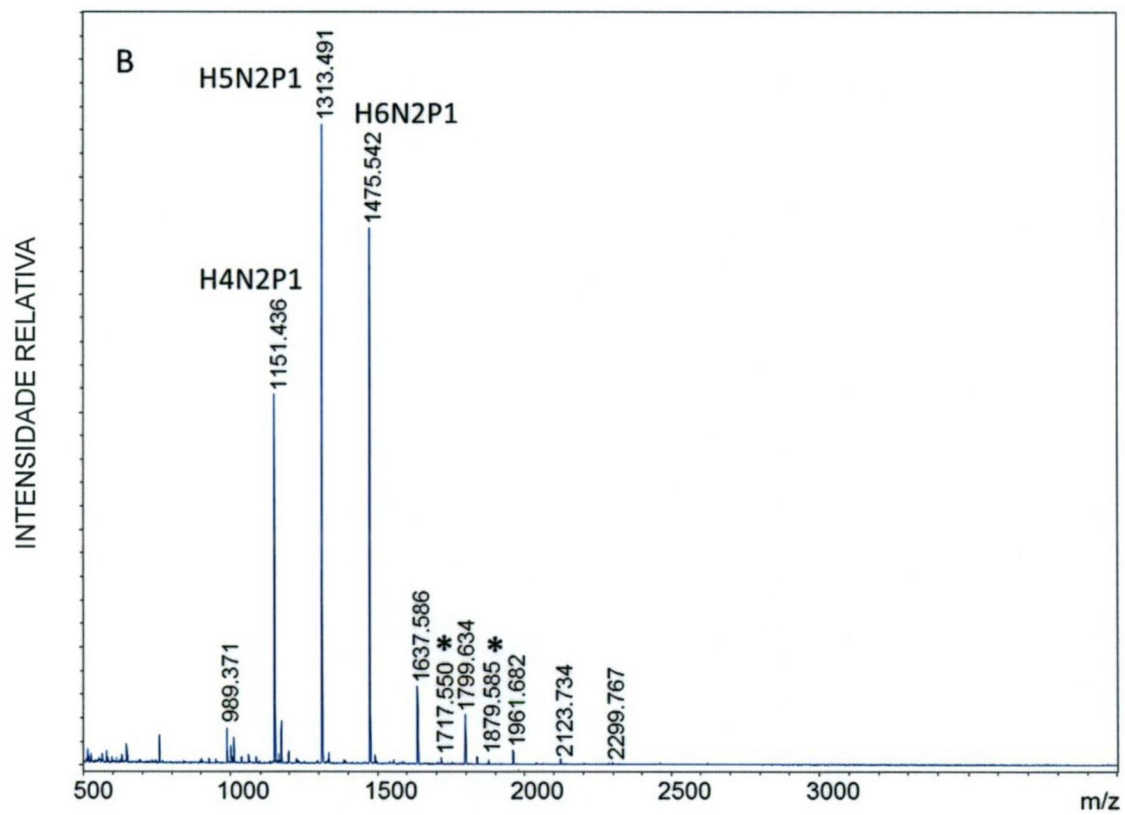
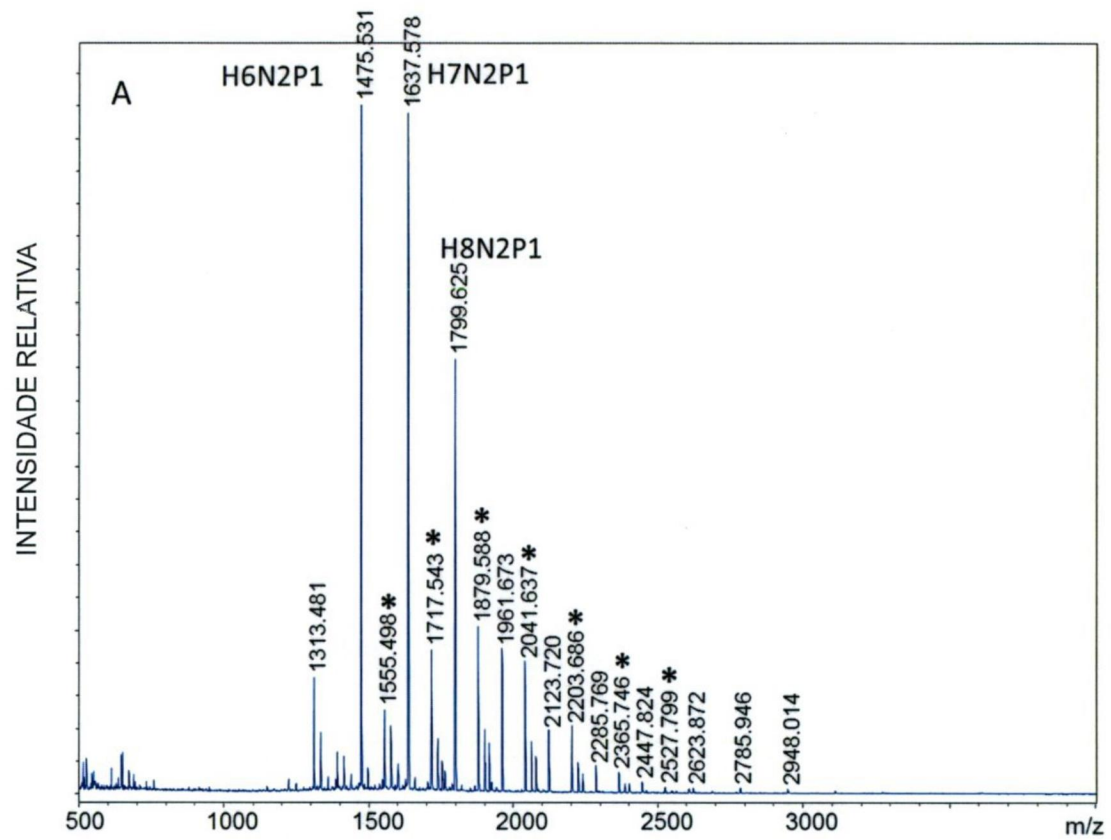


FIG. 17

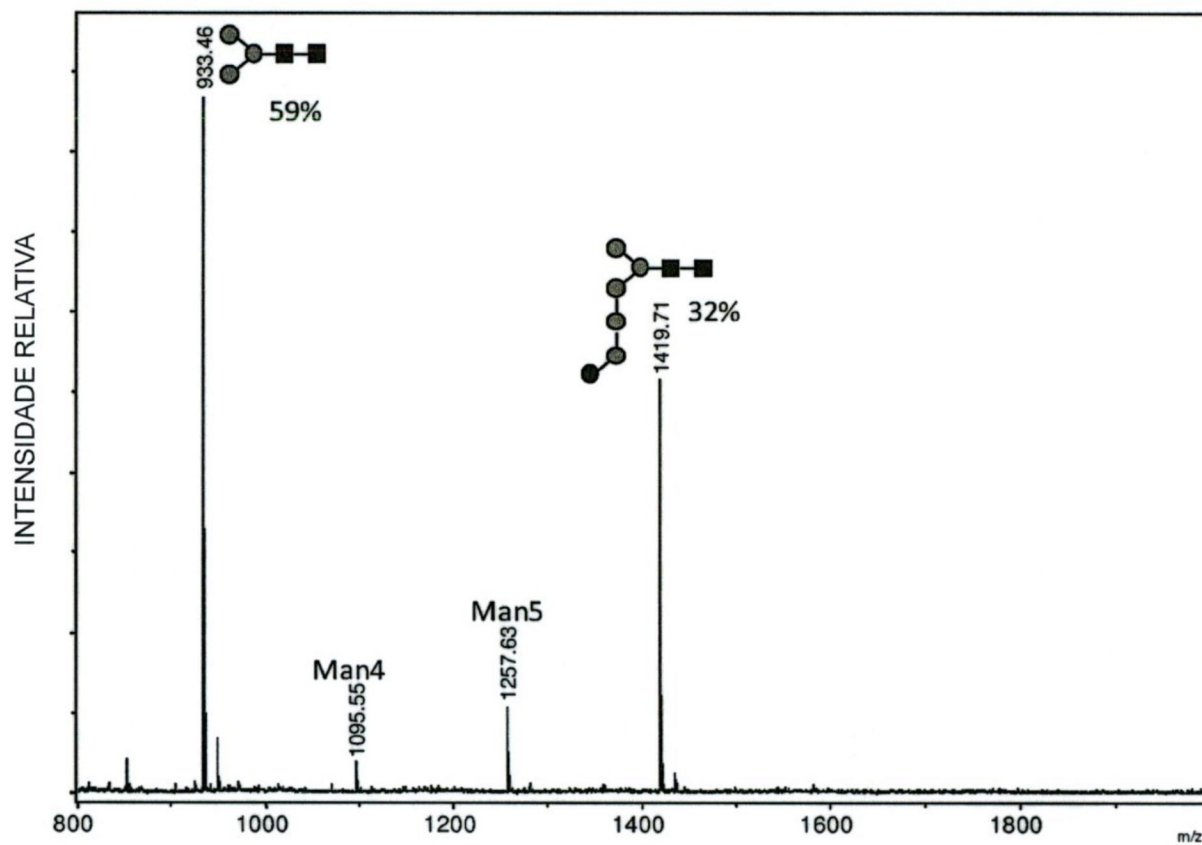


FIG. 18

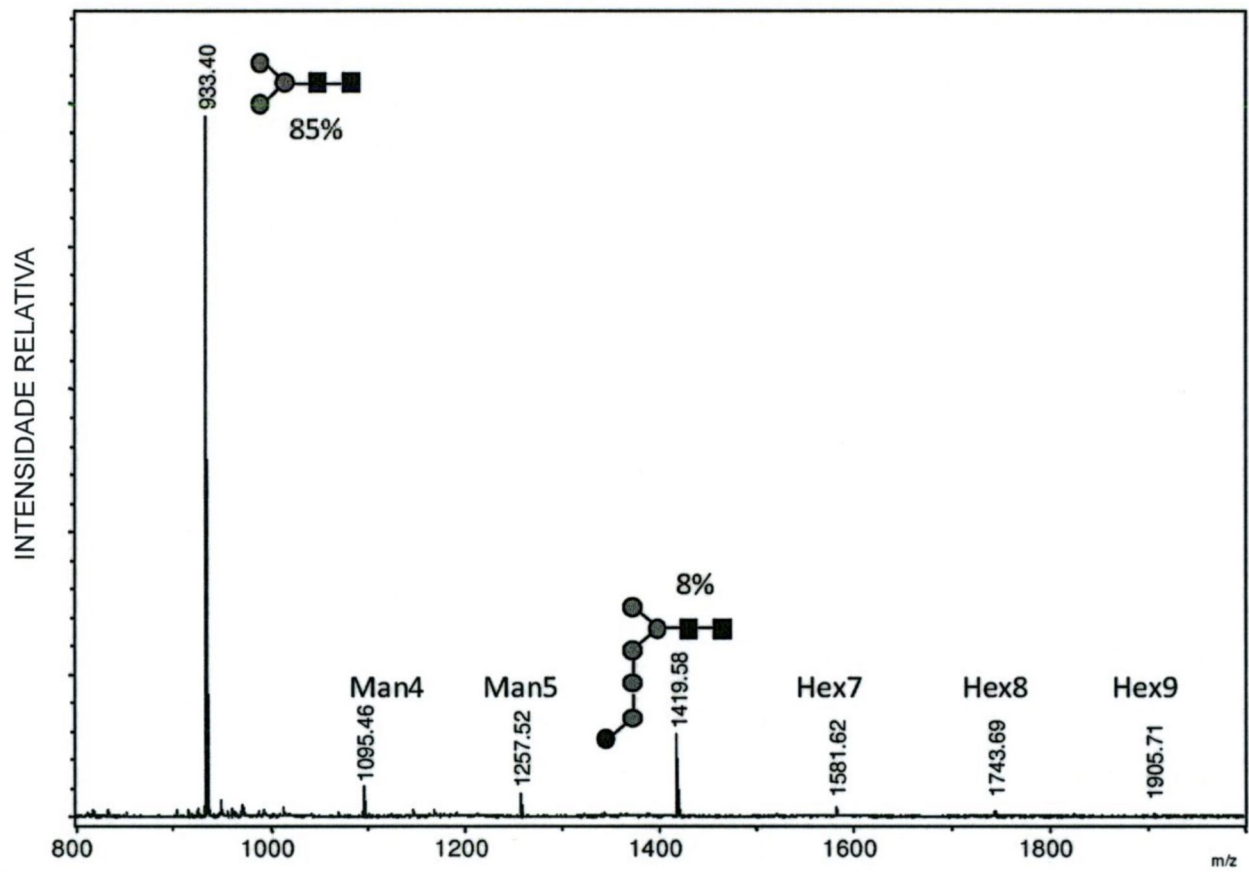


FIG. 19

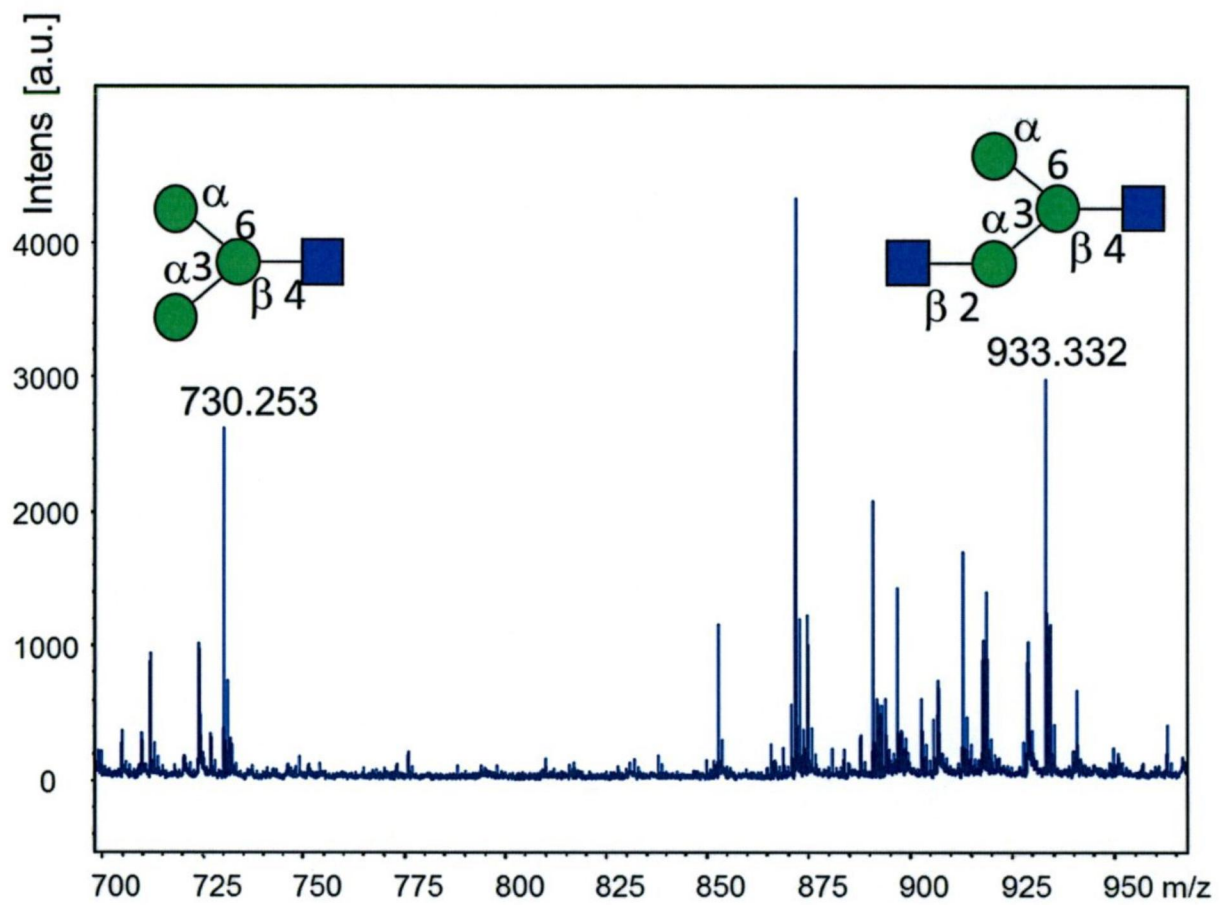


FIG. 20



FIG. 21

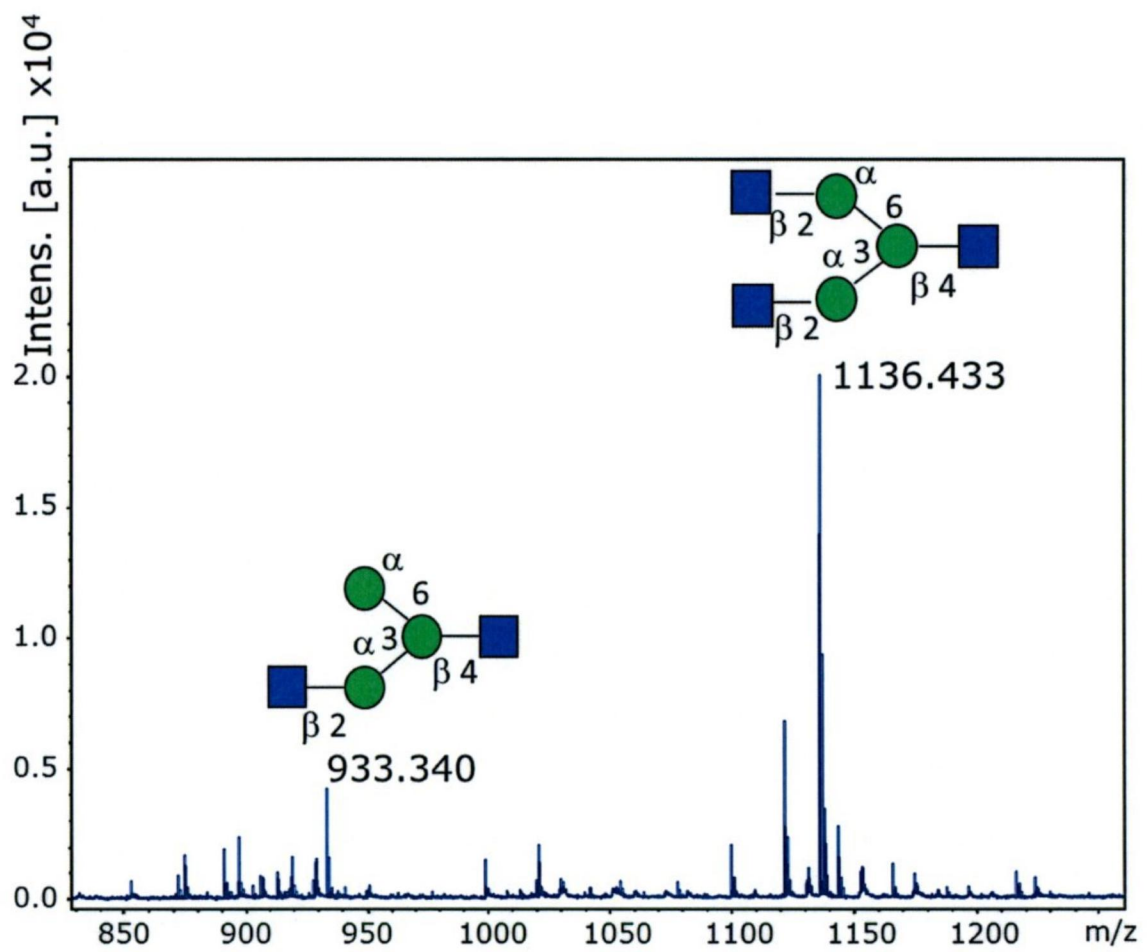


FIG. 22

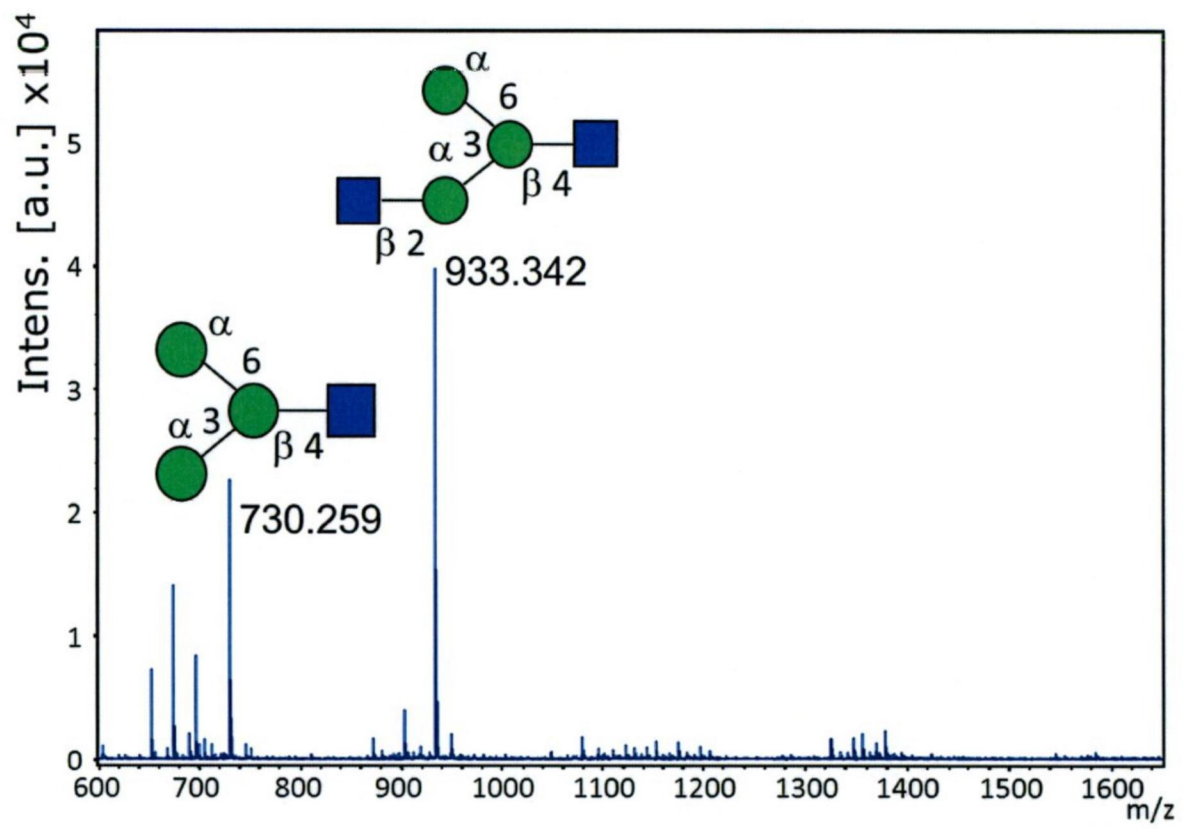




FIG. 24

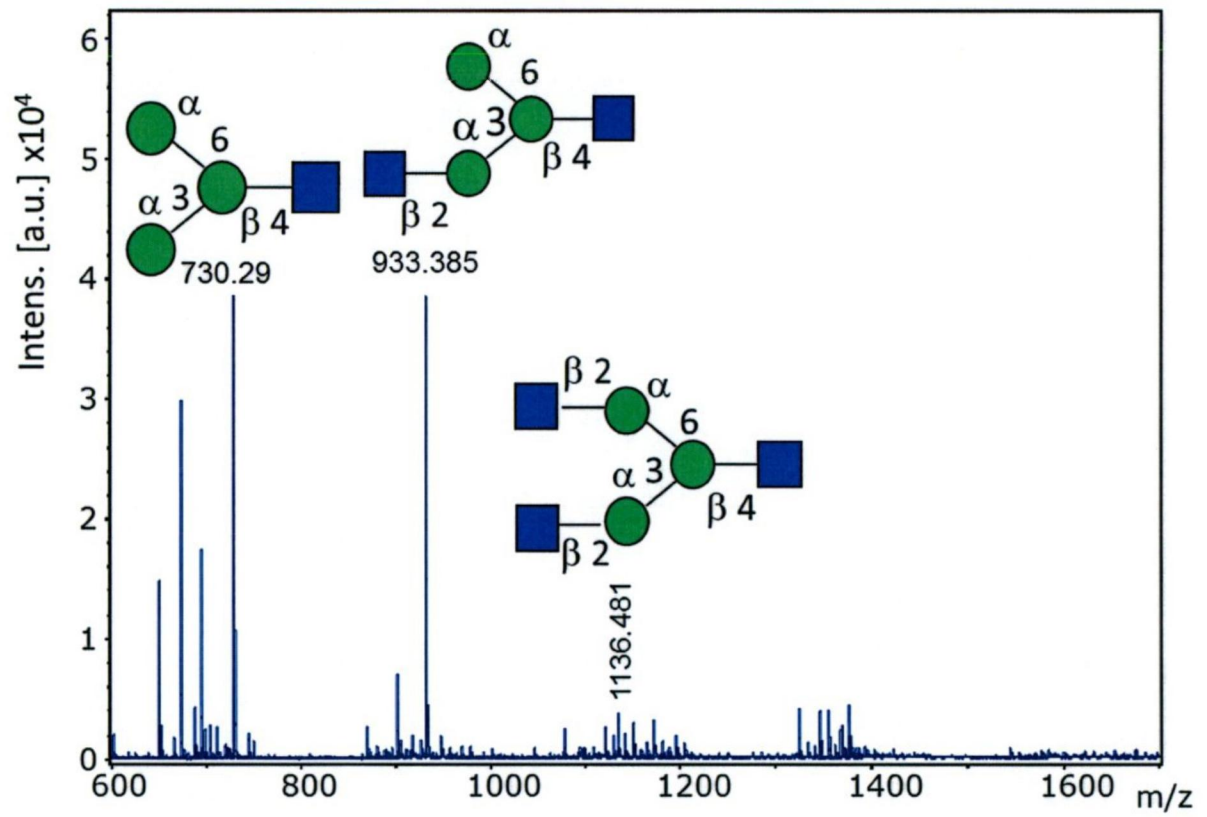


FIG. 25

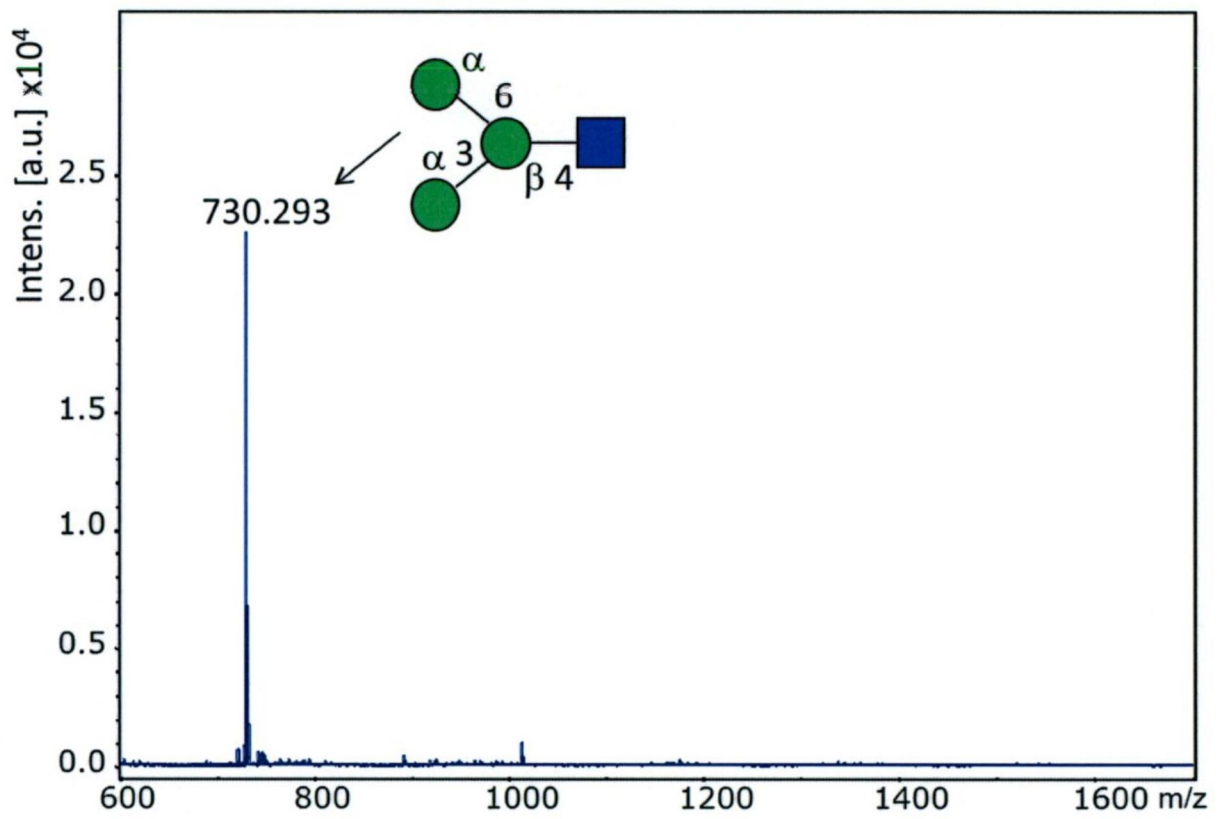


FIG. 26

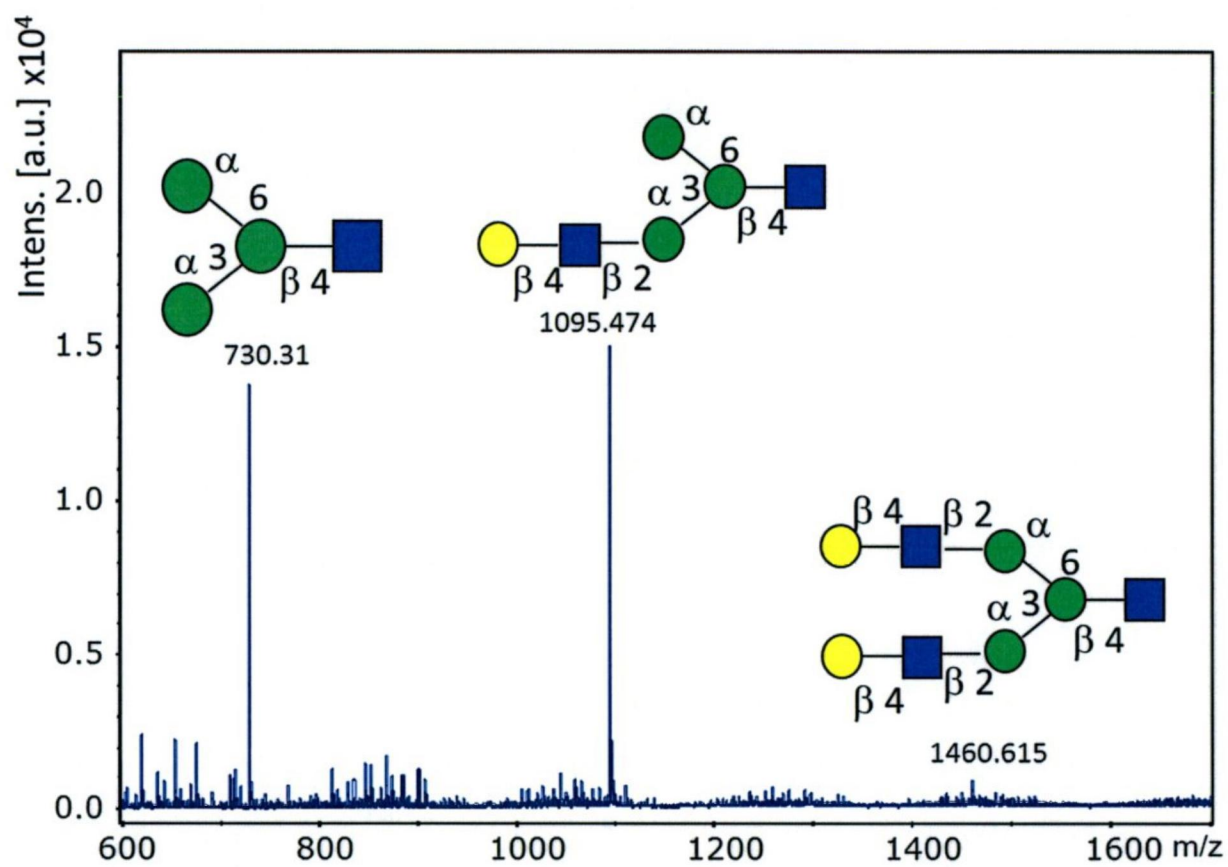
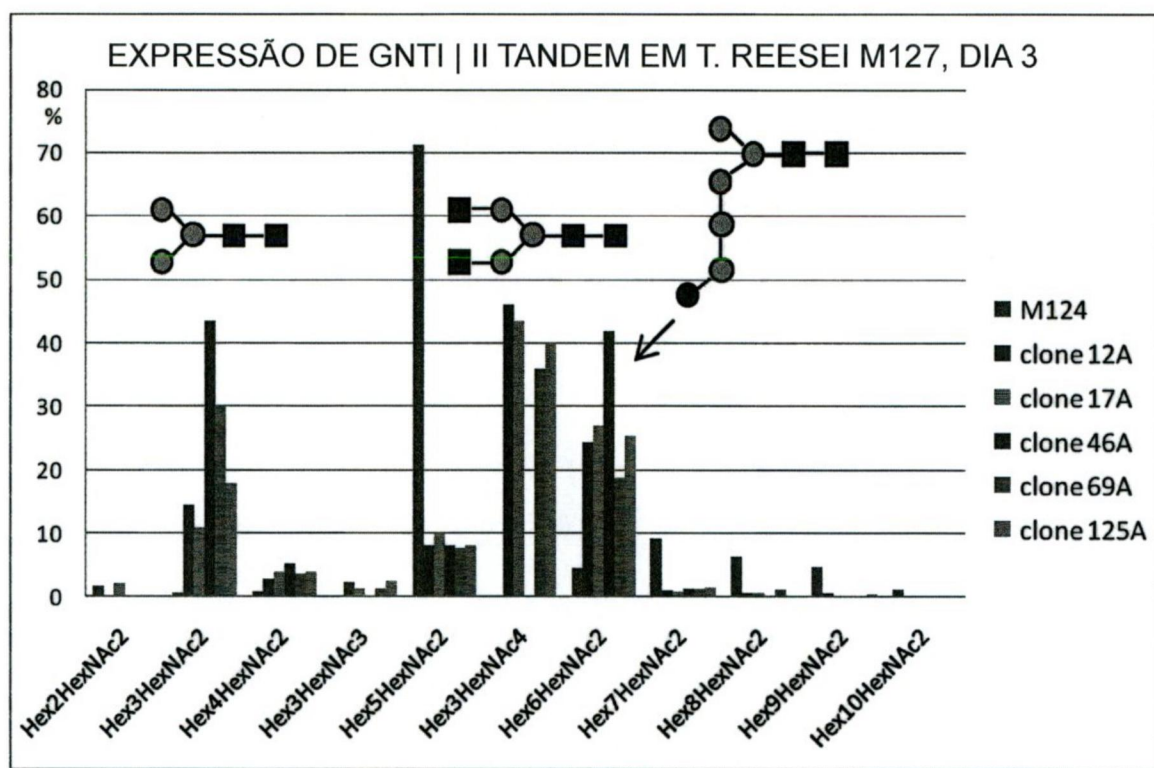
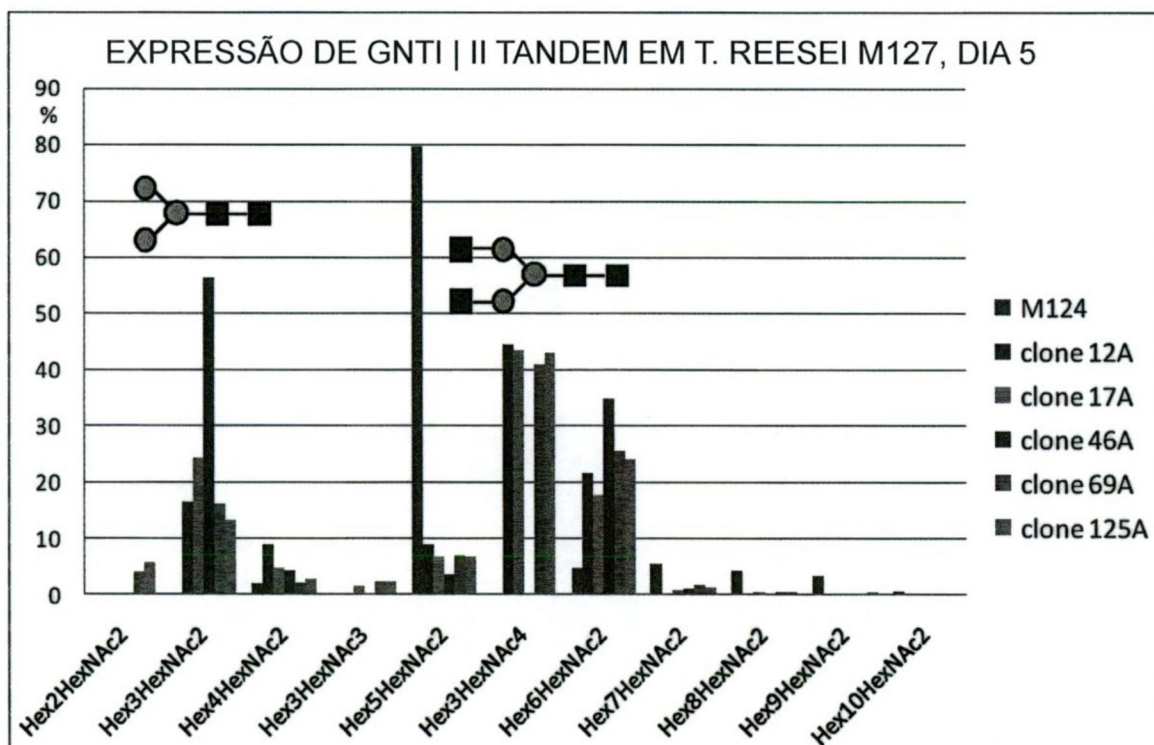


FIG. 27

A.

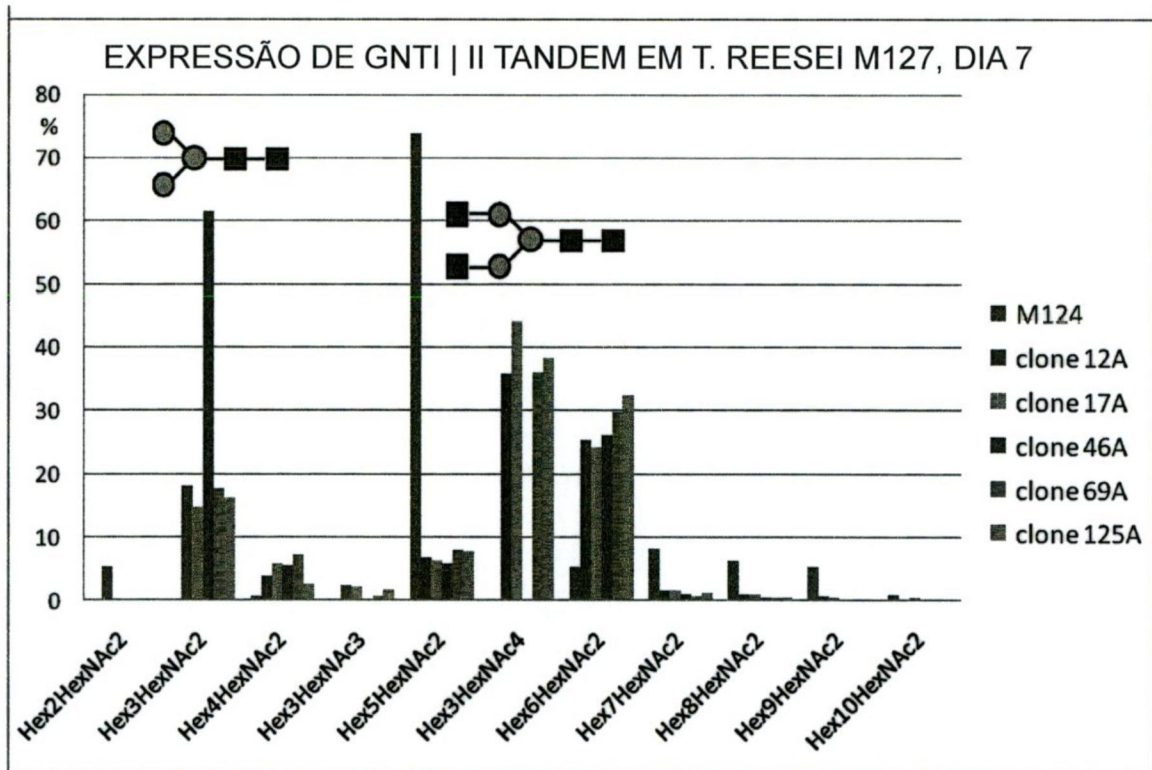


B.



C.

FIG. 27 (Continuação)



D.

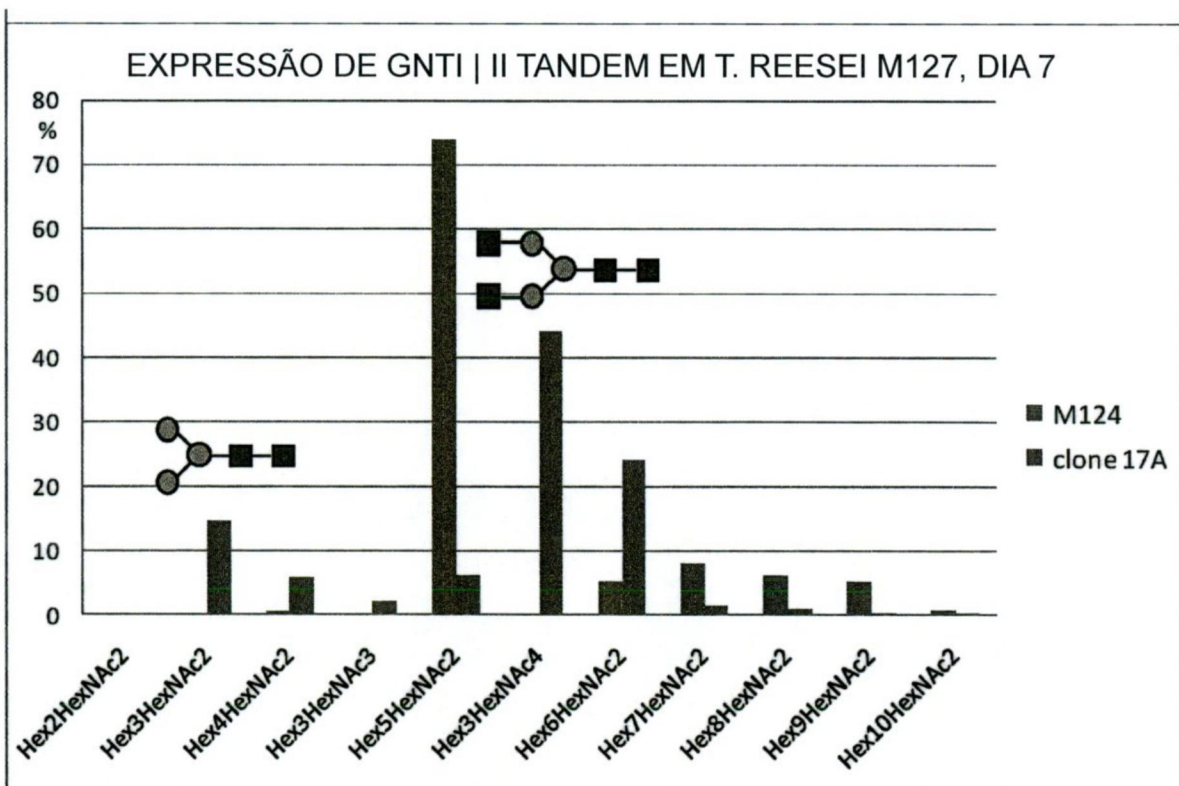


FIG. 27 (Continuação)

E.

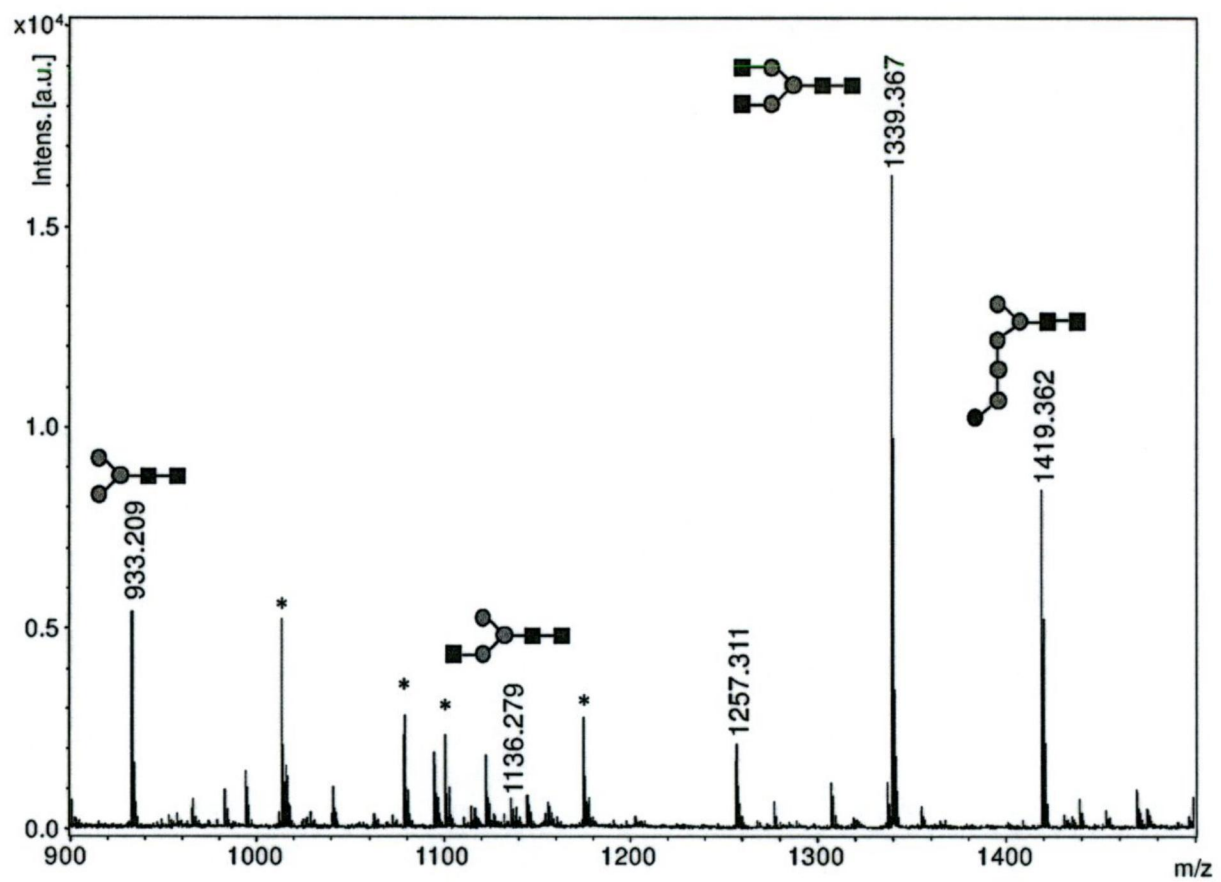
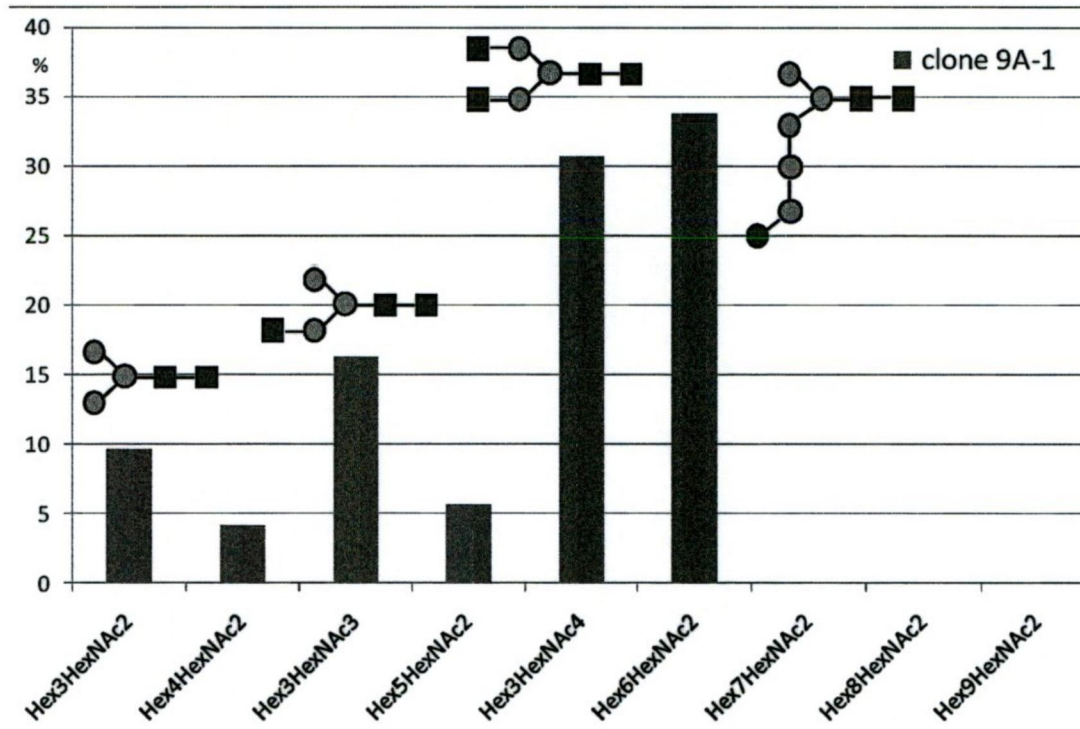


FIG. 28

A.



B.

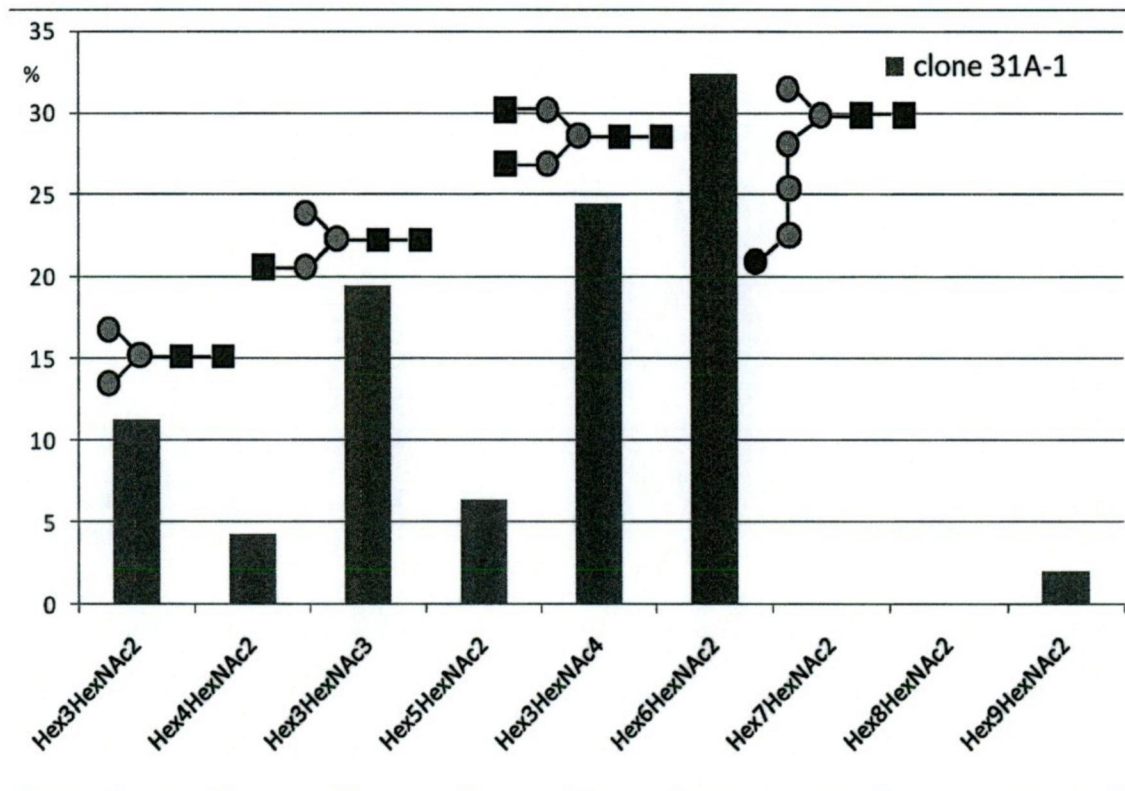


FIG. 28 (Continuação)

C.

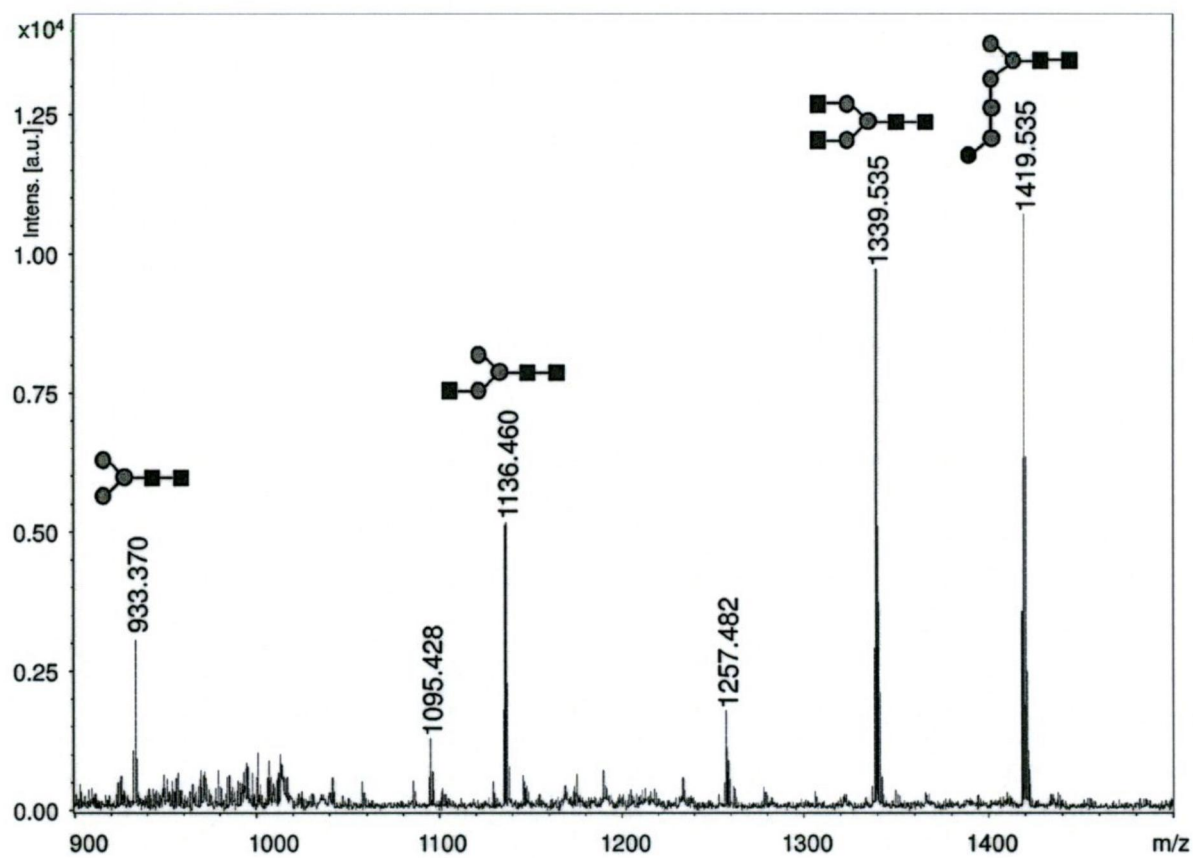


FIG. 29

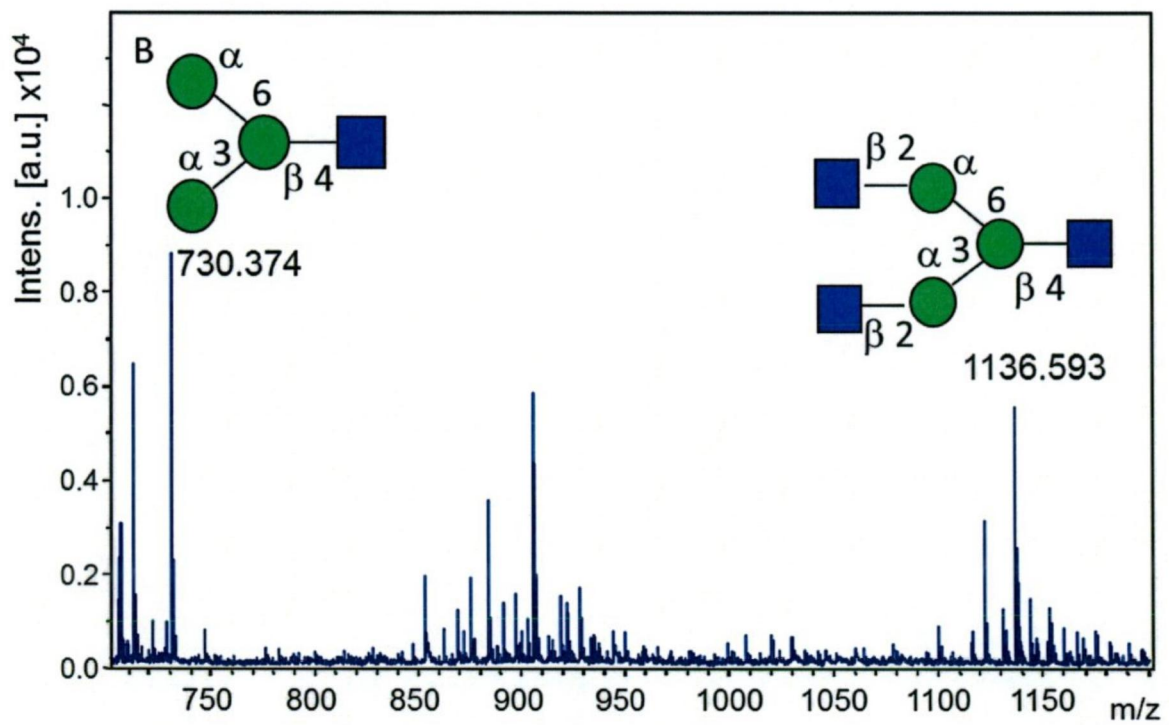
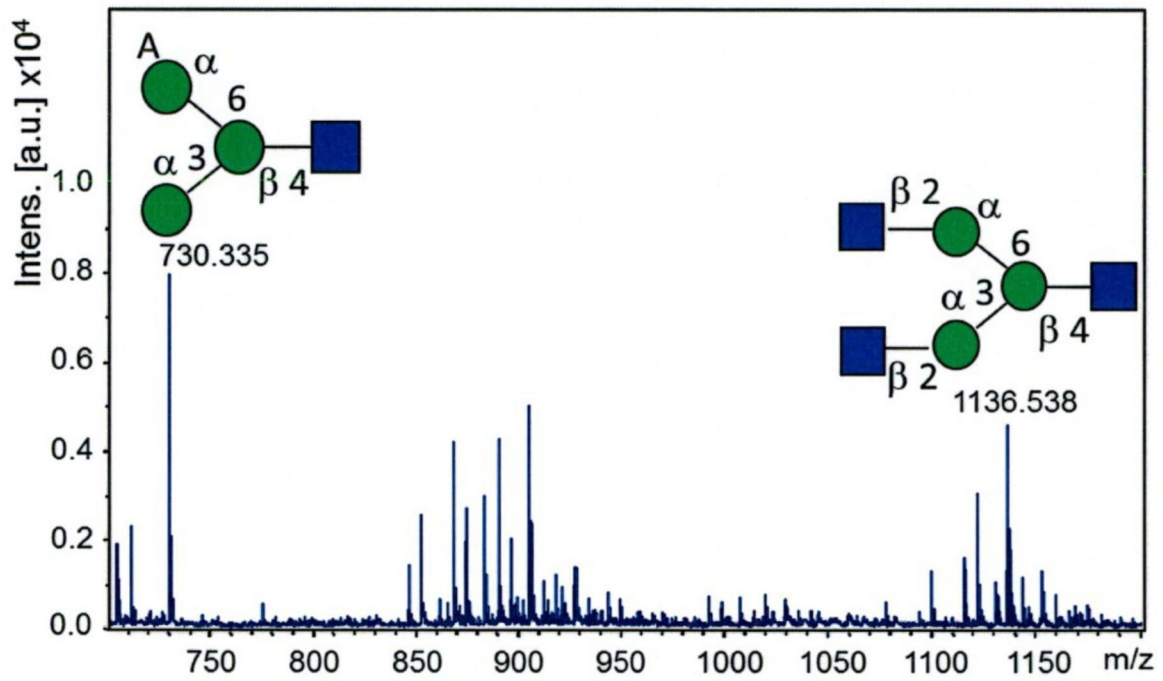


FIG. 30

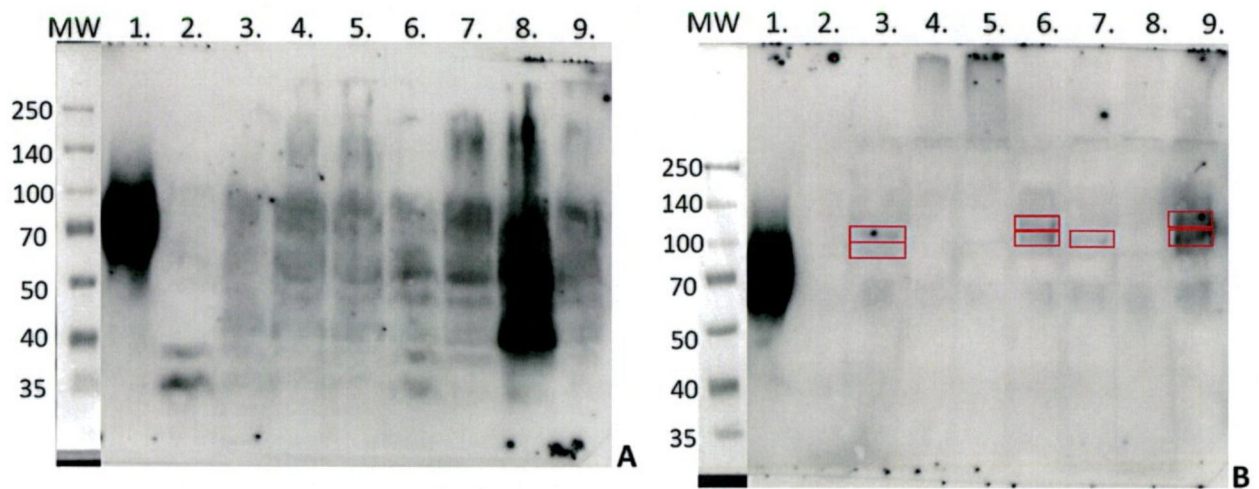
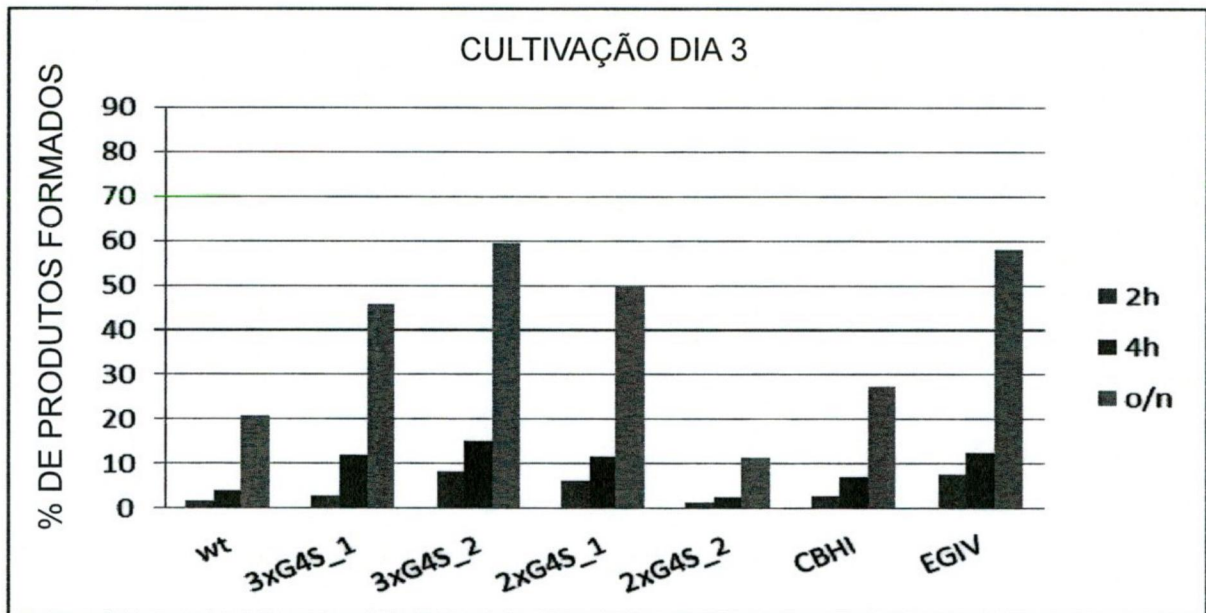


FIG. 31

A.



B.

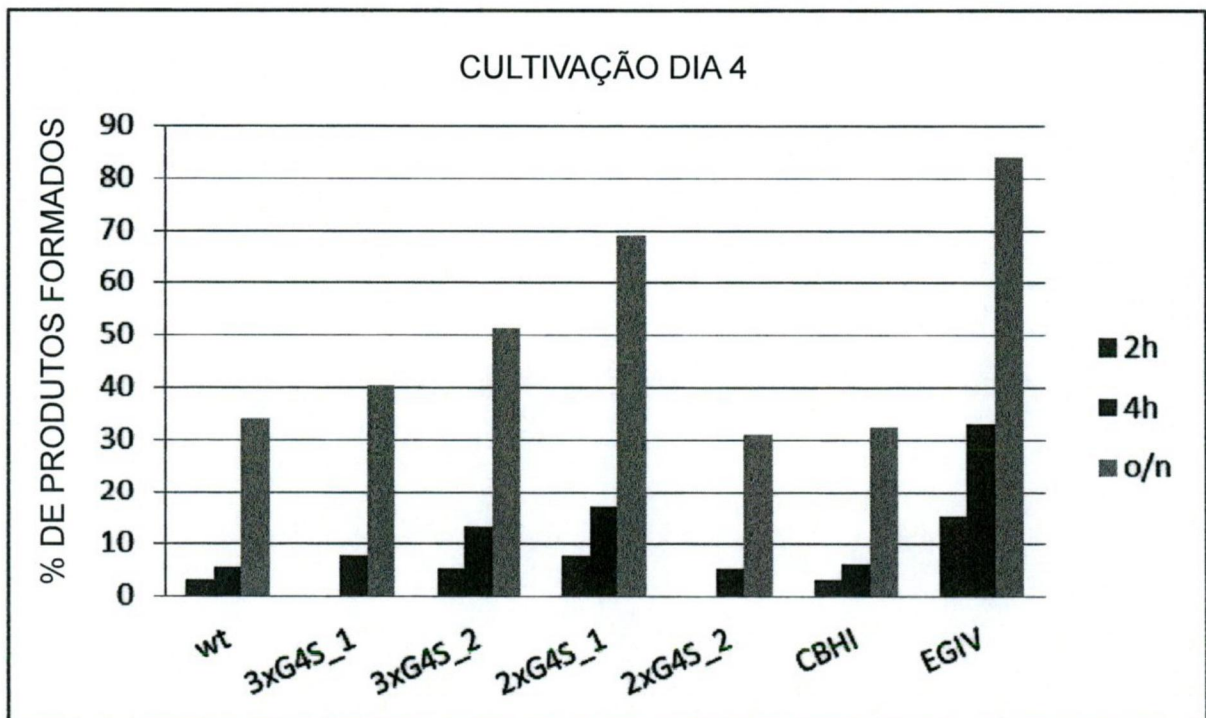


FIG. 32

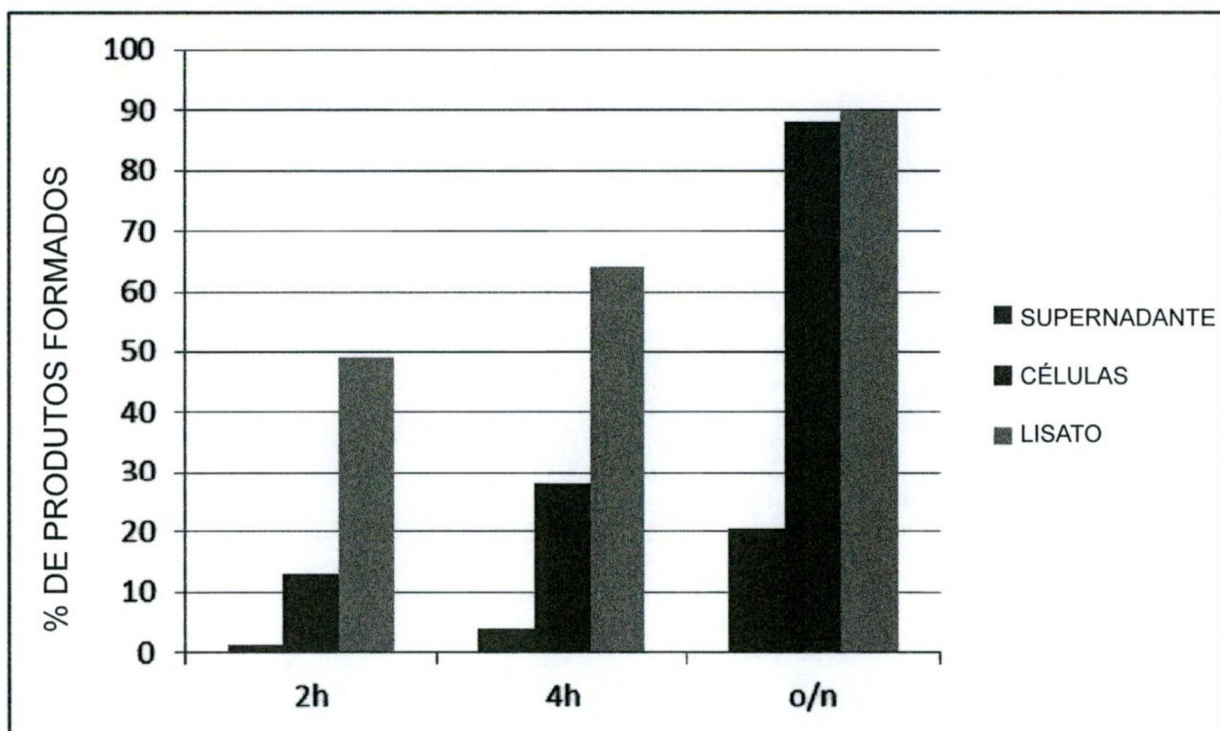


FIG. 33

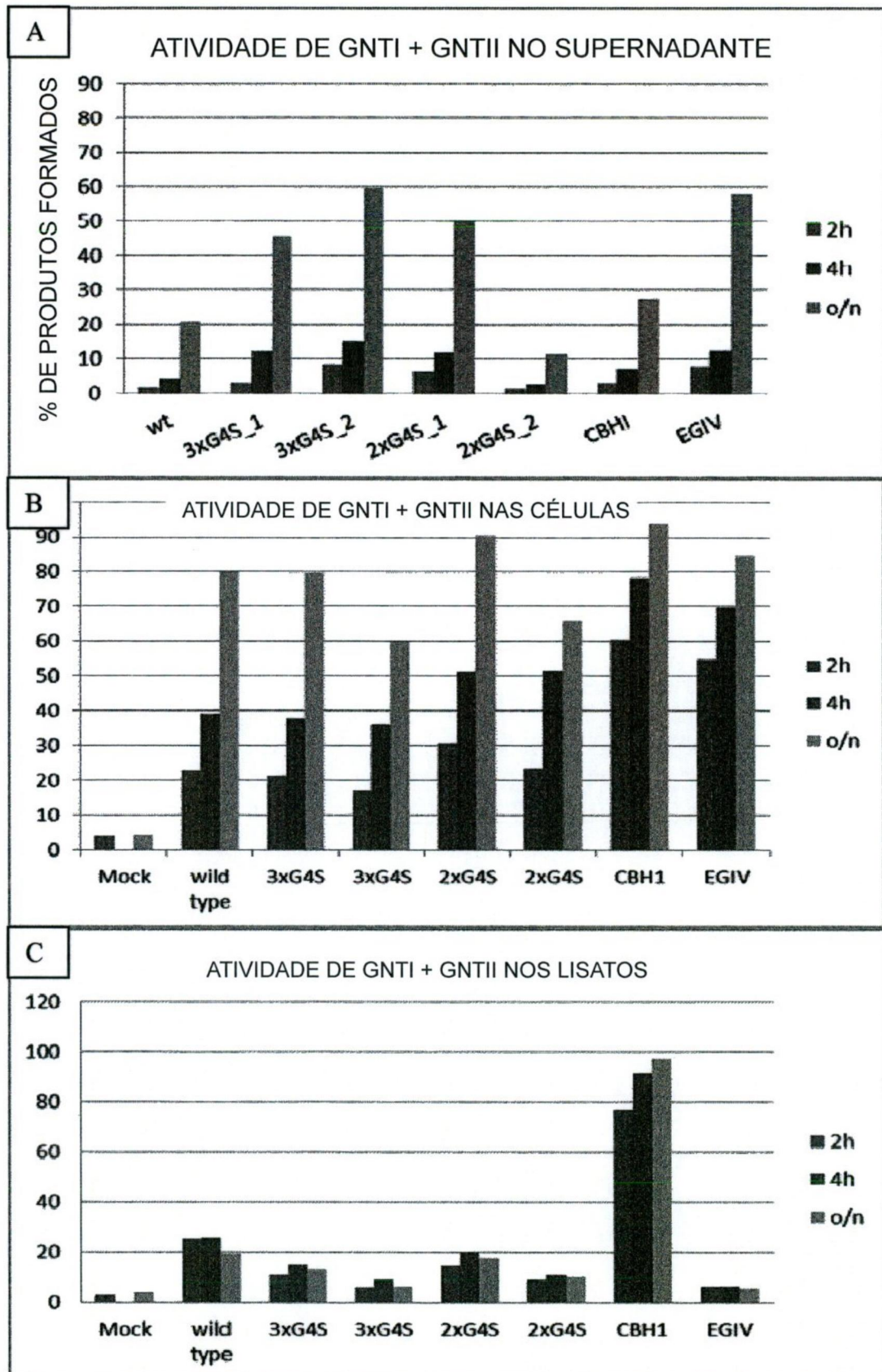


FIG. 34

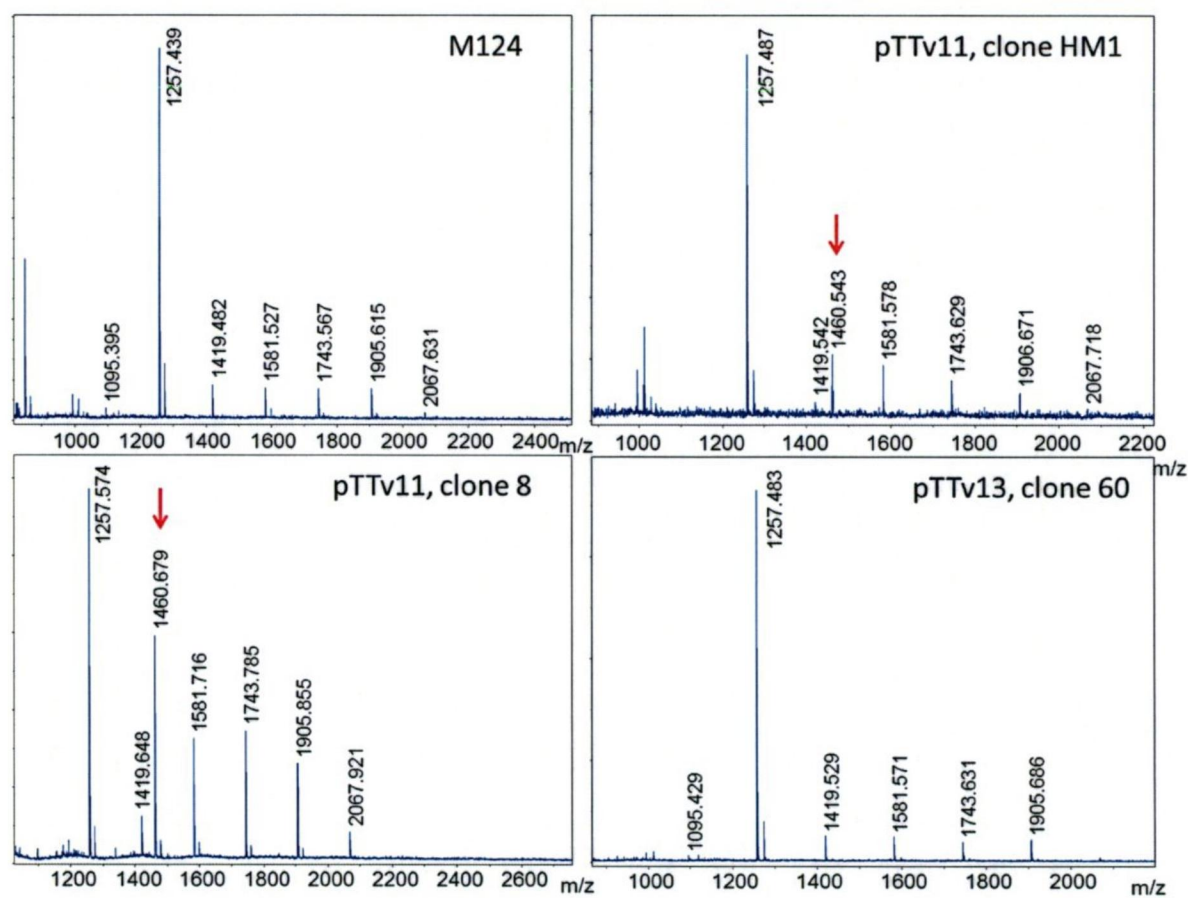


FIG. 35

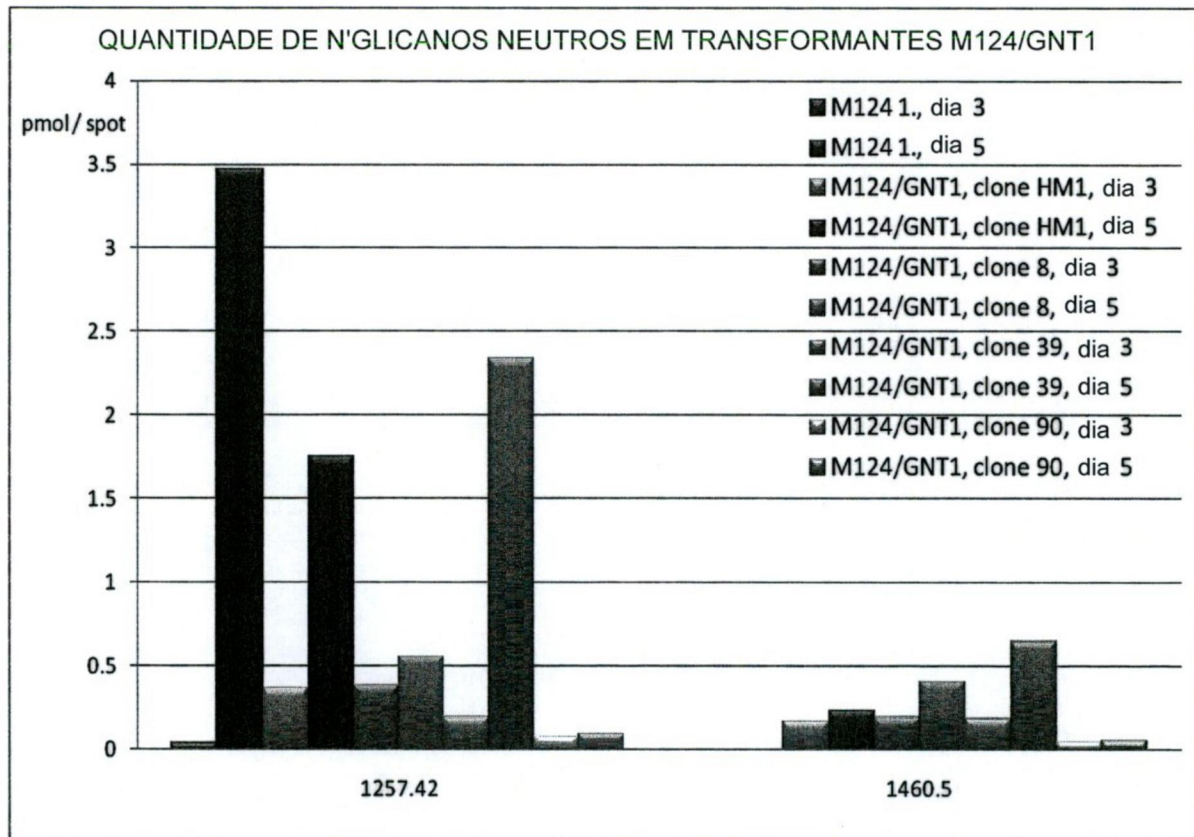


FIG. 36

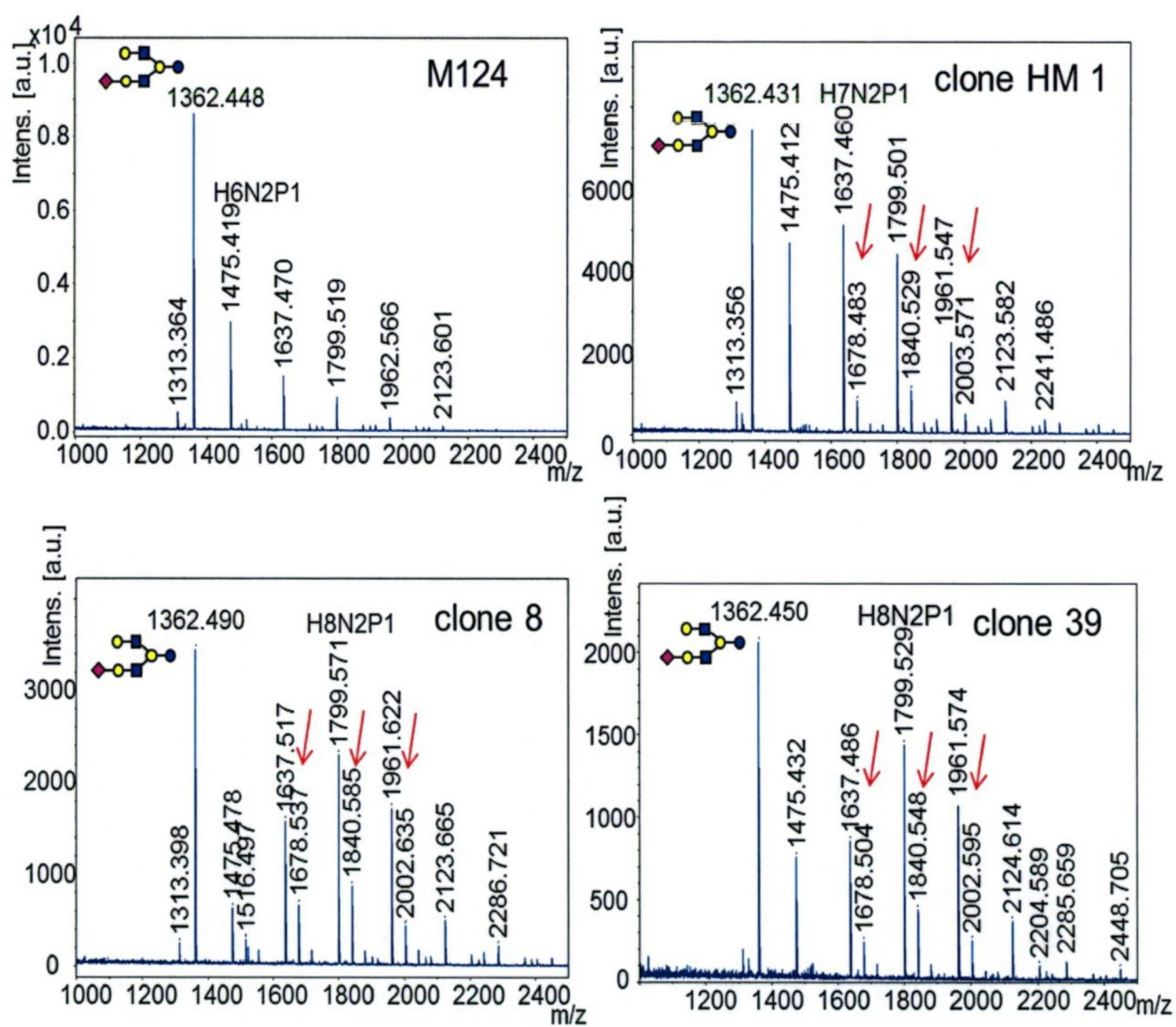
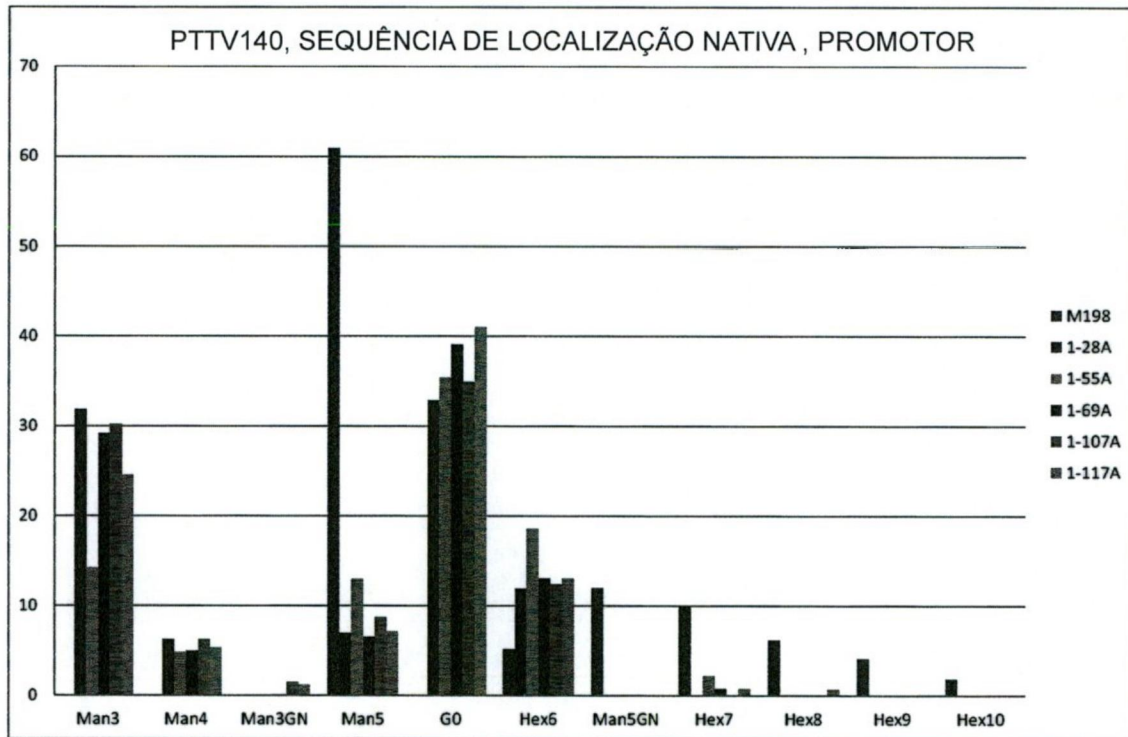


FIG. 37

A.



B.

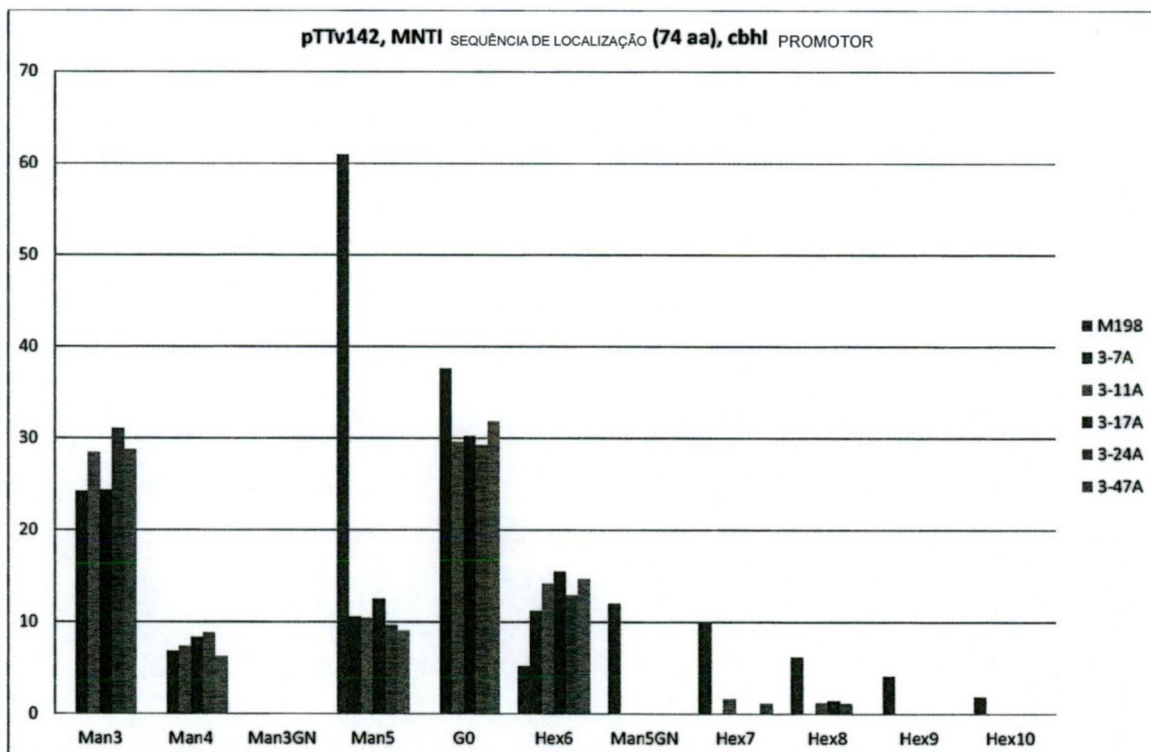
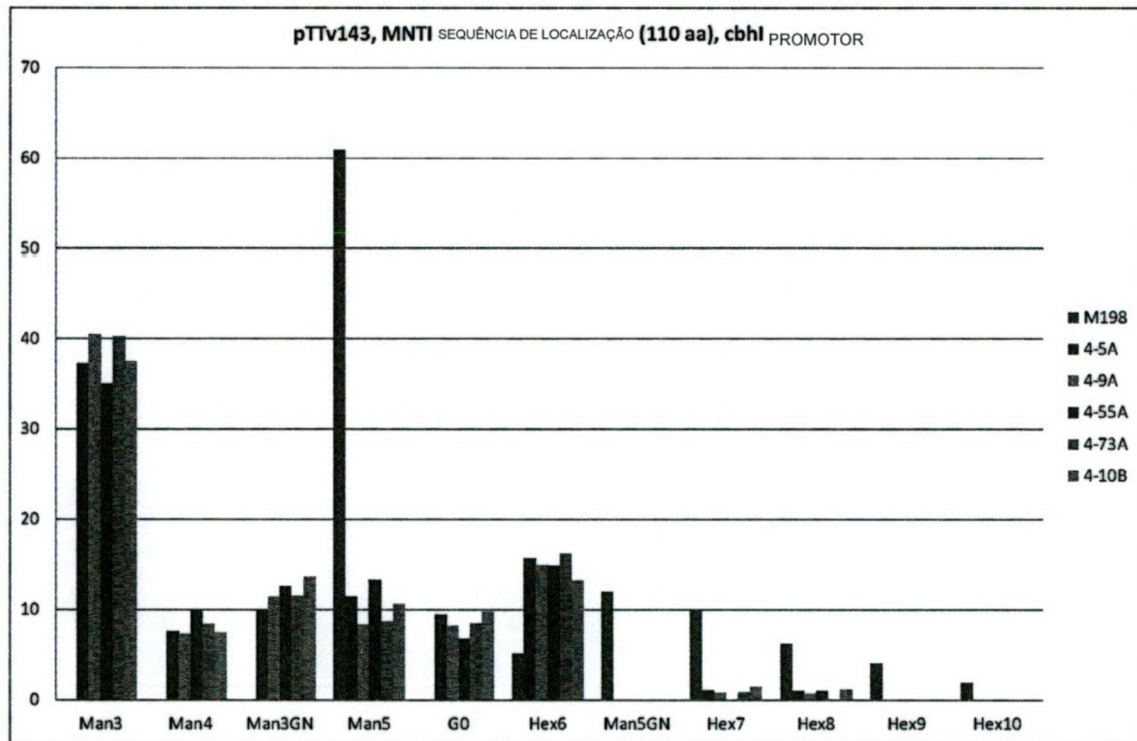


FIG. 37 (Continuação)

C.



D.

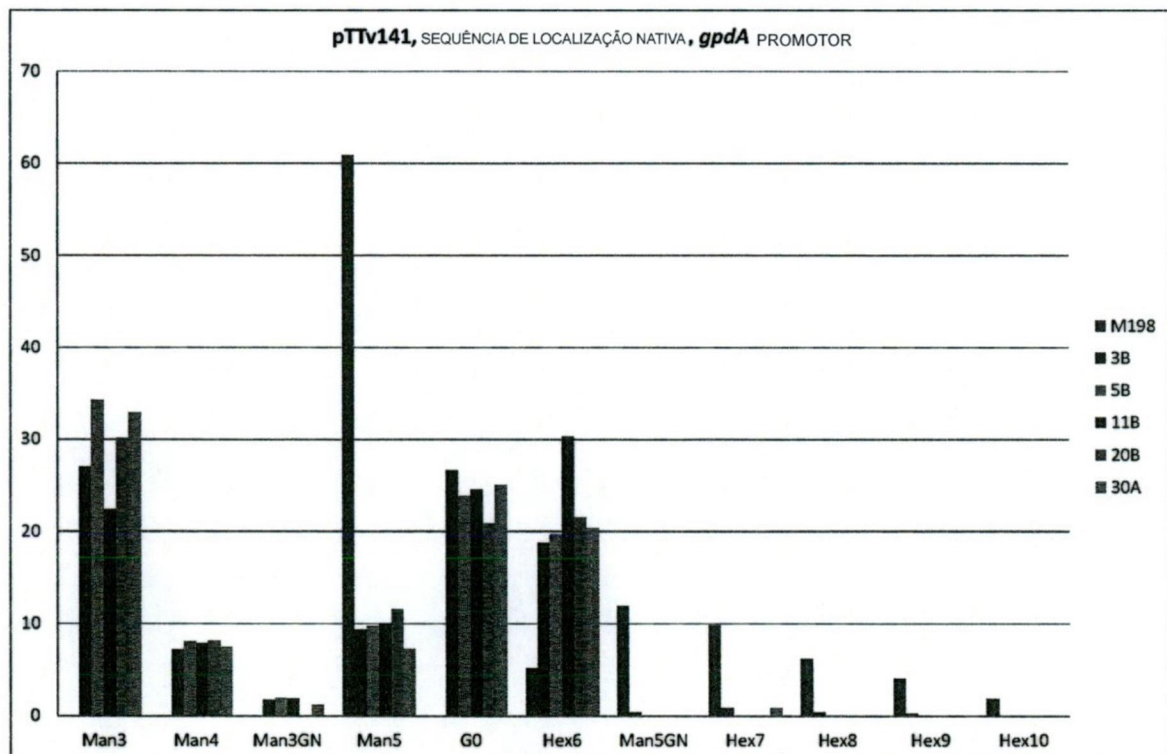
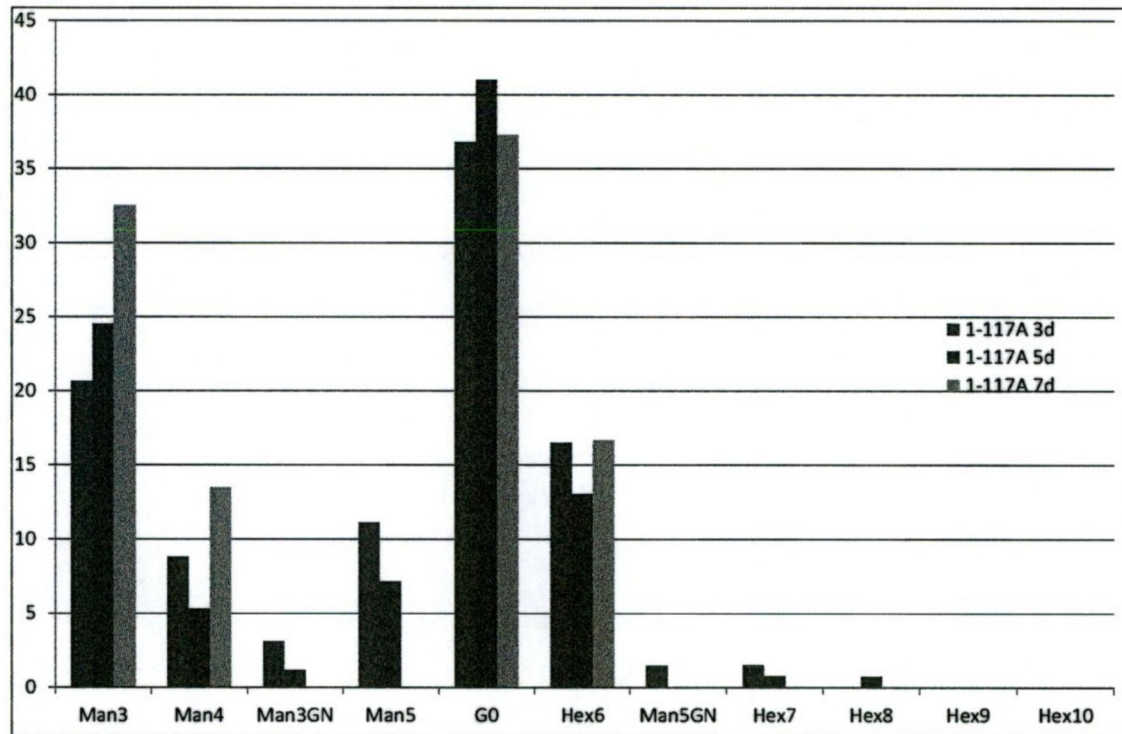
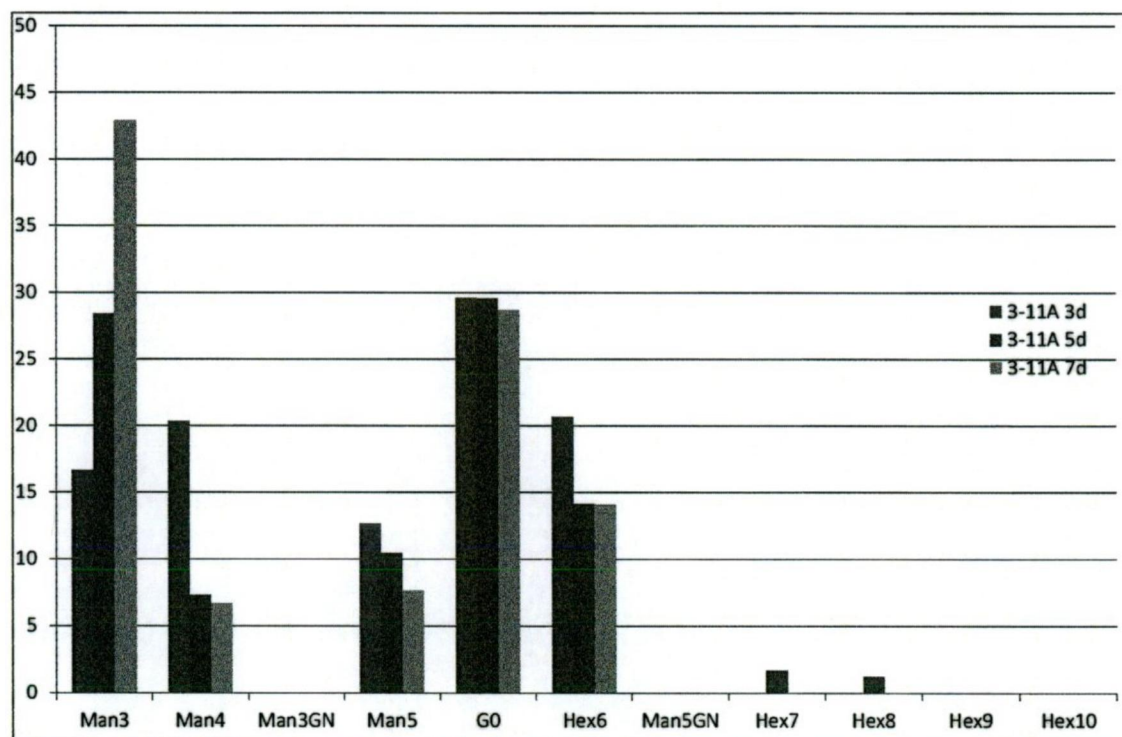


FIG. 38

A.

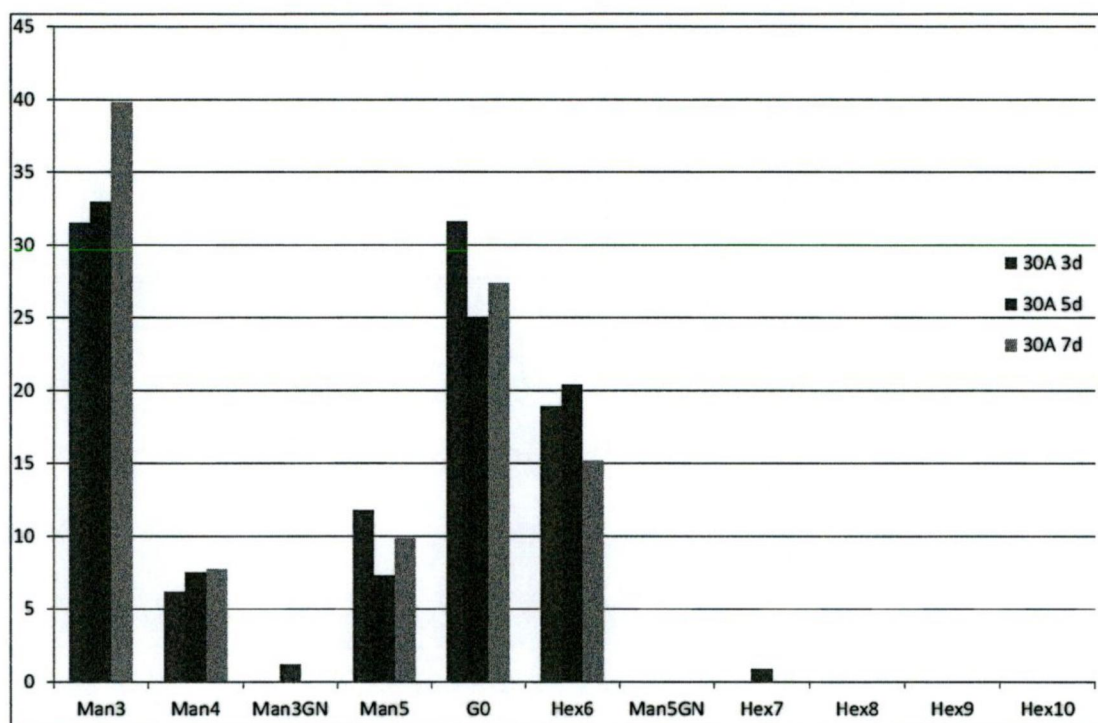


B.



C.

FIG. 38 (Continuação)



D.

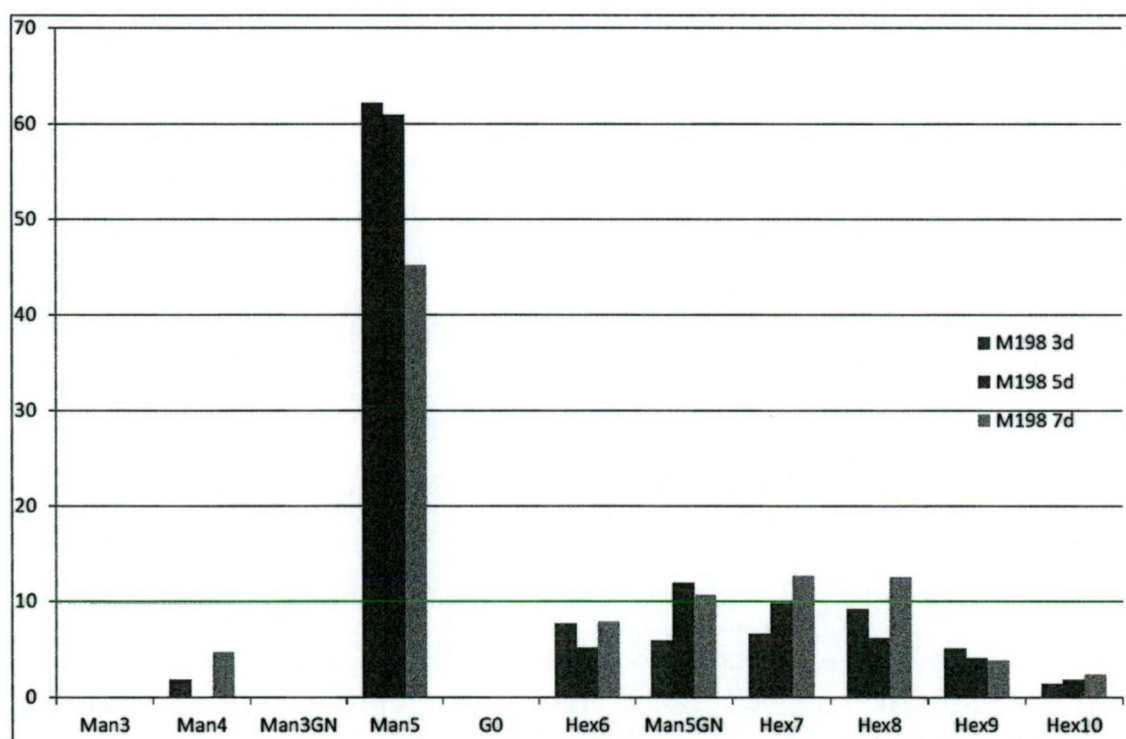


FIG. 39

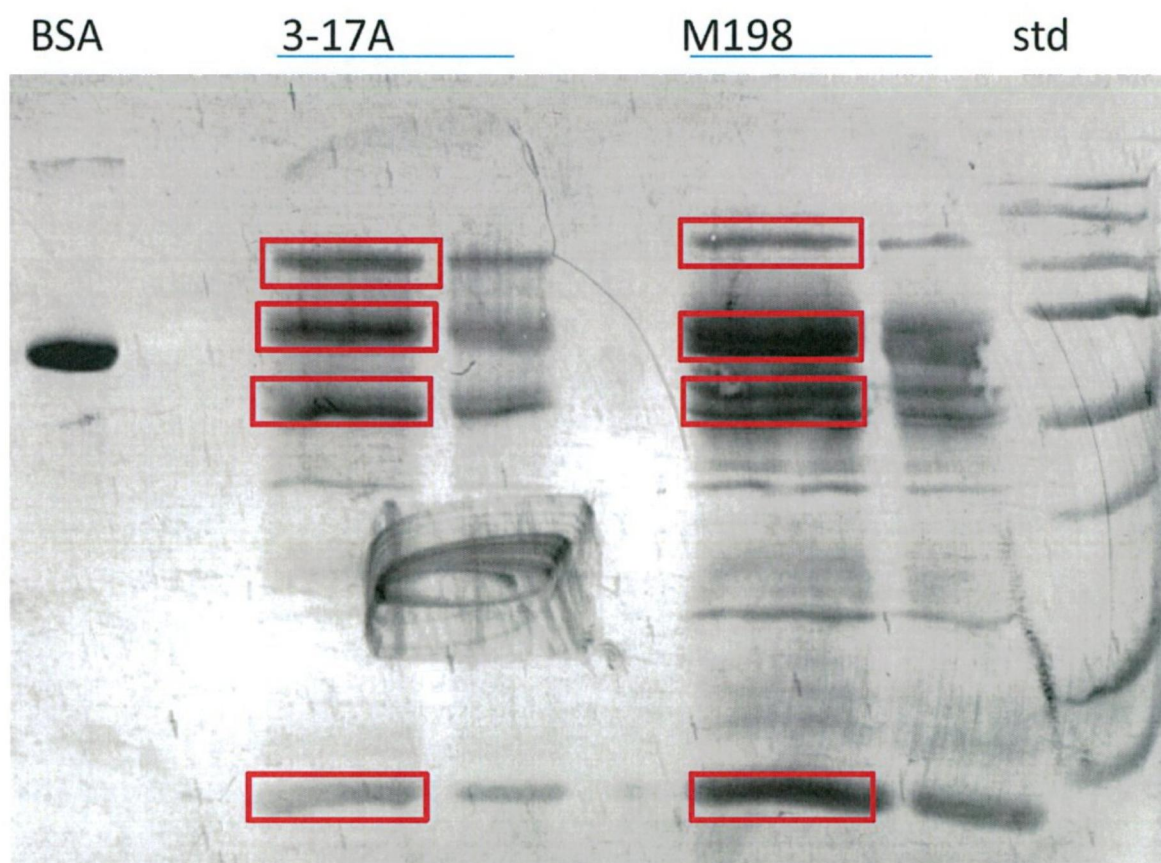
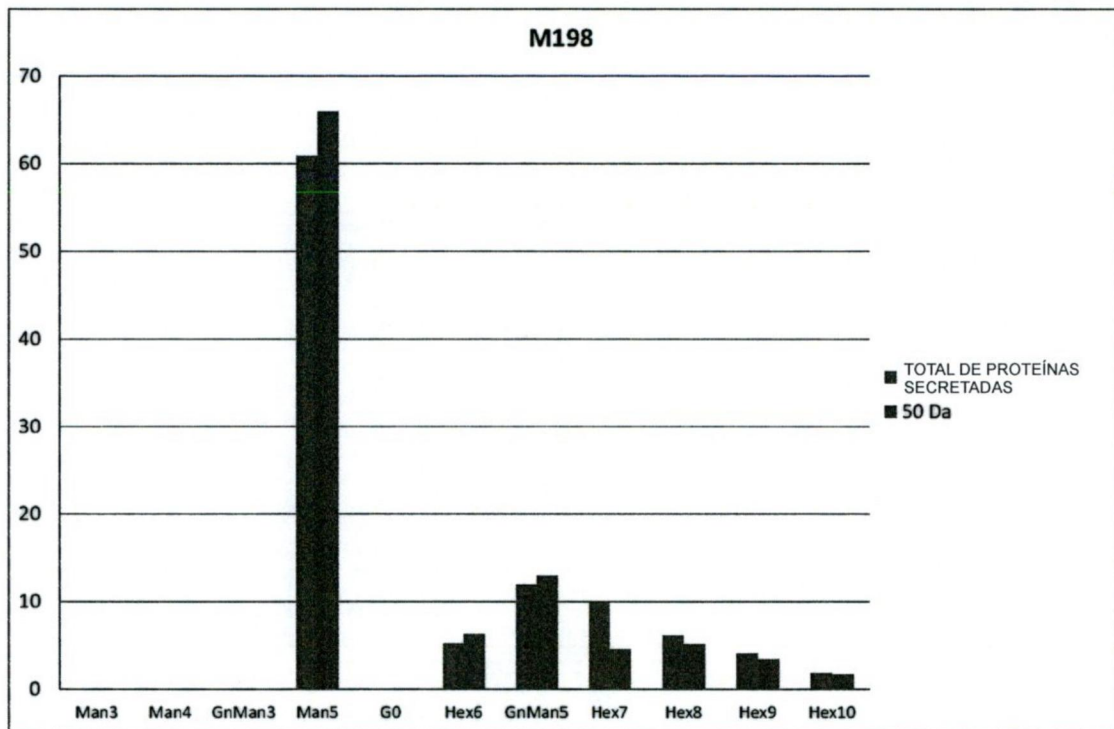


FIG. 40

A.



B.

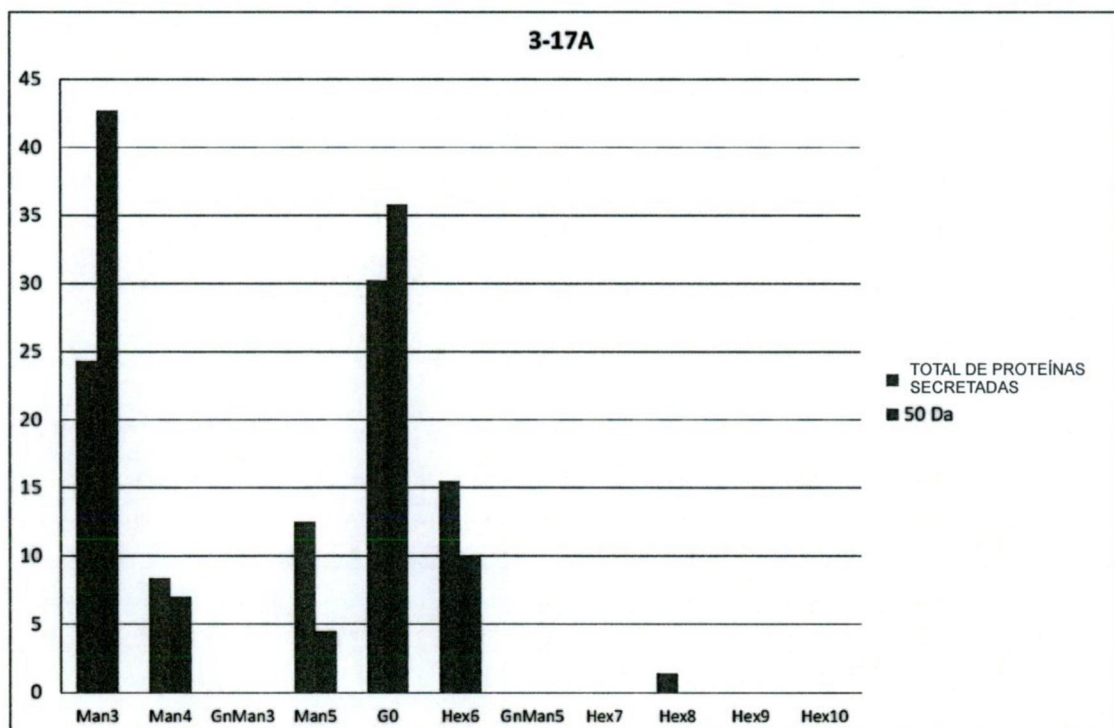


FIG. 41

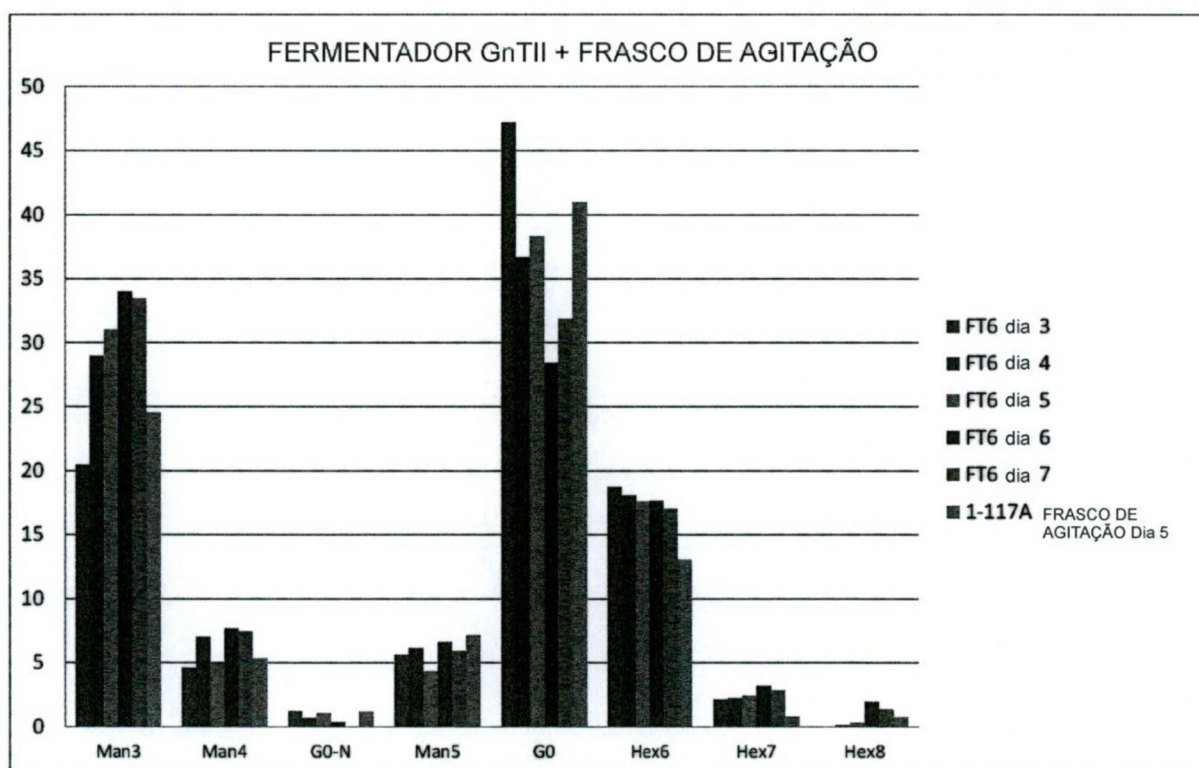


FIG. 42

Aoryzae_gi|169779159|ref|XP_00
 Aniger_gi|317037568|ref|XP_001
 Nfischeri_gi|119478936|ref|XP_001
 Treesei_jgi|Trire2|104121|fgen
 Tvirens_jgi|TriviGv29_8_2|1944
 Tatroviride_jgi|Triat2|270085|
 Foxysporum_gi|342880904|gb|EGU
 Gzea_gi|46136275|ref|XP_389829
 Mthermophila_gi|347013322|gb|A
 Ncrassa_gi|85106839|ref|XP_962
 Moryzae_gi|145611997|ref|XP_36
 Spombe_gi|19114765|ref|NP_5938

Aoryzae_gi|169779159|ref|XP_00
 Aniger_gi|317037568|ref|XP_001
 Nfischeri_gi|119478936|ref|XP_001
 Treesei_jgi|Trire2|104121|fgen
 Tvirens_jgi|TriviGv29_8_2|1944
 Tatroviride_jgi|Triat2|270085|
 Foxysporum_gi|342880904|gb|EGU
 Gzea_gi|46136275|ref|XP_389829
 Mthermophila_gi|347013322|gb|A
 Ncrassa_gi|85106839|ref|XP_962
 Moryzae_gi|145611997|ref|XP_36
 Spombe_gi|19114765|ref|NP_5938

Aoryzae_gi|169779159|ref|XP_00
 Aniger_gi|317037568|ref|XP_001
 Nfischeri_gi|119478936|ref|XP_001
 Treesei_jgi|Trire2|104121|fgen
 Tvirens_jgi|TriviGv29_8_2|1944
 Tatroviride_jgi|Triat2|270085|
 Foxysporum_gi|342880904|gb|EGU
 Gzea_gi|46136275|ref|XP_389829
 Mthermophila_gi|347013322|gb|A
 Ncrassa_gi|85106839|ref|XP_962
 Moryzae_gi|145611997|ref|XP_36
 Spombe_gi|19114765|ref|NP_5938

Aoryzae_gi|169779159|ref|XP_00
 Aniger_gi|317037568|ref|XP_001
 Nfischeri_gi|119478936|ref|XP_001
 Treesei_jgi|Trire2|104121|fgen
 Tvirens_jgi|TriviGv29_8_2|1944
 Tatroviride_jgi|Triat2|270085|
 Foxysporum_gi|342880904|gb|EGU
 Gzea_gi|46136275|ref|XP_389829
 Mthermophila_gi|347013322|gb|A
 Ncrassa_gi|85106839|ref|XP_962
 Moryzae_gi|145611997|ref|XP_36
 Spombe_gi|19114765|ref|NP_5938

-----MELKHFIHELCLN-PRHTKWIAPLLVI 26
 -----MDWMRLIRDLCFN-PRHTKWMAPLLVL 26
 -----MDLKHTLRDLN-PRHTRWVAPLLIL 26
 -----MA-----SLIKTAVDIANGRHALSRYVIFGLWL 28
 -----MA-----SLIKFASDVANGRHALSKEIFMGLWL 28
 -----MA-----SLIKFASDVATGRHALSKEIFVGLFL 28
 MPESASGTL-----QGVRFRLNVLNGRHALSKEIFIALWL 36
 MADPAPGALA-----RGRTRFVRNVLTGQHALSKEIFVALWL 36
 MTRMRSSPKPTATMADQNRPIHIRATRLVFDILNGRHVSKEIFPLVFL 50
 MAAPSSRPES-----NPPLYKQALDFALDVANGRHALSKEIFPALFL 42
 MAAERPSTLG-----KPVQFVFDVANGRHPLSRAIPMLLA 36
 MSSVET-----RNSFNPFRVLFDLGSYGWLHPSRLLL 32

GDAFLCALIIWKIPYTEIDWTTMQQIALYISGERDYTLIKGSTGPLVYP 76
 GDAFLCALIIWKVPYTEIDWATYMQQISLYLSGERDYTLIRGSTGPLVYP 76
 GDAVLCALIIWKVPYTEIDWTTMQQISLYISGERDYTLIKGSTGPLVYP 76
 ADAVLCGLIIWKVPYTEIDWVAYMEQVTQFVHGERDYPKMEGGTGPLVYP 78
 ADAVLCGLIIWKVPYTEIDWVAYMEQITQFVHGERDYPKMEGGTGPLVYP 78
 ADAVLCGLIIWKVPYTEIDWVAYMEQVTQFVNGERDYPKMEGGTGPLVYP 78
 VDALCGGLIIWKIPYTEIDWVAYMQQISQFVSGERDYTKMEGGTGPLVYP 86
 ADAVGTSLIIWKVPYTEIDWEAYMQQVSQFISGERDYTKIEGGTGPLVYP 86
 ADALLCALIIWKVPYTEIDWVAYMEQVAQILSGERDYTKIRGNTGPLVYP 100
 VDALCGGLIIWKVPYTEIDWVAYMEQVSQILSGERDYTKVGGTGPLVYP 92
 FDGLLCGLIIKKVP----- 50
 LEIPFVFAIISKVPYTEIDWVAYMEQVNSFLLGERDYKSLVGCTGPLVYP 82
 : * * *

AAHVSYMALYHLTDEGRDILFGQILFAVLYLVTLAVVMVCYRQSGAPPY 126
 AAHVSYTALYHLTDEGRDIFFGQILFAVLYLITLVVVLCCYRQSGAPPY 126
 AAHVYIFNLYHLTDEGRDIFLGQILFAVLYLATLTAMTCYRQAGAPPY 126
 AAHVYIYTGlyLTNKGTDILLAQQLFAVLYMATLAVVMTCYSKAKVPPY 128
 AAHVYIYTGlyLTNKGTDILLAQQLFAVLYMATLGVVMLCYWKAKVPPY 128
 AAHVYIYTGlyLTNKGTDILLAQQLFAVLYMATLGVVMLSYWKAKVPPY 128
 AAHVYTYTGlyITDKGNTILLAQQLFAVLYMATLAVVMMLCYWKAKVPPY 136
 AAHVYTYTGlyITNKGNTILLAQQLFAVLYMATLAVVMMLCYWKAKVPPY 136
 AAHVYIYTGlyHLTDEGRNITLAQQLFGFLYMTLAVVMMACYWKAKVPPY 150
 AAHVYIYTGlyHLTDEGRNITLAQQLFAGLYMTLAVVMGCYWKAKVPPY 142
 -----SCYRKAKVPPY 61
 GGHVFLYTLTYLTGDTNIVRAQYIFAFVYWITTAIVGYLFKIVRAPFY 132
 : * * *

LFPLLVLSKRLHSVFLRLFNDGLAVCAMWIAILLFQNKKTWAGVTAWTV 176
 LLPLLVLSKRLHSVFLRLFNDGLAALAMWVAILLFMNKKTWAGVTAWTV 176
 LLPLLVLSKRLHSVFLRLFNDGFAAYAMWVAILLFMNKKTWAGVTAWTV 176
 IFPLLVLSKRLHSVFLRNFDCFAAFWLWLCIFFFQRRWTIGALAYSI 178
 IFPLLVLSKRLHSVFLRNFDCFAAFWLWLSIFFFQRRVWTLGAIAYTI 178
 IFPLLVLSKRLHSVFLRNFDCFAAFWLWLCIYFQNRWTFGALAYTL 178
 MFIFLIASKRLHSVFLRNFDCFAVFFLWLTIFLQRRQWTVGSLVYSW 186
 MFVFLIASKRLHSVFLRNFDCFAVFFLWLSIYFFQRRNWTGSLAYTW 186
 VFPLLVLSKRLHSVFLRNFDCFAVFFLWLAIFALQRRWTFGALMYTL 200
 LFPLLVLSKRLHSVFLRNFDCFAVFFLWLAIFFFQRRNWTGALAYTL 192
 VLPLLVLSKRLHSVFLRNFDCFAVFFLWLAICYFQRRWTFGALAYTL 111
 IYVLLVLSKRLHSVFLRNFDCFAVFFLWLAIFLSSCKKVVRAVILSV 182
 : * * * * * : * * * * *

FIG. 42 (Continuação)

Aoryzae_gi|169779159|ref|XP_00
 Aniger_gi|317037568|ref|XP_001
 Nfischeri_gi|119478936|ref|XP_001
 Treesei_jgi|Trire2|104121|fgen
 Tvirens_jgi|TriviGv29_8_2|1944
 Tatroviride_jgi|Triat2|270085|
 Foxysporum_gi|342880904|gb|EGU
 Gzea_gi|46136275|ref|XP_389829
 Mthermophila_gi|347013322|gb|A
 Ncrassa_gi|85106839|ref|XP_962
 Moryzae_gi|145611997|ref|XP_36
 Spombe_gi|19114765|ref|NP_5938

Aoryzae_gi|169779159|ref|XP_00
 Aniger_gi|317037568|ref|XP_001
 Nfischeri_gi|119478936|ref|XP_001
 Treesei_jgi|Trire2|104121|fgen
 Tvirens_jgi|TriviGv29_8_2|1944
 Tatroviride_jgi|Triat2|270085|
 Foxysporum_gi|342880904|gb|EGU
 Gzea_gi|46136275|ref|XP_389829
 Mthermophila_gi|347013322|gb|A
 Ncrassa_gi|85106839|ref|XP_962
 Moryzae_gi|145611997|ref|XP_36
 Spombe_gi|19114765|ref|NP_5938

Aoryzae_gi|169779159|ref|XP_00
 Aniger_gi|317037568|ref|XP_001
 Nfischeri_gi|119478936|ref|XP_001
 Treesei_jgi|Trire2|104121|fgen
 Tvirens_jgi|TriviGv29_8_2|1944
 Tatroviride_jgi|Triat2|270085|
 Foxysporum_gi|342880904|gb|EGU
 Gzea_gi|46136275|ref|XP_389829
 Mthermophila_gi|347013322|gb|A
 Ncrassa_gi|85106839|ref|XP_962
 Moryzae_gi|145611997|ref|XP_36
 Spombe_gi|19114765|ref|NP_5938

Aoryzae_gi|169779159|ref|XP_00
 Aniger_gi|317037568|ref|XP_001
 Nfischeri_gi|119478936|ref|XP_001
 Treesei_jgi|Trire2|104121|fgen
 Tvirens_jgi|TriviGv29_8_2|1944
 Tatroviride_jgi|Triat2|270085|
 Foxysporum_gi|342880904|gb|EGU
 Gzea_gi|46136275|ref|XP_389829
 Mthermophila_gi|347013322|gb|A
 Ncrassa_gi|85106839|ref|XP_962
 Moryzae_gi|145611997|ref|XP_36
 Spombe_gi|19114765|ref|NP_5938

GVGIKMTLLLLAPAIIVVTVLSLS-LVPSIRLGILALLIQVLLAIPFLQG 225
 GVAIKMTLLLLAPAIIVVTVLSLS-LGPSVGLGVLAVLVQVLLAIPFLQN 225
 GVGIKMTLLLLAPAIIVVTVLSLS-LGPSMQLGFLAVLIQVLFIPFLQN 225
 GLGVKMSLLLVLPAVVIVLYLGRG-FKGALRLWLWMVQVQLLLAIPFIT 227
 GLGVKMSLLLVLPAVVIVLYLGRG-FKGALRLWLWMVQVQLLLAIPFIT 227
 GLGVKMSLLLVLPAVVIVLYLGRG-FKGALRLWLWMVQVQLLLAIPFIT 227
 GLGIKMSLLLVLPAIGVILFLGRG-LWPSRLRLAWLMAQVQFAIGLPPFIT 235
 GLGIKMSLLLVLPAIGVILFLGRG-FWPGRLRLAWLMAQVQFAIGLPPFIT 235
 GLGVKMSLLLVLPAVGIVLLGAG-FATSLRLAAVIGLVQVLIAPVFLSN 249
 GLGVKMTLLLLLPAVGIVLFLGSGSFVTTLQVLATMGLVQILIGVPPFLAH 242
 GLGIKMTVLLSLPAVGIVLLGRG-FGGALNVAISIMGQLQVAIGLPPFLSK 160
 ACSVKMSLLLYVPAYLVLLQLLG-PKKTWMHIFVIIIVQILFSIPFLAY 231
 . . :*: * * * : : . :* . :*: :

NPIGYVARAFELTRQFMFKWTNVNRFVGEDFLSKQFSLALLGLHIFLLG 275
 NPAGYLSRAFELTRQFMFKWTNVNRFVGEVFLSKSFSLALLAVHIVLLG 275
 NPAGYVSRAFELTRQFMFKWTNVNRFVGEELFLSRKFSALLALHILLLG 275
 NWRGYLGRAFELSRQFKFEWTVNWRMLGEDFLSRGFSITLLAFHAIFLL 277
 NWKGYLGRAFELSRQFKFEWTVNWRMLGEELFLSRGFSITLLAFHALFLL 277
 NWAGYLGRAFELSRQFKFEWTVNWRMLGEETFLSRGFSITLLTFHVVTLL 277
 NPRGYAARAFELSRQFQFKWTNVNWRMLGEEVFLSKYFALSLLACHILVLL 285
 NSRGYAARAFELSRQFKFEWTVNWRMLGEEVFLSKSFALFLACHVTALL 285
 NPWGYLGRAFELSRQFFFKWTNVNRFVGEVFLSKFSLALLGLHVAVLA 299
 YPTEYLSRAFELSRQFFFKWTNVNRFVGEEIFLSKGFALTLLALHVLVLL 292
 NAWGYLSRAFELSRQFMFKWTNVNRFVGEETFLSKPFALTLLALHASVLL 210
 FWS-YWTQAFDFGRAFDYKWTNVNRFIPRSIFESTSFSTSIILFLHVALLV 280
 * :*: * * * :*: : . . * * * : * * *

LFVTTGWLRLPSGSNVPDFLRLSLQG-----RQRTVVLSSKSFIMTV 315
 AFAVTGWLRLYRSRSLPAFIRNLLAG-----RHRTVSLPKPYIMSV 315
 LFVATVWLKPSGSDLSPLQLRIQR-----RYRTASLSKSFIMTA 315
 AFILGRWLKIRERTVLGMIPYVIRFRSPFTEQEERASINRVVTPGYVMST 327
 IFILGRWLRIKERSFLGMIPYVLRFTSPFTEHEEASISHRVVTPGYVMSA 327
 VFIAARWLKLQERSLLGIITYAVRFQSPFTEQEAKVSKKVTPRYVLAT 327
 IFISKRWIQPTGRSLYDLIPSLRLKSPFTMQEQLRISH-YVTPYAMTT 334
 VFISQRWLQPTGRPLSAMIPSLQLKSPFTLQEQRLRISH-YVTPYVMTT 334
 IFVTTRWLKPARKPVSQLIVPILLG-KSPFTEEEQRAVSRDVTPRFILT 348
 IFITTRWIKPARKSLVQLISPVLLAGKPPLTVPEHRAAARDVTPTYIMTT 342
 AFVTKRWLKPASKSIGGLIAPLLSG-RPIFTAEEAQTAAARAVTPYVMTT 259
 AFTCKHWNKLSRATPFAMVNSMLTLKP-----LPKLQLATPNFIFTA 322
 * * : : : . : :

MLTSLAIGLLCARSLHYQFFAYLSWATPCLLWRARLHPILYAIWALQEW 365
 MLSSLTVGLLCARSLHYQFFAYLSWATPFLWLWRAGFHPILLYLIWAMQEW 365
 MLSSLAIGLLCARSLHYQFFAYLACATPFLWLWQAGFHPILVYVWVAQEW 365
 ILSANVVGLLFARSLHYQFYAYLAWATPYLLWTACPNLLVVAPLWAAQEW 377
 MLSANVVGLLFARSLHYQFYAYLAWATPFLWLWTASPNLLVVVPLWAAQEW 377
 ILSANVIGLLFARSLHYQFYAYLAWATPFLWLWTAYPNLLVVVPLWLAQEW 377
 MLTANLIGLLFARSLHYQFYAYLAWATPYLLWRATEDPVIIVAIWAAQEW 384
 MLSANVIGLLFARSLHYQFYAYLAWASPYLIWRATEDPFIVLLIWAQEW 384
 ILSANVVGLLFARSLHYQFYAYLAWMTPYLLWRSVGHPIQYAIWTAQEW 398
 ILSANAVGLLFARSLHYQFYAYVAVSTPFLWLWRAGLHPVLVYLLWAVHEW 392
 MLTANIVGMLFARSLHYQFYAYLAWSTPYLLWRSIHPHLLQWGLWALQEW 309
 LATSNLIGILCARSLHYQFYAWFAWSPYLCYQASFPAPIVIGLWMLQEW 372
 : : :*: * * * :*: : * * : : : : * : *

FIG. 42 (Continuação)

Aoryzae_gi 169779159 ref XP_00	AWNVPSTNASSSVVFSVLAVQVFG-----V	391
Aniger_gi 317037568 ref XP_001	AWNTFPSTNLSSIIVVLSLATQSFG-----V	391
Nfischeri_gi 119478936 ref XP_	AWNTYPTSTNASSLVVILSLAAQVFG-----V	391
Treesei_jgi Trire2 104121 fgen	AWNVPSTPLSSSVVSVLAVTVA-----MA	403
Tvirens_jgi TriviGv29_8_2 1944	AWNVPSTPLSSNVVSVLAVTVA-----MA	403
Tatroviride_jgi Triat2 270085	AWNVPSTPLSSSVVISLVPVCLLSPQLLVSHDIYNFANCSAILRPRGIA	427
Foxysporum_gi 342880904 gb EGU	AWNVPSTDLSSSTIAVNTMLATVV-----LV	410
Gzea_gi 46136275 ref XP_389829	AWNVPSTDLSSRVTVGAMLATVV-----LA	410
Mthermophila_gi 347013322 gb A	AWNVPSTPISSGVVGVLAALTA-----LV	424
Ncrassa_gi 85106839 ref XP_962	AWNVPSTPASSAVVGVVGLVTVA-----GV	418
Moryzae_gi 145611997 ref XP_36	AWNVPSTPVSSGVVGVMAITVG-----AV	335
Spombe_gi 19114765 ref NP_5938	AWNVPSTKLSSLIACVPLITILK-----L	398
	.: ** :.:	
Aoryzae_gi 169779159	LLNSRNALSDAPPRRKGEHIQ-----	413 (SEQ ID NO: 133)
Aniger_gi 317037568	LANSASAFYTMRSNPSGKEHNQ-----	413 (SEQ ID NO: 135)
Nfischeri_gi 1194789	LGNSFSRKHLDS--SQKEHLQ-----	411 (SEQ ID NO: 134)
Treesei_jgi Trire2 1	F---AGSNPQPRETSKPKQH-----	420 (SEQ ID NO: 126)
Tvirens_jgi TriviGv2	F---VGSNPQ--RGAPKPKQL-----	419 (SEQ ID NO: 128)
Tatroviride_jgi Tria	FGQDISATLNPDGVAKPLGELENDGLRVVHLASVQVVSFGLHHAHNELGG	477 (SEQ ID NO: 127)
Foxysporum_gi 342880	YLGTARRAVPAPAAQVGNVDDKNK-----	434 (SEQ ID NO: 129)
Gzea_gi 46136275 ref	YRGARLAVPP--SQARKIEAKNK-----	432 (SEQ ID NO: 130)
Mthermophila_gi 3470	WLGAREDWEP--RRVLLKGEAAKR-----	446 (SEQ ID NO: 131)
Ncrassa_gi 85106839	WFGAREEWEPGMKSSSKKEEAAMR-----	442 (SEQ ID NO: 132)
Moryzae_gi 145611997	MVGAKAEFRP--QVPVAKKVEAKR-----	357 (SEQ ID NO: 136)
Spombe_gi 19114765 r	YTS DYRKP-----	406 (SEQ ID NO: 137)

RESUMO

Patente de Invenção: "**PROTEÍNAS RECOMBINANTES COM ATIVIDADE N-ACETILGLUCOSAMINILTRANSFERASE, POLINUCLEOTÍDEO, VETOR DE EXPRESSÃO, CÉLULAS E MÉTODOS DE PRODUÇÃO DE N-GLICANO COMPLEXO**".

A presente descrição refere-se a proteínas recombinantes tendo atividade de N-acetilglucosaminiltransferase. A presente descrição também se refere a métodos para a produção de N-glicanos complexos, incluindo as etapas de fornecimento de células hospedeiras que contêm tais proteínas recombinantes e a cultura das células hospedeiras de modo que as proteínas recombinantes sejam expressas.

Este anexo apresenta o código de controle da listagem de sequências biológicas de que trata a Resolução INPI 228 de 11/11/2009:

Código de Controle

Campo 1



7E11F2BC3910368B

Campo 2



6BA4CDE39777883D

Outras Informações:

- Nome do Arquivo: P196713.txt
- Data de Geração do Código: 23-05-2013
- Hora de Geração do Código: 14:29:01
- Código de Controle:
 - Campo 1: 7E11F2BC3910368B
 - Campo 2: 6BA4CDE39777883D