



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

(21) BR 112012011980-0 A2



(22) Data do Depósito: 19/11/2010

(43) Data da Publicação Nacional: 08/09/2021

(54) Título: MÉTODOS DE PRODUZIR CÉLULAS TRANSFORMADAS DE YARROWIA LIPOLYTICA E PROTEÍNAS-ALVO COMPREENDENDO N-GLICANOS, CÉLULAS TRANSFORMADAS DE YARROWIA LIPOLYTICA E SUAS CULTURAS, BEM COMO COMPOSIÇÃO COMPREENDENDO GLICOPROTEÍNAS

(51) Int. Cl.: C12N 9/10; C12P 21/00; C12N 9/24; C12N 15/81.

(30) Prioridade Unionista: 19/11/2009 US 61/262,828.

(71) Depositante(es): OXYRANE UK LIMITED.

(72) Inventor(es): STEVEN CHRISTIAN JOZEF GEYSENS; WOUTER VERVECKEN.

(86) Pedido PCT: PCT IB2010003154 de 19/11/2010

(87) Publicação PCT: WO 2011/061629 de 26/05/2011

(85) Data da Fase Nacional: 18/05/2012

(57) Resumo: MÉTODOS DE PRODUZIR CÉLULAS TRANSFORMADAS DE YARROWIA LIPOLYTICA E PROTEÍNAS-ALVO COMPREENDENDO N-GLICANOS, CÉLULAS TRANSFORMADAS DE YARROWIA LIPOLYTICA E SUAS CULTURAS, BEM COMO COMPOSIÇÃO COMPREENDENDO GLICOPROTEÍNAS. A presente invenção refere-se a métodos e a células fúngicas geneticamente manipuladas úteis para produção de moléculas-alvo contendo N-glicanos complexos similares de mamíferos ou contendo intermediários em uma via de glicosilação mamífera.



Relatório Descritivo da Patente de Invenção para "**MÉTODOS DE PRODUZIR CÉLULAS TRANSFORMADAS DE YARROWIA LIPOLYTICA E PROTEÍNAS-ALVO COMPREENDENDO N-GLICANOS, CÉLULAS TRANSFORMADAS DE YARROWIA LIPOLYTICA E SUAS CULTURAS,**

5 **BEM COMO COMPOSIÇÃO COMPREENDENDO GLICOPROTEÍNAS".**

REFERÊNCIA CRUZADA A RELACIONADOS

Este pedido de patente reivindica a prioridade sobre o pedido de patente U.S. No. Serial 61/262.828, depositado em 19 de novembro de 2009. A descrição do pedido de patente prévio é incorporada por referência

10 em sua totalidade.

CAMPO TÉCNICO

A invenção relaciona-se a métodos e materiais para produzir glicoproteínas em células fúngicas, e mais particularmente, para células fúngicas geneticamente manipuladas para produzir proteínas contendo N-glicanos complexos similares de mamíferos ou proteínas contendo intermediários dentro de uma via de glicosilação mamífera.

ANTECEDENTES

Sistemas de expressão de alto desempenho são necessários para produzir a maioria de produtos biofarmacêuticos (por exemplo, proteínas recombinantes) atualmente em desenvolvimento. A atividade biológica de muitos destes produtos biofarmacêuticos é dependente da sua modificação pós-tradução (por exemplo, fosforilação ou glicosilação). Um sistema de expressão baseado em levedura combina a facilidade da manipulação genética e a fermentação de um organismo microbiano com a capacidade de secretar e modificar proteínas. Entretanto, as glicoproteínas recombinantes produzidas em células de levedura exibem principalmente estruturas heterogêneas de glicano de alta manose e hipermanose, que podem ser prejudiciais para a função proteica, processamento a jusante, e uso terapêutico subsequente, particularmente onde a glicosilação desempenha um papel biologicamente significante.

25 **SUMÁRIO**

Os métodos e as células fúngicas geneticamente manipuladas descritas neste pedido podem ser usados para produzir moléculas-alvo (por exemplo, proteínas-alvo) que contêm N-glicanos similares de mamíferos ou

contêm intermediários dentro da via de glicosilação mamífera (por exemplo, ser humano). As moléculas-alvo isoladas de tais células engenheiradas podem ser usadas para aplicações biofarmacêuticas incluindo produção de anticorpos, produção de citocinas, e para o tratamento de distúrbios metabólicos, tais como distúrbios de armazenamento lisossomal.

Em um aspecto, este documento caracteriza um método de produção de uma célula fúngica (por exemplo, *Yarrowia lipolytica* ou *Arxula adeninivorans*) capaz de produzir proteínas compreendendo GlcNAc-Man₅GlcNAc₂ N-glicanos. O método inclui o fornecimento de uma célula fúngica geneticamente engenheirada para produzir proteínas compreendendo Man₅GlcNAc₂ N-glicanos; e introdução na célula de um ácido nucleico que codifica uma GlcNAc-transferase I, em que o ácido nucleico inclui uma sequência de direcionamento de nucleotídeo para direcionar a GlcNAc-transferase I codificada a um compartimento intracelular (por exemplo, aparelho de Golgi), em que a expressão de GlcNAc-transferase I na célula fúngica produz proteínas incluindo GlcNAcMan₅GlcNAc₂ N-glicanos. O método pode incluir ainda a introdução na célula de um ácido nucleico que codifica uma proteína alvo, em que a célula produz a proteína-alvo modificada para incluir GlcNAcMan₅GlcNAc₂ N-glicanos. A proteína-alvo pode ligar a um receptor Fc. A proteína-alvo pode ser um anticorpo ou fragmento do mesmo. A proteína-alvo pode ser uma glicoproteína terapêutica. A proteína-alvo pode ser Interferon-β, GM-CSF, Interferon-γ ou eritropoietina.

A célula fúngica geneticamente engenheirada para produzir proteínas contendo Man₅GlcNAc₂ N-glicanos pode ser deficiente na atividade de OCH1 e incluir um ácido nucleico que codifica α-1,2-manosidase, em que o ácido nucleico que codifica α-1,2-manosidase inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a α-1,2-manosidase codificada para o retículo endoplasmático. A sequência de direcionamento pode ser uma sequência HDEL.

O método pode incluir ainda a introdução em uma célula de um ácido nucleico que codifica uma manosidase II, em que o ácido nucleico que codifica a manosidase II inclui uma sequência nucleotídica que codifica uma

sequência de direcionamento para direcionar a manosidase II codificada para o aparelho de Golgi, em que a expressão da manosidase II na célula fúngica produz proteínas contendo GlcNAcMan₃GlcNAc₂ N-glicanos.

O método pode incluir ainda a introdução em uma célula de um ácido nucleico que codifica uma galactosiltransferase, em que o ácido nucleico que codifica a galactosiltransferase inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a galactosiltransferase codificada para o aparelho de Golgi, em que a expressão da galactosiltransferase na célula fúngica produz proteínas contendo GalGlcNAc-Man₅GlcNAc₂ ou GalGlcNAcMan₃GlcNAc₂ N-glicanos. A galactosiltransferase pode ser uma fusão de uma UDP-Glc-4-epimerase e o domínio catalítico β-1,4-galactosiltransferase I. Tal método pode incluir ainda a introdução na célula de um ácido nucleico que codifica uma proteína alvo, em que a célula produz a proteína-alvo modificada para conter GalGlcNAcMan₅GlcNAc₂ ou GalGlcNAcMan₃GlcNAc₂ N-glicanos. Os métodos podem incluir o isolamento da proteína-alvo modificada para conter GalGlcNAcMan₅GlcNAc₂ ou GalGlcNAcMan₃GlcNAc₂ N-glicanos.

Em outro aspecto, este documento caracteriza um método de produção de uma proteína-alvo contendo GlcNAcMan₃GlcNAc₂ N-glicanos. O método inclui o fornecimento de uma célula fúngica (por exemplo, *Yarrowia lipolytica* ou *Arxula adeninivorans*) geneticamente engenheirada para incluir um ácido nucleico que codifica uma GlcNAc-transferase I, uma α-1,2-manosidase e uma manosidase II, em que o ácido nucleico inclui uma sequência nucleotídica que codifica uma sequência de direcionamento, ou sequências nucleotídicas que codificam sequências de direcionamento, para direcionar cada proteína codificada para um compartimento intracelular, em que a célula fúngica é deficiente na atividade de OCH1; e introdução na célula de um ácido nucleico que codifica uma proteína alvo, em que a célula produz a proteína-alvo contendo GlcNAcMan₃GlcNAc₂ N-glicanos. O ácido nucleico que codifica α-1,2-manosidase pode incluir uma sequência de direcionamento de retículo endoplasmático para direcionar a α-1,2-manosidase codificada para o retículo endoplasmático. Por exemplo, a sequência de di-

recionamento pode ser uma sequência HDEL. O ácido nucleico que codifica a GlcNAc-transferase I e a manosidase II pode incluir uma sequência de direcionamento de Golgi, ou sequências de direcionamento de Golgi, para direcionar a GlcNAc-transferase I codificada e manosidase II para o aparelho de Golgi. A proteína-alvo pode se ligar a um receptor Fc. A proteína-alvo pode ser um anticorpo ou fragmento do mesmo. A proteína-alvo pode ser uma glicoproteína terapêutica. A proteína-alvo pode ser Interferon- β , GM-CSF, Interferon- γ ou eritropoietina.

Em algumas modalidades, o método pode incluir ainda a introdução na célula de um ácido nucleico que codifica uma galactosiltransferase, em que o ácido nucleico que codifica a galactosiltransferase inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a galactosiltransferase codificada para o aparelho de Golgi, em que a expressão da galactosiltransferase na célula fúngica produz a proteína-alvo modificada para conter GalGlcNAcMan₃GlcNAc₂ N-glicanos. A proteína-alvo modificada para conter GalGlcNAcMan₃GlcNAc₂ N-glicanos pode ser isolada da célula fúngica.

Este documento também caracteriza um método de produção de uma célula fúngica (por exemplo, *Yarrowia lipolytica* ou *Arxula adeninivorans*) capaz de produzir proteínas contendo GlcNAcMan₃GlcNAc₂ N-glicanos. O método inclui o fornecimento de uma célula fúngica geneticamente engenheirada para produzir proteínas contendo Man₃GlcNAc₂ N-glicanos; introdução na célula de um ácido nucleico que codifica uma GlcNAc-transferase I, em que o ácido nucleico inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar GlcNAc-transferase I codificada para um compartimento intracelular (por exemplo, aparelho de Golgi), em que a expressão de GlcNAc-transferase I na célula fúngica produz proteínas contendo GlcNAcMan₃GlcNAc₂ N-glicanos. O método pode incluir ainda a introdução na célula de um ácido nucleico que codifica uma proteína alvo, em que a célula produz a proteína-alvo modificada para conter GlcNAcMan₃GlcNAc₂ N-glicanos. A proteína-alvo pode se ligar a um receptor Fc. A proteína-alvo pode ser um anticorpo ou fragmento do

mesmo. A proteína-alvo pode ser uma glicoproteína terapêutica. A proteína-alvo pode ser Interferon- β , GM-CSF, Interferon- γ ou eritropoietina.

A célula fúngica geneticamente engenheirada para produzir proteínas contendo $\text{Man}_3\text{GlcNAc}_2$ N-glicanos pode ser deficiente na atividade de ALG3, e incluir um ácido nucleico que codifica α -1,2-manosidase, em que o ácido nucleico inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a α -1,2-manosidase codificada para o retículo endoplasmático. Tal célula fúngica ainda pode ser deficiente na atividade de OCH1 e/ou ainda incluir um ácido nucleico que codifica α -1,3-glicosiltransferase (por exemplo, ALG6).

O método pode incluir ainda a introdução na célula de um ácido nucleico que codifica uma GlcNAc-transferase II, em que o ácido nucleico que codifica a GlcNAc-transferase II inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a GlcNAc-transferase II codificada para um compartimento intracelular, em que a expressão de GlcNAc-transferase II na célula fúngica produz proteínas contendo $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos.

O método pode incluir ainda a introdução na célula de um ácido nucleico que codifica uma galactosiltransferase, em que o ácido nucleico que codifica a galactosiltransferase inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a galactosiltransferase codificada para o aparelho de Golgi, em que a expressão da galactosiltransferase na célula fúngica produz proteínas contendo GalGlcNAc- $\text{Man}_3\text{GlcNAc}_2$ ou Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos. A galactosiltransferase pode ser uma fusão de uma UDP-Glc-4-epimerase e o domínio catalítico β -1,4-galactosiltransferase I. O método pode incluir ainda a introdução na célula de um ácido nucleico que codifica uma proteína alvo, em que a célula produz a proteína-alvo modificada para conter GalGlcNAc $\text{Man}_3\text{GlcNAc}_2$ ou Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos.

O método pode incluir ainda a introdução na célula de um ácido nucleico que codifica as subunidades α e β de Glicosidase II, em que a expressão das subunidades α e β de Glicosidase II na célula fúngica produzem

proteínas incluindo GalGlcNAcMan₃GlcNAc₂ ou Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos.

Este documento também caracteriza um método de produção de uma proteína-alvo contendo Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos. O método inclui o fornecimento de uma célula fúngica geneticamente engenheirada para ser deficiente na atividade de ALG3 e incluindo um ácido nucleico que codifica uma GlcNAc-transferase I, uma GlcNAc-transferase II e uma galactosiltransferase, em que o ácido nucleico que codifica a GlcNAc-transferase I, a GlcNAc-transferase II e a galactosiltransferase inclui uma sequência nucleotídica que codifica uma sequência de direcionamento, ou sequências nucleotídicas que codificam sequências de direcionamento, para direcionar cada proteína codificada para um compartimento intracelular (por exemplo, o aparelho de Golgi); e introdução na célula de um ácido nucleico que codifica uma proteína alvo, em que a célula produz a proteína-alvo contendo Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos. A célula fúngica pode ser mais deficiente na atividade de OCH1 e/ou ainda incluir um ácido nucleico que codifica α-1,3-glicosiltransferase, tal como ALG6. A célula fúngica pode incluir ainda um ácido nucleico que codifica as subunidades α e β de Glicosidase II, em que a expressão das subunidades α e β de Glicosidase II na célula fúngica produzem a proteína-alvo contendo Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos.

Em outro aspecto, este documento caracteriza uma célula fúngica isolada geneticamente engenheirada para produzir proteínas contendo GlcNAcMan₃GlcNAc₂ N-glicanos. A célula fúngica pode ser deficiente na atividade de OCH1 e incluir um ácido nucleico que codifica uma α-1,2-manosidase, uma GlcNAc-transferase I, e uma manosidase II, em que o ácido nucleico que codifica α-1,2-manosidase, a GlcNAc-transferase I, e a manosidase II inclui uma sequência nucleotídica que codifica uma sequência de direcionamento, ou sequências nucleotídicas que codificam sequências de direcionamento, para direcionar cada proteína codificada para um compartimento intracelular, em que a expressão da α-1,2-manosidase, da GlcNAc-transferase I e da manosidase II na célula fúngica produz proteínas contendo GlcNAcMan₃GlcNAc₂ N-glicanos. A célula fúngica pode incluir ainda um áci-

do nucleico que codifica uma proteína alvo, em que a célula produz a proteína-alvo modificada para conter GlcNAcMan₃GlcNAc₂ N-glicanos.

Em algumas modalidades, tal célula fúngica inclui ainda um ácido nucleico que codifica uma GlcNAc-transferase II, em que o ácido nucleico que codifica GlcNAc-transferase II inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a GlcNAc-transferase II codificada para um compartimento intracelular, em que a expressão de GlcNAc-transferase II na célula fúngica produz proteínas contendo GlcNAc₂Man₃GlcNAc₂ N-glicanos.

10 Em algumas modalidades, tal célula fúngica inclui ainda um ácido nucleico que codifica uma galactosiltransferase, em que o ácido nucleico que codifica a galactosiltransferase inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a galactosiltransferase codificada para o aparelho de Golgi, em que a expressão da galactosiltransferase na célula fúngica produz proteínas contendo Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos.

Ainda em outro aspecto, este documento caracteriza uma célula fúngica isolada geneticamente engenheirada para produzir proteínas contendo GlcNAc₂Man₃GlcNAc₂ N-glicanos. A célula fúngica é geneticamente engenheirada para ser deficiente na atividade de ALG3 e inclui um ácido nucleico que codifica uma GlcNAc-transferase I e uma GlcNAc-transferase II, em que o ácido nucleico que codifica a GlcNAc-transferase I e a GlcNAc-transferase II inclui uma sequência nucleotídica que codifica uma sequência de direcionamento, ou sequências nucleotídicas que codificam sequências de direcionamento, para direcionar cada proteína codificada para um compartimento intracelular, em que a expressão de GlcNAc-transferase I, e GlcNAc-transferase II na célula fúngica produz proteínas contendo GlcNAc₂Man₃GlcNAc₂ N-glicanos. A célula fúngica geneticamente engenheirada ainda pode ser deficiente na atividade de OCH1 e/ou ainda incluir um ácido nucleico que codifica α-1,3-glicosiltransferase. Uma célula fúngica geneticamente engenheirada também pode incluir um ácido nucleico que codifica uma proteína alvo, em que a célula produz a proteína-alvo modificada

para conter GlcNAc₂Man₃GlcNAc₂ N-glicanos. Uma célula fúngica pode incluir ainda um ácido nucleico que codifica as subunidades α e β de Glicosidase II, em que a expressão das subunidades α e β de Glicosidase II na célula fúngica produz a proteína contendo GlcNAc₂Man₃GlcNAc₂ N-glicanos. A 5 célula fúngica pode incluir ainda um ácido nucleico que codifica uma galactosiltransferase, em que o ácido nucleico que codifica a galactosiltransferase inclui uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a galactosiltransferase codificada para o aparelho de Golgi, em que a expressão da galactosiltransferase na célula fúngica produz 10 proteínas contendo Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos.

Este documento também caracteriza uma cultura substancialmente pura de células de *Yarrowia lipolytica*, um número substancial do qual são geneticamente engenheirados para produzir glicoproteínas contendo Gal₂GlcNac₂Man₃GlcNAc₂ N-glicanos. As células são geneticamente engenheiradas para ser deficientes na atividade de ALG3 e incluir um ácido nucleico que codifica uma GlcNAc-transferase I, uma GlcNAc-transferase II, e uma galactosiltransferase, em que o ácido nucleico que codifica GlcNAc-transferase I, GlcNAc-transferase II e galactosiltransferase inclui uma sequência nucleotídica que codifica uma sequência de direcionamento, ou sequências nucleotídicas que codificam sequências de direcionamento, para direcionar cada proteína codificada para um compartimento intracelular, em que a expressão de GlcNAc-transferase I, GlcNAc-transferase II, e galactosiltransferase na célula produz proteínas contendo Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos. A célula fúngica geneticamente engenheirada ainda pode ser 15 deficiente na atividade de OCH1 e/ou ainda incluir um ácido nucleico que codifica α-1,3-glicosiltransferase (por exemplo, ALG6). As células ainda podem incluir um ácido nucleico que codifica as subunidades α e β de Glicosidase II, em que a expressão das subunidades α e β de Glicosidase II na célula fúngica produzem a proteína-alvo contendo Gal₂GlcNAc₂Man₃GlcNAc₂ N- 20 glicanos. A 25 célula fúngica geneticamente engenheirada ainda pode ser deficiente na atividade de OCH1 e/ou ainda incluir um ácido nucleico que codifica α-1,3-glicosiltransferase (por exemplo, ALG6). As células ainda podem incluir um ácido nucleico que codifica as subunidades α e β de Glicosidase II, em que a expressão das subunidades α e β de Glicosidase II na célula fúngica produzem a proteína-alvo contendo Gal₂GlcNAc₂Man₃GlcNAc₂ N- 30 glicanos.

Em outro aspecto, este documento caracteriza uma cultura substancialmente pura de células de *Yarrowia lipolytica*, um número substancial

das quais são geneticamente engenheiradas para produzir glicoproteínas contendo $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos, em que as células são geneticamente engenheiradas para serem deficientes na atividade de OCH1 e incluir um ácido nucleico que codifica uma α -1,2-manosidase, uma GlcNAc-transferase I, uma manosidase II, uma GlcNAc-transferase II, e uma galactosiltransferase, em que o ácido nucleico que codifica a α -1,2-manosidase, a GlcNAc-transferase I, a manosidase II, a GlcNAc-transferase II, e a galactosiltransferase inclui uma sequência nucleotídica que codifica uma sequência de direcionamento, ou sequências nucleotídicas que codificam sequências de direcionamento, para direcionar cada proteína codificada para um compartimento intracelular, em que a expressão de α -1,2-manosidase, GlcNAc-transferase I, manosidase II, GlcNAc-transferase II, e galactosiltransferase nas células produz proteínas compreendendo $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos.

Este documento também caracteriza uma composição que inclui uma glicoproteína, em que pelo menos 50% (por exemplo, pelo menos 70% ou pelo menos 85% dos N-glicanos na glicoproteína são $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos).

A menos que de outra maneira definido, todos os termos técnicos e científicos usados neste pedido têm o mesmo significado que comumente entendido por um versado ordinário na técnica à qual esta invenção pertence. Embora os métodos e materiais similares ou equivalentes aos descritos neste pedido possam ser usados na prática ou teste da presente invenção, métodos e materiais exemplares são descritos abaixo. Todas as publicações, pedidos de patentes, patentes, Números de Acesso Genbank®, e outras referências mencionadas neste pedido são incorporados por referência na sua totalidade. Em caso de conflito, o presente pedido de patente, incluindo definições, controlará. Os materiais, métodos e exemplos são ilustrativos somente e não destinados a serem limitantes.

Outras características e vantagens da invenção serão evidentes a partir da seguinte descrição detalhada, e das reivindicações.

DESCRÍÇÃO DOS DESENHOS

A figura 1A é uma representação de estruturas de $\text{Man}_5\text{GlcNAc}_2$ e $\text{Man}_3\text{GlcNAc}_2$.

A figura 1B é um diagrama esquemático do plasmídeo pYIOCH1 PUT TOPO.

5 A figura 2 é uma série de eletroferogramas representando a análise de N-glicano de proteínas secretadas obtidas de células selvagens po1d Inuga de *Yarrowia lipolytica* ou células Δoch1 po1d Inuga de *Yarrowia lipolytica*. O N-glicano principal na inativação OCH1 torna-se $\text{Man}_8\text{GlcNAc}_2$. A análise foi realizada usando eletroforese de carboidrato assistida por sequenciador, assistida por fluoróforo de DNA (DSA-FACE). "M5", "M6", "M8" e "M9", referem-se ao número de resíduos de manose conjugados à estrutura da base N-acetilglicosamina. O eixo Y representa as unidades de fluorescência relativas como uma indicação da quantidade de cada estrutura de N-glicano. O eixo X representa a mobilidade relativa de cada estrutura de N-glicano por um tubo capilar. O topo eletroferograma é uma análise de dextransa para uso como um padrão de mobilidade.

10

15

A figura 3 é um diagrama esquemático de plasmídeos pYLHUX-dL2preManHDEL e pYLTUXdL2preManHDEL.

20 A figura 4 é uma série de eletroferogramas representando o perfil de N-glicano após introdução de um ManHDEL (= α-1,2-manosidase marcada com HDEL) cassette de expressão (sob controle do promotor TEF1 ou Hp4d) na cepa G014. "Rd" significa "integração randômica" através das sequências zeta presentes nos vetores mostrados na figura 3. O N-glicano principal de acordo com a expressão de manosidase é $\text{Man}_5\text{GlcNAc}_2$. Conservando o marcador URA3 de uma destas cepas (G018, ver Tabela 2) não modifica o perfil de N-glicano.

25

30 A figura 5 é um esquemático da estratégia de construção de plasmídeos JME926 pPTLeu2-ADE2ex-Hp4dManHDEL (YI) e OXYP289 pPTAxp1-LEU2ex-Hp4dManHDEL (YI). Ver a figura 23 para a construção do vetor pYLTmAXrGnTII.

A figura 6 é uma série de eletroferogramas representando o perfil de N-glicano após introdução de um cassette de expressão ManHDEL

(= α -1,2-manosidase marcada com HDEL) (sob controle do promotor Hp4d) na cepa G014 pela integração direcionada (Tg) no locus LEU2 ou no AXP1. Man₅GlcNAc₂ torna-se o N-glicano principal.

A figura 7 é uma descrição da sequência de aminoácidos (SEQ ID NO:3) e a sequência nucleotídica otimizada por códon de *Yarrowia* (SEQ ID NO:4) da proteína de fusão entre os 100 aminoácidos N-terminais de Kre2p e o domínio catalítico de GlcNAc-transferase I humana. Em negrito: parte Kre2p da proteína de fusão; em fonte normal: parte GnT I da proteína de fusão; sublinhado: códons de início e de parada.

A figura 8 é um diagrama esquemático da estratégia de construção para plasmídeos pYLTmAXhGnTI e pYLHp4mAXhGnTI.

A figura 9 é uma série de eletroferogramas representando o perfil de N-glicano após introdução de GnT I atividade na cepa G036 pela transformação com um vetor que expressa de GnT I. "Rd" significa "integração randômica" através das sequências zeta presentes nos vetores mostrados na figura 8. O N-glicano principal de acordo com a expressão da atividade de GnT I é GlcNAcMan₅GlcNAc₂. No tratamento *in vitro* com α -1,2-manosidase não modifica o perfil significativamente, indicando que somente quantidades pequenas de N-glicanos de alta manose além de Man₅GlcNAc₂ estão presentes. Tratamento com hexosaminidase *in vitro* resulta em um deslocamento de GlcNAcMan₅GlcNAc₂ em direção a Man₅GlcNAc₂.

A figura 10 é uma representação da sequência de aminoácidos (SEQ ID NO:7) e a sequência nucleotídica otimizada por códon de *Yarrowia* (SEQ ID NO:8) da proteína de fusão entre os 36 aminoácidos N-terminais de Mnn2p e o domínio catalítico de manosidase II de *Drosophila melanogaster*. Em negrito: parte de Mnn2p de proteína de fusão; em fonte normal: parte Man II de proteína de fusão; sublinhado: códons de início e de parada.

A figura 11 é uma representação esquemática da estratégia de construção de plasmídeos pYLTmAXDmManII e pYLTmAXDmManII (LEU2ex).

A figura 12 é uma série de eletroferogramas representando o perfil de N-glicano após introdução da atividade de Man II na cepa G040 pe-

la transformação com um vetor de expressão de Man II. "Rd" significa "integração randômica" através das sequências zeta presentes nos vetores mostrados na figura 11. De acordo com a expressão da atividade de Man II, um novo pico aparece com mobilidade eletroforética mais alta, bem como um pico 'de ombro' que corre quase na mesma posição que $\text{Man}_5\text{GlcNAc}_2$. Tratamento com hexosaminidase *in vitro* resulta em um deslocamento para frente destes picos (ao lado do deslocamento observado de GlcNAc- $\text{Man}_5\text{GlcNAc}_2$ em direção a $\text{Man}_5\text{GlcNAc}_2$), indicando a presença de GlcNAc terminal e dessa forma identificando os picos como $\text{GlcNAcMan}_3\text{GlcNAc}_2$ e $\text{GlcNAcMan}_4\text{GlcNAc}_2$. O tratamento *in vitro* com α -1,2-manosidase não modifica o perfil significativamente, indicando que somente pequenas quantidades de N-glicanos de alta manose além de $\text{Man}_5\text{GlcNAc}_2$ estão presentes.

A figura 13 é a sequência de aminoácidos (SEQ ID NO:9) e a sequência nucleotídica otimizada por códon de *Yarrowia* (SEQ ID NO:10) da proteína de fusão entre os 46 aminoácidos N-terminais de Mnn2p, a proteína similar a UDP-Glc-4-epimerase de *Schizosaccharomyces pombe* e o domínio catalítico de β -1,4-galactosil transferase I humana. A parte Mnn2p da proteína de fusão é de 1-46, as sequências ligantes são de 47-49 e 405-408, sequências de epimerase da proteína de fusão são de 50-404, e a parte Man II da proteína de fusão é de 409-763 da SEQ ID NO:9. A parte Mnn2p são dos nucleotídeos 1-138, as sequências ligantes são de nucleotídeos 139-147 e 1213-1224, sequências de epimerase são dos nucleotídeos 148-1212, e a parte Man II é de 1225-2289 da SEQ ID NO:10. Códons de início e de parada estão sublinhados.

A figura 14 é uma representação esquemática da estratégia de construção de plasmídeos pYLTMAXSpGal10hGalTI e pYLTMAXSp-Gal10hGalTI (ADE2ex).

A figura 15 é uma série de eletroferogramas representando o perfil de N-glicano após introdução da atividade de Gal10-GalTI na cepa G040. O transformante resultante G044 foi cultivado em 2 meios diferentes. "Rd" significa "integração randômica" através das sequências zeta presentes nos vetores mostrados na figura 14. De acordo com a expressão da ativi-

de de Gal10-GalTI, um novo pico aparece correndo em uma posição entre Man₇GlcNAc₂ e Man₈GlcNAc₂. Tratamento com galactosidase *in vitro* resulta em um deslocamento para frente deste pico e um aumento igual de GlcNAcMan₅GlcNAc₂ (o último que é confirmado como representando este N-glicano pelo tratamento duplo com galactosidase e hexosaminidase). Isto indica a presença do terminal galactose e dessa forma identificação do novo pico do perfil de G044 como GalGlcNAcMan₅GlcNAc₂. Tratamento *in vitro* com α-1,2-manosidase indica a presença de uma grande quantidade de N-glicanos de alta manose (especialmente Man₈GlcNAc₂) que ainda não foram cortados a Man₅GlcNAc₂.

A figura 16 é uma representação esquemática do plasmídeo p-YLalg3PUT-ALG6.

A figura 17 é uma série de eletroferogramas representando o perfil de N-glicano após introdução de pYLalg3PUT-ALG6 na cepa G036. A superexpressão de ALG6 resulta em uma quantidade significante de picos de glicosilados (GlcMan₅GlcNAc₂ e Glc₂Man₅GlcNAc₂), indicando que a estrutura Glc₃Man₅GlcNAc₂ que foi transferida para a proteína nascente não é completamente cortada em direção a Man₅GlcNAc₂ por Glicosidase II. Dependendo do meio de crescimento, Man₅GlcNAc₂ é gerado parcialmente (ainda algum Man₅GlcNAc₂ e Man₄GlcNAc₂) ou quase completamente cortado em direção a Man₃GlcNAc₂ pela ação da α-1,2-manosidase de *T. reesei* localizada por ER marcada com HDEL. Os picos de Man₄GlcNAc₂ e Man₅GlcNAc₂ são identificados como tal, pela sua sensibilidade em direção à α-1,2-manosidase. Por causa da finalização com glicose, GlcMan₅GlcNAc₂ e Glc₂Man₅GlcNAc₂ são insensíveis em direção a este tratamento. A manosidase de Jack Bean é parcialmente capaz de remover α-1,6-manose livre ligada enquanto também converte Man3-5'GlcNAc₂ em Man₁GlcNAc₂.

A figura 18 é uma representação esquemática da estratégia de construção do plasmídeo pYLTmAXhGnT1 (Hygr ex).

A figura 19 é uma série de eletroferogramas representando os perfis de N-glicano após introdução da atividade de GnT I na versão não curada (G039) ou curada (G045) da cepa Δalg3-Hp4dALG6 pela transforma-

ção com um vetor expressando GnT I. A geração de GlcNAcMan₃GlcNAc₂ foi provada através de uma digestão com hexosaminidase. O novo pico desloca-se completamente de volta em direção a Man₃GlcNAc₂. Na cepa G048, a conversão em direção a GlcNAcMan₃GlcNAc₂ não foi completa uma vez que algum Man₃GlcNAc₂ ainda pode ser observado. Esta cepa também tem algum Man₅GlcNAc₂ remanescente como mostrado pela digestão com α-1,2-manosidase.

A figura 20 é uma representação esquemática da estratégia de construção do plasmídeo JME925 pPTAde2-URA3ex-Hp4dhGnTI.

A figura 21 é uma série de eletroferogramas representando perfis de N-glicano após introdução da atividade de GnT I na versão curada da cepa $\Delta alg3$ -Hp4dALG6 (=G045); a integração de uma expressão dirigida por um construto Hp4d no locus ADE2 (*Tg-ade2*). Nesta cultura, a quantidade de N-glicanos glicosilados foi alta e a conversão de Man₄₋₅GlcNAc₂ a Man₃GlcNAc₂ não foi completa. Um novo pico correndo ao lado de Man_{4'}GlcNAc₂ foi observado no transformante G057 e pode ser designado como GlcNAcMan₃GlcNAc₂ baseado no resultado da digestão com hexosaminidase: o novo pico completamente se desloca de volta em direção a Man₃GlcNAc₂.

A figura 22 é a sequência de aminoácidos (SEQ ID NO:17) e a sequência nucleotídica otimizada por códon de *Yarrowia* (SEQ ID NO:18) da proteína de fusão entre os 36 aminoácidos N-terminais de Mnn2p e o domínio catalítico da GlcNAc-transferase II de rato. Em negrito: parte de Mnn2p da proteína de fusão; em fonte normal: parte GnT II da proteína de fusão; sublinhado: códons de início e de parada.

A figura 23 é uma representação esquemática da estratégia de construção de plasmídeos pYLTmAXrGnTII e pYLTmAXrGnTII (ADE2 ex).

A figura 24 é uma série de eletroferogramas representando perfis de N-glicano após introdução da atividade de GnT II em uma cepa que sintetiza GlcNAcMan₃GlcNAc₂. As cepas resultantes foram obtidas através da transformação dupla de G045 com os construtos de expressão GnT I e GnT II ou através da transformação de G047 com o construto de expressão

GnTII. Em ambos os casos, o pico que representa GlcNAcMan₃GlcNAc₂ desapareceu quase completamente e um novo pico, sobre uma unidade de glicose maior, apareceu. O tratamento com hexosaminidase indica a presença de dois resíduos de GlcNAc terminais para o novo N-glicano; o pico desloca-se aproximadamente duas unidades de glicose à esquerda e dessa forma representa GlcNAc₂Man₃GlcNAc₂. O tratamento com α -1,2-manosidase não resulta em diferenças principais, indicando que somente são limitadas as quantidades de Man₄₋₅GlcNAc₂ presentes.

A figura 25 é um diagrama esquemático de plasmídeos pYL-TUXdL2preAnGlcII e pYLeu2ExTEFpreLip2AnGlucII β para a expressão de atividade de Glicosidase II.

A figura 26 é um esquemático da estratégia de construção de plasmídeos JME923 pPTura3-LEU2ex-TefL2preAnGlcIIa+b [alt1].

A figura 27 é um esquemático da estratégia de construção de plasmídeos JME923 pPTura3-LEU2ex-Hp4dL2preAnGlcIIa+b[alt1] e Zeta-LEU2ex-Hp4dL2preAnGlcIIa+b[alt].

A figura 28 é uma série de eletroferogramas representando perfis de N-glicano após introdução de atividade de Glicosidase II em uma cepa que sintetiza GlcNAcMan3GlcNAc2. As cepas resultantes foram obtidas através de integração randômica (G060) ou direcionadas (G061) de um construto de expressão dual das subunidades gls2 α e gls2 β . Em ambos os casos, uma redução de picos de glicosilados é observada. O tratamento com α -1,2-manosidase indica que nem todos os Man5'GlcNAc₂ gerados foram convertidos em direção a Man₃GlcNAc₂ pela α -1,2-manosidase heteróloga marcadada com HDEL. Por causa da finalização com glicose, GlcMan₅GlcNAc₂ e Glc₂Man₅GlcNAc₂ são insensíveis em direção a este tratamento. A manosidase de Jack Bean é parcialmente capaz de remover α -1,6-manose livre ligada tanto nos N-glicanos glicosilados restantes como em GlcNAc-Man₃GlcNAc₂. Além disso, este tratamento converte Man₃₋₅GlcNAc₂ em Man₁GlcNAc₂. "Rd" significa "integração randômica" através das sequências zeta presentes nos vetores mostrados na figura 27. "Tg-ade2" e "Tg-ura3" significam a integração direcionada no locus ADE2 resp. URA3.

A figura 29 é uma série de eletroferogramas representando o perfil de N-glicano de secretoma das cepas G070 e G071, que foram gerados através da introdução de GlcNAc-transferase II na cepa G061. Os N-glicanos foram tratados com qualquer α-1,2-manosidase (removendo todo os resíduos terminais de α-1,2-manoose ligada) ou hexosaminidase (que remove os resíduos terminais de β-1,2-GlcNAc ligada) para permitir a identificação dos picos nos perfis nativos de G070 e G071. Os N-glicanos contendo glicose não são sensíveis a qualquer das duas enzimas. O tratamento com α-1,2-manosidase resulta na decoração de Man5'GlcNAc2 e Man4GlcNAc2 em direção a Man3GlcNAc2. O tratamento com hexosaminidase remove resíduos terminais de β-1,2-GlcNAc ligada que foram adicionados por GlcNAc-transferase I e II para gerar Man3GlcNAc2.

A figura 30A é a sequência nucleotídica da pré-pró-Lip2-cadeia leve sintética (LC) (SEQ ID NO:32).

15 A figura 30B é a sequência de aminoácidos de pré-pró-Lip2-LC sintético (SEQ ID NO:33)

A figura 31A é a sequência nucleotídica da pré-pró-Lip2-cadeia pesada sintética (HC) (SEQ ID NO:34).

20 A figura 31B é a sequência de aminoácidos de pré-pró-Lip2-HC sintético (SEQ ID NO:35).

A figura 32 é uma série de eletroferogramas representando a análise de perfil de N-glicano de culturas de frasco de agitação SuperT/glicerol de cepas glicoengenheiradas G045, G057, G061 e G071 que foram transformadas com pYLhp4L2pré-pró-HerHC/LC (GUT2ex)-ori2. Ver a 25 Tabela 2 para uma descrição das cepas G045, G057, G061 e G071.

A figura 33 é um gráfico dos resultados a partir de um ELISA funcional em pontos de tempo diferentes na fermentação batelada alimentada de G096.

As figuras 34A e B são a série de eletroferogramas representando a análise de perfil de N-glicano do secretoma em pontos de tempo diferentes dentro da fermentação batelada alimentada de G096.

DESCRÍÇÃO DETALHADA

Como descrito neste pedido, a síntese *in vivo* dos N-glicanos complexos similares de mamíferos em glicoproteínas secretadas pela levedura pode ser baseada em uma estrutura de base de $\text{Man}_3\text{GlcNAc}_2$ ou em $\text{Man}_5\text{GlcNAc}_2$ (ver figura 1A, "Man" refere-se à manose, e "GlcNAc" refere-se a N-glicosamina). Para produzir a estrutura de base de $\text{Man}_5\text{GlcNAc}_2$, as células de levedura podem ser engenheiradas tal que a atividade de α -1,2-manosidase seja aumentada em um compartimento intracelular e a atividade de alongamento de Cadeia Exterior (OCH1) seja reduzida. Para produzir a estrutura de base de $\text{Man}_3\text{GlcNAc}_2$, a atividade da Glicosilação 3 Ligada à Asparagina (ALG3) e, em algumas modalidades, OCH1 é reduzida, atividade de α -1,2-manosidase e, em algumas modalidades, a atividade de α -1,3-glicosiltransferase é aumentada. O perfil de N-glicano de proteínas produzidas em tais células de levedura pode ser alterado pela engenharia adicional das células de levedura para conter uma ou mais das seguintes atividades:

atividade de GlcNAc transferase I (GnT I), atividade de manosidase II, atividade de GlcNAc transferase II (GnT II), atividade de Glicosidase II, e atividade de galactosiltransferase (Gal T). Por exemplo, expressando GnT I em uma célula de levedura que produz $\text{Man}_5\text{GlcNAc}_2$ ou $\text{Man}_3\text{GlcNAc}_2$ N-glicanos resulta na transferência de uma porção GlcNAc a $\text{Man}_5\text{GlcNAc}_2$ ou $\text{Man}_3\text{GlcNAc}_2$ N-glicanos tal que $\text{GlcNAcMan}_5\text{GlcNAc}_2$ ou $\text{GlcNAc-Man}_3\text{GlcNAc}_2$ N-glicanos, respectivamente, sejam produzidos. Em células que produzem $\text{GlcNAcMan}_5\text{GlcNAc}_2$ N-glicanos, expressando uma manosidase II resulta em dois resíduos de manose que são removidos de $\text{GlcNAc-Man}_5\text{GlcNAc}_2$ N-glicanos para produzir $\text{GlcNAcMan}_3\text{GlcNAc}_2$ N-glicanos. Em células que produzem $\text{GlcNAcMan}_3\text{GlcNAc}_2$ N-glicanos, expressando GnT II resulta na transferência de outra porção de GlcNAc a $\text{GlcNAcMan}_3\text{GlcNAc}_2$ N-glicanos para produzir $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos. Expressão de Gal T em células que produzem $\text{GlcNAcMan}_3\text{GlcNAc}_2$ ou $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos resulta na transferência de galactose a $\text{GlcNAcMan}_3\text{GlcNAc}_2$ ou $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos para produzir $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$ ou $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos. Em algumas modalidades, Glicosidase II (por exemplo, por expressão de subunidades α e β) pode ser expressa

para aumentar a produção da estrutura de base de $\text{Man}_3\text{GlcNAc}_2$.

Moléculas alvo

As moléculas alvo, como usado neste pedido, refere-se a qualquer molécula que sofre N-glicosilação em uma célula geneticamente engenheirada (por exemplo, uma célula fúngica, tal como *Yarrowia lipolytica*, *Arxula adeninivorans*, ou outras espécies relacionadas célula de levedura dimórfica; uma célula vegetal, ou célula animal). Em algumas modalidades, as 5 moléculas-alvo são capazes de trafegar por uma ou mais etapas da via secretória de *Yarrowia lipolytica* ou *Arxula adeninivorans* (ou outras espécies relacionadas de levedura dimórfica), resultando na sua N-glicosilação pelo 10 maquinário da célula hospedeira. As moléculas-alvo podem ser endógenas ou exógenas.

Proteínas-alvo adequadas incluem proteínas patógenas (por exemplo, toxoide tetânico; toxoide difterico; proteínas de superfície viral (por 15 exemplo, glicoproteínas de citomegalovírus (CMV) B, H e gCIII; glicoproteínas de envelope de vírus de imunodeficiência humana 1 (VIH 1); glicoproteínas de envelope de vírus de sarcoma de Rous (RSV); glicoproteínas de envelope de herpes simplex vírus (HSV); glicoproteínas de envelope de vírus Epstein Barr (EBV); glicoproteínas de envelope de vírus de varicella-zoster 20 (VZV); glicoproteínas de envelope de vírus papiloma humano (HPV); glicoproteínas de vírus de influenza; e antígenos de superfície da família de Hepatite), proteínas lisossomais (por exemplo, glicocerebrosidase, cerebrosidase, ou galactocerebrosidase), insulina, glucagon, fatores de crescimento, citocinas, quimiocinas, uma ligação de proteína a um receptor Fc, anticorpos 25 ou fragmentos dos mesmos, ou fusões de alguma das proteínas a anticorpos ou fragmentos de anticorpos (por exemplo, proteína-Fc). Fatores de crescimento incluem, por exemplo, Fator de crescimento endotelial vascular (VEGF), fator de crescimento similar à insulina (IGF), proteína morfogênica óssea (BMP), fator estimulante de colônia de granulócitos (G-CSF), fator 30 estimulante colônia de macrófago-granulócito (GM-CSF), fator de crescimento de Nervo (NGF); uma Neurotrofina, fator de crescimento derivado de Plaquetas (PDGF), Eritropoietina (EPO), Trombopoietina (TPO), Miostatina

(GDF-8), fator de Diferenciação de Crescimento 9 (GDF9), fator de crescimento de fibroblasto básico (bFGF ou FGF2), fator de crescimento Epidérmico (EGF), fator de crescimento de Hepatócito (HGF). Citocinas incluem, por exemplo, interleucinas (por exemplo, IL-1 a IL-33, tais como IL-1, IL-2, 5 IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, IL-13, ou IL-15) e interferons (por exemplo, interferon β ou Interferon- γ). Quimiocinas incluem, por exemplo, I-309, TCA-3, MCP-1, MIP-1 α , MIP-1 β , RANTES, C10, MRP-2, MARC, MCP-3, MCP-2, MRP-2, CCF18, MIP-1 γ , Eotaxina, MCP-5, MCP-4, NCC-1, Ck β 10, HCC-1, Leucotactina-1, LEC, NCC-4, TARC, PARC, ou 10 Eotaxina 2. Também incluídas são glicoproteínas tumorais (por exemplo,抗ígenos associados ao tumor), por exemplo, antígeno carcinoembrionário (CEA), mucina humanas, HER-2/neu, e antígeno específico para próstata (PSA) [Henderson e Finn, *Advances in Immunology*, 62, pp. 217-56 (1996)]. Em uma modalidade, a proteína-alvo é um anticorpo anti-HER2/neu. Em algumas modalidades, a proteína-alvo pode ser aquela associada a um distúrbio de armazenamento lisossomal, proteínas alvos que incluem, por exemplo, Alfa-L-iduronidase, beta-D-galactosidase, beta-Glicosidase, beta-hexosaminidase, beta-D-manosidase, alfa-L-fucosidase, arilsulfatase B, arilsulfatase A, alfa-N-acetilgalactosaminidase, aspartilglicosaminidase, iduronato-2-20 sulfatase, alfa-glicosaminida-N-acetyltransferase, beta-D-glicoronidase, hialuronidase, alfa-L-manosidase, alfa-neuraminidase, fosfotransferase, lipase ácida, ácido ceramidase, esfingomielinase, tioesterase, catepsina K, e lipo-proteína lipase.

As proteínas-alvo também podem ser proteínas de fusão. As 25 proteínas de fusão incluem, por exemplo, uma fusão de (i) qualquer proteína descrita neste pedido ou fragmento da mesma com (ii) um anticorpo ou fragmento do mesmo. Como usado neste pedido, o termo "fragmento de anticorpo" refere-se a (a) um fragmento de ligação a antígeno ou (b) uma parte Fc do anticorpo que pode interagir com um receptor Fc. Um fragmento de 30 ligação a antígeno pode ser, por exemplo, um fragmento Fab, F(ab')₂, Fv, e Fv de cadeia única (scFv). Um fragmento scFv é uma cadeia de polipeptídeo única que inclui as regiões de variável da cadeia tanto pesada como leve do

anticorpo do qual o scFv é derivado. Além disso, diacorpos [Poljak (1994) *Structure* 2 (12):1121-1123; Hudson *et al.* (1999) *J. Immunol. Methods* 23 (1-2):177-189] e intracorpos [Huston *et al.* (2001) *Hum. Antibodies* 10 (3-4):127-142; Wheeler *et al.* (2003) *Mol. Ther.* 8 (3):355-366; Stocks (2004) *Drug Discov. Today* 9 (22): 960-966] podem ser usados nos métodos da invenção.

As proteínas-alvo também podem ser juntadas a um ou mais de um polímero, um veículo, um adjuvante, uma imunotoxina, ou uma porção detectável (por exemplo, fluorescente, luminescente ou radioativa). Por exemplo, uma proteína-alvo pode ser juntada ao polietilenoglicol, que pode ser usado para aumentar o peso molecular de pequenas proteínas e/ou aumenta o tempo de residência de circulação.

Em algumas modalidades, a molécula-alvo pode ser, ou conter, dolicol.

Células Geneticamente Engenheiradas

As células geneticamente engenheiradas descritas neste pedido podem ser usadas para produzir moléculas-alvo que contêm N-glicanos similares de mamíferos ou moléculas-alvo que contêm intermediários dentro da via de glicosilação mamífera. Por exemplo, como descrito neste pedido, os ácidos nucleicos que codificam uma ou mais enzimas podem ser introduzidos em uma célula fúngica tal que a célula produza o N-glicano desejado (por exemplo, GlcNAcMan₅GlcNAc₂, GlcNAcMan₃GlcNAc₂, GlcNAc₂Man₃GlcNAc₂, GalGlcNAcMan₃GlcNAc₂ ou Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos). Dessa forma, em alguma das modalidades descritas neste pedido, uma célula fúngica pode conter um ácido nucleico que codifica uma enzima, ou um ácido nucleico pode codificar múltiplas enzimas. Cada tal ácido nucleico também pode conter uma sequência de direcionamento como discutido abaixo. Além disso, um ácido nucleico que codifica uma molécula-alvo também pode ser introduzido na célula fúngica tal que a molécula-alvo seja produzida e modificada para conter o N-glicano desejado (por exemplo, GlcNAcMan₅GlcNAc₂, GlcNAcMan₃GlcNAc₂, GlcNAc₂Man₃GlcNAc₂, GalGlcNAcMan₃GlcNAc₂ ou Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos).

Os termos "ácido nucleico" e "polinucleotídeo" são usados inter-

cambiavelmente neste pedido, e referem-se tanto a RNA como a DNA, incluindo cDNA, DNA genômico, DNA sintético, e DNA (ou RNA) contendo análogos de ácidos nucleicos. Os ácidos nucleicos podem ter qualquer estrutura tridimensional. Um ácido nucleico pode ser de fita dupla ou de fita simples

5 (isto é, uma fita sentido ou uma fita antissentido). Exemplos não limitantes de ácidos nucleicos incluem genes, fragmentos gênicos, exons, ítrons, RNA mensageiro (mRNA), RNA de transferência, RNA ribossômico, siRNA, micro-RNA, ribozimas, cDNA, polinucleotídeos recombinantes, polinucleotídeos ramificados, plasmídeos, vetores, DNA isolado de qualquer sequência, RNA isolado de qualquer sequência, sondas de ácidos nucleicos, e iniciadores, bem como análogos de ácidos nucleicos. "Polipeptídeo" e "proteína" são usados intercambiavelmente neste pedido e significam qualquer cadeia ligada ao peptídeo de aminoácidos, apesar de modificação de comprimento ou pós-tradução.

10 15 Um "ácido nucleico isolado" refere-se a um ácido nucleico que é separado de outras moléculas de ácidos nucleicos que estão presentes em um genoma que ocorre naturalmente, incluindo ácidos nucleicos que normalmente flanqueiam um ou ambos os lados do ácido nucleico em um genoma que ocorre naturalmente (por exemplo, um genoma de levedura). O

20 25 termo "isolado" como usado neste pedido com respeito a ácidos nucleicos também inclui qualquer sequência de ácidos nucleicos que não ocorre naturalmente, uma vez que tais sequências que não ocorrem naturalmente não são encontradas na natureza e não têm sequências imediatamente contíguas em um genoma que ocorre naturalmente.

Um ácido nucleico isolado pode ser, por exemplo, uma molécula de DNA, fornecido de uma das sequências de ácidos nucleicos normalmente encontradas imediatamente flanqueando aquela molécula de DNA em um genoma que ocorre naturalmente é removido ou ausente. Dessa forma, um ácido nucleico isolado inclui, sem limitação, uma molécula de DNA que existe como uma molécula separada (por exemplo, um ácido nucleico quimicamente sintetizado, ou um fragmento de cDNA ou DNA genômico produzido por PCR ou tratamento com endonuclease de restrição) independente de

outras sequências bem como DNA que é incorporado em um vetor, um plasmídeo de replicação autônoma, um vírus (por exemplo, qualquer paramixovírus, retrovírus, lentivírus, adenovírus, ou vírus de herpes), ou em DNA genômico de um procarioto ou eucarioto. Além disso, um ácido nucleico isolado pode incluir um ácido nucleico engenheirado, tal como uma molécula de DNA que é parte de um híbrido ou fusão de ácido nucleico. Um ácido nucleico existente entre centenas a milhões de outros ácidos nucleicos, por exemplo, dentro de bibliotecas de cDNA ou bibliotecas genômicas, ou fatias de gel que contêm um digestão de restrição de DNA genômico, não se considera 5 um ácido nucleico isolado.

10

O termo "exógeno" como usado neste pedido com referência ao ácido nucleico e uma célula hospedeira particular refere-se a qualquer ácido nucleico que não ocorre (e não pode ser obtido de) naquela célula particular como encontrado na natureza. Dessa forma, um ácido nucleico que não ocorre naturalmente é considerado ser exógeno a uma célula hospedeira uma vez introduzido na célula hospedeira. É importante observar que ácidos nucleicos que não ocorrem naturalmente podem conter subsequências de ácidos nucleicos ou fragmentos de sequências de ácidos nucleicos que são encontradas na natureza contanto que o ácido nucleico no conjunto não existe na natureza. Por exemplo, uma molécula de ácidos nucleicos que contém uma sequência de DNA genômico dentro de um vetor de expressão é o ácido nucleico que não ocorre naturalmente, e dessa forma, é exógena a uma célula hospedeira uma vez introduzida na célula hospedeira, uma vez que aquela molécula de ácidos nucleicos no conjunto (DNA genômico mais 15 DNA de vetor) não existe na natureza. Dessa forma, qualquer vetor, plasmídeo de replicação autônoma, ou vírus (por exemplo, retrovírus, adenovírus ou vírus de herpes) que no conjunto não existem na natureza são considerados ácidos nucleicos que não ocorrem naturalmente. Segue-se que fragmentos de DNA genômico produzidos por PCR ou tratamento por endonuclease de restrição bem como cDNAs são considerados ser ácido nucleico que não ocorrem naturalmente uma vez que existem como moléculas separadas não encontradas na natureza. Segue-se também que qualquer ácido 20

25

30

nucleico que contém uma sequência promotora e sequência que codifica o polipeptídeo (por exemplo, cDNA ou DNA genômico) em um arranjo não encontrado na natureza é ácido nucleico que não ocorre naturalmente. Um ácido nucleico de ocorrência natural pode ser exógeno a uma célula particular.

- 5 Por exemplo, um cromossomo inteiro isolado de uma célula da levedura x é um ácido nucleico exógeno com respeito a uma célula da levedura y uma vez que aquele cromossomo é introduzido em uma célula de levedura.

Células adequadas para engenharia genética incluem, por exemplo, células fúngicas (por exemplo, *Yarrowia lipolytica* ou qualquer outra 10 célula de levedura dimórfica relacionada descrita neste pedido), células vegetais, ou células animais. As células podem ser células primárias, células imortalizadas, ou células transformadas. As células podem ser aqueles em um animal, por exemplo, um mamífero não humano. Tais células, antes da engenharia genética como especificadas neste pedido, podem ser obtidas 15 de uma variedade de fontes comerciais e instalações de recursos de pesquisa, tais como, por exemplo, American Type Culture Collection (Rockville, MD).

A engenharia genética de uma célula pode incluir modificações genéticas, tais como: (i) deleção de um gene endógeno que codifica uma 20 proteína tendo atividade de N-glicosilação; (ii) introdução de um ácido nucleico recombinante que codifica uma forma mutante de uma proteína (por exemplo, proteína endógena ou exógena) tendo atividade de N-glicosilação (isto é, expressando uma proteína mutante tendo uma atividade de N-glicosilação); (iii) introdução ou expressão de uma molécula de RNA que 25 interfere na expressão funcional de uma proteína tendo atividade de N-glicosilação; (iv) introdução de um ácido nucleico recombinante que codifica uma proteína selvagem (por exemplo, endógeno ou exógeno) tendo atividade de N-glicosilação (isto é, expressando uma proteína tendo uma atividade de N-glicosilação); ou (v) alteração de elementos promotores ou potencializadores de um ou mais genes endógenos que codificam proteínas tendo 30 atividade de N-glicosilação, dessa forma, para alterar a expressão das suas proteínas codificadas. As moléculas de RNA incluem, por exemplo, pequeno

RNA de interferência (siRNA), RNA de grampo curto (shRNA), RNA antisentido ou microRNA (miRNA). Entende-se que o item (ii) inclui, por exemplo, a substituição de um gene endógeno por um gene que codifica uma proteína tendo maior atividade de N-glicosilação quanto ao gene endógeno assim substituído. A engenharia genética também inclui a alteração de um gene endógeno que codifica uma proteína tendo uma atividade de N-glicosilação para produzir uma proteína tendo adições (por exemplo, uma sequência heteróloga), deleções, ou substituições (por exemplo, mutações, tais como mutações pontuais; mutações conservativas ou não conservativas).

5 Mutações podem ser introduzidas especificamente (por exemplo, mutagênese sítio-dirigida ou recombinação homóloga) ou podem ser introduzidas randomicamente (por exemplo, as células podem ser quimicamente mutagenized como descrito em, por exemplo, Newman e Ferro-Novick (1987) *J. Cell Biol.* 105 (4):1587.

10

15 As modificações genéticas descritas neste pedido podem resultar em um ou mais de (i) um aumento em uma ou mais atividades de N-glicosilação na célula geneticamente modificada, (ii) uma redução em uma ou mais atividades de N-glicosilação na célula geneticamente modificada, (iii) uma modificação na localização ou distribuição intracelular de uma ou

20 mais atividades de N-glicosilação na célula geneticamente modificada, ou (iv) uma modificação na proporção de uma ou mais atividades de N-glicosilação na célula geneticamente modificada. Entende-se que um aumento na quantidade de uma atividade de N-glicosilação pode ser devido à superexpressão de uma ou mais proteínas tendo atividade de N-glicosilação,

25 um aumento no número de cópias de um gene endógeno (por exemplo, duplicação gênica), ou uma alteração no promotor ou potencializador de um gene endógeno que estimula um aumento na expressão da proteína codificada pelo gene. Uma redução em uma ou mais atividades de N-glicosilação pode ser devido à superexpressão de uma forma mutante (por exemplo, uma forma negativa dominante) de uma ou mais proteínas tendo atividades de N-glicosilação alteradas, introdução ou expressão de uma ou mais moléculas de RNA de interferência que reduzem a expressão de uma ou mais

30

proteínas tendo uma atividade de N-glicosilação, ou deleção de um ou mais genes endógenos que codificam uma proteína tendo atividade de N-glicosilação.

Os métodos de deleção ou perturbação de um ou mais genes endógenos são descritos nos Exemplos acompanhantes. Por exemplo, para perturbar um gene por recombinação homóloga, um vetor de "substituição gênica" pode ser construído de tal modo para incluir um gene com marcador selecionável. O gene com marcador selecionável pode ser operacionalmente ligado, tanto na extremidade 5' como na 3', a porções do gene de comprimento suficiente para mediar a recombinação homóloga. O marcador selecionável pode ser um de qualquer número de genes que complementam a auxotrofia da célula hospedeira ou fornecem resistência a antibióticos, incluindo genes URA3, LEU2 e HIS3. Outros marcadores selecionáveis adequados incluem o gene CAT, que confere a resistência a cloranfenicol a células de levedura, ou gene lacZ, que resulta em colônias azuis devido à expressão de β-galactosidase. Os fragmentos de DNA linearizados do vetor de substituição gênica então são introduzidos nas células usando métodos bem conhecidos na técnica (vide abaixo). A integração dos fragmentos lineares no genoma e a perturbação do gene podem ser determinadas baseada no marcador de seleção e podem ser verificadas, por exemplo, por análise de Southern blot.

Como detalhado nos exemplos acompanhantes, subsequentes ao seu uso na seleção, um marcador selecionável pode ser removido do genoma da célula hospedeira, por exemplo, por sistemas Cre-loxP (vide abaixo). Este processo da remoção de marcador é referido como "cura" ao longo dos Exemplos.

Alternativamente, um vetor de substituição gênica pode ser construído de tal modo para incluir uma porção do gene a ser perturbado, onde a porção é destituída de qualquer sequência promotora gênica endógena e não codifica nenhum, ou fragmento inativo da sequência de codificação do gene. Um "fragmento inativo" é um fragmento do gene que codifica uma proteína tendo, por exemplo, menos de aproximadamente 10% (por

exemplo, menos de aproximadamente 9%, menos de aproximadamente 8%, menos de aproximadamente 7%, menos de aproximadamente 6%, menos de aproximadamente 5%, menos de aproximadamente 4%, menos de aproximadamente 3%, menos de aproximadamente 2%, menos de aproximadamente 1% ou 0%) da atividade da proteína produzida da sequência de codificação completa do gene. Tal porção do gene é inserida em um vetor de tal modo que nenhuma sequência promotora conhecida seja operacionalmente ligada à sequência gênica, mas que um códon de parada e uma sequência de terminação de transcrição sejam operacionalmente ligados à porção da sequência gênica. Este vetor pode ser posteriormente linearizado na porção da sequência gênica e transformado em uma célula. Por meio da recombi-
5 nação homóloga única, este vetor linearizado então está integrado na con-
traparte endógena do gene.
10

Os vetores de expressão podem ser autônomos ou integrativos.
15 Um ácido nucleico recombinante pode ser introduzido na célula na forma de um vetor de expressão, tal como um plasmídeo, fago, transpon-
son, cosmídeo ou partícula de vírus. O ácido nucleico recombinante pode ser mantido extracromossomicamente ou pode estar integrado no DNA cro-
mossômico da célula de levedura. Os vetores de expressão podem conter
20 genes marcadores de seleção que codificam proteínas necessárias para a viabilidade celular sob condições selecionadas (por exemplo, URA3, que codifica uma enzima necessária para a biossíntese de uracila ou TRP1, que codifica uma enzima necessária para a biossíntese de triptofano) para permitir detecção e/ou seleção daquelas células transformadas com os ácidos nu-
25 cleicos desejados (ver, por exemplo, Patente U.S. No. 4.704.362). Os vete-
res de expressão também podem incluir uma sequência de replicação autô-
noma (ARS). Por exemplo, Patente U.S. No. 4.837.148 descreve sequências de replicação autônoma que fornecem um meio adequado para manter plasmídeos em *Pichia pastoris*.

30 Vetores integrativos são descritos, por exemplo, na Patente U.S. No. 4.882.279. Vetores integrativos geralmente incluem uma sequência ar-
ranjada em série de pelo menos um primeiro fragmento de DNA inserível,

um gene de marcador selecionável, e um segundo fragmento de DNA inserível. O primeiro e o segundo fragmentos de DNA inserível têm cada um aproximadamente 200 (por exemplo, aproximadamente 250, aproximadamente 300, aproximadamente 350, aproximadamente 400, aproximadamente 450, 5 aproximadamente 500, ou aproximadamente 1000 ou mais) nucleotídeos de comprimento e têm sequências nucleotídicas que são homólogas a porções de DNA genômico das espécies a serem transformadas. Uma sequência nucleotídica que contém um gene de interesse (por exemplo, um gene que codifica uma proteína tendo atividade de N-glicosilação) para a expressão é 10 inserida neste vetor entre o primeiro e segundo fragmentos de DNA inserível ou antes ou após o gene marcador. Os vetores integrativos podem ser linearizados antes da transformação de levedura para facilitar a integração da sequência nucleotídica de interesse no genoma de célula hospedeira.

Um vetor de expressão pode caracterizar um ácido nucleico re- 15 combinante no controle de um promotor de levedura (por exemplo, *Yarrowia lipolytica*, *Arxula adeninivorans*, ou outras espécies de levedura dimórficas relacionadas), que o permite ser expresso na levedura. Promotores de levedura adequados incluem os promotores TEF1, HP4D, GAP, POX2, ADC1, TPI1, ADH2, POX, e Gal10. Ver, por exemplo, Madzak *et al.*, (2000) *J. Mol. 20 Microbiol. Biotechnol.* 2:207-216; Guarente *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79 (23):7410. Promotores adequados adicionais são descritos em, por exemplo, Zhu e Zhang (1999) *Bioinformatics* 15 (7-8):608-611 e na Patente U.S. No. 6.265.185. Onde o vetor de expressão deve ser introduzido 25 em uma célula dos animais, tal como uma célula mamífera, o vetor de expressão pode caracterizar um ácido nucleico recombinante no controle de um promotor de célula animal adequado para a expressão na célula hospedeira de interesse. Exemplos de promotores mamíferos incluem os promotores de SV40 e de citomegalovírus (CMV).

Um promotor pode ser (condicional) constitutivo ou induzível. 30 Entende-se que um promotor constitutivo é promotor cuja expressão é constante sob condições de cultura padrão. Promotores induzíveis são promotores que são responsivos a uma ou mais sugestões de indução. Por exemplo,

um promotor induzível pode ser quimicamente regulado (por exemplo, um promotor cuja atividade transcrecional é regulada pela presença ou ausência de um agente de indução química, tal como um álcool, tetraciclina, um esteróide, um metal, ou outra pequena molécula) ou fisicamente regulado (por exemplo, um promotor cuja atividade transcrecional é regulada pela presença ou ausência de um indutor físico, tal como luz ou temperaturas altas ou baixas). Um promotor induzível também pode ser indiretamente regulado por um ou mais fatores de transcrição que são diretamente regulados por sugestões químicas ou físicas.

A engenharia genética de uma célula também inclui a ativação de um gene endógeno (por exemplo, um gene que codifica uma proteína que tem atividade de N-glicosilação) que está presente na célula hospedeira, mas não é normalmente expresso nas células ou não é expresso em níveis significantes nas células. Por exemplo, uma sequência reguladora (por exemplo, um promotor ou um potencializador gênico) de um gene endógeno pode ser modificada tal que a sequência de codificação operacionalmente ligada exiba expressão aumentada. A recombinação homóloga ou direcionamento podem ser usados para substituir ou inutilizar a região reguladora normalmente associada ao gene com uma sequência reguladora que faz que ao gene seja expresso em níveis mais altos do que evidente na célula correspondente não geneticamente engenheirada, ou faz com que o gene exiba um modelo de regulação ou indução que é diferente daquele evidente na célula correspondente não geneticamente engenheirada. Métodos adequados para introduzir alterações de uma sequência reguladora (por exemplo, um promotor ou potencializador) de um gene são descritos, por exemplo, na Publicação do Pedido de Patente U.S. No. 20030147868.

Entende-se que outras modificações geneticamente engenheiradas também podem ser condicionais. Por exemplo, um gene pode ser condicionalmente deletado usando, por exemplo, uma DNA recombinase sítio-específica, tal como o sistema Cre-loxP (ver, por exemplo, Gossen *et al.* (2002) *Ann. Rev. Genetics* 36:153-173 e Publicação do Pedido de Patente U.S. No. 20060014264).

Um ácido nucleico recombinante pode ser introduzido em uma célula descrita neste pedido usando uma variedade de métodos, tais como a técnica de esferoplastos ou o método de transformação de levedura por cloreto de lítio de célula completa. Outros métodos úteis para a transformação de plasmídeos ou vetores de ácidos nucleicos lineares em células são descritos, por exemplo, na Patente U.S. No. 4.929.555; Hinnen *et al.* (1978) *Proc. Nat. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163; Patente U.S. No. 4.879.231; e Sreekrishna *et al.* (1987) *Gene* 59:115. Eletroporação e procedimentos de transformação de célula completa com PEG1000 também podem ser usados, como descrito por Cregg e Russel, *Methods in Molecular Biology: Pichia Protocols*, Chapter 3, Humana Press, Totowa, N.J., pp. 27-39 (1998). A transfecção de células animais pode caracterizar, por exemplo, a introdução de um vetor nas células usando fosfato de cálcio, eletroporação, choque térmico, lipossomas ou reagentes de transfecção, tais como FUGENE® ou LIPOFECTAMINE®, ou contatando vetores ácidos nucleicos nus com células em solução (ver, por exemplo, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual Second Edition* vol. 1, 2 e 3. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, EUA, Nov. 1989.

Células de levedura transformadas podem ser selecionadas usando técnicas apropriadas incluindo, mas não limitadas a, cultura de células auxotróficas após transformação na ausência do produto bioquímico requerido (devido à auxotrofia da célula), seleção e detecção de um novo fenótipo, ou cultura na presença de um antibiótico que é tóxico à levedura na ausência de um gene de resistência contido nos transformantes. Os transformantes também podem ser selecionados e/ou verificados pela integração do cassete de expressão no genoma, que pode ser avaliado, por exemplo, por Southern blot ou análise por PCR.

Antes da introdução dos vetores em uma célula alvo de interesse, os vetores podem ser cultivados (por exemplo, amplificados) em células bacterianas, tais como *Escherichia coli* (*E. coli*). O DNA do vetor pode ser isolado de células bacterianas por qualquer um dos métodos conhecidos na

técnica que resultam na purificação de DNA do vetor do ambiente bacteriano. O DNA do vetor purificado pode ser extraído extensivamente com fenol, clorofórmio e éter, para assegurar que nenhuma das proteínas de *E. coli* está presente na preparação de DNA de plasmídeo, uma vez que estas proteínas podem ser tóxicas a células mamíferas.

Engenharia genética, como descrita neste pedido, pode ser usada para expressar (por exemplo, superexpressar), introduzir modificações ou deletar qualquer número de genes que codificam proteínas tendo atividade de N-glicosilação. Tais proteínas incluem, por exemplo, OCH1, ALG3, α-1,3-glicosiltransferase, GnT I, manosidase II, GnT II, Glicosidase II, ou Gal T. Os genes que codificam proteínas tendo atividade de N-glicosilação podem ser de qualquer espécie contendo tais genes. Espécies fúngicas exemplares das quais os genes que codificam proteínas tendo atividade de N-glicosilação podem ser obtidos incluem, sem limitação, *Pichia anomala*, *Pichia bovis*, *Pichia canadensis*, *Pichia carsonii*, *Pichia farinose*, *Pichia fermentans*, *Pichia fluxuum*, *Pichia membranaefaciens*, *Pichia membranaefaciens*, *Candida valida*, *Candida albicans*, *Candida ascalaphidarum*, *Candida amphixiae*, *Candida Antarctica*, *Candida atlantica*, *Candida atmosphaerica*, *Candida blattae*, *Candida carpophila*, *Candida cerambycidarum*, *Candida chauliodes*, *Candida corydalis*, *Candida dosseyi*, *Candida dubliniensis*, *Candida ergatensis*, *Candida fructus*, *Candida glabrata*, *Candida fermentati*, *Candida guilliermondii*, *Candida haemulonii*, *Candida insectamens*, *Candida insectorum*, *Candida intermedia*, *Candida jeffresii*, *Candida kefyr*, *Candida krusei*, *Candida lusitaniae*, *Candida lyxosopha*, *Candida maltosa*, *Candida membranifaciens*, *Candida milleri*, *Candida oleophila*, *Candida oregonensis*, *Candida parapsilosis*, *Candida quercitrusa*, *Candida shehatea*, *Candida temnochilae*, *Candida tenuis*, *Candida tropicalis*, *Candida tuchiya*, *Candida sinolaborantium*, *Candida sojae*, *Candida viswanathii*, *Candida utilis*, *Pichia membranaefaciens*, *Pichia silvestris*, *Pichia membranaefaciens*, *Pichia chodati*, *Pichia membranaefaciens*, *Pichia membranaefaciens*, *Pichia menbranaefaciens*, *Pichia minuscula*, *Pichia pastoris*, *Pichia pseudopolymorpha*, *Pichia quercuum*, *Pichia robertsii*, *Pichia saitoi*, *Pichia silvestrisi*, *Pichia strasburgensis*, *Pichia terricola*, *Pichia vanriji*, *Pseu-*

dozyma Antarctica, *Rhodosporidium toruloides*, *Rhodotorula glutinis*, *Saccharomyces bayanus*, *Saccharomyces bayanus*, *Saccharomyces momdshuricus*, *Saccharomyces uvarum*, *Saccharomyces bayanus*, *Saccharomyces cerevisiae*, *Saccharomyces bisporus*, *Saccharomyces chevalieri*, *Saccharomyces delbrueckii*, *Saccharomyces exiguous*, *Saccharomyces fermentati*, *Saccharomyces fragilis*, *Saccharomyces marxianus*, *Saccharomyces mellis*, *Saccharomyces rosei*, *Saccharomyces rouxii*, *Saccharomyces uvarum*, *Saccharomyces willianus*, *Saccharomycodes ludwigii*, *Saccharomycopsis capsularis*, *Saccharomycopsis fibuligera*, *Saccharomycopsis fibuligera*, *Endomyces hordei*, *Endomycopsis fibuligera*. *Saturnispora saitoi*, *Schizosaccharomyces octosporus*, *Schizosaccharomyces pombe*, *Schwanniomyces occidentalis*, *Torulaspora delbrueckii*, *Torulaspora delbrueckii*, *Saccharomyces dairensis*, *Torulaspora delbrueckii*, *Torulaspora fermentati*, *Saccharomyces fermentati*, *Torulaspora delbrueckii*, *Torulaspora rosei*, *Saccharomyces rosei*, *Torulaspora delbrueckii*, *Saccharomyces rosei*, *Torulaspora delbrueckii*, *Saccharomyces delbrueckii*, *Torulaspora delbrueckii*, *Saccharomyces delbrueckii*, *Zygosaccharomyces mongolicus*, *Dorulaspora globosa*, *Debaryomyces globosus*, *Torulopsis globosa*, *Trichosporon cutaneum*, *Trigonopsis variabilis*, *Williopsis californica*, *Williopsis saturnus*, *Zygosaccharomyces bisporus*, *Zygosaccharomyces bisporus*, *Debaryomyces disporua*. *Saccharomyces bisporas*, *Zygosaccharomyces bisporus*, *Saccharomyces bisporus*, *Zygosaccharomyces mellis*, *Zygosaccharomyces priorianus*, *Zygosaccharomyces rouxiim*, *Zygosaccharomyces rouxii*, *Zygosaccharomyces barkeri*, *Saccharomyces rouxii*, *Zygosaccharomyces rouxii*, *Zygosaccharomyces major*, *Saccharomyces rouxii*, *Pichia anomala*, *Pichia bovis*, *Pichia Canadensis*, *Pichia carsonii*, *Pichia farinose*, *Pichia fermentans*, *Pichia fluxuum*, *Pichia membranaefaciens*, *Pichia pseudopolymorpha*, *Pichia quercuum*, *Pichia robertsii*, *Pseudozyma Antarctica*, *Rhodosporidium toruloides*, *Rhodosporidium toruloides*, *Rhodotorula glutinis*, *Saccharomyces bayanus*, *Saccharomyces bayanus*, *Saccharomyces bisporus*, *Saccharomyces cerevisiae*, *Saccharomyces chevalieri*, *Saccharomyces delbrueckii*, *Saccharomyces fermentati*, *Saccharomyces fragilis*, *Saccharomycodes ludwigii*, *Schizosaccharomyces pombe*, *Schwanniomyces*

occidentalis, *Torulaspora delbrueckii*, *Torulaspora globosa*, *Trigonopsis variabilis*, *Williopsis californica*, *Williopsis saturnus*, *Zygosaccharomyces bisporus*, *Zygosaccharomyces mellis*, *Zygosaccharomyces rouxii*, ou qualquer outro fungo (por exemplo, levedura) conhecido na técnica ou descrito neste pedido. Eucariotos inferiores exemplares também incluem várias espécies de *Aspergillus* incluindo, mas não limitados a, *Aspergillus caesiellus*, *Aspergillus candidus*, *Aspergillus carneus*, *Aspergillus clavatus*, *Aspergillus deflectus*, *Aspergillus flavus*, *Aspergillus fumigatus*, *Aspergillus glaucus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus ochraceus*, *Aspergillus oryzae*, *Aspergillus parasiticus*, *Aspergillus penicilloides*, *Aspergillus restrictus*, *Aspergillus sojae*, *Aspergillus sydowi*, *Aspergillus tamari*, *Aspergillus terreus*, *Aspergillus ustus*, ou *Aspergillus versicolor*. Gêneros protozoários exemplares dos quais os genes que codificam proteínas tendo atividade de N-glicosilação podem ser obtidos incluem, sem limitação, *Blastocritidia*, *Cryptidida*, *Endotrypanum*, *Herpetomonas*, *Leishmania*, *Leptomonas*, *Phytomonas*, *Trypanosoma* (por exemplo, *T. brucei*, *T. gambiense*, *T. rhodesiense* e *T. cruzi*), e *Wallaceina*. Por exemplo, o gene que codifica GnT I pode ser obtido do ser humano (No. Suíço de Acesso de Proteína P26572), rato, *Arabidopsis*, camundongo, ou *Drosophila*; o gene que codifica Gnt II pode ser obtido do ser humano, rato (No. Suíço de Acesso de Proteína Q09326), *Arabidopsis*, ou camundongo; o gene que codifica Man II pode ser obtido de ser humano, rato, *Arabidopsis*, camundongo, *Drosophila* (No. Suíço de Acesso de Proteína Q24451); e o gene que codifica GalT pode ser obtido de ser humano (No. Suíço de Acesso de Proteína P15291), rato, camundongo ou bovino.

Em algumas modalidades, uma célula geneticamente engenheirada sem o gene OCH1 (No de Acesso GenBank: AJ563920) ou produto gênico (mRNA ou proteína) do mesmo. Em algumas modalidades, uma célula geneticamente engenheirada sem o gene ALG3 (Números de Acesso Genbank®: XM_503488, Genolevures Ref: YALI0E03190g) ou produto gênico (mRNA ou proteína) do mesmo. Em algumas modalidades, uma célula geneticamente engenheirada expressa (por exemplo, superexpressa) α-1,3-

glicosiltransferase (por exemplo, ALG6, Números de Acesso Genbank®: XM_502922, Genolevures Ref: YALI0D17028g) proteína. Em algumas modalidades, uma célula geneticamente engenheirada expressa a proteína α-1,2-manosidase (por exemplo, No. de Acesso Genbank®:AF212153). Em 5 algumas modalidades, uma célula geneticamente engenheirada expressa uma proteína GlcNAc-transferase I (por exemplo, No. Suíço de Acesso de Prot. P26572). Em algumas modalidades, uma célula geneticamente engenheirada expressa uma proteína manosidase II ou domínio catalítico da mesma (por exemplo, No. Suíço de Acesso de Prot. Q24451). Em algumas 10 modalidades, uma célula geneticamente engenheirada expressa uma proteína galactosiltransferase I ou domínio catalítico da mesma (por exemplo, No. Suíço de Acesso de Prot. P15291). Em algumas modalidades, a célula geneticamente engenheirada expressa uma proteína GlcNAc-transferase II ou domínio catalítico da mesma (por exemplo, No. Suíço de Acesso de Prot. 15 Q09326). Em algumas modalidades, a célula geneticamente engenheirada expressa uma subunidade alfa ou beta (ou tanto subunidade alfa como beta) de uma Glicosidase II, tal como Glicosidase II de *Yarrowia lipolytica*, *Trypanosoma brucei* ou *Aspergillus niger*. Uma célula geneticamente engenheirada pode ter qualquer combinação destas modificações.

20 Por exemplo, em algumas modalidades, uma célula geneticamente engenheirada pode necessitar do gene OCH1 e expressar α-1,2-manosidase, GlcNAc-transferase I, manosidase II, e uma galactosiltransferase I. Em alguma modalidade, uma célula geneticamente engenheirada pode faltar o gene ALG3, e expressar α-1,2-manosidase, GlcNAc-transferase I, 25 GlcNAc-transferase I, e uma galactosiltransferase I. Tal célula geneticamente engenheirada ainda pode expressar α-1,3-glicosiltransferase e/ou expressar subunidades alfa e beta de uma Glicosidase II e/ou sem o gene OCH1.

Uma de mais de tais proteínas pode ser proteínas de fusão que 30 contêm uma sequência de direcionamento heteróloga. Por exemplo, a α-1,2-manosidase pode ter uma sequência de aminoácidos de retenção de retículo endoplasmático (ER) HDEL (vide Exemplos). Entende-se que qualquer proteína tendo atividade de N-glicosilação pode ser engenheirada em uma pro-

teína de fusão compreendendo uma sequência HDEL. Outras proteínas podem ter sequências heterólogas que direcionam a proteína para o aparelho de Golgi. Por exemplo, os 100 primeiros aminoácidos N-terminais codificados pelo gene Kre2p de levedura, os 36 primeiros aminoácidos N-terminais (No. de Acesso Suíço de Prot. P38069) codificados pelo gene Mnn2 de *S. cerevisiae* ou os 46 primeiros aminoácidos N-terminais codificados pelo gene Mnn2p de *S. cerevisiae* podem ser usados para direcionar proteínas para o Golgi. Como tal, os ácidos nucleicos que codificam uma proteína a ser expressa em uma célula fúngica podem incluir uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a proteína codificada para um compartimento intracelular. Por exemplo, a α-1,2-manosidase pode ser direcionada ao ER, enquanto GnT I, GnTII, manosidase e Gal T podem ser direcionadas para o Golgi.

Em modalidades onde uma proteína tendo atividade de N-glicosilação é derivada de uma célula que é de um tipo diferente (por exemplo, de uma espécie diferente) daquela célula na qual a proteína deve ser expressa, um ácido nucleico que codifica a proteína pode ser códon-otimizado para a expressão na célula particular de interesse. Por exemplo, um ácido nucleico que codifica uma proteína tendo N-glicosilação de *Trypanosoma brucei* pode ser códon-otimizado para a expressão em uma célula de levedura, tal como *Yarrowia lipolytica*. Tal otimização do códon pode ser útil para aumentar a expressão da proteína na célula de interesse. Os métodos de códon-otimização de um ácido nucleico que codifica uma proteína são conhecidos na técnica e descritos, por exemplo, em Gao *et al.* (Biotechnol. Prog. (2004) 20 (2): 443 - 448), Kotula *et al.* (Nat. Biotechn. (1991) 9, 1386 - 1389), e Bennetzen *et al.* (J. Biol. Chem. (1982) 257 (6):2036-3031). A tabela 1 mostra o uso de códon de *Yarrowia lipolytica*. Os dados foram derivados de 2.945.919 códons presentes em 5.967 sequências de codificação. Os conteúdos da Tabela 1 foram obtidos de um Banco de Dados de Uso de Códon, que pode ser encontrado na rede de alcance mundial em kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=284591.

TABELA 1

Tabela de Uso de Códon de *Yarrowia lipolytica*

UUU 15.9(46804)	CU 21.8(64161)	AU 6.8(20043)	GU 6.1(17849)
UUC 23.0(67672)	CC 20.6(60695)	AC 23.1(68146)	GC 6.1(17903)
UUA 1.8(5280)	CA 7.8(22845)	AA 0.8(2494)	GA 0.4(1148)
UUG 10.4(30576)	CG 15.4(45255)	AG 0.8(2325)	GG 12.1(35555)
CUU 13.2(38890)	CU 17.4(51329)	AU 9.6(28191)	GU 6.0(17622)
CUC 22.6(66461)	CC 23.3(68633)	AC 14.4(42490)	GC 4.4(12915)
CUA 5.3(15548)	CA 6.9(20234)	AA 9.8(28769)	GA 21.7(63881)
CUG 33.5(98823)	CG 6.8(20042)	AG 32.1(94609)	GG 7.7(22606)
AUU 22.4(66134)	CU 16.2(47842)	AU 8.9(26184)	GU 6.7(19861)
AUC 24.4(71810)	CC 25.6(75551)	AC 31.3(92161)	GC 9.8(28855)
AUA 2.2(6342)	CA 10.5(30844)	AA 12.4(36672)	GA 8.4(24674)
AUG 22.6(66620)	CG 8.5(25021)	AG 46.5(136914)	GG 2.4(7208)
GUU 15.8(46530)	CU 25.5(75193)	AU 21.5(63259)	GU 16.6(48902)
GUC 21.5(63401)	CC 32.7(96219)	AC 38.3(112759)	GC 21.8(64272)
GUA 4.0(11840)	CA 11.2(32999)	AA 18.8(55382)	GA 20.9(61597)
GUG 25.7(75765)	CG 8.9(26190)	AG 46.2(136241)	GG 4.4(12883)

Os campos da tabela são mostrados como [tripletos] [frequência: por mil] ([número]).

- 5 Em algumas modalidades, as proteínas humanas podem ser introduzidas na célula e uma ou mais proteínas de levedura endógenas tendo atividade de N-glicosilação podem ser suprimidas (por exemplo, deletadas ou mutadas). As técnicas para "humanização" de uma via de glicosilação fúngica são descritas, por exemplo, em Choi *et al.* (2003) *Proc. Natl. Acad. Sci. USA* 100 (9):5022-5027; Vervecken *et al.* (2004) *Appl. Environ. Microb.* 70 (5):2639-2646; e Gerngross (2004) *Nature Biotech.* 22 (11):1410-1414.

Onde a engenharia genética envolve, por exemplo, modificações na expressão de uma proteína ou expressão de uma proteína exógena (incluindo uma forma mutante de uma proteína endógena), uma variedade de 15 técnicas pode ser usada para determinar se as células geneticamente engenheiradas expressam a proteína. Por exemplo, a presença do mRNA que codifica a proteína ou a própria proteína pode ser detectada usando, por exemplo, análise de Northern Blot ou RT-PCR ou análise de Western Blot, respectivamente. A localização intracelular de uma proteína tendo atividade 20 de N-glicosilação pode ser analisada usando uma variedade de técnicas,

incluindo o fracionamento subcelular e imunofluorescência.

Os métodos para detecção da glicosilação de uma molécula-alvo incluem DNA assistido pelo sequenciador (DSA), eletroforese de carboidrato assistida por fluoróforo (FACE) ou espectrometria de massa de tempo de voo com dessorção/ionização por laser aumentada por superfície (SELDI-TOF MS). Por exemplo, uma análise pode utilizar a DSA-FACE na qual, por exemplo, as glicoproteínas são desnaturadas seguidas pela imobilização, por exemplo, em uma membrana. As glicoproteínas então podem ser reduzidas com um agente redutor adequado, tal como ditiotreitol (DTT) ou λ -mercaptoetanol. Grupos sulfidrila das proteínas podem ser carboxilados usando um ácido, tal como ácido iodoacético. Após, N-glicanos podem ser liberados da proteína usando uma enzima, tal como N-glicosidase F. N-glicanos, opcionalmente, podem ser reconstituídos e derivados por aminação redutiva. N-glicanos derivados então podem ser concentrados. A instrumentação adequada para a análise de N-glicano inclui, por exemplo, o sequenciador de DNA ABI PRISM® 377 (Applied Biosystems). A análise de dados pode ser realizada usando, por exemplo, o programa GENSCAN® 3.1 (Applied Biosystems). Opcionalmente, manoproteínas isoladas podem ser ainda tratadas com uma ou mais enzimas para confirmar seu status de N-glicano. Métodos adicionais para análise de N-glicano incluem, por exemplo, espectrometria de massa (por exemplo, MALDI-TOF-MS), cromatografia líquida de alta pressão (HPLC) em fase normal, fase reversa e cromatografia de troca iônica (por exemplo, com detecção amperométrica pulsada quando glicanos não são marcados e com fluorescência ou absorvância de UV se os glicanos forem apropriadamente marcados). Ver também Callewaert *et al.* (2001) *Glycobiology* 11 (4):275-281 e Freire *et al.* (2006) *Bioconjug. Chem.* 17 (2):559-564.

Onde alguma das modificações genéticas da célula geneticamente engenheirada é induzível ou condicional na presença de uma sugestão de indução (por exemplo, uma sugestão química ou física), a célula geneticamente engenheirada pode ser, opcionalmente, cultivada na presença de um agente de indução antes, durante, ou subsequente à introdução do

ácido nucleico. Por exemplo, seguindo à introdução do ácido nucleico que codifica uma proteína alvo, a célula pode ser exposta a um agente de indução química que é capaz de promover a expressão de uma ou mais proteínas tendo atividade de N-glicosilação. Onde múltiplas sugestões de indução 5 induzem a expressão condicional de uma ou mais proteínas tendo atividade de N-glicosilação, uma célula pode ser contatada com múltiplos agentes de indução.

Moléculas alvo modificadas para incluir o N-glicano desejado podem ser isoladas da célula geneticamente engenheirada. A molécula-alvo 10 modificada pode ser mantida dentro da célula de levedura e liberada na lise celular ou a molécula-alvo modificada pode ser secretada em meio de cultura através de um mecanismo fornecido por uma sequência de codificação (nativa ao ácido nucleico exógeno ou engenheirada no vetor de expressão), que dirige a secreção da molécula da célula. A presença da molécula-alvo 15 modificada no lisado celular ou meio de cultura pode ser verificada por uma variedade de protocolos padrão para detectar a presença da molécula. Por exemplo, onde a molécula-alvo alterada é uma proteína, tais protocolos podem incluir, mas não são limitados a, imunoblotting ou radioimunoprecipitação com um anticorpo específico para a proteína-alvo alterada (ou a própria 20 proteína alvo), ligação a um ligante específico para a proteína-alvo alterada (ou a própria proteína alvo), ou teste para uma atividade enzimática específica da proteína-alvo modificada (ou a própria proteína alvo).

Em algumas modalidades, pelo menos aproximadamente 25% das moléculas-alvo isoladas da célula geneticamente engenheirada contêm 25 o N-glicano desejado. Por exemplo, pelo menos aproximadamente 27%, pelo menos aproximadamente 30%, pelo menos aproximadamente 35%, pelo menos aproximadamente 40%, pelo menos aproximadamente 45%, pelo menos aproximadamente 50%, pelo menos aproximadamente 55%, pelo menos aproximadamente 60%, pelo menos aproximadamente 65%, pelo 30 menos aproximadamente 70%, pelo menos aproximadamente 75%, pelo menos aproximadamente 80%, pelo menos aproximadamente 85%, pelo menos aproximadamente 90%, ou pelo menos aproximadamente 95%, ou

pelo menos aproximadamente 99% das moléculas-alvo isoladas da célula geneticamente engenheirada podem conter o N-glicano desejado.

Em algumas modalidades, nas moléculas-alvo produzidas usando os métodos descritos neste pedido, pelo menos 50% (por exemplo, pelo menos 55, 60, 65, 70, 75, 80, ou 85%) dos N-glicanos na glicoproteína podem ser GlcNAc₂Man₃GlcNAc₂ N-glicanos. A porcentagem de GlcNAc₂Man₃GlcNAc₂ N-glicanos pode ser estimada das áreas de pico em elektroferogramas de DSA-FACE. Ver o Exemplo 13.

Em algumas modalidades, as moléculas-alvo modificadas isoladas podem ser congeladas, liofilizadas, ou imobilizadas e armazenadas sob condições apropriadas, por exemplo, que permitem às moléculas-alvo alteradas conservar a atividade biológica.

Culturas de Células Engenheiradas

Este documento também fornece uma cultura substancialmente pura de qualquer uma das células geneticamente engenheiradas descritas neste pedido. Como usado neste pedido, uma "cultura substancialmente pura" de uma célula geneticamente engenheirada é uma cultura daquela célula em que menos de aproximadamente 40% (isto é, menos do que aproximadamente: 35%; 30%; 25%; 20%; 15%; 10%; 5%; 2%; 1%; 0,5%; 0,25%; 0,1%; 0,01%; 0,001%; 0,0001%; ou até menos) do número total de células viáveis na cultura são células viáveis além da célula geneticamente engenheirada, por exemplo, células bacterianas, fúngicas (incluindo a levedura), de micoplasmas ou de protozoários. O termo "aproximadamente" neste contexto significa que a porcentagem relevante pode ser 15% por cento da porcentagem especificada acima ou abaixo da porcentagem especificada. dessa forma, por exemplo, aproximadamente 20% podem ser 17% a 23%. Tal cultura de células geneticamente engenheiradas inclui as células e um meio de crescimento, armazenamento ou transporte. Meios podem ser líquidos, semissólidos (por exemplo, meios gelatinosos) ou congelados. A cultura inclui as células que crescem no meio líquido ou no semissólido ou são armazenadas ou transportadas em um meio de armazenamento ou de transporte, incluindo um meio congelado de armazenamento ou transporte. As culturas

estão em um vaso de cultura ou vaso de armazenamento ou substrato (por exemplo, uma placa, frasco, ou tubo de cultura ou um frasco ou tubo de armazenamento).

As células geneticamente engenheiradas descritas neste pedido
 5 podem ser armazenadas, por exemplo, como suspensões celulares conge-
 ladas, por exemplo, no tampão contendo um crioprotetor, tais como glicerol
 ou sacarose, como células liofilizadas. Alternativamente, podem ser armaze-
 nadas, por exemplo, como preparações celulares secas obtidas, por exem-
 plo, por secagem em leito fluidizado ou secagem por pulverização, ou qual-
 10 quer outro método de secagem adequado.

Distúrbios Tratáveis por Moléculas de N-glicosilação Alteradas

As moléculas-alvo isoladas modificadas para conter o N-glicano
 desejado podem ser usadas para tratar uma variedade de distúrbios, incluin-
 do distúrbios metabólicos, câncer, e distúrbios inflamatórios.

15 (i) Distúrbios Metabólicos

Um distúrbio metabólico é aquele que afeta a produção da ener-
 gia dentro de células humanas individuais (ou animais). A maioria dos dis-
 túrbios metabólicos são genéticos, embora alguns possam ser "adquiridos"
 em consequência de dieta, toxinas, infecções, etc. Os distúrbios metabólicos
 20 genéticos também são conhecidos como os erros inatos do metabolismo.

Em geral, os distúrbios metabólicos genéticos são causados por defeitos
 genéticos que resultam em ausência ou enzimas impropriamente construí-
 das necessárias para alguma etapa no processo metabólico da célula. As
 maiores classes de distúrbios metabólicos são distúrbios do metabolismo de
 25 carboidratos, distúrbios do metabolismo de aminoácidos, distúrbios do meta-
 bolismo de ácidos orgânicos (acidúrias orgânicas), distúrbios de oxidação de
 ácidos graxos e metabolismo mitocondrial, distúrbios do metabolismo de por-
 firinas, distúrbios de metabolismo de purinas ou pirimidinas, distúrbios de
 distúrbios de metabolismo de esteroides da função mitocondrial, distúrbios
 30 da função peroxissomal, e distúrbios de armazenamento lisossomai (LSDs).

Exemplos de distúrbios metabólicos que podem ser tratados pe-
 la administração de uma ou mais moléculas glicosiladas (ou as composições

farmacêuticas das mesmas) descritos neste pedido podem incluir hemocromatose hereditária, albinismo oculocutâneo, deficiência de proteína C, angi-oedema hereditário tipo I, deficiência de sucrase-isomaltase congênita, Cri-gler-Najjar tipo II, síndrome de Laron, Mieloperoxidase hereditária, hipotiroi-
5 dismo primário, síndrome do QT longo congênita, deficiência de globulina de ligação à tiroxina, hipercolesterolemia familiar, quilomicronemia familiar, abe-ta-lipoproteinemia, níveis baixos de lipoproteína A plasmática, enfisema he-
reditário com injúria hepática, hipotireoidismo congênito, osteogenese imper-
feita, hipofibrinogenemia hereditária, deficiência de alfa-1antitripsina,
10 diabetes nefrogênica insípida, diabetes neurohipofiseal insípida, deficiência de adenosina deaminase, doença de Pelizaeus Merzbacher, doença de von Willebrand tipo IIA, deficiência de fatores combinados V e VIII, displasia es-
pondiloepifiseal tarda, coroideremia, doença de célula I, doença de Batten,
ataxia-telangiectasias, doença renal policística dominante autossômica
15 ADPKD, doença de inclusão do microvillus, esclerose tuberosa, síndrome óculo-cérebro-renal de Lowe, esclerose lateral amiotrófica, síndrome mielo-
displásica, síndrome de linfócito Bare, doença de Tangier, colestase intra-
hepática familiar, adrenoleucodistrofia ligada ao X, síndrome de Scott, sín-
drome de Hermansky-Pudlak tipos 1 e 2, síndrome de Zellweger, condrodis-
20 plasia rizomélica puntiforme, hiperoxalúria primária autossômica recessiva,
síndrome de Mohr-Tranebjaerg, atrofia muscular espinhal e bulbar, discine-
sia ciliar primária (síndrome de Kartagener), gigantismo e acromegalia, ga-
lactorreia, doença de Addison, virilismo de glândula suprarrenal, síndrome de Cushing, cetoacidose, aldosteronismo primário ou secundário, síndrome
25 de Miller Dieker, lissencefalia, doença de neurônio motor, síndrome de U-
sher, síndrome de Wiskott-Aldrich, síndrome de Optiz, doença de Hunting-
ton, pancreatite hereditária, síndrome de antifosfolipídeo, doença de tecido
conectivo de sobreposição, síndrome de Sjögren, síndrome do homem rígi-
do, síndrome de Brugada, síndrome nefrítica congênita de tipo finlandês,
30 síndrome de Dubin-Johnson, hipofosfatose ligada ao X, síndrome de Pendred, hipoglicemia hiperinsulinêmica persistente da infância, esferocitose
hereditária, aceruloplasmidemia, lipofuscinoze ceroide neuronal infantil,

pseudoacondroplasia e epifiseal múltipla, distrofia macular similar a Stargardt, doença de Charcot-Marie-Tooth ligada ao X, retinite autossômica dominante pigmentosa, síndrome de Wolcott-Rallison, doença de Cushing, distrofia muscular do tipo cinturas, mucopolissacaridose tipo IV, amiloidose familiar hereditária finlandesa, doença de Anderson, sarcoma, leucemia mielomonocítica crônica, cardiomiopatia, faciogenital displasia, doença de Torsion, ataxias de Huntington e espinocerebelar, hiper-homositeinemia hereditária, polineuropatia, doença de neurônio motor inferior, retinite pigmentada, poliartrite soronegativa, fibrose pulmonar intersticial, fenômeno de Raynaud, 5 granulomatose de Wegner, proteinúria, CDG-Ia, CDG-Ib, CDG-Ic, CDG-IId, CDG-Ie, CDG-IIf, CDG-IIa, CDG-IIb, CDG-IIc, CDG-IId, síndrome de Ehlers-Danlos, exostoses múltiplas, síndrome de Griscelli (tipo 1 ou tipo 2), ou retardamento mental não específico ligado ao X. Além disso, os distúrbios metabólicos também podem incluir distúrbios de armazenamento lisossomal tal como, mas não limitados a, doença de Fabry, doença de Farber, doença de Gaucher, GM₁-gangliosidose, doença de Tay-Sachs, doença de Sandhoff, doença de ativador de GM₂, doença de Krabbe, leucodistrofia metacromática, doença de Niemann-Pick (tipos A, B, e C), doença de Hurler, doença de Scheie, doença de Hunter, doença de Sanfilippo, doença de Morquio, doença de Maroteaux-Lamy, deficiência de hialuronidase, aspartilglicosaminúria, fucosidose, manosidose, doença de Schindler, sialidose tipo 1, doença de Pompe, Picnodisostose, lipofuscinose ceroide, doença de armazenamento de éster de colesterol, doença de Wolman, deficiência múltipla de sulfatase, galactosialidose, mucolipidose (tipos II, III, e IV), cistinose, distúrbio de armazenamento de ácido siálico, doença de retenção de quilomicrons com síndrome de Marinesco-Sjögren, síndrome de Hermansky-Pudlak, síndrome de Chediak-Higashi, doença de Danon, ou displasia gelefísica.

Os sintomas de um distúrbio metabólico são numerosos e diversos e podem incluir um ou mais de, por exemplo, anemia, fadiga, contusão fácil, plaquetas de sangue baixas, alargamento hepático, alargamento de baço, enfraquecimento esquelético, deficiência pulmonar, infecções (por exemplo, infecções no peito ou pneumonia), deficiência renal, dano cerebral

progressivo, convulsão, meconio extraespesso, tossidela, respiração, produção excessiva de saliva ou mucosa, dificuldade de respiração, dor abdominal, intestino ou intestinos fechados, problemas de fertilidade, pólipos no nariz, baqueteamento de dedo da mão/dedo do pé e pele, dor nas mãos ou pés, angioqueratoma, perspiração reduzida, opacidade corneana e lenticular, cataratas, prolapsos e/ou regurgitação da válvula mitral, cardiomegalia, intolerância à temperatura, dificuldade de marcha, dificuldade de engolir, perda de visão progressiva, perda de audição progressiva, hipotonía, macroglossia, areflexia, lombalgia, apneia do sono, ortopneia, sonolência, lordose, ou escoliose. Entende-se que devido à natureza diversa das proteínas incorretas ou ausentes e os fenótipos de doença resultantes (por exemplo, apresentação sintomática de um distúrbio metabólico), um dado distúrbio apresentará geralmente somente a característica de sintomas àquele particular distúrbio. Por exemplo, um paciente com a doença de Fabry pode apresentar um subconjunto particular dos sintomas supracitados tais como, mas não limitados a, intolerância à temperatura, rotação corneana, dor, bolhas cutâneas, náusea, ou diarreia. Um paciente com síndrome de Gaucher pode apresentar esplenomegalia, cirrose, convulsões, hipertonia, apneia, osteoporose ou descoramento cutâneo.

Além da administração de uma ou mais moléculas descritas neste pedido, um distúrbio metabólico também pode ser tratado por nutrição adequada e vitaminas (por exemplo, terapia de cofator), terapia física, e medicações para dor.

Dependendo da natureza específica de um dado distúrbio metabólico, um paciente pode apresentar estes sintomas em qualquer idade. Em muitos casos, os sintomas podem se apresentar na infância ou na primeira idade adulta. Por exemplo, os sintomas da doença de Fabry podem ser apresentados em uma primeira idade, por exemplo, aos 10 ou 11 anos da idade.

Como usado neste pedido, um indivíduo "em risco de desenvolver um distúrbio metabólico" é um indivíduo que tem uma predisposição para desenvolver um distúrbio, isto é, uma predisposição genética para desenvol-

ver o distúrbio metabólico em consequência de uma mutação em uma enzima, tal como alfa-L-iduronidase, beta-D-galactosidase, beta-glicosidase, beta-hexosaminidase, beta-D-manosidase, alfa-L-fucosidase, arilsulfatase B, arilsulfatase A, alfa-N-acetilgalactosaminidase, aspartylglycosaminidase, iduronato-2-sulfatase, alfa-glicosaminida-N-acetyltransferase, beta-D-glicoronidase, hialuronidase, alfa-L-manosidase, alfa-neurominidase, fosfotransferase, lipase ácida, ceramidase ácida, esfingomielinase, tioesterase, catepsina K, ou lipoproteína lipase. Claramente, os indivíduos "em risco de desenvolver um distúrbio metabólico" não são todos os indivíduos dentro de uma espécie de interesse.

Um indivíduo "suspeito em ter um distúrbio" é aquele que tem um ou mais sintomas de um distúrbio, tais como qualquer um dos descritos neste pedido.

(ii) Câncer

O câncer é uma classe de doenças ou distúrbios caracterizados pela divisão descontrolada de células e a capacidade destas de disseminar-se, pelo crescimento direto no tecido adjacente através de invasão, ou por implante em sítios distantes por metástase (onde as células de câncer são transportadas pela circulação sanguínea ou sistema linfático). O câncer pode afetar pessoas em todas as idades, mas o risco tende a aumentar com a idade. Tipos de cânceres podem incluir, por exemplo, câncer de pulmão, câncer de mama, câncer de cólon, câncer pancreático, câncer renal, câncer de estômago, câncer hepático, câncer ósseo, câncer hematológico, câncer de tecido neural, melanoma, câncer de tireoide, câncer ovariano, câncer testicular, câncer de próstata, câncer cervical, câncer vaginal ou câncer de bexiga.

Como usado neste pedido, um indivíduo "em risco de desenvolver um câncer" é um indivíduo que tem uma predisposição para desenvolver um câncer, isto é, uma predisposição genética para desenvolver o câncer, tal como uma mutação em um gene supressor tumoral (por exemplo, mutação em BRCA1, p53, RB, ou APC) ou foi exposto a condições que podem resultar em câncer. Dessa forma, um indivíduo também pode ser um "em

risco de desenvolver um câncer" quando o indivíduo foi exposto a níveis mutagênicos ou cancerígenos de certos compostos (por exemplo, compostos cancerígenos na fumaça de cigarro, tais como acroleína, arsênico, benzeno, Benz{a}antraceno, Benzo{a}pireno, Polônio-210 (Radon), uretano, ou cloreto de vinila). Além disso, o indivíduo pode estar "em risco de desenvolver um câncer" quando o indivíduo foi exposto, por exemplo, a grandes doses de luz ultravioleta ou raio X, ou exposto (por exemplo, Infectado) a um vírus causador/associado a tumor, tal como papilomavírus, vírus Epstein-Barr, vírus de hepatite B, ou vírus de linfoma da leucemia de célula T humana. Do acima mencionado, será claro que os indivíduos "em risco de desenvolver um câncer" não são todos os indivíduos dentro de uma espécie de interesse.

Um indivíduo "suspeito em ter um câncer" é aquele que tem um ou mais sintomas de um câncer. Os sintomas do câncer são bem conhecidos por aqueles versados na técnica e incluem, sem limitação, agregados na mama, modificações no mamilo, cistos de mama, dor na mama, perda de peso, fraqueza, fadiga excessiva, dificuldade de comer, perda de apetite, tosse crônica, agravamento da falta de ar, tossindo sangue, sangue na urina, sangue nas fezes, náusea, vômito, metástases hepáticas, metástases de pulmão, metástases de osso, plenitude abdominal, inchaço abdominal, fluido em cavidade peritoneal, hemorragia vaginal, constipação, distensão abdominal, perfuração do cólon, peritonite aguda (infecção, febre, dor), dor, vomitando sangue, sudorese pesada, febre, alta pressão sanguínea, anemia, diarreia, icterícia, vertigem, frios, espasmos musculares, metástases do cólon, metástases do pulmão, metástases de bexiga, metástases hepáticas, metástases de osso, metástases renais, e metástases pancreáticas, dificuldade de engolir, e similares. Do acima mencionado será claro que os indivíduos "suspeitos em ter um câncer" não são todos os indivíduos dentro de uma espécie de interesse.

Além da administração de uma ou mais moléculas de N-glicosilação alteradas descritas neste pedido, um câncer também pode ser tratado por agentes quimioterápicos, radiação ionizante, agentes de imunoterapia, ou agentes hipertermoterapia. Agentes quimioterápicos incluem, por

exemplo, cisplatina, carboplatina, procarbazina, mecloretamina, ciclofosfamida, camptotecina, adriamicina, ifosfamida, melfalano, clorambucil, bissulfano, nitrosureia, dactinomicina, daunorrubicina, doxorrubicina, bleomicina, plicomicina, mitomicina, etoposídeo, verampil, podofilotoxina, tamoxifeno, 5 taxol, transplatina, 5-flurouracil, vincristina, vinblastina e metotrexato.

(iii) Distúrbios Inflamatórios

Um "distúrbio inflamatório," como usado neste pedido, refere-se a um processo no qual uma ou mais substâncias (por exemplo, substâncias que naturalmente não ocorrem no indivíduo), através da ação de células brancas do sangue (por exemplo, células B, células T, macrófagos, monócitos ou células dendríticas) impropriamente provocam uma resposta patológica, por exemplo, uma resposta imune patológica. Consequentemente, tais células envolvidas na resposta inflamatória são referidas como "células inflamatórias." A resposta inflamatória impropriamente provocada pode ser aquela onde nenhuma substância estranha (por exemplo, um antígeno, um vírus, uma bactéria, um fungo) está presente dentro ou sobre o indivíduo. A resposta impropriamente provocada pode ser aquela onde um autocomponente (por exemplo, um autoantígeno) é visado (por exemplo, um distúrbio autoimune, tal como esclerose múltipla) pelas células inflamatórias. A resposta impropriamente provocada também pode ser uma resposta que é imprópria em magnitude ou duração, por exemplo, anafilaxia. Dessa forma, a resposta impropriamente direcionada pode ser devido à presença de uma infecção microbiana (por exemplo, viral, bacteriana, ou fúngica). Tipos de distúrbios inflamatórios (por exemplo, doença autoimune) podem incluir, mas não são limitados a, osteoartrite, artrite reumatoide (RA), espondiloartropatias, síndrome de POEMS, doença de Crohn, doença de Castleman multicêntrica, lúpus sistêmico eritematoso (SLE), esclerose múltipla (MS), distrofia muscular (MD), diabetes mellitus dependente de insulina (IDDM), dermatomiosite, poliomiosite, neuropatias inflamatórias, tais como síndrome de Guillain Barre, vasculite, tais como granulomatose de Wegener, poliarterite nodosa, polimialgia reumática, arterite temporal, síndrome de Sjogren, doença de Bechet, síndrome de Churg-Strauss ou arterite de Takayasu. Também

estão incluídos em distúrbios inflamatórios certos tipos de alergias, tais como rinite, sinusite, urticária, urticária, angioedema, dermatite atópica, alergias alimentares (por exemplo, uma alergia à noz), alergias a fármacos (por exemplo, penicilina), alergias a insetos (por exemplo, alergia a uma picada de abelha), ou mastocitose. Os distúrbios inflamatórios também podem incluir colite ulcerativa e asma.

Um indivíduo "em risco de desenvolver um distúrbio inflamatório" refere-se a um indivíduo com um histórico familiar de um ou mais distúrbios inflamatórios (por exemplo, uma predisposição genética para um ou mais distúrbios inflamatórios) ou um exposto a uma ou mais condições indutoras de inflamação. Por exemplo, um indivíduo pode ter sido exposto a um superantígeno viral ou bacteriano tal como, mas não limitado a, enterotoxinas de estafilococo (SEs), uma exotoxina de *Streptococcus pyogenes* (SPE), uma toxina da síndrome de choque tóxico de *Staphylococcus aureus* (TSST-1), uma exotoxina mitogênica de estreptococo (SME) e um superantígeno de estreptococo (SSA). Do acima mencionado será claro que os indivíduos "em risco de desenvolver um distúrbio inflamatório" não são todos os indivíduos dentro de uma espécie de interesse.

Um indivíduo "suspeito em ter um distúrbio inflamatório" é aquele que apresenta um ou mais sintomas de um distúrbio inflamatório. Os sintomas de distúrbios inflamatórios são bem conhecidos na técnica e incluem, mas não são limitados a, vermelhidão, inchaço (por exemplo, articulações inchadas), articulações que estão quentes ao toque, dor nas articulações, endurecimento, perda de função das articulações, febre, frios, fadiga, perda de energia, dores de cabeça, perda de apetite, endurecimento muscular, insônia, coceira, nariz entupido, espirro, tosse, um ou mais sintomas neurológicos, tais como vertigem, convulsão ou dor. Do acima mencionado será claro que os indivíduos "suspeitos em ter um distúrbio inflamatório" não são todos os indivíduos dentro de uma espécie de interesse.

Além da administração de uma ou mais moléculas descritas neste pedido, um distúrbio inflamatório também pode ser tratado pelo fármaco anti-inflamatório não esteroidal (NSAID), um fármaco antirreumático modifi-

cador de doença (DMARD), um modificador de resposta biológica, ou um corticosteroide. Os modificadores de resposta biológica incluem, por exemplo, um agente anti-TNF. Exemplos não limitantes de agentes anti-TNF incluem um receptor de TNF solúvel ou um anticorpo específico para TNF, tal como adalimumabe, infliximabe, ou etanercept.

Os métodos adequados para tratar (por exemplo, impedindo ou melhorando um ou mais sintomas de) algum dos distúrbios descritos neste pedido usando alguma das moléculas de N-glicosilação alteradas (ou composições farmacêuticas disso) são apresentados na seguinte seção.

10 Composições Farmacêuticas e Métodos de Tratamento

Uma molécula-alvo modificada para ter o N-glicano desejado pode ser incorporada em uma composição farmacêutica contendo uma quantidade terapeuticamente eficaz da molécula e um ou mais adjuvantes, excipientes, veículos, e/ou diluentes. Diluentes, veículos e excipientes aceitáveis tipicamente não afetam adversamente a homeostase de um recipiente (por exemplo, equilíbrio de eletrólitos). Veículos aceitáveis incluem sais biocompatíveis, inertes ou bioabsorvíveis, agentes tamponantes, oligo- ou polissacarídeos, polímeros, agentes de melhora de viscosidade, conservantes e similares. Um veículo exemplar é salina fisiológica (NaCl 0,15 M, pH 7,0 a 7,4). Outro veículo exemplar é fosfato de sódio 50 mM, cloreto de sódio 100 mM. Detalhes adicionais de técnicas de formulação e administração de composições farmacêuticas podem ser encontrados, por exemplo, em Remington's Pharmaceutical Sciences (Maack Publishing Co, Easton, Pa). Os compostos ativos suplementares também podem ser incorporados nas composições.

A administração de uma composição farmacêutica contendo moléculas com N-glicanos pode ser sistêmica ou local. As composições farmacêuticas podem ser formuladas tal que sejam adequadas para a administração parenteral e/ou não parenteral. Modalidades de administração específicas incluem administração subcutânea, intravenosa, intramuscular, intraperitoneal, transdérmica, intratecal, oral, retal, bucal, tópica, nasal, oftalmica, intra-articular, intra-arterial, subaracnóidea, brônquica, linfática, vaginal, e in-

trauterina.

A administração pode ser por injeções periódicas de um bolus da composição farmacêutica ou pode ser ininterrupta ou contínua pela administração intravenosa ou intraperitoneal de um reservatório que é externo (por exemplo, uma bolsa IV) ou interno (por exemplo, um implante bioerodível, um órgão bioartificial, ou uma colônia de células de produção de molécula de N-glicosilação alteradas implantadas). Ver, por exemplo, Pat. U.S. Nos. 4.407.957, 5.798.113 e 5.800.828. A administração de uma composição farmacêutica pode ser alcançada usando meios de entrega adequados, tais como: uma bomba (ver, por exemplo, Annals of Pharmacotherapy, 27:912 (1993); Cancer, 41:1270 (1993); Cancer Research, 44:1698 (1984); microencapsulamento (ver, por exemplo, Pat. U.S. Nos. 4.352.883; 4.353.888; e 5.084.350); implantes de polímero de liberação contínua (ver, por exemplo, Sabel, Pat. U.S. No. 4.883.666); macroencapsulamento (ver, por exemplo, U.S. Pat. Nos. 5.284.761, 5.158.881, 4.976.859 e 4.968.733 e pedidos de patente PCT publicadas WO92/19195, WO 95/05452); injeção, subcutaneamente, intravenosamente, intra-arterialmente, intramuscularmente, ou em outro sítio adequado; ou administração oral, em cápsula, líquido, comprimido, pílula, ou formulação de liberação prolongada.

Exemplos de sistemas de entrega parenterais incluem partículas de copolímero de etileno-acetato de vinila, bombas osmóticas, sistemas de infusão implantáveis, entrega de bomba, entrega celular encapsulada, entrega lipossomal, injeção entregue por agulha, injeção sem agulha, nebulizador, aerossol, eletroporação e adesivo transdérmico.

Formulações adequadas para a administração parenteral convenientemente contêm uma preparação aquosa estéril da molécula de N-glicosilação alterada, que preferencialmente é isotônica com o sangue do recipiente (*por exemplo*, solução salina fisiológica). As formulações podem ser apresentadas em forma de dose única ou de multidose.

Formulações adequadas para a administração oral podem ser apresentadas como unidades discretas, tais como cápsulas, discos, comprimidos, ou pastilhas, cada um contendo uma quantidade predeterminada da

molécula de N-glicosilação alterada; ou uma suspensão em um licor aquoso ou um líquido não aquoso, tal como um xarope, um elixir, uma emulsão, ou um desenho.

Uma molécula tendo N-glicanos adequados para administração tópica pode ser administrada a um mamífero (por exemplo, um paciente humano) como, por exemplo, um creme, um borrifador, uma espuma, um gel, uma pomada, uma unguento, ou um *dry rub*. Um *dry rub* pode ser reidratado no sítio da administração. Tais moléculas também podem ser infundidas diretamente (por exemplo, embebidas em e secas) em uma bandagem, gaze ou adesivo, que então pode ser aplicado topicamente. Tais moléculas também podem ser mantidas em um estado semilíquido, gelificadas, ou totalmente líquido em uma bandagem, gaze, ou adesivo para administração tópica (ver, por exemplo, Patente U.S. No. 4.307.717).

Quantidades terapeuticamente eficazes de uma composição farmacêutica podem ser administradas a um indivíduo em necessidade da mesma em um regime de dosagem verificável por um versado na técnica. Por exemplo, uma composição pode ser administrada ao indivíduo, por exemplo, sistematicamente em uma dosagem de 0,01 µg/kg a 10.000 µg/kg de peso corporal do indivíduo, por dose. Em outro exemplo, a dosagem é de 1 µg/kg a 100 µg/kg de peso corporal do indivíduo, por dose. Em outro exemplo, a dosagem é de 1 µg/kg a 30 µg/kg de peso corporal do indivíduo, por dose, por exemplo, de 3 µg/kg a 10 µg/kg de peso corporal do indivíduo, por dose.

A fim de otimizar a eficácia terapêutica, uma molécula contendo um N-glicano pode ser primeiro administrada em regimes de dosagem diferentes. A dose da unidade e o regime dependem de fatores que incluem, *por exemplo*, a espécie do mamífero, seu status imune, o peso corporal do mamífero. Tipicamente, os níveis de tal molécula em um tecido podem ser monitorados usando ensaios de rastreamento apropriados como parte de um procedimento de teste clínico, *por exemplo*, para determinar a eficácia de um regime de tratamento dado.

A frequência da dosagem de uma molécula está dentro das ha-

bilidades e o juízo clínico de profissionais médicos (por exemplo, médicos ou enfermeiras). Tipicamente, o regime de administração é estabelecido por testes clínicos que podem estabelecer parâmetros ótimos de administração. Entretanto, o profissional pode variar tais regimes de administração de acordo com a idade do indivíduo, saúde, peso, sexo e status médico. A frequência da dosagem pode ser variada dependendo de se o tratamento é profilático ou terapêutico.

A toxicidade e a eficácia terapêutica de tais moléculas ou composições farmacêuticas das mesmas podem ser determinadas por procedimentos farmacêuticos conhecidos, por exemplo, em culturas celulares ou animais experimentais. Estes procedimentos podem ser usados, por exemplo, para determinar a LD50 (a dose letal a 50% da população) e a ED50 (a dose terapeuticamente eficaz em 50% da população). A proporção de dose entre efeitos tóxicos e terapêuticos é o índice terapêutico e pode ser expresso como a proporção LD50/ED50. As composições farmacêuticas que exibem altos índices terapêuticos são preferenciais. Embora as composições farmacêuticas que exibem efeitos colaterais tóxicos possam ser usadas, cuidado deve ser tomado para projetar um sistema de entrega que visa tais compostos ao sítio do tecido afetado a fim de minimizar o dano potencial a células normais (por exemplo, células não alvo) e, por meio disso, reduzir os efeitos colaterais.

Os dados obtidos dos ensaios de cultura celular e estudos animais podem ser usados na formulação de uma faixa da dosagem para uso em indivíduos apropriados (por exemplo, pacientes humanos). A dosagem de tais composições farmacêuticas geralmente está dentro de uma faixa de concentrações circulantes que incluem a ED50 com muito pouco ou nada de toxicidade. A dosagem pode variar dentro desta faixa dependendo da forma de dosagem empregada e da via de administração utilizada. Para uma composição farmacêutica usada como descrito neste pedido (por exemplo, para tratar um distúrbio metabólico em um indivíduo), a dose terapeuticamente eficaz pode ser estimada inicialmente de ensaios de cultura celular. Uma dose pode ser formulada em modelos animais para alcançar uma faixa de

concentração plasmática circulante que inclui a IC₅₀ (isto é, a concentração da composição farmacêutica que alcança uma inibição meio-máxima de sintomas) como determinada na cultura celular. Tal informação pode ser usada para determinar mais exatamente doses úteis em humanos. Os níveis no plasma podem ser medidos, por exemplo, pela cromatografia líquida de alto desempenho.

Como definido neste pedido, uma "quantidade terapeuticamente eficaz" de uma molécula contendo um N-glicano é uma quantidade da molécula que é capaz de produzir um resultado medicamente desejável (por exemplo, a melhora de um ou mais sintomas de um distúrbio metabólico) em um indivíduo tratado. Uma quantidade terapeuticamente eficaz (isto é, uma dosagem eficaz) pode incluir quantidades de micrograma ou miligrama do composto por quilograma do peso do indivíduo ou da amostra (por exemplo, aproximadamente 1 micrograma por quilograma a aproximadamente 500 miligramas por quilograma, aproximadamente 100 microgramas por quilograma a aproximadamente 5 miligramas por quilograma, ou aproximadamente 1 micrograma por quilograma a aproximadamente 50 microgramas por quilograma).

O indivíduo pode ser qualquer mamífero, por exemplo, um ser humano (por exemplo, um paciente humano) ou um primata não humano (por exemplo, chimpanzé, babuíno ou macaco), um camundongo, um rato, um coelho, um porquinho da Índia, um gerbo, um hamster, um cavalo, um tipo do gado (por exemplo, vaca, porco, ovelha ou cabra), um cão, um gato, ou uma baleia.

Uma molécula ou composição farmacêutica da mesma descrita neste pedido pode ser administrada a um indivíduo como uma terapia de combinação com outro tratamento, por exemplo, um tratamento de um distúrbio metabólico (por exemplo, um distúrbio de armazenamento lisossomal). Por exemplo, a terapia de combinação pode incluir a administração no indivíduo (por exemplo, um paciente humano) de um ou mais agentes adicionais que fornecem um benefício terapêutico ao indivíduo que tem, ou está em risco de desenvolvimento, (ou suspeito em ter) um distúrbio metabólico (por

exemplo, um distúrbio de armazenamento lisossomal). Dessa forma, o composto ou composição farmacêutica e um ou mais agentes adicionais podem ser administrados ao mesmo tempo. Alternativamente, a molécula pode ser administrada primeiro e um ou mais agentes adicionais administrados segundo, ou vice-versa.

Será apreciado que em exemplos onde uma terapia prévia é particularmente tóxica (por exemplo, um tratamento para um distúrbio metabólico com perfis significantes de efeitos colaterais), a administração de uma molécula descrita neste pedido pode ser usada para compensar e/ou diminuir a quantidade da terapia anterior em um nível suficiente para fornecer o mesmo benefício terapêutico ou melhorado, mas sem a toxicidade.

Qualquer uma das composições farmacêuticas descritas neste pedido pode estar incluída em um recipiente, pacote, ou dispensador em conjunto com instruções de administração.

A seguir são os exemplos da prática da invenção. Não devem ser interpretados como limitação do escopo da invenção de nenhum modo.

EXEMPLOS

A tabela 2 contém uma lista de todas as cepas usadas nos experimentos descritos abaixo. Na Tabela 2, MH = α-1,2-manosidase marcada com HDEL; ζ = integração randômica através de sequências zeta; ancoragem Δ = integração em um locus específico; e (H) = resistente à higromicina.

TABELA 2
Listagem de Cepas Usadas em Exemplos

Número	Nome curto	Descrição	Marcadores	N-Glicanos Esperados
G013	P01d Inuga Δoch1(URA3) cl 26.1	P01d Inuga Δoch1(URA3) cl 26.1 transformado com pYLOCH1PUT-TOPO digerido com SpeI/Bst1107	URA3 ⁺ leu2 ade2 ⁻ gut2 ⁻	Principalmemente Man ₅ GlcNAc ₂
G014	P01d Inuga Δoch1 (curado) cl 7	P01d Inuga Δoch1 (G013) curado do marcador URA3 usando pJUB4-Cre	ura3 ⁻ leu2 ⁻ ade2 ⁻ gut2 ⁻	Principalmemente Man ₅ GlcNAc ₂
G016	P01d Inuga Δoch1 TerfMH (S-Not) cl 1.4	P01d Inuga Δoch1 (curado) (G014) transformado com pYLUXL2preManHDEL(Y) digerido com NotI	URA3 ⁺ leu2 ade2 ⁻ gut2 ⁻	Man ₅ GlcNAc ₂
G018	P01d Inuga Δoch1 Hp4dMH (S-Not) cl 11.2	P01d Inuga Δoch1 (curado) (G014) transformado com pYLUXL2preManHDEL(Y) digerido com NotI	URA3 ⁺ leu2 ade2 ⁻ gut2 ⁻	Man ₅ GlcNAc ₂
G036	P01d Inuga Δoch1 Hp4dMH (S-Not)(curado) cl 2.2 (1 cópia ManHDEL)	P01d Inuga Δoch1 Hp4dMH (Not) cl 11.2 (G018) curado do marcador URA3 usando pRRQ2	ura3 ⁻ leu2 ⁻ ade2 ⁻ gut2 ⁻	Man ₅ GlcNAc ₂
G039	P01d Inuga Δoch1 Hp4dMH (S-curado) Δ- alg3ALG6 cl 24.1	P01d Inuga Δoch1 Hp4dMH (S-Not-curado) cl 2.2 (G036) transformado com pYALG3PUT-ALG6 digerido com NotI/PacI	URA3 ⁺ leu2 ade2 ⁻ gut2 ⁻	Glc ₁₋₂ Man ₅ GlcNAc ₂ e Man ₃ GlcNAc ₂
G040	P01d Inuga Δoch1 Hp4dMH (S-curado) TefGnTl cl 5.4 (1 cópia GnT)	P01d Inuga Δoch1 Hp4dMH (S-Not-curado) cl 2.2 (G036) transformado com pYLTmAx hGnTl digerido com NotI	URA3 ⁺ leu2 ade2 ⁻ gut2 ⁻	GlcNAcMan ₅ GlcNAc ₂
G043	P01d Inuga Δoch1 Hp4dMH (S-curado) TefGnTl TefManl cl 15	P01d Inuga Δoch1 Hp4dMH (S-curado) TefGnTl cl 5.4 (G040) transformado com pYLTmAXDm-Manl (LEU2 ex) digerido com NotI	URA3 ⁺ LEU2 ⁺ ade2 ⁻ gut2 ⁻	GlcNAcMan ₃ GlcNAc ₂
G044	P01d Inuga Δoch1 Hp4dMH (S-curado) TefGnTl TefGnTl cl 12	P01d Inuga Δoch1 Hp4dMH (S-curado) TefGnTl cl 5.4 (G040) transformado com pYLTmAXSpGal10hGnTl (ADE2 ex) digerido com NotI	URA3 ⁺ leu2 ADE2 ⁺ gut2 ⁻	GalGlcNAcMan ₅ GlcNAc ₂
G045	P01d Inuga Δoch1 Hp4dMH (S-curado) Δ- alg3ALG6 cl 2.16 (curado)	P01d Inuga Δoch1 Hp4dMH (S-curado) Δ- alg3ALG6 cl 24.1 (G039) curado do marcador URA3 usando pRRQ2	ura3 ⁻ leu2 ⁻ ade2 ⁻ gut2 ⁻	Glc ₁₋₂ Man ₅ GlcNAc ₂ e Man ₃ GlcNAc ₂
G046	P01d Inuga Δoch1 Hp4dMH (ancoragem Δleu2)	P01d Inuga Δoch1 (curado) (G014) transformado com JME926 pPTIeu2-ADE2Ex- Hp4dManHDEL (Y) digerido com NotI	ura3 ⁻ leu2 ⁻ ADE2 ⁺ gut2 ⁻	Man ₅ GlcNAc ₂

Continuação

G047	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 TerfGnTl(H) cl2 (1 cópia GnT)	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 cl 24.1 (G039) transformado com pYL-TmAXhGnTl (Hyg ^R ex) digerido com NotI	URA3 ^r leu2 ^r ade2 ^r gut2 ^r	Glc ₁₋₂ Mans ₅ GlcNAc ₂ e GlcNAcMan ₃ GlcNAc ₂
G048	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 (curado) clone TerfGnTl 7.3	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 cl 2.16 (curado) (G045) transformado com pYL-TmAXhGnTl digerido com NotI	URA3 ^r leu2 ^r ade2 ^r gut2 ^r	Glc ₁₋₂ Mans ₅ GlcNAc ₂ e GlcNAcMan ₃ GlcNAc ₂
G050	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 (curado) TerfGnTl TerfGnTl cl. 42.3	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 cl 2.16 (curado) (G045) transformado com pYL-TmAXhGnTl e pYL-TmAXhGnTl (ADE2 Ex) digerido com NotI	URA3 ^r leu2 ^r ADE2 ^r gut2 ^r	Glc ₁₋₂ Mans ₅ GlcNAc ₂ e GlcNAc ₂ Man ₃ GlcNAc ₂
G051	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 TerfGnTl(H) TerfGnTl clone 4.5	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 TerfGnTl cl2 (G047) transformado com pYL-TmAXhGnTl (ADE2 Ex) digerido com NotI	URA3 ^r leu2 ^r ADE2 ^r gut2 ^r	Glc ₁₋₂ Mans ₅ GlcNAc ₂ e GlcNAc ₂ Man ₃ GlcNAc ₂
G052	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Hp4dhGnTl (ζ) clone 16	P01d Inuga Δoch1 Hp4dMH (ζ -Not-curado) cl 2.2 (G036) transformado com pYLHp4mAhhGnTl digerido com NotI	URA3 ^r leu2 ^r ade2 ^r gut2 ^r	GlcNAcMan ₅ GlcNAc ₂
G053	P01d Inuga Δoch1 Hp4dMH (ancoragem Δaxp1)	P01d Inuga Δoch1 (curado) (G014) transformado com OXYYP289-pPTAxp1-Leu2Ex-Hp4dManHDEL (YI) digerido com NotI	ura3 ^r LEU2 ^r ade2 ^r gut2 ^r	Man ₅ GlcNAc ₂
G054	P01d Inuga Δoch1 Hp4dMH (ancoragem Δaxp1) (curado)	P01d Inuga Δoch1 Hp4dMH (ancoragem Δaxp1) (G053) curado do marcador LEU2 usando pUB4-Cre	ura3 ^r leu2 ^r ade2 ^r gut2 ^r	Man ₅ GlcNAc ₂
G055	P01d Inuga Δoch1 Hp4dMH (ancoragem Δleu2-curado)	P01d Inuga Δoch1 Hp4dMH (ancoragem Δleu2) (G046) curado do marcador ADE2 usando pR-RQ2	ura3 ^r leu2 ^r ade2 ^r gut2 ^r	Man ₅ GlcNAc ₂
G056	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 Hp4dGnTl (ζ) clone E	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 cl 2.16 (curado) (G045) transformado com pYLHp4mAhhGnTl digerido com NotI	URA3 ^r leu2 ^r ade2 ^r gut2 ^r	Glc ₁₋₂ Mans ₅ GlcNAc ₂ e GlcNAcMan ₃ GlcNAc ₂
G057	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 Hp4dGnTl (ancoragem Δade2) clone G	P01d Inuga Δoch1 Hp4dMH (ζ -curado) Δ-alg3ALG6 cl 2.16 (curado) (G045) transformado com digerido com NotI JME925 pPTade2-URAA3ex-Hp4dhGnTl	URA3 ^r leu2 ^r ade2 ^r gut2 ^r	Glc ₁₋₂ Mans ₅ GlcNAc ₂ e GlcNAcMan ₃ GlcNAc ₂

Continuação

G058	P01d Inuga Δoch1 Hp4dMH (ζ-curado) Δ-alg3ALG6 Hp4dGnT1 (ζ-curado)	P01d Inuga Δoch1 Hp4dMH (ζ-curado) Δ-alg3ALG6 Hp4dGnT1 (ζ) (G056) curado do marcador URA3 usando pRRQ2	ura3' leu2' ade2' gut2'	Glc ₁₋₂ MansGlcNAc ₂ e GlcNAcMan ₃ GlcNAc ₂
G059	P01d Inuga Δoch1 Hp4dMH (ζ-curado) Δ-alg3ALG6 Hp4dGnT1 (ancoragem Δade2-curado)	P01d Inuga Δoch1 Hp4dMH (ζ-curado) Δ-alg3ALG6 Hp4dGnT1 (ancoragem Δade2) (G057) curado do marcador URA3 usando pRRQ2	ura3' leu2' ade2' gut2'	Glc ₁₋₂ MansGlcNAc ₂ e GlcNAcMan ₃ GlcNAc ₂
G060	P01d Inuga -och1 Hp4dMH (ζ-curado) -alg3ALG6 Hp4dGnT1 (ancoragem -ade2) Hp4dGls2I/9 (ζ) clone 6	P01d Inuga -och1 Hp4dMH (ζ-curado) -alg3ALG6 Hp4dGnT1 (G057) transformado com Hp4dL2preAnGlicII a+b(alt) digerido com Not	Hp4dMH (ζ-curado) (ancoragem -ade2) com Zeta-LEU2Ex-	GlcNAcMan ₃ GlcNAc ₂
G061	P01d Inuga -och1 Hp4dMH (ζ-curado) -alg3ALG6 Hp4dGnT1 (ancoragem -ade2) Hp4d Gls2I/8 (ancoragem -ura3) clone 18	P01d Inuga -och1 Hp4dMH (ζ-curado) -alg3ALG6 Hp4dGnT1 (G057) transformado com JME923 pPTUra3-Leu2Ex-Hp4d L2preAnGlicIIa+b(alt) digerido com Not	Hp4dMH (ζ-curado) (ancoragem -ade2) com pYLP4mAXrGnTII	GlcNAcMan ₃ GlcNAc ₂
G070	P01d Inuga Δoch1 Hp4dMH (1 cópia-ζ-curado) Δalg3ALG6 (curado) Hp4dGnT1 (ancoragem Δade2) Hp4dGls2αβ (ancoragem Δura3) Hp4dGnTII (ζ) cl 6	P01d Inuga Δoch1 Hp4dMH (1 cópia-ζ-curado) Δalg3ALG6 (curado) Hp4dGnT1 (ancoragem Δade2) Hp4d Gls2 α/β (ancoragem Δura3) (G061) transformado com pYLP4mAXrGnTII (ADE2ex) digerido com Not	URA3 ⁺ LEU2 ⁺ ADE2 ⁺ gut2'	GlcNAc ₂ Man ₃ GlcNAc ₂
G071	P01d Inuga Δoch1 Hp4dMH (1 cópia-ζ-curado) Δalg3ALG6 (curado) Hp4dGnT1 (ancoragem Δade2) Hp4d Gls2 α/β (ancoragem Δura3) Hp4dGnTII (integração no locus Axp1 estava apontado para) cl 8	P01d Inuga Δoch1 Hp4dMH (1 cópia-ζ-curado) Δalg3ALG6 (curado) Hp4dGnT1 (ancoragem Δade2) Hp4d Gls2 α/β (ancoragem Δura3) (G061) transformado com OXYP2289 PPTAxp1-ADE2ex-Hp4dGnTII digerido com Not	URA3 ⁺ LEU2 ⁺ ADE2 ⁺ gut2'	GlcNAc ₂ Man ₃ GlcNAc ₂
G096	P01d Inuga Δoch1 Hp4dMH (1 cópia-ζ-curado) Δalg3ALG6 (curado) Hp4dGnT1 (ancoragem Δade2) Hp4d Gls2 α/β (ancoragem Δura3) Hp4dGnTII (integração no locus Axp1 estava apontado para) Hp4dPP-HC/LC clone 13	P01d Inuga Δoch1 Hp4dMH (1 cópia-ζ-curado) Δalg3ALG6 (curado) Hp4dGnT1 (ancoragem Δade2) Hp4d Gls2 α/β (ancoragem Δura3) Hp4dGnTII (ancoragem Δaxp1) (G071) transformado com pYLP4L2preproHerHC&LC (Gut2ex)-ori2 digerido com Not	URA3 ⁺ LEU2 ⁺ ADE2 ⁺ GUT2 ⁺	GlcNAc ₂ Man ₃ GlcNAc ₂

Exemplo 1: perturbação de OCH1 de *Yarrowia lipolytica*

A geração de uma cepa de expressão proteica glicoengenheirada foi feita na cepa de *Yarrowia lipolytica* po1d Inuga (uma cepa tendo os auxotróficos *leu2-*, *ura3-*, *gut2-e ade2-*). Uma estratégia para nocaute do gene OCH1 (Acesso GenBank No: AJ563920) em *Yarrowia lipolytica* foi baseado como descrito para o gene *LIP2* (Fickers et al., 2003 *J Microbiol Methods.* 55 (3):727-37). A estratégia de construção gênica seguida para o gene OCH1 é descrita na Publicação de Patente U.S. No. 20090069232-A1. O vetor resultante foi chamado pYIOCH1 PUT TOPO (figura 1B).

O fragmento OCH1 KO foi isolado do plasmídeo por um digestão de restrição Spel/Bst1107I e transformado a cepa de *Yarrowia lipolytica* po1d Inuga. Várias cepas uracila prototróficas foram obtidas e rastreadas por PCR em DNA genômico (gDNA) usando os iniciadores *Yloch1 prom fw* (5'-TCGCTATCACGTCTAGC-3', SEQ ID NO:1) e *Yloch1 term rev* (5'-ACTCTGTATACTTGTATGTACTGTGAGAC-3', SEQ ID NO:2) para analisar a integração genômica do plasmídeo. Um fragmento de tamanho correto (isto é, 2328 pb contra 1894 pb no selvagem) foi amplificado para vários clones testados. O nocaute do gene OCH1 também foi confirmado pela análise de N-glicano do agrupamento de glicoproteínas totais secretadas no meio de crescimento (= secretoma): a estrutura Man₈GlcNAc₂ tornou-se o N-glicano predominante dentro do perfil de açúcar (figura 2). Este perfil diferencia-se daquele da cepa selvagem, que contém uma quantidade mais alta de Man₉GlcNAc₂ – o último mais provavelmente contém uma manose adicional em consequência da atividade de OCH1p – bem como algumas estruturas com um número mesmo mais alto de resíduos de manose.

Para remover o gene *URA3*, um clone Δ och1 positivo (chamado de G013, vide a Tabela 2) foi transformado com o plasmídeo episomal pUB4-Cre (Fickers et al., 2003, supra) que contém um cassete de expressão de Cre recombinase. A remoção do gene *URA3* foi rastreada por PCR em gDNA usando os iniciadores *Yloch1 prom fw* e *Yloch1 term rev* (ver acima). Os clones nos quais o marcador URA3 não foi extirpado mais resultaram na amplificação de uma banda de 2328 pb; em vez disso um fragmento de PCR

de 1075 pb (excl. *URA3*) foi obtido. Os clones positivos foram verificados no nível de N-glicano do secretoma e mostram um perfil muito similar àquela da cepa não curada (figura 2). Uma das cepas curadas (chamada de G014, ver Tabela 2) foi selecionada para engenharia adicional de N-glicano.

5 Exemplo 2: superexpressão de uma α-1,2-manosidase retida no ER por integração randômica ou por integração direcionada/ancorada

Para permitir a geração de $\text{Man}_5\text{GlcNAc}_2$ ligado a glicoproteínas expressas por uma cepa Δoch1, α-1,2-manosidase foi expressa para clivar $\text{Man}_8\text{GlcNAc}_2$ a $\text{Man}_5\text{GlcNAc}_2$ (isto é, atividade de α-1,2-manosidase tipo Golgi). Tal manosidase deve ser direcionada para o sistema de secreção. α-1,2-manosidase de *Trichoderma reesei* (acesso Genbank no. AF212153), fusionada ao pré-pró-fator de acoplamento de *S. cerevisiae* e marcada com uma sequência HDEL (SEQ ID NO:21) para localizá-la no ER, é capaz de cortar $\text{Man}_8\text{GlcNAc}_2$ a $\text{Man}_5\text{GlcNAc}_2$ *in vivo* em *Pichia pastoris* bem como em 10 *Trichoderma reesei* e *Aspergillus niger*. Os construtos de expressão foram feitos onde uma versão códon-otimizada de α-1,2-manosidase de *T. reesei* marcada com HDEL foram fusionados à pré-sequência sinal *LIP2* de *Y. lipolytica* e colocada no controle transcracional de qualquer promotor *TEF1*, Hp4d (Madzak *et al.*, 2000, *J. Mol. Microbiol. Biotechnol.* 2:207-216), GAP ou 15 *POX2*. A estratégia de construção destes plasmídeos é descrita na Publicação de Patente U.S. No. 20090069232-A1.

Dois destes vetores, pYLHUXdL2preManHDEL e pYLTUX-dL2preManHDEL (figura 3) - com a manosidase no controle transcracional do promotor Hp4d resp. *TEF1*, foram usados para transformar a cepa G014 (derivada do Exemplo 1). Os vetores foram digeridos com *NotI* para permitir a integração randômica no genoma através das sequências zeta. Os transformantes *URA3* prototróficos foram selecionados para a análise de N-glicano. Vários transformantes mostram uma conversão clara de $\text{Man}_8\text{GlcNAc}_2$ em direção a $\text{Man}_5\text{GlcNAc}_2$ (figura 4). Uma vez que os clones que expressam 20 manosidase no promotor controle *TEF1* mostraram um fenótipo de crescimento lento e agrupado (um destes clones foi chamado de G016), as etapas adicionais na glicoengenharia foram feitas em um contexto de cepa onde o 25

gene está sob controle transcrecional de Hp4d.

Um clone positivo que o ManHDEL sob controle do promotor hp4d (G018) foi escolhido, do qual o marcador *URA3* foi curado através da transformação transiente do plasmídeo pRRQ2 (Richard *et al.*, 2001 *J. Bacteriol.* 183:3098-3107), expressando Cre-recombinase. Vários clones de *ura3* foram selecionados após o procedimento e um clone (G036), mostrando um perfil de $\text{Man}_5\text{GlcNAc}_2$ claro no secretoma, foi usado para o trabalho de engenharia adicional (figura 4). A análise Southern deste clone revelou a presença de um cassete de expressão de manosidase randomicamente integrado. Esta análise Southern foi realizada em DNA genômico digerido com Hind III usando um fragmento de PCR específico marcado com DIG para manosidase que foi gerado usando iniciadores Man para (5'-GCCT TCCA-GACCTCTTGGAACGCCTACCACC-3', SEQ ID NO:22) e Man rev (5'-GCC AGGTGGCCGCCTCGTCGAGAAGAAGATCG-3', SEQ ID NO:23).

Em uma estratégia alternativa, dois construtos foram gerados que permitem a integração direcionada do cassete de expressão de manosidase dirigida por Hp4d no locus *LEU2* ou *AXP1* do genoma de *Yarrowia*. A construção destes plasmídeos, JME926_pPTLeu2-ADE2ex-Hp4dManHDEL (YI) e OXYP289_pPTAxp1-LEU2ex-Hp4dManHDEL (YI), é descrita na figura 5. Antes da transformação para a cepa G014, ambos os construtos foram digeridos com NotI e os respectivos cassetes de expressão foram isolados. Clones *ADE2* selecionados prototróficos foram integrados potencialmente ao cassete de expressão de manosidase no locus *LEU2*, ao passo que prototróficos *LEU2* potencialmente tinham integrado o cassete no locus *AXP1*. Os transformantes foram verificados pela análise Southern para avaliar o direcionamento apropriado no genoma. Isto foi realizado em DNA genômico digerido com BamHI (integração no locus *LEU2*) ou digerido com HindIII (integração no locus *AXP1*) usando um fragmento de PCR específico marcado com DIG para a manosidase que foi gerada usando os iniciadores Man para (5'-GCCTCCAGACCTCTTGGAACGCCTACCACC-3', SEQ ID NO:22) e Man rev (5'-GCCAGGTGGCCGCCTCGTCGAGAAGAAGATCG-3', SEQ ID NO:23). Os clones selecionados também foram verificados para a natureza

dos N-glicanos sintetizados para as glicoproteínas secretadas. Na maioria dos casos, a expressão de α -1,2-manosidase dirigida por Hp4d corretamente direcionada resultou na síntese de oligossacarídeos predominantemente $\text{Man}_5\text{GlcNAc}_2$ (figura 6). Para cada locus de direcionamento, um clone de expressão de manosidase (G046 em caso de ancoramento de *LEU2*; G053 em caso de ancoramento de *AXP1*) foi selecionado para curar através da expressão transiente de Cre recombinase usando o plasmídeo pRRQ2 para a cepa G046 e pUB4-Cre da cepa G053. As cepas curadas resultantes (G055 e G054, respectivamente) foram reverificadas através de Southern blotting e seu perfil de $\text{Man}_5\text{GlcNAc}_2$ confirmado através da DSA-FACE usando análise de N-glicano.

Exemplo 3: expressão de GlcNAc-transferase I

Uma sequência códon-otimizada de *Yarrowia* foi gerada para a expressão de uma proteína de fusão consistindo dos 100 primeiros aminoácidos N-terminais da proteína Kre2 de *S. cerevisiae* (SwissProt AccNo P27809) seguida pelo domínio catalítico de GlcNAc-transferase I humana (SwissProt AccNo P26572) (figura 7, SEQ ID NO:3 e SEQ ID NO:4). A Kre2p de levedura de 100 aminoácidos N-terminais serve como um sinal de localização de Golgi para o domínio catalítico de GnT I. Deste modo, é assegurado que a proteína de fusão GnT I é localizada depois na via de secreção que α -1,2-manosidase marcada com HDEL retida no ER a fim de permitir à enzima que converte os N-glicanos ligados à proteína $\text{Man}_5\text{GlcNAc}_2$ a $\text{GlcNAcMan}_5\text{GlcNAc}_2$. O gene sintético códon-otimizado para a expressão da proteína de fusão foi colocado no controle transcripcional do promotor *TEF1* ou Hp4d, que resulta nos plasmídeos pYLTmAXhGnTI e pYLHp4mAXhGnTI. A estratégia de construção é mostrada na figura 8. A expressão funcional da proteína de fusão Kre2-GnT I deve resultar na adição de um resíduo GlcNAc β -1,2-ligado aos glicanos $\text{Man}_5\text{GlcNAc}_2$ disponíveis resultando na síntese de $\text{GlcNAcMan}_5\text{GlcNAc}_2$.

Os plasmídeos pYLTmAXhGnTI e pYLHp4mAXGnTI foram digeridos com NotI antes da transformação para a cepa G036 (confirmar no Exemplo 2), conhecido por produzir $\text{Man}_5\text{GlcNAc}_2$ N-glicanos nas suas proteí-

nas secretadas. Os transformantes foram selecionados para uracila prototrófica. A análise do perfil de N-glicosilação no secretoma de vários destes clones mostrou uma modificação clara no modelo de N-glicano: o $\text{Man}_5\text{GlcNAc}_2$ foi significativamente reduzido e um novo pico, representando um N-glicano com peso molecular mais alto (aproximadamente uma unidade de glicose extra), apareceu. Tratamento dos N-glicanos isolados com Jack Bean β -N-acetil-hexosaminidase, uma enzima capaz de remover os resíduos GlcNAc terminais β ligados, indicou que o novo N-glicano é $\text{GlcNAcMan}_5\text{GlcNAc}_2$: o novo pico desapareceu e foi completamente convertido em $\text{Man}_5\text{GlcNAc}_2$ (figura 9). Dependendo do método de cultura usado, aproximadamente 70% do agrupamento de N-glicano total resultou ser $\text{GlcNAcMan}_5\text{GlcNAc}_2$ (com aproximadamente 77% do $\text{Man}_5\text{GlcNAc}_2$ disponível que é convertido).

Um transformante que expressa a proteína de fusão Kre2-GnT I sob controle do promotor *TEF1* foi denominado cepa G040 e selecionado para uso adicional. A análise genômica desta cepa através de Southern blot indicou a presença de um cassete de expressão. A análise Southern foi feita em DNA genômico digerido com BamHI usando um fragmento de PCR marcado com DIG específico para GnT I que foi gerado usando iniciadores 5'-GGATGATCACACAATGGCCCTGTTCTG-3' (SEQ ID NO:5) e 5'-TGCTCTAGACTAGTTCCAAGAGGGGTC-3' (SEQ ID NO:6). A análise do perfil de glicosilação no secretoma da cepa G040 contra cepas carregando 1 a 3 cópias (confirmadas pelo mesmo southern blot) do cassete de expressão Kre2-GnT I dirigido por Hp4d, não mostrou diferença significante na capacidade de transferência de GlcNAc.

25 Exemplo 4: expressão de manosidase II

Uma sequência códon-otimizada de *Yarrowia* foi gerada para a expressão da proteína de fusão consistindo dos 36 primeiros aminoácidos N-terminais da proteína Mnn2 de *S. cerevisiae* (SwissProt AccNo P38069) seguida pelo domínio catalítico de manosidase II de *Drosophila melanogaster* (SwissProt AccNo Q24451) (figura 10, SEQ ID NO:7 e SEQ ID NO:8). A Mnn2 de levedura de 36 aminoácidos N-terminais serve como um sinal de localização Golgi do domínio catalítico de Man II. Deste modo, é assegurado

que a proteína de fusão Mnn2-Man II está localizada na mesma ou até uma posição posterior na via de secreção do que a proteína de fusão Kre2-GnT I e é, por isso, capaz de converter GlcNAcMan₅GlcNAc₂ em GlcNAc-Man₃GlcNAc₂. O gene sintético códon-otimizado de *Yarrowia* para a expressão da proteína de fusão foi colocado sob controle transcripcional do promotor *TEF1*, que resulta nos plasmídeos pYLTmAXDmManII e pYLTmAXDmManII (LEU2ex). A estratégia de construção é mostrada na figura 11.

O plasmídeo pYLTmAXDmManII (LEU2ex) foi digerido com NotI antes da transformação para a cepa G040 (ver o Exemplo 3), que era conhecido por produzir GlcNAcMan₅GlcNAc₂ N-glicanos nas suas proteínas secretadas. Os transformantes foram selecionados para a prototrofia de leucina. A análise do perfil de N-glicosilação no secretoma de vários destes clones mostrou uma modificação no modelo de N-glicano: um novo pico que representa um N-glicano com um peso molecular mais baixo de aproximadamente duas unidades de glicose apareceu, que pode indicar a formação de GlcNAcMan₃GlcNAc₂ e, dessa forma, atividade de manosidase II parcial. Também outro pico aparece, correndo quase na mesma posição que Man₅GlcNAc₂ (isto é um ombro ao pico), potencialmente representando GlcNAcMan₄GlcNAc₂. A última estrutura pode ser o resultado de um evento de corte parcial, onde a atividade de manosidase II somente removeu um resíduo de manose em vez de dois. O tratamento dos N-glicanos isolados com Jack Bean β -N-acetil-hexosaminidase resultou em um deslocamento à esquerda do modelo de glicano com aproximadamente uma unidade de glicose e dessa forma uma mobilidade eletroforética mais alta devido à perda de um resíduo de GlcNAc terminal (figura 12). Isto ainda confirma a geração de GlcNAcMan₄GlcNAc₂ e GlcNAcMan₃GlcNAc₂ de GlcNAcMan₅GlcNAc₂ devido à expressão de uma atividade de manosidase II funcional. Dependendo do método de cultura usada, aproximadamente 15% do agrupamento de N-glicano total resultou ser GlcNAcMan₃GlcNAc₂: aproximadamente 35% do GlcNAcMan₅GlcNAc₂ disponível perdeu 1 ou 2 resíduos de manose, com 20% que são completamente cortados em direção a GlcNAcMan₃GlcNAc₂.

Exemplo 5: expressão de galactosiltransferase I

A síntese de N-glicanos com resíduos de galactose terminais não somente depende da presença de uma galactosiltransferase funcional e bem localizada dentro da via de secreção, mas também na disponibilidade da UDP-Gal, o substrato doador que é usado pela enzima. Embora se pense geralmente que UDP-Glc e UDP-GlcNAc estão suficientemente disponíveis no aparelho de Golgi de organismos de levedura, que é menos conhecido para UDP-Gal. Para superar a deficiência de UDP-Gal potencial durante a glicoengenharia, tentativas foram feitas anteriormente em *Pichia pastoris* para visar uma proteína de fusão UDP-Glc-4-epimerase de *Schizosaccharomyces pombe* (codificado pelo *GAL10* similar ao gene SPBC365.14c - SwissProt AccNo Q9Y7X5) e o domínio catalítico de β-1,4-galactosiltransferase I humana (GalT I) (SwissProt AccNo P15291) no aparelho de Golgi da levedura (Bobrowicz *et al.*, *Glycobiology* 14 (9):757-766, 2004). A localização da proteína de fusão Gal10p-GalT I dentro da via de secreção, preferencialmente em uma posição onde atividade de manosidase II e de transferência de GlcNAc já atuou sobre os N-glicanos das proteínas destinadas à secreção, foi realizada usando os primeiros 46 aminoácidos N-terminais de Mnn2p de *S. cerevisiae* como sinal de direcionamento N-terminal.

Por isso, uma sequência códon-otimizada de *Yarrowia* foi gerada para a expressão da proteína de fusão consistindo dos primeiros 46 aminoácidos N-terminais da proteína Mnn2 de *S. cerevisiae*, seguida pela proteína similar a Gal10 de *S. pombe* e o domínio catalítico de GalT I humana (figura 13). O gene sintético resultante foi colocado sob controle transcripcional do promotor *TEF1*, que resulta nos plasmídeos pYLtmAXSpGal10hGalTI e p-YLtmAXSpGal10hGalTI (ADE2ex). A estratégia de construção é mostrada na figura 14.

O plasmídeo pYLtmAXSpGal10hGalTI (ADE2ex) foi digerido com *NotI* antes da transformação para a cepa G040 (ver o Exemplo 3), conhecido por produzir GlcNAcMan₅GlcNAc₂ N-glicanos nas suas proteínas secretadas. Os transformantes foram selecionados para sua prototrofia de adenina. A análise do perfil de N-glicosilação no secretoma de vários destes

clones mostrou uma modificação no modelo de N-glicano: um novo pico aparece, correndo em uma posição entre $\text{Man}_7\text{GlcNAc}_2$ e $\text{Man}_8\text{GlcNAc}_2$ (figura 15). O tratamento dos N-glicanos com β -1,4-galactosidase de *Streptococcus pneumonia* indica que o pico representa $\text{GalGlcNAcMan}_5\text{GlcNAc}_2$ uma vez que esta digestão *in vitro* resulta no desaparecimento deste novo pico e igualmente um grande aumento em $\text{GlcNAcMan}_5\text{GlcNAc}_2$.

Usando esta organização e dependendo das condições de crescimento, aproximadamente 75% de $\text{GlcNAcMan}_5\text{GlcNAc}_2$ foram convertidos em $\text{GalGlcNAcMan}_5\text{GlcNAc}_2$. A soma total da estrutura galactosilada contabilizava aproximadamente 25% do agrupamento de N-glicano total. De uma digestão de α -1,2-manosidase *in vitro*, é claro, entretanto, que uma quantidade significante de N-glicanos de alta manose não foi convertida em $\text{Man}_5\text{GlcNAc}_2$ (figura 15). Dependendo do meio de cultura usado, a taxa de conversão de $\text{Man}_5\text{GlcNAc}_2$ em direção a $\text{GlcNAcMan}_5\text{GlcNAc}_2$ também é mais baixa do que aquela observado na cepa parental G040. Isto ainda está mais provavelmente relacionado à taxa de crescimento mais lenta observada para transformantes desta proteína de fusão Mnn2-Gal10-GaIT I.

Exemplo 6: nocaute de YIALG3 e superexpressão simultânea de YIALG6

Para permitir a geração de uma plataforma $\text{Man}_3\text{GlcNAc}_2$, o gene *ALG3* da cepa G036 (*Doch1* de *po1d* Inuga + α -1,2-manosidase dirigida por *Hp4d*) tem que ser inativado. Isto resulta na perda da atividade de *Alg3p* α -1,6-manosiltransferase localizada no ER e modifica a composição da estrutura de precursor de N-glicano ligado ao lipídio. A transferência desta estrutura para um sítio de N-glicosilação de uma cadeia de polipeptídeo nascente permite converter o perfil de glicosilação de levedura em estruturas N-glicano similares de mamíferos sem a necessidade de expressar Manosidase II. Entretanto, uma vez que esta nova estrutura ligada ao lipídio não é transferida eficientemente para polipeptídeos nascentes, o gene *ALG6* de *Yarrowia* (codificando uma *Alg6p* α -1,3-glicosil transferase localizada no ER) tem que ser superexpresso simultaneamente para reduzir a potencial subglicosilação de proteína tanto quanto possível.

Um vetor chamado de pYLalg3PUT-ALG6 (figura 16) foi constru-

ído anteriormente para permitir o nocaute simultâneo de *YIALG3* e superexpressão de *YIALG6* dirigido por Hp4d. Ver a Publicação de Patente U.S. No. 20090069232-A1. Um fragmento NotI/PacI deste vetor, contendo este casse-te de ativação/desativação, foi transformado em G036 de *Yarrowia lipolytica*

5 e os transformantes foram selecionados baseados na sua prototrofia de uracila. Os clones que tinham se integrado corretamente ao construto foram diretamente rastreados através da análise de N-glicano no secretoma. Dos 80 clones rastreados, 2 clones mostraram um perfil de N-glicosilação que pode ajustar-se com a inativação de *YIALG3* em uma cepa que expressa

10 uma α-1,2-manosidase localizada no ER. Separado de uma fração de $\text{Man}_3\text{GlcNAc}_2$ glicanos, havia ainda $\text{Man}_4\text{GlcNAc}_2$ e $\text{Man}_5\text{GlcNAc}_2$ bem como uma quantidade significante de N-glicanos glicosilados ($\text{Glc}-\text{Man}_5\text{GlcNAc}_2$ e $\text{Glc}_2\text{Man}_5\text{GlcNAc}_2$). Os últimos são o resultado de um corte ineficiente por Glicosidase II (Grinna e Robbins, *J. Biol. Chem.* 255, 2255-2258, 1980). A natureza das estruturas de $\text{Man}_4\text{GlcNAc}_2$ e $\text{Man}_5\text{GlcNAc}_2$ foi confirmada pelo tratamento *in vitro* dos N-glicanos com α-1,2-manosidase

15 (figura 17). Dependendo das condições de crescimento usadas, o nível de $\text{Man}_3\text{GlcNAc}_2$ pode aumentar até 60% do agrupamento de N-glicano total, com os picos de glicosilados que são insensíveis em direção à α-1,2-manosidase e somente levemente sensíveis em direção ao tratamento com Jack Bean α-manosidase (α-manosidase inespecífica que pode atuar nos resíduos de α-1,2-, α-1,3- e α-1,6 ligados à manose). Ao contrário, a última enzima converte $\text{Man}_3\text{GlcNAc}_2$ gerado em $\text{Man}_1\text{GlcNAc}_2$ (figura 17).

20

25

30

Um dos dois transformantes positivos foi chamado de G039 e usado para trabalho de glicoengenharia adicional. A cepa foi transformada transientemente com pRRQ2 expressando Cre-recombinase para permitir a conservação do marcador *URA3* que foi introduzido na transformação de G036 com o vetor pYLalg3PUT-ALG6. A análise mostra que o perfil de glicosilação permanece o mesmo após conservação. Uma cepa curada foi selecionada para uso adicional e designada G045.

Exemplo 7: expressão de GlcNAc-transferase I em uma cepa de produção de $\text{Man}_3\text{GlcNAc}_2$

Similar ao que foi feito no exemplo 3, a introdução de uma atividade de GnT I foi realizada através da expressão da proteína de fusão Kre2-GnT I. A integração randômica de tal construto de expressão para GnT I foi realizado de três modos: **1)** a cepa não curada G039 (ver o Exemplo 6) foi transformada com o vetor digerido com NotI pYLTmAXhGnTI (Hygr ex) e clones expressando GnT I foram inicialmente selecionados com base na sua capacidade de sobreviver a 300 µg/ml de higromicina adicionados às placas de seleção, **2)** a cepa curada G045 (ver o Exemplo 6) foi transformada com o vetor digerido com NotI pYLTmAXhGnTI (ver também o Exemplo 3) e clones expressando GnT I foram inicialmente selecionados com base em sua prototrofia de uracila ou **3)** a cepa curada G045 (ver o Exemplo 6) foi transformado com o vetor digerido com NotI pYLHp4mAXhGnTI e clones expressando GnT I foram inicialmente selecionados com base em sua prototrofia de uracila. A estratégia de construção de pYLTmAXhGnTI (Hygr ex) é mostrada na figura 18. Usando plasmídeos pYLTmAXhGnTI (Hygr ex) e pYLT-mAXhGnTI, a expressão de GnT I estava sob o controle transcrional do promotor *TEF1*; usando plasmídeo pYLHp4mAXhGnTI, a expressão de GnT I estava sob o controle do promotor Hp4d.

A transformação de G039 com pYLTmAXhGnTI (Hygr ex) resultou em três clones que somente emergiram nas placas de cultura após um período de incubação mais longo do que o que foi esperado. Entretanto, a análise do perfil de N-glicosilação do secretoma destes clones mostrou uma modificação clara no modelo de N-glicano: $\text{Man}_3\text{GlcNAc}_2$ presente na cepa de G039 não transformada foi significativamente reduzido ou quase completamente ausente enquanto um novo pico, representando um N-glicano com peso molecular mais alto (aproximadamente uma unidade de glicose extra), apareceu. Tratamento dos N-glicanos isolados com Jack Bean β -N-acetylhexosaminidase, uma enzima capaz de remover os resíduos GlcNAc terminais β ligados, indicou que o novo N-glicano de fato é $\text{GlcNAcMan}_3\text{GlcNAc}_2$. O novo pico desapareceu e foi completamente convertido em $\text{Man}_3\text{GlcNAc}_2$ (figura 19). Um dos transformantes avaliados foi usado para o trabalho de glicoengenharia adicional e denominado G047. Resultados similares tam-

bém foram obtidos quando a cepa curada G045 foi transformada com p-YLTmAXhGnT1 (G048) ou com pYLHp4mAXhGnT1 (G056). A Cepa G056 foi selecionada para curar através da expressão transiente de Cre recombinase usando o plasmídeo pRRQ2. A cepa resultante foi chamada G058.

5 Dependendo do método de cultura usado, aproximadamente 70% do agrupamento de N-glicano total da cepa G047 resultou ser GlcNAc-Man₃GlcNAc₂ com um pouco de Glc₁₋₂Man₅GlcNAc₂ restante e quase nenhum Man₃GlcNAc₂ estava presente (taxa de conversão >>90%) (figura 19). Apesar da alta taxa de conversão, somente uma cópia do cassete de expressão de GnT I pode ser identificada nesta cepa através de Southern blot. A análise Southern foi feita em DNA genômico digerido com BamHI usando um fragmento de PCR marcado com DIG específico para GnT I que foi gerado usando iniciadores 5'-GGATGATCACACAATGGCCCTGTTCTG-3' (SEQ ID NO:11) e 5'-TGCTCTAGACTAGTTCCAAGAGGGGTC-3' (SEQ ID NO:12).

10 Em uma estratégia alternativa, um construto JME925 pPTAde2-URA3ex-Hp4dhGnT1 foi gerado para permitir a integração direcionada do cassete de expressão de GnT I dirigido por Hp4d no locus *ADE2* do genoma de *Yarrowia*. A estratégia de construção é representada na figura 20. Antes 15 da transformação para a cepa G045, o plasmídeo foi digerido com NotI e o cassete de direcionamento/expressão foi isolado. Os transformantes foram selecionados baseados em sua prototrofia de adenina. A integração correta do cassete de expressão no locus *ADE2* foi verificada através de PCR usando o iniciador de sentido direto Ver1Ade2 (5'-CGACGATAGAGCAGGTCTCA 20 CTGTTGGGAATGCTG-3', SEQ ID NO:13) iniciador de sentido reverso Ver2Ade2 (5'-CTACACTGACGAAGTGGACATCCGGCTTGGACTG-3', SEQ ID NO:14) e ainda confirmado através de Southern blot. Isto foi feito em DNA genômico digerido com BamHI/Spel usando um fragmento de PCR marcado com DIG específico para GnT I que foi gerado usando iniciadores 5'-GGATGATCACACAATGGCCCTGTTCTG-3' (SEQ ID NO:15) e 5'-TGC 25 TCTAGACTAGTTCCAAGAGGGGTC-3' (SEQ ID NO:16). A síntese de GlcNAc-Man₃GlcNAc₂ para o secretoma foi confirmada através da análise de N-

glicano e tratamento *in vitro* com Jack Bean β -N-acetil-hexosaminidase (figura 21). Um transformante expressando GnT I (chamado de G057) foi selecionado para curar através da expressão transiente de Cre recombinase usando o plasmídeo pRRQ2. A cepa resultante foi chamada G059.

5 Exemplo 8: expressão de GlcNAc-transferase II

Uma sequência códon-otimizada de *Yarrowia* foi gerada para a expressão da proteína de fusão consistindo dos 36 primeiros aminoácidos N-terminais da proteína Mnn2 de *S. cerevisiae* (SwissProt AccNo P38069) seguida pelo domínio catalítico de GlcNAc-transferase II de rato (GnT II) (SwissProt AccNo Q09326) (figura 22, SEQ ID NO:17 e SEQ ID NO:18). Os 36 aminoácidos N-terminais de Mnn2 da levedura servem como um sinal de localização Golgi para o domínio catalítico de GnT II. Deste modo, foi assegurado que a proteína de fusão Mnn2-GnT II foi localizada no mesmo ou até uma posição posterior na via de secreção do que a proteína de fusão Kre2-GnT I (e Mnn2-Man II) e foi, por isso, capaz de converter GlcNAc-Man₃GlcNAc₂ em GlcNAc₂Man₃GlcNAc₂. O gene sintético da expressão da proteína de fusão foi colocado no controle transcripcional do promotor *TEF1*, que resulta nos plasmídeos pYLTmAXrGnTII e pYLTmAXrGnTII (ADE2ex). A estratégia de construção é mostrada na figura 23.

20 Uma cepa que expressa a atividade de GnT II foi gerada de dois modos diferentes: 1) cepa G045 (ver o Exemplo 6) foi transformada simultaneamente com pYLTmAXhGnTII digerido com NotI e pYLTmAXrGnTII digerido com NotI (ADE2 ex) e os transformantes foram selecionados baseados em sua prototrofia de uracila e adenina ou 2) cepa G047 (Exemplo 7) foi transformada com pYLTmAXrGnTII digerido com NotI (ADE2 ex) e os transformantes foram selecionados baseados em sua prototrofia de adenina. A integração dos cassetes de expressão foi verificada usando iniciador de sentido direto TefPromFW 5'-GTCCCCGAATTACCTTCC-3' (SEQ ID NO:19) e iniciador de sentido reverso Lip2TermRV 5'-AGGTAGAAGTTGAAAGAGTG-3' (SEQ ID NO:20). A análise de N-glicano no secretoma em combinação com o tratamento *in vitro* dos açúcares isolados com Jack Bean β -N-acetilhexosaminidase indicou que vários transformantes foram capazes de produ-

zir GlcNAc₂Man₃GlcNAc₂ e dessa forma expressando uma atividade de GnT II funcional (figura 24). Em uma condição selecionada, aproximadamente 40% do agrupamento de N-glicano total consistiu de GlcNAc₂Man₃GlcNAc₂. A taxa de conversão do substrato GlcNAcMan₃GlcNAc₂ a GlcNAc₂Man₃GlcNAc₂ foi 90%. As cepas selecionadas a partir do final foram chamadas G050 (a transformação dupla de G045) e G051 (expressão de GnT II em G047).

Exemplo 9: expressão de glicosidase II alfa e subunidades de beta (Gls2α e Gls2β)

Baseado nos experimentos descritos nos Exemplos 6 a 8, a estratégia que envolve o nocaute de *YIALG3* e superexpressão simultânea de *YIALG6* resulta na geração de N-glicanos carregando um ou dois resíduos de glicose terminal (Glc1-2Man5'GlcNAc2). A presença destes resíduos de glicose impede a conversão em direção a Man₃GlcNAc₂ pela α-1,2-manosidase marcada com HDEL localizada no ER. Para os resíduos de glicose a serem removidos, a atividade de Glicosidase II dentro do ER tem que ser aumentada. Em um contexto sem expressão de α-1,2-manosidase, a superexpressão da subunidade alfa e beta de Glicosidase II de *Aspergillus niger* resultaram em conversão mais alta de Glc₁₋₂Man₅GlcNAc₂ em Man5'GlcNAc₂ (Publicação de Patente U.S. No. 20090069232-A1). Os construtos da superexpressão de subunidades gls2 de *A. niger* foram produzidos como se segue: 1) um códon-otimizado de *Yarrowia* cDNA foi gerado para a expressão das subunidade gls2α e gls2β maduras (sem o peptídeo sinal) de *A. niger*; 2) os cDNA's foram clonados estruturados para a pré-sequência *LIP2* de *Y. lipolytica*; 3) as sequências resultante *LIP2pre-gls2α* e *LIP2pre-gls2β* foram clonadas sob controle transcripcional do promotor *TEF1* constitutivo. Os plasmídeos resultantes foram chamados pYLTUXdL2preAnGlcIIα e pYLeu2ExTEFpreLip2AnGluclβ (figura 25).

Baseado nestes plasmídeos, os novos construtos foram gerados para a superexpressão simultânea de subunidades gls2α e gls2β de *A. niger* sob controle do promotor *TEF1* (vetor JME923 pPTura3-LEU2ex-TefL2preAnGlcIIα+b[alt1] para a integração direcionada - figura 26) ou con-

trole do promotor Hp4d (vetor JME923 pPTura3-LEU2ex-Hp4dL2preAnGlcIIa+b [alt1] para integração direcionada e vetor Zeta-LEU2ex-Hp4dL2preAnGlcIIa+b [alt] para a integração randômica - figura 27).

A cepa G057 (ver o exemplo 7) foi transformada com plasmídeos digeridos com NotI JME923 pPTura3-LEU2ex-Hp4dL2preAnGlcIIa+b [alt1] e Zeta-LEU2ex-Hp4dL2preAnGlcIIa+b[alt] e transformantes foram selecionados baseados em sua prototrofia de leucina. Vários clones foram analisados genomicamente através de PCR e análise Southern para avaliar a integração do cassete de expressão de gls2 α e gls2 β . A geração de sonda DIG e análise por PCR da subunidade gls2 α foi feita usando os iniciadores AnGls2 α -FW (5'-GCTGGACTCTTCTTCTATCC-3') (SEQ ID NO:24) e AnGls2 α -RV (5'-GGTCTCCTTCAGAGACAGG-3') (SEQ ID NO:25); para o subunidade gls2 β fizemos o uso de iniciadores AnGls2 β -FW (5'-CCAAGTTCTACAAGGACACC-3') (SEQ ID NO:26) e AnGlc2 β -RV (5'-CCCTTGACGACCTTAGAGG-3') (SEQ ID NO:27). A análise Southern para verificar a integração direcionada do cassete de expressão dual Hp4dGls2 α / β foi feita em gDNA digerido com Eco47III usando a sonda gls2 α , e em Spel/Sfil digerido com gDNA usando a sonda gls2 β . A maioria dos clones selecionados mostrou integração correta do cassete de expressão dual no locus *URA3*. A análise Southern da integração randômica do cassete de expressão dual Hp4dGls2 α / β foi verificada em gDNA digerido com Pvul com ambas as sondas. Em todos os casos, somente uma cópia do cassete de expressão dual estava integrada.

Após, a análise de N-glicano foi realizada em vários clones confirmados tendo o cassete de expressão dual Hp4dGls2 α / β (corretamente) integrado. A N-glicosilação foi examinada na proteína secretada total após três dias de cultura falcon. Vários clones mostraram uma redução significante dos açúcares glicosilados e um aumento de Man₃GlcNAc₂ e GlcNAc-Man₃GlcNAc₂. Os perfis de um clone que integrou o cassete de expressão dual randomicamente (= cepa G060) de um lado e de um modo direcionado (= cepa G061) no outro, são mostrados na figura 28. Os dois picos menores representam Man₄GlcNAc₂ e Man₅GlcNAc₂, uma vez que deslocam

Man₃GlcNAc₂ resp. Man₁GlcNAc₂ no tratamento com α-1,2-manosidase e manosidase de Jack Bean. O último tratamento também resulta em uma conversão parcial do Glc₁₋₂Man₅GlcNAc₂ restante em Glc1-2Man₄GlcNAc₂ e de GlcNAcMan₃GlcNAc₂ em GlcNAcMan₂GlcNAc₂. A presença de Man₄GlcNAc₂ 5 e Man₅GlcNAc₂ entretanto indica conversão incompleta em direção a Man₃GlcNAc₂ por α-1,2-manosidase marcada com HDEL heterologamente coexpressa. Similarmente a presença de Man₃GlcNAc₂ indica a transferência incompleta de um GlcNAc-resíduo por GnT I humana recombinante para obter GlcNAcMan₃GlcNAc₂. Entretanto, baseado nos resultados descritos 10 acima (por exemplo, cultura de G047 no Exemplo 7, figura 19), é claro que as diferenças em condições de cultura podem aumentar as taxas de conversão significativamente e dessa forma melhorar o resultado final.

Exemplo 10: expressão de GlcNAc-transferase II na cepa G061 de produção de GlcNAcMan₃GlcNAc₂

15 Como descrito no Exemplo 8, uma sequência códon-otimizada *Yarrowia* foi gerada para a expressão da proteína de fusão consistindo dos 36 primeiros aminoácidos N-terminais da proteína Mnn2 de *S. cerevisiae* (SwissProt AccNo P38069) seguido pelo domínio catalítico da GlcNAc-transferase II de rato (GnT II) (SwissProt AccNo Q09326) (figura 22, SEQ ID 20 NO:17 e SEQ ID NO:18, respectivamente). Os 36 aminoácidos N-terminais de Mnn2 de levedura servem como um sinal para localização de Golgi para o domínio catalítico de GnT II. Deste modo, foi assegurado que a proteína de fusão Mnn2-GnT II estava localizada no mesmo ou até uma posição posterior na via de secreção do que a proteína de fusão Kre2-GnT I e foi, por isso, 25 capaz de converter GlcNAcMan₃GlcNAc₂ em GlcNAc₂Man₃GlcNAc₂. O gene sintético da expressão da proteína de fusão foi colocado sob controle transcripcional do promotor Hp4d que resulta em plasmídeo pYLHp4mA_XrGnTII, que foi usado para a integração randômica do cassete de expressão GnT II dirigido por Hp4d no genoma de *Yarrowia*. Em uma estratégia alternativa, o 30 construto OXYP289 pPTAxp1-ADE2ex-Hp4dhGnTII foi gerado para permitir a integração direcionada do cassete de expressão GnT II dirigido por Hp4d no locus *AXP1* do genoma de *Yarrowia*.

Antes da transformação da cepa G061 (ver o Exemplo 9), os plasmídeos foram digerido com NotI e o cassete de direcionamento/expressão foi isolado. Os transformantes foram selecionados baseados na sua prototrofia de adenina. A integração correta do cassete de expressão no locus *ADE2* foi confirmada pela análise Southern blot após digerir DNA genômico com XbaI. Uma sonda marcada com DIG com a especificidade da sequência de codificação de GnT II foi gerada usando iniciador de sentido direto rGnTII-FW (5'-GACCAGATGCTGCGAAACG-3') (SEQ ID NO: 28) e iniciador de sentido reverso rGnTII-RV (5'-CTTGACGTCCACCTTGTG-3') (SEQ ID NO: 29). Esta estratégia produz uma banda de 3172 pb quando o gene está integrado com sucesso no locus *Axp1*.

Em uma estratégia alternativa, a integração correta no locus *Axp1* pode ser examinada através de uma reação PCR em DNA genômico usando o iniciador de sentido direto AXPVer1b (5'-GCCTGAACGGCACGA
15 TG CGATCGTGGCAATCC-3') (SEQ ID NO: 30) e o iniciador reverso AXP-
Ver2b (5'-CAAGAACGCTCAGGCTCGGCGAATCTCCATC-3') (SEQ ID NO:
31). Em caso do direcionamento correto no locus *Axp1*, espera-se um fragmento de PCR de 6489 pb.

A análise de N-glicano no secretoma, em combinação com o tratamento *in vitro* dos açúcares isolados com Jack Bean β -N-acetilhexosaminidase ou α -1,2-manosidase de *T. reesei* indicou que vários transformantes foram capazes de produzir GlcNAc₂Man₃GlcNAc₂ e dessa forma de expressar uma atividade de GnT II funcional (figura 29). As análises indicaram que aproximadamente 25 a 30% do agrupamento de N-glicano total
25 consistiu de GlcNAc₂Man₃GlcNAc₂, com um GlcNAcMan₃GlcNAc₂ à taxa de conversão GlcNAc₂Man₃GlcNAc₂ de aproximadamente 90%. As cepas selecionadas a partir do final foram chamadas G070 (a integração de pY-LHp4mAXrGnTII em G061) e G071 (a integração de OXYP289 pPTAxp1-
ADE2ex-Hp4dhGnTII em G061).

30 Exemplo 11: construção de um plasmídeo de tandem para expressão dirigida por Hp4d simultânea da cadeia pesada (HC) e cadeia leve (LC) de anti-HER2 em *Yarrowia lipolytica*

As sequências de aminoácido das cadeias pesadas e leves do anticorpo anti-HER2 foram obtidas de Carter *et al.*, *Proc Natl Acad Sci USA.*, 89 (10): 4285-4289 (1992); e Ward *et al.*, *Appl Environ Microbiol.*, 70 (5): 2567-2576 (2004). As sequências de aminoácido relevantes foram traduzidas reversamente, códon-otimizadas para *Yarrowia lipolytica*, e sintetizadas por GenArt, Regensburg, Alemanha. As regiões de conteúdo de GC muito alto (> 80%) ou muito baixo (<30%) foram evitadas onde possível. Durante os processos de otimização, os seguintes motivos de sequência cis atuante foram evitados: caixas internas TATA, sítios chi e sítios de entrada ribossônica, estiramentos de sequência ricos em AT ou ricos em GC, sequências de repetição e estruturas secundárias de RNA bem como sítios doadores e aceptores de splice (críptico). A fim de permitir a secreção das proteínas ectópicas, a sequência de codificação do 'pré-pró-' sinal proteico Lip2 (seguido por aquele de um ligante de peptídeo 'GGG') foi adicionada à região 5' das sequências de codificação. 'GGG' foi adicionado para aumentar as modificações do processamento de Kex2 correto. A figura 30A contém a sequência nucleotídica do pré-pró-Lip2-LC sintético (= 750 pb) (SEQ ID NO: 32). A figura 30B contém a sequência de aminoácidos do pré-pró-Lip2-LC (= 250 Aa; PM = 27.011 Da; pl = 8,46) (SEQ ID NO: 33). A figura 31A contém a sequência nucleotídica do pré-pró-Lip2-HC sintético (= 1458 pb) (SEQ ID NO: 34). A figura 31B contém a sequência de aminoácidos do pré-pró-Lip2-HC (= 486 Aa; PM = 52.853 Da; pl = 8,65) (SEQ ID NO: 35). As sequências de codificação de pré-pró-Lip2-HC e-LC foram introduzidas no mesmo vetor, chamado pYLHp4L2pré-pró-HerHC/LC (GUT2ex)-ori2.

25 Exemplo 12: expressão de HC e LC do anticorpo anti-HER2 em cepa de *Yarrowia lipolytica* com um grau variado de glicoengenharia

O plasmídeo pYLHp4L2pré-pró-HerHC&LC (GUT2ex)-ori2 foi digerido com NotI e o cassete de expressão HC-/LC-tandem foi isolado antes de transformar as cepas G045, G057, G061 e G071 de *Yarrowia lipolytica* (ver a Tabela 2). Os transformantes que contêm o cassete de expressão de HC-/LC randomicamente integrado foram selecionados baseados na sua capacidade de crescer em glicerol como a única fonte de carbono. A análise

de expressão de HC e LC foi feita através de western blotting após uma cultura de frasco de agitação de 4 dias dos transformantes selecionados em meio rico contendo glicerol como a única fonte de carbono (meio SuperT/glicerol: extrato de levedura 0,5%; extrato de malte 2%; triptona 1%; glicerol 1,5%; tampão de fosfato 200 mM pH 6,8). A detecção de LC foi realizada usando um camundongo monoclonal para Cadeias leves Livres Kappa (4C11) (Abcam) enquanto a detecção de HC foi feita usando IgG de camundongo anti-humano monoclonal (cadeia γ específica) (Sigma).

Os N-glicanos do secretoma das cepas de produção de anticorpo anti-HER2 mostrou um perfil similar como as cepas glicoengenheiradas correspondentes que não expressavam nenhum HC e LC (figura 32). As porcentagens de N-glicanos em cepas com contexto G045, G057, G061 e G071 foram determinadas após uma cultura de frasco de agitação de 6 dias no meio SuperT/glicerol. Em um contexto G045, 54,6% dos N-glicanos eram Man₃GlcNAc₂. No contexto G057, 47,5% dos N-glicanos eram GlcNAc₁Man₃GlcNAc₂. Em um contexto G061, 58,9% dos N-glicanos eram GlcNAc₁Man₃GlcNAc₂. Em um contexto G071, 37,6% dos N-glicanos eram GlcNAc₂Man₃GlcNAc₂.

Exemplo 13: fermentação da cepa G096 de *Yarrowia*, uma cepa sintetizada de GlcNAc₂Man₃GlcNAc₂ expressando HC e LC do anticorpo anti-HER2

Vários transformantes pYLhp4L2pré-pró-HerHC&LC (GUT2ex)-ori2 de G071 de *Yarrowia lipolytica*, uma cepa capaz de sintetizar GlcNAc₂Man₃GlcNAc₂, foram analisados para níveis de expressão de HC e LC. Um destes clones, G096, foi escolhido para análise adicional.

A fermentação foi feita em um biorreator de tanque agitado de 14 litros (MAVAG AG) munido de um controle do processo e sistema de gestão (Lucillus PIMS). A pressão de oxigênio parcial relativa no meio, as concentrações de CO₂ e O₂ no gás de escape, valor de pH, temperatura, sobrepressão do reator, peso do reator, peso da alimentação e peso base foi tudo monitorado em tempo real. A geração de espuma foi neutralizada adicionando polipropilenoglicol como agente antiespumante (PPG). Os ajustes no pH foram feitos pela adição de uma solução de amônia 25% ou por uma solução

de ácido fosfórico 8,5%.

Uma cultura semente de G096 foi cultivada a 28°C em um frasco de agitação contendo meio rico. A cultura semente foi inoculada no fermentador contendo meio mineral para começar uma fase de batelada a 28°C com crescimento irrestrito, usando somente glicerol como fonte de carbono. Esta fase foi usada para alcançar rapidamente uma alta concentração de biomassa. Daquele ponto para frente, o processo foi deslocado à batelada alimentada exponencial de glicerol (com glicerol como única fonte de carbono e energia; pH 6), com uma taxa de crescimento constante de 0,02. Como um exemplo, os resultados de uma fermentação de batelada alimentada a 28°C são descritos abaixo.

A fase de batelada alimentada durou 148 horas. Em pontos de tempo diferentes da fermentação, as amostras foram tomadas para acompanhar os seguintes parâmetros: 1) expressão de esqueletos proteicos de LC e HC através de western blot; 2) expressão de anticorpo anti-HER2 funcional através de um ELISA; e 3) evolução do perfil de N-glicosilação do secretoma. O nível de expressão de HC completo alcançou um máximo em torno do ponto de tempo 7 (39 horas) e permanece aproximadamente igual dali para a frente. A expressão de LC alcançou um máximo entre os pontos de tempo 7 (39 horas) e 10 (73 horas), mas diminuiu um pouco nos pontos de tempo posteriores. Alguns dímeros LC foram produzidos entre pontos de tempo 5 (25 horas) e 9 (62 horas), mas desapareceram novamente daquele ponto para a frente.

Um ELISA funcional foi desenvolvido para medir a produção do anticorpo anti-HER2 que tem pelo menos um domínio de ligação a antígeno funcional. As placas foram recobertas com uma variante recombinante do antígeno HER2 natural, a quimera ErbB2/Fc humana recombinante (R&D Systems). Então uma diluição do meio, coletado em pontos de tempo diferentes, foi adicionada às placas recobertas. A avaliação da quantidade da proteína de ligação a antígeno foi feita usando um anticorpo kappa anti-humano HRP-conjugado LC (Sigma). A evolução da quantidade da proteína de ligação à quimera ErbB2/Fc (uma medida da quantidade do anticorpo

anti-HER2 funcional secretado) dentro da fermentação de batelada alimentada é mostrada na figura 33. Os dados mostram um aumento gradual nos níveis do anticorpo anti-HER2, com um máximo de 10 a 12 mg/L no fim da fase de produção.

5 A análise de N-glicano foi feita em amostras tomadas em vários pontos de tempo durante a fermentação de batelada alimentada. Os resultados são mostrados nas figuras 34A e 34B. No início da fase de batelada alimentada, houve uma quantidade significante de N-glicanos presentes contendo glicose. Do ponto de tempo 6 progressivamente (34 horas após partida
10 da alimentação exponencial), o nível de N-glicanos glicosilados reduziu significativamente com apenas algum deixado no momento da coleta (ponto de tempo 18, 148 horas). Isto indicou que as proteínas que originalmente carregam N-glicanos contendo glicose, foram diluídas até ao fim da fermentação.
15 Naquele ponto aproximadamente 86% dos N-glicanos isolados do secretoma tinham a estrutura GlcNAc₂Man₃GlcNAc₂.

OUTRAS MODALIDADES

Embora a invenção tenha sido descrita em conjunto com a descrição detalhada da mesma, a descrição precedente é destinada para ilustrar e não limitar o escopo da invenção, que é definida pelo escopo das reivindicações acrescentadas. Outros aspectos, vantagens e modificações estão dentro do escopo das seguintes reivindicações.
20

REIVINDICAÇÕES

1. Método de produzir uma célula transformada de *Yarrowia lipolytica* capaz de produzir proteínas compreendendo GlcNAcMan₃GlcNAc₂ N-glicanos, caracterizado pelo fato de que compreende:

5 a) fornecer uma célula transformada de *Yarrowia lipolytica* geneticamente manipulada para produzir proteínas compreendendo Man₃GlcNAc₂ N-glicanos; e

10 b) introduzir na dita célula um ácido nucleico que codifica uma GlcNAc-transferase I, em que o dito ácido nucleico compreende uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a GlcNAc-transferase I codificada para um compartimento intracelular, em que a expressão da dita GlcNAc-transferase I na dita célula fúngica produz proteínas compreendendo GlcNAcMan₃GlcNAc₂ N-glicanos.

15 2. Método, de acordo com a reivindicação 1, caracterizado pelo fato de que apresenta qualquer uma das características, de (i) a (xv), abaixo:

20 (i) o método compreende ainda a introdução na dita célula de um ácido nucleico codificando uma proteína-alvo, em que a dita célula produz a dita proteína-alvo modificada para compreender os ditos GlcNAc-Man₃GlcNAc₂ N-glicanos;

25 (ii) a dita proteína-alvo liga-se a um receptor Fc;
 (iii) a dita proteína-alvo é um anticorpo ou fragmento do mesmo;
 (iv) a dita proteína-alvo é uma glicoproteína terapêutica;
 (v) a dita proteína-alvo é Interferon-β, GM-CSF, Interferon-γ ou eritropoietina;

30 (vi) o dito compartimento intracelular é o aparelho de Golgi;
 (vii) a dita célula geneticamente manipulada para produzir proteínas compreendendo Man₃GlcNAc₂ N-glicanos é deficiente na atividade de ALG3 e compreende um ácido nucleico que codifica uma α-1,2-manosidase, em que o dito ácido nucleico compreende uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a α-1,2-manosidase codificada para o retículo endoplasmático;
 (viii) o método, de acordo com o item (vii) acima, em que a célula

que é geneticamente manipulada para a produção de proteínas compreendendo $\text{Man}_3\text{GlcNAc}_2$ N-glicanos é deficiente ainda na atividade OCH1;

5 (ix) o método, de acordo com o item (vii) ou (viii) acima, em que a célula que é geneticamente manipulada para a produção de proteínas compreendendo $\text{Man}_3\text{GlcNAc}_2$ N-glicanos compreende ainda um ácido nucleico codificando uma α -1,3-glicosiltransferase;

(x) a α -1,3-glicosiltransferase é ALG6;

10 (xi) o método compreende ainda a introdução na dita célula de um ácido nucleico que codifica uma GlcNAc-transferase II, em que o dito ácido nucleico que codifica a dita GlcNAc-transferase II compreende uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a GlcNAc-transferase II codificada para um compartimento intracellular, em que a expressão da dita GlcNAc-transferase II na dita célula produz proteínas compreendendo $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos;

15 (xii) o método compreende ainda a introdução na dita célula de um ácido nucleico que codifica uma galactosiltransferase, em que o dito ácido nucleico que codifica a dita galactosiltransferase compreende uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a galactosiltransferase para o aparelho de Golgi, em que a expressão da dita galactosiltransferase na dita célula fúngica produz proteínas compreendendo $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$ ou $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos;

20 (xiii) o método, de acordo com o item (xii) acima, em que o referido método compreende ainda a introdução na dita célula de um ácido nucleico que codifica uma proteína-alvo, em que a dita célula produz a dita proteína-alvo modificada compreendendo o dito $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$ ou $\text{Gal}_2\text{GlcNac}_2\text{Man}_3\text{GlcNac}_2$ N-glicanos;

25 (xiv) o método, de acordo com o item (xii) acima, em que o dita galactosiltransferase é uma fusão de uma UDP-Glc-4-epimerase e um domínio catalítico de uma β -1,4-galactosiltransferase I; ou

30 (xv) o método, de acordo com o item (xii) acima, em que o referido método compreende ainda a introdução na dita célula de um ácido nu-

cleico que codifica as subunidades α e β de uma glicosidase II, em que a expressão das ditas subunidades α e β da dita glicosidase II na dita célula produz proteínas compreendendo GalGlcNAcMan₃GlcNAc₂ ou Gal₂GlcNAc₂Man₃GlcNac₂ N-glicanos.

5 3. Método de produzir um célula transformada de *Yarrowia lipolytica* capaz de produzir proteínas compreendendo GlcNAcMan₅GlcNac₂ N-glicanos, caracterizado pelo fato de que compreende:

a) fornecer uma célula *Yarrowia lipolytica* geneticamente manipulada para produzir proteínas compreendendo Man₅GlcNAc₂ N-glicanos; e

10 b) introduzir na dita célula um ácido nucleico codificando uma GlcNAc-transferase I, em que o dito ácido nucleico compreende uma sequência de nucleotídeos que codifica uma sequência de direcionamento para direcionar a GlcNac-transferase I codificada para um compartimento intracellular, em que a expressão da dita GlcNAc-transferase I na dita célula produz proteínas compreendendo GlcNAcMan₅GlcNAc₂ N-glicanos.

15 4. Método, de acordo com a reivindicação 3, caracterizado pelo fato de que apresenta qualquer uma das características, de (i) a (x), abaixo:

(i) o método compreende ainda introduzir na dita célula um ácido nucleico que codifica uma proteína-alvo, em que a dita célula produz a dita proteína-alvo modificada que compreende a dita GlcNAcMan₅GlcNAc₂ N-glicanos;

(ii) o método, de acordo com o item (i) acima, em que a proteína-alvo apresenta qualquer uma das características, de a) a d), abaixo:

a) liga-se a um receptor Fc;

25 b) é um anticorpo ou fragmento deste;

c) é uma glicoproteína terapêutica; ou

d) é um interferon- β , GM-CSF, interferon γ ou eritropoietina;

(iii) o compartimento intracelular é o aparelho de Golgi;

30 (iv) a célula que é geneticamente manipulada para produzir proteínas compreendendo Man₅GlcNAc₂ N-glicanos é deficiente na atividade OCH1 e compreende um ácido nucleico codificando uma α -1,2-manosidase, em que o dito ácido nucleico que codifica a dita α -1,2-manosidase compre-

ende uma sequência de nucleotídeos que codifica uma sequência de direcionamento para direcionar a α-1,2-manosidase para o retículo endoplasmático;

- 5 (v) o método, de acordo com o item (iv) acima, em que a dita sequência alvo é uma sequência HDEL;

(vi) o método compreende ainda a introdução na dita célula de um ácido nucleico que codifica uma manosidase II, em que o dito ácido nucleico que codifica a dita manosidase II compreende uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a ma-

- 10 nosidase II codificada para o aparelho de Golgi, em que a expressão da dita manosidase II na dita célula fúngica produz proteínas compreendendo GlcNAcMan₃GlcNAc₂ N-glicanos;

(vii) o método compreende ainda a introdução na dita célula de um ácido nucleico que codifica uma galactosiltransferase, em que o dito áci-

- 15 do nucleico que codifica a dita galactosiltransferase compreende uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a galactosiltransferase codificada para o aparelho de Golgi, em que a expressão da dita galactosiltransferase na dita célula produz a dita proteína-alvo modificada para compreender GalGlcNAcMan₅GlcNAc₂ ou
- 20 GalGlcNAcMan₃GlcNAc₂ N-glicanos;

(viii) o método, de acordo com o item (vii) acima, compreendendo ainda introduzir na dita célula um ácido nucelico codificando uma proteína alvo, em que a dita célula produz a dita proteína-alvo modificada que compreende os ditos GalGlcNAcMan₅GlcNAc₂ ou GalGlcNAcMan₃GlcNAc₂ N-

- 25 glicanos;

(ix) o método, de acordo com o item (vii) acima, em que a dita galactosiltransferase é uma fusão de uma UDP-Glc-4-epimerase e o domínio catalítico de β-1,4-galactosiltransferase I; ou

- 30 (ix) o método, de acordo com o item (viii) acima, compreendendo ainda isolar a dita proteína-alvo modificada para compreender os ditos GalGlcNAcMan₅GlcNAc₂ ou GalGlcNAcMan₃GlcNAc₂ N-glicanos.

5. Método de produzir uma proteína-alvo compreendendo Glc-

NAcMan₃GlcNAc₂ N-glicanos, caracterizado pelo fato de que compreende:

- a) fornecer uma célula transformada de *Yarrowia lipolytica* geneticamente manipulada para compreender um ácido nucleico codificando uma GlcNAc-transferase I, uma α-1,2-manosidase e uma manosidase II, em que
- 5 o dito ácido nucleico compreende uma sequência de nucleotídeo codificando a sequência direcionamento para direcionar cada proteína codificada para um compartimento intracelular, em que a dita célula é deficiente na atividade OCH1; e
- b) introduzir na dita célula um ácido nucleico codificando uma
- 10 proteína-alvo, em que a dita célula produz a dita proteína-alvo compreendendo a dita GlcNAcMan₃GlcNAc₂ N-glicanos.

6. Método, de acordo com a reivindicação 5, caracterizado pelo fato de que apresenta qualquer uma das características, de (i) a (ix), abaixo:

- (i) o ácido nucleico codifica a dita α-1,2-manosidase compreende
- 15 um sequência de direcionamento para o reticulo endoplasmático que direciona α-1,2-manosidase codificada para o reticulo endoplasmático;
- (ii) o método, de acordo com o item (iv) acima, em que a dita sequência de direcionamento é uma sequência HDEL;
 - (iii) o ácido nucleico que codifica a dita GlcNAc-transferase I e a
- 20 dita manosidase II compreende as sequências de direcionamento para o aparelho de Golgi que direciona a GlcNAc-transferase I e a dita manosidase II codificadas para o aparelho de Golgi;
- (iv) a proteína-alvo liga-se a um receptor Fc;
 - (v) a proteína-alvo é um anticorpo ou fragmento deste;
- 25 (vi) a proteína-alvo é uma glicoproteína terapêutica;
- (vii) a proteína-alvo é um interferon-β, GM-CSF, interferon γ ou eritropoietina;
 - (viii) o método ainda compreende a introdução na dita célula de um ácido nucleico que codifica uma galactosiltransferase, em que o dito ácido nucleico que codifica a dita galactosiltransferase compreende uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a galactosiltransferase codificada para um aparelho de Golgi, em

que a expressão da dita galactosiltransferase na dita célula produz proteínas modificadas compreendendo GalGlcNAcMan₃GlcNAc₂ N-glicanos; ou

(ix) o método, de acordo com o item (viii) acima, compreende ainda isolar a dita proteína-alvo modificada para compreender o dito GalGlc-
5 NAcMan₃GlcNAc₂ N-glicanos.

7. Método de produção de uma proteína-alvo compreendendo Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos, caracterizado pelo fato de que compreende:

a) fornecer uma célula transformada de *Yarrowia lipolytica* geneticamente manipulada sendo deficiente na atividade ALG3 e compreendendo um ácido nucleico codificando uma GlcNAc-transferase I, GlcNAc-transferase II e uma galactosiltransferase, em que o dito ácido nucleico codifica GlcNac-transferase I, GlcNac-transferase II e uma galactosiltransferase compreende uma sequência de nucleotídeo codificando uma sequência de direcionamento para direcionar cada proteína codificada para um compartimento intracelular; e
10

b) introduzir na dita célula um ácido nucleico codificando uma proteína-alvo, em que a dita célula produz a dita proteína-alvo compreendendo a dita Gal₂GlcNAc₂Man₃GlcNAc₂ N-glicanos.

20 8. Método, de acordo com a reivindicação 7, caracterizado pelo fato de que apresenta qualquer uma das características, de (i) a (iv) abaixo:

(i) a dita célula é deficiente ainda na atividade de OCH1;
25 (ii) o método, de acordo com a reivindicação 7 ou item (i) acima, em que a dita célula compreende ainda um ácido nucleico que codifica a-1,3-glicosiltransferase;

(iii) o método, de acordo com o item (ii) acima, em que o dito ácido nucleico que codifica a dita α-1,3-glicosiltransferase é ALG6; ou

30 (iv) o método, de acordo com a reivindicação 7 ou item (i) ou (ii) acima, em que a dita célula compreende ainda um ácido nucleico que codifica as subunidades α e β de uma Glicosidase II, em que a expressão das ditas subunidades α e β da dita Glicosidase II na dita célula fúngica produz a dita proteína-alvo compreendendo os ditos Gal₂GlcNAc₂Man₃GlcNAc₂ N-gli-

canos.

9. Célula transformada de *Yarrowia lipolytica*, caracterizada pelo fato de que é geneticamente manipulada para produzir glicoproteínas compreendendo GlcNAcMan₃GlcNAc₂ N-glicanos, em que a dita célula fúngica é 5 deficiente na atividade OCH1 e compreende um ácido nucleico codificando um α-1,2-manosidase, um GlcNAc-transferase I e manosidase II, em que o dito ácido nucleico que codifica a dita α -1,2-manosidase, um GlcNAc-transferase I e manosidase II compreende uma sequência de nucleotídeo que codifica uma sequência de direcionamento para direcionar cada proteína 10 codificada para um compartimento celular, em que a expressão de α -1,2-manosidase, GlcNAc-transferase I e manosidase II na dita célula produz proteínas compreendendo GlcNAcMan₃GlcNAc₂ N-glicanos.

10. Célula, de acordo com a reivindicação 9, caracterizada pelo fato de que apresenta qualquer uma das características, de (i) a (v), abaixo:

15 (i) compreende ainda um ácido nucleico codificando uma proteína-alvo, em que a dita célula produz a dita proteína-alvo modificada que compreende a dita GlcNAcMan₃GlcNAc₂ N-glicanos;

20 (ii) compreende ainda um ácido nucleico codificando uma GlcNAc-transferase II, em que o dito ácido nucleico que codifica a dita GlcNAc-transferase II compreende uma sequência de nucleotídeo que codifica uma sequência de direcionamento para direcionar a GlcNAc-transferase II codificada para um compartimento intracelular, em que a expressão da dita GlcNAc-transferase II na dita célula produz proteínas compreendendo GlcNAc₂Man₃GlcNAc₂ N-glicanos;

25 (iii) é a célula, de acordo com o item (ii) acima, em que a dita célula compreende ainda um ácido nucleico que codifica uma galactosil-transferase, em que o dito ácido nucleico que codifica a dita galactosiltransferase compreende uma sequência nucleotídica que codifica uma sequência de direcionamento para direcionar a galactosiltransferase codificada para um 30 aparelho de Golgi, em que a expressão da dita galactosiltransferase na dita célula produz proteínas modificadas compreendendo Gal₂NAc₂Man₃GlcNAc₂ N-glicanos;

(iv) é a célula, de acordo com o item (ii) acima, em que pelo menos 50%, pelo menos 70%, ou pelo menos 85% dos N-glicanos da proteína produzida pela célula são GlcNAc₂Man₃GlcNAc₂ N-glicanos; ou
(v) está na foram isolada.

5 11. Célula transformada de *Yarrowia lipolytica*, caracterizada pelo fato de que é geneticamente manipulada para produzir glicoproteínas compreendendo GlcNAc₂Man₃GlcNAc₂ N-glicanos, em que as ditas células são geneticamente manipuladas para serem deficientes na atividade ALG3 e compreendem um ácido nucleico codificando uma GlcNAc-transferase I e
10 uma GlcNAc-transferase II, em que o dito ácido nucleico que codifica a dita GlcNAc-transferase I e uma GlcNAc-transferase II compreende uma sequência de nucleotídeo codificando uma sequência de direcionamento para direcionar cada a proteína codificada para um compartimento celular, em que a expressão de GlcNAc-transferase I e uma GlcNAc-transferase II na
15 dita célula produz proteínas compreendendo GlcNAc₂Man₃GlcNAc₂ N-glicanos.

12. Célula, de acordo com a reivindicação 11, caracterizada pelo fato de que apresenta qualquer uma das características, de (i) a (vii), abaixo:

- (i) é deficiente na atividade OCH1;
- 20 (ii) é uma célula, de acordo com a reivindicação 11 ou item (i) acima, em que a dita célula geneticamente manipulada compreende ainda um ácido nucleico que codifica α-1,3-glicosiltransferase;
- (iii) é uma célula, de acordo com a reivindicação 11, ou item (i) ou (ii) acima, em que a dita célula geneticamente manipulada compreende ainda um ácido nucleico que codifica uma proteína-alvo, em que a dita célula produz a proteína-alvo modificada compreendendo a dita Glc-
25 NAc₂Man₃GlcNAc₂ N-glicanos;
- (iv) é uma célula, de acordo com a reivindicação 11 ou qualquer um dos itens (i) a (iii) acima, em que a dita célula compreende ainda um ácido nucleico que codifica as subunidades α e β de uma glicosidase II, em que a expressão das ditas subunidades α e β da dita glicosidase II na dita célula produz proteínas compreendendo GlcNAc₂Man₃GlcNAc₂;

(v) é uma célula, de acordo com o item (iv) acima, em que a dita célula comprehende um ácido nucleico que codifica a glicosiltransferase, em que o dito ácido nucleico que codifica a dita galactosiltransferase comprehendo uma sequência de nucleotídeo que codifica uma sequência de direcionamento para direcionar a glicosiltransferase codificada para o aparelho de Golgi, em que a expressão da dita glicosiltransferase na dita célula produz proteínas comprehendendo $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNac}_2$ N-glicanos;

5 (vi) é uma célula, de acordo com a reivindicação 11 ou qualquer um dos itens (i) a (iv) acima, em que pelo menos 50%, pelo menos 70%, ou 10 pelo menos 85% dos N-glicanos das proteínas produzidas pela célula são $\text{GlcNAc}_2\text{Man}_3\text{GlcNac}_2$ N-glicanos; ou

15 (vii) está na forma isolada.

13. Cultura substancialmente pura de células de *Yarrowia lipolytica*, caracterizada pelo fato de que é geneticamente manipuladas para produzir glicoproteínas comprehendendo $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos, em que as ditas células são geneticamente manipuladas para serem deficientes na atividade de ALG3 e compreender um ácido nucleico que codifica uma GlcNAc-transferase I, uma GlcNAc-transferase II e uma galactosiltransferase, em que o dito ácido nucleico que codifica a dita uma GlcNAc-transferase I, uma GlcNAc-transferase II e uma galactosiltransferase comprehende sequências nucleotídicas que codificam sequências de direcionamento para direcionar cada proteína codificada para um compartimento intracelular, em que expressão da dita GlcNAc-transferase I, GlcNAc-transferase II e galactosiltransferase nas ditas células produz proteínas comprehendendo $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos.

20 14. Cultura substancialmente pura, de acordo com a reivindicação 13, caracterizada pelo fato de que apresenta qualquer uma das características, de (i) a (iii), abaixo:

25 (i) as células são deficientes ainda na atividade OCH1;
 (ii) a cultura substancialmente pura é uma cultura, de acordo com a reivindicação 13 ou com o item (i) acima, em que a dita célula comprehende ainda um ácido nucleico que codifica α -1,3-glicosiltransferase; ou

(iii) a cultura substancialmente pura, de acordo com a reivindicação 13 ou os itens (i) ou (ii) acima, em que a dita célula comprehende ainda um ácido nucleico que codifica as subunidades α e β da glicosidase II, em que a expressão das ditas subunidades α e β da glicosidase II na dita célu-
5 las fúngica produz a dita proteína comprehendendo a dita $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos.

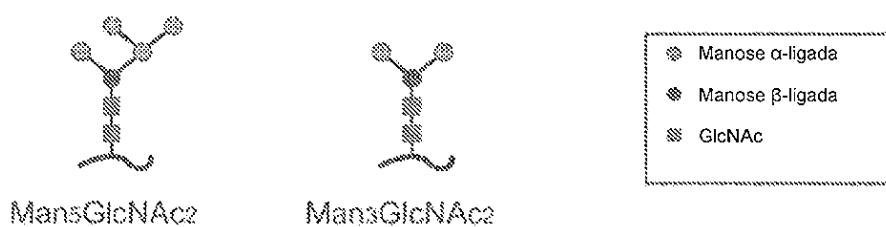
15. Cultura substancialmente pura de células de *Yarrowia lipolytica*, caracterizada pelo fato de que são geneticamente manipuladas para produzir glicoproteínas comprehendendo $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos, em que a dita célula é geneticamente manipulada para ser deficiente na atividade de OCH1 e comprehende um ácido nucleico que codifica a-
10 1,2-manosidase, uma GlcNAc-transferase I, uma manosidase II, uma Glc-
NAc-transferase II e uma galactosiltransferase, em que o dito ácido nucleico
que codifica a dita α-1,2-manosidase, a dita GlcNAc-transferase I, a dita ma-
nosidase II, a dita GlcNAc-transferase II e a dita galactosiltransferase compreende sequências nucleotídicas que codificam sequências de direciona-
15 mento para direcionar cada proteína codificada para um compartimento in-
tracelular, em que expressão da dita α-1,2 manosidase, a dita GlcNAc-
transferase I, a dita manosidase II, a dita GlcNAc-transferase II, e a dita ga-
20 lactosiltransferase nas ditas células produz proteínas comprehendendo
 $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos.

16. Composição, caracterizada pelo fato de que comprehende uma glicoproteína, pelo menos 50%, pelo menos 70%, ou pelo menos 85% dos N-glicanos na dita glicoproteína são $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ N-glicanos e
25 em que as glicoproteínas foram produzidas usando uma célula transformada de *Yarrowia lipolytica* geneticamente manipulada.

17. Invenção, em quaisquer formas de suas concretizações ou em qualquer categoria aplicável de reivindicação, por exemplo, de produto ou de processo ou uso englobadas pela matéria inicialmente descrita, reve-
30 lada ou ilustrada no pedido de patente.

FIG. 1

A.



B.

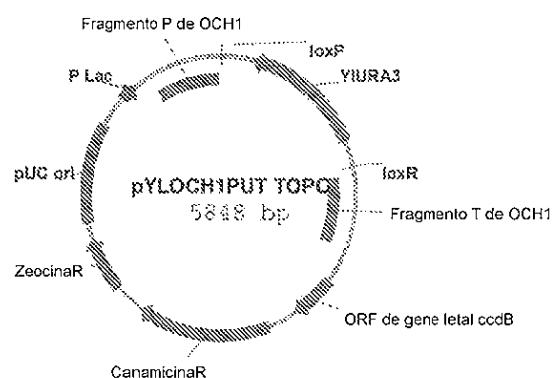


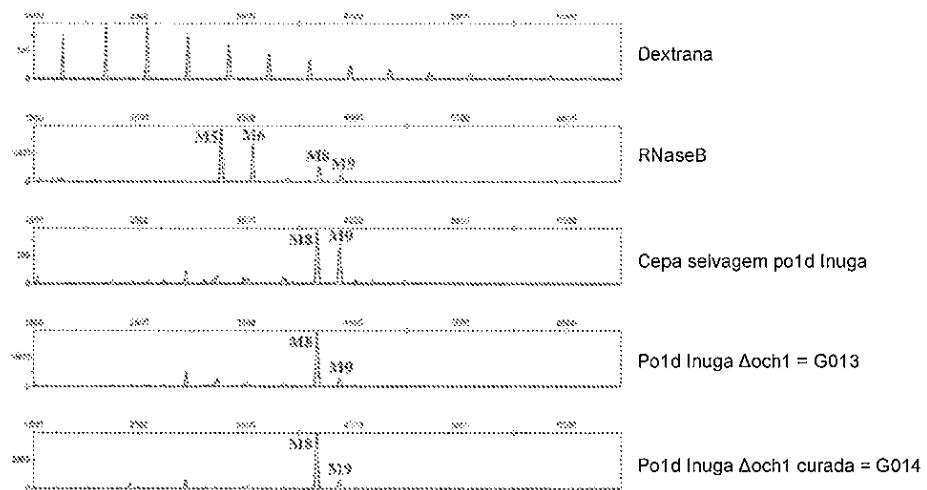
FIG. 2

FIG. 3

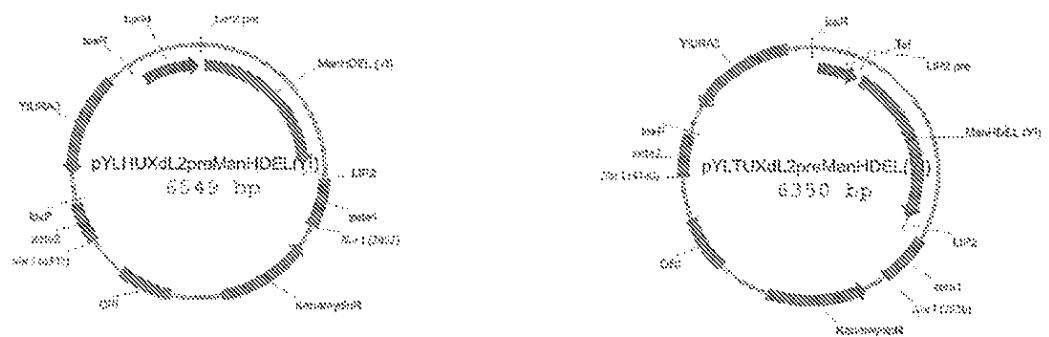
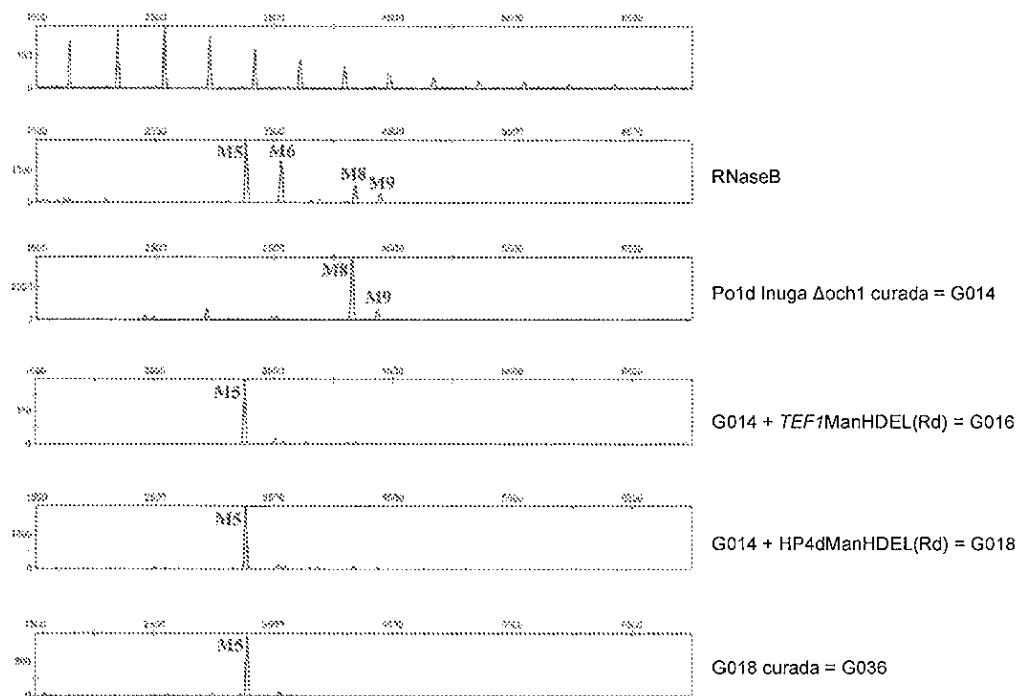
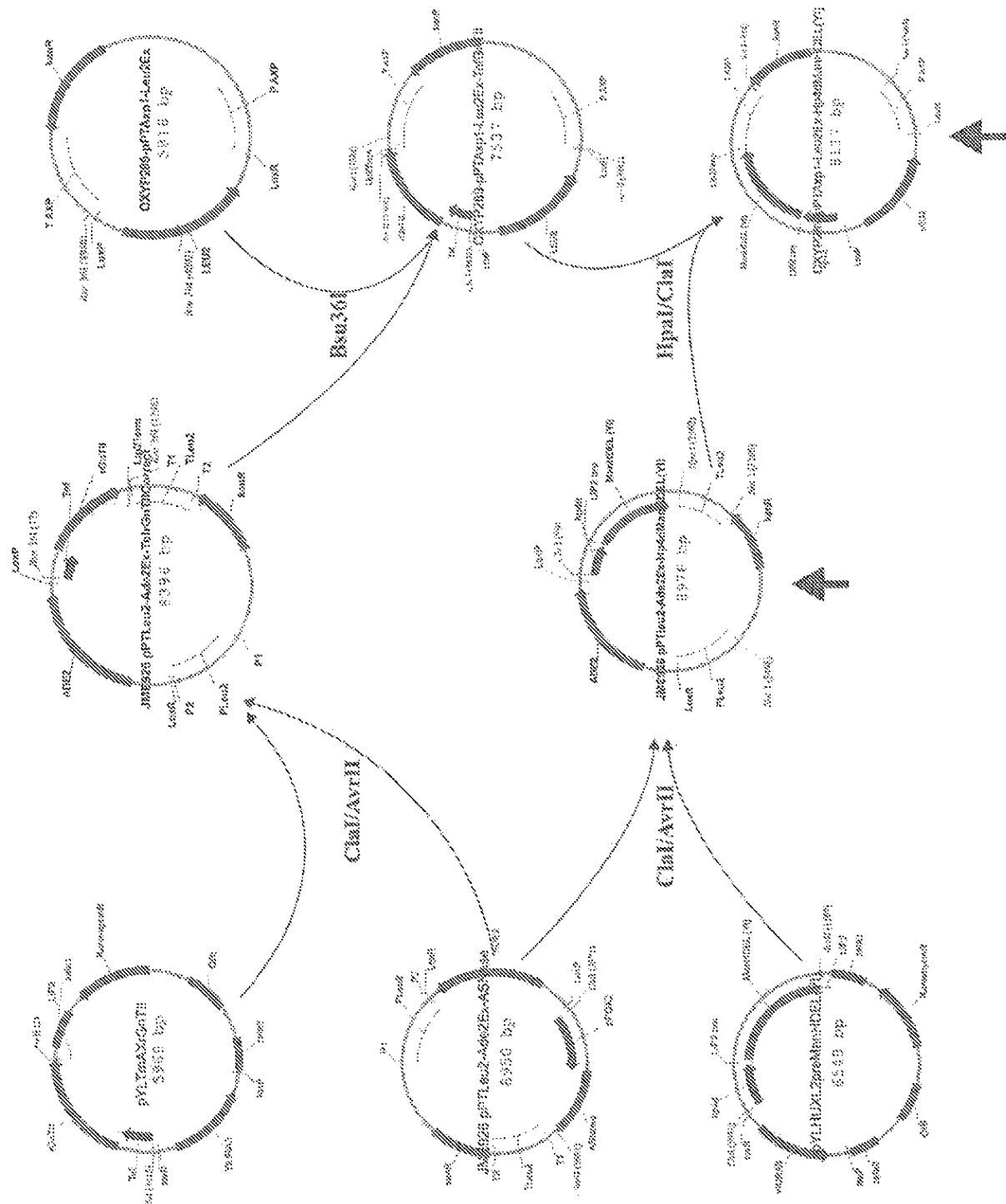


FIG. 4



5
FIG.

FIG. 6

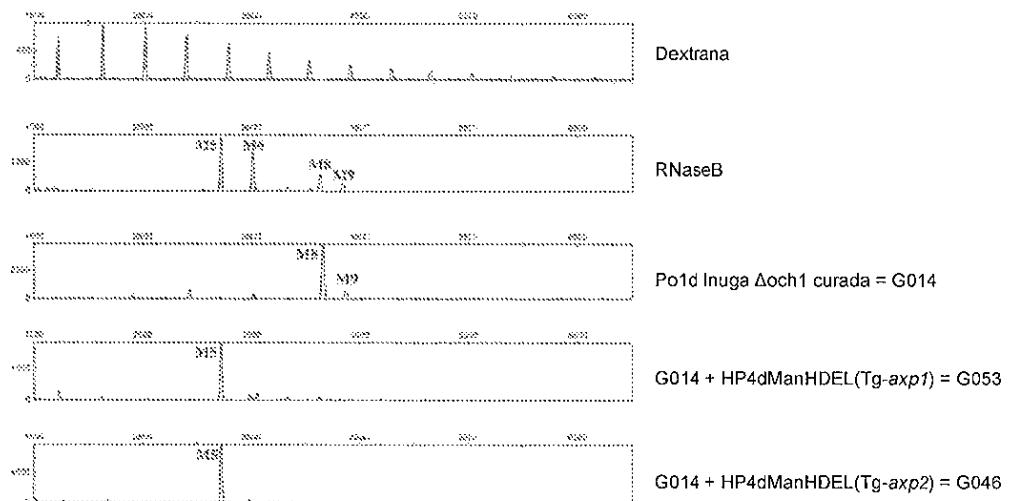


FIG. 7

FIG. 8

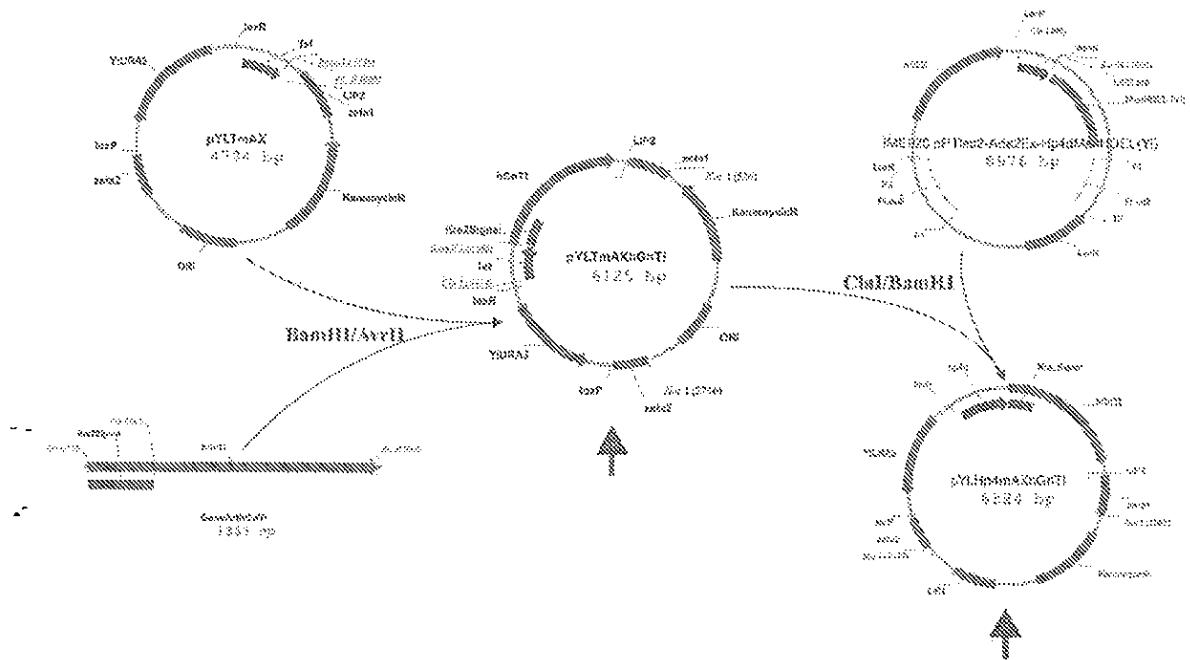


FIG. 9

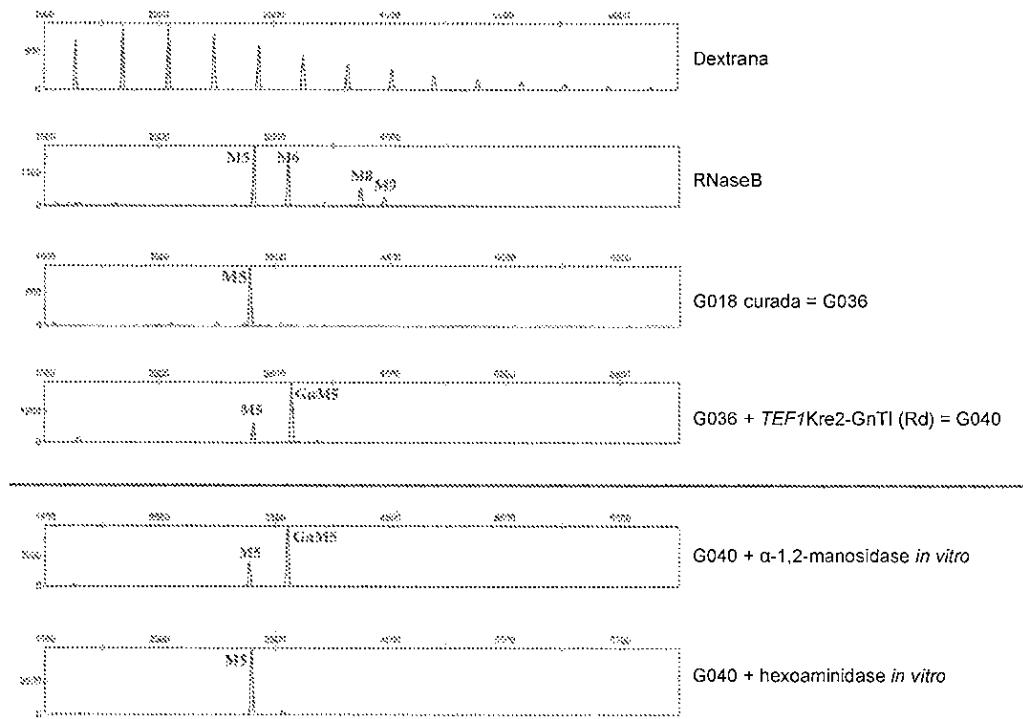


FIG. 10

MILTKRFGKLFKLTFLVUTLCLGFVITNKGYMDENTSRRDPFIRPPLKVARSPPGQCQDVV
 QDVPNVDVQMLLELYDRMSFKDIQGGVWKQGWNIKYDPLKYNAAHKLKVFFVPHSHNDPGW
 IQTFFEEYYQHDTKHILSNALRHLHDNPEMKFIWAEISYFARFYHDLGENKKLQMKSIVKN
 GQLEFVTGGWVMPDEANSWRNVLQLTEGQTLKQFMNVPTASWAIOPFGHSPTMPYI
 LQKSGFKNMLIQRTHYSVKKELAQOROLEFLWRQIWWDNKGDITALFTHMMPFYSYDIPHTC
 GDPDKVCCQFDKRMGSFGLSCPWKPPRTISDQNVAARS DLLVDQWKKKAELYRTNVLL
 IPILGDDDFRKQNTEDVQRVNYERLFEHINSQAHQAFNVAQFGTIQLQEYFDAVHQAEAGQA
 EFPFLSGDFFTYADRSNDYWSGYTSRPYHCRMIRVLMHYVRAAEMLS AWHSDGMARIE
 ERLEQARRELSLPQHHDDGITGTAKTHVVVDYEQRMQEALKACQMVMQQS VYRLLTKPSIY
 SPDFSF SYFTLDDSRWP GSGVEDSRTTI ILGEDILPSKHVVMHNTLPHWREQLVDFYVSS
 FFVSVTDLANNPVEAQVSPVWSWHHDTLTKTIHPQGSTPKYRIIEFKARVPPMGLATYVLT
 ISDSKPEHTSYASNLLRKNP TSLPQYEPDVKF G DPREISLRVGNNGPTLAFSEQGLK
 SIQLTQDSPHVPVHF KFLKYGVRS HGRSGAYLFLPNGPASPVELGQPVVLVTKGKLESS
 VSVGLPSV VHQTIMRGGAPEIRNLVDIGSLDNTIEIVMRLETHIDSGDI F YTDLNGLQFIK
 RRLDKLPLQANYYPI PSCMFIEDANTRLTLLTGQPLGGSSLASGELE IMQDRRIASDDE
 RGLGQGVLDNKPVLIHYRIVLEKVNNCVRPSKLHPAGYLTSAAHKASQSLLDPLDKFIFA
 ENEWIGAQGQF GGHDHPSAREDLDVSVMRRLTKSSAKTQRVGYVLHRTNLMQCGTPEEHTQ
 KLDVCHLLPNVARCERTLTFQNL EHLDGMVAPEVCPMETAAYVSSHS (SEQ ID
 NO: 7)

ATGCTGCTGACCAACCGAATTCTCTTACCTGTTCAAGCTGACCTTCATGCTGCTGATCCCTGTTGCTGCTG
 ATCACCAACACGTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
 CCCCCGACCCGGCCAGTGTCAAGGACGTGGTCAAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
 CGAATGCTTTCAGGACATCGAOGGCCGGTGTGAAGCAGGGCTGGAACATCAAGTACGACCCCCCTGAAGTAC
 AACGCCACCACAGCTGAAGGTGTTCTGTGGTCCCCACTCTCACACGACCCCCCTGGATTCAAGCTTCGAG
 GAGTACTACCAGCAACGACACCAAGCACATCTCTAACGCCCTGCGACACCTGCAACGACAACCTGAGATGAAG
 TTATCTGGCCGAGATCTCTTACTCTGCCGATTCTACACGACCTGGCGAGAACLAAGAACTGCAAGATGAAG
 TCTATCTGTGAAGAACGGCCAGCTGGAGTTCTGTGACCGGGCGCTGGGTGATGCCGACGAGGCCACTCTACTGG
 CGAACACGTGCTGCTGCAGCTGACCCGAGGGCCAGACCTGGCTGAAGCAGTTCATGAACGTGACCCCCACUGCTCT
 TGGCCATCGACCCCTTCGCCACTCTCCCACCATGCCCTACATCTGCAAGAAGTCTGGCTCAAGAACATGCTG
 ATCCAGCGAACCCACTACTCTGTGAAGAAGGGAGCTGGCCAGCAGCGACAGCTGGAGTTCTGTGGGAGACAGATC
 TGGGACAAACAGGGCGACACGCCCTGTTCAACACATGATGCCCTCTACTCTACGACATCCCCAACACCTGT
 GGGCCCGACCCCAAGGTGTGTCAGTTCAAGCGAATGGCTCTTCGGCCCTGTCTTGCCCTGGGAAG
 CTGCCCCCTCGAACCATCTCTGCTGATCCCCCTGGCGACGACTTCCGATTCAAGCAGAACACCGAGTGG
 GACGTGCAGCGAGTGAACCTACGGAGCGACTTTGAGACATCAACTCTCAGGCCACTTCAACGTGCAGGCTCAG
 TTGGCACTCTGCAAGGAGTACTTCGAGCGACTCTGCAAGCGCTCCACCAAGGGCGAGCGAGCCGGCAGGCTGG
 TCTGGCACTTTTCACCTACCGCGACCGATCTGACAAACTACTGGCTCTGGCTACTACACCTCTGACCCCTACCA
 AACCGAATGGACCGAGTGTGATCCACTACGAGCCGGCAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTG
 ATGGCCCGAATCGAGGAGCGACTGGAGCAGGGCCGACGAGAGCTGCTCTGTTCCAGCACCACGGCATCACC
 GGCACCGCCAAGACCCACGTGGTGGACTACGAGCAGCGAATGCAAGGAGGCCCTGAAGGCTGTCAAGTGG
 ATGCAGCAGTCTGTCTACCGACTCTGACTAAGCCCTATCTACTCTCCGACTTCTCTTCTCTACTTCACC
 CTGGACGACTCTGATGGCCCGCTCTGGCTGGAGGACTCTGCAACCAACCATCATCTGGCGAGGACATCTG
 CCCTCTAAGCACGTGGTGTGACAACACCCCTGCCCACTGGCGAGGAGCAGCTGGTCACTTACGTGTCT
 CCCTCGTGTGTGACCGACCTGGCCAAACACCCCGTGGAGGCCAGGTGTCTCCGTGTGGCTTGGCACCAC

GACACCCGTGACCAAGACCATCCACCCCCCAGGGCTCTACCAAGTACCGAATCATCTCAAGGCCCGAGTGCCC
CCCATGGGCCTGGGCACCTACGTGCTGACCATCTCCGACTCTAAGCCGAGCACACCTCTTACGCCCTPAACCTG
CTGCTGGCAAAGAACCCCCACCTCTCTGCCCCCTGGGCCAGTACCCCGAGGACGTGAAGTTCGGCGACCCCCGAGAG
ATCTCTCTGCGAGTGGGCAACGGCCCCACCCCTGGCCTTCTGAGCAGGGCCTGCTGAAGTCTATCCAGCTGACC
CAGGACTCTCCCCACGTGCCCCGTGCACTTCAAGTTCTGAGTACGGCGTGCATCTCACGGGACCCGATCTGGC
GCCTACCTGTTCTGCCAACGGACCGGCTCTCCCGTGGAGCTGGGACAGCCCGTGGTGCCTGGTGAACCAAGGGC
AAGCTGGAGTCTCTGTCCTGTCAGCTGGCACCCAGACCATCATGGGAGGGGAGCCCCCGAG
ATCCGAAACCTGGTGGACATCGGATCTCTGGACAACACCGAGATCGTGAATGCGACTGGAGACCCACATCGACTCT
GGCGACATCTTCAACCCGACCTGAAACGGCTTGCAAGTTCATCAAGCGACGACGACTGGACAAGCTGCCCTGCA
GCCAAGTACTACCCCCATCCCCCTGGCATGTTCATCGAGGACGCCAACACCCGACTGACCCCTGCTGACGGGCCAG
CCCCCTGGCGGATCTTCTGCCCCCTGGAGCTGGAGATCATGCAGGACCGACGACTGGCCTCTGACGACGAG
CGAGGCCCTGGGCCAGGGCGTGGACAACAAGCCCGTGGCTGCACATCTACCGACTGGTGCCTGGAGAAGGTGAAC
AACTGTGTGCGACCCCTCAAGCTGCAACCCGCTGGCTACCTGACCTCTGCCGCCACAAGGCCCTCTGAGTCTG
CTGGACCCCCCTGGACAACCTCATCTTCCCGAGAACCGAGTGGATGCCGCCCCAGGGCCAGTTCGGAGGGCACCC
CCCTCTGCCCGAGAGGGACCTGACCTGCTGTGATGCCGACGACTGACCAAGTCTGCAAGACCCAGGGAGTG
GGCTACGTGCTGCACCGAACCAACCTGATGCCGACCCCCCGAGGAGCACACCCAGAACGCTGGACGGTCTG
CACCTGCTGCCAACGTCGCCGATGTGAGCGAACCCACCCGACTGACCTTCTGCAAGAACCTGGAGCACCTGGACGGC
ATGGTGGCCCCCGAGGTGTCCCATGGAGACCGCCGCTACGTGTCGTCCCACCTCTCTTAG (SEQ ID
NO:8)

FIG. 10 cont.

FIG. 11

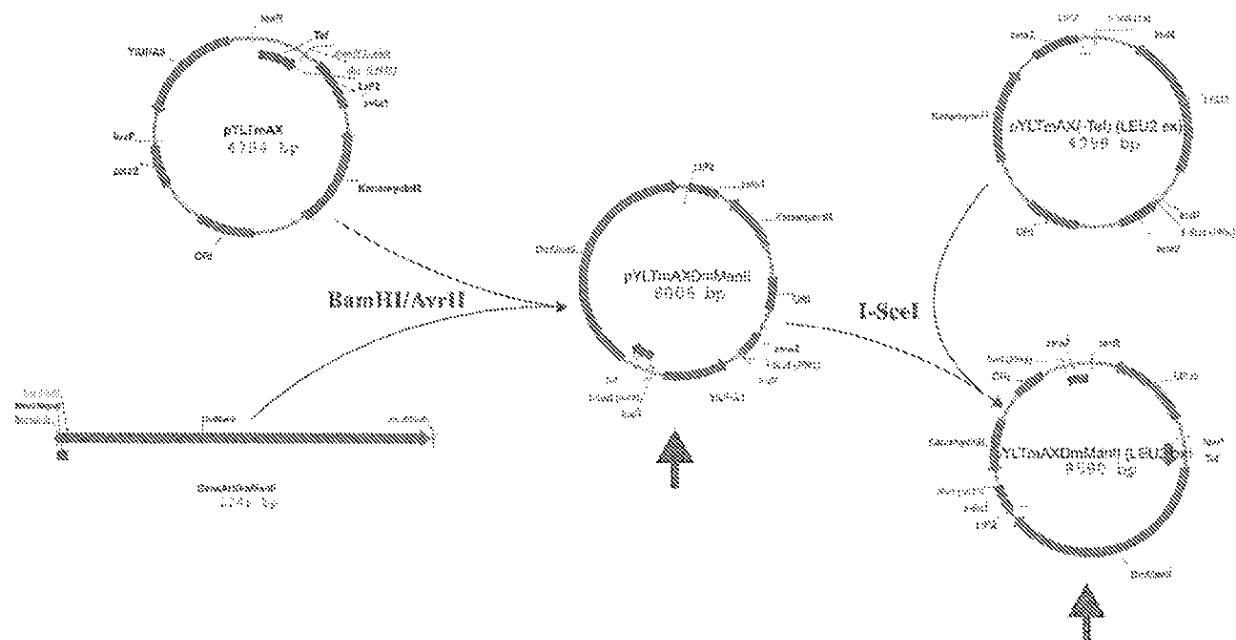


FIG. 12

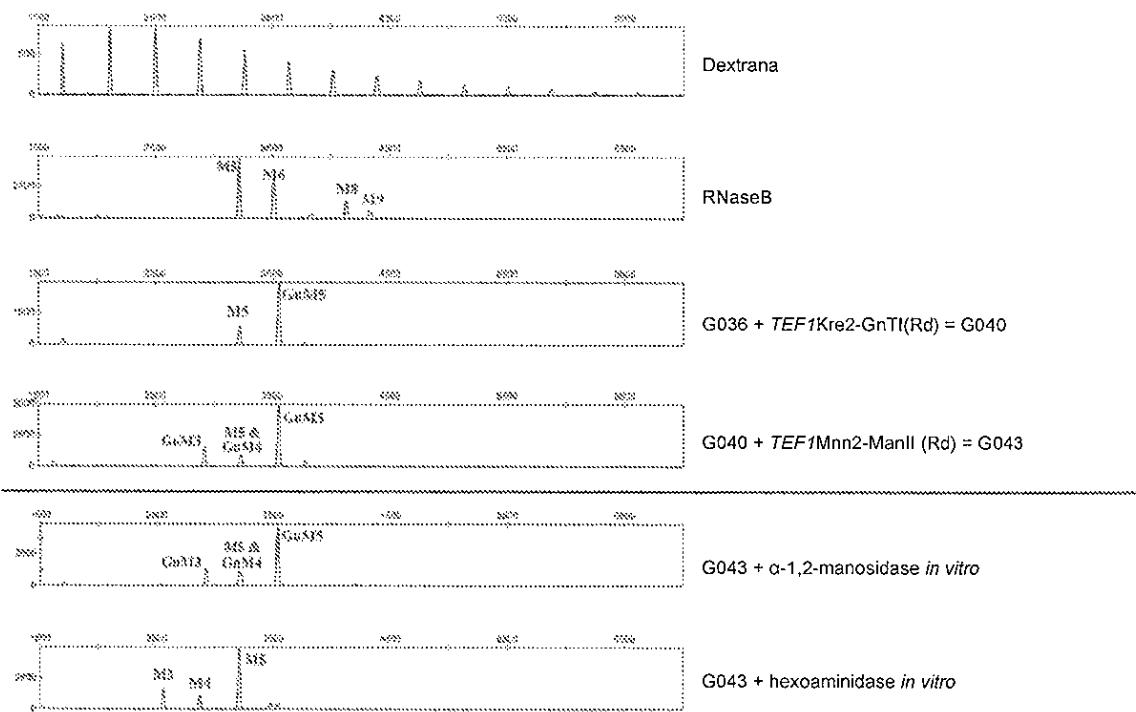


FIG. 13

FIG. 13 conti

FIG. 14

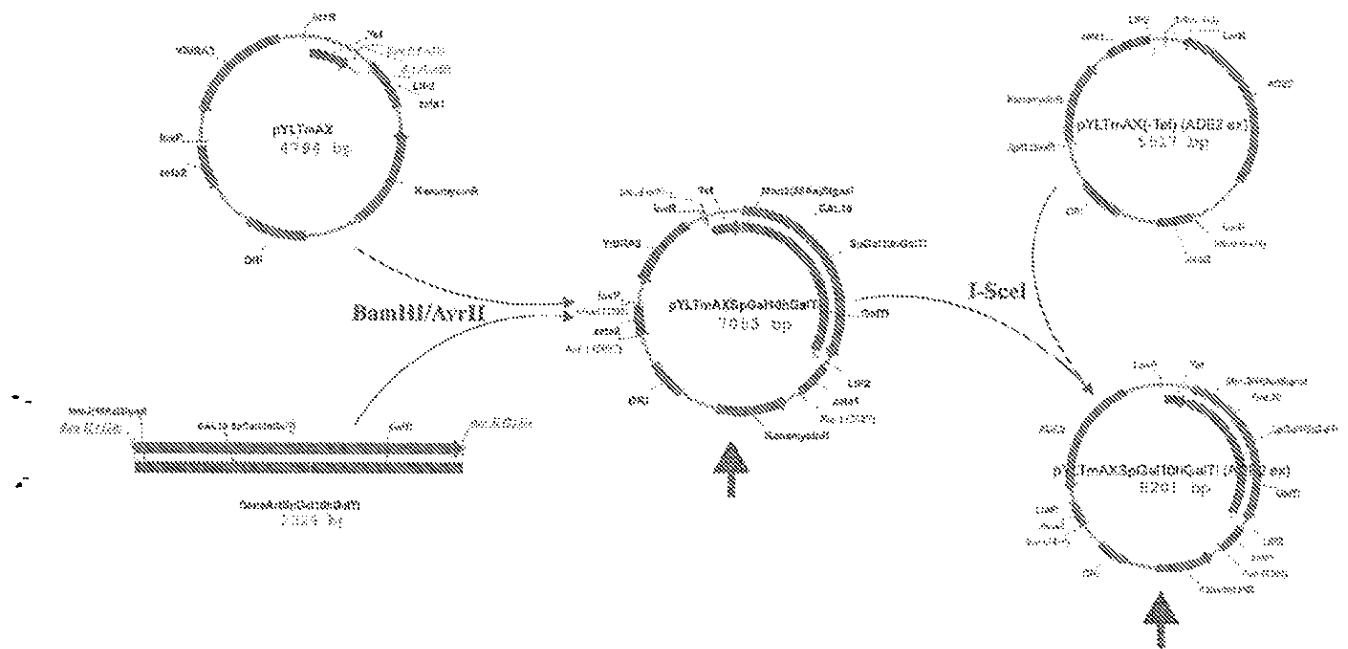


FIG. 15

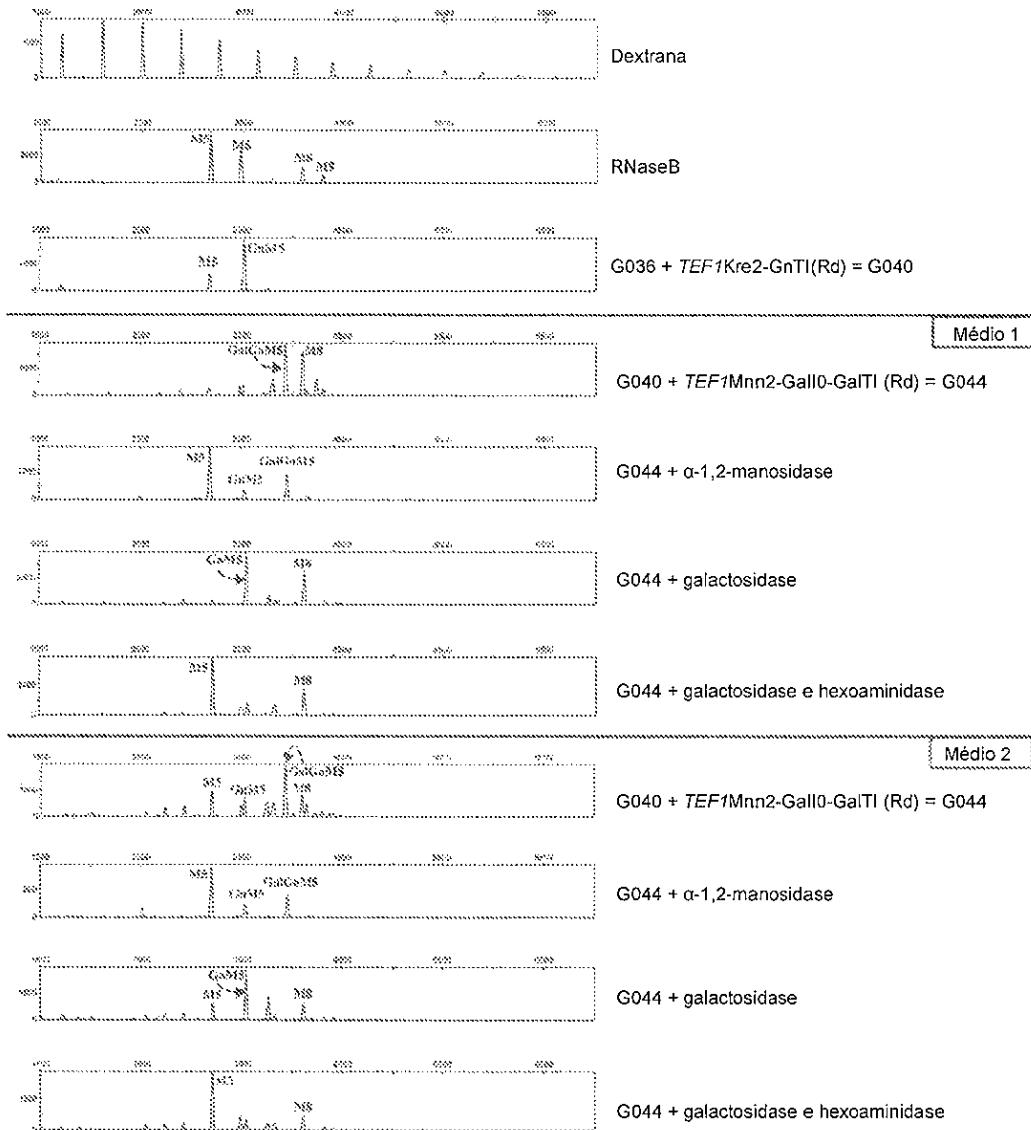


FIG. 16

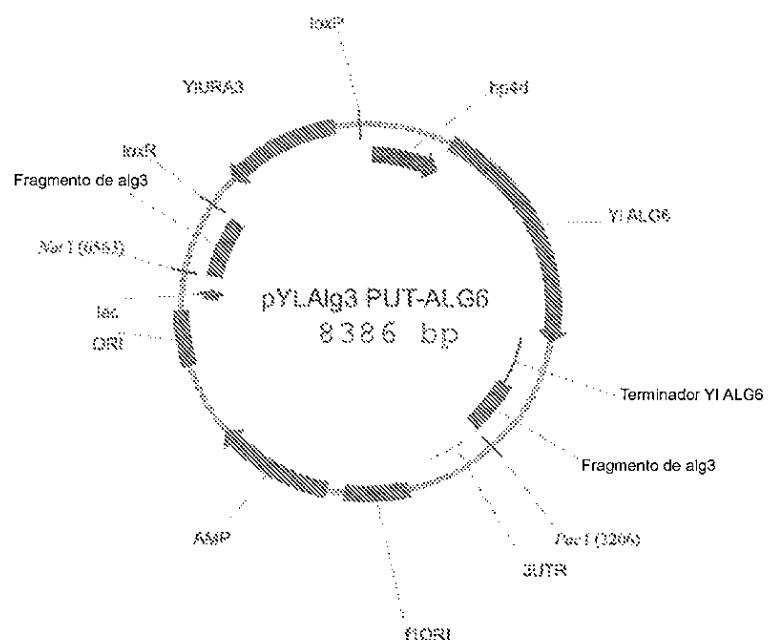


FIG. 17

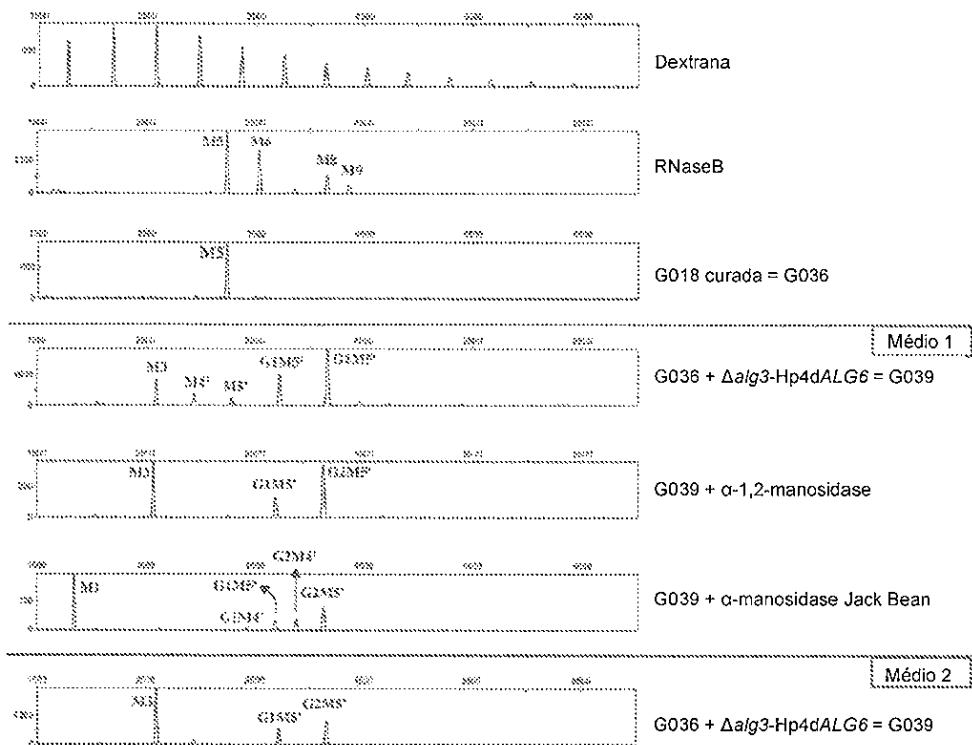
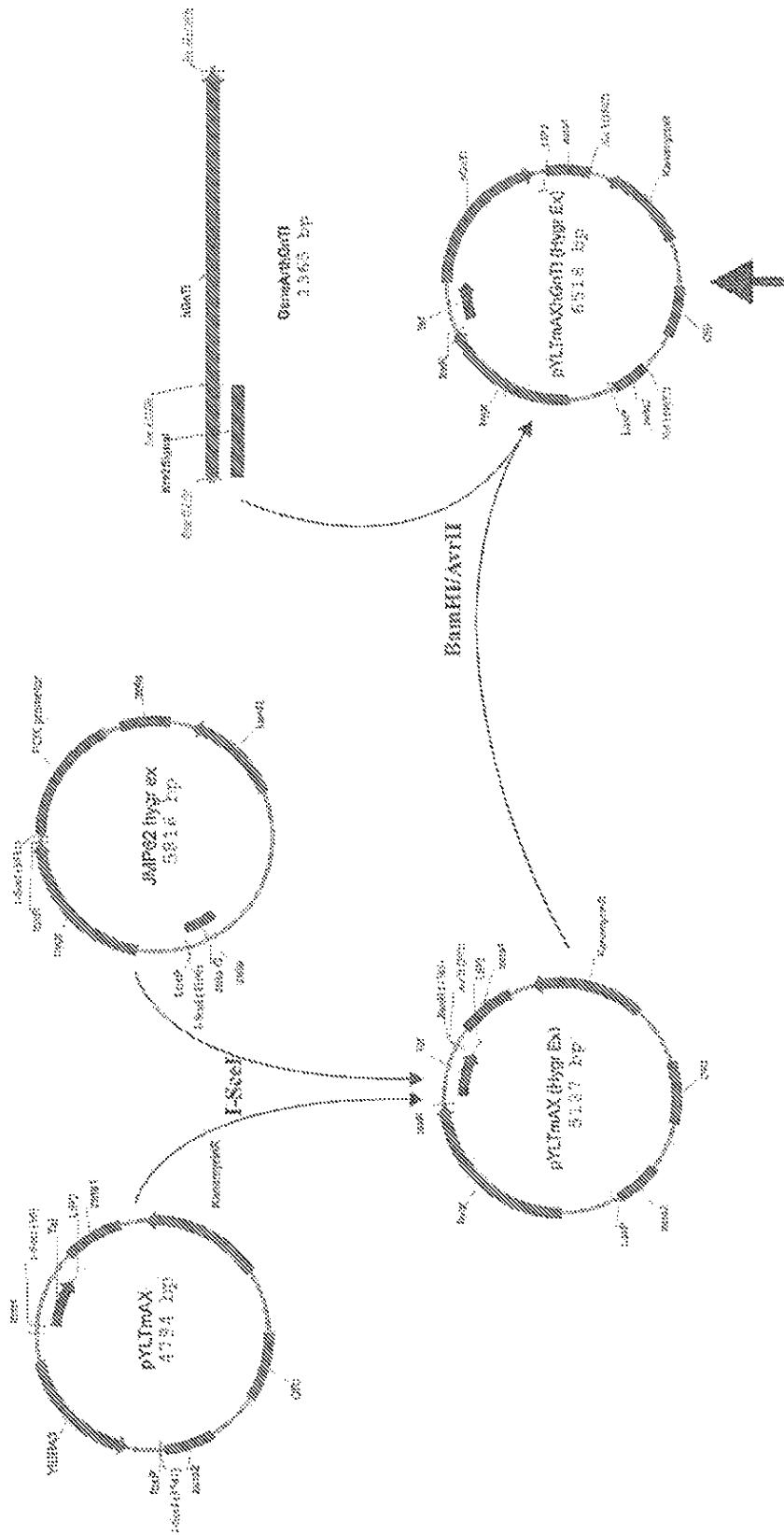


FIG. 18



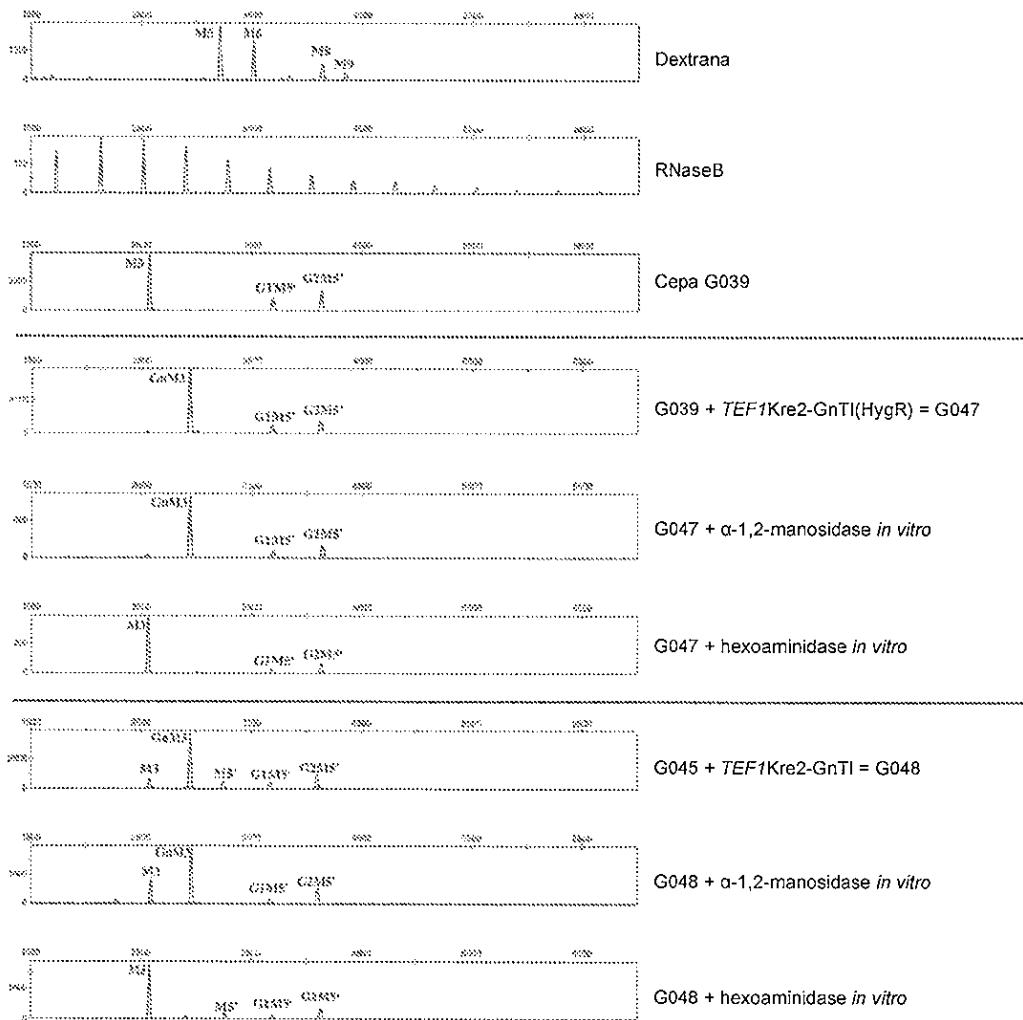
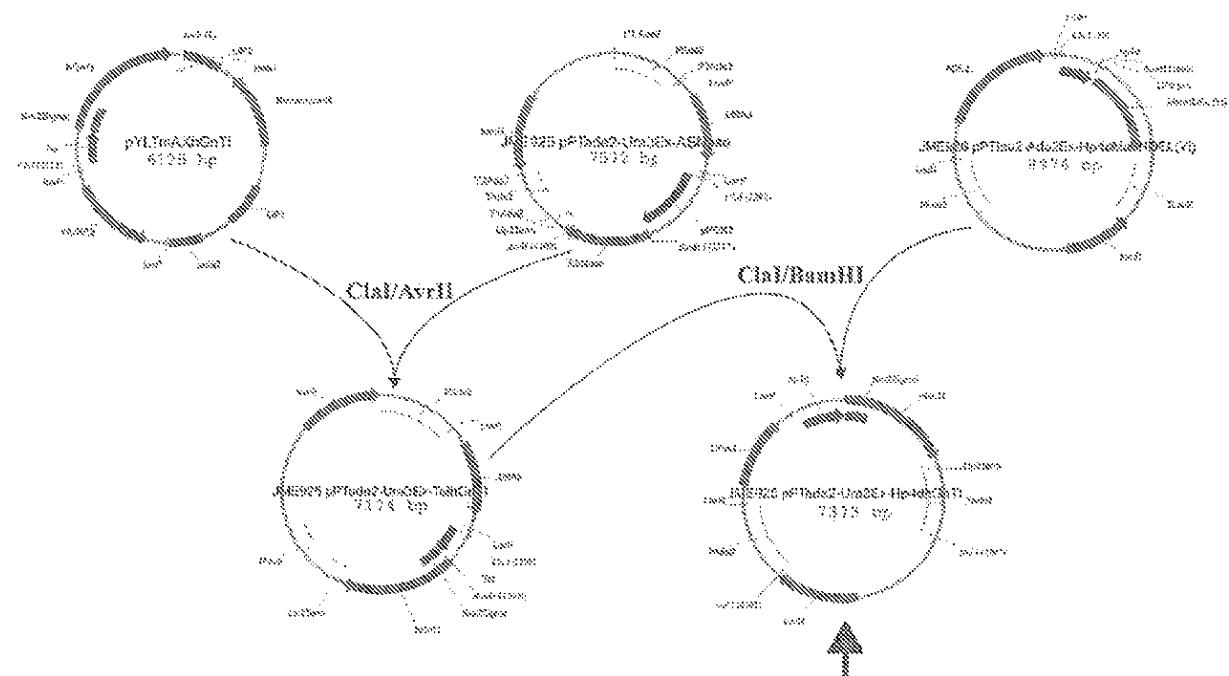


FIG. 19

FIG. 20



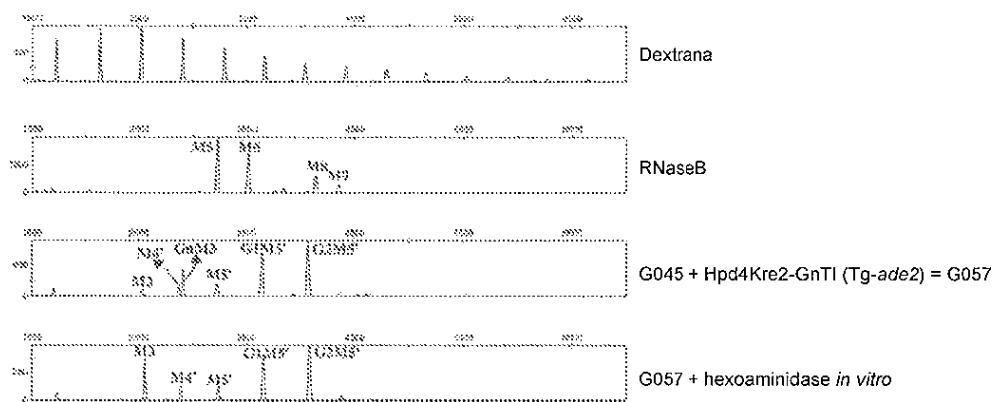


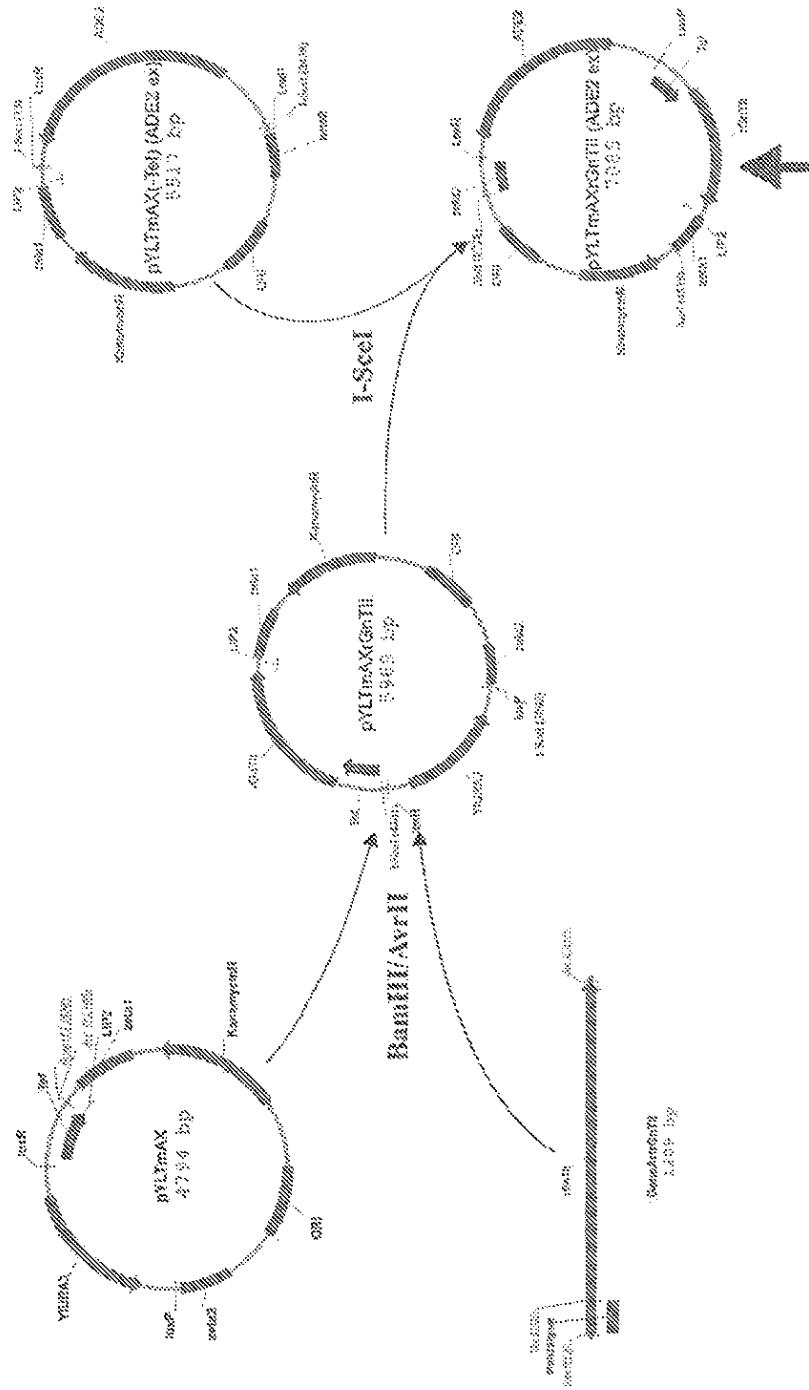
FIG. 21

FIG. 22

MLLTKRESKLFKLTFIVLILCGLFVITNKYMDENTESLVYQLNFQQLRNVDKDGTWSPG
 ELVLVVQVHNRPPEYLRLLIDSRLKAQGIREVLVIFSHDFWSAEINSLISSVDFCPVLQVF
 FPFSTIOLYPSEFPGSDPRDCPDLKKNAALKLCCINAEPDSFGHYREAKFSQTKHWWWW
 KLMFVWERVKVLQDYTGLILFLEEDHLYAPDFYHVFKKMWLKQQECPGCDVLSLGTYTT
 IRSFYGIADKVVDVTKSTEHNMGLALTRDAYQKLIECTDTFCYDDYNWDWTLQYLTIA
 CLPKVWKVLPQAPRIFHAGDCGMHHKKTCRPSTQSAQIESLLNNNQYLFPETLVIGEK
 FPMAAISPPRKNGGWGDIRDHELCKSYRRLQ (SEQ ID NO:17)

AAGCTGGTGAACCAAGCGATTCTCTAAAGCTGTTAACAGCTGACCTTCATCGTGCTGATCCTGTGTCGGCTTGTT
 CGTGATCACCAACAAGTACATGGACGGAGAACACCTCGAGCCTGGTACCCAGCTGAACCTCGACCAGAT
 GCTGGAAACGTGGACAAGGACGGCACCTGGCTCCCGCGAGCTGGTGCCTGGTGCAGGTGCACAA
 CCGACCCGAGTACCTGCGACTGCTGATCGACTCTCTCGAAAGGCCAGGCATCCGAGAGGTGCTGGT
 GATCTTCTCTCACCGACTCTGGTCTGCCAGATCAACCTCCCTGATCTCTCTGTGGACTCTGTCCCTGCT
 GCAGGTGTTCTCCCATTACGACATCCAGCTGTACCCCTCTGAGTTCCCCGGCTGTGACCCCCGAGACTGTC
 CCCGAGACCTGAAGAAGAACGCCGCCCTGAAGCTGGCTGTATCAACGCCAGTACCCGACTCTTCG
 GUCACTACCGAGAGGCCAACCTCTCTCAGACCAAGCACCACCTGGTGGTGAAGCTGCACCTCGTGTGG
 AGCGAGTGAAGGTGCTGCAGGACTACACCCGGCTGATCTGTTCTGAGGAGGACACTACCTGGCC
 CCGACCTCTACCACTGTTCAAGAAGATGTGGAAAGCTGAAGCAGCAGGAGTGTCCCCGGCTGCGACGTGCT
 GTCTCTGGGACCTACACCACCATCCGATCTTCTACGGCATCGCCGACAAAGGTGGACGTCAAGACCTGG
 AAGTCTACCGAGCACACATGGGCTGCCCCCTGACCCGAGATGCCCTACCGAAAGCTGATCGAGTGTACC
 GACACCTCTGTAACCTACGACGACTACAACCTGGGACTGGACTCTGCGAGTACCTGACCCGGCTGCTGC
 CCAAGGTGTTGAAGGTGCTGGGCCAGGCCCCCTGAATCTTCACGCCGGACTGTGGCATGACCA
 CAAGAAGACCTGTCGACCCCTCACCCAGTCTGCCCCAGATCGAGTCTUTGCTGAACAACAAGCAGAAC
 CTGTTCCCTGAGACCCCTGGTATGGCGAGAAGTCCCCATGGGCCATCTGCCCTCCCCGAAAGAAC
 GCGGCTGGGGCGACATCCGAGACCACTCTGTAAGTCTTACCGACGACTGCAGTAG (SEQ ID
 NO:18)

FIG. 23



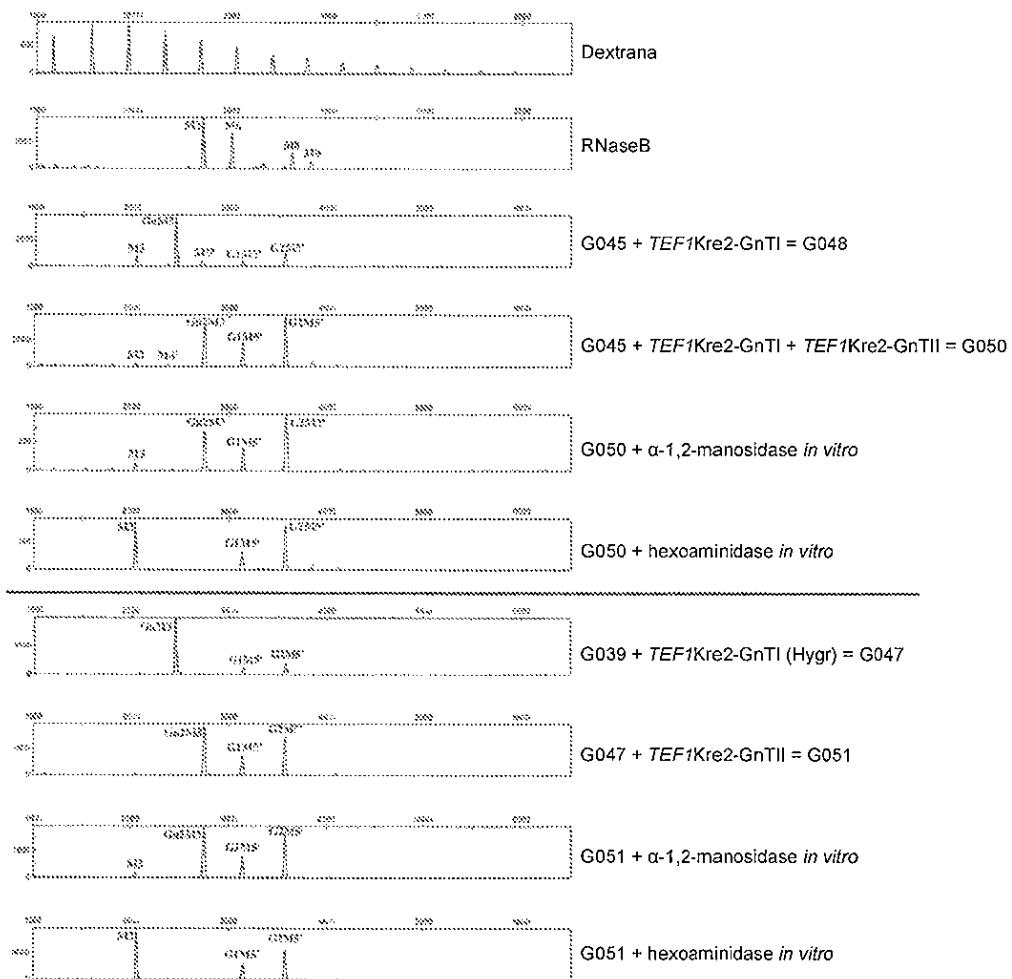


FIG. 24

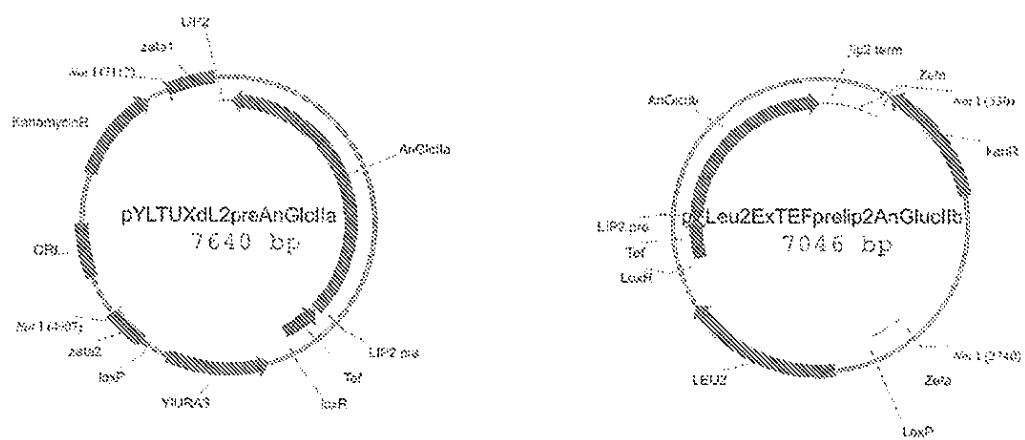


FIG. 25

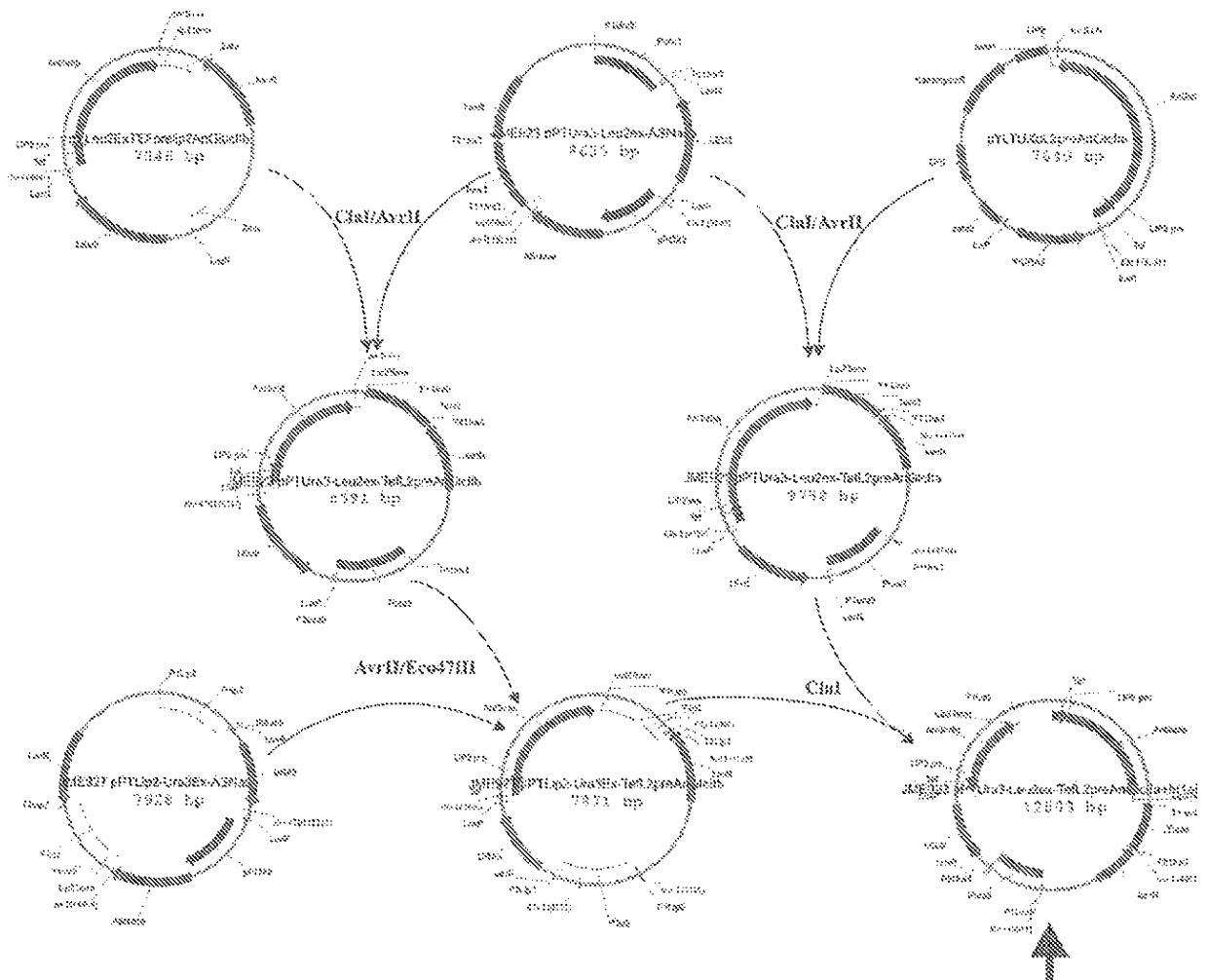


FIG. 26

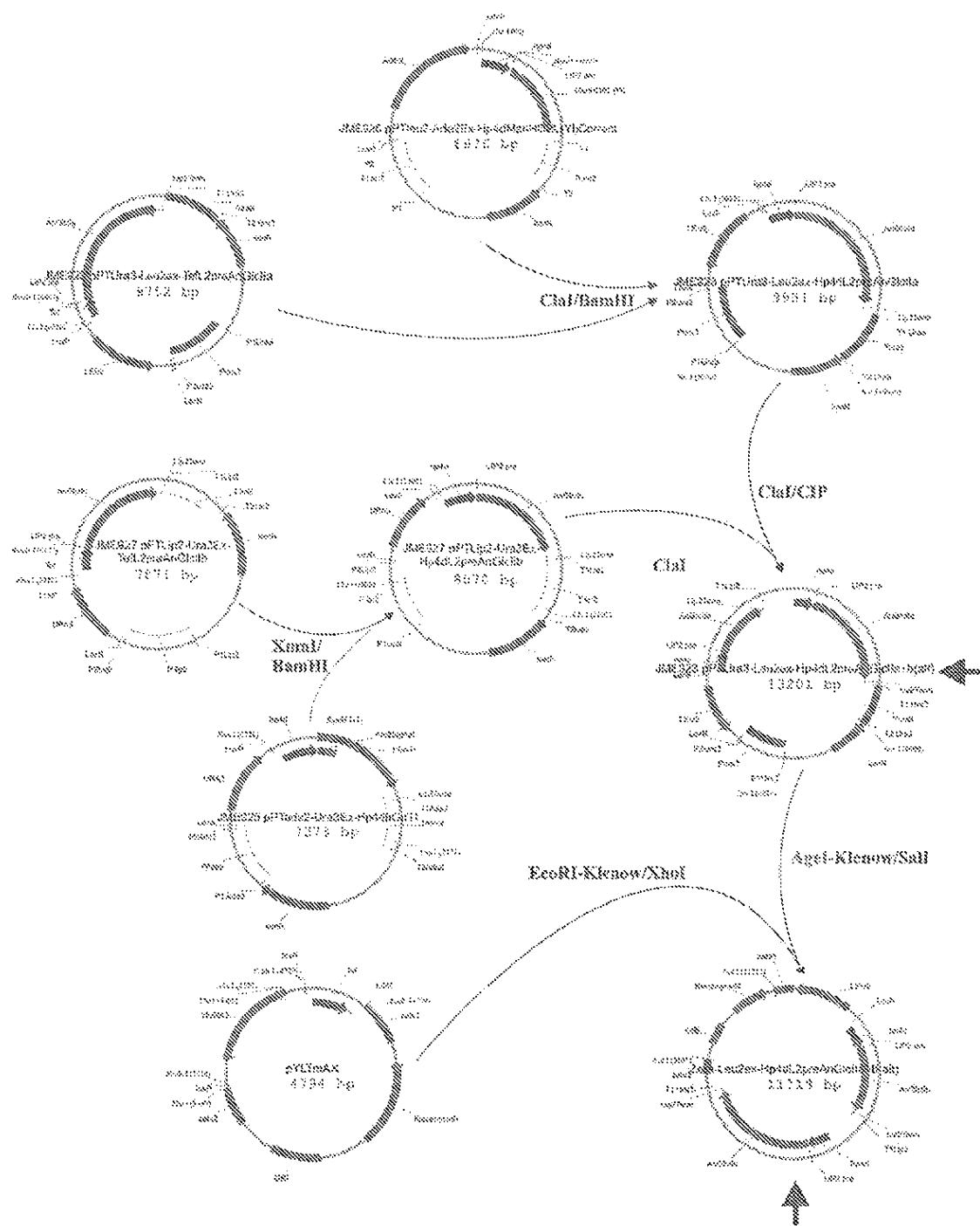


FIG. 27

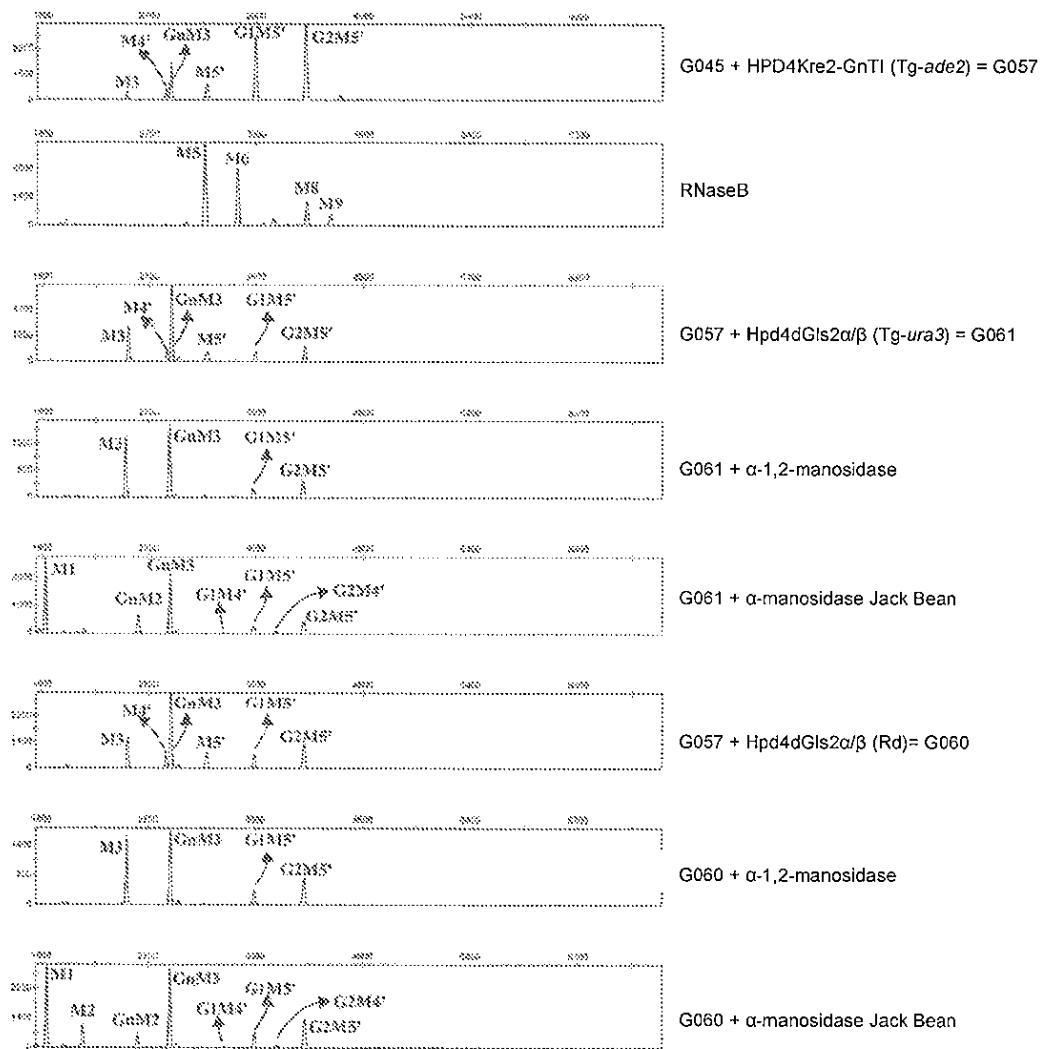


FIG. 28

31/39

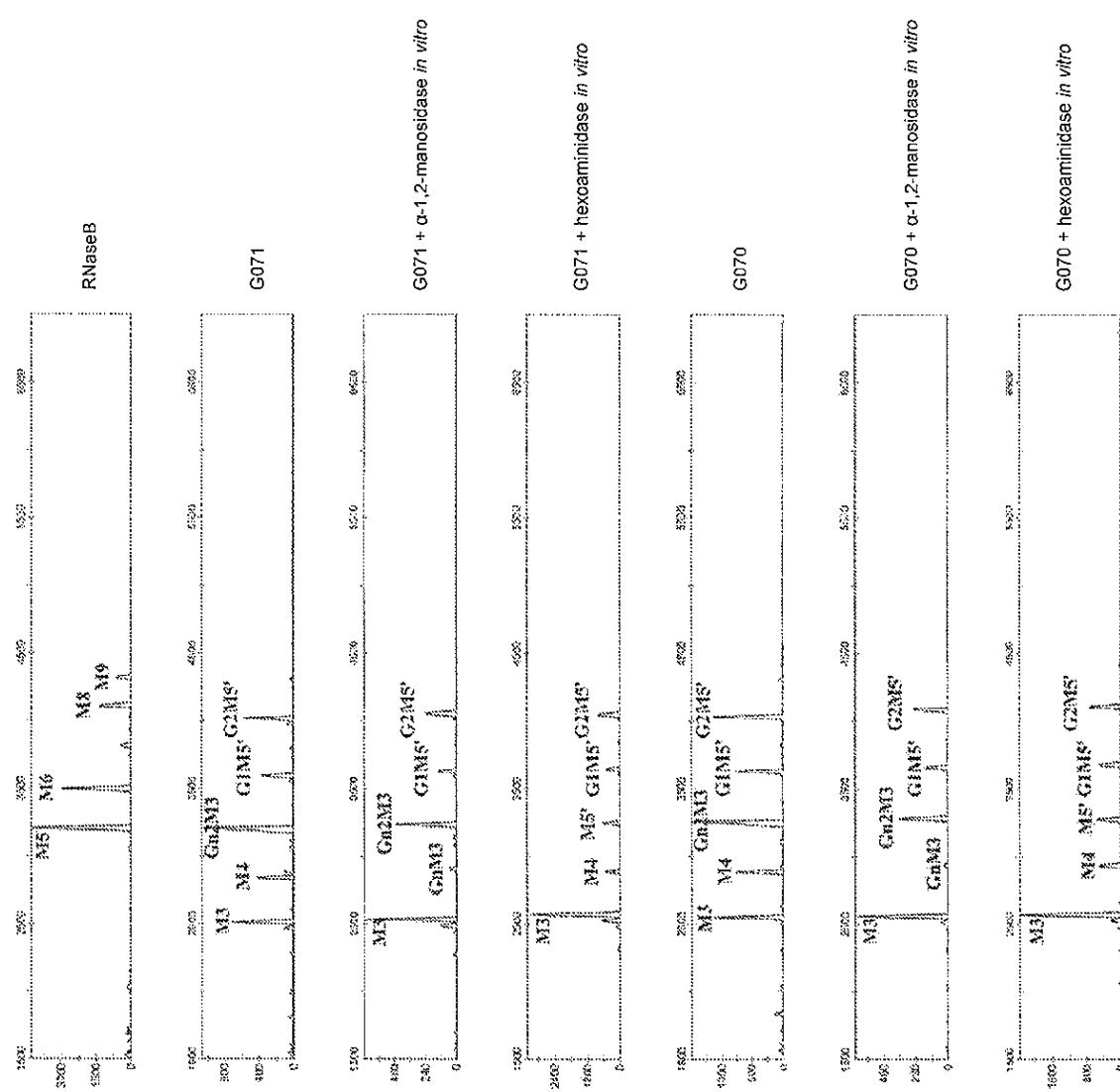


FIG. 29

ATGAAGCTTCCACCATCCTTTCACAGCCTGCCTACCCCTGGCTGCCGCC
CCCTTCCCCCATCACTCCTCTGAGGCCCGCAGTTCTCCAGAAGCGAGGCCCG
GCGACATTCAAGATGACTCAGTCTCCCTCTCTGTCTGCTCTGTGGGTGA
CGAGTGACCATTAACCTGTCGAGCTCTCAGGACGTGAACACTGCTGTTGCTTG
GTATCAGCAGAAGCCTGGAAGGCTCTAAGCTGCTGATCTACTCTGCCCTT
TCCTGTACTCTGGCGTGCCTCTCGATTTCTGGCTCTCGATCTGGAACCGACT
TCACCCCTGACCATTTCTCTGAGCCTGAGGACTTTGCTACCTACTACTGTC
AGCAGCATTACACCACCCCTCTACTTTGGACAGGGACCAAGGTTGAGAT
TAAGCGAACCGTGGCTGCTCCCTCTGTGTTCATTTCCCCCCCCTGACGAGC
AGCTGAAGTCTGGAACTGCTCTGTTGTCCTGCTGAACAACCTTTACCCC
CGAGAGGCTAAGGTTCACTGGAAGGTGGACAACGCTCTGCACTCTGGAACACT
CTCAGGAGTCTGTTACTGAGCAGGACTCTAAGGACTCGACCTACTCTCTCT
TCTACCCCTGACCCCTGCTAAGGCTGACTACGAGAAGCATAAGGTGTACGCTT
GTGAGGTIAACCATCAGGGACTGTCCTCTCCCGTACCCAAGTCTTTAACCGA
GGCGAGTGCTAA

FIG. 30A

MKLSTILFTACATLAAALPSIPTPSEAAVLQKRGGGDIQMKTQSPSSLSASVGDRVITICR
ASQDVNTAVAWYQQKPGKAPKLLIYSASFYSGVPSRFSGSRSGTDFTLTISSLQPEDFA
TYYCQQHYTTPPTFGQGTKEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLNNFYPREA
KVQWKVDNALQSGNSQESVTEQDSKDSTYLSSTTLSKADYEKHKVYACEVTHQGLSSP
VTKSFNRGEC

FIG. 30B

ATGAAGCTTCCACCATCCTTCACAGCCTGCCTACCCCTGGCTGCCGCC
CCCTCCCCCATCACTCCTCTGAGGCCAGTTCTCCAGAAGCGAGGCCG
GCGAGGTTCACTGGTTGAGTCTGGTGAGGACTGGTTCACTGGCTGGATC
TCTGCGACTGTCTGTGCTGCTCTGGCTCAACATCAAGGACACCTACATT
ATTGGGTCCGACAGGCTCCCGAAAGGGACTGGAGTGGTTGCCGAATCTA
CCCTACCAACGGCTACACTCGATACTGACTCTGTGAAGGGACGATTCA
ATTCTGCCGACACCTCTAACAACTGCCTACCTGCAGATGAACACTCTGCG
AGCTGAGGACACTGCTGTACTACTGTTCTCGATGGGGAGGTGACGGTTTT
ACGCCATGGACTACTGGGGACAGGAACTCTGGTGACCGTTCTCTGCTTCT
ACCAAGGGACCTCTGTGTTCTCTCTGGCCCCCTCTCTAACACCTCCCTGCTG
GGAACCTGCTCTGGATGTCCTGGTAAGGACTACTCTCTGAGGCTGTGAC
TGTGCTTGAACCTGCGCTCTGACTCTGGTGTTCACACCTCCCTGCTG
TCTGCAGTCCTCTGGACTGTACTCTCTCTCTGTGACCGTGCCTCTTC
TTCTCTGGGAACCCAGACCTACATCTGTAACGTGAACCACAAGCCCTCTAAC
ACTAAGGTGGACAAGAAGGTGGAGCCTAACGTTGACAAAGACCCATACCT
GTCCCCCTTGTCTCTGAGCTGCTGGGAGGACCCCTGTTTCTGTTCC
CCCCCAAGCTAACGGACACCTGATGATTCTGAACCCCTGAGGTGACCTG
TGTTGTGGTGGACGTTCTCATGAGGACCCCTGAGGTGAAGTTAACGGTACG
TGGACGGTGTGAGGTTACAACGCTAACACTAACGCCCCGAGAGGGAGCAGTA
CAACTCTACTTACCGAGTGGTCTGTGCTGACTGTTCTGCATCAGGACTGGC
TGAACGGAAAGGAATACAAGTGTAAAGGTCTCCAACAAGGCTCTGCCTGCTCC
TATTGAAAAGACCATCTAACGGCTAACGGACAGCCCAGAGAGGCTCAGGTT
TACACTCTGCCCTTCCGAGAGGGAGTGACCAAGAACCAAGGCTCTGCCTG
CTTGTCTGGTCAAGGGATTCTACCCCTCTGACATTGCTGTTGAGTGGGAGTCT
AACGGACAGCCTGAGAACAACTACAAGAACCAACCCCTCTGTTCTGGACTCTG
ACGGCTCTTCTCCTGTACTCTAACGCTGACCGTGGACAAGTCTCGATGGCAG
CAGGGAAACGTGTTCTCTGTTCCGTGATGCATGAGGCTCTGCACAAACCA
CACCCAGAACGTCTGTCTGTCTCCGGCAAGTAA

FIG. 31A

MKLSTILFTACATLAAALPSITPSEAAVLQKRGGGEVQLVESGGGLVQPGGSLRLSCAA
SGFNKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMN
SLRAEDTAVYYCSRWGFDGYAMDYWGQGTLTVSSASTKGPSVFPLAPSSKSTSGGTAA
LGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSVVTPSSLGTQTYICN
VNHKPSNTKVDDKVEPKSCDKTHTCPCTPAPELLGGPSVFLFPPKPDTLMISRTPEVTC
VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC
KVSNKALPAPIEKTISKAKGQPREPQVTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEW
ESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL
SLSPGK

FIG. 31B

36/39

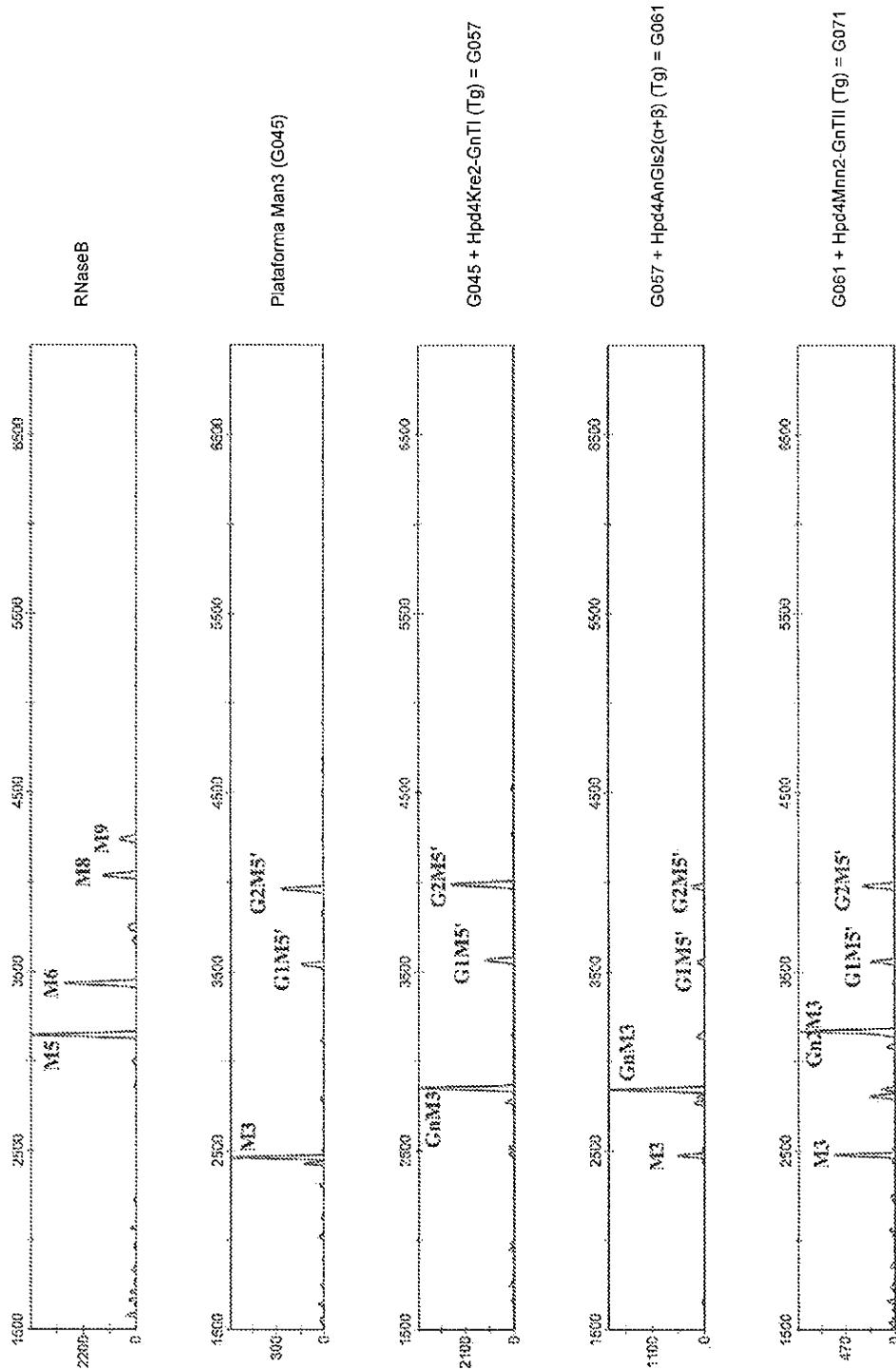
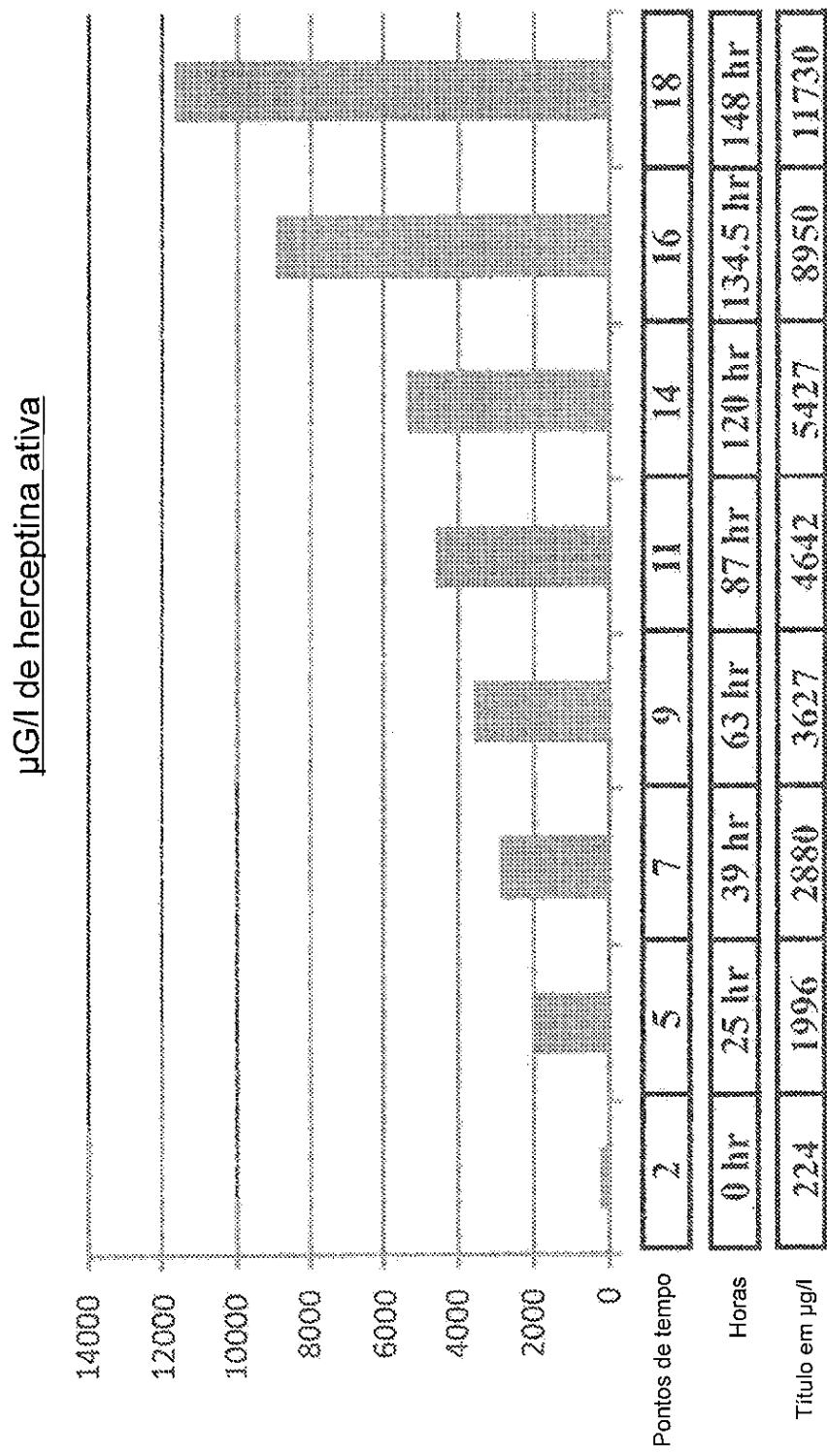


FIG. 32

**FIG. 33**

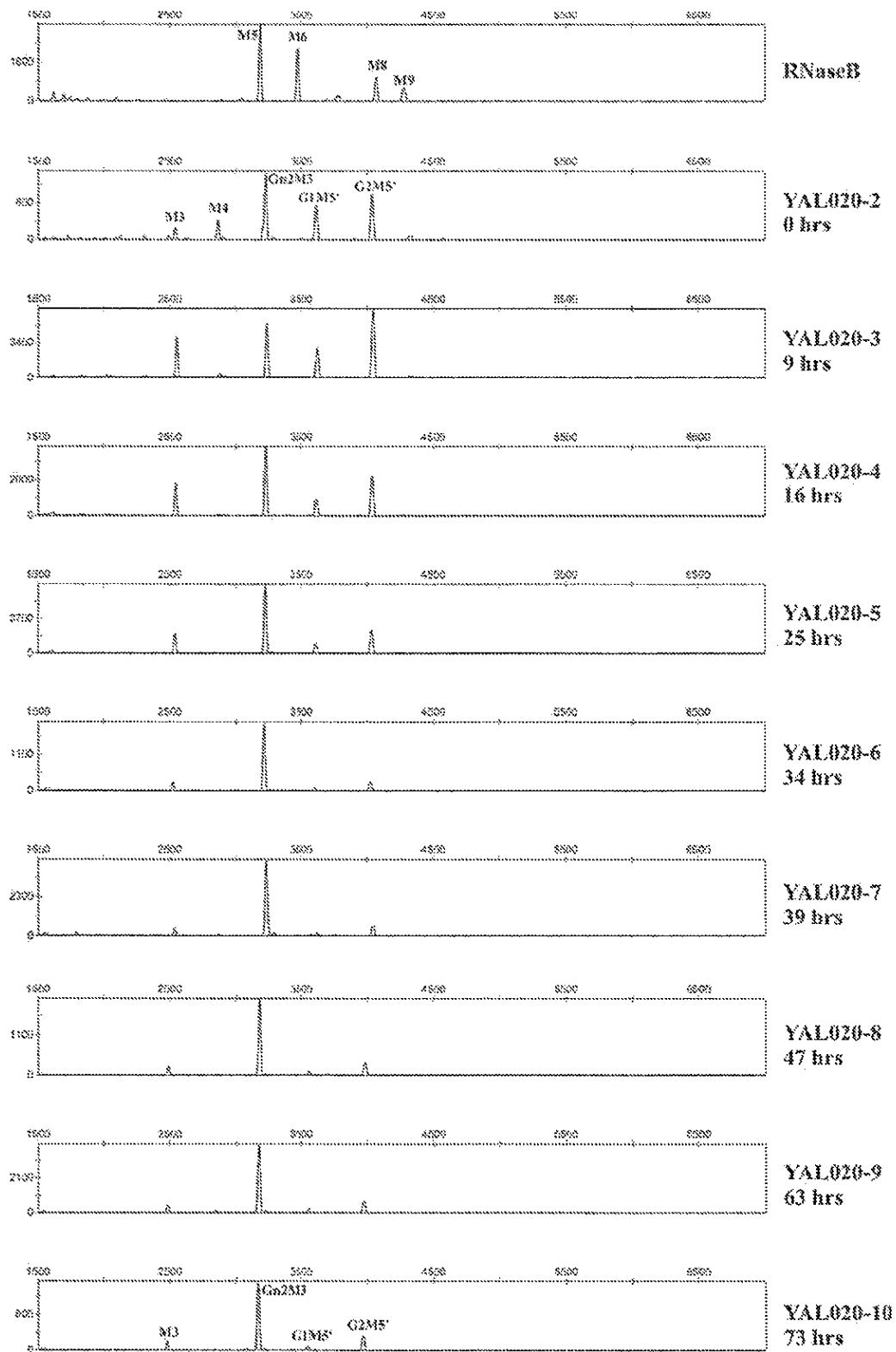


FIG. 34A

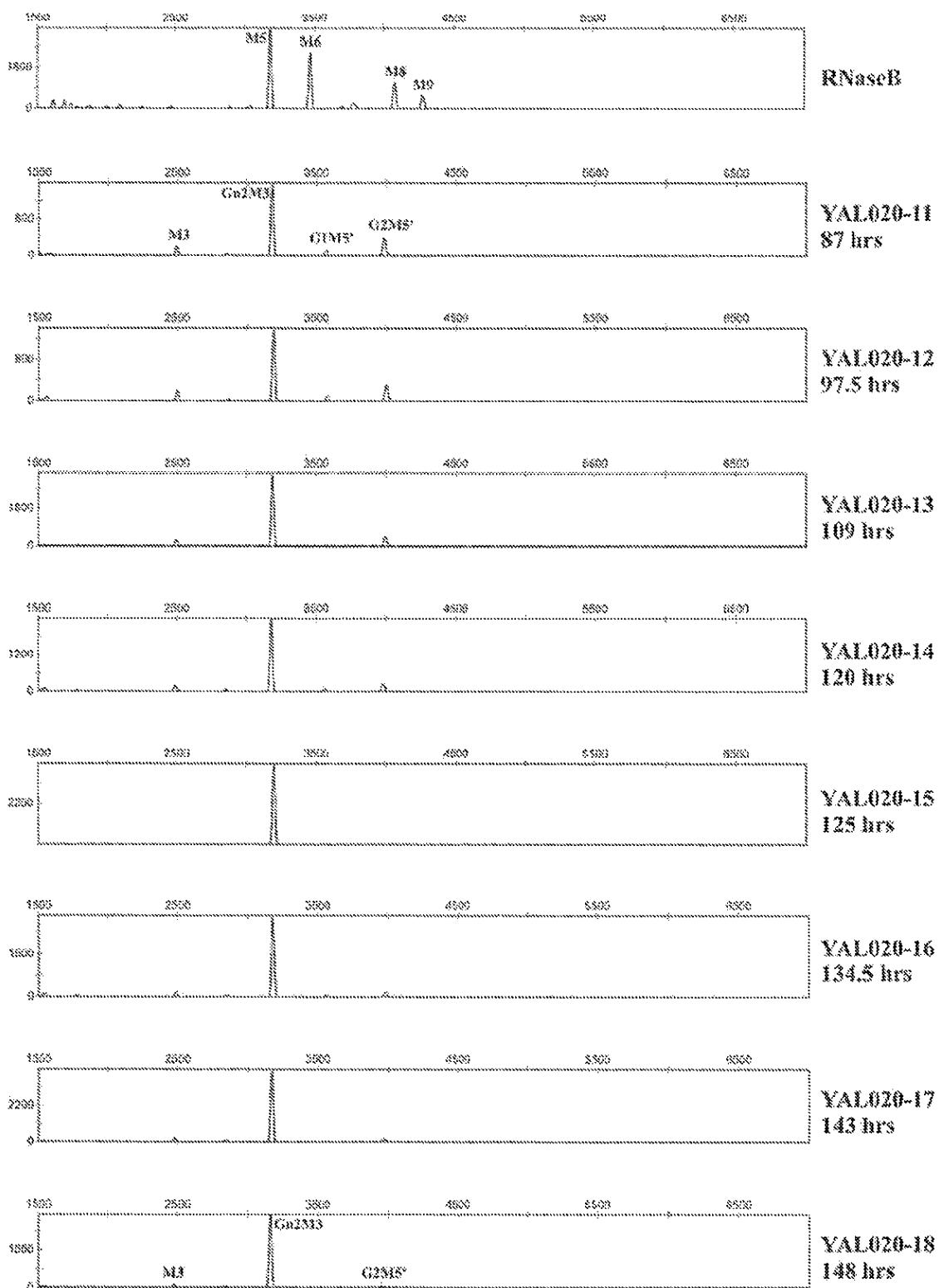


FIG. 34B

RESUMO

Patente de Invenção: "**MÉTODOS DE PRODUZIR CÉLULAS TRANSFORMADAS DE YARROWIA LIPOLYTICA E PROTEÍNAS-ALVO COMPREENDENDO N-GLICANOS, CÉLULAS TRANSFORMADAS DE YARROWIA LIPOLYTICA E SUAS CULTURAS, BEM COMO COMPOSIÇÃO COMPREENDENDO GLICOPROTEÍNAS**".

A presente invenção refere-se a métodos e a células fúngicas geneticamente manipuladas úteis para produção de moléculas-alvo contendo N-glicanos complexos similares de mamíferos ou contendo intermediários em uma via de glicosilação mamífera.

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



50F4FD9CF82D9665

Campo 2



5B94486EA6B55E86

Outras Informações:

- Nome do Arquivo: p186434.TXT
- Data de Geração do Código: 18-05-2012
- Hora de Geração do Código: 14:51:03
- Código de Controle:
 - Campo 1: 50F4FD9CF82D9665
 - Campo 2: 5B94486EA6B55E86