



República Federativa do Brasil
Ministério da Indústria, Comércio Exterior
e Serviços
Instituto Nacional da Propriedade Industrial

(11) PI 0708949-0 B1

(22) Data do Depósito: 09/03/2007

(45) Data de Concessão: 04/04/2017



* B R P I 0 7 0 8 9 4 9 B 1 *

(54) Título: MÉTODOS PARA DEGRADAR OU CONVERTER UM MATERIAL CELULÓSICO, PARA PRODUZIR UMA SUSBTÂNCIA, E PARA PRODUZIR UM POLIPEPTÍDEO TENDO ATIVIDADE ENDOGLUCANASE, E, CÉLULA MICROBIANA RECOMBINANTE

(51) Int.Cl.: C12N 9/42; C12N 15/56; C12N 15/11; C12N 5/10

(30) Prioridade Unionista: 20/03/2006 US 60/784,088

(73) Titular(es): NOVOZYMES, INC.. NOVOZYMES A/S

(72) Inventor(es): PAUL HARRIS; ELENA VLASENKO; SOREN FLENSTED LASSEN; KRISTIAN BERTEL ROMER MORKEBERG

“MÉTODOS PARA DEGRADAR OU CONVERTER UM MATERIAL CELULÓSICO, PARA PRODUZIR UMA SUSBTÂNCIA, E PARA PRODUZIR UM POLIPEPTÍDEO TENDO ATIVIDADE ENDOGLUCANASE, E, CÉLULA MICROBIANA RECOMBINANTE”

5 Referência a uma Listagem de Seqüência

Este pedido contém uma Listagem de Seqüências em forma legível por computador. A forma legível por computador é incorporada aqui por referência.

Referência a um Depósito de Material Biológico

10 Este pedido contém uma referência a depósitos de material biológico que foram feitos no Northern Regional Research Center (NRRL) sob o Tratado de Budapeste e receberam os números de acesso NRRL B-3Q900N, NRRL B-30902, NRRL B-30903, e NRRL 8-30904, cujos depósitos microbianos são incorporados aqui por referência.

15 Antecedente da Invenção

Campo da Invenção

A presente invenção refere-se a polipeptídeos isolados tendo atividade de endoglucanase e polinucleotídeos isolados codificando os polipeptídeos. A invenção também refere-se a construções de ácido nucleico, vetores, e células hospedeiras compreendendo os polinucleotídeos assim como a métodos para
20 produzir e usar os polipeptídeos.

Descrição da Arte Relacionada

Celulose é um polímero da glicose de açúcar simples covalentemente ligada por beta-1,4-ligações. Muitos microorganismos produzem enzimas que hidrolisam glucanos beta-ligados. Estas enzimas incluem endoglucanases, celobiohidrolases, e beta-glucosidases. Endoglucanases digerem o polímero de celulose em locais aleatórios, abrindo o mesmo ao ataque por celobiohidrolases. Celobiohidrolases sequencialmente liberam moléculas de celobiose das extremidades do polímero de celulose. Celobiohidrolase I é uma atividade de 1,4-beta-D-glucano celobiohidrolase (E.G. 3.2.1.91) que catalisa a hidrólise de ligações 1,4-beta-D-glucosídicas em celulose, celotetrose, ou qualquer polímero contendo glicose beta-1,4-ligada, liberando celobiose das extremidades redutoras da cadeia. Celobiohidrolase II é uma atividade de 1,4-D-glucano celobiohidrolase (E.G. 3.2.1.91) que catalisa a hidrólise de ligações 1,4-beta-D-glucosídicas em celulose, celotetrose, ou qualquer polímero contendo glicose beta-1,4-ligada, liberando celobiose das extremidades não redutoras da cadeia. Celobiose é um dímero de glicose beta-1,4-ligado solúvel em água. Beta-glucosidases hidrolisam celobiose em glicose.

A conversão de cargas de alimentação celulósica em etanol tem as vantagens da pronta disponibilidade de grandes quantidades de cargas de alimentação, a característica desejável de evitar queimadas ou terra presa aos materiais, e a limpeza do combustível etanol. Madeira, resíduos agrícolas, culturas herbáceas, e refugos sólidos municipais foram considerados como cargas de alimentação para a produção de etanol. Estes materiais primariamente consistem de celulose, hemicelulose, e lignina. Uma vez que a celulose é convertida em glicose, a glicose é facilmente fermentada por levedura em etanol.

Roy *et al.*, 1990, *Journal of General Microbiology* 136; 1967 - 1972, descrevem a purificação e as propriedades de uma endoglucanase

extracelular de *Myceliophthora thermophila* ATCC 48104, Chemoglazov *et al.*, 1988, *Biokhimiya* 53; 475 - 482, descrevem o isolamento, a purificação, e especificidade de substrato de uma endoglucanase de *Myceliophthora thermophila*. Klyosov *et al.*, 1988, *Biotechnology Letters* 10: 351 - 354, descrevem uma endoglucanase termoestável de *Myceliophthora thermophila*, Guzhova e Loginova, 1987, *Prikladnaya Biokhimiya I Mikrobiologiya* 23: 820 - 825, descrevem enzimas celulolíticas de *Myceliophthora thermophila*, Rabinovich *et al.*, 1986, *Bioorganicheskaya Khimiya* 12: 1549 - 1560, descrevem a purificação e caracterização de uma endoglucanase de *Myceliophthora thermophila*. Svistova *et al.*, 1986, *Mikrobiologiya* 55: 49 - 54, descrevem a regulação de biossíntese de celulose em *Myceliophthora thermophila*. Bhat e Maheshwari, 1987, *Applied and Environmental Microbiology* 53; 2175 - 2182, descrevem a atividade de componentes do sistema de celulose extracelular de *Myceliophthora thermophila*. Klyosov *et al.*, 1987, *Prikladnaya Biokhimiya I Mikrobiologiya* 23: 44 - 50, descrevem uma endoglucanase termoestável de *Myceliophthora thermophila*. Jorgensen *et al.*, 2003, *Enzyme and Microbial Technology* 32: 851 - 861, e Thygesen *et al.*, 2003, *Enzyme and Microbial Technology* 32: 606 - 615, descrevem as enzimas degradando celulose *Penicillium brasilianum* IBT 20888.

Seria uma vantagem na técnica identificar novas endoglucanases tendo propriedades melhoradas, como taxa de hidrólise melhorada, melhor estabilidade térmica, absorção reduzida para lignina, e capacidade para hidrolisar componentes não celulósicos de biomassa, como hemicelulose, além de hidrolisar celulose. Endoglucanases com uma faixa ampla de atividades laterais em hemicelulose podem ser especialmente benéficas para melhorar o rendimento global de hidrólise de substratos de biomassa ricos em hemicelulose complexa.

É um objeto da presente invenção prover polipeptídeos melhorados tendo atividade de endoglucanase e polinucleotídeos codificando

os polipeptídeos.

Sumário da Invenção

A presente invenção refere-se a polipeptídeos isolados tendo atividade de endoglucanase selecionados dentre o grupo consistindo de:

- 5 (a) um polipeptídeo compreendendo uma seqüência de aminoácido que tem pelo menos 80% de identidade com o polipeptídeo maduro de SEQ ID NO. 4 ou SEQ ID NO: 10, pelo menos 85% de identidade com o polipeptídeo maduro de SEQ. ID NO: 6, ou pelo menos 75% de identidade com o polipeptídeo maduro de SEQ ID NO: 8;
- 10 (b) um polipeptídeo que é codificado por um polinucleotídeo que hibridiza sob condições de alta estringência pelo menos com (i) a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9, (ii) a seqüência de cDNA contida em uma seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3
- 15 ou SEQ ID NO: 9 ou a seqüência de DNA genômico compreendendo a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 5 ou SEQ ID NO: 7, ou (iii) um filamento complementar de comprimento completo de (i) ou (ii);
- (c) um polipeptídeo que é codificado por um polinucleotídeo
- 20 tendo pelo menos 80% de identidade com a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3 ou SEQ ID NO: 9, pelo menos 85% de identidade com a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 5, ou pelo menos 75% de identidade com a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 7; e
- 25 (d) uma variante compreendendo uma substituição, deleção, e/ou inserção de um ou mais aminoácidos do polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10.

A presente invenção também refere-se a polinucleotídeos isolados codificando polipeptídeos tendo atividade de endoglucanase,

selecionados dentre o grupo consistindo de:

(a) um polinucleotídeo codificando um polipeptídeo compreendendo uma seqüência de aminoácido que tem pelo menos 80% de identidade com o polipeptídeo maduro de SEQ ID NO: 4 ou SEQ ID NO: 10, pelo menos 85% de identidade com o polipeptídeo maduro de SEQ ID NO: 6, ou pelo menos 75% de identidade com o polipeptídeo maduro de SEQ ID NO: 8;

(b) um polinucleotídeo que hibridiza sob condições pelo menos de alta estringência com (i) a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9, (ii) a seqüência de cDNA contida em uma seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3 ou SEQ ID NO: 9 ou a seqüência de DNA genômico compreendendo a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 5 ou SEQ ID NO: 7, ou (iii) um filamento complementar de (i) ou (ii);

(c) um polinucleotídeo tendo pelo menos 80% de identidade com a seqüência de codificação de polipeptídeo maduro de SEQ. ID NO: 3 ou SEQ ID NO: 9, pelo menos 85% de identidade com a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 5, ou pelo menos 75% de identidade com a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 7; e

(d) um polinucleotídeo codificando uma variante compreendendo uma substituição, deleção, e/ou inserção de um ou mais aminoácidos do polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10,

Em um aspecto preferido, o polipeptídeo maduro contém os aminoácidos 17 a 389 de SEQ ID NO: 4. Em outro aspecto preferido, o polipeptídeo maduro contém os aminoácidos 16 a 397 de SEQ ID NO: 6. Em outro aspecto preferido, o polipeptídeo maduro contém os aminoácidos 22 a

429 de SEQ ID NO: 8. Em outro aspecto preferido, o polipeptídeo maduro contém os aminoácidos 25 a 421 de SEQ ID NO: 10. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 67 a 1185 de SEQ ID NO: 3. Em outro aspecto preferido, a
5 sequência de codificação de polipeptídeo maduro contém os nucleotídeos 84 a 1229 de SEQ ID NO: 5. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 77 a 1300 de SEQ ID NO: 7. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 73 a 1468 de SEQ ID NO: 9.

10 A presente invenção refere-se também à construção de ácido nucleico, vetores de expressão recombinantes, células hospedeiras recombinantes compreendendo os polinucleotídeos, e métodos de produzir um polipeptídeo tendo atividade de endoglucanase.

A presente invenção também refere-se a métodos de inibir a
15 expressão de um polipeptídeo em uma célula, compreendendo administrar à célula ou expressar na célula uma molécula de RNA filamento duplo (dsRNA), em que o dsRNA compreende uma subsequência de um polinucleotídeo da presente invenção. O presente também refere-se a tal molécula de RNA filamento duplo inibidor (dsRNA), em que opcionalmente
20 o dsRNA é uma molécula de siRNA ou de miRNA.

A presente invenção também refere-se a métodos de usar os polipeptídeos tendo atividade de endoglucanase na conversão de celulose em glicose e várias substâncias.

A presente invenção também refere-se às plantas
25 compreendendo um polinucleotídeo isolado codificando este polipeptídeo tendo atividade de endoglucanase.

A presente invenção também refere-se a métodos para produzir este um polipeptídeo tendo atividade de endoglucanase, compreendendo: (a) cultivar uma planta transgênica ou uma célula planta

compreendendo um polinucleotídeo codificando este polipeptídeo tendo atividade de endoglucanase sob condições que conduzem à produção do polipeptídeo; e (b) recuperar o polipeptídeo.

A presente invenção ainda refere-se à construção de ácido nucleico compreendendo um gene codificando uma proteína, em que um gene é ligado operativamente a uma seqüência de nucleotídeos codificando um peptídeo de sinal compreendendo ou consistindo de aminoácidos 1 a 16 de SEQ ID NO: 4, aminoácidos 1 a 15 de SEQ ID NO: 6, aminoácidos 1 a 21 de SEQ ID NO: 8, ou aminoácidos 1 a 16 de SEQ ID NO: 10 e uma segunda seqüência de nucleotídeos codificando um pró-peptídeo compreendendo ou consistindo de aminoácidos 17 a 24 de SEQ ID NO: 10, em que um gene é estranho para a primeira e segunda seqüências de nucleotídeos.

Breve Descrição das Figuras

Figura 1 mostra um mapa de restrição de pCIC161.
Figura 2 mostra a seqüência de DNA genômico e a seqüência de aminoácido deduzida de endoglucanase de *Myceliophthora thermophila* CBS 117.65 (SEQ ID NOs: 3 e 4, respectivamente).

Figura 3 mostra um mapa de restrição de pA2C161.
Figura 4 mostra um mapa de restrição de pCIC453.
Figura 5 mostra um mapa de restrição de pA2C453.
Figura 6 mostra a seqüência de cDNA e a seqüência de aminoácido deduzida de uma endoglucanase de basidiomicete CBS 495.95 (SEQ ID NOs: 5 e 6, respectivamente).

Figura 7 mostra um mapa de restrição de pCIC486.
Figura 8 mostra um mapa de restrição de pA2C486.
Figura 9 mostra a seqüência de cDNA e a seqüência de aminoácido deduzida de uma endoglucanase de basidiomicete CBS 494.95 (SEQ ID NOs: 7 e 8, respectivamente).

Figuras 10A e 10B mostram a seqüência de DNA genômico e

a sequência de aminoácido deduzida de uma endoglucanase de *Penicillium brasilianum* cepa IBT 20888 (SEQ ID NOs: 9 e 10, respectivamente).

Figura 11 mostra um mapa de restrição de pKBK03.

Figura 12 mostra um mapa de restrição de pPBCel5C.

5 Figura 13 mostra a atividade específica de endoglucanase de *Penicillium brasilianum* IBT 20888 CEL5C em valores de pH diferentes e 50°C (n=2).

10 Figura 14 mostra a atividade específica de endoglucanase de *Penicillium brasilianum* IBT 20888 CEL5C em temperaturas diferentes e pH 4,8 (n=2).

Figura 15 mostra a atividade residual de endoglucanase de *Penicillium brasilianum* IBT 20888 CEL5C após 20 horas de incubação em valores de pH diferentes e 25°C e 50°C (n=2).

15 Figura 16 mostra a atividade relativa em PASC (2 mg/ml) como uma função de temperatura para basidiomicete CBS 494.95 e basidiomicete CBS 495.95 a pH 5.0.

20 Figura 17 mostra a conversão relativa de PASC (2 mg/ml) como uma função de temperatura após 45 horas de hidrólise com basidiomicete CBS 494.95 e basidiomicete CBS 495.95 (0,5 mg de proteína por g de PASC) a pH 5,0.

25 Figura 18 mostra uma comparação de endoglucanases de *Myceliophthora thermophila* CBS 117.65, basidiomicete CBS 494.95, basidiomicete CBS 495.95, e *Trichoderma reesei* para produção de açúcares redutores a partir de beta-glucano (1% p/v) após 2 horas de reação de hidrólise a pH 5,5 e 60°C.

Figura 19 mostra uma comparação de endoglucanases de *Myceliophthora thermophila* CBS 117.65, basidiomicete CBS 494.95, basidiomicete CBS 495.95, e *Trichoderma reesei* para a produção de açúcares redutores a partir de beta-glucano (1% p/v) após 2 horas de reação de

hidrólise a pH 5,5 e 60°C.

Definições

Atividade de endoglucanase: O termo "atividade de endoglucanase" é definido aqui como uma endo-1,4-beta-D-glucano 4-glucano-hidrolase (E.C. No. 3.2.1.4) que catalisa a endo-hidrólise de ligações 1,4-beta-D-glicosídicas em celulose, derivados de celulose (como carboximetil celulose e hidroxietil celulose), liquenina, ligações beta-1,4 em misturas de beta-1,3 glucanos como beta-D-glucanos ou xiloglucanos de cereais, e outro material de planta contendo componentes celulósicos. Para os fins da presente invenção, a atividade de endoglucanase é determinada usando hidrólise de carboximetil celulose (CMC) de acordo com o procedimento de Ghose, 1987, *Pure and Appl. Chem.* 59: 257 - 268. Uma unidade de atividade de endoglucanase é definida como 1,0 µmol de açúcares redutores produzidos por minuto a 50°C, pH 4,8.

Em um aspecto preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase ainda apresentam atividade de enzima para um ou mais substratos selecionados dentre o grupo consistindo de xilano, xiloglucano, arabinoxilano, 1,4-beta-D-manano, e galactomanano. A atividade dos polipeptídeos tendo atividade de endoglucanase nestes substratos de polissacarídeo é determinada como a porcentagem do substrato hidrolisado em açúcares redutores após incubar o substrato (5 mg por ml) com um polipeptídeo tendo atividade de endoglucanase da presente invenção (5 mg de proteína por g de substrato) durante 24 horas com agitação intermitente a pH 5,0 (50 mM de acetato de sódio) e 50°C. Açúcares redutores em misturas de hidrólise são determinados pelo teste de hidrazida de ácido p-hidroxibenzóico(PHBAH).

Em um aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase ainda têm atividade de enzima para xilano. Em outro aspecto mais preferido os polipeptídeos da presente

invenção tendo atividade de endoglucanase ainda têm atividade de enzima para xiloglucano. Em outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase ainda têm atividade de enzima para arabinosilano. Em outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase ainda têm atividade de enzima para 1,4-beta-D-manano. Em outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase ainda têm atividade de enzima para galactomanano. Em outro aspecto mais preferido, os polipeptídeos da presente invenção tendo atividade de endoglucanase ainda têm atividade de enzima para xilano, xiloglucano, arabinosilano, 1,4-beta-D-manano, e/ou galactomanano.

Os polipeptídeos da presente invenção têm pelo menos 20%, preferivelmente pelo menos 40%, mais preferivelmente pelo menos 50%, mais preferivelmente pelo menos 60%, mais preferivelmente pelo menos 70%, mais preferivelmente pelo menos 80%, ainda mais preferivelmente pelo menos 90%, o mais preferivelmente pelo menos 95%, e ainda o mais preferivelmente pelo menos 100% da atividade de endoglucanase de polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10.

Família 5 de glicosídeo hidrolase ou família GH5: O termo "família 5 de glicosídeo hidrolase" ou "família GH5" é definido aqui como um polipeptídeo dentro na família 5 glicosídeo hidrolase de acordo com Henrissat B., 1991, A classification of glycosyl hydrolases based on amino-acid sequence similarities, *Biochem. J.* 280: 309 - 316, e Henrissat B., e Bairoch A., 1996, Updating the sequence-based classification of glycosyl hydrolases, *Biochem, J.* 316: 695 -696.

Polipeptídeo isolado: O termo "polipeptídeo isolado" como usado aqui refere-se a um polipeptídeo que é pelo menos 20% puro, preferivelmente pelo menos 40% puro, mais preferivelmente pelo menos 60%

puro, ainda mais preferivelmente pelo menos 80% puro, o mais preferivelmente pelo menos 90% puro, e ainda o mais preferivelmente pelo menos 95% puro, como determinado por SDS-PAGE.

Poli-peptídeo substancialmente puro: O termo "poli-peptídeo substancialmente puro" denota aqui uma preparação de poli-peptídeo que contém no máximo 10%, preferivelmente no máximo 8%, mais preferivelmente no máximo 6%, mais preferivelmente no máximo 5%, mais preferivelmente no máximo 4%, mais preferivelmente no máximo 3%, ainda mais preferivelmente no máximo 2%, o mais preferivelmente no máximo 1%, e ainda o mais preferivelmente no máximo 0,5% em peso de outro material de poli-peptídeo com que é naturalmente ou recombinantemente associado. Assim, prefere-se que o poli-peptídeo substancialmente puro é pelo menos 92% puro, preferivelmente pelo menos 94% puro, mais preferivelmente pelo menos 95% puro, mais preferivelmente pelo menos 96% puro, mais preferivelmente pelo menos 96% puro, mais preferivelmente pelo menos 97% puro, mais preferivelmente pelo menos 98% puro, ainda mais preferivelmente pelo menos 99%, o mais preferivelmente pelo menos 99,5% puro, e ainda o mais preferivelmente 100% puro em peso do material de poli-peptídeo total presente na preparação.

Os poli-peptídeos da presente invenção estão preferivelmente em uma forma substancialmente pura. Em particular, prefere-se que os poli-peptídeos estejam na "forma essencialmente pura", isto é, que a preparação de poli-peptídeo seja essencialmente livre de outro material de poli-peptídeo com o qual ele é naturalmente ou recombinantemente associado. Isto pode ser obtido, por exemplo, por preparação do poli-peptídeo por meio de métodos recombinantes bem conhecidos ou por métodos clássicos de purificação.

Aqui, o termo "poli-peptídeo substancialmente puro" é sinônimo de termos "poli-peptídeo isolado" e "poli-peptídeo na forma isolada."

Polipeptídeo maduro: O termo "polipeptídeo maduro" é definido aqui como um polipeptídeo tendo atividade de endoglucanase que é em sua forma final após tradução e quaisquer modificações pós-tradução tais, como processamento N-terminal, truncagem C-terminal, glicosilação, etc.

5 **Seqüência de codificação de polipeptídeo maduro:** O termo "seqüência de codificação de polipeptídeo maduro" é definido aqui como uma seqüência de nucleotídeo que codifica um polipeptídeo maduro tendo atividade de endoglucanase.

10 **Identidade:** O parentesco entre duas seqüências de aminoácido ou entre duas seqüências de nucleotídeo é descrito pelo parâmetro "identidade".

Para os fins da presente invenção, o grau de identidade entre duas seqüências de aminoácido é determinado usando o algoritmo Needleman-Wunsch (Needleman e Wunsch, 1970, *J. Mol. Biol.* 48: 443 - 453) como implementado no programa Needle de EMBOSS com penalidade de abertura de espaço de 10, penalidade de extensão de espaço de 0,5, e a matriz de EBLOSUM62. A saída de Needle rotulada "identidade mais longa" é usada como a porcentagem de identidade e é calculada como a seguir:

(Resíduos idênticos x 100)/(Comprimento de Alinhamento - Número de espaços em alinhamento)

20 Para os fins da presente invenção, o grau de identidade entre duas seqüências de nucleotídeo é determinado usando o algoritmo Needleman-Wunsch (Needleman e Wunsch, 1970, *J. Mol. Biol.* 48: 443 - 453) como implementado no programa de Needle de EMBOSS com penalidade de abertura de espaço de 10, penalidade de extensão de espaço de 0,5, e a matriz de EDNAFULL. A saída de Needle rotulada "identidade mais longa" é usada como a porcentagem de identidade e é calculada como a seguir:

(Resíduos idênticos x 100)/(Comprimento de Alinhamento - Número de

Espaços em Alinhamento)

Seqüência homóloga: O termo "seqüência homóloga" é definido aqui como uma proteína prevista que dá um valor E (ou escore de expectativa) de menos que 0,001 em uma busca **fasta** (Pearson, W. R., 1999, in *Bioinformatics Methods and Protocols*, S. Misener e S. A. Krawetz, ed, pp. 185 - 219) com a endoglucanase madura de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10.

Fragmento de polipeptídeo: O termo "fragmento de polipeptídeo" é definido aqui como um polipeptídeo tendo um ou mais aminoácidos deletados do término amino e/ou carbóxi do polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10; ou uma seqüência homóloga do mesmo, em que um fragmento tem atividade de endoglucanase. Em um aspecto preferido, um fragmento contém pelo menos 295 resíduos de aminoácido, mais preferivelmente pelo menos 315 resíduos de aminoácido, e o mais preferivelmente pelo menos 335 resíduos de aminoácido do polipeptídeo maduro de SEQ ID NO: 4 ou uma seqüência homóloga do mesmo. Em outro aspecto preferido, um fragmento contém pelo menos 320 resíduos de aminoácido, mais preferivelmente pelo menos 340 resíduos de aminoácido, e o mais preferivelmente pelo menos 360 resíduos de aminoácido do polipeptídeo maduro de SEQ ID NO: 6 ou uma seqüência homóloga do mesmo. Em outro aspecto preferido, um fragmento contém pelo menos 325 resíduos de aminoácido, mais preferivelmente pelo menos 345 resíduos de aminoácido, e o mais preferivelmente pelo menos 365 resíduos de aminoácido do polipeptídeo maduro de SEQ ID NO: 8 ou uma seqüência homóloga do mesmo. Em outro aspecto preferido, um fragmento contém pelo menos 335 resíduos de aminoácido, mais preferivelmente pelo menos 355 resíduos de aminoácido, e o mais preferivelmente pelo menos 375 resíduos de aminoácido do polipeptídeo maduro de SEQ ID NO: 10 ou uma seqüência homóloga do mesmo.

Subseqüência: O termo "subseqüência" é definido aqui como uma seqüência de nucleotídeo tendo um ou mais nucleotídeos deletados da extremidade 5' e/ou 3' da seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9; ou uma
5 seqüência homóloga do mesmo; em que a subseqüência codifica um fragmento de polipeptídeo tendo atividade de endoglucanase. Em um aspecto preferido, uma subseqüência contém pelo menos 885 nucleotídeos, mais preferivelmente pelo menos 945 nucleotídeos, e o mais preferivelmente pelo menos 1005 nucleotídeos da seqüência de codificação de polipeptídeo maduro
10 de SEQ. ID NO: 3 ou uma seqüência homóloga do mesmo. Em outro aspecto preferido, uma subseqüência contém pelo menos 960 nucleotídeos, mais preferivelmente pelo menos 1020 nucleotídeos, e o mais preferivelmente pelo menos 1080 nucleotídeos da seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 5 ou uma seqüência homóloga do mesmo. Em outro aspecto
15 preferido, uma subseqüência contém pelo menos 975 nucleotídeos, mais preferivelmente pelo menos 1035 nucleotídeos, e o mais preferivelmente pelo menos 1095 nucleotídeos da seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 7 ou uma seqüência homóloga do mesmo. Em outro aspecto preferido, uma subseqüência contém pelo menos 1005 nucleotídeos, mais
20 preferivelmente pelo menos 1065 nucleotídeos, e o mais preferivelmente pelo menos 1125 nucleotídeos da seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 9 ou uma seqüência homóloga do mesmo.

Variante alélica: O termo "variante alélica" denota aqui qualquer uma dentre duas ou mais formas alternativas de um gene ocupando o
25 mesmo *locus* cromossômico. Variação alélica surge naturalmente através da mutação, e pode resultar em polimorfismo dentro das populações. Mutações de gene podem ser silentes (sem mudança no polipeptídeo codificado) ou podem codificar polipeptídeos tendo seqüências de aminoácido alteradas. Uma variante alélica de um polipeptídeo é um polipeptídeo codificado por

uma variante alélica de um gene.

Polinucleotídeo isolado: O termo “polinucleotídeo isolado” como usado aqui refere-se a um polinucleotídeo que é pelo menos 20% puro, preferivelmente pelo menos 40% puro, mais preferivelmente pelo menos 60% puro, ainda mais preferivelmente pelo menos 80% puro, o mais preferivelmente pelo menos 90% puro, e ainda o mais preferivelmente pelo menos 95% puro, como determinado por eletroforese agarose.

Polinucleotídeo substancialmente puro: O termo “polinucleotídeo substancialmente puro” como usado aqui refere-se a uma preparação de polinucleotídeo livre de outros nucleotídeos estranhos ou indesejados e em uma forma apropriada para usar em sistemas de produção de proteína geneticamente engenheirados. Assim, um polinucleotídeo substancialmente puro contém no máximo 10%, preferivelmente no máximo 8%, mais preferivelmente no máximo 6%, mais preferivelmente no máximo 5%, mais preferivelmente no máximo 4%, mais preferivelmente no máximo 3%, ainda mais preferivelmente no máximo 2%, o mais preferivelmente no máximo 1%, e ainda o mais preferivelmente no máximo 0,5% em peso de outro material de polinucleotídeo com o qual ele é naturalmente ou recombinantemente associado. Um polinucleotídeo substancialmente puro pode, no entanto, incluir regiões 5' e 3' de ocorrência natural não traduzidas, como promotores e terminadores. Prefere-se que o polinucleotídeo substancialmente puro seja pelo menos 90% puro, preferivelmente pelo menos 92% puro, mais preferivelmente pelo menos 94% puro, mais preferivelmente pelo menos 95% puro, mais preferivelmente pelo menos 96% puro, mais preferivelmente pelo menos 97% puro, ainda mais preferivelmente pelo menos 98% puro, o mais preferivelmente pelo menos 99%, e ainda o mais preferivelmente pelo menos 99,5% puro em peso. Os polinucleotídeos da presente invenção estão preferivelmente em uma forma substancialmente pura. Em particular, prefere-se que os polinucleotídeos descritos aqui estejam

na "forma essencialmente pura", isto é, que a preparação de polinucleotídeo seja essencialmente livre de outro material de polinucleotídeo com o qual ele é naturalmente ou recombinantemente associado. Aqui, o termo "polinucleotídeo substancialmente puro" é sinônimo dos termos

5 "polinucleotídeo isolado" e "polipeptídeo na forma isolada." Os polinucleotídeos podem ser de origem genômica, cDNA, RNA, semi-sintética, sintética, ou quais combinações dos mesmos.

Seqüência de codificação: Quando usado aqui o termo "seqüência de codificação" significa uma seqüência de nucleotídeo, que

10 diretamente especifica a seqüência de aminoácido de seu produto de proteína. Os limites da seqüência de codificação geralmente são determinados por uma matriz de leitura aberta, que geralmente começa com o códon de partida ATG ou códons de partida alternativos como GTG e TTG e terminado com um códon de parada como TAA, TAG, e TGA. A seqüência de codificação pode

15 ser uma seqüência de DNA, cDNA, ou de nucleotídeo recombinante.

Seqüência de codificação de polipeptídeo maduro: O termo "seqüência de codificação de polipeptídeo maduro" é definido aqui como uma seqüência de nucleotídeos que codifica um polipeptídeo maduro tendo atividade de endoglucanase.

20 **cDNA:** O termo "cDNA" é definido aqui como uma molécula de DNA que pode ser preparada por transcrição reversa de uma molécula de mRNA maduro, emendada, obtida de uma célula eucariótica. cDNA não tem as seqüências de íntrons que estão geralmente presentes no DNA genômico correspondente. O transcrito de RNA primário inicial é um precursor para

25 mRNA que é processado através de uma série de etapas antes de aparecer como um mRNA emendado maduro. Estas etapas incluem a remoção de seqüências de íntron por um processo chamado "emenda". cDNA derivado de mRNA não apresenta, assim, quaisquer seqüências de íntron.

Construção de ácido nucleico: O termo "construção de ácido

nucleico" como usado aqui refere-se a uma molécula de ácido nucleico, de filamento único ou duplo, que é isolada de um gene de ocorrência natural ou que é modificada para conter segmentos de ácidos nucleicos em um modo que não iria de outra forma existir na natureza. O termo construção de ácido nucleico é sinônimo do termo "cassete de expressão" quando a construção de ácido nucleico contém as seqüências de controle requeridas para expressão de uma seqüência de codificação da presente invenção.

Seqüência de controle: O termo "seqüência de controle" é definido aqui para incluir todos componentes, que são necessários ou vantajosos para a expressão de um polinucleotídeo codificando um polipeptídeo da presente invenção. Cada seqüência de controle pode ser nativa ou estranha para a seqüência de nucleotídeos codificando o polipeptídeo ou nativa ou estranha para cada outro. Tais seqüências de controle incluem, mas não são limitados a um líder, seqüência de poliadenilação, seqüência de pró-peptídeo, promotor, seqüência de peptídeo sinal, e terminador de transcrição. Em um mínimo, as seqüências de controle incluem um promotor, e sinais de parada transcripcional e translacional. As seqüências de controle podem ser providas com ligadores para o fim de introduzir sítios de restrição específicos facilitando a ligação das seqüências de controle com a região codificada da seqüência de nucleotídeos codificando um polipeptídeo.

Ligado operativamente: O termo "ligado operativamente" denota aqui uma configuração em que uma seqüência de controle é colocada em uma posição apropriada relativa na seqüência de codificação da seqüência de polinucleotídeo de modo que a seqüência de controle dirige a expressão da seqüência de codificação de um polipeptídeo.

Expressão: O termo "expressão" inclui qualquer etapa envolvida na produção do polipeptídeo incluindo, mas não limitado a transcrição, modificação pós-transcripcional, tradução, modificação pós-

translacional, e secreção.

Vetor de expressão: O termo "vetor de expressão" é definido aqui como uma molécula de DNA linear ou circular que compreende um polinucleotídeo codificando um polipeptídeo da presente invenção, e que é
5 ligado operativamente a nucleotídeos adicionais que provê sua expressão.

Célula hospedeira: O termo "célula hospedeira", como usado aqui, inclui qualquer tipo de célula que é suscetível para transformação, transfecção, transdução, e outros com uma construção de ácido nucleico ou vetor de expressão compreendendo um polinucleotídeo da presente invenção.

Modificação: O termo "modificação" significa aqui qualquer
10 modificação química do polipeptídeo consistindo do polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10; ou uma sequência homóloga do mesmo; assim como manipulação genética do DNA codificando este polipeptídeo. A modificação pode incluir substituições,
15 deleções e/ou inserções de um ou mais aminoácidos assim como substituições de uma ou mais cadeias laterais de aminoácido.

Variante artificial: Quando usado aqui, o termo "variante artificial" significa um polipeptídeo tendo atividade de endoglucanase produzida por um organismo expressando uma sequência de nucleotídeo
20 modificada da sequência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9; ou uma sequência homóloga do mesmo. A sequência de nucleotídeo modificada é obtida através de intervenção humana por modificação da sequência de nucleotídeo descrita na SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9; ou uma
25 sequência homóloga do mesmo.

Descrição Detalhada da Invenção

Polipectídeos tendo atividade de endoglucanase

Em um primeiro aspecto, a presente invenção refere-se a polipeptídeos isolados compreendendo uma sequência de aminoácido que tem

um grau de identidade no polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10, de pelo menos 60%, preferivelmente pelo menos 65%, mais preferivelmente pelo menos 70%, mais preferivelmente pelo menos 75%, mais preferivelmente pelo menos 80%, mais preferivelmente pelo menos 85%, ainda mais preferivelmente pelo menos 90%, o mais preferivelmente pelo menos 95%, e ainda o mais preferivelmente pelo menos 96%, 97%, 98%, ou 99%, que tem atividade de endoglucanase (abaixo "polipeptídeos homólogos"). Em um aspecto preferido, os polipeptídeos homólogos têm uma seqüência de aminoácido que difere por dez aminoácidos, preferivelmente por cinco aminoácidos, mais preferivelmente por quatro aminoácidos, ainda mais preferivelmente por três aminoácidos, o mais preferivelmente por dois aminoácidos, e ainda o mais preferivelmente por um aminoácido do polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10.

Um polipeptídeo da presente invenção preferivelmente compreende a seqüência de aminoácido de SEQ ID NO: 4 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em um aspecto preferido, um polipeptídeo compreende a seqüência de aminoácido de SEQ ID NO: 4. Em outro aspecto preferido, um polipeptídeo compreende o polipeptídeo maduro de SEQ ID NO: 4. Em outro aspecto preferido, um polipeptídeo compreende aminoácidos 17 a 389 de SEQ ID NO: 4, ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo compreende aminoácidos 17 a 389 de SEQ ID NO: 4. Em outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido de SEQ ID NO: 4 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido de SEQ ID NO: 4. Em outro aspecto preferido, um polipeptídeo consiste do polipeptídeo maduro de

SEQ ID NO: 4. Em outro aspecto preferido, um polipeptídeo consiste de aminoácidos 17 a 389 de SEQ ID NO: 4 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo consiste de aminoácidos 17 a 389 de SEQ ID NO: 4.

Um polipeptídeo da presente invenção preferivelmente também compreende a seqüência de aminoácido de SEQ ID NO: 6 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em um aspecto preferido, um polipeptídeo compreende a seqüência de aminoácido de SEQ ID NO: 6. Em outro aspecto preferido, um polipeptídeo compreende o polipeptídeo maduro de SEQ ID NO: 6. Em outro aspecto preferido, um polipeptídeo compreende aminoácidos 16 a 397 de SEQ ID NO: 6, ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo compreende aminoácidos 16 a 397 de SEQ ID NO: 6. Em outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido de SEQ ID NO: 6 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido de SEQ ID NO: 6. Em outro aspecto preferido um polipeptídeo consiste do polipeptídeo maduro de SEQ ID NO: 6. Em outro aspecto preferido, um polipeptídeo consiste de aminoácidos 16 a 397 de SEQ ID NO: 6 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo consiste de aminoácidos 16 a 397 de SEQ ID NO: 6.

Um polipeptídeo da presente invenção preferivelmente também compreende a seqüência de aminoácido de SEQ ID NO: 8 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em um aspecto preferido, um polipeptídeo compreende a

seqüência de aminoácido de SEQ ID NO: 8. Em outro aspecto preferido, um polipeptídeo compreende o polipeptídeo maduro de SEQ ID NO: 8. Em outro aspecto preferido, um polipeptídeo compreende aminoácidos 22 a 429 de SEQ ID NO: 8, ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo compreende aminoácidos 22 a 429 de SEQ ID NO: 8. Em outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido de SEQ ID NO: 8 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido de SEQ ID NO: 8. Em outro aspecto preferido, um polipeptídeo consiste do polipeptídeo maduro de SEQ ID NO: 8. Em outro aspecto preferido, um polipeptídeo consiste de aminoácidos 22 a 429 de SEQ ID NO: 8 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo consiste de aminoácidos 22 a 429 de SEQ ID NO: 8.

Um polipeptídeo da presente invenção preferivelmente também compreende a seqüência de aminoácido de SEQ ID NO: 10 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em um aspecto preferido, um polipeptídeo compreende a seqüência de aminoácido de SEQ ID NO: 10. Em outro aspecto preferido, um polipeptídeo compreende o polipeptídeo maduro de SEQ ID NO: 10. Em outro aspecto preferido, um polipeptídeo compreende aminoácidos 25 a 421 de SEQ ID NO: 10, ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo compreende aminoácidos 25 a 421 de SEQ ID NO: 10. Em outro aspecto preferido, um polipeptídeo consiste da seqüência de aminoácido de SEQ ID NO: 10 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um

polipeptídeo consiste da seqüência de aminoácido de SEQ ID NO: 10. Em outro aspecto preferido, um polipeptídeo consiste do polipeptídeo maduro de SEQ ID NO: 10. Em outro aspecto preferido, um polipeptídeo consiste de aminoácidos 25 a 421 de SEQ ID NO: 10 ou uma variante alélica do mesmo; ou um fragmento do mesmo que tem atividade de endoglucanase. Em outro aspecto preferido, um polipeptídeo consiste de aminoácidos 25 a 421 de SEQ ID NO: 10.

Em um segundo aspecto, a presente invenção refere-se a polipeptídeos isolados tendo atividade de endoglucanase que são codificados por polinucleotídeos que hibridizam sob condições de estringência muito baixa, preferivelmente condições de estringência baixa, mais preferivelmente condições de estringência média, mais preferivelmente condições de estringência média a alta, ainda mais preferivelmente condições de alta estringência, e o mais preferivelmente condições de estringência muito alta com (i) a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9, (ii) a seqüência de cDNA contida em uma seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3 ou SEQ ID NO: 9 ou a seqüência de DNA genômico compreendendo a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 5 ou SEQ ID NO: 7, (iii) uma subsequência de (i) ou (ii), ou (iv) um filamento complementar de (i), (ii), ou (iii) (J. Sambrook, E, F, Fritsch, e T, Maniatis, 1989, *Molecular Cloning, A Laboratory Manual*, 2^a. edição, Cold Spring Harbor, New York). Uma subsequência da seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9 contém pelo menos 100 nucleotídeos contíguos ou preferivelmente pelo menos 200 nucleotídeos contíguos. Além disso, a subsequência pode codificar um fragmento de polipeptídeo que tem atividade de endoglucanase. Em um aspecto preferido, a seqüência de codificação de polipeptídeo maduro contém os nucleotídeos 67 a 1185 de SEQ ID NO: 3. Em

outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 84 a 1229 de SEQ ID NO: 5. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 77 a 1300 de SEQ ID NO: 7. Em outro aspecto preferido, a
5 sequência de codificação de polipeptídeo maduro contém os nucleotídeos 73 a 1488 de SEQ ID NO: 9. Em outro aspecto preferido, o filamento complementar é o filamento complementar de comprimento completo da sequência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9.

10 A sequência de nucleotídeos de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9; ou a subsequência do mesmo; assim como a sequência de aminoácido de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10; ou um fragmento do mesmo; pode ser usada para projetar uma sonda de ácido nucleico para identificar e clonar DNA
15 codificando polipeptídeos tendo atividade de endoglucanase de cepas de gêneros ou espécies diferentes de acordo com os métodos bem conhecidos na arte. Em particular, tais sondas podem ser usadas para hibridização com o genômico ou cDNA do gênero ou espécie de interesse, após os procedimentos padrões de Southern blotting, a fim de identificar e isolar o gene
20 correspondente no mesmo. Tais sondas podem ser consideravelmente mais curtas do que a sequência completa, mas devem ter pelo menos 14, preferivelmente pelo menos 25, mais preferivelmente pelo menos 35, e o mais preferivelmente pelo menos 70 nucleotídeos de comprimento. No entanto, prefere-se que a sonda de ácido nucleico seja de pelo menos 100 nucleotídeos
25 de comprimento. Por exemplo, a sonda de ácido nucleico pode ter pelo menos 200 nucleotídeos, preferivelmente pelo menos 300 nucleotídeos, mais preferivelmente pelo menos 400 nucleotídeos, ou o mais preferivelmente pelo menos 500 nucleotídeos de comprimento. Outras sondas mais longas podem ser usadas, por exemplo, sondas de ácido nucleico que tem pelo menos 600

nucleotídeos, pelo menos preferivelmente pelo menos 700 nucleotídeos, mais preferivelmente pelo menos 800 nucleotídeos, ou o mais preferivelmente pelo menos 900 nucleotídeos de comprimento. Ambas as sondas de DNA e RNA podem ser usadas. As sondas são tipicamente rotuladas para detectar o gene correspondente (por exemplo, com ^{32}P , ^3H , ^{35}S , biotina, ou avidina). Tais sondas são englobadas pela presente invenção.

Uma biblioteca de DNA ou cDNA genômico preparado a partir destes outros organismos pode, assim, ser triada para DNA que hibridiza com as sondas descritas acima e que codifica um polipeptídeo tendo atividade de endoglucanase. DNA genômico ou outro de tais outros organismos pode ser separado por agarose ou eletroforese gel de poliacrilamida, ou outras técnicas de separação, DNA das bibliotecas ou o DNA separado pode ser transferido para e imobilizado em nitrocelulose ou outro material de veículo apropriado. A fim de identificar um clone ou DNA que é homólogo com SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9; ou uma subsequência do mesmo; o material de veículo é preferivelmente usado em um Southern blot.

Para os fins da presente invenção, a hibridização indica que a sequência de nucleotídeos hibridiza para uma sonda de ácido nucleico rotulada correspondente a sequência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9; a sequência de cDNA contida em uma sequência de codificação de polipeptídeo maduro de SEQ ID NO: 3 ou SEQ ID NO: 9 ou a sequência de DNA genômico compreendendo a sequência de codificação de polipeptídeo maduro de SEQ ID NO: 5 ou SEQ ID NO: 7; seu filamento complementar; ou uma subsequência do mesmo; sob condições de estrigência muito alta para muito baixa. Moléculas em a sonda de ácido nucleico hibridiza sob estas condições podem ser detectadas usando, por exemplo, filme de raios X.

Em um aspecto preferido, a sonda de ácido nucleico é a

seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3. Em outro aspecto preferido, a sonda de ácido nucleico contém os nucleotídeos 67 a 1185 de SEQ ID NO: 3. Em outro aspecto preferido, a sonda de ácido nucleico é uma seqüência de polinucleotídeo que codifica o polipeptídeo de SEQ ID NO: 4, ou a subsequência do mesmo. Em outro aspecto preferido, a sonda de ácido nucleico é SEQ ID NO: 3. Em outro aspecto preferido, a sonda de ácido nucleico é a seqüência de polinucleotídeo contida em plasmídeo pCIC161 que está contida em *E. coli* NRRL B-30902, em que a seqüência de polinucleotídeo do mesmo codifica um polipeptídeo tendo atividade de endoglucanase. Em outro aspecto preferido, a sonda de ácido nucleico é a região de codificação de polipeptídeo maduro contida em plasmídeo pCIC161 que está contido em *E. coli* NRRL B-30902.

Em outro aspecto preferido, a sonda de ácido nucleico é a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 5. Em outro aspecto preferido, a sonda de ácido nucleico contém os nucleotídeos 84 a 1229 de SEQ ID NO: 5. Em outro aspecto preferido, a sonda de ácido nucleico é uma seqüência de polinucleotídeo que codifica o polipeptídeo de SEQ ID NO: 6, ou uma subsequência do mesmo. Em outro aspecto preferido, a sonda de ácido nucleico é SEQ ID NO: 5. Em outro aspecto preferido, a sonda de ácido nucleico é a seqüência de polinucleotídeo contida em plasmídeo pCIC453 que está contido em *E. coli* NRRL B-30903, em que a seqüência de polinucleotídeo do mesmo codifica um polipeptídeo tendo atividade de endoglucanase. Em outro aspecto preferido, a sonda de ácido nucleico é a região de codificação de polipeptídeo maduro contido em plasmídeo pCIC453 que está contido em *E. coli* NRRL B-30903.

Em outro aspecto preferido, a sonda de ácido nucleico é a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 7. Em outro aspecto preferido, a sonda de ácido nucleico contém os nucleotídeos 77 a 1300 de SEQ ID NO: 7. Em outro aspecto preferido, a sonda de ácido

nucleico é uma seqüência de polinucleotídeo que codifica o polipeptídeo de SEQ ID NO: 8, ou uma subsequência do mesmo. Em outro aspecto preferido, a sonda de ácido nucleico é SEQ ID NO: 7. Em outro aspecto preferido, a sonda de ácido nucleico é a seqüência de polinucleotídeo contido em

5 plasmídeo pCIC486 que está contido em *E. coli* NRRL B-30904, em que a seqüência de polinucleotídeo do mesmo codifica um polipeptídeo tendo atividade de endoglucanase. Em outro aspecto preferido, a sonda de ácido nucleico é a região de codificação de polipeptídeo maduro contida em plasmídeo pCIC486 que está contido em *E. coli* NRRL B-30904.

10 Em outro aspecto preferido, a sonda de ácido nucleico é a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 9. Em outro aspecto preferido, a sonda de ácido nucleico contém os nucleotídeos 73 a 1468 de SEQ ID NO: 9. Em outro aspecto preferido, a sonda de ácido nucleico é uma seqüência de polinucleotídeo que codifica o polipeptídeo de

15 SEQ ID NO: 10, ou uma subsequência do mesmo. Em outro aspecto preferido, a sonda de ácido nucleico é SEQ ID NO: 9. Em outro aspecto preferido, a sonda de ácido nucleico é a seqüência de polinucleotídeo contida em plasmídeo pPBCeI5C que está contido em *E. coli* NRRL B-30900N, em que a seqüência de polinucleotídeo do mesmo codifica um polipeptídeo tendo

20 atividade de endoglucanase. Em outro aspecto preferido, a sonda de ácido nucleico é a região de codificação de polipeptídeo maduro contida em plasmídeo pPBCeI5C que está contido em *E. coli* NRRL B-30900N.

Para sondas longas de pelo menos 100 nucleotídeos em comprimento, condições de estringência muito alta a muito baixa são

25 definidas como pré-hibridização e hibridização a 42°C em 5X SSPE, 0,3% de SDS, 200 µg/ml de DNA de esperma de salmão desnaturado e cisalhado, e ou 25% de formamida para estringências baixa e muito baixa, 35% de formamida para estringências média e meio alta, ou 50% de formamida para estringências alta e muito alta, seguindo procedimentos padrões de Southern blotting

durante 12 a 24 horas de modo ótimo.

Para sondas longas de pelo menos 100 nucleotídeos em comprimento, o material de veículo é finalmente lavado três vezes cada durante 15 minutos usando 2X SSC, 0,2% de SDS preferivelmente pelo menos a 45°C (estringência muito baixa), mais preferivelmente pelo menos a 50°C (estringência baixa), mais preferivelmente pelo menos a 55°C (estringência média), mais preferivelmente pelo menos a 60°C (estringência média a alta), ainda mais preferivelmente pelo menos a 65°C (estringência alta), e o mais preferivelmente pelo menos a 70°C (estringência muito alta).

Para sondas curtas que têm de cerca de 15 nucleotídeos e cerca de 70 nucleotídeos em comprimento, condições de stringência são definidas como pré-hibridização, hibridização, e lavagem pós-hibridização a cerca de 5°C a cerca de 10°C abaixo do T_m calculado usando o cálculo de acordo com Bolton e McCarthy (1962, *Proceedings of the National Academy of Sciences USA* 48:1390) em 0,9 M de NaCl, 0,09 M de Tris-HCl pH 7,6, 6 mM de EDTA, 0,5% de NP-40, 1X de solução de Denhardt, 1 mM pirofosfato de sódio, 1 mM fosfato monobásico de sódio, 0,1 mM ATP, e 0,2 mg de RNA de levedura por ml seguindo procedimentos padrões de Southern blotting durante 12 a 24 horas de modo ótimo.

Para sondas curtas que têm cerca de 15 nucleotídeos a cerca de 70 nucleotídeos de comprimento, o material de veículo é lavado uma vez em 6X SSC mais 0,1% de SDS durante 15 minutos e duas vezes cada durante 15 minutos usando 6X SSC a 5°C a 10°C abaixo o T_m calculado.

Em um terceiro aspecto, a presente invenção refere-se a polipeptídeos isolados codificados por polinucleotídeos compreendendo ou consistindo de seqüências de nucleotídeos que têm um grau de identidade para a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9 de pelo menos 60%, preferivelmente pelo menos 65%, mais preferivelmente pelo menos 70%,

mais preferivelmente pelo menos 75%, mais preferivelmente pelo menos 80%, mais preferivelmente pelo menos 85%, mais preferivelmente pelo menos 90%, ainda mais preferivelmente pelo menos 95%, e o mais preferivelmente pelo menos 97% de identidade, que codifica um polipeptídeo ativo. Em um aspecto preferido, a seqüência de codificação de polipeptídeo maduro contém os nucleotídeos 67 a 1185 de SEQ ID NO: 3. Em outro aspecto preferido, a seqüência de codificação de polipeptídeo maduro contém os nucleotídeos 84 a 1229 de SEQ ID NO: 5. Em outro aspecto preferido, a seqüência de codificação de polipeptídeo maduro contém os nucleotídeos 77 a 1300 de SEQ ID NO: 7. Em outro aspecto preferido, a seqüência de codificação de polipeptídeo maduro contém os nucleotídeos 73 a 1468 de SEQ ID NO: 9. Ver seção de polinucleotídeo aqui.

Em um quarto aspecto, a presente invenção refere-se a variantes artificiais compreendendo uma substituição, deleção, e/ou inserção de um ou mais aminoácidos do polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10; ou uma seqüência homóloga do mesmo. Preferivelmente, as mudanças de aminoácido são de uma natureza menor, isto é as substituições de aminoácido conservadoras ou inserções que não afetam significativamente da duplicação e/ou atividade da proteína; deleções pequenas, tipicamente de um a cerca de 30 aminoácidos; extensões pequenas de amino ou carboxil-terminal, como um resíduo de metionina amino-terminal; um peptídeo ligador pequeno de até cerca de 20-25 resíduos ou uma extensão pequena que facilita a purificação por mudança da carga líquida ou outra função, como um trato poli-histidina, um epítipo antigênico ou um domínio de ligação.

Exemplos de substituições conservadoras estão dentro do grupo de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutâmico e ácido aspártico), aminoácidos polares (glutamina e asparagina), aminoácidos hidrofóbicos (leucina, isoleucina e valina),

aminoácidos aromáticos (fenilalanina, triptofano e tirosina), e aminoácidos pequenos (glicina, alanina, serina, treonina e metionina). Substituições de aminoácido que não alteram geralmente a atividade específica são conhecidas na técnica e são descritas, por exemplo, por H. Neurath e R. L. Hill, 1979, *In*,
 5 *The Proteins*, Academic Press, New York. As mudanças de ocorrência mais comuns são Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Asn/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, e Asp/Gly.

Além dos 20 aminoácidos padrões, aminoácidos não padrões
 10 (como 4-hidroxiprolina, 6-N-metil lisina, ácido 2-aminoisobutírico, isovalina, e alfa-metil serina) podem ser substituídos com resíduos de aminoácido de um polipeptídeo de tipo selvagem. Um número limitado de aminoácidos não conservadores, aminoácidos que não são codificados pelo código genético, e aminoácidos desnaturados podem ser substituídos por resíduos de
 15 aminoácido. "Aminoácidos desnaturados" foram modificados após a síntese de proteína, e/ou têm uma estrutura química em sua(s) cadeia (s) lateral (ais) diferente (s) das dos aminoácidos padrões. Aminoácidos desnaturados podem ser quimicamente sintetizados, e preferivelmente, são comercialmente disponíveis, e incluem ácido piperídico, ácido tiazolidina carboxílico,
 20 dehidroprolina, 3- e 4-metilprolina, e 3,3-dimetilprolina.

Alternativamente, as mudanças de aminoácido são de tal natureza que as propriedades físico-químicas do polipeptídeos são alteradas. Por exemplo, as mudanças de aminoácido podem melhorar a estabilidade térmica do polipeptídeo, alterar a especificidade do substrato, mudar o pH
 25 ótimo, e outros.

Aminoácidos essenciais no polipeptídeo parental podem ser identificados de acordo com os procedimentos conhecidos na arte, como mutagênese sítio-dirigida ou mutagênese de varredura de alanina (Cunningham e Wells, 1989, *Science* 244: 1081 - 1085). Na última técnica,

mutações de alanina única são introduzidas em cada resíduo na molécula, e as moléculas mutantes resultantes são testadas para atividade biológica (isto é, atividade de endoglucanase) para identificar resíduos de aminoácido que são críticos para a atividade da molécula. Ver também, Hilton *et al.*, 1996, *J. Biol. Chem.* 271: 4699 - 4708. O sítio ativo da enzima ou outra interação biológica também pode ser determinado por análise física de estrutura, como determinado por tais técnicas como ressonância magnética nuclear, cristalografia, difração de elétrons, ou rotulação de fotoafinidade, em conjunto com mutação de aminoácidos de sítio de contato putativo. Ver, por exemplo, de Vos *et al.*, 1992, *Science* 255: 306 - 312; Smith *et al.*, 1992, *J. Mol. Biol.* 224: 899 - 904; Wlodaver *et al.*, 1992, *FEBS Lett.* 309: 59 - 64. As identidades de aminoácidos essenciais também podem ser deduzidas da análise de identidades com polipeptídeos que estão relacionados com um polipeptídeo de acordo com a invenção.

Substituições de aminoácido único ou múltiplo podem ser feitas e testadas usando métodos conhecidos de mutagênese, recombinação, e/ou embaralhamento, seguido por um procedimento de triagem relevante, como os descritos por Reidhaar-Olson e Sauer, 1988, *Science* 241: 53 - 57; Bowie e Sauer, 1989, *Proc. Natl. Acad. Sci. USA* 86: 2152 - 2156; WO 95/17413; ou WO 95/22625. Outros métodos que podem ser usados incluem PCR propenso a erros, apresentação na superfície de fagos (por exemplo, Lowman *et al.*, 1991, *Biochem.* 30: 10832 - 10837; patente U.S. 5.223.409; WO 92/06204), e mutagênese região -dirigida (Derbyshire *et al.*, 1986, *Gene* 46: 145; Ner *et al.*, 1988, *DNA* 7: 127).

Métodos de mutagênese/ embaralhamento podem ser combinados com métodos de triagem automatizados, de elevada produção, para detectar a atividade de polipeptídeos mutagenizados clonados expressados por células hospedeiras (Ness *et al.*, 1999, *Nature Biotechnology* 17: 893 - 896). Moléculas de DNA mutagenizadas que codificam

polipeptídeos ativos podem ser recuperadas das células hospedeiras e rapidamente seqüenciadas usando métodos padrões na arte. Estes métodos permitem a determinação rápida da importância de resíduos de aminoácido individuais em um polipeptídeo de interesse, e podem ser aplicados em

5 polipeptídeos de estrutura desconhecida.

O número total de substituições de aminoácido, deleções e/ou inserções do polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10, como aminoácidos 17 a 389 de SEQ ID NO: 2, aminoácidos 16 a 397 de SEQ ID NO: 6, aminoácidos 22 a 429 de SEQ ID

10 NO: 8, ou aminoácidos 25 a 421 de SEQ ID NO: 10, é 10, preferivelmente 9, mais preferivelmente 8, mais preferivelmente 7, mais preferivelmente no máximo 6, mais preferivelmente 5, mais preferivelmente 4, ainda mais preferivelmente 3, o mais preferivelmente 2, e ainda o mais preferivelmente 1.

Fontes de Polipeptídeos Tendo Atividade de Endoglucanase

Um polipeptídeo da presente invenção pode ser obtido a partir de microorganismos de qualquer gênero. Para os fins da presente invenção, o termo "obtido a partir de" como usado aqui em conexão com uma dada fonte deve significar que o polipeptídeo codificado por uma seqüência de nucleotídeo é produzido pela fonte ou por uma cepa em que a seqüência de

15 nucleotídeos a partir da fonte foi inserida. Em um aspecto preferido, o polipeptídeo obtido de uma fonte dada é secretado extracelularmente.

Um polipeptídeo tendo atividade de endoglucanase da presente invenção pode ser um polipeptídeo bacteriano. Por exemplo, o polipeptídeo pode ser um polipeptídeo bacteriano gram-positivo como um polipeptídeo de

25 *Bacillus*, *Streptococcus*, *Streptomyces*, *Staphylococcus*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Clostridium*, *Geobacillus*, ou *Oceanobacillus* tendo atividade [enzima], ou um polipeptídeo bacteriano gram-negativo como um de *E. coli*, *Pseudomonas*, *Salmonella*, *Campylobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Ilyobacter*, *Neisseria*, ou polipeptídeo de

Ureaplasma tendo atividade de endoglucanase.

Em um aspecto preferido, o polipeptídeo é um polipeptídeo de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus ciauxii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus*
 5 *lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, ou *Bacillus thuringiensis* tendo atividade de endoglucanase.

Em outro aspecto preferido, o polipeptídeo é um polipeptídeo de *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*,
 10 ou *Streptococcus equi* subsp. *Zooepidemicus* tendo atividade de endoglucanase.

Em outro aspecto preferido, o polipeptídeo é um polipeptídeo de *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, ou *Streptomyces lividans* tendo atividade de
 15 endoglucanase.

Um polipeptídeo tendo atividade de endoglucanase da presente invenção também pode ser um polipeptídeo de fungo, e mais preferivelmente um polipeptídeo de levedura como um polipeptídeo de *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, ou *Yarrowia*
 20 tendo atividade de endoglucanase; ou mais preferivelmente um polipeptídeo de fungo filamentoso como um polipeptídeo de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Chrysosporium*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*, *Schizophyllum*,
 25 *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, ou *Trichoderma* tendo atividade de endoglucanase.

Em um aspecto preferido, o polipeptídeo é um polipeptídeo de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*,

Saccharomyces norbensis, ou *Saccharomyces oviformis* tendo atividade de endoglucanase.

Em outro aspecto preferido, o polipeptídeo é um polipeptídeo de *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*,
 5 *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium tropicum*, *Chrysosporium*
merdarium, *Chrysosporium inops*, *Chrysosporium pannicola*,
Chrysosporium queenslandicum, *Chrysosporium zonatum*, *Fusarium*
 10 *bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium*
 15 *torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia spededonium*, *Thielavia*
 20 *setosa*, *Thielavia subthermophifa*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, ou *Trichoderma viride* tendo atividade de endoglucanase.

Em outro aspecto preferido, o polipeptídeo é um polipeptídeo de *Penicillium brasilianum*, *Penicillium camembertii*, *Penicillium*
 25 *capsulatum*, *Penicillium chrysogenum*, *Penicillium citreonigrum*, *Penicillium citrinum*, *Penicillium claviforme*, *Penicillium corylophilum*, *Penicillium crustosum*, *Penicillium digitatum*, *Penicillium expansum*, *Penicillium funiculosum*, *Penicillium glabrum*, *Penicillium granulatum*, *Penicillium griseofulvum*, *Penicillium islandicum*, *Penicillium italicum*, *Penicillium*

janthinellum, *Penicillium lividum*, *Penicillium megasporum*, *Penicillium melinii*, *Penicillium notatum*, *Penicillium oxalicum*, *Penicillium puberulum*, *Penicillium purpurescens*, *Penicillium pururogenum*, *Penicillium roquefortii*, *Penicillium rugulosum*, *Penicillium spinulosum*, *Penicillium waksmanii*, ou

5 *Penicillium* sp. tendo atividade de endoglucanase.

Em um aspecto mais preferido, o polipeptídeo é um polipeptídeo de *Myceliophthora thermophila*, e o mais preferivelmente um polipeptídeo *Myceliophthora thermophila* CBS 111.65, por exemplo, o polipeptídeo de SEQ ID NO: 4, ou o polipeptídeo maduro do mesmo.

10 Em outro aspecto mais preferido, o polipeptídeo é um polipeptídeo de basidiomicete CBS 495.95, por exemplo, o polipeptídeo de SEQ ID NO: 6, ou o polipeptídeo maduro do mesmo.

Em outro aspecto mais preferido, o polipeptídeo é um polipeptídeo de basidiomiceto CBS 494.95, por exemplo, o polipeptídeo de

15 SEQ ID NO: 8, ou o polipeptídeo maduro do mesmo.

Em outro aspecto mais preferido, o polipeptídeo é um polipeptídeo de *Penicillium brasilianum*, e o mais preferivelmente um polipeptídeo de *Penicillium brasilianum* IBT 20888, por exemplo, o polipeptídeo de SEQ ID NO: 10, ou o polipeptídeo maduro do mesmo.

20 Será entendido que para a espécie anteriormente mencionada a invenção engloba tanto os estados perfeitos como os imperfeitos, e outros equivalentes taxonômicos, por exemplo, anamorfos, sem levar em conta o nome da espécie pela qual eles são conhecidos. O versado na técnica irá facilmente reconhecer a identidade de equivalentes apropriados.

25 Cepas destas espécies são facilmente acessíveis para o público em várias coleções de cultura, como American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM), Centraalbureau Voor Schimmelcultures (CBS), e Agricultural Research Service Patent Culture Collection, Northern Regional Research

Center (NRRL).

Além disso, tais polipeptídeos podem ser identificados e obtidos de outras fontes incluindo microorganismos isolados de natureza (por exemplo, solo, compostagem, água, etc.) usando as sondas mencionadas
5 acima. Técnicas para isolar os microorganismos de habitats naturais são bem conhecidas na arte. O polinucleotídeo pode então ser obtido por uma triagem similar de um cDNA genômico ou biblioteca deste microorganismo. Uma vez uma seqüência de polinucleotídeo codificando um polipeptídeo foi detectada com a (s) sonda (s), o polinucleotídeo pode ser isolado ou clonado por
10 utilização de técnicas que são bem conhecidas para o versado na técnica (ver, por exemplo, Sambrook *et al.*, 1989, *supra*).

Polipeptídeos da presente invenção também incluem polipeptídeos fusionados ou polipeptídeos de fusão cliváveis em que outro polipeptídeo é fusionado no N-terminal ou o C-término do
15 polipeptídeo ou fragmento do mesmo. Um polipeptídeo fusionado é produzido por fusão de uma seqüência de nucleotídeo (ou uma porção do mesmo) codificando outro polipeptídeo para uma seqüência de nucleotídeo (ou uma porção do mesmo) da presente invenção. Técnicas para produzir os polipeptídeos de fusão são conhecidas na arte, e incluem ligar a seqüência de
20 codificação codificando os polipeptídeos de modo elas ficam na armação e a que expressão do polipeptídeo fusionado está sob o controle do mesmo promotor (s) e terminador.

Um polipeptídeo de fusão pode ainda compreender um sítio de clivagem. Quando da secreção da proteína de fusão, o sítio é clivado
25 liberando o polipeptídeo tendo atividade de endoglucanase da proteína de fusão.

Exemplos de sítios de clivagem incluem, mas não são limitados a um sítio de Kex2 que codifica o dipeptídeo Lys-Arg (Martin *et al.*, 2003, *J. Ind. Microbiol. Biotechnol.* 3: 568 - 76; Svetina *et al.*, 2000, *J.*

Biotechnol. 76: 245 - 251; Rasmussen-Wilson *et al.*, 1997, *Appl. Environ. Microbiol.* 63: 3488 - 3493; Ward *et al.*, 1995, *Biotechnology* 13: 498 - 503; e Contreras *et al.*, 1991, *Biotechnology* 9: 378 - 381), um sítio Ile-(Glu ou Asp)-Gly-Arg, que é clivado por uma protease de fator Xa após o resíduo de arginina (Eaton *et al.*, 1986, *Biochem.* 25: 505 - 512); um sítio Asp-Asp-Asp-Asp-Lys, que é clivado por uma enteroquinase após a lisina (Collins-Racie *et al.*, 1995, *Biotechnology* 13: 982 - 987); um sítio His-Tyr-Glu ou sítio His-Tyr-Asp, que é clivado por Genenase I (Carter *et al.*, 1989, *Proteins: Structure, Function, and Genetics* 6: 240 - 248); um sítio Leu-Val-Pro-Arg-Gly-Ser, que é clivado por trombina após o Arg (Stevens, 2003, *Drug Discovery World* 4: 35- 48); um sítio Glu-Asn-Leu-Tyr-Phe-Gln-Gly, que é clivado por protease TEV após o Gln (Stevens, 2003, *supra*); e um sítio Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro, que é clivado por uma forma geneticamente engenheirada de 3C protease de rinovírus humano após o Gln (Stevens, 2003, *supra*).

Polinucleotídeos

A presente invenção também refere-se a um polinucleotídeo isolado compreendendo ou consistindo de uma sequência de nucleotídeo que codifica um polipeptídeo da presente invenção tendo atividade de endoglucanase.

Em um aspecto preferido, a sequência de nucleotídeo compreende ou consiste de SEQ ID NO: 3. Em outro aspecto mais preferido, a sequência de nucleotídeo compreende ou consiste da sequência contida em plasmídeo pCIC161 que está contido em *E. coli* NRRL B-30902. Em outro aspecto preferido, a sequência de nucleotídeo compreende ou consiste da região de codificação de polipeptídeo maduro de SEQ ID NO: 3. Em outro aspecto preferido, a sequência de nucleotídeo compreende ou consiste de nucleotídeos 67 a 1185 de SEQ ID NO: 3. Em outro aspecto mais preferido, a sequência de nucleotídeo compreende ou consiste da região de codificação de

polipeptídeo maduro contida em plasmídeo pCIC161 que está contido em *E. coli* NRRL B-30902. A presente invenção também engloba as seqüências de nucleotídeos que codificam polipeptídeos compreendendo ou consistindo da seqüência de aminoácido de SEQ ID NO: 4 ou o polipeptídeo maduro do mesmo, que difere de SEQ ID NO: 3 ou a seqüência de codificação de polipeptídeo maduro do mesmo por virtude da degeneração do código genético. A presente invenção também refere-se as subsequências de SEQ ID NO: 3 que codifica fragmentos de SEQ ID NO: 4 que tem atividade de endoglucanase.

Em outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste de SEQ ID NO: 5. Em outro aspecto mais preferido, a seqüência de nucleotídeo compreende ou consiste da seqüência contida em plasmídeo pCIC453 que está contido em *E. coli* NRRL B-30903. Em outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste da região de codificação de polipeptídeo maduro de SEQ ID NO: 5. Em outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste de nucleotídeos 84 a 1229 de SEQ ID NO: 5. Em outro aspecto mais preferido, a seqüência de nucleotídeo compreende ou consiste da região de codificação de polipeptídeo maduro contida em plasmídeo pCIC453 que está contido em *E. coli* NRRL B-30903. A presente invenção também engloba as seqüências de nucleotídeos que codifica polipeptídeos compreendendo ou consistindo da seqüência de aminoácido de SEQ ID NO: 6 ou o polipeptídeo maduro do mesmo, que difere de SEQ ID NO: 5 ou a seqüência de codificação de polipeptídeo maduro do mesmo por virtude da degeneração do código genético. A presente invenção também refere-se às subsequências de SEQ ID NO: 5 que codifica fragmentos de SEQ ID NO: 6 que tem atividade de endoglucanase.

Em outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste de SEQ ID NO: 7. Em outro aspecto mais preferido,

a seqüência de nucleotídeo compreende ou consiste da seqüência contida em plasmídeo pCIC486 que está contido em *E. coli* NRRL B-30904. Em outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste da região de codificação de polipeptídeo maduro de SEQ ID NO: 7. Em outro

5 aspecto preferido, a seqüência de nucleotídeo compreende ou consiste de nucleotídeos 77 a 1300 de SEQ ID NO: 7. Em outro aspecto mais preferido, a seqüência de nucleotídeo compreende ou consiste da região de codificação do polipeptídeo maduro contida em plasmídeo pCIC486 que está contido em *E. coli* NRRL B-30904. A presente invenção também engloba seqüências de

10 nucleotídeos que codificam polipeptídeos compreendendo ou consistindo da seqüência de aminoácido de SEQ ID NO: 8 ou o polipeptídeo maduro do mesmo, que difere de SEQ ID NO: 7 ou a seqüência de codificação de polipeptídeo maduro do mesmo por virtude da degeneração do código genético. A presente invenção também refere-se às subsequências de SEQ ID

15 NO: 7 que codificam fragmentos de SEQ ID NO: 8 quem tem atividade de endoglucanase.

Em outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste de SEQ ID NO: 9. Em outro aspecto mais preferido, a seqüência de nucleotídeo compreende ou consiste da seqüência contida em

20 plasmídeo pPBCel5C que está contido em *E. coli* NRRL B-30900N. Em outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste da região de codificação de polipeptídeo maduro de SEQ ID NO: 9. Em outro aspecto preferido, a seqüência de nucleotídeo compreende ou consiste de nucleotídeos 73 a 1488 de SEQ ID NO: 9. Em outro aspecto mais preferido, a

25 seqüência de nucleotídeo compreende ou consiste da região de codificação de polipeptídeo maduro contida em plasmídeo pPBCel5C que está contido em *E. coli* NRRL B-30900N. A presente invenção também engloba seqüências de nucleotídeos que codificam polipeptídeos compreendendo ou consistindo da seqüência de aminoácido de SEQ ID NO: 10 ou o polipeptídeo maduro do

mesmo, que difere de SEQ ID NO: 9 ou a seqüência de codificação de polipeptídeo maduro do mesmo por virtude da degeneração do código genético. A presente invenção também refere-se às subsequências de SEQ ID NO: 9 que codificam fragmentos de SEQ ID NO: 10 que tem atividade de endoglucanase.

A presente invenção também refere-se a polinucleotídeos mutantes compreendendo ou consistindo de pelo menos uma mutação em uma seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9, em que a seqüência de nucleotídeo mutante codifica o polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10. Em um aspecto preferido, o polipeptídeo maduro contém os aminoácidos 17 a 389 de SEQ ID NO: 4. Em outro aspecto preferido, o polipeptídeo maduro contém os aminoácidos 16 a 397 de SEQ ID NO: 6. Em outro aspecto preferido, o polipeptídeo maduro contém os aminoácidos 22 a 429 de SEQ ID NO: 8. Em outro aspecto preferido, o polipeptídeo maduro contém os aminoácidos 25 a 421 de SEQ ID NO: 10.

As técnicas usadas para isolar ou clonar um polinucleotídeo codificando um polipeptídeo são conhecidas e incluem isolamento de DNA genômico, preparação de cDNA, ou uma combinação dos mesmos. A clonagem dos polinucleotídeos da presente invenção a partir deste DNA genômico pode se efetuada, por exemplo, por uso da reação de cadeia polimerase bem conhecida (PCR) ou triagem de anticorpo de bibliotecas de expressão para detectar fragmentos de DNA clonados com aspectos estruturais compartilhados. Ver, por exemplo, Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Outros procedimentos de amplificação de ácido nucleico como reação de cadeia ligase (LCR), transcrição ativada ligada (LAT) e amplificação com base na seqüência de nucleotídeo (NASBA) podem ser usados. Os polinucleotídeos podem ser clonados a partir de uma cepa de *Myceliophthora thermophila*

CBS 117.65, basidiomicete CBS 494.95, ou basidiomicete CBS 495.95, ou outra ou um organismo relacionado e assim, por exemplo, pode ser uma espécie alélica ou variante da região de codificação de polipeptídeo da sequência de nucleotídeo.

5 A presente invenção também refere-se a polinucleotídeos isolados compreendendo ou consistindo de sequências de nucleotídeos que tem um grau de identidade para a sequência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9 de pelo menos 60%, preferivelmente pelo menos 65%, mais preferivelmente
10 pelo menos 70%, mais preferivelmente pelo menos 75%, mais preferivelmente pelo menos 80%, mais preferivelmente pelo menos 85%, mais preferivelmente pelo menos 90%, ainda mais preferivelmente pelo menos 95%, e o mais preferivelmente pelo menos 97% de identidade, que codifica um polipeptídeo ativo. Em um aspecto preferido, a sequência de
15 codificação de polipeptídeo maduro contém os nucleotídeos 67 a 1185 de SEQ ID NO: 3. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 84 a 1229 de SEQ ID NO: 5. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 77 a 1300 de SEQ ID NO: 7. Em outro aspecto
20 preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 73 a 1488 de SEQ ID NO: 9.

 Modificação de uma sequência de nucleotídeo codificando um polipeptídeo da presente invenção pode ser necessária para a síntese de polipeptídeos substancialmente similares para o polipeptídeo. O termo
25 "substancialmente similar" para o polipeptídeo refere-se às formas não de ocorrência natural do polipeptídeo. Estes polipeptídeos podem diferir em algum modo engenheirado do polipeptídeo isolado de sua fonte nativa, por exemplo, variantes artificiais que diferem na atividade específica, termoestabilidade, pH ótimo, ou outros. A sequência variante pode ser

construída com base na seqüência de nucleotídeos apresentada como a região de codificação de polipeptídeo de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9, por exemplo, uma subsequência do mesmo, e/ou por introdução de substituições de nucleotídeo que não dão lugar a outra seqüência de aminoácido do polipeptídeo codificado pela seqüência de nucleotídeo, mas que corresponde ao uso de códon do organismo hospedeiro destinado para a produção da enzima, ou por introdução de substituições de nucleotídeo que podem dar lugar a uma seqüência de aminoácido diferente. Para uma descrição geral de substituição de nucleotídeo, ver, por exemplo, Ford *et al.*, 1991, *Protein Expression and Purification* 2: 95 - 107.

Será evidente para o versado na técnica que tais substituições podem ser feitas fora das regiões críticas na função da molécula e assim resultar em um polipeptídeo ativo. Resíduos de aminoácido essenciais para a atividade do polipeptídeo codificado por um polinucleotídeo isolado da invenção, e assim preferivelmente não sujeitos à substituição, podem ser identificados de acordo com os procedimentos conhecidos na arte, como mutagênese sítio-dirigida ou mutagênese de varredura de alanina (ver, por exemplo, Cunningham e Wells, 1989, *supra*). Na última técnica, mutações são introduzidas em cada resíduo positivamente carregado na molécula, e as moléculas mutantes resultantes são testadas para atividade de endoglucanase para identificar os resíduos de aminoácido que são críticos para a atividade da molécula. Sítios de interação de substrato-enzima também podem ser determinados por análise da estrutura tri-dimensional como determinado por tais técnicas como análise de ressonância magnética nuclear, cristalografia ou rotulação de fotoafinidade (ver, por exemplo, de Vos *et al.*, 1992, *supra*; Smith *et al.*, 1992, *supra*; Wlodaver *et al.*, 1992, *supra*).

A presente invenção também refere-se a polinucleotídeos isolados codificando um polipeptídeo da presente invenção, que hibridiza sob condições de estringência muito baixa, preferivelmente condições de

estringência baixa, mais preferivelmente condições de estringência média, mais preferivelmente condições de estringência média a alta, ainda mais preferivelmente condições de estringência alta, e o mais preferivelmente condições de estringência muito alta com (i) a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9, (ii) a seqüência de cDNA contida em uma seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3 ou SEQ ID NO: 9 ou a seqüência de DNA genômico compreendendo a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 5 ou SEQ ID NO: 7, ou (iii) um filamento complementar de (i) ou (ii); ou variantes alélicas e subsequências dos mesmos (Sambrook *et al.*, 1989, *supra*), como definido aqui. Em um aspecto preferido, a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3 contêm os nucleotídeos 67 a 1185. Em outro aspecto preferido, a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 5 contêm os nucleotídeos 84 a 1229. Em outro aspecto preferido a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 7 contêm os nucleotídeos 77 a 1300. Em outro aspecto preferido, a seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 9 contêm os nucleotídeos 73 a 1768.

A presente invenção também refere-se a polinucleotídeos isolados obtidos por (a) hibridizar uma população de DNA sob condições de estringência muito baixa, baixa, média, média a alta, alta, ou muito alta com (i) a seqüência de codificação de polipeptídeo maduro SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, ou SEQ ID NO: 9, (ii) a seqüência de cDNA contida em uma seqüência de codificação de polipeptídeo maduro de SEQ ID NO: 3 ou SEQ ID NO: 9 ou a seqüência de DNA genômico compreendendo a seqüência de codificação de polipeptídeo maduro SEQ ID NO: 5 ou SEQ ID NO: 7, ou (iii) um filamento complementar de (i) ou (ii); e (b) isolamento do polinucleotídeo de hibridização, que codifica um polipeptídeo tendo atividade de endoglucanase. Em um aspecto preferido, a seqüência de codificação de

polipeptídeo maduro de SEQ ID NO: 3 contêm os nucleotídeos 67 a 1185. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro de SEQ ID NO: 5 contêm os nucleotídeos 84 a 1229. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro de SEQ ID NO: 7 contêm os nucleotídeos 77 a 1300. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro de SEQ ID NO: 9 contêm os nucleotídeos 73 a 1768.

Construções de Ácido Nucleico

A presente invenção também refere-se a construções de ácido nucleico compreendendo um polinucleotídeo isolado da presente invenção ligado operativamente a uma ou mais sequências de controle que dirigem a expressão da sequência de codificação em uma célula hospedeira apropriada sob condições compatíveis com as sequências de controle.

Um polinucleotídeo isolado codificando um polipeptídeo da presente invenção pode ser manipulado em uma variedade de modos para prover a expressão do polipeptídeo. Manipulação da sequência de polinucleotídeo antes de sua inserção em um vetor pode ser desejável ou necessária dependendo do vetor de expressão. As técnicas para modificar as sequências de polinucleotídeos utilizando métodos de DNA recombinantes são bem conhecidas na arte.

A sequência de controle pode ser uma sequência promotora apropriada, uma sequência de nucleotídeo que é reconhecida por uma célula hospedeira para expressão de um polinucleotídeo codificando um polipeptídeo da presente invenção. A sequência promotora contém sequências de controle transcripcional que mediam a expressão do polipeptídeo. O promotor pode ser qualquer sequência de nucleotídeo que mostra atividade transcripcional na célula hospedeira de escolha incluindo promotores mutantes, truncados, e híbridos, e pode ser obtido a partir de genes codificando polipeptídeos extracelulares ou intracelulares ou homólogos ou

heterólogos na célula hospedeira.

Exemplos de promotores apropriados para dirigir a transcrição das construções de ácido nucleico da presente invenção, especialmente em uma célula bacteriana hospedeira, são os promotores obtidos a partir do operon de *E. coli lac*, gene de agarase (dagA) de *Streptomyces coelicolor*, gene de levansucrase (sacB) de *Bacillus subtilis*, gene de alfa-amilase (amyL) de *Bacillus licheniformis*, gene de amilase maltogênica (amyM) de *Bacillus stearothermophilus*, gene de alfa-amilase (amyQ) de *Bacillus amyloliquefaciens*, gene de penicilinase (penP) de *Bacillus licheniformis*, genes de xylA e xylB de *Bacillus subtilis*, e gene de beta-lactamase procariótica (Villa-Kamaroff *et al.*, 1978, *Proceedings of the National Academy of Sciences USA* 75: 3727 - 3731), assim como o promotor *tac* (DeBoer *et al.*, 1983, *Proceedings of the National Academy of Sciences USA* 80: 21 - 25). Outros promotores são descritos em "Useful proteins from recombinant bacteria" in *Scientific American*, 1980, 242: 74 - 94; e em Sambrook *et al.*, 1989, *supra*.

Exemplos de promotores apropriados para dirigir a transcrição das construções de ácido nucleico da presente invenção em uma célula hospedeira de fungo filamentoso são promotores obtidos a partir dos genes de amilase TAKA de *Aspergillus oryzae*, proteinase aspártica de *Rhizomucor miehei*, alfa-amilase neutra de *Aspergillus niger*, alfa-amilase de ácido estável de *Aspergillus niger*, glucoamilase (glaA) de *Aspergillus niger* ou *Aspergillus awamori*, lipase de *Rhizomucor miehei*, protease alcalina de *Aspergillus oryzae*, triose fosfato isomerase de *Aspergillus oryzae*, acetamidase de *Aspergillus nidulans*, amiloglucosidase de *Fusarium venenatum* (WO 00/56900), Daria de *Fusarium venenatum* (WO 00/56900), Quinn de *Fusarium venenatum* (WO 00/56900), protease de tipo tripsina de *Fusarium oxysporum* (WO 96/00787), beta-glucosidase de *Trichoderma reesei*, celobiohidrolase I de *Trichoderma reesei*, celobiohidrolase II de *Trichoderma*

reesei, endoglucanase I de *Trichoderma reesei*, endoglucanase II de *Trichoderma reesei*, endoglucanase III de *Trichoderma reesei*, endoglucanase IV de *Trichoderma reesei*, endoglucanase V de *Trichoderma reesei*, xilanase I de *Trichoderma reesei*, xilanase II de *Trichoderma reesei*, beta-xilosidase de *Trichoderma reesei*, assim como o promotor NA2-tpi (um híbrido dos promotores dos genes para alfa-amilase neutra de *Aspergillus niger* e triose fosfato isomerase de *Aspergillus oryzae*); e promotores mutante, truncados, e híbridos dos mesmos.

Em um hospedeiro de levedura, promotores utilizáveis são obtidos a partir dos genes para enolase de *Saccharomyces cerevisiae* (ENO-1), galactoquinase (GAL1) de *Saccharomyces cerevisiae*, desidrogenase/gliceraldeído-3-fosfato desidrogenase de álcool de *Saccharomyces cerevisiae* (ADH1, ADH2/GAP), triose fosfato isomerase (TPI) de *Saccharomyces cerevisiae*, metalotioneína (CUP1) de *Saccharomyces cerevisiae*, e 3-fosfoglicerato quinase de *Saccharomyces cerevisiae*. Outros promotores utilizáveis para células hospedeiras de levedura são descritos por Romanos *et al.*, 1992, *Yeast* 8: 423 - 488.

A seqüência de controle também pode ser uma seqüência de terminador de transcrição apropriada, uma seqüência reconhecida por uma célula hospedeira para terminar a transcrição. A seqüência do terminador é ligada operativamente no término 3' da seqüência de nucleotídeo codificando o polipeptídeo. Qualquer terminador que é funcional na célula hospedeira de escolha pode ser usado na presente invenção.

Os terminadores preferidos para células hospedeiras de fungo filamentoso são obtidos dos genes para TAKA amilase de *Aspergillus oryzae*, glucoamilase de *Aspergillus niger*, antranilato sintase de *Aspergillus nidulans*, alfa-glucosidase de *Aspergillus niger*, e protease de tipo tripsina de *Fusarium oxysporum*.

Os terminadores preferidos para células hospedeiras de

levedura são obtidos dos genes para enolase de *Saccharomyces cerevisiae*, citocromo C (CYC1) de *Saccharomyces cerevisiae*, e gliceraldeído-3-fosfato dehidrogenase de *Saccharomyces cerevisiae*. Outros terminadores utilizáveis para células hospedeiras de levedura são descritos por Romanos *et al.*, 1992, *supra*.

A seqüência de controle também pode ser uma seqüência líder apropriada, uma região não traduzida de um mRNA que é importante para a tradução pela célula hospedeira. A seqüência líder é ligada operativamente no término 5' da seqüência de nucleotídeo codificando o polipeptídeo. Qualquer seqüência líder, isto é funcional na célula hospedeira de escolha, pode ser usada na presente invenção.

Os líderes preferidos para células hospedeiras de fungo filamentosos são obtidos a partir dos genes para TAKA amilase de *Aspergillus oryzae* e triose fosfato isomerase de *Aspergillus nidulans*.

Os líderes apropriados para células hospedeiras de levedura são obtidos a partir dos genes para enolase de *Saccharomyces cerevisiae* (ENO-1), 3-fosfoglicerato quinase de *Saccharomyces cerevisiae*, alfa-fator de *Saccharomyces cerevisiae*, e dehidrogenase/gliceraldeído-3-fosfato dehidrogenase de álcool (ADH2/GAP) de *Saccharomyces cerevisiae*.

A seqüência de controle também pode ser uma seqüência de poliadenilação, uma seqüência ligada operativamente no término 3' da seqüência de nucleotídeo e que, quando transcrita, é reconhecida pela célula hospedeira como um sinal para adicionar resíduos poliadenosina para mRNA transcrito. Qualquer seqüência de poliadenilação que é funcional na célula hospedeira de escolha pode ser usada na presente invenção.

As seqüências de poliadenilação preferidas para células hospedeiras de fungos filamentosos são obtidas dos genes para TAKA amilase de *Aspergillus oryzae*, glucoamilase de *Aspergillus niger*, antranilato sintase de *Aspergillus nidulans*, protease de tipo tripsina de *Fusarium*

oxysporum, e alfa-glucosidase de *Aspergillus niger*.

As seqüências de poliadenilação utilizáveis para células hospedeiras de levedura são descritas por Guo e Sherman, 1995, *Molecular Cellular Biology* 15: 5983 - 5990.

5 A seqüência de controle também pode ser uma região de codificação de peptídeo sinal que codifica para uma seqüência de aminoácido ligada no término amino de um polipeptídeo e dirige o polipeptídeo codificado na via secretória da célula. A extremidade 5' da seqüência de codificação da seqüência de nucleotídeo pode inerentemente conter uma
10 região de codificação de peptídeo sinal naturalmente ligado em uma matriz de leitura da tradução com o segmento da região de codificação que codifica o polipeptídeo secretado. Alternativamente, a extremidade 5' da seqüência de codificação pode conter uma região de codificação de peptídeo sinal que é estranho para a seqüência de codificação. A região de codificação de peptídeo
15 sinal estranho pode ser requerida onde a seqüência de codificação não contém naturalmente uma região de codificação de peptídeo sinal. Alternativamente, a região de codificação de peptídeo sinal estranho pode simplesmente substituir a região de codificação de peptídeo sinal natural a fim de melhorar a secreção do polipeptídeo. No entanto, qualquer região de codificação de
20 peptídeo sinal que dirige o polipeptídeo expresso na via secretória de uma célula hospedeira de escolha, isto é, secretado em um meio de cultura, pode ser usado na presente invenção.

 As regiões eficazes de codificação de peptídeo sinal para células hospedeiras bacterianas são as regiões de codificação de peptídeo sinal
25 obtidas dos genes para amilase maltogênica de *Bacillus* NCIB 11837, alfa-amilase de *Bacillus stearothermophilus*, subtilisina de *Bacillus licheniformis*, beta-lactamase de *Bacillus licheniformis*, proteases neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM) e prsA de *Bacillus subtilis*. Outros peptídeos de sinal são descritos por Simonen e Palva, 1993, *Microbiological*

Reviews 57: 109 - 137.

As regiões eficazes de codificação de peptídeo sinal para células hospedeiras de fungo filamentoso são as regiões de codificação de peptídeo sinal obtidas dos genes para TAKA amilase de *Aspergillus oryzae*,
 5 amilase neutra de *Aspergillus niger*, glucoamilase de *Aspergillus niger*,
 proteinase aspártica de *Rhizomucor miehei*, celulase de *Humicola insolens*,
 endoglucanase V de *Humicola insolens*, e lipase de *Humicola lanuginosa*.

Peptídeos de sinal utilizáveis para células hospedeiras de levedura são obtidos dos genes para fator-alfa de *Saccharomyces cerevisiae* e
 10 invertase de *Saccharomyces cerevisiae*. Outras regiões de codificação de peptídeo de sinal utilizáveis são descritas por Romanos *et al.*, 1992, *supra*.

Em um aspecto preferido, o peptídeo sinal compreende ou consiste de aminoácidos 1 a 16 de SEQ ID NO: 4. Em outro aspecto preferido, a região de codificação de peptídeo sinal contém os nucleotídeos 19
 15 a 69 de SEQ ID NO: 3.

Em outro aspecto preferido, o peptídeo sinal compreende ou consiste de aminoácidos 1 a 15 de SEQ ID NO: 6. Em outro aspecto preferido, a região de codificação de peptídeo sinal compreende ou consiste de nucleotídeos 39 a 83 de SEQ ID NO: 5.

20 Em outro aspecto preferido, o peptídeo sinal compreende ou consiste de aminoácidos 1 a 21 de SEQ ID NO: 8. Em outro aspecto preferido, a região de codificação de peptídeo sinal compreende ou consiste de nucleotídeos 14 a 76 de SEQ ID NO: 7.

25 Em outro aspecto preferido, o peptídeo sinal compreende ou consiste de aminoácidos 1 a 16 de SEQ ID NO: 10. Em outro aspecto preferido, a região de codificação de peptídeo sinal compreende ou consiste de nucleotídeos 1 a 48 de SEQ ID NO: 9.

A sequência de controle também pode ser uma região de codificação de pró-peptídeo que codifica para uma sequência de aminoácido

posicionada no término amino de um polipeptídeo. O polipeptídeo resultante é conhecido como uma pró-enzima ou pró-polipeptídeo (ou um zimogênio em alguns casos). Um pró-peptídeo geralmente é inativo e pode ser convertido para um polipeptídeo ativo maduro por clivagem catalítica ou autocatalítica do pró-peptídeo do pró-polipeptídeo. A região codificada de pró-peptídeo pode ser obtida a partir dos genes para protease alcalina de *Bacillus subtilis* (*aprE*), protease neutra de *Bacillus subtilis* (*nprT*), fator-alfa de *Saccharomyces cerevisiae*, proteinase aspártica de *Rhizomucor miehei*, e lacase de *Myceliophthora thermophila* (WO 95/33836).

Em um aspecto preferido, o pró-peptídeo compreende ou consiste de aminoácidos 17 a 24 de SEQ ID NO: 10. Em outro aspecto preferido, a região de codificação de pró-peptídeo compreende ou consiste de nucleotídeos 49 a 72 de SEQ ID NO: 9.

Onde ambas regiões de pró-peptídeo e peptídeo sinal são apresentadas no término amino de um polipeptídeo, a região de pró-peptídeo está posicionada próximo ao término amino de um polipeptídeo e a região de peptídeo sinal está posicionada próximo ao término amino da região de pró-peptídeo.

Também pode ser desejável adicionar seqüências reguladoras que permitem a regulação da expressão do polipeptídeo relativo ao crescimento da célula hospedeira. Exemplos de sistemas reguladores são os que causam a expressão do gene a ser ligado e desligado em resposta a um estímulo químico ou físico, incluindo a presença de um composto regulador. Os sistemas reguladores em sistemas procarióticos incluem os sistemas operadores *lac*, *tac*, e *trp*. Em levedura, o sistema ADH2 ou sistema GAL1 também pode ser usado. Em fungos filamentosos, o promotor TAKA alfa-amilase, promotor glucoamilase de *Aspergillus niger*, e promotor glucoamilase de *Aspergillus oryzae* também podem ser usados como seqüências reguladoras. Outros exemplos de seqüências reguladoras são as

que permitem a amplificação do gene. Em sistemas eucarióticos, estes incluem o gene redutase di-hidrofolato que é amplificado na presença de metotrexato, e os genes metalotioneína que são amplificados com metais pesados. Nestes casos, a sequência de nucleotídeo codificando o polipeptídeo
5 deve ser ligada operativamente com a sequência reguladora.

Vetores de Expressão

A presente invenção também refere-se aos vetores de expressão recombinante compreendendo um polinucleotídeo da presente invenção, um promotor, e sinais de parada de transcrição e translação. Os
10 vários ácidos nucleicos e sequências de controle descritos aqui podem ser unidos juntos para produzir um vetor de expressão recombinante que pode incluir um ou mais sítios de restrição convenientes para permitir a inserção ou substituição da sequência de nucleotídeo codificando o polipeptídeo em tais sítios. Alternativamente, uma sequência de polinucleotídeo da presente
15 invenção pode ser expressa ao inserir a sequência de nucleotídeo ou uma construção de ácido nucleico compreendendo a sequência em um vetor apropriado para expressão. Ao criar o vetor de expressão, a sequência de codificação está localizada no vetor de modo que a sequência de codificação é ligada operativamente com as sequências de controle apropriadas para
20 expressão.

O vetor de expressão recombinante pode ser qualquer vetor (por exemplo, um plasmídeo ou vírus) que pode ser convenientemente submetido a procedimentos de DNA recombinantes e pode ocasionar a expressão da sequência de nucleotídeo. A escolha do vetor tipicamente irá
25 depender da compatibilidade do vetor com a célula hospedeira em que o vetor deve ser introduzido. Os vetores podem ser lineares ou plasmídeos circulares fechados.

O vetor pode ser um vetor autonomamente replicante, isto é, um vetor que existe como uma entidade extracromossômica, cuja replicação é

independente da replicação cromossômica, por exemplo, um plasmídeo, um elemento extracromossômico, um minicromossômico, ou um cromossomo artificial. O vetor pode conter qualquer meio para assegurar uma auto-replicação. Alternativamente, o vetor pode ser um que, quando introduzido na

5 célula hospedeira, é integrado no genoma e replicado junto com o cromossomo (s) em que ele foi integrado. Além disso, um vetor único ou plasmídeo ou dois ou mais vetores ou plasmídeos que juntos contém o DNA total a ser introduzido no genoma da célula hospedeira, ou um transposon pode ser usado.

10 Os vetores da presente invenção preferivelmente contêm um ou mais marcadores selecionáveis que permitem uma seleção fácil de transformadores, transfectados, transduzidos, ou células semelhantes. Um marcador selecionável é um gene cujo produto provê resistência a biocidas ou viral, resistência a metais pesados, fototrofia a autotrofos, e semelhantes.

15 Exemplos de marcadores bacterianos selecionáveis são os genes *dal* de *Bacillus subtilis* ou *Bacillus licheniformis*, ou marcadores que conferem resistência a antibióticos como ampicilina, canamicina, cloranfenicol, ou resistência a tetraciclina. Os marcadores apropriados para células hospedeiras de levedura são ADE2, HIS3, LEU2, LYS2, MET3,

20 TRP1, e URA3. Os marcadores selecionáveis para uso em uma célula hospedeira de fungo filamentoso incluem, mas não são limitados a, *amdS* (acetamidase), *argB* (ornitina carbamoiltransferase), *bar* (fosfinotricina acetiltransferase), *hph* (higromicina fosfotransferase), *niaD* (nitrato redutase), *pyrG* (orotidina-5'- fosfato decarboxilase), *sC* (sulfato adeniltransferase), e

25 *trpC* (antranilato sintase), assim como equivalentes dos mesmos. Preferidos para uso em uma célula de *Aspergillus* são os genes *amdS* e *pyrG* de *Aspergillus nidulans* ou *Aspergillus oryzae* e o gene *bar* de *Streptomyces hygroscopicus*.

Os vetores da presente invenção preferivelmente contêm um

elemento (s) que permite a integração do vetor no genoma de célula hospedeira ou replicação autônoma do vetor na célula independente do genoma.

Para integração no genoma de célula hospedeira, o vetor pode se basear na sequência de polinucleotídeo codificando o polipeptídeo ou qualquer outro elemento do vetor para integração no genoma por recombinação homóloga ou não homóloga. Alternativamente, o vetor pode conter sequências de nucleotídeos adicionais para dirigir a integração por recombinação homóloga no genoma da célula hospedeira em um local (s) preciso no (s) cromossomo (s). Para aumentar a probabilidade de integração em um local preciso, os elementos integracionais devem conter preferivelmente um número suficiente de ácidos nucleicos, como 100 a 10.000 pares de base, preferivelmente 400 a 10.000 pares de base, e o mais preferivelmente 800 a 10.000 pares de base, tendo um grau elevado de identidade com a sequência de marcação correspondente para melhorar a probabilidade de recombinação homóloga. Os elementos integracionais podem ser qualquer sequência, isto é, homóloga com a sequência de marcação no genoma da célula hospedeira. Além disso, os elementos integracionais podem ser as sequências de nucleotídeo de não codificação e de codificação. Por outro lado, o vetor pode ser integrado no genoma da célula hospedeira por recombinação não homóloga.

Para replicação autônoma, o vetor pode ainda compreender uma origem de replicação permitindo ao vetor replicar autonomamente na célula hospedeira em questão. A origem de replicação pode ser qualquer replicador de plasmídeo mediando a replicação autônoma que funciona em uma célula. O termo “origem de replicação ” ou “replicador de plasmídeo” é definido aqui como uma sequência de nucleotídeo que permite a um plasmídeo ou vetor replicar *in vivo*.

Exemplos de origens bacterianas de replicação são as de

origens de replicação de plasmídeos pBR322, pUC19, pACYC177, e pACYC184 permitindo a replicação em *E. coli*, e pUB110, pE194, pTA1060, e pAM β 1 permitindo a replicação em *Bacillus*.

Exemplos de origens de replicação para uso em uma célula hospedeira de levedura são a origem de replicação de 2 microns, ARS1, ARS4, a combinação de ARS1 e CEN3, e a combinação de ARS4 e CEN6.

Exemplos de origens de replicação utilizáveis em uma célula de fungo filamentoso são AMA1 e ANS1 (Gems *et al.*, 1991, *Gene* 98: 61 - 67; Cullen *et al.*, 1987, *Nucleic Acids Research* 15: 9163 - 9175; WO 00/24883). Isolamento do gene AMA1 e construção de plasmídeo ou vetores compreendendo o gene podem ser alcançados de acordo com os métodos descritos em WO 00/24883.

Mais do que uma cópia de um polinucleotídeo da presente invenção pode ser inserida em uma célula hospedeira para aumentar a produção do produto de gene. Um aumento no número de cópia do polinucleotídeo pode ser obtido por integração de pelo menos uma cópia adicional da sequência no genoma de célula hospedeira ou por inclusão de um gene de marcador selecionável amplificável com o polinucleotídeo onde células contendo cópias amplificadas do gene de marcador selecionável, e assim cópias adicionais do polinucleotídeo, podem ser selecionadas por cultivo das células na presença do agente selecionável apropriado.

Os procedimentos usados para ligar os elementos descritos acima para construir os vetores de expressão recombinante da presente invenção são bem conhecidos do versado na técnica (ver, por exemplo, Sambrook *et al.*, 1989, *supra*).

Células Hospedeiras

A célula hospedeira pode ser qualquer célula utilizável na produção recombinante de um polipeptídeo da presente invenção, por exemplo, uma procariótica ou uma eucariótica.

A célula hospedeira pró-cariótica pode ser qualquer bactéria gram positiva ou uma bactéria gram negativa. Bactérias gram positivas incluem, mas não limitadas a *Bacillus*, *Streptococcus*, *Streptomyces*, *Staphylococcus*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Clostridium*,
 5 *Geobacillus*, e *Oceanobacillus*. Bactérias gram negativas incluem, mas não são limitadas a, *E. coli*, *Pseudomonas*, *Salmonella*, *Campylobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Llyobacter*, *Neisseria*, e *Ureaplasma*.

A célula bacteriana hospedeira pode ser qualquer célula de
 10 *Bacillus*. As células de *Bacillus* utilizáveis na prática da presente invenção incluem, mas não são limitadas a células de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*. *Bacillus*
 15 *stearothermophilus*, *Bacillus subtilis*, e *Bacillus thuringiensis*.

Em um aspecto preferido, a célula bacteriana hospedeira é uma célula de *Bacillus amyloliquefaciens*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus* ou *Bacillus subtilis*. Em um aspecto mais preferido, a célula bacteriana hospedeira é uma célula de *Bacillus*
 20 *amyloliquefaciens*. Em outro aspecto mais preferido, a célula bacteriana hospedeira é uma célula de *Bacillus clausii*. Em outro aspecto mais preferido, a célula bacteriana hospedeira é uma célula de *Bacillus licheniformis*. Em outro aspecto mais preferido, a célula bacteriana hospedeira é uma célula de *Bacillus subtilis*.

25 A célula bacteriana hospedeira também pode ser qualquer célula de *Streptococcus*. As células *Streptococcus* utilizáveis na prática da presente invenção incluem, mas não são limitadas a, *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, e *Streptococcus equi* subsp. *Zooepidemicus*.

Em um aspecto preferido, a célula bacteriana hospedeira é uma célula de *Streptococcus equisimilis*. Em outro aspecto preferido, a célula bacteriana hospedeira é uma célula de *Streptococcus pyogenes*. Em outro aspecto preferido, a célula bacteriana hospedeira é uma célula de *Streptococcus uberis*. Em outro aspecto preferido, a célula bacteriana hospedeira é uma célula de *Streptococcus equi* subsp. *Célula Zooepidemicus*.

A célula bacteriana hospedeira também pode ser qualquer célula de *Streptomyces*. As células de *Streptomyces* utilizáveis na prática da presente invenção incluem, mas não são limitadas a, *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, e *Streptomyces lividans*.

Em um aspecto preferido, a célula bacteriana hospedeira é uma célula de *Streptomyces achromogenes*. Em outro aspecto preferido, a célula bacteriana hospedeira é uma célula de *Streptomyces avermitilis*. Em outro aspecto preferido, a célula bacteriana hospedeira é uma célula de *Streptomyces coelicolor*. Em outro aspecto preferido, a célula bacteriana hospedeira é uma célula de *Streptomyces griseus*. Em outro aspecto preferido, a célula bacteriana hospedeira é uma célula de *Streptomyces lividans*.

A introdução de DNA em uma célula de *Bacillus* pode, por exemplo, ser efetuada por transformação de protoplastos (ver, por exemplo, Chang e Cohen, 1979, *Molecular General Genetics* 168: 111 - 115), por uso em células competentes (ver, por exemplo, Young e Spizizen, 1961, *Journal of Bacteriology* 81: 823 - 829, ou Dubnau e Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56: 209 - 221), por eletroporação (ver, por exemplo, Shigekawa e Dower, 1988, *Biotechniques* 6: 742 - 751), ou por conjugação (ver, por exemplo, Koehler e Thome, 1987, *Journal of Bacteriology* 169: 5271 - 5278). A introdução de DNA em uma célula de *E coli* pode, por exemplo, ser efetuada por transformação de protoplastos (ver, por exemplo, Hanahan, 1983, *J. Mol. Biol.* 166: 557 - 580) ou eletroporação (ver, por

exemplo, Dower *et al.*, 1988, *Nucleic Acids Res.* 16: 6127 - 6145). A introdução de DNA em uma célula de *Streptomyces* pode, por exemplo, ser efetuada por transformação de protoplastos e eletroporação (ver, por exemplo, Gong *et al.*, 2004, *Folia Microbiol. (Praha)* 49: 399 - 405), por conjugação (ver, por exemplo, Mazodier *et al.*, 1989, *J. Bacteriol.* 171: 3583 - 3585), ou por transdução (ver, por exemplo, Burke *et al.*, 2001, *Proc. Natl. Acad. Sci. USA* 98: 6289 - 6294). A introdução de DNA em uma célula de *Pseudomonas* pode, por exemplo, ser efetuada por eletroporação (ver, por exemplo, Choi *et al.*, 2006, *J. Microbiol. Methods* 64: 391 - 397) ou por conjugação (ver, por exemplo, Pinedo e Smets, 2005, *Appl. Environ. Microbiol.* 71: 51 - 57). A introdução de DNA em uma célula de *Streptococcus* pode, por exemplo, ser efetuada por competência natural (ver, por exemplo, Perry e Kuramitsu, 1981 *Infect. Immun.* 32: 1295 - 1297), por transformação de protoplastos (ver, por exemplo, Catt e Jollick, 1991, *Microbios*, 68: 189 - 2070, por eletroporação (ver, por exemplo, Buckley *et al.*, 1999, *Appl. Environ. Microbiol.* 65: 3800 - 3804) ou por conjugação (ver, por exemplo, Clewell, 1981, *Microbiol. Rev.* 45: 409 - 436). No entanto, qualquer método conhecido para introduzir DNA em uma célula hospedeira pode ser usado.

A célula hospedeira também pode ser uma eucariótica, como uma célula de mamífero, inseto, planta, ou fungo.

Em um aspecto preferido, a célula hospedeira é uma célula de fungo. "Fungos", como usado aqui, incluem os filos Ascomycota, Basidiomycota, Chytridiomycota, e Zygomycota (como definidos por Hawksworth *et al.*, In, *Ainsworth and Bisby's Dictionary of The Fungi*, 8^a. edição, 1995, CAB International, University Press, Cambridge, UK) assim como Oomycota (como citado em Hawksworth *et al.*, 1995, *supra*, página 171) e todos os fungos mitospóricos (Hawksworth *et al.*, 1995, *supra*).

Em um aspecto mais preferido, a célula hospedeira de fungo é uma célula de levedura. "Levedura" como usada aqui inclui levedura

ascosporogenosa (Endomicetales), levedura basidiosporogenosa, e levedura pertencendo aos Fungi imperfecti (Blastomicetes). Porque a classificação de levedura pode mudar no futuro, para os fins desta invenção, levedura deve ser definida como descrito em *Biology and Activities of Yeast* (Skinner, F. A.,
 5 Passmore, S. M., e Davenport, R. R., eds, *Soc. App. Bacteriol. Symposium* série nº. 9, 1980).

Em um aspecto ainda mais preferido, a célula hospedeira de levedura é uma célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, ou *Yarrowia*.

10 Em um aspecto o mais preferido, a célula hospedeira de levedura é uma célula de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, ou *Saccharomyces oviformis*. Em outro aspecto o mais preferido, a célula hospedeira de levedura
 15 é uma célula de *Kluyveromyces lactis*. Em outro aspecto o mais preferido, a célula hospedeira de levedura é uma célula de *Yarrowia lipolytica*.

Em outro aspecto mais preferido, a célula hospedeira de fungo é uma célula de fungo filamentoso. "Fungos filamentosos" incluem todas as formas filamentosas da subdivisão Eumycota e Oomycota (como definido por
 20 Hawksworth *et al.*, 1995, *supra*). Os fungos filamentosos geralmente são caracterizados por uma parede miceliana composta de quitina, celulose, glucano, quitosana, manano, e outros polissacarídeos complexos. Crescimento vegetativo é por alongamento das hifas e catabolismo de carbono é obrigatoriamente aeróbico. Em contraste, crescimento vegetativo por
 25 leveduras como *Saccharomyces cerevisiae* é por formação de broto de um talo unicelular e catabolismo de carbono pode ser fermentativo.

Em um aspecto ainda mais preferido, a célula hospedeira de fungos filamentosos é uma célula de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*,

Coriolus, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*,
Mucor, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*,
Penicillium, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*,
Talaromyces, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, ou
5 *Trichoderma*.

Em um aspecto o mais preferido, a célula hospedeira de fungos
filamentosos é uma célula de *Aspergillus awamori*, *Aspergillus fumigatus*,
Aspergillus foetidus, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus*
niger ou *Aspergillus oryzae*. Em outro aspecto o mais preferido, a célula
10 hospedeira de fungos filamentosos é uma célula de *Fusarium bactridioides*,
Fusarium cerealis, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium*
graminearum, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium*
negundi, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*,
Fusarium sambucinum, *Fusarium sarcochroum*, *Fusarium sporotrichioides*,
15 *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, ou
Fusarium venenatum. Em outro aspecto o mais preferido, a célula hospedeira
de fungos filamentosos é uma célula de *Bjerkandera adusta*, *Ceriporiopsis*
aneirina, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis*
gilvescens, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis*
20 *subrufa*, *Ceriporiopsis subvermispora*, *Chrysosporium keratinophilum*,
Chrysosporium lucknowense, *Chrysosporium tropicum*, *Chrysosporium*
merdarium, *Chrysosporium inops*, *Chrysosporium pannicola*, *Chrysosporium*
queenslandicum, *Chrysosporium zonatum*, *Coprinus cinereus*, *Coriolus*
hirsutus, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*,
25 *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium brasilianum*,
Penicillium purpurogenum, *Phanerochaete chrysosporium*, *Phlebia radiata*,
Pleurotus eryngii, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*,
Trichoderma harzianum, *Trichoderma koningii*, *Trichoderma*
longibrachiatum, *Trichoderma reesei*, ou *Trichoderma viride*.

As células de fungo podem ser transformadas por um processo envolvendo formação de protoplastos, transformação dos protoplastos, e regeneração da parede celular em um modo conhecido por si. Os procedimentos apropriados para a transformação de células hospedeiras de *Aspergillus* e *Trichoderma* são descritos em EP 238 023 e Yelton *et al.*, 1984, *Proceedings of the National Academy of Sciences USA* 81: 1470 - 1474. Os métodos apropriados para transformar a espécie *Fusarium* são descritos por Malardier *et al.*, 1989, *Gene* 78: 147 - 158, e WO 96/00787. A levedura pode ser transformada usando os procedimentos descritos por Becker e Guarente, In Abelson, J. N. e Simon, M. I., editores, *Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology*, volume 194, pp 182 - 187, Academic Press, Inc., New York; Ito *et al.*, 1983, *Journal of Bacteriology* 153: 163; e Hinnen *et al.*, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1920.

15 **Métodos de Produção**

A presente invenção também refere-se a métodos para produzir um polipeptídeo da presente invenção, compreendendo: (a) cultivar uma célula, que, em sua forma de tipo selvagem, é capaz de produzir o polipeptídeo, sob condições que conduzem para produção do polipeptídeo; e (b) recuperar o polipeptídeo. Em um aspecto preferido, a célula é do gênero *Myceliophthora*. Em um aspecto mais preferido, a célula é *Myceliophthora thermophila*. Em um aspecto o mais preferido, a célula é *Myceliophthora thermophila* CBS 117.65. Em outro aspecto preferido, a célula é basidiomicete CBS 494.95. Em outro aspecto preferido, a célula é basidiomicete CBS 495.95. Em outro aspecto preferido, a célula é do gênero *Penicillium*. Em outro aspecto mais preferido, a célula é *Penicillium brasilianum*. Em outro aspecto o mais preferido, a célula é *Penicillium brasilianum* IBT 20888.

A presente invenção também refere-se a métodos para

produzir um polipeptídeo da presente invenção, compreendendo: (a) cultivar uma célula hospedeira sob condições que conduzem para produção do polipeptídeo; e (b) recuperar o polipeptídeo.

5 A presente invenção também refere-se a métodos para produzir um polipeptídeo da presente invenção, compreendendo: (a) cultivar uma célula hospedeira sob condições que conduzem para produção do polipeptídeo, em que a célula hospedeira compreende uma sequência mutante de nucleotídeos tendo pelo menos uma mutação em uma sequência de codificação de polipeptídeo maduro de SEQ ID NO: 3, SEQ ID NO: 5, SEQ
10 ID NO: 7, ou SEQ ID NO: 9, em que a sequência mutante de nucleotídeos codifica um polipeptídeo que compreende ou consiste do polipeptídeo maduro de SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10, e (b) recuperar o polipeptídeo.

Em um aspecto preferido, o polipeptídeo maduro de SEQ ID
15 NO: 4 contém os aminoácidos 17 a 389. Em outro aspecto preferido, o polipeptídeo maduro de SEQ ID NO: 6 contém os aminoácidos 16 a 397. Em outro aspecto preferido, o polipeptídeo maduro de SEQ ID NO: 8 contém os aminoácidos 22 a 429. Em outro aspecto preferido, o polipeptídeo maduro de SEQ ID NO: 10 contém os aminoácidos 25 a 421. Em outro aspecto
20 preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 87 a 1185 de SEQ ID NO: 3. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 84 a 1229 de SEQ ID NO: 5. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 77 a 1300 de
25 SEQ ID NO: 7. Em outro aspecto preferido, a sequência de codificação de polipeptídeo maduro contém os nucleotídeos 73 a 1488 de SEQ ID NO: 9.

Nos métodos de produção da presente invenção, as células são cultivadas em um meio nutriente apropriado para produção do polipeptídeo usando métodos bem conhecidos na arte. Por exemplo, a célula pode ser

cultivada por cultivo em frasco de agitação, e fermentação em grande escala ou pequena escala (incluindo fermentações em estado contínuo, batelada, batelada alimentada, ou sólido) em laboratório ou fermentadores industriais realizada em um meio apropriado e sob condições permitindo ao polipeptídeo ser expresso e/ou isolado. O cultivo ocorre em um meio nutriente apropriado compreendendo carbono e fontes de nitrogênio e sais inorgânicos, usando procedimentos conhecidos na arte. Os meios apropriados são disponíveis de fornecedores comerciais ou podem ser preparados de acordo com as composições publicadas (por exemplo, em catálogos do American Type Culture Collection). Se o polipeptídeo for secretado no meio nutriente, o polipeptídeo pode ser recuperado diretamente do meio. Se o polipeptídeo não for secretado no meio, ele pode ser recuperado dos lisados de células.

Os polipeptídeos podem ser detectados usando métodos conhecidos na técnica que são específicos para os polipeptídeos. Estes métodos de detecção podem incluir uso de anticorpos específicos, formação de um produto de enzima, ou desaparecimento de um substrato de enzima. Por exemplo, um teste de enzima pode ser usado para determinar a atividade do polipeptídeo como descrito aqui.

O polipeptídeo resultante pode ser recuperado usando métodos conhecidos na arte. Por exemplo, o polipeptídeo pode ser recuperado do meio nutriente por procedimentos convencionais incluindo, mas não limitados a, centrifugação, filtração, extração, secagem por pulverização, evaporação, ou precipitação.

Os polipeptídeos da presente invenção podem ser purificados por uma variedade de procedimentos conhecidos na técnica incluindo, mas não limitados a, cromatografia (por exemplo, troca iônica, afinidade, hidrofóbico, cromatofocalização, e exclusão por tamanho), procedimentos eletroforéticos (por exemplo, focalização isoeletrica preparativa), solubilidade diferencial (por exemplo, precipitação por sulfato de amônio), SDS-PAGE, ou

extração (ver, por exemplo, *Protein Purification*, J. -C, Janson e Lars Ryden, editores, VCH Publishers, New York, 1989) para obter polipeptídeos substancialmente puros.

Plantas

5 A presente invenção também refere-se a plantas, por exemplo, uma planta transgênica, parte de planta, ou célula de planta, compreendendo um polinucleotídeo isolado codificando um polipeptídeo tendo atividade de endoglucanase da presente invenção de modo a expressar e produzir o polipeptídeo em quantidades recuperáveis. O polipeptídeo pode ser
10 recuperado da planta ou parte de planta. Alternativamente, a planta ou parte de planta contendo o polipeptídeo recombinante pode ser usada como tal para melhorar a qualidade de um alimento ou ração, por exemplo, melhorando o valor nutricional, palatabilidade, e propriedades reológicas, ou para destruir um fator antinutritivo.

15 A planta transgênica pode ser dicotiledônea (um dicótile) ou monocotiledônea (um monocótile). Exemplos de plantas monocotiledôneas são gramíneas, como gramas do prato (grama azul, *Poa*), forragens como *Festuca*, *Lolium*, grama temperada, como *Agrostis*, e cereais, por exemplo, trigo, aveias, centeio, cevada, arroz, sorgo, e milho gráudo (milho).

20 Exemplos de plantas dicotiledôneas são tabaco, legumes, como tremoços, batata, beterraba, ervilha, feijão e soja, e plantas crucíferas (família *Brassicaceae*), como couve-flor, semente de colza, e o organismo de modelo intimamente relacionado *Arabidopsis thaliana*.

25 Exemplos de parte de plantas incluem talo, calo, folhas, raiz, frutos, sementes, e tubérculos assim como os tecidos individuais compreendendo estas partes, por exemplo, epiderme, mesófilos, parênquima, tecidos vasculares, meristemas. Compartimentos de célula de planta específicos, como cloroplastos, apoplastos, mitocôndrias, vacúolos, peroxissomas e citoplasma também são considerados como sendo uma parte

de planta. Além disso, qualquer célula de planta, qualquer que seja a origem do tecido, é considerada para sendo uma parte de planta. Do mesmo modo, partes de planta como tecidos específicos e células isoladas para facilitar a utilização da invenção também são consideradas partes de plantas, por exemplo, embriões, endospermas, aleurona e capas de sementes.

Também estão incluídas dentro do escopo da presente invenção as progênes de tais plantas, partes de plantas, e células de plantas.

A planta transgênica ou célula de planta expressando um polipeptídeo da presente invenção pode ser construída de acordo com os métodos conhecidos na arte. Em resumo, a planta ou célula de planta é construída por incorporação de uma ou mais construções de expressão codificando um polipeptídeo da presente invenção no genoma hospedeiro de planta ou genoma de cloroplasto e propagando a planta modificada resultante ou célula de planta em uma planta transgênica ou célula de planta.

A construção de expressão é convenientemente uma construção de ácido nucleico que compreende um polinucleotídeo codificando um polipeptídeo da presente invenção ligado operativamente com seqüências reguladoras apropriadas requeridas para a expressão da seqüência de nucleotídeos na planta ou parte de planta de escolha. Além disso, a construção de expressão pode compreender um marcador selecionável utilizável para identificar as células hospedeiras em que a construção de expressão foi integrada e seqüências de DNA necessárias para a introdução da construção na planta em questão (a última depende do método de introdução de DNA a ser usado).

A escolha de seqüências reguladoras, como promotor e seqüências de terminador e opcionalmente seqüências de sinal ou de transito é determinada, por exemplo, com base em quando, onde, e como o se deseja que o polipeptídeo seja expresso. Por exemplo, a expressão do gene codificando um polipeptídeo da presente invenção pode ser constitutiva ou

indutível, ou pode ser específica do desenvolvimento, estágio ou tecido, e o produto de gene pode ser marcado em um tecido específico ou parte de planta como sementes ou folhas. As seqüências reguladoras são, por exemplo, descritas por Tague *et al.*, 1988, *Plant Physiology* 86: 506.

- 5 Para expressão constitutiva, 35S-CaMV, a ubiquitina de milho 1, e o promotor 1 actina de arroz podem ser usados (Franck *et al.*, 1980, *Cell* 21: 285 - 294, Christensen *et al.*, 1992, *Plant Mo. Biol.* 18: 675 - 689; Zhang *et al.*, 1991, *Plant Cell* 3: 1155 - 1165), promotores de órgão específicos podem ser, por exemplo, um promotor de tecidos imersos em armazenamento
- 10 como sementes, tubérculos de batata, e frutas (Edwards & Coruzzi, 1990, *Ann, Rev. Genet.* 24: 275 - 303), ou de tecidos de imersão metabólicos como meristemas (Ito *et al.*, 1994, *Plant Mol. Biol.* 24: 863 -878), um promotor específico de semente como o promotor glutelina, prolamina, globulina, ou albumina de arroz (Wu *et al.*, 1998, *Plant and Cell Physiology* 39: 885 - 889),
- 15 um promotor de *Vicia faba* da legumina B4 e o gene de proteína de semente desconhecida de *Vicia faba* (Conrad *et al.*, 1998, *Journal of Plant Physiology* 152: 708 - 711), um promotor de uma proteína do corpo oleoso da semente (Chen *et al.*, 1998, *Plant and Cell Physiology* 39: 935 - 941), o promotor napA de proteína de armazenamento de *Brassica napus*, ou qualquer outro
- 20 promotor específico de semente conhecido na arte, por exemplo, como descrito em WO 91/14772. Além disso, o promotor pode ser um promotor específico de folha como o promotor *rbcs* de arroz ou tomate (Kyoizuka *et al.*, 1993, *Plant Physiology* 102: 991 - 1000, o promotor de gene metiltransferase adenina de vírus clorela (Mitra e Higgins, 1994, *Plant Molecular Biology* 26:
- 25 85 - 93), ou o promotor de gene *aldP* de arroz (Kagaya *et al.*, 1995, *Molecular and General Genetics* 248: 668 - 874), ou um promotor indutível por ferida como o promotor de batata pin2 (Xu *et al.*, 1993, *Plant Molecular Biology* 22: 573 - 588). Do mesmo modo, o promotor pode ser indutível por tratamentos abióticos como temperatura, seca, ou alterações em salinidade ou

induzido por substâncias aplicadas de modo exógeno que ativam o promotor, por exemplo, etanol, estrógenos, hormônios de planta como etileno, ácido abscísico, e ácido giberélico, e metais pesados.

Um elemento melhorador do promotor também pode ser usado para obter uma maior expressão de um polipeptídeo da presente invenção na planta. Por exemplo, o elemento melhorador do promotor pode ser um íntron que é colocado entre o promotor e a seqüência de nucleotídeos codificando um polipeptídeo da presente invenção. Por exemplo, Xu *et al.*, 1993, *supra*, descrevem o uso do primeiro íntron do gene actina 1 de arroz para melhorar a expressão.

O gene marcador selecionável e quaisquer outras partes da construção de expressão podem ser escolhidos a partir dos disponíveis na arte.

A construção de ácido nucleico é incorporada no genoma de planta de acordo com as técnicas convencionais conhecidas na arte, incluindo transformação mediada por *Agrobacterium*, transformação mediada por vírus, microinjeção, bombardeio de partículas, transformação biolística, e eletroporação (Gasser *et al.*, 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto *et al.*, 1989, *Nature* 338: 274).

Atualmente, a transferência de gene mediada por *Agrobacterium tumefaciens* é o método de escolha para gerar dicotiledôneas transgênicas (para um estudo, ver Hooykas e Schilperoort, 1992, *Plant Molecular Biology* 19: 15 - 38) e também pode ser usado para transformar monocotiledôneas, apesar de outros métodos de transformação serem usados com frequência para estas plantas. Atualmente, o método de escolha para gerar monocotiledôneas transgênicas é bombardeio de partículas (partículas microscópicas de ouro ou de tungstênio revestidas com o DNA transformante) de calos embriônicos ou embriões em desenvolvimento (Christou, 1992, *Plant Journal* 2: 275 - 281; Shimamoto, 1994, *Current Opinion Biotechnology* 5: 158 - 162; Vasil *et al.*, 1992, *Bio/Technology* 10: 667 - 674). Um método

alternativo para transformação de monocotiledôneas é com base na transformação de protoplasto como descrito por Omirulleh *et al.*, 1993, *Plant Molecular Biology* 21: 415 - 428.

5 Após a transformação, os transformantes tendo incorporado a construção de expressão são selecionados e regenerados em todas as plantas de acordo com os métodos bem conhecidos na arte. Com frequência, o procedimento de transformação é projetado para a eliminação seletiva de genes de seleção ou durante a regeneração ou nas gerações seguintes por uso, por exemplo, de co-transformação com duas construções de T-DNA
10 separadas ou excisão específica de sítio do gene de seleção por uma recombinase específica.

A presente invenção também refere-se a métodos para produzir um polipeptídeo da presente invenção compreendendo: (a) cultivar uma planta transgênica ou uma célula de planta compreendendo um
15 polinucleotídeo codificando um polipeptídeo tendo atividade de endoglucanase da presente invenção sob condições que conduzem para produção do polipeptídeo; e (b) recuperar o polipeptídeo.

Remoção ou Redução de Atividade de Endoglucanase

A presente invenção também refere-se a métodos para
20 produzir um mutante de uma célula parental, que compreende desregular ou deletar uma sequência de polinucleotídeo, ou uma porção do mesmo, codificando um polipeptídeo da presente invenção, que resulta na célula mutante produzir menos do polipeptídeo do que a célula parental quando cultivada sob as mesmas condições.

25 A célula mutante pode ser construída por redução ou eliminação da expressão de uma sequência de nucleotídeos codificando um polipeptídeo da presente invenção usando métodos bem conhecidos na arte, por exemplo, inserções, desregulações, substituições, ou deleções. Em um aspecto preferido, a sequência de nucleotídeos é inativa. A sequência de

nucleotídeos a ser modificada ou inativada pode ser, por exemplo, a região de codificação ou uma parte da mesma essencial para atividade, ou um elemento regulador requerido para a expressão da região de codificação. Um exemplo desta seqüência de controle ou reguladora pode ser uma seqüência de promotor ou uma parte funcional da mesma, isto é, uma parte isto é suficiente para afetar a expressão da seqüência de nucleotídeo. Outras seqüências de controle para possível modificação incluem, mas não são limitadas a, um líder, seqüência de poliadenilação, seqüência de pró-peptídeo, seqüência de peptídeo sinal, terminador de transcrição, e ativador transcripcional.

Modificação ou inativação da seqüência de nucleotídeos pode ser realizada ao se submeter a célula parental a mutagênese e selecionar para células mutantes em que a expressão da seqüência de nucleotídeos foi reduzida ou eliminada. A mutagênese, que pode ser específica ou aleatória, pode ser realizada, por exemplo, pelo uso de um agente mutagenizante químico ou físico, pelo uso de um oligonucleotídeo apropriado, ou submetendo a seqüência de DNA a mutagênese gerada por PCR. Além disso, a mutagênese pode ser realizada pelo uso de qualquer combinação destes agentes mutagenizante.

Exemplos de um agente mutagenizante químico ou físico apropriado para o presente fim incluem irradiação ultravioleta (UV), hidroxilamina, N-metil-N'-nitro-N-nitrosoguanidina (MNNG), O-metil hidroxilamina, ácido nitroso, sulfonato de etil metano (EMS), bissulfito de sódio, ácido fórmico, e análogos de nucleotídeos.

Quando tais agentes são usados, a mutagênese é tipicamente realizada por incubação da célula parental a ser mutagenizada na presença do agente mutagenizante de escolha sob condições apropriadas, e triando e/ou selecionando as células mutantes demonstrando reduzida ou nenhuma expressão do gene.

Modificação ou inativação da seqüência de nucleotídeos pode

ser efetuada por introdução, substituição, ou remoção de um ou mais nucleotídeos no gene ou um elemento regulador requerido para a transcrição ou tradução do mesmo. Por exemplo, nucleotídeos podem ser inseridos ou removidos de modo a resultar na introdução de um códon de parada, a remoção do códon de partida, ou uma mudança na matriz de leitura aberta. Tal modificação ou inativação pode ser efetuada por mutagênese sítio-dirigida ou mutagênese PCR gerada de acordo com os métodos conhecidos na arte. Apesar de, em princípio, a modificação poder ser realizada *in vivo*, isto é, diretamente na célula expressando a seqüência de nucleotídeos a ser modificada, prefere-se que modificação seja realizada *in vitro* como exemplificado abaixo.

Um exemplo de um modo conveniente para eliminar ou reduzir a expressão de uma seqüência de nucleotídeos por uma célula é com base nas técnicas de substituição de gene, deleção de gene, ou desregulação de gene. Por exemplo, no método de desregulação de gene, uma seqüência de ácido nucleico correspondente na seqüência de nucleotídeos endógeno é mutagenizada *in vitro* para produzir uma seqüência de ácido nucleico defeituosa que é então transformada na célula parental para produzir um gene defeituoso. Por recombinação homóloga, a seqüência de ácido nucleico defeituosa substitui a seqüência de nucleotídeo endógena. Pode ser desejável que a seqüência de nucleotídeos defeituosa também codifica um marcador que pode ser usado para seleção de transformantes em que a seqüência de nucleotídeos foi modificada ou destruída. Em um aspecto particularmente preferido, a seqüência de nucleotídeos é desregulada com um marcador selecionável como o descrito aqui.

Alternativamente, a modificação ou inativação da seqüência de nucleotídeo pode ser realizada por técnicas de RNAi ou anti-sentido estabelecidas ou usando uma seqüência complementar na seqüência de nucleotídeo. Mais especificamente, a expressão da seqüência de nucleotídeos

por uma célula pode ser reduzida ou eliminada por introdução de uma sequência complementar para sequência de nucleotídeos do gene que pode ser transcrito na célula e é capaz de hibridizar no mRNA produzido na célula. Sob condições permitindo à sequência de nucleotídeos anti-sentido complementar hibridizar em mRNA, a quantidade de proteína traduzida é assim reduzida ou eliminada.

A presente invenção ainda refere-se a uma célula mutante de uma célula parental que compreende uma desregulação ou deleção de uma sequência de nucleotídeos codificando o polipeptídeo ou uma sequência de controle do mesmo, que resulta na célula mutante produzir menos do polipeptídeo ou nenhum polipeptídeo comparado na célula parental.

As células mutantes deficientes em polipeptídeos assim criadas são particularmente utilizáveis como células hospedeiras para a expressão de polipeptídeos homólogos e/ou heterólogos. Assim, a presente invenção ainda refere-se a métodos para produzir um polipeptídeo homólogo ou heterólogo compreendendo: (a) cultivar célula mutante sob condições que conduzem para produção do polipeptídeo; e (b) recuperar o polipeptídeo. O termo "polipeptídeo heterólogo" é definido aqui como polipeptídeos que não são nativos na célula hospedeira, uma proteína nativa em que as modificações foram feitas para alterar a sequência nativa, ou uma proteína nativa cuja expressão é quantitativamente alterada como um resultado de uma manipulação da célula hospedeira por técnicas de DNA recombinantes.

Em um outro aspecto, a presente invenção refere-se a um método para produzir um produto de proteína essencialmente livre de atividade de endoglucanase por fermentação de uma célula que produz tanto um polipeptídeo da presente invenção assim como o produto de proteína de

interesse por adição de uma quantidade eficaz de um agente capaz de inibir a atividade de endoglucanase no caldo de fermentação antes, durante, ou depois da fermentação estar completada, recuperar o produto de

interesse do caldo de fermentação, e opcionalmente submeter o produto recuperado a outra purificação.

Em um outro aspecto, a presente invenção refere-se a um método para produzir um produto de proteína essencialmente livre de
5 atividade de endoglucanase por cultivo da célula sob condições permitindo a expressão do produto, submetendo o caldo de cultura resultante a um tratamento de pH e de temperatura combinado de modo a reduzir substancialmente a atividade de endoglucanase, e recuperar o produto do
10 caldo de cultura. Alternativamente, o tratamento combinado de pH e de temperatura pode ser realizado na preparação de enzima recuperada do caldo de cultura. O tratamento combinado de pH e de temperatura opcionalmente pode ser usado em combinação com um tratamento com um inibidor de endoglucanase.

De acordo com este aspecto da invenção, é possível remover
15 pelo menos 60%, preferivelmente pelo menos 75%, mais preferivelmente pelo menos 85%, ainda mais preferivelmente pelo menos 95%, e o mais preferivelmente pelo menos 99% da atividade de endoglucanase. Remoção completa de atividade de endoglucanase pode ser obtida pelo uso deste método.

20 O tratamento combinado de pH e de temperatura é preferivelmente realizado em um pH na faixa de 2-3 ou 10-11 e uma temperatura na faixa de pelo menos 75-85°C durante um período suficiente de tempo para atingir o feito desejado, onde tipicamente, 1 a 3 horas são suficientes.

25 Os métodos usados para cultivo e purificação do produto de interesse podem ser realizados por métodos conhecidos na arte.

Os métodos da presente invenção para produzir um produto essencialmente livre de endoglucanase são de particular interesse na produção de polipeptídeos eucarióticos, em particular, proteínas de fungo como as

enzimas. A enzima pode ser selecionada dentre, por exemplo, uma enzima amilolítica, enzima lipolítica, enzima proteolítica, enzima celulítica, oxidoreductase, ou enzima degradando a parede da célula de planta. Exemplos de tais enzimas incluem aminopeptidase, amilase, amiloglucosidase, 5 carbohidrase, carboxipeptidase, catalase, celulase, quitinase, cutinase, ciclodextrina glicosiltransferase, deoxiribonuclease, esterase, galactosidase, beta-galactosidase, glucoamilase, glicose oxidase, glucosidase, haloperoxidase, hemicelulase, invertase, isomerase, lacase, ligase, lipase, liase, manosidase, oxidase, enzima pectinolítica, peroxidase, fitase, 10 fenoxidase, polifenoxidase, enzima proteolítica, ribonuclease, transferase, transglutaminase, ou xilanase. As células deficientes em endoglucanase também podem ser usadas para expressar proteínas heterólogas de interesse farmacêutico como hormônios, fatores de crescimento, receptores, e outros.

Deve-se entender que o termo "polipeptídeos eucarióticos" 15 inclui não somente polipeptídeos nativos, mas também os polipeptídeos, por exemplo, enzimas, que foram modificadas por substituições de aminoácido, deleções ou adições, ou outras tais modificações para aumentar a atividade, termoestabilidade, tolerância a pH e outros.

Em um outro aspecto, a presente invenção refere-se a um 20 produto de proteína essencialmente livre de atividade de endoglucanase que é produzido por um método da presente invenção.

Métodos de Inibir a Expressão de um Polipeptídeo

A presente invenção também refere-se aos métodos de inibir a expressão de um polipeptídeo em uma célula, compreendendo administrar à 25 célula ou expressar na célula uma molécula de RNA (dsRNA) de filamento duplo, em que o dsRNA compreende uma subsequência ou porção de um polinucleotídeo da presente invenção. Em um aspecto preferido, o dsRNA é de cerca de 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 ou mais nucleotídeos duplex em comprimento. Em outro aspecto preferido, o polipeptídeo tem

atividade de endoglucanase.

O dsRNA é preferivelmente um RNA interferência pequeno (siRNA) ou um RNA micro (miRNA). Em um aspecto preferido, o dsRNA é RNA interferência pequeno (siRNAs) para inibir a transcrição. Em outro
5 aspecto preferido, o dsRNA é RNA micro (miRNAs) para inibir a tradução.

A presente invenção também refere-se a tais moléculas de RNA filamento duplo (dsRNA) para inibir a expressão de um polipeptídeo em uma célula, em que o dsRNA compreende uma subsequência ou porção de um polinucleotídeo codificando o polipeptídeo maduro de SEQ ID NO: 4,
10 SEQ ID NO: 6, SEQ ID NO: 8, ou SEQ ID NO: 10. Apesar da presente invenção não ser limitada por qualquer mecanismo particular de ação, o dsRNA pode entrar em uma célula e causar a degradação de um RNA filamento único (ssRNA) de seqüências similares ou idênticas, incluindo mRNAs endógenos. Quando uma célula é exposta a dsRNA, mRNA do gene
15 homólogo é seletivamente degradado por um processo chamado RNA interferência (RNAi).

Os dsRNAs da presente invenção podem ser usados em terapêuticas de silenciamento de genes. Em um aspecto, a invenção provê métodos para seletivamente degradar o RNA usando os dsRNAs da presente
20 invenção. O processo pode ser praticado *in vitro*, *ex vivo* ou *in vivo*. Em um aspecto, as moléculas de dsRNA podem ser usadas para gerar uma mutação de perda de função em uma célula, um órgão ou um animal. Métodos para fabricar e usar moléculas de dsRNA para degradar seletivamente RNA são bem conhecidos na arte, ver, por exemplo, patente U.S. 6.506.559; patente
25 U.S. 6.511.824; patente U.S. 6.515.109; e patente U.S. 6.489.127.

Composições

A presente invenção também refere-se às composições compreendendo um polipeptídeo da presente invenção. Preferivelmente, as composições são enriquecidas em tal polipeptídeo. O termo "enriquecido"

indica que a atividade de endoglucanase da composição foi aumentada, por exemplo, com um fator de enriquecimento de pelo menos 1.1.

A composição pode compreender um polipeptídeo da presente invenção como o componente enzimático principal, por exemplo, uma
 5 composição mono-componente. Alternativamente, a composição pode compreender atividades enzimáticas múltiplas, como um aminopeptidase, amilase, carbohidrase, carboxipeptidase, catalase, celulase, quitinase, cutinase, ciclodextrina glicosiltransferase, deoxiribonuclease, esterase, alfa-galactosidase, beta-galactosidase, glucoamilase, alfa-glucosidase, beta-
 10 glucosidase, haloperoxidase, invertase, lacase, lipase, manosidase, oxidase, enzima pectinolítica, peptidoglutaminase, peroxidase, fitase, polifenoloxidase, enzima proteolítica, ribonuclease, transglutaminase, ou xilanoase. A(s) enzima (s) adicional(ais) pode(m) ser produzida(s), por exemplo, por um microorganismo pertencendo no gênero *Aspergillus*, preferivelmente
 15 *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, ou *Aspergillus oryzae*; *Fusarium*, preferivelmente *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium*
 20 *heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sulphureum*, *Fusarium toruloseum*, *Fusarium trichothecioides*, ou *Fusarium venenatum*; *Humicola*, preferivelmente *Humicola insolens* ou *Humicola lanuginosa*; ou *Trichoderma*,
 25 preferivelmente *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, ou *Trichoderma viride*.

As composições de polipeptídeo podem ser preparadas de acordo com os métodos conhecidos na técnica e podem estar na forma de um líquido ou uma composição em pó. Por exemplo, a composição de

polipeptídeo pode estar na forma de um granulado ou um microgranulado. O polipeptídeo a ser incluído na composição pode ser estabilizado de acordo com os métodos conhecidos na arte.

Exemplos são dados abaixo de usos preferidos das composições de polipeptídeo da invenção. A dosagem da composição de polipeptídeo da invenção e outras condições sob as quais a composição é usada podem ser determinadas com base em métodos conhecidos na arte.

Usos

A presente invenção também refere-se a métodos para a degradação ou conversão de um material celulósico, compreendendo: tratar o material celulósico com uma composição compreendendo uma quantidade eficaz de um polipeptídeo tendo atividade de endoglucanase da presente invenção. Em um aspecto preferido, o método ainda compreende a recuperação do material celulósico degradado ou convertido.

Os polipeptídeos e as células hospedeiras da presente invenção podem ser usados na produção de monossacarídeos, dissacarídeos, e polissacarídeos como cargas de alimentação de fermentação ou químicas a partir de biomassa celulósica para a produção de etanol, plásticos, outros produtos ou intermediários. A composição compreendendo o polipeptídeo tendo atividade de endoglucanase pode estar na forma de um caldo de fermentação bruto com ou sem as células removidas ou na forma de uma preparação de enzima purificada ou semi-purificada. A composição também pode compreender outras proteínas e enzimas utilizáveis no processamento de biomassa, por exemplo, celobiohidrolase, beta-glucosidase, enzimas hemicelulolíticas, melhoradores (WO 2005/074647, WO 2005/074656), etc. Alternativamente, a composição pode compreender uma célula hospedeira da presente invenção como uma fonte do polipeptídeo tendo atividade de endoglucanase em um processo de fermentação com a biomassa. Em particular, os polipeptídeos e as células hospedeiras da presente invenção

podem ser usados para aumentar o valor de resíduos de processamento (grãos secos de destiladores, grãos gastos de cervejaria, bagaço de cana de açúcar, etc.) por degradação completa ou parcial de celulose ou hemicelulose. A célula hospedeira também pode conter genes heterólogos ou nativos que

5 codificam outras proteínas e enzimas, mencionadas acima, utilizáveis no processamento de biomassa.

Nos métodos da presente invenção, qualquer material celulósico, como biomassa, pode ser usado. Deve-se entender aqui que o termo "material celulósico" engloba lignocelulose. Biomassa pode incluir,

10 mas não é limitada a recursos de madeira, refugo sólido municipal, papel usado, colheitas, e resíduos de colheitas (ver, por exemplo, Wiselogle *et al.*, 1995, in Handbook on Bioethanol (Charles E. Wyman, editor), pp. 105 - 118, Taylor & Francis, Washington D. C.; Wyman, 1994, *Bioresource Technology* 50: 3 - 16; Lynd, 1990, *Applied Biochemistry and Biotechnology* 24/25: 695 -

15 719; Mosier *et al.*, 1999, Recent Progress in Bioconversion of Lignocellulosics, in *Advances in Biochemical Engineering/Biotechnology*, T. Scheper, editor responsável, volume 65, pp. 23 - 40, Springer-Verlag, New York).

O polissacarídeo predominante na parede celular primária de

20 biomassa é celulose, o segundo o mais abundante é hemicelulose, e o terceiro é pectina. A parede celular secundária, produzida depois da célula ter parado de crescer, também contém polissacarídeos e é reforçada por lignina polimérica covalentemente reticulada em hemicelulose. Celulose é um homopolímero de anidrocetobiose e assim um beta-(1-4)-D-glucano linear,

25 enquanto hemiceluloses incluem uma variedade de compostos, como xilanos, xiloglucanos, arabinoxilanos, e mananos em estruturas ramificadas complexas com um espectro de substituintes. Apesar de geralmente polimorfa, celulose é encontrada em tecido de planta primariamente como uma matriz cristalina insolúvel de cadeias glucano paralelas. Hemiceluloses geralmente ligam por

ponte hidrogênio a celulose, assim como a outras hemiceluloses, que ajudam a estabilizar a matriz da parede celular.

Três classes maiores de enzimas são usadas para a ruptura da biomassa celulósica:

5 (1) As "endo-1,4-beta-glucanases" ou 1,4-beta-D-glucano-4-glucano-hidrolases (EC 3.2.1.4), que atuam aleatoriamente em substratos 1,4-beta- glucano insolúvel e solúvel.

 (2) As "exo-1,4-beta-D-glucanases" incluindo tanto as 1,4-beta-D-glucano glucohidrolases (EC 3.2.1.74), que liberam D-glicose de 1,4-
10 beta-D-glucanos e lentamente hidrolisam D-celobiose, como celobiohidrolases (1,4-beta-D-glucano celobiohidrolases, EC 3.2.1.91), que liberam D-celobiose de 1,4-beta-glucanos.

 (3) As "beta-D-glucosidases" ou beta-D-glucosídeo glucohidrolases (EC 3.2.1.21), que atuam para a liberação de unidades de D-
15 glicose de celobiose e celodextrinas solúveis, assim como um arranjo de glicosídeos.

Os polipeptídeos tendo atividade de endoglucanase da presente invenção são preferivelmente usados em conjunto com outras proteínas celulolíticas, por exemplo, exo-1,4-beta-D-glucanases e beta-D-glucosidases,
20 para degradar o componente de celulose do substrato de biomassa, (ver, por exemplo, Brigham *et al.*, 1995, in *Handbook on Bioethanol* (Charles E. Wyman, editor), PP. 119 - 141, Taylor & Francis, Washington D. C; Lee, 1997, *Journal of Biotechnology* 56: 1 - 24).

As exo-1,4-beta-D-glucanases e beta-D-glucosidases podem
25 ser produzidas por qualquer método conhecido na técnica (ver, por exemplo, Bennett, J. W. e LaSure, L. (eds.), *More Gene Manipulations in Fungi*, Academic Press, CA, 1991).

As quantidades ótimas de um polipeptídeo tendo atividade de endoglucanase e outras proteínas celulolíticas dependem de vários fatores

incluindo, mas não limitados a mistura de proteínas de componentes celulolíticos, o substrato celulósico, a concentração de substrato celulósico, o pré-tratamento (s) do substrato celulósico, temperatura, tempo, pH, e inclusão de organismo de fermentação (por exemplo, levedura para sacarificação e fermentação simultâneas). O termo "proteína celulolíticas" é definido aqui como as proteínas ou misturas de proteínas mostradas como sendo capazes de hidrolisar ou converter ou degradar celulose sob as condições testadas.

Em um aspecto preferido, a quantidade de polipeptídeo tendo atividade de endoglucanase por g de material celulósico é de cerca de 0,5 a cerca de 50 mg, preferivelmente de cerca de 0,5 a cerca de 40 mg, mais preferivelmente de cerca de 0,5 a cerca de 25 mg, mais preferivelmente de cerca de 0,75 a cerca de 20 mg, mais preferivelmente de cerca de 0,75 a cerca de 15 mg, ainda mais preferivelmente de cerca de 0,5 a cerca de 10 mg, e o mais preferivelmente de cerca de 2,5 a cerca de 10 mg por g de material celulósico.

Em outro aspecto preferido, a quantidade de proteínas celulolíticas por g de material celulósico é de cerca de 0,5 a cerca de 50 mg, preferivelmente de cerca de 0,5 a cerca de 40 mg, mais preferivelmente de cerca de 0,5 a cerca de 25 mg, mais preferivelmente de cerca de 0,75 a cerca de 20 mg, mais preferivelmente de cerca de 0,75 a cerca de 15 mg, ainda mais preferivelmente de cerca de 0,5 a cerca de 10 mg, e o mais preferivelmente de cerca de 2,5 a cerca de 10 mg por g de material celulósico.

Nos métodos da presente invenção, a composição pode ser suplementada por uma ou mais atividades de enzima adicionais para melhorar a degradação do material celulósico. As enzimas adicionais preferidas são hemicelulases, esterases (por exemplo, lipases, fosfolipases, e/ou cutinases), proteases, lacases, peroxidases, ou misturas dos mesmos.

Nos métodos da presente invenção, a(s) enzima (s) adicional(ais) podem ser adicionada(s) antes de ou durante a fermentação,

incluindo durante ou após a propagação do(s) microorganismo (s) de fermentação.

As enzimas podem ser derivadas ou obtidas de qualquer origem apropriada, incluindo, origem de bactérias, fungos, leveduras ou mamíferos. O termo "obtido" significa aqui que a enzima pode ter sido isolada de um organismo que naturalmente produz a enzima como uma enzima nativa. O termo "obtido" também significa aqui que a enzima pode ter sido produzida recombinantemente em um organismo hospedeiro, em que a enzima recombinantemente produzida é ou nativa ou estranha para o organismo hospedeiro ou tem uma seqüência de aminoácido modificada, por exemplo, tendo um ou mais aminoácidos que são deletados, inseridos e/ou substituídos, isto é, uma enzima recombinantemente produzida que é um mutante e/ou um fragmento de uma seqüência de aminoácido nativa ou uma enzima produzida por processos de embaralhamento de ácido nucleico conhecidos na arte. Estão englobadas dentro do significado de uma enzima nativa as variantes naturais e dentro do significado de uma enzima estranha as variantes recombinantemente obtidas, como por mutagênese sítio-dirigido ou embaralhamento.

As enzimas também podem ser purificadas. O termo "purificado" como usado aqui cobre as enzimas livres de outros componentes do organismo das quais ele é derivado. O termo "purificado" também cobre as enzimas livres de componentes do organismo nativo a partir das quais ele é obtido. As enzimas podem ser purificadas, com somente quantidades menores de outras proteínas estando presentes. A expressão "outras proteínas" refere-se em particular a outras enzimas. O termo "purificado" como usado aqui também refere-se à remoção de outros componentes, particularmente outras proteínas e o mais particularmente outras enzimas presentes na célula de origem da enzima da invenção. A enzima pode ser "substancialmente pura," isto é, livre de outros componentes do organismo em que é produzida, isto é,

por exemplo, um organismo hospedeiro para enzimas recombinantemente produzidas. Em um aspecto preferido, as enzimas são pelo menos 75% (p/p), preferivelmente pelo menos 80%, mais preferivelmente pelo menos 85%, mais preferivelmente pelo menos 90%, mais preferivelmente pelo menos 95%, mais preferivelmente pelo menos 96%, mais preferivelmente pelo menos 97%, ainda mais preferivelmente pelo menos 98%, ou o mais preferivelmente pelo menos 99% puras. Em outro aspecto preferido, a enzima é 100% pura.

As enzimas usadas na presente invenção podem estar em qualquer forma apropriada para uso nos processos descritos aqui, como, por exemplo, um caldo de fermentação bruto com ou sem células, um pó seco ou granulado, um granulado não pulverulento, um líquido, um líquido estabilizado, ou uma enzima protegida. Granulados podem ser produzidos, por exemplo, como descrito na patente U.S. 4.106.991 e 4.661.452, e opcionalmente podem ser revestidos por processos conhecidos na arte. Preparações de enzima líquida podem, por exemplo, ser estabilizadas por adição de estabilizadores como um açúcar, um álcool de açúcar ou outro poliol, e/ou ácido láctico ou outro ácido orgânico de acordo com o processo estabelecido. As enzimas protegidas podem ser preparadas de acordo com o processo descrito em EP 238.216.

Os métodos da presente invenção podem ser usados para processar um material celulósico para muitos produtos orgânicos, químicos e combustíveis utilizáveis. Além disso, para etanol, algumas mercadorias e especialmente produtos químicos que podem ser produzidos a partir de celulose incluem xilose, acetona, acetato, glicina, lisina, ácidos orgânicos (por exemplo, ácido láctico), 1,3-propanodiol, butanodiol, glicerol, etileno glicol, furfural, poli-hidroxicanoatos, cis, ácido cis-mucônico, e ração animal (Lynd, L. R., Wyman, C. E., e Gerngross, T. U., 1999, *Biocommodity Engineering, Biotechnol, Prog.*, 15: 777 - 793; Philippidis, G. P., 1996,

Cellulose bioconversion technology, in *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179 - 212, e Ryu, D. D. Y., e Mandels, M., 1980, Cellulases: biosynthesis and applications, *Enz. Microb. Technol.*, 2: 91 - 102). Benefícios de co-produção
 5 potenciais se estendem além da síntese de produtos orgânicos múltiplos de carboidrato fermentável. Resíduos ricos em lignina restantes após processamento biológico podem ser convertidos em produtos químicos derivados de lignina, ou usados para a produção de energia.

Métodos convencionais usados para processar o material
 10 celulósico de acordo com os métodos da presente invenção são bem entendidos pelo versado na técnica. Os métodos da presente invenção podem ser implementados usando qualquer aparelho de processamento de biomassa convencional configurado para operar de acordo com a invenção.

Este aparelho pode incluir um reator agitado de batelada, um
 15 reator agitado de fluxo contínuo com ultrafiltração, um reator de coluna de fluxo tampão contínuo (Gusakov, A. V., e Sinitsyn, A. P., 1985, Kinetics of the enzymatic hydrolysis of cellulose: 1. A mathematical model for a batch reactor process, *Enz. Microb. Technol.* 7: 346 - 352), um reator por atrito (Ryu, S. K., e Lee, J. M., 1983, Biconversion of waste cellulose by using an
 20 attrition bioreactor, *Biotechnol. Bioeng.* 25: 53 - 65), ou um reator com agitação intensiva induzida por um campo eletromagnético (Gusakov, A. V., Sinitsyn, A. P., Davydkin, I. Y., Davydkin, V. Y., Protas, O. V., 1996, Enhancement of enzymatic cellulose hydrolysis using a novel type of bioreactor with intensive stirring induced by electromagnetic field, *Appl. Biochem. Biotechnol.* 56: 141 - 153).
 25

Os métodos convencionais incluem, mas não são limitados a, sacarificação, fermentação, hidrólise e fermentação separadas (SHF), sacarificação e fermentação simultâneas (SSF), sacarificação e co-fermentação simultâneas (SSCF), hidrólise e fermentação híbridas (HHF), e

conversão microbiana direta (DMC).

SHF usa etapas de processo separadas para primeiro hidrolisar enzimaticamente a celulose em glicose e então fermentar a glicose em etanol. Em SSF, a hidrólise enzimática de celulose e a fermentação de glicose em etanol é combinada em uma etapa (Philippidis, G. P., 1996, Cellulose bioconversion technology, in *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179 - 212). SSCF inclui a co-fermentação de açúcares múltiplos (Sheehan, J., e Himmel, M., 1999, Enzymes, energy and the environment: A strategic perspective on the U.S. Department of Energy's research and development activities for bioethanol, *Biotechnol. Prog.* 15: 817 - 827). HHF inclui duas etapas separadas realizadas no mesmo reator mas em temperaturas diferentes, isto é, sacarificação enzimática de temperatura elevada seguido por SSF em uma temperatura inferior que a cepa de fermentação pode tolerar. DMC combina todos os três processos (produção de celulase, hidrólise de celulose, e fermentação) em uma etapa (Lynd, L. R., Weimer, P. J., van Zyl, W. H., e Pretorius, I. S., 2002, Microbial cellulose utilization: Fundamentals and biotechnology, *Microbiol. Mol. Biol. Reviews* 66: 506 - 577).

"Fermentação" ou "processo de fermentação" refere-se a qualquer processo de fermentação ou qualquer processo compreendendo uma etapa de fermentação. Um processo de fermentação inclui, sem limitação, processos de fermentação usados para produzir produtos de fermentação incluindo álcoois (por exemplo, arabinitol, butanol, etanol, glicerol, metanol, 1,3-propanodiol, sorbitol, e xilitol); ácidos orgânicos (por exemplo, ácido acético, ácido acetônico, ácido adípico, ácido ascórbico, ácido cítrico, ácido 2,5-diceto-D-glucônico, ácido fórmico, ácido fumárico, ácido glucárico, ácido glucônico, ácido glucurônico, ácido glutárico, ácido 3-hidroxipropiônico, ácido itacônico, ácido láctico, ácido málico, ácido malônico, ácido oxálico, ácido propiônico, ácido succínico, e ácido xilônico); cetonas (por exemplo,

acetona); aminoácidos (por exemplo, ácido aspártico, ácido glutâmico, glicina, lisina, serina, e treonina); gases (por exemplo, metano, hidrogênio (H_2), dióxido de carbono (CO_2), e monóxido de carbono (CO)). Processos de fermentação também incluem processos de fermentação usados na indústria de álcool consumível (por exemplo, cerveja e vinho), indústria de laticínios (por exemplo, laticínios fermentados), indústria de couro, e indústria de fumo.

A presente invenção ainda refere-se a métodos de produzir uma substância, compreendendo: (a) sacarificar um material celulósico com uma composição compreendendo uma quantidade eficaz de um polipeptídeo tendo atividade de endoglucanase; (b) fermentar o material celulósico sacarificado de etapa (a) com um ou mais microorganismos de fermentação; e (c) recuperar a substância da fermentação. A composição compreendendo o polipeptídeo tendo atividade de endoglucanase pode estar na forma de um caldo de fermentação bruto com ou sem as células removidas ou na forma de uma preparação de enzima semi-purificada ou purificada ou a composição pode compreender uma célula hospedeira da presente invenção como uma fonte do polipeptídeo tendo atividade de endoglucanase em um processo de fermentação com a biomassa.

A substância pode ser qualquer substância derivada da fermentação. Em uma forma de realização preferida, a substância é um álcool. Deve-se entender que o termo "álcool" engloba uma substância que contém uma ou mais porções de hidroxila. Em uma forma de realização mais preferida, o álcool é arabinitol. Em outra forma de realização mais preferida, o álcool é butanol. Em outra forma de realização mais preferida, o álcool é etanol. Em outra forma de realização mais preferida, o álcool é glicerol. Em outra forma de realização mais preferida, o álcool é metanol. Em outra forma de realização mais preferida, o álcool é 1,3-propanodiol. Em outra forma de realização mais preferida, o álcool é sorbitol. Em outra forma de realização mais preferida, o álcool é xilitol. Ver, por exemplo, Gong, C. S., Cao, N, J.,

- Du, J., e Tsao, G. T., 1999, Ethanol production from renewable resources, in *Advances in Biochemical Engineering/Biotechnology*, Scheper, T., ed., Springer-Verlag Berlin Heidelberg, Alemanha, 65: 207 - 241; Silveira, M. M., e Jonas, R., 2002, The biotechnological production of sorbitol, *Appl. Microbiol. Biotechnol.* 59: 400 - 408; Nigam, P., e Singh, D., 1995, Processes for fermentative production of xylitol - a sugar substitute, *Process Biochemistry* 30 (2): 117 - 124; Ezeji, T. C., Qureshi, N. e Blaschek, H. P., 2003, Production of acetone, butanol and ethanol by *Clostridium beijerinckii* BA101 and *in situ* recovery by gas stripping, *World Journal of Microbiology and Biotechnology* 19 (6): 595 - 603.

Em outra forma de realização preferida, a substância é um ácido orgânico. Em outra forma de realização mais preferida, o ácido orgânico é ácido acético. Em outra forma de realização mais preferida, o ácido orgânico é ácido acetônico. Em outra forma de realização mais preferida, o ácido orgânico é ácido adípico. Em outra forma de realização mais preferida, o ácido orgânico é ácido ascórbico. Em outra forma de realização mais preferida, o ácido orgânico é ácido cítrico. Em outra forma de realização mais preferida, o ácido orgânico é ácido 2,5-diceto-D-glucônico. Em outra forma de realização mais preferida, o ácido orgânico é ácido fórmico. Em outra forma de realização mais preferida, o ácido orgânico é ácido fumárico. Em outra forma de realização mais preferida, o ácido orgânico é ácido glucárico. Em outra forma de realização mais preferida, o ácido orgânico é ácido glucônico. Em outra forma de realização mais preferida, o ácido orgânico é ácido glucurônico. Em outra forma de realização mais preferida, o ácido orgânico é ácido glutárico. Em outra forma de realização preferida, o ácido orgânico é ácido 3-hidroxipropiônico. Em outra forma de realização mais preferida, o ácido orgânico é ácido itacônico. Em outra forma de realização mais preferida, o ácido orgânico é ácido láctico. Em outra forma de realização mais preferida, o ácido orgânico é ácido málico. Em

outra forma de realização mais preferida, o ácido orgânico é ácido malônico. Em outra forma de realização mais preferida, o ácido orgânico é ácido oxálico. Em outra forma de realização mais preferida, o ácido orgânico é ácido propiônico. Em outra forma de realização mais preferida, o ácido orgânico é ácido succínico. Em outra forma de realização mais preferida, o ácido orgânico é ácido xilônico. Ver, por exemplo, Chen, R., e Lee, Y. Y., 1997, Membrane-mediated extractive fermentation for lactic acid production from cellulosic biomass, *Appl. Biochem, Biotechnol.* 63-65: 435 - 448.

Em outra forma de realização preferida, a substância é uma cetona. Deve-se entender que o termo "cetona" engloba uma substância que contém uma ou mais porções de cetona. Em outra forma de realização mais preferida, a cetona é acetona. Ver, por exemplo, Qureshi e Blaschek, 2003, *supra*.

Em outra forma de realização preferida, a substância é um aminoácido. Em outra forma de realização mais preferida, o ácido orgânico é ácido aspártico. Em outra forma de realização mais preferida, o aminoácido é ácido glutâmico. Em outra forma de realização mais preferida, o aminoácido é glicina. Em outra forma de realização mais preferida, o aminoácido é lisina. Em outra forma de realização mais preferida, o aminoácido é serina. Em outra forma de realização mais preferida, o aminoácido é treonina. Ver, por exemplo, Richard, A., e Margaritis, A., 2004, Empirical modeling of batch fermentation kinetics for poly (glutamic acid) production and other microbial biopolymers, *Biotechnology and Bioengineering* 87 (4): 501 - 515.

Em outra forma de realização preferida, a substância é um gás. Em outra forma de realização mais preferida, o gás é metano. Em outra forma de realização mais preferida, o gás é H₂. Em outra forma de realização mais preferida, o gás é CO₂. Em outra forma de realização mais preferida, o gás é CO. Ver, por exemplo, Kataoka, N., A. Miya, e K. Kiriya, 1997, Studies on hydrogen production by continuous culture system of hydrogen-producing

anaerobic bacteria, *Water Science and Technology* 38 (6-7): 41 - 47; e Gunaseelan V. N. in *Biomass and Bioenergy*, vol. 13 (1-2), pp. 83 - 114, 1997, Anaerobic digestion of biomass for methane production: A review.

Produção de uma substância de material celulósico tipicamente requer quatro etapas principais. Estas quatro etapas são pré-tratamento, hidrólise enzimática, fermentação, e recuperação. Exemplificado abaixo é um processo para produzir etanol, mas deve-se entender que processos similares podem ser usados para produzir outras substâncias, por exemplo, as substâncias descritas acima.

Pré-tratamento. No pré-tratamento ou etapa de pré-hidrólise, o material celulósico é aquecido para romper a lignina e estrutura de carboidrato, solubilizar a maior parte da hemicelulose, e tornar a fração de celulose acessível para enzimas celulolíticas. O aquecimento é realizado ou diretamente com vapor ou em suspensão onde um catalisador também pode ser adicionado no material para acelerar as reações. Os catalisadores incluem ácidos fortes, como ácido sulfúrico e SO₂, ou alcalino, como hidróxido de sódio. O propósito do estágio de pré-tratamento é para facilitar a penetração das enzimas e microorganismos. A biomassa celulósica também pode ser submetida a um pré-tratamento de explosão a vapor hidrotérmico (ver pedido de patente 20020164730).

Sacarificação. Na etapa de hidrólise enzimática, também conhecida como sacarificação, as enzimas como descrito aqui são adicionadas no material pré-tratado para converter a fração de celulose para glicose e/ou outros açúcares. A sacarificação geralmente é realizada em reatores de tanques agitados ou fermentadores sob condições controladas de pH, temperatura, e mistura. Uma etapa de sacarificação pode durar 200 horas. A sacarificação pode ser realizada em temperaturas de cerca de 30°C a cerca de 65°C, em particular, em torno de 50°C, e a um pH na faixa entre cerca de 4 a cerca de 5, especialmente em torno de pH 4,5. Para produzir glicose que pode

ser metabolizada por levedura, a hidrólise é tipicamente realizada na presença de uma beta-glucosidase.

Fermentação. Na etapa de fermentação, açúcares, liberados do material celulósico como um resultado das etapas de hidrólise enzimática e pré-tratamento, são fermentados em etanol por um organismo de fermentação, como levedura. A fermentação também pode ser realizada simultaneamente com a hidrólise enzimática no mesmo vaso, novamente sob condições controladas de pH, temperatura, e mistura. Quando a sacarificação e fermentação são realizadas simultaneamente no mesmo vaso, o processo geralmente é chamado sacarificação e fermentação simultâneas ou SSF.

Qualquer substrato apropriado celulósico ou matéria prima pode ser usado em um processo de fermentação da presente invenção. O substrato geralmente é selecionado com base no produto de fermentação desejado, isto é, a substância a ser obtida da fermentação, e o processo empregado, como é bem conhecido na arte. Exemplos de substratos apropriados para usar nos métodos da presente invenção incluem materiais contendo celulose, como resíduos de madeira ou planta ou açúcares de baixo peso molecular DP1-3 obtidos a partir de material celulósico processado que pode ser metabolizado pelo microorganismo de fermentação, e que pode ser fornecido por adição direta ao meio de fermentação.

O termo "meio de fermentação" será entendido como fazendo referência a um meio antes do(s) microorganismo (s) de fermentação ser adicionado(s), como, um meio resultante de um processo de sacarificação, assim como um meio usado em um processo de sacarificação e fermentação simultâneas (SSF).

"Microorganismo de fermentação" refere-se a qualquer microorganismo apropriado para uso em um processo de fermentação desejado. Os microorganismos de fermentação apropriados de acordo com a invenção são capazes de fermentar, isto é, converter, açúcares, como glicose,

xilose, arabinose, manose, galactose, ou oligossacarídeos diretamente ou indiretamente no produto de fermentação desejado. Exemplos de microorganismos de fermentação incluem organismos de fungos, como leveduras. As leveduras preferidas incluem cepas de *Saccharomyces spp.*, e em particular, *Saccharomyces cerevisiae*. A levedura comercialmente disponível inclui, por exemplo, Red Star^{®/TM}/Lesaffre Ethanol Red (disponível de Red Star/Lesaffre, USA) FALI (disponível de Fleischmann's Yeast, a division of Burns Philp Food Inc., USA), SUPERSTART (disponível de Alltech), GERT STRAND (disponível de Gert Strand AB, Suécia) e FERMIOL (disponível de DSM Specialties).

Em uma forma de realização preferida, a levedura é *Saccharomyces spp.* Em uma forma de realização mais preferida, a levedura é *Saccharomyces cerevisiae*. Em outra forma de realização mais preferida, a levedura é *Saccharomyces distaticus*. Em outra forma de realização mais preferida, a levedura é *Saccharomyces uvarum*. Em outra forma de realização preferida, a levedura é *Kluyveromyces*. Em outra forma de realização mais preferida, a levedura é *Kluyveromyces marxianus*. Em outra forma de realização mais preferida, a levedura é *Kluyveromyces fragilis*. Em outra forma de realização preferida, a levedura é *Candida*. Em outra forma de realização mais preferida, a levedura é *Candida pseudotropicalis*. Em outra forma de realização mais preferida, a levedura é *Candida brassicae*. Em outra forma de realização preferida, a levedura é uma *Clavispora*. Em outra forma de realização mais preferida, a levedura é *Clavispora lusitaniae*. Em outra forma de realização mais preferida, a levedura é *Clavispora opuntiae*. Em outra forma de realização preferida, a levedura é *Pachysolen*. Em outra forma de realização mais preferida, a levedura é *Pachysolen tannophilus*. Em outra forma de realização preferida, a levedura é *Bretannomyces*. Em outra forma de realização mais preferida, a levedura é *Bretannomyces clausenii* (Philippidis, G. P., 1996, Cellulose bioconversion technology, in *Handbook*

on *Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179 - 212).

Bactérias que podem eficientemente fermentar glicose em etanol incluem, por exemplo, *Zymomonas mobilis* e *Clostridium thermocellum* (Philippidis, 1996, *supra*).

É bem conhecido na técnica que os organismos descritos acima também podem ser usados para produzir outras substâncias, como descrito aqui.

A clonagem de genes heterólogos em *Saccharomyces cerevisiae* (Chen, Z., Ho, N. W. Y., 1993, Cloning and improving the expression of *Pichia stipitis* xylose reductase gene in *Saccharomyces cerevisiae*, *Appl. Biochem. Biotechnol.* 39 - 40: 135 - 147; Ho, N. W. Y., Chen, Z, Brainard, A. P., 1998, Genetically engineered *Saccharomyces* yeast capable of effectively co-fermenting glicose and xylose, *Appl. Environ. Microbiol.* 64: 1852 - 1859), ou em bactérias como *Escherichia coli* (Beall, D. S., Qhta, K., Ingram, L. O., 1991, Parametric studies of ethanol production from xylose and other sugars by recombinant *Escherichia coli*, *Biotech. Bioeng.* 38: 296 - 303), *Klebsiella oxytoca* (Ingram, L. O., Gomes, P. F., Lai, X., Moniruzzaman, M., Wood, B. E., Yomano, L. P., York, S. W., 1998, Metabolic engineering of bacteria for ethanol production, *Biotechnol. Bioeng.* 58: 204 - 214), e *Zymomonas mobilis* (Zhang, M., Eddy, C., Deanda, K., Finkelstein, M., e Picataggio, S., 1995, Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobitis*, *Science* 267: 240 - 243; Deanda, K., Zhang, M., Eddy, C., e Picataggio, S., 1996, Development of an arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering, *Appl. Environ. Microbiol.*, 62: 4465 - 4470) tem levado à construção de organismos capazes de converter hexoses e pentoses em etanol (co-fermentação).

Levedura ou outro microorganismo tipicamente é adicionado à

celulose ou hidrolisado degradado e a fermentação prossegue durante cerca de 24 a cerca de 96 horas, como cerca de 35 e a cerca de 60 horas. A temperatura é tipicamente entre cerca de 26°C a cerca de 40°C, em particular de cerca de 32°C, a cerca de pH 3 a cerca de pH 6, em particular em torno de pH 4-5.

- 5 Em uma forma de realização preferida, levedura ou outro microorganismo é aplicado na celulose degradada ou hidrolisada e a fermentação prossegue durante cerca de 24 a cerca de 96 horas, como tipicamente 35-60 horas. Em uma forma de realização preferida, a temperatura geralmente é entre cerca de 26 a cerca de 40°C, em particular de
- 10 cerca de 32°C, e o pH geralmente é de cerca de pH 3 a cerca de pH 6, preferivelmente em torno de pH 4-5. Levedura ou outro microorganismo é preferivelmente aplicado em quantidades de aproximadamente 10^5 a 10^{12} , preferivelmente de aproximadamente 10^7 a 10^{10} , especialmente aproximadamente 5×10^7 contagem viável por ml de caldo de fermentação.
- 15 Durante uma fase de produção de etanol, a contagem de célula de levedura deve preferivelmente estar na faixa de aproximadamente 10^7 a 10^{10} , especialmente em torno de aproximadamente 2×10^8 . Outra diretriz com relação ao uso de levedura para fermentação pode ser encontrada em, por exemplo, "The Alcohol Textbook" (Editores K. Jacques, T. P. Lyons e D. R.
- 20 Kelsall, Nottingham University Press, United Kingdom 1999), que é incorporado aqui por referência.

- O processo na técnica mais amplamente usado é o processo de sacarificação e fermentação simultâneas (SSF) onde não se tem um estágio de manutenção para a sacarificação, significando que a levedura e a enzima são
- 25 adicionadas juntas.

 Para a produção de etanol, após a fermentação, a massa é destilada para extrair o etanol. O etanol obtido de acordo com o processo da invenção pode ser usado como, por exemplo, etanol de combustível; etanol para bebidas, isto é, álcool neutro potável, ou etanol industrial.

Um estimulador de fermentação pode ser usado em combinação com qualquer um dos processos enzimáticos descritos aqui para ainda melhorar o processo de fermentação, e em particular, o desempenho do microorganismo de fermentação, como, melhora da taxa e rendimento de etanol. Um "estimulador de fermentação" refere-se a estimuladores para crescimento dos microorganismos de fermentação, em particular, levedura. Estimuladores de fermentação preferidos para crescimento incluem vitaminas e minerais. Exemplos de vitaminas incluem multivitaminas, biotina, pantotenato, ácido nicotínico, meso-inositol, tiamina, piridoxina, ácido para-aminobenzóico, ácido fólico, riboflavina, e vitaminas A, B, C, D, e E. Ver, por exemplo, Alfenore *et al.*, Improving ethanol production and viability of *Saccharomyces cerevisiae* by a vitamin feeding strategy during fed-batch process, Springer-Verlag (2002), que é incorporado aqui por referência. Exemplos de minerais incluem minerais e sais minerais que podem suprir nutrientes compreendendo P, K, Mg, S, Ca, Fe, Zn, Mn, e Cu.

Recuperação. O álcool é separado do material celulósico fermentado e purificado por métodos convencionais de destilação. Etanol com uma pureza de até cerca de 96% em volume de etanol pode ser obtido, que pode ser usado como, por exemplo, etanol combustível, etanol para bebidas, isto é, álcool neutro potável, ou etanol industrial.

Para outras substâncias, qualquer método conhecido na técnica pode ser usado incluindo, mas não limitado a, cromatografia (por exemplo, troca iônica, afinidade, hidrofóbico, cromatofocalização, e exclusão por tamanho), procedimentos eletroforéticos (por exemplo, focalização isoeletrica preparativa), solubilidade diferencial (por exemplo, precipitação com sulfato de amônio), SDS-PAGE, destilação, ou extração.

Nos métodos da presente invenção, o polipeptídeo tendo atividade de endoglucanase e outra proteína (s) celulolítica (s) pode ser suplementado por uma ou mais atividades de enzima adicionais para melhorar

a degradação do material celulósico. As enzimas adicionais preferidas são hemicelulases, esterases (por exemplo, lipases, fosfolipases, e/ou cutinases), proteases, lacases, peroxidases, ou misturas dos mesmos.

Nos métodos da presente invenção, a (s) enzima (s) adicional
5 (ais) pode (m) ser adicionada (s) antes ou durante a fermentação, incluindo durante ou após a propagação do (s) microorganismo (s) de fermentação.

Pró-peptídeos e Peptídeos de Sinal

A presente invenção também refere-se às construções de ácido
nucleico compreendendo um gene codificando uma proteína, em que o gene é
10 ligado operativamente a uma ou ambas dentre uma primeira seqüência de nucleotídeos codificando um peptídeo de sinal compreendendo ou consistindo de aminoácidos 1 a 16 de SEQ ID NO: 4, aminoácidos 1 a 15 de SEQ ID NO: 6, aminoácidos 1 a 21 de SEQ ID NO: 8, ou aminoácidos 1 a 16 de SEQ ID NO: 10 e uma segunda seqüência de nucleotídeos codificando um pró-peptídeo compreendendo ou consistindo de aminoácidos 17 a 24 de SEQ ID
15 NO: 10, em que um gene é estranho para uma primeira e segunda seqüência de nucleotídeos.

Em um aspecto preferido, a primeira seqüência de nucleotídeos compreende ou consiste de nucleotídeos 19 a 69 de SEQ ID NO:
20 3. Em outro aspecto preferido, a primeira seqüência de nucleotídeos compreende ou consiste de nucleotídeos 39 a 83 de SEQ ID NO: 5. Em outro aspecto preferido, a primeira seqüência de nucleotídeos compreende ou consiste de nucleotídeos 14 a 76 de SEQ ID NO: 7. Em outro aspecto preferido, a primeira seqüência de nucleotídeos compreende ou consiste de
25 nucleotídeos 1 a 48 de SEQ ID NO: 9. Em outro aspecto preferido, a segunda seqüência de nucleotídeos compreende ou consiste de nucleotídeos 49 a 72 de SEQ ID NO: 9.

A presente invenção também refere-se a vetores de expressão recombinantes e células hospedeiras recombinantes compreendendo tais

construções de ácido nucleico.

A presente invenção também refere-se a métodos para produzir uma proteína compreendendo (a) cultivar esta célula hospedeira recombinante sob condições apropriadas para produção da proteína; e (b)
5 recuperar a proteína.

A proteína pode ser nativa ou heteróloga para uma célula hospedeira. O termo "proteína" não se destina aqui a fazer referência a um comprimento específico do produto codificado e, assim, engloba peptídeos, oligopeptídeos, e proteínas. O termo "proteína" também engloba dois ou mais
10 polipeptídeos combinados para formar o produto codificado. As proteínas também incluem polipeptídeos híbridos que compreendem uma combinação de seqüências de polipeptídeo completa ou parcial obtidas dentre pelo menos duas proteínas diferentes em que uma ou mais pode ser heteróloga ou nativa na célula hospedeira. Proteínas ainda incluem variações alélicas de ocorrência
15 natural e engenheiradas das proteínas e proteínas híbridas mencionadas acima.

Preferivelmente, a proteína é um hormônio ou variante do mesmo, enzima, receptor ou porção do mesmo, anticorpo ou porção do mesmo, ou repórter. Em um aspecto mais preferido, a proteína é um oxidoreductase, transferase, hidrolase, liase, isomerase, ou ligase. Em um
20 aspecto ainda mais preferido, a proteína é um aminopeptidase, amilase, carbohidrase, carboxipeptidase, catalase, celulase, quitinase, cutinase, ciclodextrina glicosiltransferase, deoxiribonuclease, esterase, alfa-galactosidase, beta-galactosidase, glucoamilase, alfa-glucosidase, beta-glucosidase, invertase, lacase, outro lipase, manosidase, mutanase, oxidase,
25 enzima pectinolítica, peroxidase, fitase, polifenoloxidase, enzima proteolítica, ribonuclease, transglutaminase ou xilanoase.

O gene pode ser obtido a partir de qualquer procariótico, eucariótico, ou outra fonte.

A presente invenção é ainda descrita pelos seguintes exemplos

que não devem ser considerados como limitando o escopo da invenção.

Exemplos

Materiais

Os produtos químicos usados como tampões e substratos
5 foram produtos comerciais de pelo menos um tipo reagente.

Cepas

Myceliophthora thermophila CBS 117.65, basidiomicete CBS
494.95, basidiomicete CBS 495.95, e *Penicillium brasilianum* cepa IBT
20888 (IBT Culture Collection of Fungi, Technical University of Dinamarca,
10 Copenhagen, Dinamarca) foram usadas como fontes dos genes de
endoglucanase. *Saccharomyces cerevisiae* Cepa W3124 (MATa; ura 3-52; leu
2-3, 112; his 3-D200; pep 4-1137; prc1::HIS3; prb1:: LEU2; cir⁺) foi usada
para a triagem de bibliotecas de expressão de *Myceliophthora thermophila*
CBS 117.65 para a atividade de endoglucanase. *Aspergillus oryzae* HowB104
15 (alfa-amilase-negativa) foi usada para expressão dos genes *cel5a*.

Meios de Cultura e Soluções

Meio LB foi composto, por litro, de 10 g de triptona, 5 g de
extrato de levedura, e 5 g de cloreto de sódio.

Meio de ampicilina LB foi composto, por litro, de 10 g de
20 triptona, 5 g de extrato de levedura, 5 g de cloreto de sódio, e 50 µg de
ampicilina por ml (filtro esterilizado, adicionado após autoclavagem).

Placas de ampicilina LB foram compostas, por litro, de meio
de ampicilina LB e 15 g de ágar bacto.

Meio YPD foi composto de extrato de levedura a 1%, peptona
25 a 2%, e glicose a 2% de filtro esterilizado adicionado após autoclavagem.

Meio YPM foi composto de 1% extrato de levedura, peptona a
2%, e maltodextrina a 2% de filtro esterilizado adicionado após
autoclavagem.

Meio SC-URA com galactose foi composto, por litro, de 100

ml de 10X sais Basal, 28 ml de ácidos casamino a 20% sem vitaminas, 10 ml de 1% triptofano, 3,6 ml de treonina a 5% (filtro esterilizado, adicionado após autoclavagem), e 100 ml de galactose a 20% (filtro esterilizado, adicionado após autoclavagem).

5 Meio SC-URA com glicose foi composto, por litro, de 100 ml de 10X solução de sais Basal, 28 ml de ácidos casamino a 20% sem vitaminas, 10 ml de triptofano a 1%, 3,6 ml de treonina a 5% (filtro esterilizado, adicionado após autoclavagem), e 100 ml de glicose a 20% (filtro esterilizado, adicionado após autoclavagem).

10 Solução 10X de sais Basal foi composta, por litro, de 75 g de base de nitrogênio de levedura, 113 g de ácido succínico, e 68 g de NaOH.

SC-ágar foi composto, por litro, de meio SC-URA (com glicose ou galactose como indicado) e 20 g de ágar.

15 Placas de agar SC de celulose AZCL HE a 0,1% com galactose foram compostas, por litro, de meio SC-URA com galactose, 20 g de ágar, e celulose AZCL HE a 0,1% (Megazyme, Wicklow, Irlanda).

20 Meio BA foi composto, por litro, de 10 g de material seco de licor de milho, 10 g de NH_4NO_3 , 10 g de KH_2PO_4 , 0,75 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,1 ml de plurônico, e 0,5 g de CaCO_3 . O pH foi ajustado para 6,5 antes da autoclavagem.

Placas COVE foram compostas, por litro, de 342,3 g de sacarose, 25 g de ágar Noble, 20 ml de solução de sais COVE, 10 mM de acetamida, e 20 mM de CsCl. A solução foi ajustada a pH 7,0 antes da autoclavagem.

25 Solução de sais COVE foi composta, por litro, de 26 g de KCl, 26 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 76 g de KH_2PO_4 , e 50 ml de metais de traço COVE.

Solução de metais de traço COVE foi composta, por litro, de 0,04 g de $\text{NaB}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$, 0,4 g de $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 1,2 g de $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0,7 g de $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0,8 g de $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, e 10 g de $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$.

TE foi composto de 10 mM de Tris pH 7,4 e 0,1 mM de EDTA.

Exemplo 1: Construção de bibliotecas de expressão de cDNA de *Myceliophthora thermophila* CBS 117.65 em *Saccharomyces cerevisiae*

5 *Myceliophthora thermophila* CBS 117.65 foi cultivado em 200 ml de meio BA a 30° C durante cinco dias a 200 rpm. Micélios da cultura do frasco de agitação foram coletadas por filtração dos conteúdos através de um funil forrado com Miracloth™ (CalBiochem, San Diego, CA, USA). Micélios foram então ensaduichadas entre dois pedaços de Miracloth™ e secadas com
10 toalhas de papel absorvente. A massa miceliana foi então transferida para tubos plásticos de centrifuga Falcon 1059 e congelada em nitrogênio líquido. Micélios congeladas foram armazenadas em um congelador a -80°C até usar.

A extração de RNA total foi realizada com tiocianato de guanidina seguido por ultracentrifugação através de uma almofada 5,7 M de
15 CsCl, e isolamento de poli(A)+RNA foi realizado por cromatografia de afinidade de oligo(dT)-celulose, usando os procedimentos descritos em WO 94/14953.

cDNA filamento duplo foi sintetizado a partir de 5 µg de poli(A)⁺ RNA pelo método RNase H (Gubler e Hoffman, 1983, *Gene* 25: 263
20 - 289, Sambrook *et al.*, 1989, *Molecular cloning: A laboratory manual*, Cold Spring Harbor lab., Cold Spring Harbor, NY, USA). O poli(A)⁺ RNA (5 µg em 5 µl de água tratada com DEPC (0,1% de dietilpirocarbonato)) foi aquecido a 70°C durante 8 minutos em um pré-siliconizado, tubo Eppendorf isento de RNase-, resfriado bruscamente em gelo, e combinado em um
25 volume final de 50 µl com tampão transcriptase reverso composto de 50 mM de Tris-HCl, pH 8,3, 75 mM de KCl, 3 mM de MgCl₂, 10 mM de ditioneitol (DTT) (Bethesda Research Laboratories, Bethesda, MD, USA), 1 mM de dATP, dGTP e dTTP, e 0,5 mM de 5-metil-dCTP (Pharmacia, Uppsala, Suécia), 40 unidades de inibidor de ribonuclease placentar humano (RNasin,

Promega, Madison, WI, USA), 1,45 µg de iniciador oligo(dT)₁₈-Not I (Pharmacia), e 1000 unidades de transcriptase reversa de SuperScript II RNase H (Bethesda Research Laboratories). cDNA de primeiro filamento foi sintetizado por incubação da mistura de reação a 45 °C durante 1 hora. Após
5 síntese, a mistura híbrida de mRNA:cDNA foi filtrada em gel através de uma coluna giratória MicroSpin S-400 HR (Pharmacia) de acordo com as instruções dos fabricantes.

Após a filtração em gel, os híbridos foram diluídos em 250 µl de tampão do segundo de filamento (20 mM de Tris-HCl, pH 7,4, 90 mM de
10 KCl, 4,6 mM de MgCl₂, 10 mM (NH₄)₂SO₄, 0,16 mM de NAD) contendo 200 µM de cada dNTP, 60 unidades de *E. coli* DNA polimerase I (Pharmacia), 5,25 unidades de RNase H (Promega), e 15 unidades de *E. coli* DNA ligase (Boehringer Mannheim, Mannheim, Alemanha). A síntese de cDNA de segundo filamento foi realizada por incubação do tubo de reação a 16 °C
15 durante 2 horas e um adicional de 15 minutos a 25 °C. A reação foi parada por adição de EDTA em uma concentração final de 20 mM seguido por extrações com fenol e clorofórmio.

O cDNA filamento duplo foi precipitado a -20°C durante 12 horas por adição de 2 volumes de 96% de etanol e 0,2 volume de 10 M de
20 acetato de amônio, recuperado por centrifugação a 13.000 x g, lavado em 70% de etanol, secado, e recolocado em suspensão em 30 µl de tampão nuclease de feijão Mung (30 mM de acetato de sódio pH 4,6, 300 mM de NaCl, 1 mM de ZnSO₄, 0,35 mM de DTT, 2% de glicerol) contendo 25 unidades de nuclease de feijão Mung (Pharmacia). O DNA de filamento único
25 "grampo de cabelo" foi preso por incubação da reação a 30°C durante 30 minutos, seguido por adição de 70 µl de 10 mM de Tris-HCl-1 mM EDTA pH 7,5, extração por fenol, e precipitação com 2 volumes de 96% de etanol e 0,1 volume de 3 M de acetato de sódio pH 5,2 em gelo durante 30 minutos.

Os cDNAs filamento duplo foram recuperados por

centrifugação a 13.000 x g e terminados de modo obtuso em 30 µl de tampão T4 DNA polimerase (20 mM de Tris-acetato, pH 7,9, 10 mM de acetato de magnésio, 50 mM de acetato de potássio, 1 mM de DTT) contendo 0,5 mM de cada dNTP e 5 unidades de T4 DNA polimerase (New England Biolabs, Ipswich, MA, USA) por incubação da mistura de reação a 16 °C durante 1 hora. A reação foi parada por adição de EDTA em uma concentração final de 20 mM, seguido por extrações com fenol e clorofórmio, e precipitação durante 12 horas a -20°C por adição de 2 volumes de etanol a 96% e 0,1 volume de 3 M de acetato de sódio pH 5,2.

Após a reação de suprimento, os cDNAs foram recuperados por centrifugação a 13.000 x g, lavados em 70% de etanol, e secados. O grânulo de cDNA foi recolocado em suspensão em 25 µl de tampão de ligação (30 mM de Tris-HCl, pH 7,8, 10 mM de MgCl₂, 10 mM de DTT, 0,5 mM de ATP) contendo 2,5 µg de adaptadores não palindrômicos *Bst* XI (Invitrogen, Carlsbad, CA, USA), mostrados abaixo, e 30 unidades de T4 ligase (Promega), e então incubados a 16 °C durante 12 horas. A reação foi parada por aquecimento a 65 °C durante 20 minutos e então resfriada em gelo durante 5 minutos.

5'-CTTTCAGCACA-3' (SEQ ID NO: 1)

3'-GAAAGGTC-5' (SEQ ID NO. 2)

O cDNA adaptado foi digerido com *Not* I, seguido por incubação durante 2,5 horas a 37 °C. A reação foi parada por aquecimento a 65 °C durante 10 minutos. Os cDNAs foram fracionados em tamanho por eletroforese gel em um agarose gel a 0,8% SeaPlaque GTG de temperatura de fusão baixa (Cambrex Corporation, East Rutherford, NJ, USA) em 44 mM de Tris Base, 44 mM de ácido bórico, 0,5 mM de tampão EDTA (TBE) para separar adaptadores não ligados e cDNAs pequenos. O cDNA foi selecionado por tamanho com um corte a 0,7 kb e resgatado do por uso de β-Agarase (New England Biolabs, Ipswich, MA, USA) de acordo com as instruções dos

fabricantes e precipitado durante 12 horas a -20°C por adição de dois volumes de etanol a 96% e 0,1 volume de acetato de sódio 3 M pH 5,2.

O cDNA direcional, selecionado por tamanho, foi recuperado por centrifugação a 13.000 x g, lavado em 70% de etanol, secado, e então
5 recolocado em suspensão em 30 µl de 10 mM de Tris-HCl-1 mM de EDTA pH 7,5. Os cDNAs foram dessalinizadas por filtração em gel através de uma coluna giratória MicroSpin S-300 HR de acordo com as instruções dos fabricantes. Três ligações de teste foram realizadas em 10 µl de tampão de ligação (30 mM de Tris-HCl, pH 7,8, 10 mM de MgCl₂, 10 mM de DTT, 0,5
10 mM de ATP) contendo 5 µl de cDNA filamento duplo (tubos de reação #1 e #2), 15 unidades de T4 ligase (Promega), e 30 ng (tubo #1), 40 ng (tubo #2), e 40 ng (tubo #3, controle de fundo de vetor) de vetor *Bst* XI-*Not* I clivado pYES2,0 (Invitrogen, Carlsbad, CA, USA). As reações de ligação foram realizadas por incubação a 16 °C durante 12 horas, então aquecimento a 70°C
15 durante 20 minutos, e finalmente adição de 10 µl de água em cada tubo. Um µl de cada mistura de ligação foi electroporado em 40 µl de células eletrocompetentes de *E. coli* DH 10B (Bethesda Research Laboratories) como descrito por Sambrook *et al.*, 1989, *supra*.

A biblioteca de cDNA de *Myceliophthora thermophila* CBS
20 117.65 foi estabelecida em *E. coli* DH 10B consistindo de massas. Cada massa foi feita por espalhamento de *E. coli* transformado em placas de ampicilina LB, dando 15.000 - 30.000 colônias/placa após incubação a 37 °C durante 24 horas. Vinte ml de meio de ampicilina LB foram adicionados na placa e as células foram colocadas em suspensão dentro da mesma. A
25 suspensão de célula foi agitada a 100 rpm em um tubo de 50 ml durante 1 hora a 37 °C.

A biblioteca de cDNA de *Myceliophthora thermophila* CBS 117,65 resultante consistia de aproximadamente 10⁶ clones individuais, com um vetor de fundo de 1%. DNA de plasmídeo de algumas das massas da

biblioteca foi isolado usando um kit de Plasmídeo Midi (QIAGEN Inc., Valencia, CA, USA), de acordo com as instruções dos fabricantes, e armazenado a -20°C .

Exemplo 2: Triagem de bibliotecas de expressão de *Myceliophthora thermophila* CBS 117.65 para atividade de endoglucanase

Alíquotas de 1 ml de plasmídeo de DNA purificado (100 ng/ml) de algumas das massas da biblioteca (exemplo 1) foram transformadas em *Saccharomyces cerevisiae* W3124 por eletroporação (Becker e Guarante, 1991, *Methods Enzymol.* 194: 182 - 187) e os transformantes foram colocados em ágar SC contendo 2% de glicose e incubados a 30°C . No total, 50 - 100 placas contendo 250 - 400 colônias de levedura foram obtidas de cada massa.

Após 3-5 dias de incubação, as placas de ágar SC foram colocadas em placas de réplica obre um conjunto de placas de agar de celulose SC de AZCL URA a 0,1% com galactose. As placas foram incubadas durante 2-4 dias a 30°C e colônias positivas de endoglucanase foram identificadas como colônias circundadas por um halo azul.

Exemplo 3: Caracterização do gene *Myceliophthora thermophila* CBS 117.65 *cel5a*

Colônias de levedura expressando endoglucanase foram inoculadas em 20 ml de meio YPD em 50 ml de tubos de teste de vidro. Os tubos foram agitados a 200 rpm durante 2 dias a 30°C . As células foram coletadas por centrifugação durante 10 minutos a 3000 rpm em uma centrífuga Heraeus Megafuge 1,0R com um rotor de 75002252 (Hanau, Alemanha).

O DNA foi isolado de acordo com WO 94/14953 e dissolvido em 50 de água deionizada. O DNA foi transformado em células de *E. coli* DH10B por procedimentos padrões de acordo com Sambrook *et al.*, 1989, *supra*. Um transformante de *E. coli* mostra subsequente conter o gene *Myceliophthora thermophila* CBS 117.65 *cel5a* foi designado pCIC161

(figura 1) e usado como o material para depósito de material biológico. A cepa de *E. coli* pCIC161 foi depositada como *E. coli* NRRL B-30902 em 23 de fevereiro de 2006.

O DNA plasmídeo foi isolado dos transformantes de *E. coli* usando procedimentos padrões de acordo com Sambrook *et al.*, 1989, *supra*. A sequência de cDNA de comprimento completo do gene *cel5a* de *Myceliophthora thermophila* CBS 117.65 foi seqüenciada com um Kit Taq DyeDeoxy Terminator Cycle Seqüencing (Perkin Elmer, Wellesley, MA, USA) e iniciadores de oligonucleotídeo sintético usando um seqüenciador de DNA da Applied Biosystems ABI PRISM™ 377 (ABI, Foster City, CA, USA) de acordo com as instruções dos fabricantes.

A sequência de nucleotídeos (SEQ ID NO: 3) e sequência de aminoácido deduzida (SEQ ID NO: 4) do gene *cel5a* de *Myceliophthora thermophila* são mostrados na figura 2. A sequência de codificação tem 1170 bp incluindo o códon de parado. A proteína prevista da codificação contém 389 aminoácidos. O %G+C da região de codificação do gene é 63,6% e a região de codificação de polipeptídeo maduro é 63,6%. Usando o programa SignalP, versão 3 (Nielsen *et al.*, 1997, *Protein Engineering* 10: 1-6), um peptídeo de sinal de 16 resíduos foi previsto. A proteína madura prevista contém 373 aminoácidos com uma massa molecular de 40,9 kDa.

Análise da sequência de aminoácido deduzida do gene *cel5a* com o programa Interproscan (Zdobnov e Apweiler, 2001, *Bioinformatics* 17: 847 - 848) mostrou que a proteína CEL5A continha a sequência de núcleo típica de uma família 5 da glicosil hidrolase, se estendendo de aproximadamente resíduo 77 de aminoácido a resíduo 350 do polipeptídeo maduro previsto. A proteína CEL5A também continha a assinatura de sequência de um domínio de ligação de celulose de fungo tipo I (CBMI). Esta assinatura de sequência conhecida como Prosite padrão PS00562 (Sigrist *et al.*, 2002, *Brief Bioinform.* 3: 265 - 274) estava presente do resíduo 8 de

aminoácido a resíduo 35 do polipeptídeo maduro previsto.

Um alinhamento global à feição de pares comparativos de seqüências de aminoácidos foi determinado usando o algoritmo Needleman-Wunsch (Needleman e Wunsch, 1970, *J. Mol. Biol.* 48: 443 - 453) como implementado no programa de Needle de EMBOSS com penalidade de abertura de espaço de 10, penalidade de extensão de espaço de 0,5, e a matriz EBLOSUM62. O alinhamento mostrou que a seqüência de aminoácido deduzida do gene de *Myceliophthora thermophila* codificando o polipeptídeo maduro CEL5A compartilhou 75% e 72% de identidade (excluindo os espaços) para as seqüências de aminoácidos deduzidas de duas proteínas de Família 5 da glicosil hidrolase de *Neurospora crassa* e *Humicola insolens*, respectivamente (números de acesso Q7SDR1 e G12624, respectivamente).

Exemplo 4: Expressão de gene *cel5a* de *Myceliophthora thermophila* CBS 117.65 em *Aspergillus oryzae*

O gene *cel5a* de *Myceliophthora thermophila* CBS 117.65 foi excisado do vetor pYES2.0 usando *Hind* III e *Xba* I, e ligado no vetor de expressão de *Aspergillus* pHD414 (EP 238 023, WO 93/11249) usando métodos padrões (Sambrook *et al.*, 1989, *supra*). O vetor de expressão de *Aspergillus* pHD414 é um derivado de p775 (EP 238 023). O plasmídeo resultante foi designado pA2C181 (figura 3).

Protoplastos de *Aspergillus oryzae* HowB104 foram preparados como descrito em WO 95/02043. Cem microlitros de suspensão de protoplastos foram misturados com 5-25 µg do vetor de expressão de *Aspergillus* pA2C161 em 10 µl de composto STC de 1,2 M de sorbitol, 10 mM de Tris-HCl, pH 7,5, 10 mM de CaCl₂) e ainda misturado com 5-25 µg de p3SR2, um plasmídeo transportando um gene *amdS* de *Aspergillus nidulans* (Christensen *et al.*, 1988, *Bio/Technology* 6: 1419 - 1422). A mistura foi deixada em temperatura ambiente durante 25 minutos. Duzentos microlitros de PEG 4000 a 80% (BDH, Poole, England) (polietileno glicol,

peso molecular 4.000), 10 mM de CaCl₂, e 10 mM de Tris-HCl pH 7,5 foram adicionados e suavemente misturados e finalmente 0,85 ml da mesma solução foi adicionado e suavemente misturado. A mistura foi deixada em temperatura ambiente durante 25 minutos, centrifugada a 2.500 x g durante 15 minutos, e o grânulo foi recolocado em suspensão em 2 ml de 1,2 M de sorbitol. Este processo de sedimentação foi repetido, e os protoplastos foram espalhados em placas COVE. Após incubação durante 4-7 dias a 37 °C esporos foram tomados e espalhados a fim de isolar as colônias únicas. Este procedimento foi repetido e esporos de uma colônia única após o segundo reisolamento foram armazenados.

Cada um dos transformantes foi inoculado em 10 ml de meio YPM. Após 2-5 dias de incubação a 30°C , 200 rpm, o sobrenadante foi removido. A atividade de endoglucanase foi identificada por aplicação de 20 µl de caldo de cultura de orifícios de 4 mm de diâmetro perfurados em uma placa de SC-ágar de celulose AZCL HE a 0,1% e incubação durante a noite a 30°C . A atividade de endoglucanase foi então identificada por um halo azul em torno de uma colônia. Vários caldos transformantes tem atividade de endoglucanase significativamente maior do que o caldo de um fundo de controle de *Aspergillus oryzae* não transformado, que demonstrou expressão eficiente da CEL5A endoglucanase de *Myceliophthora thermophila* CBS 117.65 em *Aspergillus oryzae*.

Exemplo 5: Construção de uma biblioteca de expressão de cDNA de basidiomicete CBS 495.95 em *Saccharomyces cerevisiae*

Uma biblioteca de cDNA de basidiomicete CBS 495.95, consistindo de aproximadamente 10⁶ clones individuais foi construída em *E. coli* como descrito no exemplo 1, com um fundo de vetor de 1%.

Exemplo 6: Triagem de bibliotecas de expressão de cDNA de basidiomicete CBS 495,95 para atividade de endoglucanase

Triagem da biblioteca de cDNA (exemplo 5) foi realizada

como descrito no exemplo 2.

Exemplo 7: Caracterização de um gene codificando CEL5A de basidiomicete CBS 495.95

Clonagem do gene *cel5a* de basidiomicete CBS 495.95 foi realizada como descrito no exemplo 3. Um transformante *E. coli* subsequente mostrado como contendo um gene *cel5a* foi designado pCIC453 (figura 4) e usado como o material para depósito de material biológico. A cepa pCIC453 de *E. coli* foi depositada como *E. coli* NRRL B-30903 em 23 de fevereiro de 2006.

O gene *cel5a* de basidiomicete CBS 495.95 foi excisado de pCIC453 usando *Hind* III e *Xba* I e ligado no vetor de expressão pHD414 de *Aspergillus*. O plasmídeo resultante foi designado pA2C453 (figura 5).

A sequência de nucleotídeos (SEQ ID NO: 5) e a sequência de aminoácido deduzida (SEQ ID NO: 6) do gene *cel5a* CBS 495.95 são mostradas na figura 6. A sequência de codificação é 1194 bp incluindo o códon de parada. A %G+C da região de codificação do gene é de 59,8% e a região de codificação de polipeptídeo maduro é de 60,1%. A proteína prevista de codificação contém 397 aminoácidos. Usando o programa SignalP, versão 3 (Nielsen *et al.*, 1997, *supra*), um peptídeo de sinal de 15 resíduos foi previsto. A proteína madura prevista contém 382 aminoácidos com uma massa molecular de 40,1 kDa.

Análise da sequência de aminoácido deduzida do gene *cel5a* com o programa Interproscan (Zdobnov e Apweiler, 2001, *supra*) mostrou que a proteína CEL5A continha a sequência de núcleo típica de uma Família 5 da glicosil hidrolase, se estendendo a partir de aproximadamente resíduos 81 a 359 do polipeptídeo maduro previsto. A proteína CEL5A também contida uma assinatura de sequência de um domínio de ligação de celulose de fungo de tipo I (CBMI). Esta assinatura de sequência conhecida como Prosite PS00562 padrão (Sigrist *et al.*, 2002, *supra*) estava presente do resíduo de

aminoácido 8 a resíduo 36 do polipeptídeo maduro previsto.

Um alinhamento global à feição de pares comparativos de seqüências de aminoácidos foi determinado usando o algoritmo Needleman-Wunsch (Needleman e Wunsch, 1970, *supra*) como implementado no programa de Needle EMBOSS com penalidade de abertura de espaço de 10, penalidade de extensão de espaço de 0,5, e a matriz EBLOSUM62. O alinhamento mostrou que a seqüência de aminoácido deduzida do gene CBS 495.95 codificando o polipeptídeo maduro CEL5A compartilhou 82% e 79% de identidade (excluindo espaços) para as seqüências de aminoácidos deduzidas de duas proteínas de Família 5 da glicosil hidrolase de *Irpex lacteus* e *Trametes hirsuta*, respectivamente (números de acesso Q5W7K4 e Q75UV6, respectivamente).

Exemplo 8: Expressão de gene *cel5a* de basidiomicete CBS 495.95 em *Aspergillus oryzae*

Expressão do gene *cel5a* de basidiomicete CBS 495.95, e análise de atividade de endoglucanase foi realizada como descrito no exemplo 4.

Exemplo 9: Construção de uma biblioteca de expressão de cDNA de basidiomicete CBS 494.95 em *Saccharomyces cerevisiae*

Uma biblioteca de cDNA de basidiomicete CBS 494.95, consistindo de aproximadamente 10^6 clones individuais foi construída em *E. coli* como descrito no exemplo 1, com um fundo de vetor de 1%.

Exemplo 10: Triagem de bibliotecas de expressão de cDNA de basidiomicete CBS 494.95 para atividade de endoglucanase

Triagem da biblioteca de cDNA (exemplo 9) foi realizada como descrito no exemplo 2.

Exemplo 11: Caracterização de um gene de codificação de CEL5B de basidiomicete CBS 494.95

Clonagem do gene *cel5b* de basidiomicete CBS 494.95 foi

realizada como descrito no exemplo 3. Um transformante de *E. coli* subsequentemente mostrado como contendo o gene CBS 494.95 *cel5b* foi designado pCIC486 (figura 7) e usada como o material para depósito de material biológico. A cepa *E. coli* pCIC486 foi depositada como *E. coli* NRRL B-30904 23 de fevereiro de 2006.

O gene *cel5b* de basidiomicete CBS 494.95 foi excisado do vetor pYES2.0 usando *Kpn* I e *Xho* I e ligado no vetor de expressão de *Aspergillus* pHD423 (Lassen *et al.*, 2001, *Appl Environ Microbiol* 67: 4701 - 4707), um derivado de pHD414 com um sítio *Kpn* I no poliligador. O plasmídeo resultante foi designado pA2C486 (figura 8). A sequência de nucleotídeos (SEQ ID NO: 7) e sequência de aminoácido deduzidas (SEQ ID NO: 8) do gene *cel5b* de CBS 494.95 são mostradas na figura 9. A sequência de codificação é 1290 bp incluindo o códon de parada. %G+C da região de codificação do gene é de 56,0% G+C e a região de codificação de polipeptídeo maduro é de 58,1%. A proteína prevista de codificação contém 429 aminoácidos. Usando o programa SignalP, versão 3 (Nieisen *et al.*, 1997, *supra*), um peptídeo de sinal de 21 resíduos foi previsto. A proteína madura prevista contém 408 aminoácidos com uma massa molecular de 43,1 kDa.

Análise da sequência de aminoácido deduzida do gene *cel5b* com o programa Interproscan (Zdobnov e Apweiler, 2001, *supra*) mostrou que a proteína CEL5B contida a sequência de núcleo típica de uma Família 5 da glicosil hidrolase, se estendendo de aproximadamente resíduo de aminoácido 106 a resíduo 385 do polipeptídeo maduro previsto. A proteína CEL5A também continha a assinatura de sequência de um domínio de ligação de celulose de fungo de tipo I (CBMS). Esta assinatura de sequência conhecida como Prosife PS00562 padrão (Sigrist *et al.*, 2002, *supra*) estava presente do resíduo de aminoácido 7 ao resíduo 34 do polipeptídeo maduro previsto.

Um alinhamento global à feição de pares comparativos de

seqüências de aminoácido foi determinado usando o algoritmo Needleman-Wunsch (Needleman e Wunsch, 1970, *supra*) como implementado no programa Needle de EMBOSS com penalidade de abertura de espaço de 10, penalidade de extensão de espaço de 0,5, e a matriz EBLOSUM62. O alinhamento mostrou que a seqüência de aminoácido deduzida do gene CBS 495.95 codificando o polipeptídeo maduro CEL5A compartilhou 69% e 67% de identidade (excluindo espaços) para as seqüências de aminoácido deduzidas de duas proteínas de Família 5 da glicosil hidrolase de *Irpex lacteus* e *Trametes hirsuta*, respectivamente (números de acesso Q5W7K4 e Q75UV6, respectivamente).

Exemplo 12: Expressão de CEL5B de basidiomicete CBS 494.95 em *Aspergillus oryzae*

Expressão do gene *cel5B* de basidiomicete CBS 494.95, e análise de atividade de endoglucanase foi realizada como descrito no exemplo 4.

Exemplo 13: Purificação de endoglucanases recombinantes de *Myceliophthora thermophila* CBS 117.65, basidiomicete CBS 494.95, e basidiomicete CBS 495.95

As endoglucanases de *Myceliophthora thermophila* CBS 117.65, basidiomicete CBS 494.95, e basidiomicete CBS 495.95, produziram recombinantemente em *Aspergillus oryzae* como descrito nos exemplos 4, 8, e 12, foram purificadas até homogeneidade usando um protocolo essencialmente como descrito por Otzen *et al.*, 1999, *Protein Sci.* 8: 1878 - 87.

A concentração de proteína nas preparações de enzima foi determinada usando o teste de microplaca com ácido bicinonínico (BCA) de acordo com as instruções dos fabricantes para um kit de reagente de teste de Proteína BCA (Pierce Chemical Co., Rockford, IL. USA).

Diluições de enzima foram preparadas logo antes de cada

experimento de soluções de enzima de carga, que foram armazenadas a - 20°C.

Exemplo 14: Isolamento de DNA genômico de *Penicillium brasilianum* IBT 20888

5 Esporos de cepa *Penicillium brasilianum* IBT 20888 foram propagados em arroz de acordo com Carlsen, 1994, Ph. D. thesis, Department of Biotechnology, The Technical University of Dinamarca. Os esporos foram recuperados com 20 ml de 0,1% de Tween 20 e inoculados em uma concentração de 1×10^6 esporos por ml em 100 ml de meio Mandels e Weber
10 (Mandels e Weber, 1969, *Adv. Chem. Ser.* 95: 394 - 414) contendo 1% de glicose suplementada por litro com 0,25 g of extrato de levedura e 0,75 g de Bactopeptona em um frasco de agitação com anteparos de 500 ml. Micélios de fungo foram colhidos após 24 horas de crescimento aeróbico a 30°C , 150 rpm.

15 Micélios foram coletados por filtração através de um filtro Nalgene DS0281-5000 (Nalge Nunc International Corporation, Rochester, NY, USA) até secura e congelados em nitrogênio líquido. Os micélios congelados foram triturados um pó em um almofariz resfriado com gelo seco e distribuídos em um tubo com tampa de rosca. O pó foi colocado em
20 suspensão em um volume total de 40 ml de 50 mM de tampão CAPS (ácido 3-(ciclo-hexilamino)-1-propanossulfônico)-NaQH pH 11 contendo 0,5% de sulfato de dodecil de lítio e 0,5 mM de EDTA. A suspensão foi colocada a 60°C durante 2 horas e periodicamente recolocada em suspensão por inversão. À suspensão foi adicionado um volume igual de fenol:clorofórmio
25 (1:1 v/v) neutralizado com 0,1 M Tris base, e o tubo foi misturado em uma roda girando a 37 °C durante 2 horas. Após centrifugação a 2500 rpm durante 10 minutos em um rotor Sorvall H1000B, a fase aquosa (fase de topo) foi re-extraída novamente com fenol:clorofórmio (1:1 v/v) e centrifugada a 15.000 x g durante 5 minutos. A fase aquosa da segunda extração foi levada a 2,5 M de

acetato de amônio (carga 10 M) e colocada a -20°C até congelar. Após descongelar, o extrato foi centrifugado a 15.000 x g durante 20 minutos em um rotor frio. O grânulo (primariamente rRNA) foi descartado e os ácidos nucleicos no sobrenadante foram precipitados por adição de 0,7 volumes de isopropanol. Após centrifugação a 15.000 x g durante 15 minutos, o grânulo foi enxaguado três vezes com 5 ml de 70% de etanol (sem recolocação em suspensão), secado ao ar quase que completamente, e dissolvido em 1,0 ml de 0,1 X TE. O grânulo dissolvido foi transferido para dois tubos de microcentrífugas de 1,5 ml. Os ácidos nucleicos foram precipitados por adição de acetato de amônio (0,125 ml) para 2,0 M e etanol a 63% (1,07 ml) e centrifugados a velocidade máxima durante 10 minutos em uma microcentrífuga Sorvall MC 12V (Kendro Laboratory Products, Asheville, NC, USA). O grânulo foi enxaguado duas vezes com 70% de etanol, secado ao ar completamente, e dissolvido em 500 µl de 0,1 X TE.

Exemplo 15: Preparação de uma biblioteca de DNA genômico de *Penicillium brasilianum* IBT 20888

Bibliotecas genômicas foram construídas usando um Kit TOPO Shotgun Subcloning (Invitrogen, Carlsbad, CA, USA). Brevemente, DNA celular total foi cisalhado por nebulização sob 0,703 kg/cm² de nitrogênio durante 15 segundos e fracionado em tamanho em géis agarose a 1% usando 40 mM de tampão Tris base-20 mM de acetato de sódio-1 mM dissódico EDTA (TAE). Fragmentos de DNA migrando na faixa de tamanho 3-6 kb foram excisados e eluídos usando um Kit de extração com gel MiniElute™ (QIAGEN Inc. Valencia, CA, USA). Os fragmentos eluídos foram de fracionados em tamanho novamente usando um agarose gel a 1% como acima e fragmentos de DNA migrando na faixa de tamanho 3-6 kb foram excisados e eluídos usando um Kit de extração com gel MiniElute™.

Os fragmentos de DNA eluídos foram reparados na extremidade obtusa e desfosforilados usando fosfatase alcalina de camarão

(Roche Applied Science, Mannheim, Alemanha). Os fragmentos de DNA de extremidade obtusa foram clonados no vetor pCR4Blunt-TOPO (Invitrogen, Carlsbad, CA, USA) de acordo com as instruções dos fabricantes, transformados em células de *E. coli* TOP10 eletrocompetentes por eletroporação, e colocados em placas de ampicilina LB. A eletroporação resultou em 15.300 clones.

Exemplo 16: Purificação de endoglucanase CEL5C nativa de *Penicillium brasilianum* IBT 20888

A endoglucanase foi purificada e testada como descrito em Jørgensen *et al.*, 2003 (*Enzyme Microb. Technol.* 32: 851 - 861). O substrato foi azo-carboximetil celulose (Megazyme International Ireland Ltd., Bray, Irlanda) e o tempo de incubação foi de 15 minutos. A enzima purificada foi armazenada congelada a -20°C .

Exemplo 17: Seqüenciamento N-terminal de endoglucanase CEL5C de *Penicillium brasilianum* IBT 20888

A amostra purificada de endoglucanase CEL5C de *Penicillium brasilianum* (exemplo 16) foi descongelada. Uma alíquota de 100 µl da amostra foi adicionada a 100 µl de tampão de amostra SDS-PAGE (4 ml de 0,5 M de TRIS-HCl, pH 6,8, 20 ml de 10% de SDS, 20 ml de glicerol (87%), 56 ml de água ultrapura Milli-Q®, e 15 grão de azul de bromofenol) em um tubo Eppendorf e aquecida a 95 °C durante 4 minutos. Após o aquecimento, quatro alíquotas de 20 µl de da amostra diluída foram aplicadas separadamente a um gel SDS poliacrilamida a 4-20% pré-moldado (Invitrogen, Carlsbad, CA, USA). Além das quatro trilhas contendo a amostra, uma mistura padrão de proteína Mark 12 (Invitrogen, Carlsbad, CA, USA) foi aplicada ao gel.

O gel foi colocado em um aparelho de gel Xcell SureLock™ (Invitrogen, Carlsbad, CA, USA) durante 90 minutos com um ajuste de potência de 40 mA no máximo 135 V. Após a eletroforese, o gel foi incubado

durante 5 minutos em uma solução de blotting consistindo de 10 mM de CAPS pH 11 contendo 6% de metanol. Uma membrana ProBlott (Applied Biosystems, Foster City, CA, USA) foi umedecida durante 1 minuto em metanol puro antes de ser colocada na solução de blotting durante 5 minutos a fim de saturar a membrana com 10 mM de CAPS pH 11 contendo 6% de metanol.

Electroblotting foi realizado em um aparelho Semi Dry Blotter II (KemEnTec, Copenhagen, DK) como a seguir. Seis pedaços de papel Whatman no. 1 umedecido na solução de blotting foram colocados no eletrodo positivo do aparelho de blotting seguido pela membrana ProBlott, o gel poliacrilamida, e seis pedaços de papel Whatman no. 1 umedecido na solução de blotting. O aparelho de blotting foi montado, assim colocando o eletrodo negativo em contato com a pilha superior de papel Whatman no. 1. Um peso de 11,3 kg foi colocado no topo do aparelho de blotting. Electroblotting foi realizado em uma corrente de 175 mA durante 180 minutos.

Após electroblotting, a membrana ProBlott foi colorida durante 1 minuto em 0,1% (p/v) Azul Brilhante Coomassie R-250 dissolvido em 60% de metanol, 1% de ácido acético, 39% de H₂O. A des-coloração da membrana ProBlott foi realizada em 40% de metanol aquoso durante 5 minutos antes da membrana ser enxaguada em água deionizada. Finalmente a membrana ProBlott foi secada ao ar.

Para seqüenciar o aminoácido N-terminal, dois pedaços da membrana ProBlott consistindo de uma banda de 65 kDa foram cortados e colocados no cartucho de blotting de um seqüenciador de proteína Procise da Applied Biosystems (Applied Biosystems, Foster City, CA, USA). O seqüenciamento N-terminal foi realizado usando o método "run file!" para amostras de membrana PVDF (PVDF de líquido pulsado) de acordo com as instruções dos fabricantes.

A sequência de aminoácido N-terminal foi deduzida dos cromatogramas resultantes por comparação do tempo de retenção dos picos nos cromatogramas nos tempos de retenção dos aminoácidos - PTH no cromatograma padrão.

5 A sequência de aminoácido N-terminal da endoglucanase de *Penicillium brasilianum* CEL5C purificada foi determinada diretamente usando um sistema de seqüenciamento Procise 494 HT (Applied Biosystems, Foster City, CA, USA). A sequência N-terminal foi determinada como sendo Ala-Ser-Ser-Phe-Val-Trp-Phe-Gly-Thr-Ser-Glu-Ser-Gly-Ala-Glu-Phe-Gly-
10 Asn-Gln-Asn (aminoácidos 25 a 44 em SEQ ID NO: 10).

Exemplo 18: Amplificação PCR do gene endoglucanase *cel5c* de *Penicillium brasilianum* IBT 20888

Com base na sequência de aminoácido N-terminal da endoglucanase de *Penicillium brasilianum* purificada, um iniciador dianteiro
15 foi designado usando a estratégia CODEHOP (Rose *et al.*, 1998, *Nucleic Acids Res.* 26: 1628 - 35). A partir do banco de dados, a informação sobre outras endoglucanases, dois iniciadores reversos foram designados como mostrado abaixo usando a estratégia CODEHOP.

Iniciador dianteiro, Fwd:

20 5'-TTCGGTACCTCTGAGTCTGGNGCNGARTT-3' (SEQ ID NO: 11)

Iniciador reverso, Rev1:

5'-TGATCCATATCGTGGTACTCGTTRTTNGTRTCRAA-3' (SEQ ID NO: 12)

Iniciador reverso, Rev2:

25 5'-CCGTTGTAGCGACCGTARTTRTGNGGRTC-3' (SEQ ID NO: 13)

onde R=A ou G, Y = C ou T, K = G ou T e N = A, C, G ou T

Reações de amplificação (30 µl) foram preparadas usando aproximadamente 1 µg de DNA genômico de *Penicillium brasilianum* como gabarito. Além disso, cada reação continha os seguintes componentes: 30

pmol do iniciador dianteiro, 30 pmol do iniciador reverso, 200 μ M cada de dATP, dCTP, dGTP, e dTTP, tampão 1X AmpliTaq polimerase (Applied Biosystems, Foster City, CA, USA), e 0,5 unidade de AmpliTaq polimerase (5,0 U/ μ l, Applied Biosystems, Foster City, CA, USA). As reações foram incubadas em um Robocycler (Stratagene, La Jolla, CA, USA) programado para 1 ciclo a 96 °C durante 3 minutos e a 72 °C durante 3 minutos; 34 ciclos cada a 95 °C durante 0,5 minuto, 56 °C durante 0,5 minutos, e 72 °C durante 1,5 minutos; 1 ciclo a 72 °C durante 7 minutos; e um ciclo de embebiamento a 6 °C. Taq polimerase foi adicionado a 72 °C no primeiro ciclo.

Produtos de reação PCR foram separados em um gel agarose a 2% (Amresco, Solon, OH, USA) usando o tampão TAE. Uma banda de aproximadamente 650 bp (iniciadores Fwd e Rev1) e bandas de aproximadamente 320 e 380 bp (iniciadores Fwd e Rev2) foram excisadas do gel e purificadas usando um Kit de extração de gel MiniElute™ (QIAGEN Inc., Valencia, CA, USA) de acordo com as instruções dos fabricantes. Os produtos PCR purificados foram subsequentemente analisados por seqüenciamento de DNA. O produto 320 bp foi verificado como codificando uma porção de um polipeptídeo de Família 5 da glicosil hidrolase que foi designado CEL5C.

Exemplo 19: Triagem de biblioteca genômica de *Penicillium brasilianum* IBT 20888

Retiradas da colônia da biblioteca descrita no exemplo 15 foram realizadas (Maniatis *et al.*, 1982, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, New York) e o DNA foi reticulado sobre membranas Hybond N+ (Amersham, Arlington Heights, IL, USA) durante 2 horas a 80°C . A membranas das retiradas das colônias foram pré-umedecidas usando 0,2X de SSC (30 mM de NaCl, 3 mM de citrato de sódio), 0,2% de SDS. Os filtros pré-umedecidos foram colocados em um béquer com 7,5 ml de 6X de SSPE (0,9 M de NaCl, 0,06 M de

NaH₂PO₄, e 6 mM de EDTA), 7% de SDS) por filtro a 68 °C em um banho de água agitando durante 30 minutos.

Aproximadamente 40 ng do produto PCR descrito no exemplo 18 foram rotulados com iniciador aleatório usando um kit de iniciador-It II Stratagene (Stratagene, La Jolla, CA) de acordo com as instruções dos fabricantes. O fragmento de gene radio-rotulado foi separado de nucleotídeo não incorporado usando um kit de purificação MinElute PCR (QIAGEN Inc., Valencia, CA, USA).

A sonda radioativa foi desnaturada por adição de 5,0 M de NaOH em uma concentração final de 0,5 M, e adicionada na solução de hibridização em uma atividade de aproximadamente 0,5 x 10⁶ cpm por ml de solução de hibridização. A mistura foi incubada durante 10 horas a 68 °C em um banho de água com agitação. Após a incubação, as membranas foram lavadas três vezes em 0,2X de SSC, 0,2% de SDS a 68 °C. As membranas foram então secadas em papel blotting durante 15 minutos, enroladas em SaranWrap™, e expostas em filme de raio X durante a noite a -80°C com telas intensificadoras (Kodak, Rochester, NY, USA).

Colônias produzindo sinais de hibridização com a sonda foram inoculadas em 1 ml de meio de ampicilina LB e cultivadas durante a noite a 37 °C. Diluições de cada solução foram feitas e 100 µl foram colocadas sobre placas de ampicilina LB. A diluição para cada positivo que produziu cerca de 40 colônias por placa foi escolhida para retiradas secundárias. As retiradas foram preparadas, hibridizadas e sondadas como acima. Duas colônias de cada placa positiva foram inoculadas em 3 ml de meio de ampicilina LB e cultivadas durante a noite a 37 °C.

Miniprep DNA foi preparado em cada colônia usando um Bio Robot 9600 (QIAGEN Inc, Valencia, CA, USA) de acordo com o protocolo dos fabricantes. O tamanho de cada inserto foi determinado por restrição *Eco* RI de eletroforese de gel agarose. Dois clones continham cada um inserto de

aproximadamente 5,5 kb. O seqüenciamento revelou que os clones eram idênticos, e eles são a seguir referidos como pKKAH1 (ver exemplo 20).

Exemplo 20: Caracterização da seqüência genômica de *cel5c* codificando a endoglucanase CEL5C de *Penicillium brasilianum* IBT 20888

5 Seqüenciamento de DNA do gene de endoglucanase de *Penicillium brasilianum* de pKKAH1 foi realizado com um Seqüenciador de DNA automatizado Applied Biosystems Modelo 3700 (Applied Biosystems, Foster City, CA, USA) usando uma técnica *walking* de iniciador com química de terminado de corante (Giesecke *et al.*, 1992, *J. Virol. Methods* 38: 47- 60).

10 A seqüência de nucleotídeos (SEQ ID NO: 9) e seqüência de aminoácido deduzida (SEQ ID NO: 10) do gene *cel5c* de *Penicillium brasilianum* são mostradas nas figuras 10A e 10B. A seqüência de codificação genômica de 1471 bp (incluindo códon de parada) codifica um polipeptídeo de 421 aminoácidos, interrompido por 4 íntrons de 51 bp (89-139 bp), 47 bp (352-398 bp), 55 bp (464-518 bp), e 52 bp (617-668 bp). O teor em %G+C da região de codificação do gene é 51,2% e a região de codificação de polipeptídeo maduro é 50,8%. Usando o programa SignalP (Nielsen *et al.*, 1997, *Protein Engineering* 10: 1-6), um peptídeo de sinal de 16 resíduos foi previsto. Com base na seqüência N-terminal da endoglucanase, resíduos 17 a 20 24 parecem constituir uma pró-região isto é proteoliticamente clivada durante a maturação. A proteína madura prevista contém 397 aminoácidos e tem uma massa prevista de 42,6 kDa.

25 Análise da seqüência de aminoácido deduzida do gene *cel5c* com o programa Interproscan (Zdobnov e Apweiler, 2001, *supra*) mostrou que a proteína CEL5C continha a seqüência de núcleo típica de Família 5 da glicosil hidrolase, se estendendo aproximadamente de resíduos 32 a 307 do polipeptídeo de comprimento completo previsto. A proteína CEL5C também continha a assinatura de seqüência de um domínio de ligação de celulose de fungo de tipo I (CBMI). Esta assinatura de seqüência conhecida como padrão

Prosite PS00562 (Sigrist *et al.*, 2002, *supra*) estava presente de resíduo de aminoácido 393 a resíduo 420 do polipeptídeo previsto.

Um alinhamento global à feição de pares comparativos de seqüências de aminoácido em banco de dados público foi determinado usando o algoritmo de Needleman-Wunsch (Needleman e Wunsch, 1970, *supra*) como implementado no programa de Needle EMBOSS com penalidade de abertura de espaço de 10, penalidade de extensão de espaço de 0,5, e a matriz EBLOSUM62. O alinhamento mostrou que a seqüência de aminoácido deduzida do gene *cel5c* de *Penicillium brasilianum* codificando o polipeptídeo maduro CEL5C compartilhava 74,9% e 74,5% de identidade (excluindo espaços) para as seqüências de aminoácido deduzida de duas proteínas da Família 5 da glicosil hidrolase prevista de *Neosartorya fischeri* e *Aspergillus fumigatus*, respectivamente (números de acesso A1DAP7 e Q4WM09, respectivamente).

Exemplo 21: Construção de um plasmídeo de expressão *Aspergillus oryzae* para o gene *cel5c* de endoglucanase de *Penicillium brasilianum* IBT 20888

O plasmídeo de expressão *Aspergillus* pJaL721 (WO 03/008575) consiste de um cassete de expressão com base no promotor de amilase II neutra de *Aspergillus niger* fusionado na seqüência líder não traduzida de triose fosfato isomerase de *Aspergillus nidulans* (NA2-tpi) e o terminador de amiloglicosidade de *Aspergillus niger*. Também está presente no plasmídeo o marcador seletivo *amdS* de *Aspergillus nidulans* permitindo o crescimento em acetamida como a única fonte de nitrogênio e o marcador URA3 de *Saccharomyces cerevisiae* permitindo o crescimento da cepa *pyrF* defeituosa de *Escherichia coli* DB6507 (ATCC 35673). Transformação em *E. coli* DB6507 foi realizada usando o gene URA3 de *Saccharomyces cerevisiae* como marcador seletivo como descrito abaixo.

E. coli DB6507 foi tornado competente pelo método de

Mandel e Higa, 1970, *J. Mol. Biol.* 45: 154. Os transformantes foram selecionados em meio M9 sólido (J. Sambrook, E. F. Fritsch, e T. Maniatis, 1989, *Molecular Cloning, A Laboratory Manual*, segunda edição, Cold Spring Harbor, New York) suplementados por litro com 1 g de ácidos casamino, 500 µg de tiamina, e 10 mg de canamicina.

O gene endoglucanase foi clonado em pJaL721 como descrito abaixo. O gene *cel5c* de endoglucanase de *Penicillium brasilianum* foi amplificado por PCR usando os seguintes dois iniciadores de oligonucleotídeo:

10 PCR Dianteiro: 5'-AATTGGATCCACCAT GAAATACCCT CTACTCCTG GCAAC-3' (SEQ ID NO: 14)

PCR reverso:

5'-TTAACTCGAGTTACAGACACTGCGAATAATACGCATTC-3' (SEQ ID NO: 15)

15 Para facilitar a clonagem, um sítio de enzima de restrição foi inserido na extremidade 5' de cada iniciador onde o iniciador dianteiro continha um sítio *Bam* HI e o iniciador reverso continha um sítio *Xho* I.

O DNA genômico preparado como no exemplo 14 foi usado como gabarito na reação PCR. A reação foi realizada em um volume de 50 µl contendo 1,0 unidade de Phusion (Finnzymes Oy, Espoo, Finlândia), tampão 1X Phusion HF (Finnzymes Oy, Espoo, Finlândia), 500 ng de genômico, 250 µM de cada dNTP, e 50 pmol de cada dos dois iniciadores descritos acima. A amplificação foi realizada em um ciclador térmico PTC-220 DNA Engine Dyad Peltier (MJ Research, Inc., Waltham, MA, USA) programado para 1
25 ciclo a 95 °C durante 5 minutos; 24 ciclos cada a 94 °C durante 0,5 minuto, 58 °C durante 0,5 minuto, e 68 °C durante 4,0 minutos; e 1 ciclo a 68 °C durante 15 minutos. A técnica PCR de partida rápida (Chou *et al.*, 1992, *Nucleic Acids Res.* 20: 1717) foi usada e Phusion polimerase foi adicionada após 1 minuto do primeiro ciclo.

A reação PCR produziu um fragmento de DNA único de aproximadamente 1500 bp em comprimento. O fragmento foi digerido com *Bam* HI e *Xho* I e isolado por eletroforese agarose gel, purificado, e clonado em pJaL721 digerido com *Bam* HI e *Xho* I, resultando em um plasmídeo designado pKBK03 (figura 11). A sequência do gene de endoglucanase em pKBK03 foi verificada por seqüenciamento com um analisador de DNA 3730xl da Applied Biosystems.

A fim de criar um plasmídeo para depósito de material biológico, pKBK03 foi digerido com *Bam* HI e *Xho* I e purificado. O fragmento foi reparado na terminação obtusa usando enzima Klenow (Roche Applied Science, Mannheim, Alemanha) durante 30 minutos a 25 °C. O plasmídeo pUC13 foi digerido com *Sma* I e desfosforilado com Protease Intestinal de bezerro (Roche Applied Science, Mannheim, Alemanha; CIP) durante 1 hora a 37 °C e o CIP foi inativado por aquecimento da amostra a 80°C durante 15 minutos.

O fragmento reparado na extremidade obtusa e o fragmento pUC13 desfosforilado foram ligados durante a noite a 16 °C usando T4 DNA ligase (Roche Applied Science, Mannheim, Alemanha). Uma amostra de 0,25 µg do produto ligado foi transformada em DH5α *Escherichia coli* (Invitrogen, Carlsbad, CA, USA). Após incubação durante a noite a 37 °C em placas de ampicilina LB, os transformantes foram transferidos para 2 ml de meio LB e incubados a 37 °C. Um plasmídeo designado pPBCel5C (figura 12) foi purificado usando Jetquick Plasmid Miniprep (Genomed, Löhne, Alemanha). A sequência do gene de endoglucanase foi verificada por seqüenciamento com um analisador de DNA 3730xl de Applied Biosystems. As células de *E. coli* TOP10 (Invitrogen, Carlsbad, CA) contendo plasmídeo pPBCel5C (cepa de designação PBCel5C) foram depositadas junto ao Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, como NRRL B-30900N, com uma

data de depósito de 23 de fevereiro de 2006.

Exemplo 22: Expressão de endoglucanase CEL5C de *Penicillium brasilianum* IBT 20888 em *Aspergillus oryzae*

Aspergillus oryzae BECh2 (WO 00/30322) foi transformado com 5 µg de pKBK03 como descrito por Christensen *et al.*, 1988, *Biotechnology* 6: 1419 - 1422.

Os transformantes foram cultivados em 50 ml de tubos durante 4 dias a 30°C em 10 ml de meio YPM. Os caldos completos foram centrifugados a 12.100 x g e os sobrenadantes removidos. Os sobrenadantes foram analisados por SDS-PAGE usando um gel pré-moldado Criterion XT, 10% gel Bis-Tris em um tampão XT MES (BioRad Laboratories, Hercules, CA, USA) de acordo com as instruções dos fabricantes. Um volume de 10 µl de sobrenadante foi misturado com 9 µl de tampão de amostra (0,125 M de Tris-HCl pH 6,8, 20% de glicerol, e 4,6% de SDS), e 1 µl de 1 M de ditioneitol, e aquecido a 96 °C durante 5 minutos. Em 16 dentre os 20 sobrenadantes, uma banda de aproximadamente 65 kDa foi visível na faixa dos padrões 35 kDa a 150 kDa por SDS-PAGE. Os sobrenadantes resultantes em uma banda visível após SDS-PAGE também continham atividade de endoglucanase, testados como descrito no exemplo 3. Quanto maior a intensidade da banda, maior a atividade de endoglucanase medida no mesmo sobrenadante.

Um transformante foi designado *Aspergillus oryzae* KBK03.

Exemplo 23: Produção e purificação de endoglucanase CEL5C *Penicillium brasilianum* IBT 20888 recombinante

O transformante de *Aspergillus oryzae* KBK03 foi cultivado em vinte frascos de agitação de 500 ml com 200 ml de meio YPM.

A biomassa foi removida de 4,0 litros de caldo de fermentação por centrifugação e filtração. Análise SDS-PAGE foi realizada como descrito no exemplo 9. A solução endoglucanase foi carregada sobre uma coluna XK

50 (Amersham Biosciences, Uppsala, Suécia) contendo 110 g de Avicel Ph 101 (Merck KGaA, Darmstadt, Alemanha) pré-equilibrada com 25 mM de Tris pH 7,5, antes a carga e a enzima ligada foram eluídas com 25 mM de Tris, 1% de trietanolamina a pH 11,6. Eluição da endoglucanase foi monitorada a 280 nm. A proteína eluída contendo frações foram reunidas imediatamente e o pH ajustado a 7,5. Frações contendo a endoglucanase foram reunidas.

O teor de proteína foi determinado da absorbância a 280 nm e o coeficiente de extinção calculado da estrutura primária da endoglucanase.

A purificação foi seguida por SDS-PAGE. As amostras foram fervidas durante 2 minutos com um volume igual de 2X tampão de amostra e 1/5 volume de 1% de PMSF e carregadas sobre um gel a 4-20% de tris-glicina (Invitrogen, Carlsbad, CA, USA). O gel foi colorido com reagente de coloração GelCode Blue (Pierce, Rockford, IL, USA) e descolorido com água. SDS-PAGE revelou uma banda de aproximadamente 65 kDa.

Exemplo 24: Caracterização de endoglucanase CEL5C de *Penicillium brasilianum* IBT 20888 recombinante purificado

A endoglucanase CEL5C de *Penicillium brasilianum* recombinante purificada como descrito no exemplo 23 foi caracterizada com relação a pH ótimo, temperatura ótima e estabilidade da temperatura. A atividade de endoglucanase foi medida como descrito no exemplo 16 em temperaturas de 20°C a 80°C e em valores pH de 3,0 a 10,0. A endoglucanase purificada foi diluída em água Milli-Q® ultrapura (Millipore, Billerica, MA, USA) para assegurar que a atividade estava na curva padrão. Para o pH ótimo, o substrato foi dissolvido em tampão Britton-Robinson (50 mM de ácido bórico, 50 mM de ácido acético, 50 mM de ácido fosfórico) ajustado no pH desejado. A estabilidade da temperatura foi determinada durante 20 horas a 50°C no pH na faixa de 4,0 a 6,0. Todos testes experimentais foram realizados em duplicata.

Os resultados da determinação do pH ótimo são mostrados na figura 13. O pH ótimo estava próximo de 4,0 a 50°C com atividade muita pequena a pH 3,0 e aproximadamente 80% de atividade de pico a pH 5,0.

5 Os resultados da determinação ótima de temperatura são mostrados na figura 14. A temperatura ótima a pH 4,8 foi de aproximadamente 70°C com mais que 75% de atividade de pico de 60°C a 80°C .

10 Os resultados da determinação da estabilidade da temperatura são mostrados na figura 15. Quando pré-incubada na ausência de substrato durante 20 horas a 25 °C e 50°C na faixa de pH de 4,0 a 6,0, a endoglucanase reteve mais que 80% de sua atividade de partida.

Exemplo 25: Preparação de substratos

Forragem de milho pré-tratada (PCS) foi preparada pelo U.S. Department of Energy National Renewable Energy Laboratory (NREL) usando ácido sulfúrico diluído. As seguintes condições foram usadas para o pré-tratamento: 1,4% em peso de ácido sulfúrico a 165 °C e 7,5 kg/cm² durante 8 minutos. Análise da composição foi realizada pelo NREL. Celulose e hemicelulose foram determinadas por uma hidrólise de ácido sulfúrico de estágio dois com análise subsequente de açúcares por cromatografia de líquido de desempenho elevado (HPLC) usando Procedimento Analítico Padrão do NREL #002. Lignina foi determinada gravimetricamente após hidrolisar as frações de celulose e hemicelulose com ácido sulfúrico (Procedimento Analítico Padrão do NREL #003). Sólidos insolúveis em água no pré-tratado com a forragem (PCS) foram determinados para serem 56,5% de celulose, 4,6% de hemicelulose, e 28,4% de lignina.

25 A PCS foi lavada com um volume grande de água deionizada em um funil Kimax com um filtro de vidro de porosidade grossa (Fisher Scientific, Pittsburg, PA, USA). PCS lavada em água foi moída moedor de café e adicionalmente lavado com água deionizada em um filtro Millipore de

22 µm com uma membrana expressa 6P (Millipore, Bedford, MA, USA).
Peso seco do PCS moído foi de 32,4%.

Uma suspensão de carga de 10 mg/ml de celulose intumescida em ácido fosfórico PASC) em água deionizada foi preparada usando o
5 seguinte procedimento. Cento e cinquenta ml de ácido o-fosfórico a 85% resfriado em gelo foram adicionados a 5 g de Avicel PH101 (FMC Corp., Philadelphia, PA, USA) umedecido com água. A suspensão foi agitada lentamente em um banho de gelo durante uma hora, e 100 ml de acetona resfriada com gelo foram adicionados na suspensão com agitação constante.
10 A suspensão foi transferida em um funil Kimax com um filtro de vidro de porosidade grossa, lavada três vezes com 100 ml de acetona resfriada com gelo, e drenada tão completamente como possível após cada lavagem. Finalmente, a suspensão foi lavada duas vezes com 500 ml de água, e novamente drenada tão completamente como possível após cada lavagem. O
15 PASC foi misturado com água em um volume total de 500 ml. Azida sódica foi adicionada em uma concentração final de 0,02% para evitar crescimento microbiano. A suspensão foi homogeneizada usando um misturador e armazenada a 4 °C durante até um mês.

Carboximetilcelulose (CMC, sal de sódio, tipo 7L2) com um
20 grau médio de substituição (DS) de 0,7 foi obtido de Aqualon Division of Hercules Inc., Wilmington, DE, USA. Uma solução de 6,25 mg/ml de CMC em 50 mM de acetato de sódio pH 5,0 foi preparada por adição lentamente de CMC ao tampão vigorosamente agitado. A suspensão foi aquecida a aproximadamente 60°C sob agitação contínua até o CMC estar
25 completamente dissolvido.

Celulose bacteriana (BC) foi preparada de "Nata de Coco", uma celulose comercial de tipo alimentício (Fujicco Co., Kobe, Japão), como descrito em Boisset *et al.*, 1999, *Biochemical Journal*, 340: 829 - 835. Uma suspensão de 1 mg/ml de celulose bacteriana em água deionizada com 0,01%

(p/v) de azida de sódio foi armazenada a 4 °C.

Avicel PH101 foi obtido de FMC Corporation, Philadelphia, PA, USA.

Xilano de bétula foi obtido de Sigma, St. Louis, MO, USA.

- 5 Xiloglucano de semente de tamarindo (amilóide, lote 00401), arabinoxilano de trigo (viscosidade média, 27 cSt, lote 90601), 1,4-beta-D-manano (boroidreto reduzido, Man:Gal=97:3, grau de polimerização DP ~ 15, lote 90302), e galactomanano caroba (viscosidade baixa, boroidreto reduzido, lote 90301) foram obtidos de Megazyme, Bray, Irlanda.

10 **Exemplo 26: Teste de hidrazida de ácido p-hidroxibenzóico para determinação de açúcares redutores**

Açúcares redutores (RS) foram determinados por um teste de hidrazida de ácido p-hidroxibenzóico (PHBAH) (Lever, 1972, *Anal. Biochem.* 47: 273 - 279), que foi adaptado em um formato de microplaca 96 cavidades.

- 15 Uma alíquota de 90- μ l da amostra diluída foi colocada em cada cavidades de uma microplaca 96-cavidades de fundo cônico (Costar, policarbonato transparente, Corning Inc., Acton, MA, USA). O teste foi iniciado por adição de 60 μ l de 1,25% de PHBAH em 2% de hidróxido de sódio em cada cavidade. A placa não coberta foi aquecida em um bloco de aquecimento feito de acordo com o pedido, durante 10 minutos a 95 °C. Após
20 o aquecimento, a microplaca foi resfriada em temperatura ambiente, e 35 μ l de água deionizada foi adicionada em cada cavidade. Uma alíquota de 100 μ l foi removida de cada cavidade e transferida para uma placa de fundo plano de 96-cavidades (Costar, poliestireno de ligação média, Corning Inc., Acton,
25 MA, USA). A absorbância a 410 nm (A_{410}) foi medida usando uma leitora de microplaca SpectraMAX (Molecular Devices, Sunnyvale, CA, USA). O valor A_{410} foi traduzido em equivalentes de glicose usando uma curva padrão.

A curva padrão foi obtida com seis padrões de glicose (0,005, 0,010, 0,025, 0,050, 0,075, e 0,100 mg/ml), que foram tratados similarmente

nas amostras. Padrões de glicose foram preparados por diluição de 10 mg/ml de solução de glicose de carga com água deionizada.

Para todos os substratos exceto para xilano e arabinoxilano, o grau de conversão (%) foi calculado usando a seguinte equação:

$$\text{Conversão (\%)} = \text{RS}_{(\text{mg/ml})} \times 100 \times 162 / (\text{concentração de substrato inicial}_{(\text{mg/ml})} \times 180) = \text{RS}_{(\text{mg/ml})} \times 100 / (\text{concentração de substrato inicial}_{(\text{mg/ml})} \times 1,111)$$

5 Para xilano e arabinoxilano, porcentagem de substrato hidrolisado para RS foi calculado usando a seguinte equação:

$$\text{Conversão (\%)} = \text{RS}_{(\text{mg/ml})} \times 100 \times 132 / (\text{concentração de substrato inicial}_{(\text{mg/ml})} \times 150) = \text{RS}_{(\text{mg/ml})} \times 100 / (\text{concentração de substrato inicial}_{(\text{mg/ml})} \times 1,136)$$

10 Nestas equações, RS é a concentração de açúcares redutores em solução medida em equivalentes de glicose (mg/ml), e os fatores 1,111 e 1,136 refletem o ganho de peso na conversão de polissacarídeos correspondentes em açúcares hexose (MW 180) ou pentose (MW 150).

Exemplo 27: Atividade relativa de endoglucanases em carboximetilcelulose a 50 °C

15 Tabela 1 mostra a atividade relativa das endoglucanases purificadas de *Myceliophthora thermophila* CBS 117.85, basidiomicete CBS 494.95, e basidiomicete CBS 495.95 para o sal de sódio solúvel de carboximetilcelulose (CMC). A atividade relativa é mostrada como a porcentagem da atividade de endoglucanase de basidiomicete CBS 495.95. A atividade foi determinada por medida da concentração de açúcares redutores (RS) produzidos de CMC (5 mg/ml) após 30 minutos de hidrólise em 50 mM
20 de acetato de sódio pH 5,0 a 50°C . A hidrólise foi realizada sem agitação na presença de 0,5 mg/ml de albumina de soro bovino (BSA, Sigma, St. Louis, MO, USA). Açúcares redutores foram determinados usando teste de hidrazida de ácido p-hidroxibenzóico (PHBAH) descrito no exemplo 26.

Tabela 1. Atividade relativa de endoglucanases em carboximetilcelulose (5 mg/ml) a pH 5,0 e 50°C

Endoglucanase	Atividade em CMC, %
<i>Myceliophthora thermophila</i> Cel5A	28
basidiomicete CBS 494.95 Cel5B	44
basidiomicete CBS 495.95 Cel5A	100

Exemplo 28: Estabilidade térmica de endoglucanases a 40°C - 80°C

A estabilidade térmica das endoglucanases purificadas de *Myceliophthora thermophila* CBS 117.65, basidiomicete CBS 494.95, e basidiomicete CBS 495.95 foi determinada por incubação de soluções de enzima em cinco temperaturas (40°C , 50°C , 60°C , 70°C , e 80°C), e medindo a atividade residual de enzimas em carboximetilcelulose (CMC).

As enzimas foram diluídas em 50 mM de acetato de sódio pH 5,0, que continham 3,0 mg/ml de BSA, e incubadas durante 3 horas em microtubos de 1,1ml ImmunoWare dispostos em um formato de microplaca de 8x12 (Pierce, Rockford, IL, USA). BSA foi adicionado a fim de evitar possível absorção de enzima sobre as paredes plásticas de microtubos. A concentração de proteína nas misturas de incubação foi escolhida de modo que cada enzima deve dar menos que 1% de conversão de CMC em teste subsequente para atividade CMCase.

Após uma incubação de 3 horas, alíquotas de 15 µl foram removidas usando um pipetador de 8 canais, e adicionadas a 75 µl de solução de CMC (6 mg/ml em 50 mM de acetato de sódio pH 5,0) em uma microplaca de 36 cavidades de fundo cônico (Costar, polycarbonato transparente, Corning Inc., Acton, MA, USA). A atividade de CMCase residual foi então medida como descrito no exemplo 27, e expressa como uma percentagem da atividade CMCase inicial (tabela 2).

A 40°C e 50°C , todas as três endoglucanases eram estáveis e retinham 98-100% da atividade de CMCase inicial após 3 horas de incubação. A 60°C e 70°C, Cel5A de *Myceliophthora thermophila* mostrou melhor

estabilidade do que as duas outras endoglucanases, e reteve 100% e 49,3% da atividade CMCase inicial após uma incubação de 3 horas, respectivamente. Nenhuma das endoglucanases eram estáveis a 80°C .

Tabela 2. Atividade de CMCase residual de endoglucanases após incubação durante três horas a pH 5,0 e 40-80°C

Endoglucanase	Atividade CMCase residual, % de atividade inicial				
	40°C	50°C	60°C	70°C	80°C
<i>Myceliophthora thermophila</i> Cel5A	100	100	100	49,3	3,8
basidiomicete CBS 494.95 Cel5B	100	98	28,0	5,0	4,1
basidiomicete CBS 495.95 Cel5A	100	99	6,4	2,4	0,8

Exemplo 29: Atividade relativa de endoglucanases em celulose intumescida em ácido fosfórico a 40 - 70°C

A atividade das endoglucanases purificadas de basidiomicete CBS 494.95 e basidiomicete CBS 495.95 em celulose intumescida com ácido fosfórico (PASC) foi determinada por medida da concentração dos açúcares reduzidos (RS) liberados durante a hidrólise inicial de PASC (2 mg/ml) em 50 mM de acetato de sódio pH 5,0. A hidrólise foi realizada sem agitação na presença de 0,5 mg/ml de albumina de soro bovino (BSA, Sigma, St Louis, MO, USA). As enzimas foram diluídas de modo que uma concentração de RS iria aumentar linearmente durante o início de 30 a 90 minutos de hidrólise, e o grau de conversão PASC não devendo exceder 2% durante este tempo. Os açúcares reduzidos foram determinados usando o teste de hidrazida de ácido p-hidroxibenzóico (PHBAH) como descrito no exemplo 26.

A atividade relativa como uma função de temperatura das endoglucanases de basidiomicete CBS 494.95 e basidiomicete CBS 495.95 é mostrada na figura 16. A atividade é mostrada como a percentagem da atividade da endoglucanase de basidiomicete CBS 495.95 a 60°C . Para ambas endoglucanases, a atividade em PASC atingiu o valor máximo a $T_{opt.} = 60^{\circ}\text{C}$.

Exemplo 30: Hidrólise a longo prazo de celulose intumescida com ácido

fosfórico a 40 - 70°C

Hidrólise de celulose intumescida com ácido fosfórico (PASC, 2 mg/ml) pelas endoglucanases purificadas de basidiomicete CBS 494.95 e basidiomicete CBS 495.95 foi realizada durante 45 horas em 50 mM de acetato de sódio pH 5,0 contendo 0,5 mg/ml de BSA a quatro temperaturas, 40°C , 50°C , 60°C , e 70°C . As endoglucanases foram usadas em três cargas de proteínas, 0,056, 0,167, ou 0,5 mg por g de PASC. As reações com o volume inicial de 1 ml foram cicladas sem agitação em microtubos de 1,1-ml ImmunoWare dispostos em um formato de microplaca 8x12 (Pierce, Rockford. IL. USA).

Cem alíquotas de microlitro foram removidas das reações em pontos de tempo diferentes (1, 1,5, 3, 6, 21, 27, e 45 horas) usando um pipetador de 8 canais, e adicionado a 25 µl de 2% de NaOH em uma placa de filtração MultiScreen HV de 96 cavidades (Millipore, Bedford, MA, USA). As amostras coletadas foram filtradas a vácuo em uma microplaca de fundo plano para remover o resíduo PASC. Os filtrados foram analisados para reduzir os açúcares pelo teste de hidrazida de ácido p-hidroxibenzóico (PHBAH) como descrito no exemplo 26.

Figura 17 mostra a conversão relativa de PASC após 45 horas de incubação com as endoglucanases de basidiomicete CBS 494.95 e basidiomicete CBS 495.95 (0,5 mg de proteína por g de PASC) como uma função de temperatura. A conversão relativa é mostrada como uma percentagem da conversão obtida após 45 horas de incubação com basidiomicete CBS 495.95 a 50°C . Perfis de temperatura obtidos em duas outras cargas de proteína, 0,056 e 0,167 mg de proteína por g de PASC, tem formas similares. Para ambas as endoglucanases, a temperatura ótima para hidrólise a longo prazo de PASC foi de 50 °C.

Exemplo 31: Caracterização de endoglucanases em vários substratos de polissacarídeo a 50°C

As endoglucanases purificadas de *Myceliophthora thermophila* CBS 117.65, basidiomicete CBS 494.95, e basidiomicete CBS 495.95 foram avaliadas na hidrólise de vários polissacarídeos a pH 5,0 (50 mM de tampão de acetato de sódio) e 50°C . Os resultados foram comparados com os para endoglucanase de *Trichoderma reesei* Cel7B recombinante (EGI). Endoglucanase *Trichoderma reesei* Cel7B recombinante (EGI) produzida por *Aspergillus oryzae* pode ser preparada de acordo com Takashima *et al.*, 1998, *Journal of Bacteriology* 65: 163 - 171.

Os polissacarídeos incluíam forragem de milho pré-tratada (PCS), celulose intumescida com ácido fosfórico (PASC), carboximetilcelulose (CMC), celulose bacteriana (BC), Avicel, xilano, xiloglucano, arabinoxilano, manano e galactomanano. Todos os substratos foram usados a 5 mg/ml, com a exceção de celulose bacteriana, que foi usada a 0,9 mg/ml.

Reações com um volume inicial de 1 ml foram realizadas durante 24 horas com agitação intermitente em placas Eppendorf 96 DeepWell (1,2 ml, VWR Scientific, West Chester, PA, USA) tampadas com esteiras Eppendorf 96 DeepWell (VWR Scientific, West Chester, PA, USA). Salvo indicado em contrário, as enzimas foram carregadas a 5 mg de proteína por g de sólidos.

Após 24 horas, alíquotas de 20 µl foram removidas das reações de hidrólise usando um pipetador de 8 canais, e adicionadas a 180 µl de Na₂CO₃ 102 mM - de NaHCO₃) 58 mM em uma placa de filtração MultiScreen HV de 96 cavidades (Millipore, Bedford, MA, USA) para terminar a hidrólise. As amostras foram filtradas a vácuo em uma microplaca de fundo plano. Após a diluição apropriada, os filtrados foram analisados para açúcares redutores usando o teste de hidrazida de ácido p-hidroxibenzóico (PHBAH) como descrito no exemplo 26.

Tabela 3 mostra a conversão relativa de vários polissacarídeos

pelas endoglucanases após 24 horas de incubação. A conversão relativa foi calculada como uma percentagem de conversão obtida após 24 horas de hidrólise de 1,4- β -D-manano por endoglucanase Cel5A de basidiomicete CBS 495.95. Endoglucanases de família 5 glicosídeo hidrolase (GH) tem atividade relativamente elevada em manano e galactomanano, mas atividade baixa em xilano, xiloglucano e arabinoxilano. Em contraste, Cel7B de *Trichoderma reesei* tem atividade relativamente elevada em xilano, xiloglucano e arabinoxilano, mas atividade baixa em manano e galactomanano. Os endoglucanases GH5 mostraram melhor hidrólise de PASC (celulose amorfa não substituída insolúvel) do que CMC (derivado de celulose substituído solúvel). As endoglucanases GH5 tem atividade baixa em substratos insolúveis com grau elevado de cristalinidade: celulose bacteriana, Avicel, e PCS.

Tabela 3. Conversão relativa de vários substratos de polissacarídeo (5 mg/ml) por endoglucanases (5 mg de proteína por g de sólidos); pH 5,0, 50°C , 24 horas

Substrato	Cel5A de <i>Myceliophthora thermophila</i>	Cel5B de basidiomicete CBS 494.95	Cel5A de basidiomicete CBS 495.95	Cel7B de <i>Trichoderma reesei</i>
Forragem de milho pré-tratado (PCS)	2	6	6	10
Celulose intumescida com ácido fosfórico (PASC)**	10	40	70	38
Carboximetilcelulose (CMC)**	12	12	12	14
Celulose bacteriana (BC)*	1	5	5	5
Avicel (celulose microcristalina)	1	2	3	5
xilano de bétula	0	3	2	51
xiloglucano de tamarindo	0	1	1	87
Arabinoxilano de trigo	6	7	8	81
1,4- β -manano	70	73	100	2
Galactomanano caroba	52	54	62	3

* Concentração inicial de celulose bacteriana foi de 0,9 mg/ml

** Todas as endoglucanases foram usados a 0,25 mg de proteína por g de

sólidos para hidrólise de PASC e CMC

Exemplo 32: Hidrólise de beta-glucano solúvel de cevada por endoglucanases a 60°C

A atividade das endoglucanases de *Myceliophthora thermophila* CBS 117.65, basidiomicete CBS 494.95, e basidiomicete CBS 495.95 em beta-glucano solúvel de cevada (viscosidade do meio de cultura, 230 kDa, Megazyme International Ireland Ltd., Bray, Irlanda) foi determinada a pH 5,5 (50 mM de acetato de sódio com 0,02% de azida sódica) e 60°C . Os resultados foram comparados com os para endoglucanase Cel7B de *Trichoderma reesei* (EGI). Endoglucanase Cel7B de *Trichoderma reesei* recombinante (EGI) pode ser preparada como descrito no exemplo 31.

A concentração inicial de beta-glucano nas reações de hidrólise foi de 1,0% (p/v). Um ml de reações foi ciclado sem agitação em placas Eppendorf 96 DeepWell (1,2 ml, VWR Scientific, West Chester, PA, USA). As enzimas foram usadas em três cargas de proteína, 0,05, 0,1, e 0,2 mg por g de glucano. Em um controle de tampão, as endoglucanases foram substituídas com 50 mM de acetato de sódio pH 5,5 contendo 0,02% de azida de sódio.

As alíquotas foram removidas das reações de hidrólise em 2 horas e 24 horas, diluídas com água deionizada, e analisadas para açúcares redutores usando o teste de hidrazida de ácido p-hidroxibenzóico (PHBAH) como descrito no exemplo 28. A conversão relativa de beta-glucano como uma função de carga de proteína em dois tempos de incubação, 2 horas e 24 horas, como mostrado nas figuras 18 e 19, respectivamente. A conversão relativa é mostrada como uma porcentagem de conversão obtida após 24 horas de hidrólise de beta-glucano por endoglucanase Cel5A de *Myceliophthora thermophila* CBS 117.65 (0,2 mg de proteína por g de glucano).

As endoglucanases de basidiomicete CBS 494,95 e

basidiomicete CBS 495.95 mostraram desempenho similar em hidrólise de beta-glucano, e não produziram um aumento adicional na redução da concentração de açúcar após 2 horas de hidrólise. Em contraste, as endoglucanases de *Myceliophthora thermophila* e *Trichoderma reesei* continuaram a produzir novos grupos terminais redutores além do tempo de incubação de 2 horas. A endoglucanase de *Myceliophthora thermophila* mostrou melhor desempenho em hidrólise de beta-glucano do que a endoglucanase Cel5B de basidiomicete CBS 494.95 e endoglucanase Cel5A de basidiomicete CBS 495.95.

10 Depósito de Material Biológico

Os seguintes materiais biológicos foram depositados sob os termos do Tratado de Budapeste junto ao Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, e receberam os seguintes números de acesso:

Depósito	Número de acesso	Dados de depósito
<i>E. coli</i> Cepa PBCel5C	NRRL B-30900N	23 de fevereiro de 2006
<i>E. coli</i> Cepa pCIC161	NRRL B-30902	23 de fevereiro de 2006
<i>E. coli</i> Cepa pCIC453	NRRL B-30903	23 de fevereiro de 2006
<i>E. coli</i> Cepa pCIC486	NRRL B-30904	23 de fevereiro de 2006

As cepas foram depositadas sob condições que asseguram que o acesso às culturas será disponível durante a pendência deste pedido de patente para uma pessoa determinada pelo Comissariado de Patentes e Marcas Registradas a ser intitulado sob 37 C.F.R. §1,14 e 35 U.S.C. §122. Os depósitos representam culturas substancialmente puras das cepas depositadas. Os depósitos são disponíveis como requerido por leis de patentes no estrangeiro em países em que as contrapartes do presente pedido, ou sua descendência, são depositadas. No entanto, será entendido que a disponibilidade de um depósito não constitui uma licença para praticar a invenção em questão em derrogação aos direitos de patentes concedidos pela ação governamental.

A invenção descrita e reivindicada aqui não deve ser limitada em escopo aos aspectos específicos aqui descritos, porque estes aspectos são destinados como ilustrações dos vários aspectos da invenção. Quaisquer aspectos equivalentes são destinados a estarem dentro do escopo desta invenção. De fato, várias modificações da invenção além das mostradas e aqui descritas se tornarão evidentes aos versados na arte a partir da descrição acima. Estas modificações também se destinam a estar dentro do escopo das reivindicações anexas. No caso de conflito, a presente descrição, incluindo as definições, irá gerenciar o controle.

Várias referências são citadas aqui, cujas descrições são incorporadas por referência em suas totalidades.

REIVINDICAÇÕES

1. Método para degradar ou converter um material celulósico, caracterizado pelo fato de compreender: tratar o material celulósico com uma composição compreendendo uma quantidade eficaz de um polipeptídeo isolado tendo atividade endoglucanase, o referido polipeptídeo compreendendo o polipeptídeo maduro de SEQ ID NO:4.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que o polipeptídeo maduro são os aminoácidos 17 a 389 de SEQ ID NO: 4.

3. Método de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de ainda ter atividade enzimática para um ou mais substratos selecionados dentre o grupo consistindo de xilano, xiloglucano, arabinoxilano, 1,4-beta-D-mannano, e galactomanano.

4. Método de acordo com a reivindicação 1, caracterizado pelo fato de que a composição ainda compreende uma quantidade eficaz de uma endo-1,4-beta-glucanase, exo-1,4-beta-D-glucanase, e/ou beta-D- glucosidase.

5. Método de acordo com a reivindicação 1 ou 4, caracterizado pelo fato de que a composição ainda compreende uma quantidade eficaz de uma ou mais enzimas selecionadas dentre o grupo consistindo em uma hemicelulase, esterase, protease, lacase, e peroxidase.

6. Método de acordo com qualquer uma das reivindicações 1-5, caracterizado pelo fato de que o método é um processo de pré-tratamento: uma etapa em um processo de sacarificação e fermentação simultâneas (SSF); ou uma etapa em uma hidrólise de híbrido e processo de fermentação (HHF).

7. Método para produzir uma substância, caracterizado pelo fato de compreender: (a) sacarificar um material celulósico com uma composição compreendendo uma quantidade eficaz do polipeptídeo isolado tendo atividade endoglucanase, o referido polipeptídeo compreendendo o polipeptídeo maduro de SEQ ID NO:4, (b) fermentar o material celulósico

sacarificado de etapa (a) com um ou mais microorganismos de fermentação; e (c) recuperar a substância da fermentação, em que a substância é um álcool, um ácido orgânico, uma cetona, um aminoácido ou um gás.

5 8. Método de acordo com a reivindicação 1, caracterizado pelo fato de que o polipeptídeo maduro são os aminoácidos 17 a 389 de SEQ ID NO: 4.

10 9. Método de acordo com a reivindicação 7 ou 8, caracterizado pelo fato do polipeptídeo ainda ter atividade enzimática para um ou mais substratos selecionados dentre o grupo consistindo de xilano, xiloglucano, arabinoxilano, 1,4-beta-D-mannano, e galactomanano.

10 10. Método de acordo com a reivindicação 7, caracterizado pelo fato de que a composição ainda compreende uma quantidade eficaz de uma endo-1,4-beta-glucanase, exo-1,4-beta-D-glucanase, e/ou beta-D-glucosidase.

15 11. Método de acordo com a reivindicação 7 ou 10, caracterizado pelo fato de que a composição ainda compreende uma quantidade eficaz de uma ou mais enzimas selecionadas dentre o grupo consistindo em uma hemicelulase, esterase, protease, lacase, e peroxidase.

20 12. Método de acordo com as reivindicações 7-11, caracterizado pelo fato de que nas etapas (a) e (b) são realizadas simultaneamente, em uma sacarificação e fermentação simultâneas.

25 13. Célula hospedeira microbiana recombinante, caracterizada pelo fato de compreender a construção de ácido nucleico compreendendo o polinucleotídeo apresentado em SEQ ID NO:3, em que o referido polipeptídeo codifica um polipeptídeo tendo atividade endoglucanase, o referido polipeptídeo compreendendo o polipeptídeo maduro de SEQ ID NO:4.

14. Célula hospedeira microbiana recombinante de acordo com a reivindicação 13, caracterizado pelo fato de que o polipeptídeo maduro

são os aminoácidos 17 a 389 de SEQ ID NO: 4.

15 15. Método de acordo com a reivindicação 13 ou 14, caracterizado pelo fato do polipéptido ainda ter atividade enzimática para um ou mais substratos selecionados dentre o grupo consistindo de xilano, xiloglucano, arabinosilano. 1,4-beta-D-mannano, e galactomanano.

10 16. Método para produzir um polipeptídeo tendo atividade endoglucanase, o referido polipeptídeo compreendendo o polipeptídeo maduro de SEQ ID NO:4, o método, caracterizado pelo fato de compreender: (a) cultivar a célula hospedeira microbiana recombinante como definida na reivindicação 13 sob condições que conduzem a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

17. Método de acordo com a reivindicação 16, caracterizado pelo fato de que o polipeptídeo maduro são os aminoácidos 17 a 389 de SEQ ID NO: 4.

15 18. Método de acordo com a reivindicação 16 ou 17, caracterizado pelo fato do polipeptídeo ainda ter atividade enzimática para um ou mais substratos selecionados dentre o grupo consistindo em xilano, xiloglucano, arabinosilano. 1,4-beta-D-mannano, e galactomanano.

20 19. Método para produzir um polipeptídeo tendo atividade endoglucanase, o referido polipeptídeo compreendendo o polipeptídeo maduro de SEQ ID NO:4, o método, caracterizado pelo fato de compreender: (a) cultivar uma planta ou célula de planta transgênica compreendendo um polinucleotídeo de SEQ ID NO:3, que codifica um polinucleotídeos tendo atividade endoglucanase, sob condições que conduzem a produção do polipeptídeo; e (b) recuperar o polipeptídeo.

25

20. Método de acordo com a reivindicação 19, caracterizado pelo fato de que o polipeptídeo maduro são os aminoácidos 17 a 389 de SEQ ID NO: 4.

21. Método de acordo com a reivindicação 19 ou 20,

caracterizado pelo fato do polipeptídeo ainda ter atividade enzimática para um ou mais substratos selecionados dentre o grupo consistindo em xilano, xiloglucano, arabinoxilano, 1,4-beta-D-mannano, e galactomanano.

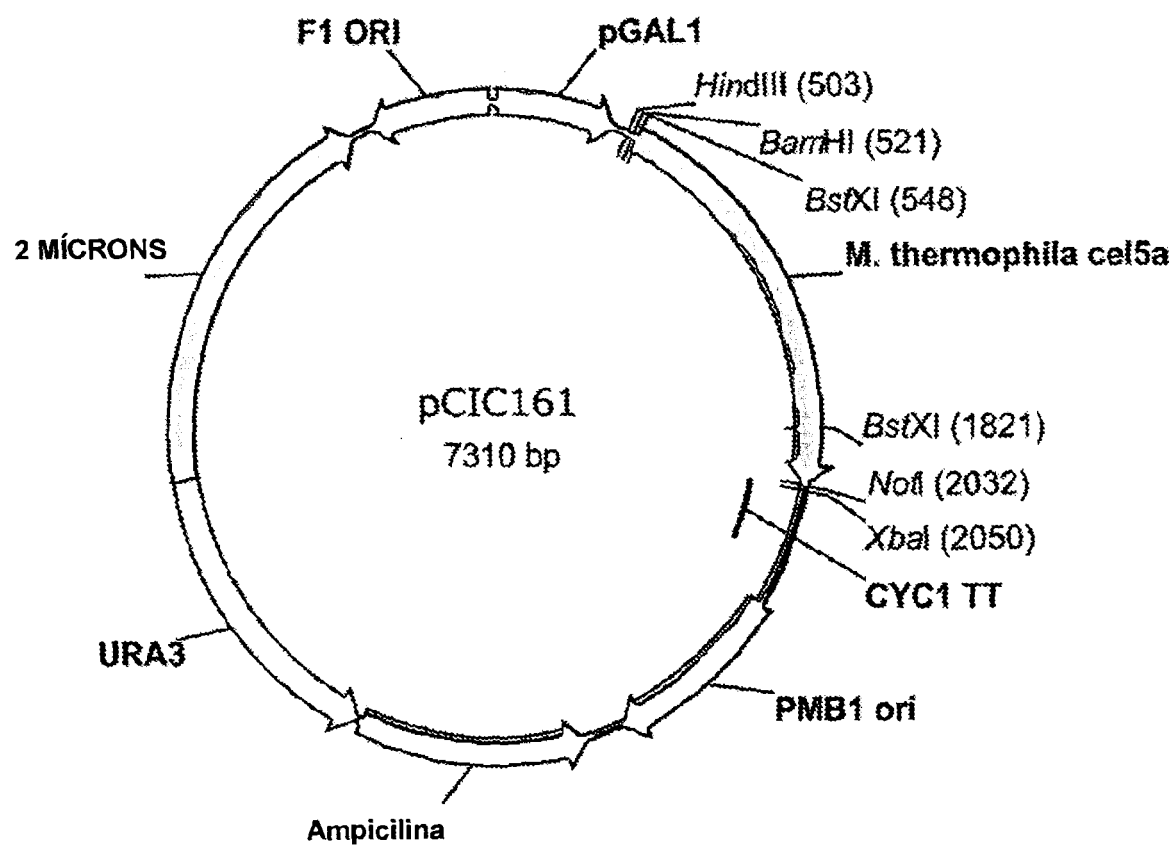


Fig. 1

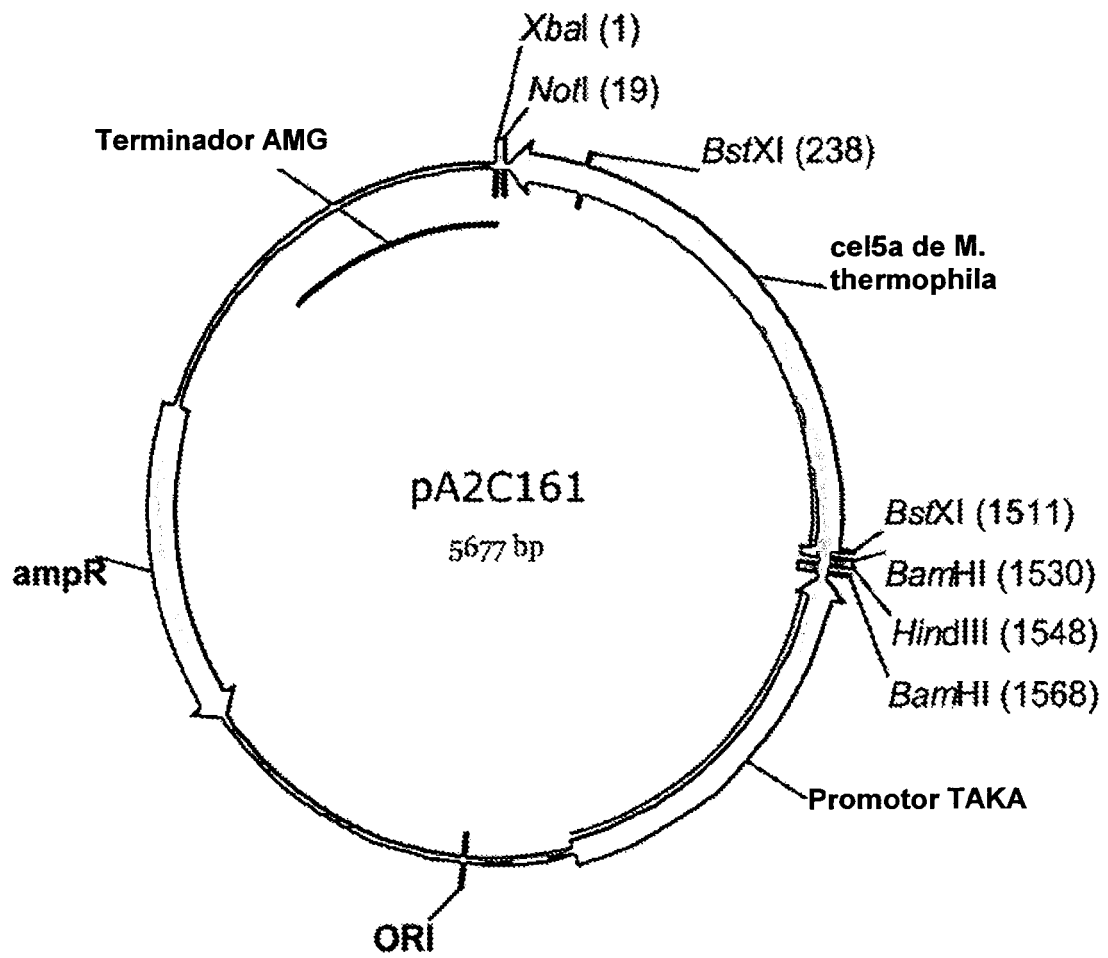


Fig. 3

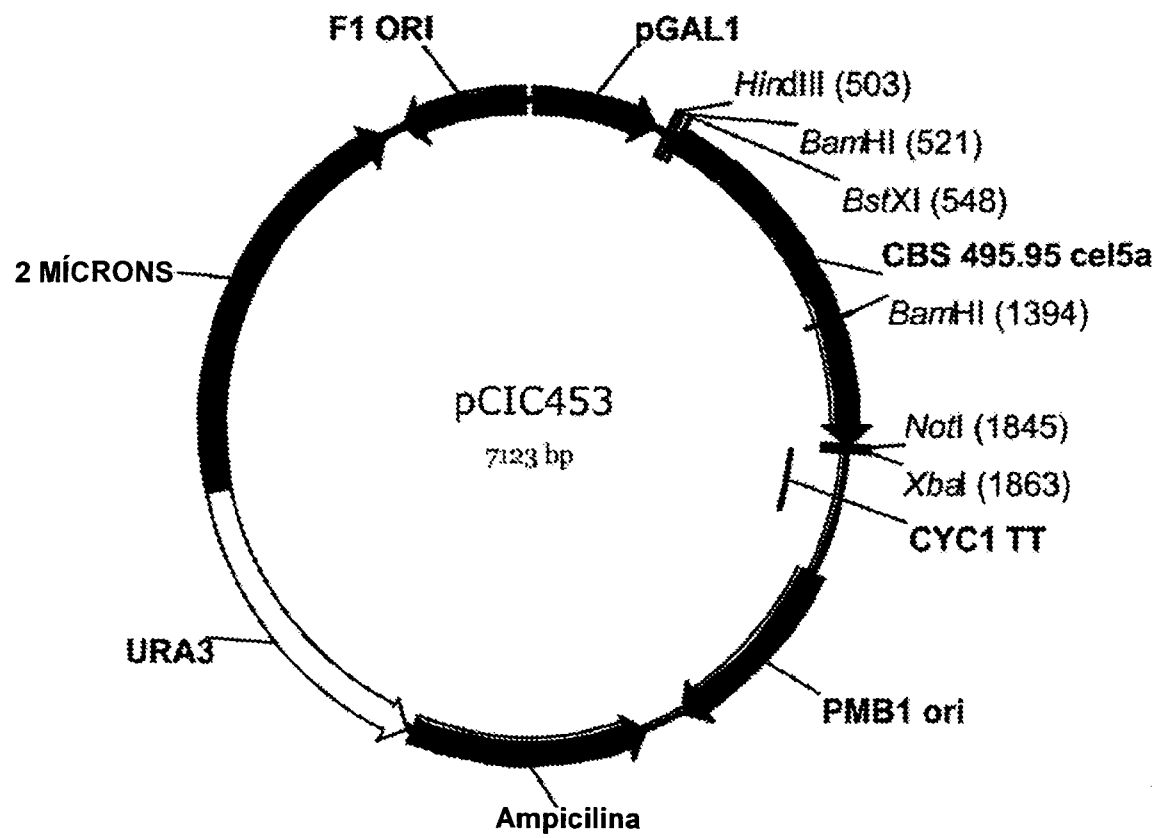


Fig. 4

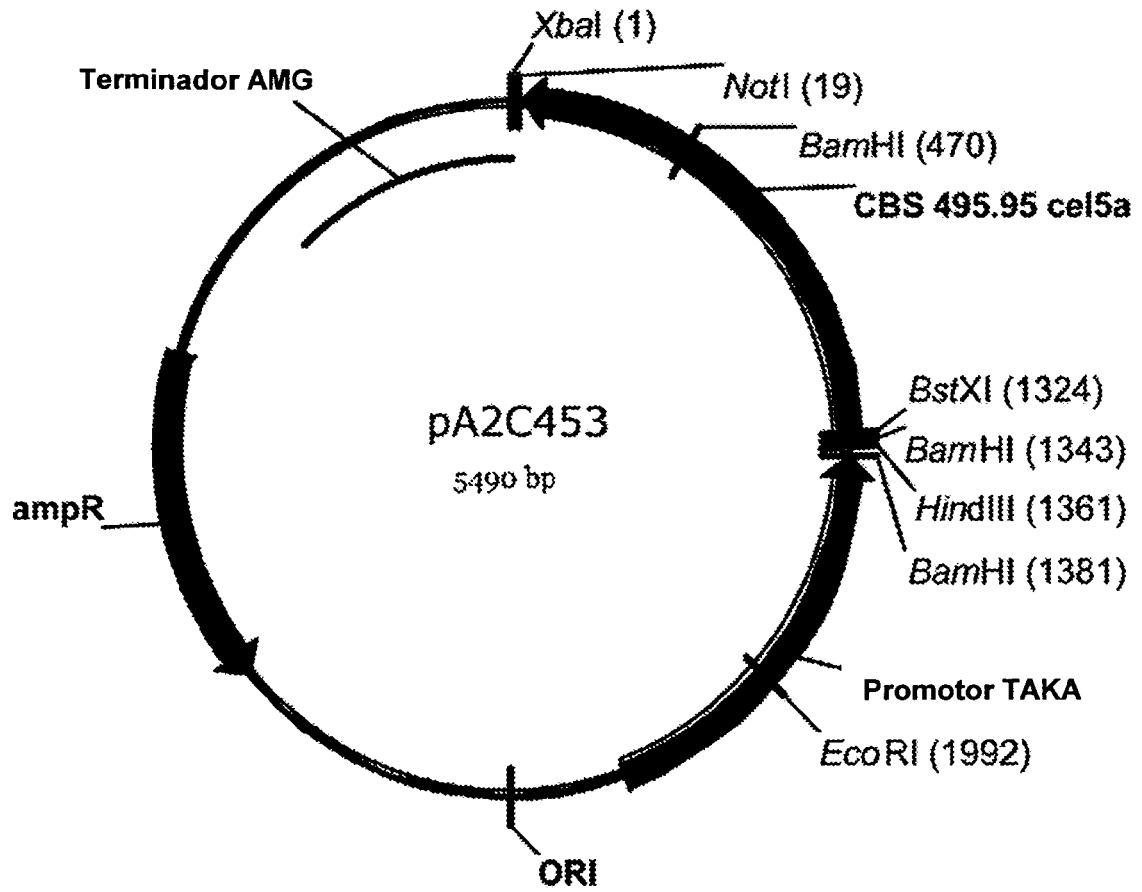


Fig. 5

GGATCCACTTAGTAACGGCCGCCAGTGTGCTGGAAAGCATGAAGTCTCTCTTCTGTCACTTGTAGCGACCGTCGCGCTCAGCTCGC
M K S L F L S L V A T V A L S S
CAGTATTCTCTGTGCGCAGTCTGGGGCAATGCGGCGGCATTGGCTTCAGCGGAAGCACCGTCTGTGATGCAGGCGCCGGCTGTGTGA
P V F S V A V W G Q C G G I G F S G S T V C D A G A G C V
AGCTCAACGACTATTACTCTCAATGCCAACCCGGCGCTCCCACTGCTACATCCGCGGCGCCAAGTAGCAACGCACCGTCCGGCACTT
K L N D Y Y S Q C Q P G A P T A T S A A P S S N A P S G T
CGACGGCCTCGGCCCCCTCCTCCAGCCTTTGCTCTGGCAGCCGCACGCCGTTCCAGTTCTTCGGTGTCAACGAATCCGGCGCGGAGT
S T A S A P S S S L C S G S R T P F Q F F G V N E S G A E
TCGGCAACCTGAACATCCCCGGTGTTCGGGCACCGACTACACCTGGCCGTCGCCATCCAGCATTGACTTCTTCATGGGCAAGGGAA
F G N L N I P G V L G T D Y T W P S P S S I D F F M G K G
TGAATACCTTCCGTATTCCGTTCCCTCATGGAGCGTCTTGTCCTCCCTGCCACTGGCATCACAGGACCTCTCGACCAGACGTACTTGG
M N T F R I P F L M E R L V P P A T G I T G P L D Q T Y L
GCGGCCTGCAGACGATTGTCAACTACATCACCGGCAAAGGCGGCTTTGCTCTCATTGACCCGCACAACCTTTATGATCTACAATGGCC
G G L Q T I V N Y I T G K G G F A L I D P H N F M I Y N G
AGACGATCTCCAGTACCAGCGACTTCCAGAAGTTCTGGCAGAACCTCGCAGGAGTGTTTAAATCGAACAGTCACGTCATCTTCGATG
Q T I S S T S D F Q K F W Q N L A G V F K S N S H V I F D
TTATGAACGAGCCTCACGATATTCCCGCCAGACCGTGTTCCAACTGAACCAAGCCGCTGTCAATGGCATCCGTGCGAGCGGTGCGA
V M N E P H D I P A Q T V F Q L N Q A A V N G I R A S G A
CGTCGCAGCTCATTCTGGTCGAGGGCACAAGCTGGACTGGAGCTGGACCTGGACGACCTCTGGCAACAGCGATGCATTCCGGTGCCA
T S Q L I L V E G T S W T G A W T W T T S G N S D A F G A
TTAAGGATCCCAACAACAACGTCGCGATCCAGATGCATCAGTACCTGGATAGCGATGGCTCTGGCACTTCGCAGACCTGCGTGTCTC
I K D P N N N V A I Q M H Q Y L D S D G S G T S Q T C V S
CCACCATCGGTGCCGAGCGGTTGCAGGCTGCGACTCAATGGTTGAAGCAGAACAACCTCAAGGGCTTCCTGGGCGAGATCGGCGCCG
P T I G A E R L Q A A T Q W L K Q N N L K G F L G E I G A
GCTCTAACTCCGCTTGTCATCAGCGCTGTGCAGGGTGCGTTGTGTTTCGATGCAGCAATCTGGTGTGTGGCTCGGCGCTCTCTGGTGGG
G S N S A C I S A V Q G A L C S M Q Q S G V W L G A L W W
CTGCGGGCCCGTGGTGGGGCGACTACTACAGTCCATCGAGCCGCCCTCTGGCCCGGCGGTGTCCGCGATCCTCCCGCAGGCCCCTGC
A A G P W W G D Y Y Q S I E P P S G P A V S A I L P Q A L
TGCCGTTTCGCGTAA
L P F A .

Fig. 6

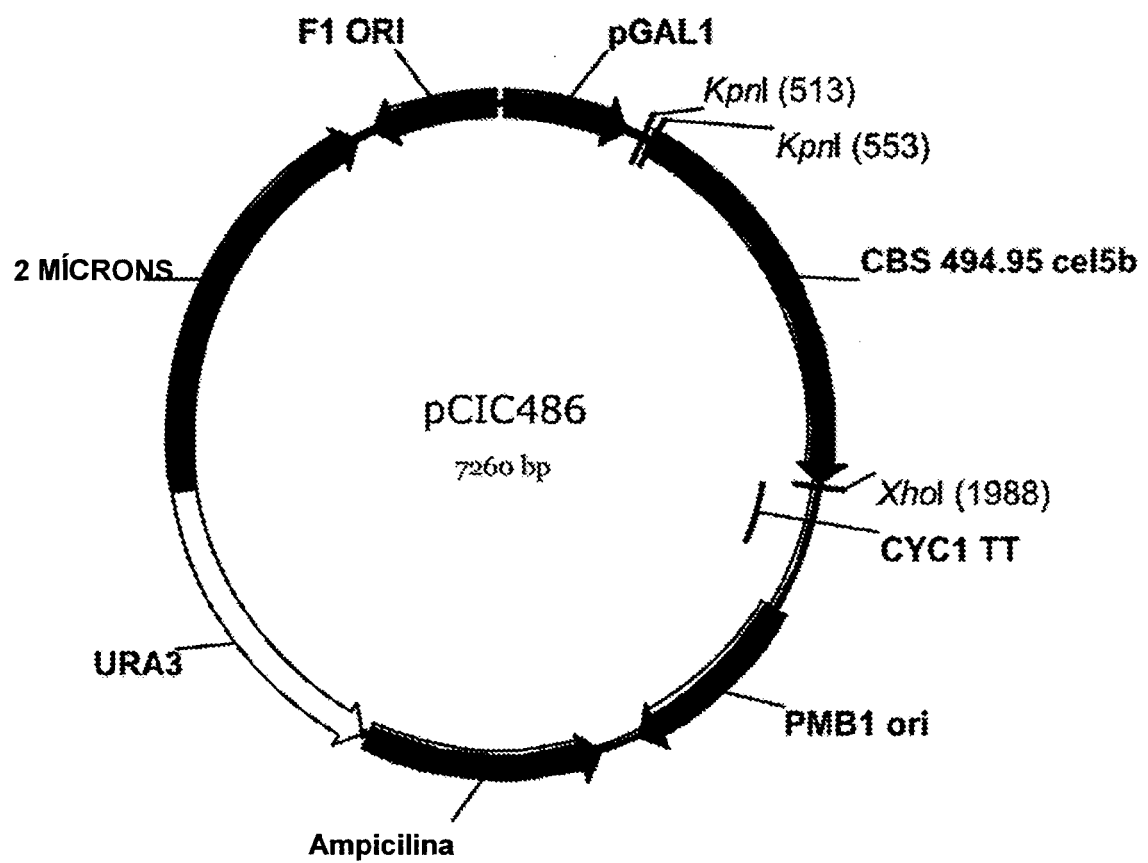


Fig. 7

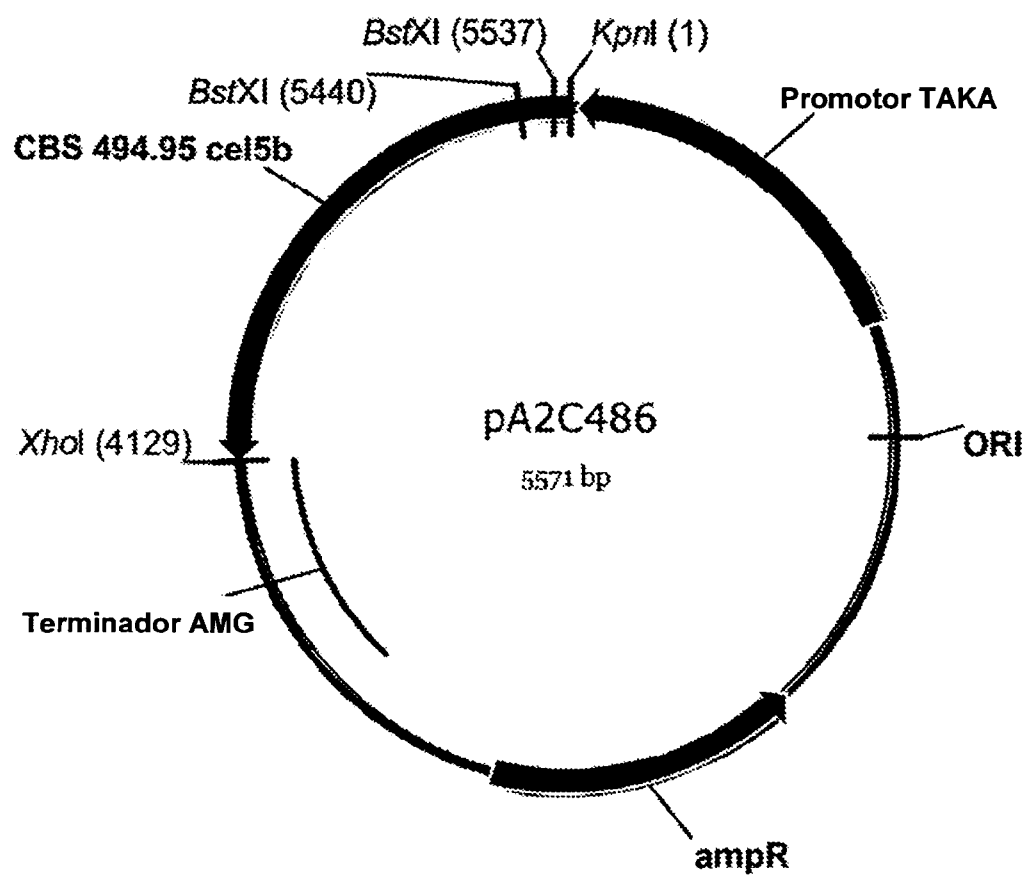


Fig. 8

GGAAAGCGTCAGTATGGTGAAATTTGCGCTTGTGGCAACTGTCGGCGCAATCTTGAGCGCTTCTGCGGCCAATGCGGGCTTCTATCTA
M V K F A L V A T V G A I L S A S A A N A A S I Y
CCAGCAATGTGGAGGCATTGGATGGTCTGGGTCCACTGTTTGCGACGCCGGTCTCGCTTGCCTTATCCTCAATGCGTACTACTTTCA
Q Q C G G I G W S G S T V C D A G L A C V I L N A Y Y F Q
GTGCTTGACGCCCCGCCGCGGGCCAGACAACGACGGGCTCGGGCGCACCGGGCGTCAACATCAACCTCTCACTCAACGGTCACTACGGG
C L T P A A G Q T T T G S G A P A S T S T S H S T V T T G
GAGCTCACACTCAACAACCGGGACGACGGCGACGAAAACAACCTACCACTCCGTCGACCACCACGACCCTACCCGCCATCTCTGTGTC
S S H S T T G T T A T K T T T T P S T T T T L P A I S V S
TGGTCGCGTCTGCTCTGGCTCCAGGACGAAGTTCAAGTTCTTCGGTGTGAATGAAAGCGGGCGCCGAATTCGGGAACACTGCTTGGCC
G R V C S G S R T K F K F F G V N E S G A E F G N T A W P
AGGGCAGCTCGGGAAGACTATACATGGCCTTCGCCTAGCAGCGTGGACTACTTCATGGGGGCTGGATTCAATACATTCCGTATCAC
G Q L G K D Y T W P S P S S V D Y F M G A G F N T F R I T
CTTCTTGATGGAGCGTATGAGCCCTCCGGCTACCGGACTCACTGGCCCATTCAACCAGACGTACCTGTCGGGCCCTCACCACCATTGT
F L M E R M S P P A T G L T G P F N Q T Y L S G L T T I V
CGACTACATCACGAACAAAGGAGGATACGCTCTTATTGACCCCCACAACCTTCATGCGTTACAACAACGGCATAATCAGCAGCACATC
D Y I T N K G G Y A L I D P H N F M R Y N N G I I S S T S
TGACTTCGCGACTTGGTGGAGCAATTTGGCCACTGTATTCAAATCCACGAAGAACGCCATCTTCGACATCCAGAACGAGCCGTACGG
D F A T W W S N L A T V F K S T K N A I F D I Q N E P Y G
AATCGATGCGCAGACCGTATACGAACCTGAATCAAGCTGCCATCAATTCGATCCGCGCCGCTGGCGCTACGTCACAGTTGATTCTGGT
I D A Q T V Y E L N Q A A I N S I R A A G A T S Q L I L V
TGAAGGAACGTCATACACTGGAGCTTGGACGTGGGTCTCGTCCGGAAACGGAGCTGCTTTTCGCGGCCGTTACGGATCCTTACAACAA
E G T S Y T G A W T W V S S G N G A A F A A V T D P Y N N
CACGGCAATTGAAATGCACCAATACCTCGACAGCGACGGTTCTGGGACAAACGAAGACTGTGTCTCCTCCACCATTGGGTGCGCAACG
T A I E M H Q Y L D S D G S G T N E D C V S S T I G S Q R
TCTCCAAGCTGCCACTGCGTGGCTGCAACAAACAGGACTCAAGGGATTCTCGGAGAGACGGGTGCTGGGTGGAATTCCCAGTGCAT
L Q A A T A W L Q Q T G L K G F L G E T G A G S N S Q C I
CGACGCCGTGTTTCGATGAACTTTGCTATATGCAACAGCAAGGCGGCTCCTGGATCGGTGCACTCTGGTGGGCTGCGGGTCCCTGGTG
D A V F D E L C Y M Q Q Q G G S W I G A L W W A A G P W W
GGGCACGTACATTTACTCGATTGAACCTCCGAGCGGTGCCGCTATCCAGAAGTCCTTCCTCAGGGTCTCGCTCCATTCTCTAG
G T Y I Y S I E P P S G A A I P E V L P Q G L A P F L .

Fig. 9

ATGAAATACCTCTACTCCTGGCAACCAGCGCTGCCCTGGCTCTTGCTGGGCCTCAAGGCTTCTCCAAGCGAGCCTCTTCTTTTGTGTATG
 M K Y P L L L A T S A A L A L A G P Q G F S K R A S S F V
 TGTAGAGCCCATAAGAATAAGAGCTGGAACCTGACCTGAATTTTTCAGGTTTGGTACCAGCGAGTCTGGGGCGGAGTTTGGCAATCAGAACATT
 W F G T S E S G A E F G N Q N I
 CCGGGTGTTTTGGGGACGGACTACATTTGGCCCCGACACTTCAGCCATTCAACGCTACGCAACGCGGGTATGAACATCTTTCGGGTGGCATT
 P G V L G T D Y I W P D T S A I Q T L R N A G M N I F R V A F
 TTGATGGAGCGATTGGTTCACCACATTGACCTCAACTCCGGATTTCGACGTATCTGCAAGACCTGAAGAGTGTACGTTATCCTCGTCAACAA
 L M E R L V P T T L T S T P D S T Y L Q D L K S V
 TGCAACCCGTGCTAAACGATCTCCAGACAGTTGACTACATCACGTGACCGGCGCATATGCTATTGTGGACCCCCATAACTTTGGACGATAGT
 L D Y I T S T G A Y A I V D P H N F G R Y
 AAGTTGAACCTTTCATCTAACCAATATTGAAAATCTCCTCACGGGTGTATTAGCTATGGCAATATAATCAACTCCACAAGTGACTTTGCTGCA
 Y G N I I N S T S D F A A
 TTTTGGACCACCGTCGCAAGCAGTTTGCATCGAATGACAAGGTCATCTTTGACACGAGTGAGCATCACTACTTCACTAACTTTGTTCCCCCA
 F W T T V A K Q F A S N D K V I F D T
 CTAAGTGACATCTGCAGACAACGAATTCAACACAGAGGACCAACGCTTGTACTGAACCTCAATCAAGCGGCCATTAATGCCATCCGAGCTGC
 N N E F N T E D Q T L V L N L N Q A A I N A I R A A
 CGGAGCCACCTCTCAGTATATCTTTGTGGAAGGGAATTCGTGGAGTGGTGCCTGGACGTGGACGTCCGTCAATACCAATCTCGTCAGCTTGAC
 G A T S Q Y I F V E G N S W S G A W T W T S V N T N L V S L T
 GGACCCCAACAACAAGATAGTCTACGAGATGCACCAAGTATCTTGACTCGGATGGATCTGGCACATCCGACACATGTGTCAGCTCGACCATCGG
 D P N N K I V Y E M H Q Y L D S D G S G T S D T C V S S T I G
 CCAGGAGCGTGTACAATCCGCGACCGAGTGGCTGAAAAGCAATGGCAAACCTTGATTCTTGGGCGAATTTCGCGGCGGTGCTAACTCGGTCTG
 Q E R V Q S A T E W L K S N G K L G F L G E F A G G A N S V C
 TCAGAGCGCCGTGACTGGAATGCTGGACTATATGCAAGCGAATAGTGACGTCTGGCTCGGCGCATCCTGGTGGGCGAGCAGGACCATGGTGGGG
 Q S A V T G M L D Y M Q A N S D V W L G A S W W A A G P W W G
 AACCTATATATATTTCGATTGAGCCGCCATCGGGGACTGCTTATTCTTACTATCTCAACATCTTGTCTGCCTACTTCCCCTCCAGCTCGGGGAG
 T Y I Y S I E P P S G T A Y S Y Y L N I L S A Y F P S S S G S

Fig. 10A

TTCCACAACCACAACCTACCTCCACTACCACCCGCTCTACATCGACAAGCACTACAGTATCCACCACAAAATCAACAAGCACCCTACAAGCGC
S T T T T T S T T T R S T S T S T T V S T T K S T S T T T S A
CACGAAATCGACAAGCACCACGACAAGCACCACCAGCACCGGGTCTACTGCTACTGCAACAGCATCTCACTGGGCACAGTGTGGCGGCATTGG
T K S T S T T T S T T S T G S T A T A T A S H W A Q C G G I G
CTGGACGGGGGCGACGACGTGTGCCAGCCCGTATACCTGCCAGGTCCAGAATGCGTATTATTTCGCAGTGTCTGTAATGCAGAAGTATCAGAAA
W T G A T T C A S P Y T C Q V Q N A Y Y S Q C L .

Fig. 10B

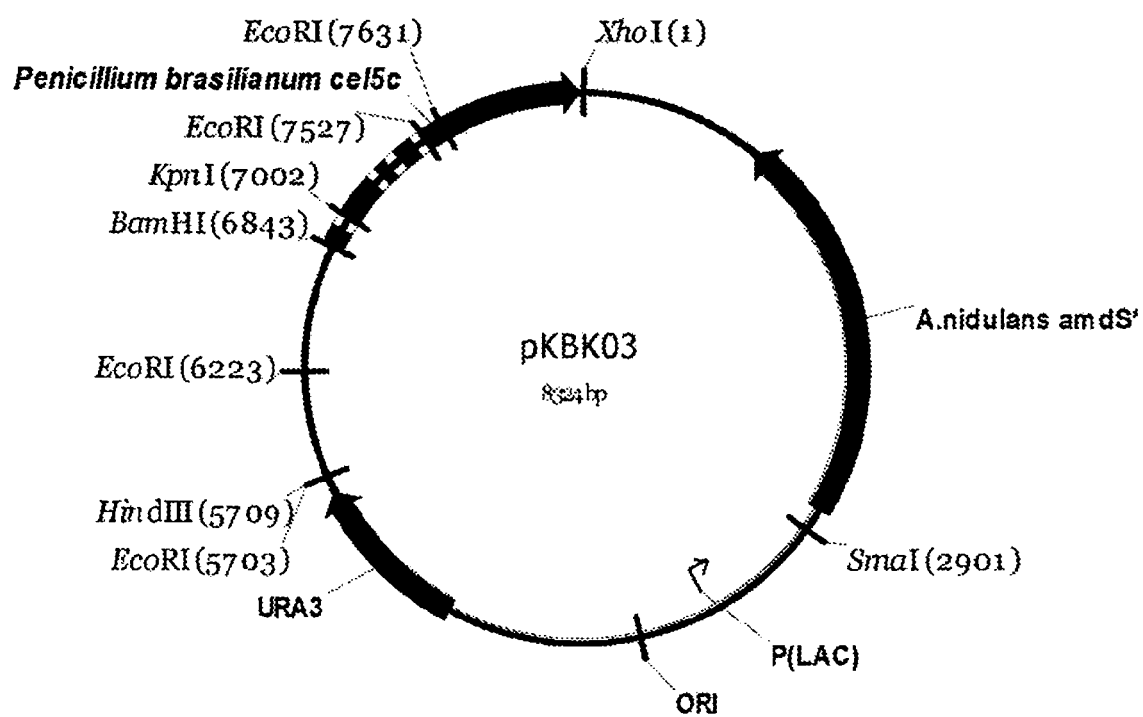


Fig. 11

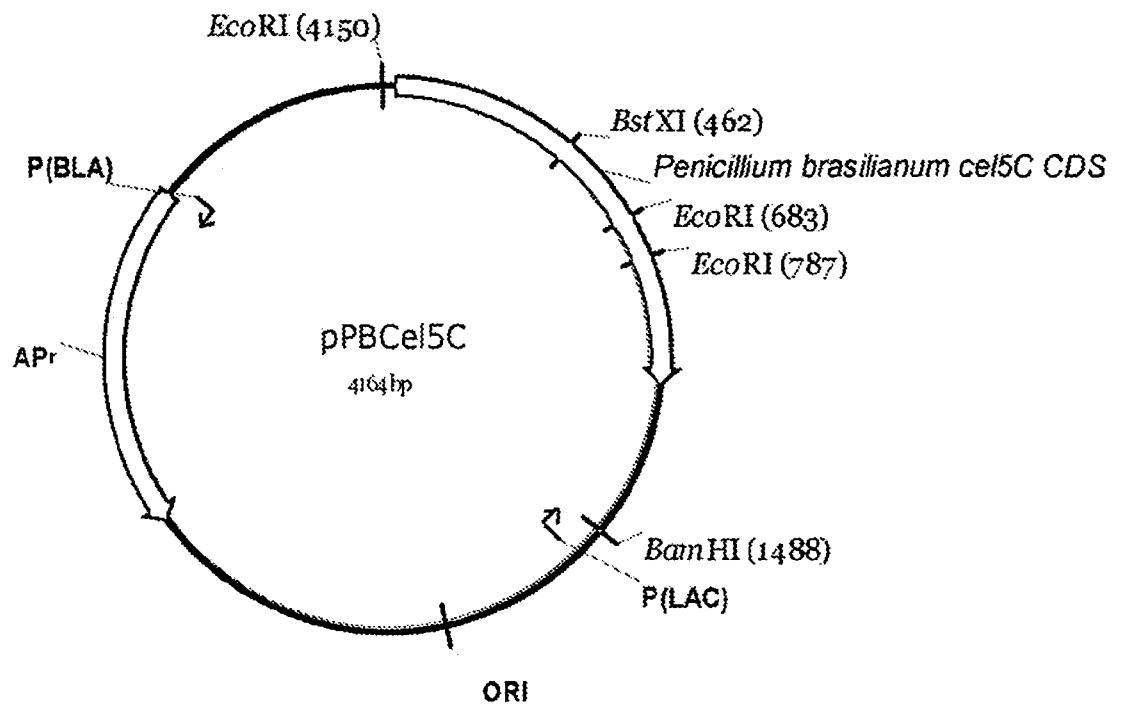


Fig. 12

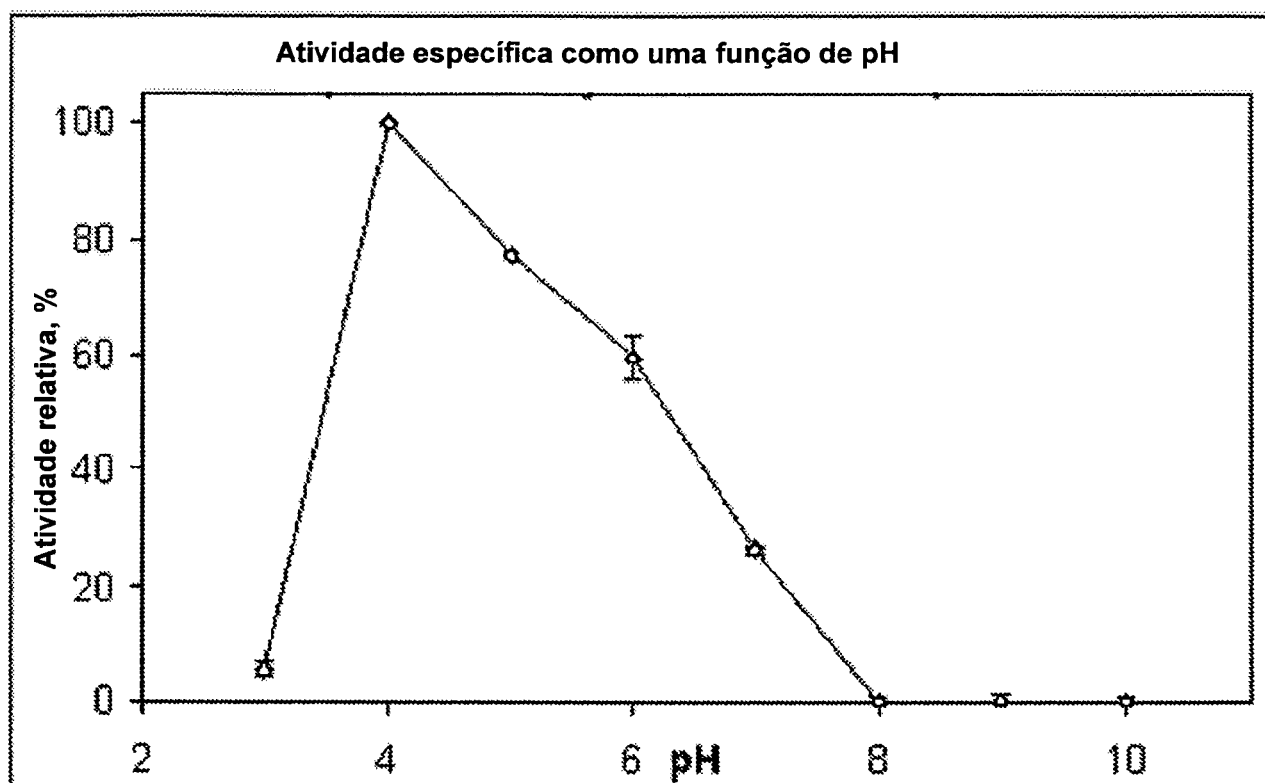


Fig.13

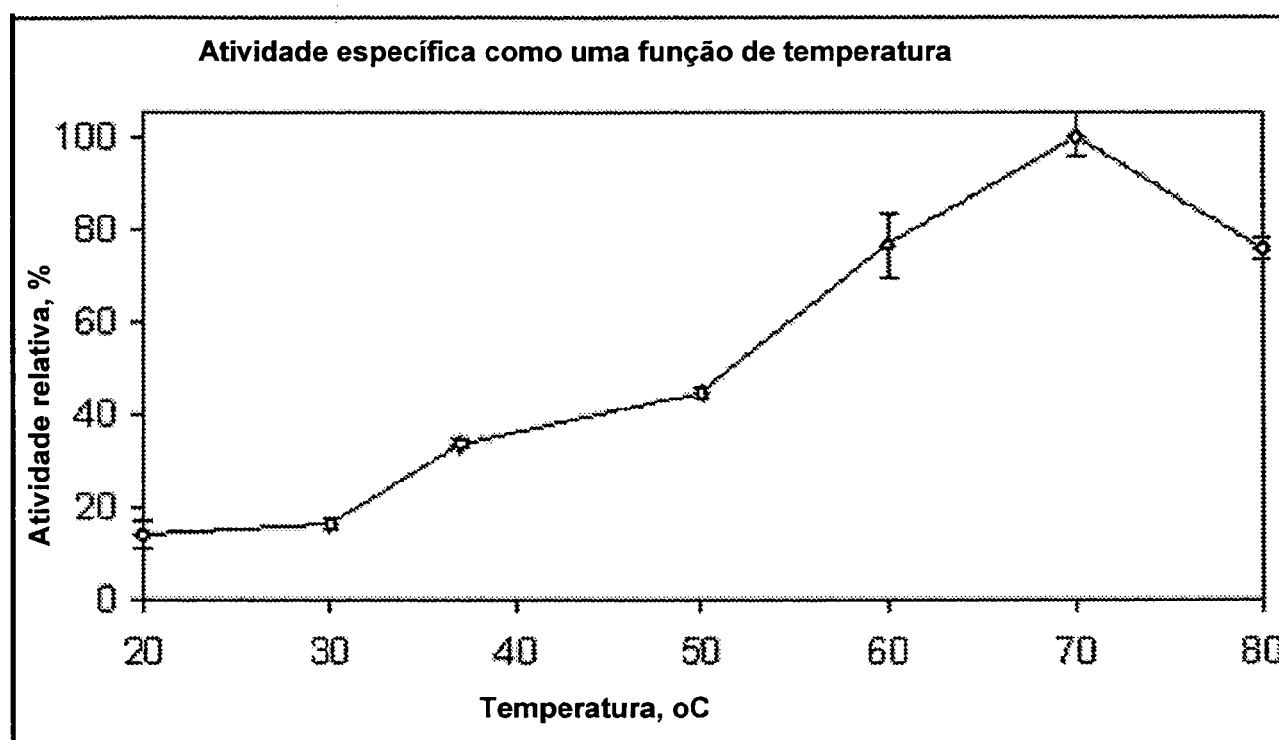


Fig.14

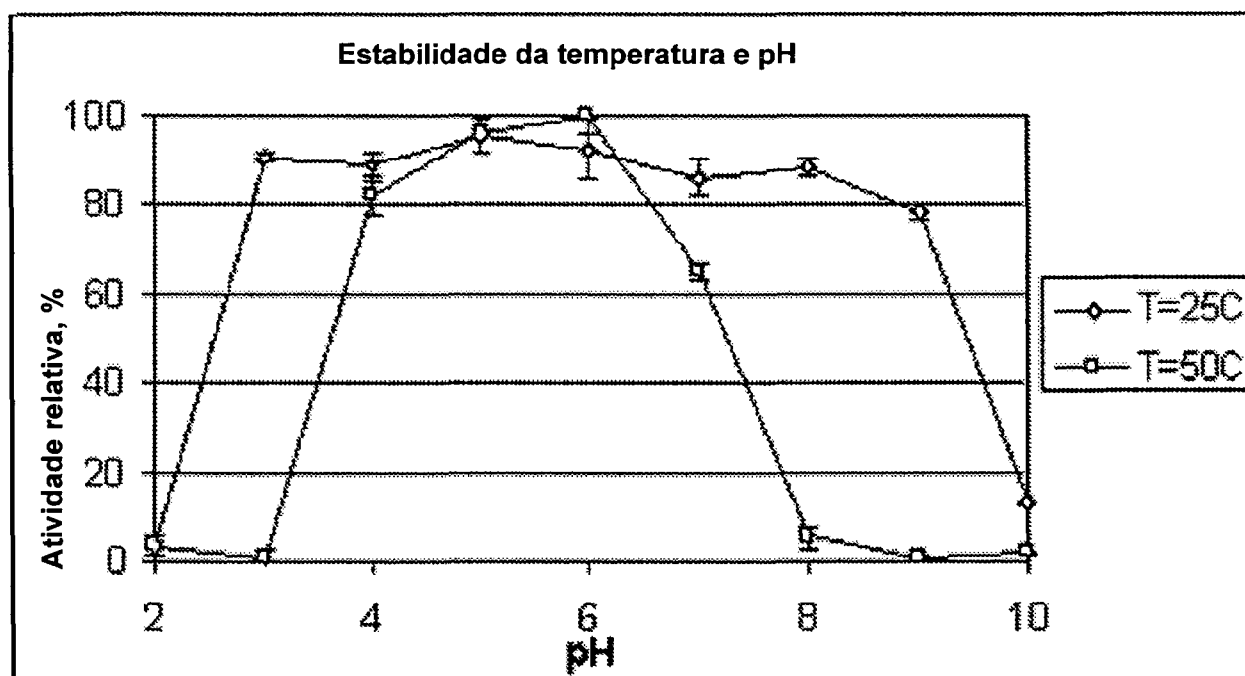


Fig.15

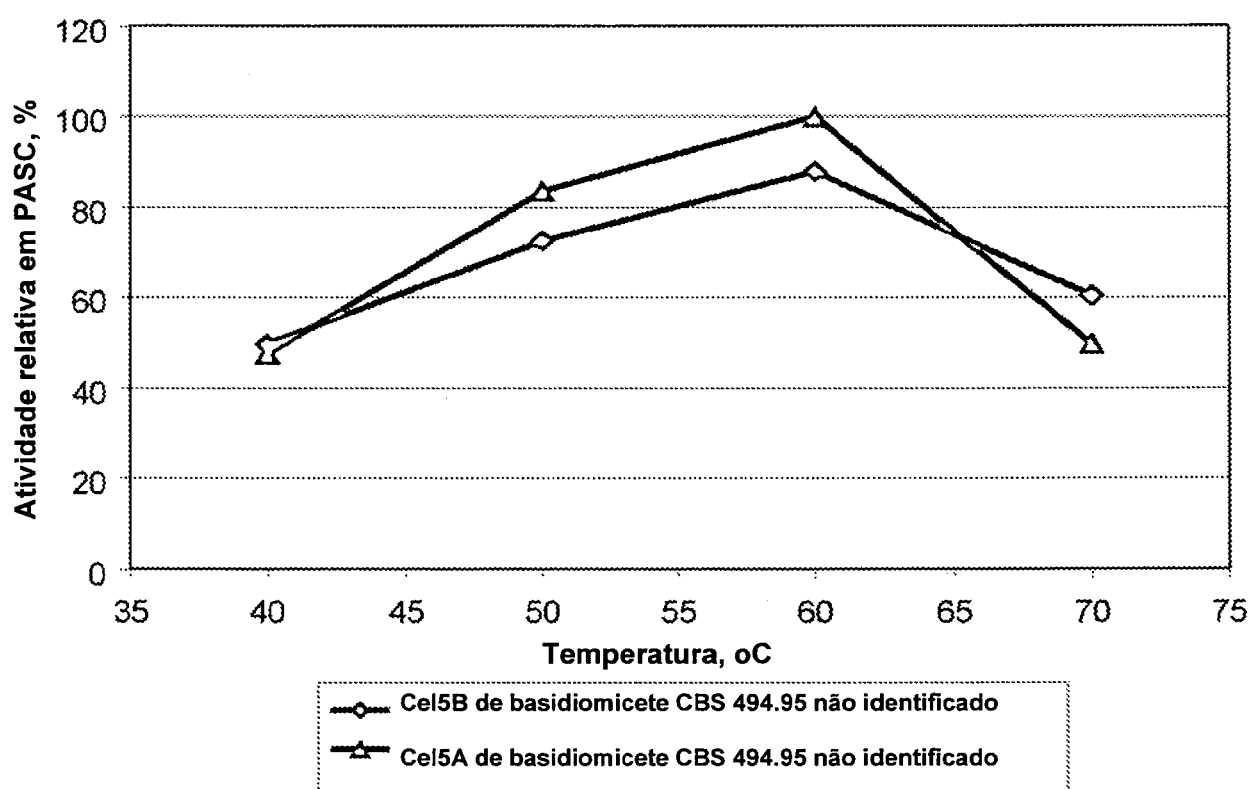


Fig.16

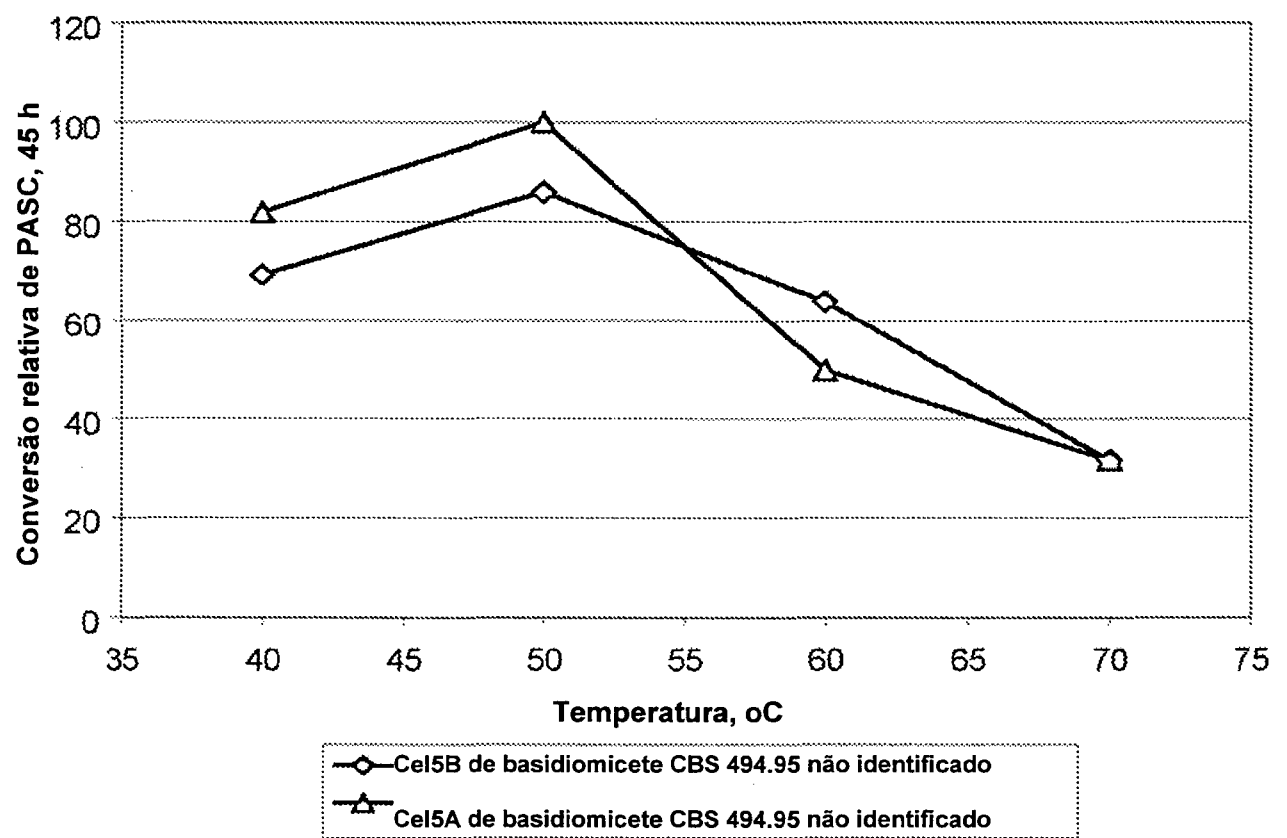


Fig.17

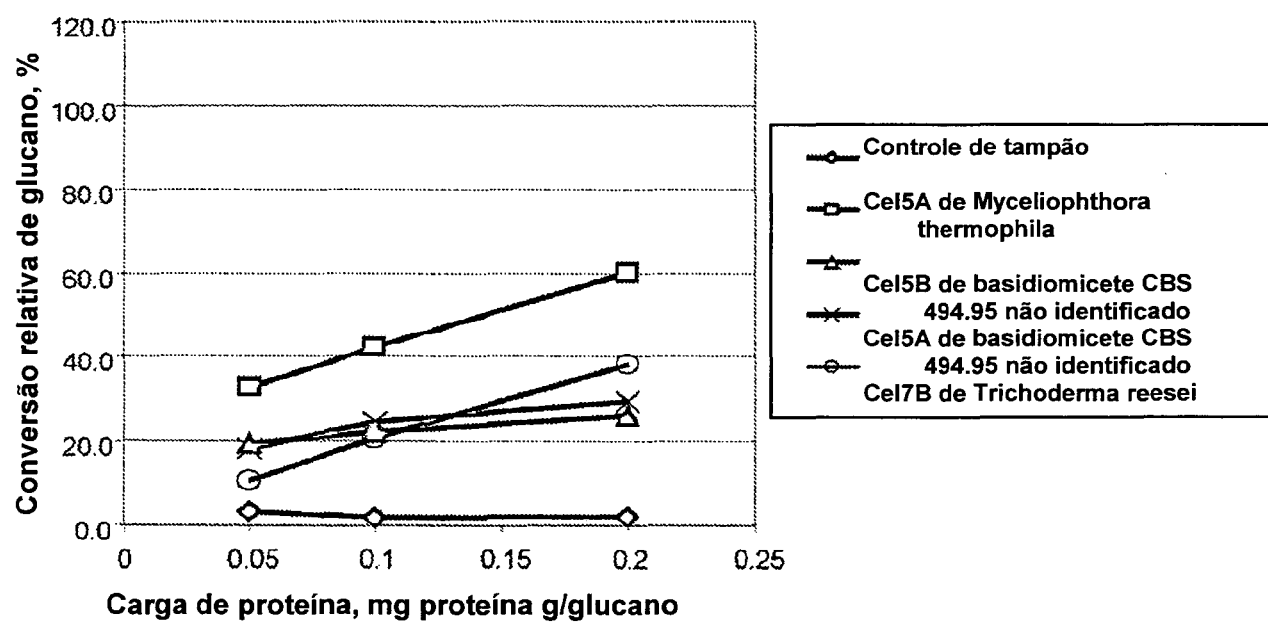


Fig.18

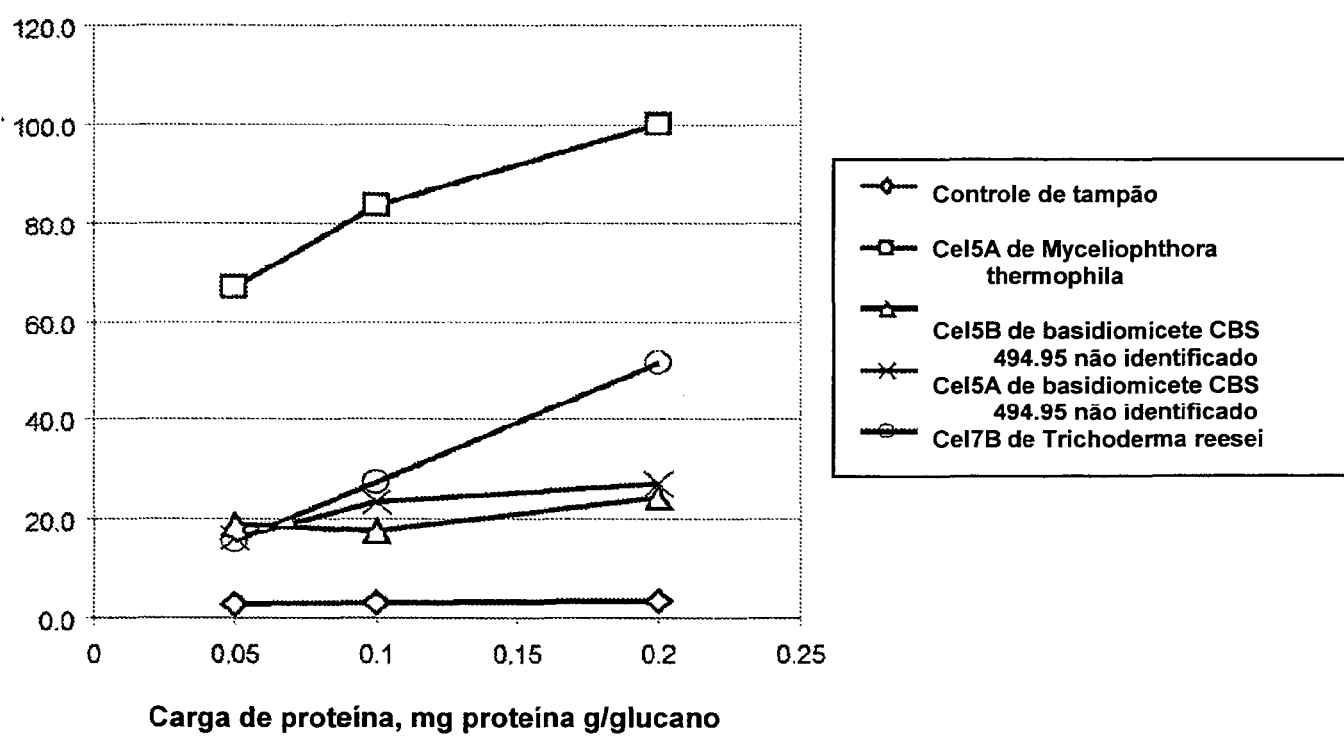


Fig.19