

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-504012

(P2004-504012A)

(43) 公表日 平成16年2月12日(2004.2.12)

(51) Int.Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	4 B O 2 4
C 1 2 N 1/15	C 1 2 N 1/15	4 B O 5 0
C 1 2 N 9/24	C 1 2 N 9/24	4 B O 6 3
C 1 2 P 19/14	C 1 2 P 19/14	4 B O 6 4
C 1 2 Q 1/68	C 1 2 Q 1/68	4 B O 6 5
		審査請求 未請求 予備審査請求 有 (全 92 頁)

(21) 出願番号	特願2001-577490 (P2001-577490)	(71) 出願人	502373684
(86) (22) 出願日	平成13年4月17日 (2001.4.17)		エマルファアープ, マーク・アーロン
(85) 翻訳文提出日	平成14年10月15日 (2002.10.15)		アメリカ合衆国フロリダ州33477ジュ
(86) 国際出願番号	PCT/NL2001/000301		ピター・スパイグラスコート193
(87) 国際公開番号	W02001/079507	(74) 代理人	100060782
(87) 国際公開日	平成13年10月25日 (2001.10.25)		弁理士 小田島 平吉
(31) 優先権主張番号	00201343.1	(72) 発明者	エマルファアープ, マーク・アーロン
(32) 優先日	平成12年4月13日 (2000.4.13)		アメリカ合衆国フロリダ州33477ジュ
(33) 優先権主張国	欧州特許庁 (EP)		ピター・スパイグラスコート193
		(72) 発明者	ブント, ベーター・ヤン
			オランダ・エヌエルー3994エツクス
			イ ホーテン・ボークドラツカースギルデ
			1
			最終頁に続く

(54) 【発明の名称】 糸状菌の分野における新規発現調節配列および発現産物

(57) 【要約】

本発明は、配列番号2および4のアミノ酸配列とそれぞれ70および75%の最少アミノ酸同一性を表すファミリー7および10のクリソスポリウム (C h r y s o s p o r i u m) のグリコシル ヒドロラーゼに対応する新規タンパク質、および配列番号6の部分的アミノ酸配列と少なくとも86%のアミノ酸同一性を表すクリソスポリウム (C h r y s o s p o r i u m) のグリセルアルデヒド リン酸 デヒドロゲナーゼに対応するタンパク質に関する。さらに本発明は、これらのタンパク質をコードする核酸配列、特に対応する遺伝子の発現を調節するプロモーター配列に関する。これらの遺伝子を発現するために好適な宿主は、真菌、特にクリソスポリウム (C h r y s o s p o r i u m) 株である。

【特許請求の範囲】

【請求項 1】

配列番号 2 のアミノ酸配列と [B L A S T アルゴリズムで決定した時に] 少なくとも 75 % のアミノ酸同一性を表すクリソスポリウム (C h r y s o s p o r i u m) のグリコシルヒドロラーゼ ファミリー 7 に対応するタンパク質、または配列番号 2 のアミノ酸配列 1-246 または 396-526 の対応する部分に同一である少なくとも 20 個の連続するアミノ酸を有するそれらの部分。

【請求項 2】

配列番号 4 のアミノ酸配列と [B L A S T アルゴリズムで決定した時に] 少なくとも 70 % のアミノ酸同一性を表すクリソスポリウム (C h r y s o s p o r i u m) のグリコシルヒドロラーゼ ファミリー 10 に対応するタンパク質、または配列番号 4 のアミノ酸配列 1-133 または 252-383 の対応する部分に同一である少なくとも 20 個の連続するアミノ酸を有するそれらの部分。

【請求項 3】

配列番号 4 のアミノ酸配列と [B L A S T アルゴリズムで決定した時に] 少なくとも 65 % のアミノ酸同一性を表すクリソスポリウム (C h r y s o s p o r i u m) のグリコシルヒドロラーゼ ファミリー 10 に対応するタンパク質、または配列番号 4 のアミノ酸配列 1-383 の対応する部分に同一である少なくとも 20 個の連続するアミノ酸を有するそれらの部分。

【請求項 4】

グリコシルヒドロラーゼ ファミリー 10 に対応し、そして配列番号 4 のアミノ酸配列 22-53 と少なくとも 55 % のアミノ酸同一性を有し、かつ/またはアミノ酸配列 22-53 から少なくとも 20 個の連続するアミノ酸を有し、かつ/または相対位置 44 にアミノ酸 F、W および Y の 1 つを有するセルロース結合ドメインを含んで成るタンパク質。

【請求項 5】

配列番号 6 の部分的アミノ酸配列と [B L A S T アルゴリズムで決定した時に] 少なくとも 86 % のアミノ酸同一性を表すクリソスポリウム (C h r y s o s p o r i u m) のグリセルアルデヒド リン酸 デヒドロゲナーゼ に対応するタンパク質、または配列番号 6 のアミノ酸配列 1-277 の対応する部分に同一である少なくとも 20 個の連続するアミノ酸を有するそれらの部分。

【請求項 6】

請求項 1 に記載の酵素の使用を含んで成る、β-グルコシド結合の加水分解法。

【請求項 7】

請求項 2、3 または 4 に記載の酵素の使用を含んで成る、β-キシロシド結合の加水分解法。

【請求項 8】

請求項 1 ないし 5 のいずれか 1 項に記載のタンパク質をコードする核酸配列。

【請求項 9】

配列番号 1、3 および 5 の任意の 1 つの核酸配列の 5' - 非コード領域に含まれる少なくとも 70 % のヌクレオチドを含んで成る核酸配列。

【請求項 10】

目的のポリペプチドをコードする核酸配列に操作可能に連結された配列番号 1、3 および 5 の任意の 1 つの核酸配列の 5' - 非コード領域に含まれる、クリソスポリウム (C h r y s o s p o r i u m) に由来する核酸発現-調節領域を含んで成る核酸構築物。

【請求項 11】

請求項 8 ないし 10 のいずれか 1 項に記載の核酸配列を含み、そしてコード核酸配列によりコードされるポリペプチドを発現することができる、組換え微生物株、好ましくは真菌株。

【請求項 12】

請求項 10 に記載の構築物または請求項 11 に記載の微生物株を使用するポリペプチドの

10

20

30

40

50

生産法。

【請求項 13】

配列番号 1、3 および 5 のいずれか 1 つの核酸配列の少なくとも 15 個の連続するヌクレオチド、またはその相補体を含んで成るオリゴヌクレオチドプローブ。

【請求項 14】

プローブが 20～50 ヌクレオチド長である、請求項 13 に記載のオリゴヌクレオチドプローブ。

【請求項 15】

プローブが検出可能な標識により標識されている、請求項 13 または 14 に記載のオリゴヌクレオチドプローブ。

10

【発明の詳細な説明】

【0001】

(技術分野)

発明の分野

本発明は糸状菌、特にクリソスポリウム属 (*Chrysosporium*) の株に由来する新規酵素、およびこれら酵素のコード配列および発現・調節配列に関する。本発明は本優先日前には公開されていない 1999 年 10 月 6 日に出願された国際公開第 00/20555 号 (PCT/NL99/00618) 明細書を拡張するものである。この開示は引用により本明細書に編入する。

発明の背景

20

遺伝子発現用の多数の宿主および形質転換法は従来技術で開示されてきた。細菌、例えば大腸菌 (*E. coli*) が挙げられることは多いが、大腸菌 (*E. coli*) は多くのタンパク質またはポリペプチドを分泌できず、しかもそれ自体は宿主細胞として工業レベルでのタンパク質またはポリペプチドの生産には望ましくない。一般に細菌には有効である大腸菌 (*E. coli*) のさらなる欠点は、原核生物は活性形で生産されるべき真核生物の多数のタンパク質またはポリペプチドに必要なさらなる修飾を提供できない点にある。タンパク質のグリコシル化およびタンパク質の正しい折り畳みは、活性なタンパク質またはポリペプチドが確実に生産されるために必要なプロセッシングの例である。そのようなプロセッシングを確実にするために、しばしば哺乳動物細胞が使用される；しかしそのような細胞の欠点は、それらを維持することが難しく、そして高価な培地を必要とする点である。したがってそのような形質転換系は、工業レベルでのタンパク質またはポリペプチドの生産には現実的ではない。それらは比較的少量の、しかし明らかに工業用ではない酵素を必要とする高価な医薬化合物から経済効果はあるかもしれない。

30

【0002】

多数の真菌発現系 (例えばアスペルギルス ニガー (*Aspergillus niger*)、アスペルギルス アワモリ (*Aspergillus awamori*)、アスペルギルス ニドランス (*Aspergillus nidulans*)、トリコデルマ レッセイ (*Trichoderma reesei*)) が開発された。多くの他の真菌が推薦されたが、様々な理由から広く認められ、または使用されることはなかった。一般的な意味で理想的な宿主は多くの基準を満たさなければならない：

40

- 理想的な宿主は廉価な培地を使用して容易に発酵されなければならない。
- 理想的な宿主は培地を効率的に使用すべきである。
- 理想的な宿主はポリペプチドまたはタンパク質を高収量で生産し、すなわち高いタンパク質対バイオマス比を表さなければならない。
- 理想的な宿主はタンパク質またはポリペプチドの効率的分泌ができるべきである。
- 理想的な宿主は所望のタンパク質またはポリペプチドの容易な単離および精製が可能でなければならない。
- 理想的な宿主は、さらなる活性化または修飾工程を必要とせずに活性形で生産されるように、所望するタンパク質またはポリペプチドをプロセッシングしなければならない。
- 理想的な宿主は容易に形質転換されるべきである。

50

－理想的な宿主は広範な発現調節要素の使用を可能とし、これにより容易な応用および汎用性を確実とする。

－理想的な宿主は廉価に使用される容易に選択可能なマーカーの使用を可能とするべきである。

－理想的な宿主は、安定な形質転換体を生成するべきである。

－理想的な宿主は、タンパク質またはポリペプチド産物に有害ではない条件下での培養を可能とするべきである（例えば低粘度、低剪断）。

【0003】

ノボ ノルディスク (Novo Nordisk) への国際公開第96/02563号明細書および米国特許第5,602,004号、同第5,604,129号および同第5,695,985号明細書は、アスペルギルス (Aspergillus) およびトリコデルマ (Trichoderma) 系の欠点を記載し、そして他の真菌に関する培養条件は大規模タンパク質生産により適し得ると示唆した。形質転換した培養物について提供された唯一の例は、ミセリオファソーラ サーモフィラ (Myceliophthora thermophila)、アクレモニウム アラバメンシス (Acremonium alabamense)、チエラビア テレストリス (Thielavia terrestris) およびスポロトリクム セルロフィルム (Sporotrichum cellulophilum) 株に関するものである。スポロトリクム属 (Sporotrichum) の株は溶解し、そして発酵条件下で他の株ではもたらされない緑の色素を生産すると報告されている。チエラビア テレストリス (Thielavia terrestris) の非一胞子形成突然変異体は、その形態により選択される微生物であると記載されている。しかしチエラビア (Thielavia) およびアクレモニウム (Acremonium) (これにより使用されたアクレモニウム属 (Acremonium) の株は使用したチエラビア属 (Thielavia) の株の不完全状態であった) のプロトプラスト形成効率は低く、そしてヒグロマイシンは選択マーカーとして有用ではないとも述べている。他の大部分はそれらの形態から有用となる可能性があるとし唆しているが、それらの形質転換は記載されていない。提案された株はコリナスクス属 (Corynascus)、サーモアスクス属 (Thermoascus)、ケタマカビ属 (Chaetomium)、シテノマイセス属 (Ctenomyces)、スキタリジウム属 (Scytalidium) およびアオカビ (Talaromyces) に属するものである。形質転換した宿主は、低レベルの導入したフミコーラ (Humicola) キシラナーゼを生産するだけで、チエラビア (Thielavia) は最低量を生産し；しかしこの情報は曖昧であり、そして実際にはチエラビア (Thielavia) が最高の態様であったと推定できた。この技術文献の命名は、1994年の工業用真菌のATCC名に基づく。このように高度なヘテロログス発現は達成されず、そして実際には推定される形態と発現の程度との間から引き出すことができた陽性の相関は無いことが明らかである。もし相関を作ることができれば、より容易に陰性になるであろう。1996年のATCCの真菌分類によれば、スポロトリクム サーモフィリウム (Sporotrichum thermophilum) ATCC 20493は、ミセリオファソーラ サーモフィラ (Myceliophthora thermophila) 株である。現在でもこの株はミセリオファソーラ サーモフィラ (Myceliophthora thermophila) と同定されている。この技術の予想不可能性は、これらの最近の開示から明らかである。

【0004】

ノボ ノルディスク (Novo Nordisk) の国際公開第97/26330号明細書は、ヘテロログスなポリペプチドの生産のために改良された特性を有する糸状菌の親細胞の突然変異体を得る方法を提案する。この方法は最初に特別に改変した形態を見だし、続いて形質転換体が親よりもより多くのヘテロログスがポリペプチドを生産するかどうかを評価することを含んで成る。この方法はフザリウム (Fusarium) A3/5およびアスペルギルス オリーゼ (Aspergillus oryzae) の株についてのみ具体的に説明している。この方法はアスペルギルス (Aspergillus)、トリコ

デルマ (Trichoderma)、チエラビア (Thielavia)、フザリウム (Fusarium)、ニューロスポーラ (Neurospora)、アクレモニウム (Acremonium)、トリプロカジウム (Tolytlocadium)、フミコーラ (Humicola)、スキタリジウム (Scytalidium)、ミセリオファソーラ (Myceliophthora) およびケカビ (Mucor) に応用できると示唆している。上記のように、この技術の予想不可能性および引用した応用の方法の予想不可能性は、成功への合理的な予想の元に一般に応用可能な教示を提供しない。

【0005】

国際公開第00/020555号明細書で、我々は上記の要件を満たす上に挙げたアスペルギルス (Aspergillus) およびトリコデルマ (Trichoderma) の簡単な使用による別の真菌発現系を記載した。この新規な系は、しばしば使用されるトリコデルマ レッセイ (Trichoderma reesei) の系よりも形質転換率が高いというさらなる利点を提供する。さらに培養条件はポリペプチド産物に有利となるさらなるボーナス (bonus) を提供する。

10

発明の詳細な記述

我々はクリソスポリウム属 (Chrysosporium) の株に由来する多数の工業的に興味深い酵素を完全な配列情報と共に記載する。また我々は、クリソスポリウム属 (Chrysosporium) の株に由来し、そしてホモログスおよびヘテロログスな遺伝子を発現するために有用な新規プロモーター系を記載する。

【0006】

本発明は、特にアミノ酸配列ならびにこれら酵素タンパク質に由来するペプチドにより同定される、ファミリー7 (例えばセロビオヒドラーゼ) および10 (例えばキシラナーゼ) のグリコシルヒドロラーゼおよびグリセルアルデヒドリン酸デヒドロゲナーゼ、およびこれらペプチドおよびタンパク質をコードする核酸配列、ならびに特にこれら遺伝子に関連する調節配列に関する。

20

【0007】

特に本発明は、上記に示した3種の単離された、または組換え酵素的タンパク質またはそれらの活性部分に関し、更なる開示および特許請求の範囲で特定される少なくともある程度の同一性を有するそれらの突然変異体、ならびにこれらタンパク質またはそれらの部分をコードする核酸配列、および/またはそれらの発現を調節する核酸配列を含む。これらの酵素は特に：(1) 配列番号2の配列と少なくとも75%、好ましくは少なくとも80%、またはさらに少なくとも85%のアミノ酸同一性を有するファミリー7のグリコシルヒドロラーゼ (セロビオヒドラーゼ)；(2) 配列番号4の配列と少なくとも70%、好ましくは少なくとも75%、またはさらに少なくとも80%のアミノ酸同一性を有するファミリー10のグリコシルヒドロラーゼ (エンドーキシラナーゼ XYL1)；および(3) 配列番号6の配列と少なくとも86%、好ましくは少なくとも90%、またはさらに少なくとも93%のアミノ酸同一性を有するグリセルアルデヒドリン酸デヒドロゲナーゼである。配列番号2、4および6の少なくとも20個、好ましくは少なくとも30個の連続するアミノ酸を有するポリペプチドおよびこれらポリペプチドをコードする核酸も、本発明の好適な部分である。対応するヌクレオチド配列はそれぞれ配列番号1 (c b h 1)、3 (x y l 1) および5 (g p d 1) に表す。

30

40

【0008】

組換え酵素は本質的に完全なタンパク質または酵素活性の少なくとも一部を有する短縮タンパク質を含んで成る。そのような短縮化された部分は触媒ドメインであるか、またはそれらの少なくとも約75%のアミノ酸であることができる。例として、本発明によるCBH1の触媒ドメインは、配列番号2のアミノ酸配列のアミノ酸20-495を含んで成り、そして本発明によるXYL1の触媒ドメインは、配列番号4のアミノ酸配列のアミノ酸54-384を含んで成る。この触媒ドメインは、他のタンパク質に由来するシグナル配列および/または他の酵素的タンパク質に由来する炭水化物-結合ドメインと合わせても合わせなくてもよい。あるいは本発明の酵素のセルロース結合ドメイン (CBH1 および

50

X Y L 1) は、他の酵素的タンパク質を触媒部分に融合することができる。

【0009】

本発明の核酸配列は、完全なタンパク質コード領域またはオリゴヌクレオチド、または好ましくは発現－調節配列であることができる。オリゴヌクレオチドは配列番号1、3および5の遺伝子に同一ではないが対応する遺伝子を同定するためのプローブとして使用することもでき；これらの遺伝子は本明細書に定める同一性の基準の百分率を満たす時、ならびにそれらのコードおよび非コード部分およびそれらの発現産物も本発明の一部である。オリゴヌクレオチドは好ましくは15～75、最も好ましくは20～50個のヌクレオチド長である。

【0010】

また本発明は、別の目的のタンパク質をコードする遺伝子に融合した任意の3種類のタンパク質のいずれかの発現－調節領域（プロモーターを含む）、あるいは別の発現調節領域に融合した任意の3種のタンパク質のコード領域、または発現－調節領域およびこれら新規タンパク質のタンパク質－コード領域の両方のいずれかを含んで成る発現系（カセット）に関する。発現－調節領域は、配列番号1、3および5の5'－非コード領域の少なくとも60%、好ましくは少なくとも70%、より好ましくは少なくとも75%、またはさらに80%、特にこれら5'－非コード領域の少なくとも40個の連続するヌクレオチドを含んで成る。ホモログスまたはヘテロログスな遺伝子のいずれかと組み合わせた、本発明の遺伝子の3'－非コード領域に同様に由来する終結配列も発現カセットにおいて有用である。

【0011】

本発明のポリヌクレオチドおよびオリゴヌクレオチドは、配列番号1、3および5の対応する配列と最少の配列同一性を有することができ、あるいはこれらの与えられた配列と緊縮条件下でハイブリダイズすることができる。緊縮なハイブリダイゼーション条件は当該技術分野で理解されている条件であり、例えば6×SSC（1000mlあたり20×SSC：175.3g NaCl、107.1g クエン酸ナトリウム、5H₂O、pH7.0）、0.1% SDS、0.05% ピロリン酸ナトリウム、5×デンハーツ溶液および20μg/ml 変性ニシン精子DNA中で56℃にて18～24時間のハイブリダイゼーション、続いて5×SSC、0.1% SDS中で56℃にて30分づつ2回の洗浄、そして2×SSC、0.1% SDS中で56℃にて30分づつ2回の洗浄。

【0012】

これらの発現系は、クリソスポリウム ロックンオウンズ属 (Chrysosporium lucknowense) の宿主のようなクリソスポリウム属 (Chrysosporium) の宿主、または他の非真菌または好ましくは真菌宿主で得ることがことができる。他の真菌宿主の例は、他のクリソスポリウム属 (Chrysosporium) の種または株、フザリウム属 (Fusarium) の種、アスペルギルス属 (Aspergillus) の種である。目的のタンパク質の回収が簡易になるように、そのような宿主は有利にはそれ自体が本来は、または培養条件の結果により目的のタンパク質に対応するタンパク質を生産することはできない宿主であり得る。

【0013】

本明細書および前述の特許請求の範囲で、本発明の発現系の産物として「ポリペプチド」または「ペプチド」または「目的のポリペプチド」または「目的のペプチド」と言う場合、この用語はタンパク質、すなわち特定の機能および／または2次および／または3次構造を有するポリペプチドも含んで成る。アミノ酸同一性の割合と言う場合、そのような同一性は、完全なタンパク質または通常使用されるBLASTアルゴリズムにより決定される最初の、および最後のアミノ酸番号により定める特定の部分に関する。

【0014】

国際公開第00/20555号明細書に記載された真菌発現系では、培養基のpHは中性またはアルカリ性であることができ、これによりもはや生産されたタンパク質またはポリペプチドを過激な、そして不活性化する可能性がある酸性pHにさらさない。タンパク質

10

20

30

40

50

またはポリペプチドが酸性の環境により適している場合、pH 4 のような酸性 pH で培養することも可能である。適当な培養は 4.0 ~ 10.0 の間の pH で起こることができる。しかし好適であるのは、宿主株が例えば 6 から 9 の間の pH のような pH でより良い成長を現すので中性からアルカリ性の pH である。pH 8 から上、そしてさらに 10 でもあり得るアルカリ性 pH での成長も、場合により良好な選択である。そのような宿主株の培養温度は、生産されるポリペプチドの種類の安定性に有利となる。培養温度は適当には 23 ~ 43 °C の温度である。明らかにそのような条件は、哺乳動物のポリペプチドの生産に特に興味深い。選択される温度は培養の経済的効果およびポリペプチドまたは培養株の感度に依るだろう。

【0015】

粘度に関連するバイオマスおよび生産されるタンパク質の量が、クリソスポリウム属 (Chrysosporium) の宿主に非常に好適であることが確認された。トリコデルマ ロンギブラキアタム (Trichoderma longibrachiatum) (正式にはトリコデルマ レッセイ (Trichoderma reesei) として知られる) およびアスペルギルス ニガー (Aspergillus niger) を用いて比較を行った。それぞれの最適条件下でトリコデルマ ロンギブラキアタム (Trichoderma longibrachiatum) は、2.5 ~ 5 g / リットルのバイオマスを与え、アスペルギルス ニガー (Aspergillus niger) は 5 ~ 10 g / リットルのバイオマスを与え、そしてクリソスポリウム属 (Chrysosporium) の宿主は 0.5 ~ 1 g / リットルのバイオマスを与えた。これは市販の使用する株よりも 5 ~ 10 倍の向上を与える。本発明はヘテロログスなタンパク質またはポリペプチドをコードする核酸配列を含んで成る発現系を対象とし、該核酸配列は以下に記載の発現調節領域および任意に分泌シグナルをコードする配列および／またはキャリアタンパク質をコードする配列に操作可能に連結されている。好ましくは本発明の組換え株は、目的のポリペプチドを分泌する。これは目的のポリペプチドを単離するために細胞を破壊する必要性を回避し、そしてまた宿主細胞の他の成分による発現産物の分解の危険性を最少とする。

【0016】

クリソスポリウム (Chrysosporium) は、Barnett and Hunter 1972、不完全菌類の属の説明 (Illustrated Genera of Imperfect Fungi)、第3版、ブルゲス (Burgess) 出版社に開示された形態と一致する形態により定めることができる。クリソスポリウム属 (Chrysosporium) の真菌の分類に関する詳細を提供する他の情報源は知られており、例えば Sutton の分類 (Van Oorschot, C. A. N. (1980) 「クリソスポリウム属 (Chrysosporium) の改定および近縁属 (A revision of Chrysosporium and allied genera)」、CBS の菌類の研究の第20で、バーン (Baarn) The Netherlands p1-36) である。CBS はブタベスト条約の寄託機関の1つである。これらの教示に従い、クリソスポリウム (Chrysosporium) 属はヒホミケス目 (Hyphomycetales) に属するモニリア (Moniliaceae) ファミリー内にある：以下の系統をクリソスポリウム (Chrysosporium) と定めるが、クリソスポリウム (Chrysosporium) はこれらの系統に限らない：C. ボトリオイデス (botryoides)、C. カルミケーリ (carmichaelii)、C. クラッシツニカタム (Crassitunicatum)、C. オイローパ (europae)、C. エボルセアンヌイ (evoluceannui)、C. ファリニコラ (farinicola)、C. ファスチジウム (fastidium)、C. フィリホルメ (filiforme)、C. ゲオルギエ (georgiae)、C. グロビフェルム (globiferum)、C. グロビフェルム変種アーチクラタム (globiferum var. articulatum)、C. グロビフェルム変種ニベウム (globiferum var. niveum)、C. ヒルンド (hirundo)、C. ヒスパニクム (hispanicum)、C. ホルミー (holmii)、C. インジクム (indicum)

10

20

30

40

50

)、C. イノプス (i n o p s)、C. ケラチノフィラム (k e r a t i n o p h i l u m)、C. クレイセリー (k r e i s e l l i)、C. クズロビアヌム (k u z u r o v i a n u m)、C. リグノラム (l i g n o r u m)、C. ロバタム (l o b a t u m)、C. ロックンオウンズ (l u c k n o w e n s e)、C. ロックンオウンズ (l u c k n o w e n s e) G a r g 27K、C. メディウム (m e d i u m)、C. メディウム変種スピッセセンス (m e d i u m v a r. s p i s s e s c e n s)、C. メフィチクム (m e p h i t i c u m)、C. メルダリウム (m e r d a r i u m)、C. メルダリウム変種ロゼウム (m e r d a r i u m v a r. r o s e u m)、C. マイナー (m i n o r)、C. パンニコラ (p a n n i c o l a)、C. パルブム (p a r v u m)、C. パルブム変種クレッセンス (p a r v u m v a r. c r e s c e n s)、C. ピロスム (p i l o s u m)、C. シュードメルダリウム (p s e u d o m e r d e r i u m)、C. ピリホルミス (p y r i f o r m i s)、C. クイーンズランジクム (q u e e n s l a n d i c u m)、C. シグレリ (s i g l e r i)、C. スルフレウム (s u l f u r e u m)、C. シンクロヌム (s y n c h r o n u m)、C. トロピクム (t r o p i c u m)、C. アンドラタム (u n d u l a t u m)、C. バルレナレンズ (v a l l e n a r e n s e)、C. ベスベルチリウム (v e s p e t r i l i u m)、C. ゾナタム (z o n a t u m)。

10

【0017】

C. ロックンオウンズ (l u c k n o w e n s e) は、セルラーゼタンパク質の自然な高い生産体を提供したので、特に興味を持たれたクリソスポリウム (C h r y s o s p o r i u m) 種の1つの形成する (国際公開第98/15633号明細書および関連する米国特許第5, 811, 381号明細書)。このクリソスポリウム ロックンオウンズ (C h r y s o s p o r i u m l u c k n o w e n s e) の特徴は：

20

S a b o u r a u d グルコース寒天上で14日に55mmの直径に達するコロニーは、クリーム色のフェルト状および綿毛状；集密および3～5mmの高さ；境界を明瞭に示す、整然として毛縁がある：後ろは淡い黄色からクリーム色。菌糸はヒアリンで、滑らか—そして薄い壁があり、少し分枝している。気生菌糸はほとんど受精能力があり、そして約1～3.5μmの幅の接近した状態で離れている。潜入菌糸は受精能力が無く、約1～4.5μmの幅で、薄い菌糸はしばしば回旋状である。分生子は末端および横であり、ほとんどが着性または短く、そしてしばしば円錐状に突出するか、または短い側枝がある。分生子は単生であるが、互いに近位にあり、1～4個の分生子が1つの菌糸細胞から発生し、半透明 (s u b h y a l i n e) の、わずかに薄い—そして滑らかな壁があり、ほとんどがほぼ球状 (s u b g l o b o s e) の、またバット形のオロボボイド (o r o b o v o i d)、1細胞化、2.5～11×1.5～6μm、広い基底痕がある。節間に分生子は無い。厚膜胞子は無い。ATCC44006、CBS251.72、CBS143.77およびCBS272.77はクリソスポリウム ロックンオウンズ (C h r y s o s p o r i u m l u c k n o w e n s e) 株の例であり、そして他の例は国際公開第98/15633号明細書 (米国特許第5, 811, 381号明細書) に提供されている。

30

【0018】

さらなる株はセルラーゼのより高い生産能力を持つ種から単離された。この株は内部での命名によりC1と呼ばれ、そしてブタペスト条約に従う寄託として、1996年8月29日に113184、ロシア、モスクワ、バクルシナ通り8のロシアアカデミーオブサイエンス (R u s s i a n A c a d e m y o f S c i e n c e) の微生物の国際寄託機関であるオール ロシアン コレクションに、寄託番号VKM F-3500Dに寄託した。これはクリソスポリウム ロックンオウンズ (C h r y s o s p o r i u m l u c k n o w e n s e) G a r g 27Kと呼ぶ。このC1株の特徴は以下の通り：

40

ジャガイモ—デキストロース寒天上で約7日に約55～66mmの直径に達するコロニーは；白—クリーム色のフェルト状、中心が2～3μmの高さ；境界を明瞭に示す、整然として毛縁がある：後ろは淡クリーム色。菌糸はヒアリンで、滑らか—そして薄い壁があり、少し分枝している。気生菌糸は受精能力があり、そして約2～3mmの幅で離れている

50

。潜入菌糸は受精能力が無い。分生子は末端および横であり；着性または短い側枝がある；無い；単生であるが、互いに近位にあり、ガラス質の、薄く－そして滑らかな壁があり、ほぼ球状の、またバット形または倒卵形の、1－細胞化、4－10 μm。厚膜孢子は無い。節間に分生子は無い。

【0019】

C1株の単離法は、国際公開第98/15633号明細書、米国特許第5,811,381号および同第6,015,707号明細書に記載されている。クリソスポリウム (Chrysosporium) の定義内には、自然にまたは誘導した突然変異のいずれかにより幾らか突然変異したものを含むクリソスポリウム (Chrysosporium) の祖先に由来する株を含む。クリソスポリウム (Chrysosporium) の突然変異体は、誘導した突然変異、特に照射および化学的突然変異誘発法の組み合わせにより得ることができる。

10

【0020】

例えばC1株はそれを紫外線光にさらして突然変異させ、UV13-6株を生成した。この株は続いてさらにN-メチル-N'-ニトロソグアニジンで突然変異させてNG7C-19を生成した。この株を次いで紫外線光にさらして突然変異させ、UV18-25株を生成した。この突然変異工程中に、形態学的特徴は液体またはプレート上の培養ならびに顕微鏡下で幾分変動した。各々の連続する突然変異誘発法で、培養物はコロニーが平らで、そしてマット状の外観に達するまで、クリソスポリウム (Chrysosporium) の特徴として記載したプレート上の毛羽だった、そしてフェルト状の外観の喪失を示した。幾つかの培地中で野生型の株で観察された茶色い色素も、突然変異した株では少なかった。液体培養では、突然変異体UV18-25は野生型C1株および突然変異体UV13-6およびNG7C-19よりも顕著に粘性が低かった。すべての株はクリソスポリウム属 (Chrysosporium) のおおざっぱな顕微鏡的特徴を維持する一方、菌糸は各連続的突然変異で狭くなり、そしてUV18-25で顕著な菌糸の断片化が観察できた。この菌糸の断片化はUV18-25の培養に付随する低い粘度が原因と思われる。株が孢子形成する能力は、各突然変異工程で低下した。上記はクリソスポリウム属 (Chrysosporium) に属する株が上記の形態学的定義から幾らかずれていることを説明している。各突然変異工程で、セルラーゼおよび細胞外タンパク質の生産はさらに上昇したが、幾つかの突然変異はプロテアーゼ発現の低下をもたらした。菌類の分類学で決定できる基準は、例えばCBS、VKMFおよびATCCから入手可能である。クリソスポリウム属 (Chrysosporium) のC1株、UV13-6株、NG7C-19株およびUV18-25と国内で命名されている株は、ブタペスト条約に従い、モスクワのオールロシアンコレクション (VKM) 寄託機関に寄託した。野生型C1株は29-08-1996に番号VKMF-3500Dで寄託され、C1 UV13-6突然変異体はVKMF-3632D (02-09-1998) として寄託され、C1 NG7c-19突然変異体はVKMF-3633D (02-09-1998) として寄託され、そしてC1 UV18-25突然変異体はVKMF-3631D (02-09-1998) として寄託された。

20

30

【0021】

当該技術分野で番号が知られている非-毒性のクリソスポリウム属 (Chrysosporium) の株を使用することは、大規模生産で環境への危険性を下げ、そして簡便な生産手法は同時に経費の低減になるので好ましい。

40

【0022】

発現-調節領域は宿主のクリソスポリウム属 (Chrysosporium) の株が発現のために認識するDNA配列である。これは発現すべきポリペプチドをコードする核酸配列に操作可能に連結されたプロモーター配列を含んで成る。プロモーターは発現する配列の開始コドンに関係する配置が発現を可能とするように連結される。プロモーター配列は構成的または誘導性であることができる。クリソスポリウム属 (Chrysosporium) の株からポリペプチドの発現を可能とする任意の発現調節配列またはそれらの組み合

50

わせを構想する。発現調節配列は、適当には真菌の発現調節領域、例えば子囊菌類の調節領域である。適当には真菌の発現調節領域は、以下の真菌の属に由来する調節領域に由来する：アスペルギルス (Aspergillus)、トリコデルマ (Trichoderma)、クリソスポリウム (Chrysosporium)、ハンセニューラ (Hansenula)、ケカビ (Mucor)、ピヒア (Pichia)、ニューロスポーラ (Neurospora)、トリプロカジウム (Tolyplocadium)、リゾムコール (Rizomucor)、フザリウム (Fusarium)、ペニシリウム (Penicillium)、サッカロミセス (Saccharomyces)、アオカビ (Talaromyces)、あるいはエメリセラ (Emericella)、ハイポクレア (Hypocrea) のようなそれらの性的形態、例えばトリコデルマ (Trichoderma) に由来するセロビオヒドラーゼプロモーター、アスペルギルス (Aspergillus) に由来するグルコアミラーゼプロモーター、(Aspergillus) に由来するグリセルアルデヒドリン酸デヒドロゲナーゼプロモーター、(Aspergillus) に由来するアルコールデヒドロゲナーゼAおよびアルコールデヒドロゲナーゼRプロモーター、(Aspergillus) に由来するTAKAアミラーゼプロモーター、ニューロスポーラ (Neurospora) のホスホグリセレートおよびクロスパスウェイ (cross-pathway) 制御プロモーター、リゾムコール ミーハイ (Rizomucor miehei) のアスパラギン酸プロティナーゼプロモーター、リゾムコール ミーハイ (Rizomucor miehei) のリパーゼプロモーターおよびペニシリウム カネセンス (Penicillium canescens) のベーターガラクトシダーゼプロモーター。宿主株と同じ属に由来する発現調節配列は、特定の宿主に特別に最も適しているので、極めて適する。すなわち好ましくは発現調節配列は、クリソスポリウム属 (Chrysosporium) の株に由来するものである。

10

20

30

40

【0023】

好ましくは選択した宿主中での高い発現を可能とする発現－調節領域が適用される。これは好ましくは本発明によるクリソスポリウム (Chrysosporium) に由来する発現－調節領域である。当該技術分野で周知なヘテロログスな宿主に由来する高い発現－調節領域でもよい。大量に発現することが知られ、したがって本発明に適当な発現調節配列を提供するタンパク質の具体的な例は、限定するわけではないがヒドロホビン (hydrophobin)、プロテアーゼ、アミラーゼ、キシラナーゼ、ペクチナーゼ、エステラーゼ、ベーターガラクトシダーゼ、セルラーゼ (例えばエンドーグルカナーゼ、セロビオヒドラーゼ) およびポリガラクトツロナーゼである。高生産は固体状態および液内発酵条件の両方で確認された。そのようなタンパク質の存在または生産を評価するためのアッセイは、当該技術分野で周知である。例えばシグマ (Sigma) およびメガザイム (Megazyme) のカタログは、多数の例を提供する。メガザイムがアイルランドのウィックロー カントリー、ブレイのブレイビジネスパークにある。シグマアルドリッチ (Sigma Aldrich) は、世界中に多くの系列会社があり、例えばそれは米国、ミズーリ州、セントルイスの私書箱14508である。セルラーゼに関して、CMCaseアッセイ、エンドビスコメトリック (endoviscometric) アッセイ、Avicelaseアッセイ、ベーターグルカナーゼアッセイ、RBBCMCCaseアッセイ、Cellazymeアッセイのような市販の利用できるアッセイを言う。キシラナーゼアッセイも市販されている (例えば、DNSおよびメガザイム)。代替は当業者に周知であり、そして主題に関連する一般的技術文献から見いだすことができ、そしてそのような情報は引用により本明細書に編入する。例として我々は、“Methods in Enzymology” 第1巻、1955右から1988年の297～299巻を示す。適当にクリソスポリウム (Chrysosporium) プロモーター配列を適用して、宿主によるそれらの良い認識を確実にする。

【0024】

我々は、ヘテロログスな発現調節配列が自然なクリソスポリウム (Chrysosporium) 配列としてクリソスポリウム (Chrysosporium) 中で効率的に働く

50

ことを見いだした。これはクリソスポリウム (C h r y s o s p o r i u m) の形質転換に使用される周知の構築物およびベクターを可能とし、ならびにこの新規な発現および分泌宿主中で良好な発現率を可能とする他のベクターを構築する多くの可能性を提供する。例えば標準的なアスペルギルス (A s p e r g i l l u s) の形質転換法は、例えば C h r i s t i a n s e n e t a l の B i o / T e c h n o l . 6 : 1 4 1 9 - 1 4 2 2 (1 9 8 8) に記載されているように使用することができる。アスペルギルス (A s p e r g i l l u s) の形質転換ベクターの詳細を提供する他の文書、例えば米国特許第 4, 8 1 6, 4 0 5 号、同第 5, 1 9 8, 3 4 5 号、同第 5, 5 0 3, 9 9 1 号、同第 5, 3 6 4, 7 7 0 号および同第 5, 5 7 8, 4 6 3 号明細書、欧州特許第 2 1 5, 5 9 4 号明細書 (トリコデルマ (T r i c h o d e r m a) にも関する) およびそれらの内容は、引用により本明細書に編入する。セルラーゼに関する大変高い発現率が、クリソスポリウム属 (C h r y s o s p o r i u m) の株で確認され、そのようなタンパク質の発現調節領域は特に好適である。我々は具体例として既に挙げた寄託したクリソスポリウム属 (C h r y s o s p o r i u m) の株に言及する。

10

20

30

40

50

【0025】

クリソスポリウム (C h r y s o s p o r i u m)、好ましくはクリソスポリウム ロックンOWNズ (C h r y s o s p o r i u m l u c k n o w e n s e) に由来する核酸発現調節領域またはそれらの誘導体を含んで成る核酸構築物は、発現されるポリペプチドをコードする遺伝子に操作可能に連結されるそのような核酸を含んで成る突然変異体のクリソスポリウム (C h r y s o s p o r i u m) のように、本発明の好適な態様を形成する。そのような核酸構築物は以下に詳細に記載するように、セルラーゼまたはキシラナーゼ発現、好ましくはセロビオヒドラーゼ発現またはダリセルアルデヒドリン酸デヒドロゲナーゼ発現に係するクリソスポリウム (C h r y s o s p o r i u m) に由来する発現調節領域となるだろう。本発明の核酸配列はクリソスポリウム属 (C h r y s o s p o r i u m) の株から適切に得ることができ、そのような株は本明細書中のいたるところで特定されている。プロモーター配列を決定できる様式は多数であり、そして当該技術分野では周知である。関連する遺伝子の始めの A T G コドンの上流領域のヌクレアーゼ欠失実験は、そのような配列を提供するだろう。また例えば、コンセンサス配列の分析により、目的遺伝子を見いだすことを導くことができる。ハイブリダイゼーションおよび増幅技法を使用して、当業者は対応するプロモーター配列に容易に到達することができる。

【0026】

C 1 エンドヌクレアーゼのプロモーター配列は、対応する遺伝子をクローニングすることによりこの様式で同定した。酵素が独特のプロモーターで高レベルで発現されるので、本発明に従い好適なプロモーターは 5 5 k D a セロビオヒドラーゼ (C B H 1 1) プロモーター、3 0 K D a キシラナーゼ (X y 1 1) プロモーターであり、およびグリセルアルデヒドリン酸デヒドロゲナーゼプロモーターである。対応するプロモーター配列は、それぞれ配列番号 1 (C B H 1 について) および配列番号 3 (X y 1 1 について) に与える配列情報を使用して、国際公開第 0 0 / 2 0 5 5 5 号明細書に記載されているようにクローニングすることにより簡単な様式で同定される。クリソスポリウム (C h r y s o s p o r i u m) の炭水化物分解酵素のプロモーター、特に C 1 プロモーターは、宿主生物、特に真菌または他の微生物宿主中で所望するポリペプチドの発現に有利に使用することができる。配列番号 1、3 および 5 に与える配列と、または他のクリソスポリウム (C h r y s o s p o r i u m) の遺伝子について見いだされる配列と、少なくとも 6 5 %、好ましくは少なくとも 7 0 %、最も好ましくは少なくとも 7 5 % のヌクレオチド配列同一性を有するプロモーター配列は、本発明の一部である。

【0027】

本発明の組換え株および核酸配列の特定の態様についても、我々は実施例で言及する。我々はまた、組換え株について高い発現プロモーター配列、例えばアスペルギルス (A s p e r g i l l u s) およびトリコデルマ (T r i c h o d e r m a) について開示されるような特に真菌での高い発現を提供する配列を記載する従来技術も言及する。従来技術は

、多数の発現調節領域をアスペルギルス (Aspergillus) で使用するために提供する (例えばノボ (Novo) の米国特許第 5, 252, 726 号明細書および Uni lever の米国特許第 5, 705, 358 号明細書)。そのような従来技術の内容は引用により本明細書に編入する。

【0028】

ハイドロホビン遺伝子は、高度に発現される真菌の遺伝子である。これは好ましくはクリソスポリウム (Chrysosporium) に由来するハイドロホビン遺伝子のプロモーター配列が、本発明の適切な態様において発現調節配列として適当に適用できることを示唆している。ハイドロホビンに関するトリコデルマ レッセイ (Trichoderma reesei) およびトリコデルマ ハージアナム (Trichoderma harzianum) の遺伝子配列は、例えば従来技術に開示され、ならびにアスペルギルス フミガタス (Aspergillus fumigatus) およびアスペルギルス ニドランス (Aspergillus nidulans) の遺伝子配列および関連する配列情報は引用により本明細書に編入する (Munoz et al, Curr. Genet. 1997, 32 (3) : 225-230; Nakarai-Setälä T. et al, Eur. J. Biochem. 1996 15: 235 (1-2) : 248-255, M. Parta et al, Infect. Immun. 1994 62 (10) : 4389-4395 および Stringer M. A. et al, Mol. Microbiol. 1995 16 (1) : 33-34)。この配列情報を使用して、当業者はすでに上記に示唆したような標準技法に従い実験せずにクリソスポリウム (Chrysosporium) のハイドロホビン遺伝子の発現調節配列を得ることができる。本発明の組換えクリソスポリウム属 (Chrysosporium) の株は、目的のポリペプチドをコードする配列に操作可能に連結されたハイドロホビン調節領域を含んで成ることができる。

【0029】

発現調節配列はさらにエンハンサーまたはサイレンサーを含んで成ることができる。これらは従来技術において周知であり、そして通常はプロモーターから幾分離れて配置される。発現調節配列は、アクチベーター結合部位およびリプレッサー結合部位を含むプロモーターも含んで成ることができる。場合により、そのような部位を修飾してこの種の調節を排除してもよい。creA 部位が存在する糸状菌プロモーターが記載された。そのような creA 部位を突然変異させて、正常には非-突然変異部位の存在から生じるグルコース抑制を確実に排除することができる。Gist-Brocades の国際公開第 94/13820 号明細書は、この原理を具体的に説明する。そのようなプロモーターの使用により、グルコースの存在下でプロモーターにより調節される核酸配列によりコードされるポリペプチドの生産が可能となる。同じ原理は、国際公開第 97/09438 号明細書からも明らかである。これらのプロモーターはそれらの creA 部位を用いて、または用いずに使用することができる。creA 部位が突然変異した突然変異体は、本発明に従い発現調節配列として組換え株中で使用することができ、そしてそれを調節する核酸配列が次いでグルコースの存在下で発現し得る。そのようなクリソスポリウム属 (Chrysosporium) のプロモーターは、国際公開第 97/09438 号明細書で具体的に説明した様式と類似する様式で抑制解除を確実にする。creA 部位の同一性は、従来技術から知られている。あるいは突然変異していない CreA 結合部位を持つプロモーターを、いずれかの抑制系 (例えば creA 遺伝子自体) に突然変異を有する宿主株に適用して、creA 結合部位の存在が際立つこの株がグルコースの存在下でタンパク質またはポリペプチドを生産できるようにすることも可能である。

【0030】

ターミネーター配列も発現調節配列であり、そして発現する配列の 3' 末端に操作可能に連結されている。任意の真菌ターミネーターが本発明の宿主クリソスポリウム属 (Chrysosporium) の株中で機能的となるだろう。例は A. ニドランス (nidulans) trpC ターミネーター (1)、A. ニガー (niger) アルファ-グルコシダーゼターミネーター (2)、A. ニガー (niger) グルコアミラーゼ ターミネ

10

20

30

40

50

ーター (3)、ムコール ミーハイ (Mucor miehei) カルボキシプロテアーゼターミネーター (米国特許第 5, 578, 463 号明細書)、およびトリコデルマ レッセイ (Trichoderma reesei) セロビオヒドラーゼターミネーターである。当然、クリソスポリウム (Chrysosporium) ターミネーター配列は、クリソスポリウム (Chrysosporium) 中で十分に機能し、そして適当である (例えば CBH1 ターミネーター)。

【0031】

本発明に従い使用される適当な組換えクリソスポリウム属 (Chrysosporium) の株は、シグナル配列として定められるアミノ酸配列をコードする配列に操作可能に連結された発現すべき核酸配列を有する。シグナル配列は、発現するポリペプチドのアミノ酸発現に操作可能に連結された時、宿主真菌からそれらの分泌を可能とするアミノ酸配列である。そのようなシグナル配列はヘテロロガスなポリペプチドに通常付随する配列か、または宿主に自然な配列であり得る。また宿主およびポリペプチドの両方に対して外来であることもできる。シグナル配列をコードする核酸配列は、枠内に配置されてシグナル配列およびヘテロロガスなポリペプチドの翻訳を可能としないなければならない。クリソスポリウム属 (Chrysosporium) の株からポリペプチドの分泌を可能とすることができ任意のシグナル配列を構想する。そのようなシグナル配列は適当な真菌シグナル配列、好ましくは子囊菌類のシグナル配列である。

【0032】

シグナル配列の適当な例は、一般的な酵母または任意の以下の真菌の具体的な属に由来することができる：アスペルギルス (Aspergillus)、トリコデルマ (Trichoderma)、クリソスポリウム (Chrysosporium)、ピヒア (Pichia)、ニューロスポーラ (Neurospora)、リゾムコール (Rizomucor)、ハンセニューラ (Hansenula)、フミコーラ (Humicola)、ケカビ (Mucor)、トリプロカジウム (Tolytlocadium)、フザリウム (Fusarium)、ペニシリウム (Penicillium)、サッカロミセス (Saccharomyces)、アオカビ (Talaromyces)、あるいはエメリセラ (Emericella)、ハイポクレア (Hypocrea) のようなそれらの性的形態。特に有用なシグナル配列は、しばしば以下のタンパク質、セロビオヒドラーゼ、エンドグルカナーゼ、ベーターガラクトシダーゼ、キシラナーゼ、ペクチナーゼ、エステラーゼ、ハイドロホビン、プロテアーゼまたはアミラーゼに自然に付随する配列である。例にはアスペルギルス (Aspergillus) またはハンセニューラ (Hansenula) のアミラーゼまたはグルコアミラーゼ (4)、アスペルギルス オリーゼ (Aspergillus oryzae) の TAKA アミラーゼ、アスペルギルス ニガー (Aspergillus niger) のアルファアミラーゼ、ケカビ (Mucor) のカルボキシペプチダーゼ (米国特許第 5, 578, 463 号明細書)、リゾムコール ミーハイ (Rizomucor miehei) のリパーゼまたはプロティナーゼ、トリコデルマ (Trichoderma) のセロビオヒドラーゼ (5)、ペニシリウム カネッセン (Penicillium canescens) のベーターガラクトシダーゼおよびサッカロミセス (Saccharomyces) のアルファ接合因子を含む。

【0033】

あるいはシグナル配列は、バチルス属 (Bacillus) の株のアミラーゼまたはスブチリシン遺伝子に由来することができる。宿主株と同じ属に由来するシグナル配列は、特定の宿主に特別に最も適合すると思われるので大変適しており、すなわち好ましくはシグナル配列は上記で言うクリソスポリウム (Chrysosporium)、特にクリソスポリウム属 (Chrysosporium) の C1 株、UV13-6 株、NG7C-19 株および UV18-25 株のシグナル配列である。糸状菌、酵母および細菌に由来するシグナル配列も有用である。非-真菌起源、特に細菌、植物および哺乳動物のシグナル配列も有用であると考えられる。

【0034】

10

20

30

40

50

本発明の任意の態様に従い使用される組換え宿主は、さらに選択可能なマーカーを含んで成ることができる。そのような選択可能なマーカーは、形質転換した、またはトランスフェクトした細胞の容易な選択を可能とする。選択可能なマーカーはしばしば遺伝子産物をコードし、非形質転換株に対して外来の特別な種類の耐性を提供する。この細胞は一般に重金属、抗生物質および殺生物剤に対して耐性であることができる。原栄養性も非-抗生物質の種類の有用な選択可能マーカーである。非-抗生物質の選択可能マーカーは、目的のタンパク質またはポリペプチドが食品または医薬品に使用される場合、そのような産物のより迅速または複雑ではない規制の認可の観点から好適であり得る。G R A S 指標がそのようなマーカーに使用されることが大変多い。当業者は多数のそのようなマーカーを利用することができる。例えば F D A はそのリストを提供している。最もよく使用されるのは、薬剤に対する耐性または栄養的欠失の解除を付与する群、例えば a m d S (アセトアミダーゼ)、h p h (ヒグロマイシン ホスホトランスフェラーゼ)、p y r G (オルニチン-5'-リン酸デカルボキシラーゼ)、t r p C (アントラニレート シンターゼ)、a r g B (オルニチン カルバモイルトランスフェラーゼ)、s C (スルフェート アデニルトランスフェラーゼ)、b a r (ホスフィノトリシン アセチルトランスフェラーゼ)、グルホシネート耐性、n i a D (硝酸レダクターゼ)、ブレオマイシン耐性遺伝子、より詳細には S h b l e、スルホニルウレア耐性、例えばアセトラクテートシンターゼ突然変異 i l v 1 から成る群から選択される選択可能マーカーである。選択マーカーが別のベクター上にある場合、または選択マーカーが目的のポリペプチドについてポリペプチドをコードする配列と同じ核酸フラグメント上にある場合、選択はまた同時形質転換により行うこともできる。

【0035】

本明細書で使用する用語であるヘテロロガスなポリペプチドとは、本発明に従い発現に使用するクリソスポリウム属 (C h r y s o s p o r i u m) の宿主株により通常は発現または分泌されないタンパク質またはポリペプチドである。このポリペプチドは、宿主株中で生じない限り植物または動物(脊椎動物または無脊椎動物)起源、例えば哺乳動物、魚、昆虫または微生物起源であることができる。哺乳動物にはヒトを含むことができる。微生物にはウイルス、細菌、古細菌、および真菌、すなわち糸状菌および酵母を含んで成る。B e r g e y の細菌決定学の手引き (M a n u a l f o r B a c t e r i a l D e t e r m i n o l o g y) は細菌および古細菌の十分なリストを提供している。医薬的目的には、ヒトのタンパク質が好まれることが大変多く、すなわち好適な態様を形成する本発明の組換え体宿主は、ポリペプチドがヒト起源である宿主となるだろう。食品製造のような目的には、適当にはヘテロロガスなポリペプチドは動物、植物または藻類起源のものとなるだろう。したがってそのような態様は本発明の適当な例であると考えられる。あるいは有用である態様には、任意の細菌、酵母、ウイルス、古細菌および真菌起源のヘテロロガスなポリペプチドを含む。真菌起源が最も好適である。

【0036】

本発明の適当な態様は、適合したコドン使用頻度を含むヘテロロガスな核酸配列を含んで成る。そのような配列は宿主に由来する自然なアミノ酸配列をコードするが、異なる核酸配列を有し、すなわち中の特定のコドンが発現のために宿主株によりより容易に使用される同じアミノ酸配列をコードする他のコドンに置き換えられている。これによりヘテロロガスな核酸配列のより良い発現を導くことができる。これは当業者には一般的な事柄である。この適合したコドン使用頻度は、非-真菌のコドンの利用に比べて真菌の既知のコドン使用頻度に基づき行うことができる。さらに特別にクリソスポリウム (C h r y s o s p o r i u m) 自体のコドン使用頻度に適合させることもできる。この類似性はトリコデルマ (T r i c h o d e r m a)、フミコーラ (H u m i c o l a) およびアスペルギルス (A s p e r g i l l u s) で観察されるコドン使用頻度が、コドン使用頻度を適合せずにそのような微生物の配列の交換を可能にできるはずとなるものである。当業者はこれら真菌のコドン使用頻度に関する詳細を入手することができ、そしてそれは引用により本明細書に編入する。

【0037】

本発明は上に挙げた組換え体のクリソスポリウム属 (C h r y s o s p o r i u m) の株に限定されず、クリソスポリウム属 (C h r y s o s p o r i u m) の株のホモログスなタンパク質をコードする核酸配列を含んで成る組換え体のクリソスポリウム属 (C h r y s o s p o r i u m) の株も網羅し、該核酸配列は発現調節配列に操作可能に連結され、そして該組換え体株は同じ条件下では対応する非一組換え体株よりもより多く該タンパク質を発現する。目的のホモログスなポリペプチドの場合、それは好ましくは本明細書にいたるところに記載しているヒドロラーゼ、プロテアーゼまたは炭水化物分解酵素のような中性またはアルカリ性酵素である。ポリペプチドは酸性でよい。好ましくは組換え体株は非一組換え体株よりも大量にポリペプチドを発現するだろう。ヘテロログスなポリペプチドに関して述べたすべての解説は、ホモログスなポリペプチドであるセルラーゼにも有効（個々の違いは考慮して）である。

10

【0038】

このように本発明は遺伝的に操作された微生物株も網羅、ここで導入される配列はクリソスポリウム (C h r y s o s p o r i u m) 起源である。しかしそのような株は、例えばクリソスポリウム (C h r y s o s p o r i u m) を形質転換またはトランスフェクトするために使用する核酸配列中に存在するヘテロログスな配列により、または目的のポリペプチドをコードする配列の多数のコピーが存在するという事実により、またはこれらが同一条件下で操作していない株を上回る量で発現されるという事実により、または発現が通常は非発現の条件で起こるという事実により、自然に存在する株と識別することができる。非組換え体の状況、または操作していない株で誘導する因子とは別の因子が発現を誘導する場合と対照的に、これは誘導性プロモーターが目的配列を調節する場合であり得る。本発明は古典的な遺伝子技法または遺伝子操作法のいずれかを使用した操作を通して誘導された株を対象とする。

20

【0039】

本発明に従い発現系およびそれらを含む宿主株は、炭水化物分解酵素（セルラーゼ、キシラナーゼ、マンナーゼ、マンノシダーゼ、ペクチナーゼ、アミラーゼ、例えばグルコアミラーゼ、 α -アミラーゼ、 α -および β -ガラクトシダーゼ、 α -および β -グルコシダーゼ、 β -グルカナナーゼ、キチナーゼ、キタナーゼ）、プロテアーゼ（エンドプロテアーゼ、アミノプロテアーゼ、アミド-およびカルボキシーペプチダーゼ、ケラチナーゼ）、他のヒドロラーゼ（リパーゼ、エステラーゼ、フィターゼ（p h y t a s e））、酸化還元酵素（カタラーゼ、グルコース-オキシダーゼ）およびトランスフェラーゼ（トランスグリコシラーゼ、トランスグルタミナーゼ、イソメラーゼおよびインベルターゼ）から選択されるヘテロログスなタンパク質をコードする核酸配列を含んで成ることができる。

30

【0040】

本発明に従い生産される最も興味深い産物は、セルラーゼ、キシラナーゼ、ペクチナーゼ、リパーゼおよびプロテアーゼであり、ここでセルラーゼおよびキシラナーゼは、ベーター1, 4-結合を開裂し、そしてセルラーゼはエンドグルカナナーゼ、セロビオヒドラーゼおよびベーターグルコシダーゼを含んで成る。これらのタンパク質は当該技術分野で知られている種々の工業的プロセスに大変有用である。特にセルラーゼに関して、我々は例えばセロビオヒドラーゼおよびエンドグルカナナーゼの使用を記載する国際公開第98/15633号明細書を言及する。該出願内容は引用により本明細書に編入する。

40

【0041】

本発明による組換え体は、目的のポリペプチドをコードする核酸配列を有することができ、酸性pH、すなわち6未満のpH、さらにpH5.5未満、より適当にはpH5未満、そしてさらにpH4以下のpHで不活性化または不安定であるポリペプチドをコードする。一般に開示されている真菌の発現系は中性からアルカリ性の条件下では培養されず、酸性pHで培養されるので、これは特に興味深い態様である。すなわち本発明による系は、酸性pHでは不活性化されるか、または不安定になり易いタンパク質またはポリペプチドの安全な真菌発現系を提供する。

50

【0042】

本発明による任意の態様で定める極めて特別な組換え株は（ここで目的のポリペプチドをコードする核酸配列は、5より高いpHで、好ましくは中性もしくはアルカリ性のpHで（すなわち7より高い）かつ／または6より高いpHで最適活性および／または安定性を表すタンパク質またはポリペプチドをコードする）、本発明の好適な態様であると考えられる。そのようなpH値で、50%より多く、70%より多く、そしてさらに90%より多い最適活性が特に有用な態様であると思われる。培養条件下で発現するポリペプチドは、培養条件で活性である必要は無く、実際には活性状態が宿主に対して害となり得るので、不活性である培養条件下で培養することが有利となり得る。しかし必要なことは、培養条件下でタンパク質またはポリペプチドが安定であるということである。安定性は、熱安定性であり得る。これはまた例えば目的のポリペプチドまたはタンパク質の生産または応用の組成物または方法に存在するような、特別な組成物または化学物質に対する安定性でもあり得る。セルラーゼまたはリパーゼ等を含んで成る界面活性剤組成物中のLASは、タンパク質にしばしば有害な化学物質である。応用で使用する期間は、短期から長期の暴露で変動するので、安定性は応用毎に変わる変動する時間の長さにわたる。当業者は場合毎に正しい条件を確認することができるだろう。種々の酵素産物の最適活性を決定するために、多数の市販されているアッセイを利用することができる。例えばシグマおよびメガザイムのカatalogにはそれらが示されている。具体的な試験例は、本明細書のいたるところで挙げている。製造元は応用に関する指針を提供する。

【0043】

クリソスポリウム属 (Chrysosporium) の株は、発現すべき目的の配列で形質転換またはトランスフェクトするために適当に使用されることができ、そしてそのような株は比較的低いバイオマスを現す。我々は、発酵の終わりに200~600cPの粘度までに培養した時、クリソスポリウム属 (Chrysosporium) の株はトリコデルマ レッセイ (Trichoderma reesei) よりも2~5倍低いバイオマスを有し、そして対応する条件下で1500~2000cPの粘度までに培養した時、アスペルギルス ニガー (Aspergillus niger) よりも10~20倍低いバイオマスを現し、すなわちそれらの各々の最適培養条件が高レベルの発現を提供できることを見いだした。この発現レベルは大変少ないバイオマスおよび大変低い粘度で2種の市販の参照株をはるかに上回る。これはそのようなクリソスポリウム属 (Chrysosporium) の株の発現の収量が、アスペルギルス ニガー (Aspergillus niger) およびトリコデルマ レッセイ (Trichoderma reesei) の収量よりもかなり高いことを意味する。そのように形質転換またはトランスフェクトしたクリソスポリウム属 (Chrysosporium) の株は、本発明の適当な態様を形成する。

【0044】

上記の条件下で、トリコデルマ レッセイ (Trichoderma reesei) に関して2, 5~5, 0g/リットル、そしてアスペルギルス ニガー (Aspergillus niger) に関して5~10のバイオマスに対して、クリソスポリウム属 (Chrysosporium) のC1 (18-25) 株に関しては0, 5~1.0g/リットルのバイオマスを見いだした。実施例ではこの方法の詳細を提供する。

【0045】

適当な態様では、対応する、または同一の条件、すなわち各々の最適培養条件下で、組換え体クリソスポリウム属 (Chrysosporium) の株はUV13-6またはC-19により生産されるセルラーゼのリットルあたりのモル生産に少なくとも等しい量で、そして最も好ましくはUV18-25株により生産される量に少なくとも等しいか、より多くのタンパク質またはポリペプチドを生産する。

【0046】

我々は発現および分泌率が、UV18-25株の菌糸の形態、すなわち分断した短い菌糸を現すクリソスポリウム属 (Chrysosporium) の株を使用した時に非常に高

いことも見いだした。このように本発明の組換え株は好ましくはそのような形態を現す。しかし本発明はこの新規および強い特徴を現す非一組換え株または他の操作したクリソスポリウム属 (*Chrysosporium*) の株も網羅する。また本発明はさらに、等しい発酵条件下で、C1に比べて、好ましくはUV13-6株よりも低い、好ましくはNG7C-19よりも低い、好ましくはUV18-25よりも低い、減少した孢子形成を現す本発明により記載される任意の態様の組換えクリソスポリウム属 (*Chrysosporium*) の株も網羅する。また本発明はさらに、等しい発酵条件下で少なくともC1に匹敵する、好ましくは任意のUV13-6、NG7C-19およびUV18-25に匹敵するバイオマスに対するタンパク質生産量比を現す本発明に従い記載された任意の態様における組換えクリソスポリウム属 (*Chrysosporium*) の株を網羅する。しかし本発明は、そのままで、または他の任意の他の態様と組み合わせてこの新規および強い特徴を現す非一組換え株あるいは他の操作したクリソスポリウム属 (*Chrysosporium*) の株も網羅する。

10

20

30

40

50

【0047】

本発明の別の魅力的な態様はまた、対応するもしくは同一の醗酵槽条件下で株NG7C-19のものより下、好ましくはUV18-25のものより下の粘度を表す組換えクリソスポリウム属 (*Chrysosporium*) の株も包含する。本発明は、しかしながら、それ自体もしくは他の態様のいずれかと組合せてこの新規かつ発明の特徴を表すクリソスポリウム属 (*Chrysosporium*) の非組換え株もしくは別の方法で工作された株もまた包含する。われわれは、UV18-25の培養物の粘度が、200~600cPの次数のものであるトリコデルマ レエセイ (*Trichoderma reesei*) のものと対照的に10cPより下である (黒色アスペルギルス (*Aspergillus niger*) のものは醗酵の終了時にそれらのそれぞれの至適の培養条件下で1500~2000cPの次数のものである) ことを決定した。こうした決定に使用された方法を実施例に提供する。

【0048】

粘度は多くの場合に視覚的モニタリングにより評価することができる。物質の流動性は、それがほぼ固体、ソース様もしくは液体であることができるくらい大きな程度まで変動することができる。粘度はまた、ブルックフィールド回転粘度測定、運動学的粘度管の使用、落下球体粘度計もしくはカップ型粘度計により容易に確かめることもできる。こうした低粘度培養物からの収量は、時間単位および細胞あたりで、商業的な既知のより高粘度の培養物からより大きい。

【0049】

本発明のこうした低粘度の培養物の加工は、とりわけ培養物がスケールアップされる場合に有利である。低粘度をもつ主題のクリソスポリウム属 (*Chrysosporium*) の株は、約150,000リットルの培養物までの培養物中で非常に良好に機能する。従って、150,000リットルまでのいかなる培養物の大きさも、本発明の有用な一態様を提供する。いかなる他の慣習的な大きさの醗酵も、本発明の株を用いて良好に実施されるはずである。これの後ろにある論法は、大スケール生産において、濃厚すぎる菌糸体を有しかつ／もしくは不均一に分布される凝集物の形成を伴う問題を生じる可能性があることである。培地は、結果として、培養の間効率的に利用されることができず、従って、とりわけ大スケールの醗酵において、すなわち150,000リットルを越えて、非効率的な生産工程につながる。通気および混合が問題になり、酸素および栄養素の窮乏、そして従って培養の間の生産的バイオマスの低下された濃度およびポリペプチドの低下された収量につながり、かつ／またはより長い醗酵時間をもたらす可能性がある。加えて、高粘度および高剪断は商業的醗酵工程で望ましくなく、そして、現在の商業的工程において、それらは生産制限因子である。全部のこれらの負の局面は、この点で商業的に使用されるトリコデルマ レエセイ (*Trichoderma reesei*)、黒色アスペルギルス (*Aspergillus niger*) および麹菌 (*Aspergillus oryzae*) よりずっとより良好な特徴を表す、すなわちより良好なタンパク質産生レベルお

よび粘度の特性ならびにバイオマスの形状を表す、本発明のクリソスポリウム属 (*Chrysosporium*) の宿主により克服することができる。

【0050】

本発明の前述の態様のいずれかのクリソスポリウム属 (*Chrysosporium*) の株 (前記株は、上に挙げられた炭水化物分解酵素、プロテアーゼ、他の加水分解酵素、酸化還元酵素および転移酵素から選択される真菌酵素の1種もしくはそれ以上の産生をさらに表す) は、本発明のとりわけ有用な一態様を考慮される。最も興味深い産物は、とりわけ、セルラーゼ、キシラナーゼ、ペクチナーゼ、リパーゼおよびプロテアーゼである。しかしながら、中性もしくはアルカリ性の至適の安定性および／または活性、好ましくはアルカリ性の至適の安定性および／もしくは活性を表す1種もしくはそれ以上の真菌酵素の産生を表すクリソスポリウム属 (*Chrysosporium*) の株もまた本発明の態様として有用であり、前記酵素は、炭水化物分解酵素、加水分解酵素およびプロテアーゼ、好ましくは加水分解酵素および炭水化物分解酵素から選択される。非組換えクリソスポリウム属 (*Chrysosporium*) の場合には、こうした酵素は、適しては国際特許出願第WO 98/15633号明細書に開示されるところのセルラーゼ以外である。とりわけ興味深い酵素は、キシラナーゼ、プロテアーゼ、エステラーゼ、 α -ガラクトシダーゼ、 β -ガラクトシダーゼ、 β -グルカナーゼおよびペクチナーゼである。酵素は前述されたものに制限されない。記述中の別の場所の安定性および活性に関する批評はここでもまた妥当である。

10

【0051】

本発明はまた目的のポリペプチドの製造方法も包含し、前記方法は、ポリペプチドの発現および好ましくはその分泌を可能にする条件下で、本発明の態様のいずれかにおいて、宿主株 (例えば、クリソスポリウム属 (*Chrysosporium*)、アスペルギルス属 (*Aspergillus*)、トリコデルマ属 (*Trichoderma*)、ハンセンラ属 (*Hansenula*)、ケカビ属 (*Mucor*)、ピキア属 (*Pichia*)、ニューロスポラ属 (*Neurospora*)、トリポクラジウム属 (*Tolyposcladium*)、リゾムコール属 (*Rhizomucor*)、フサリウム属 (*Fusarium*)、ペニシリウム属 (*Penicillium*) のもののような真菌の、もしくは細菌の、または他の微生物の) を培養すること、ならびにその後目的の産生されたポリペプチドを回収することを含んで成る。

20

30

【0052】

タンパク質もしくはポリペプチドが挙げられる場合、非突然変異体の活性を表す変異体および突然変異体、例えば天然に存在するタンパク質の置換、挿入もしくは欠失突然変異体を包含することを意図している。同じことが対応する核酸配列に関して妥当である。遺伝子シャッフリング、タンパク質工学、ならびに進化分子工学 (*directed evolution*) 部位特異的突然変異誘発およびランダム突然変異誘発のような方法は、それによりこうしたポリペプチド、変異体もしくは突然変異体を得ることができる方法である。米国特許第US 5, 223, 409号、同第US 5, 780, 279号および同第US 5, 770, 356号明細書は進化分子工学の教示を提供する。この方法を使用すれば、例えば *error prone PCR* を介する遺伝子シャッフリングにより創製された無作為に突然変異された遺伝子配列のライブラリーは、いかなる細胞型中でも生じる。各遺伝子は、分泌領域、および生じるタンパク質が分泌されかつ宿主表面に固定されて留まるようなそれに結合された固定化領域を有する。その後、特定のポリペプチドの生物学的活性を必要とする条件が創製される。これが多数の周期の間起こり、最終的に所望の特徴をもつ最終的な遺伝子に至る。言い換えれば進化の加速された指図された過程。米国特許第US 5, 763, 192号明細書もまた、確率的に生成された配列を結合する合成ポリペプチドによってDNA、RNA、ペプチド、ポリペプチドもしくはタンパク質を得る方法、宿主へのその導入、次いで対応する予め決められた特徴をもつ宿主細胞の選択を記述する。

40

【0053】

50

本発明の方法の別の応用は「進化分子工学」の過程にあり、ここで、新規タンパク質をコードするDNA配列が生成され、コードされたタンパク質が宿主細胞中で発現され、そして所望の特徴を有するタンパク質をコードするそれらの配列が突然変異されかつ再度発現される。該過程を、所望の特徴をもつタンパク質が得られるまで多数の周期の間反復する。遺伝子シャッフリング、タンパク質工学、error-prone PCR、部位特異的突然変異誘発、ならびにコンビナトリアルおよびランダム突然変異誘発は、それにより外因性タンパク質をコードする新規DNA配列を生成させることができる方法の例である。米国特許第5,223,409号、同第5,780,279号および同第5,770,356号明細書は進化分子工学の教示を提供する。KuchnerとArnold、Trends in Biotechnology、15:523-530(1997); Schmidt-DannertとArnold、Trends in Biotech.、17:135-136(1999); ArnoldとVolkov、Curr. Opin. Chem. Biol.、3:54-59(1999); Zhaoら、Manual of Industrial Microbiology and Biotechnology、第2版(DemainとDavies編)pp. 597-604、ASM プレス(ASM Press)、ワシントンDC、1999; ArnoldとWintrod、Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis, and Bioseparation、(FlickingerとDrew編)pp. 971-987、ジョン Wiley アンド サンズ(John Wiley & Sons)、ニューヨーク、1999; およびMinschullとStemmer、Curr. Opin. Chem. Biol.、3:284-290もまた参照されたい。

【0054】

コンビナトリアル突然変異誘発の一応用が、Huら、Biochemistry、1998 37:10006-10015に開示される。米国特許第5,763,192号明細書は、合成配列を確率的に生成させること、それらを宿主中に導入すること、および所望の特徴をもつ宿主細胞を選択することによる、新規タンパク質をコードするDNA配列を得る方法を記述する。人工的遺伝子組換えを遂げる方法(DNAシャッフリング)は、ランダムプライミング組換え(Z. Shaoら、Nucleic Acids Res.、26:681-683(1998))、付着伸長プロセス(H. Zhaoら、Nature Biotech.、16:258-262(1998))、およびヘテロ二重鎖組換え(A. Volkovら、Nucleic Acids Res. 27:e18(1999))を包含する。error-prone PCRはなお別のアプローチである(SongとRhee、Appl. Environ. Microbiol. 66:890-894(2000))。

【0055】

進化分子工学の方法での選択段階の2つの広範に実施される実施方法が存在する。一方法において、目的のタンパク質活性は、なんらかの形で宿主細胞の生存に不可欠にされる。例えば、所望の活性がpH8でのセルラーゼ活性である場合、セルラーゼ遺伝子を突然変異させかつ宿主細胞中に導入することができる。形質転換体を、単独の炭素源としてセルロースを用いて成長させ、そして数個のみの生存体が残存するまでpHを徐々に上昇させる。比較的高いpHで活性のセルラーゼをおそらくコードする、生存体からの突然変異されたセルラーゼ遺伝子を、もう1回の突然変異にかけ、そしてpH8でセルロースで成長することができる形質転換体得られるまで該方法を反復する。酵素の熱安定性の変異体を、遺伝子突然変異および宿主細胞の高温培養の周期により、同様に進化させることができる(Liaoら、Proc. Natl. Acad. Sci. USA 83:576-580(1986); Giverら、Proc. Natl. Acad. Sci. USA 95:12809-12813(1998))。

【0056】

大きく平行な「最も適合するものの生存(survival of the fittest)

s t) 」アプローチへの一代替は連続スクリーニングである。このアプローチにおいて、個々の形質転換体を、表示培地上で成長するコロニーの周囲の浄化された (c l e a r e d) もしくは着色された領域の観察、比色もしくは蛍光測定的酵素アッセイ、イムノアッセイ、結合アッセイなどのような伝統的方法によりスクリーニングする。例えば、J o o ら、N a t u r e 399:670-673 (1999) (補助因子としてNADHを必要としないチトクロームP450モノオキシゲナーゼを突然変異およびスクリーニングの周期により進化させた) ; M a y ら、N a t u r e B i o t e c h . 18:317-320 (2000) (逆転された立体選択性のヒダントイナーゼを類似の様式で進化させた) ; ならびにM i y a z a k i ら、J . M o l . B i o l . 297:1015-1026 (2000) (熱安定性のズブチリシンを進化させた) を参照されたい。

10

【0057】

標準的なクローニングおよびタンパク質もしくはポリペプチド単離技術を使用して、必要とされる配列情報に到達することができる。既知の配列の一部を、他の属および株中の他の相同物を単離するためのプローブとして使用することができる。例えば、特定の酵素活性をコードする核酸配列を使用して、クリソスポリウム属 (C h r y s o s p o r i u m) のライブラリーをスクリーニングすることができる。当業者はどのハイブリダイゼーション条件が適切であるかを十分に理解することができる。ライブラリーの核酸ハイブリダイゼーションの構築のための慣習的方法およびクローニング技術は、“M o l e c u l a r C l o n i n g . A L a b o r a t o r y M a n u a l ” コールド スプリング ハーバー (C o l d S p r i n g H a r b o r) 、プレス プレインビュー (P r e s s P l a i n v i e w) 、ニューヨーク中、S a m b r o o k ら (編) (1989) 、およびA u s u b e l ら (編) “ C u r r e n t P r o t o c o l s i n M o l e c u l a r B i o l o g y ” (1987) ジョン ワイリー アンド サンズ (J o h n W i l e y a n d S o n s) 、ニューヨーク中に記述される。関連する情報は、より最近の手引書および特許、ならびに当該分野の多様な商業的に入手可能なキットに由来することもまたできる。

20

【0058】

代替の一態様において、前記方法は、タンパク質もしくはポリペプチドまたはそれらの前駆体の発現および好ましくは分泌を可能にする条件下で本発明の株を培養すること、ならびにその後、産生されたポリペプチドを回収すること、ならびに、場合によっては前駆体を付加的な単離および精製段階にかけて目的のポリペプチドを得ることを含んで成る。こうした方法は、適しては、前駆体の目的のポリペプチドもしくは前駆体への切断段階を含んでよい。切断段階は、例えばプロテアーゼ切断部位が十分に分泌されるタンパク質キャリアおよび目的のポリペプチドを連結する場合は、K e x - 2 様プロテアーゼ、いずれかの塩基性アミノ酸と対にされるプロテアーゼ、もしくはK e x - 2 での切断であることができる。こうしたものについてのコンセンサス配列の詳細が入手可能であり、また、多数の代替 (例えばフリン) が既に開示されているため、当業者はK e x - 2 様プロテアーゼ配列を容易に見出すことができる。

30

【0059】

適しては、本発明の態様のいずれかのポリペプチドの製造方法において、培養は5、好ましくは5-10、より好ましくは6-9より高いpHで起こる。適しては、こうした方法において、培養は、25~43℃の間の温度、好ましくは30~40℃で起こる。本発明の方法で使用される株は、極めて適しては、組換えクリソスポリウム属 (C h r y s o s p o r i u m) 株もしくは他の真菌または非真菌株である。本発明の方法は、こうした場合に、本発明の組換え株の製造の段階によりさらに先行されることができる。適切な条件の選択は発現されるべきポリペプチドの性質に依存することができ、そして、こうした選択は当業者の通常の活動の範囲内に十分にある。

40

【0060】

本発明の組換え株の製造方法もまた主題の発明の一部である。該方法は、異種もしくは同種のポリペプチドをコードする核酸配列を適する宿主株に安定に導入することを含んで成

50

り、前記核酸配列は発現を制御する領域に操作可能に連結され、前記導入は糸状菌の形質転換についてそれ自体既知の様式で起こる。上に述べられたとおり、これについての多数の参考文献が入手可能であり、かつ、少数の選択されたものが引用されている。提供される情報は、当業者が過度の負担なしで該方法を実施することを可能にするのに十分である。該方法は、それ自体もしくは組合せで本発明の組換え株の多様な態様で記述された核酸要素のいずれかを含んで成る核酸配列の導入を含んで成る。

【0061】

例として、導入はプロトプラスト形質転換法を使用して起こることができる。該方法を実施例に記述する。代替のプロトプラストもしくはスフェロプラスト形質転換法が既知であり、かつ、他の糸状菌について従来技術に記述されているとおり使用することができる。こうした方法の詳細は引用された参考文献の多くで見出すことができ、そして従って引用することにより組み込まれる。本発明の方法は、適しては、目的のポリペプチドをコードする所望の配列の導入のための出発原料として非組換え株を使用することを含んで成る。

10

【0062】

主題の発明はまた、クリソスポリウム属 (*Chrysosporium*) の酵素の製造方法も包含し、前記方法は、中性およびアルカリ性の pH 範囲の例として 5、好ましくは 5～10、より好ましくは 6～9、適しては 6～7.5、7.5～9 より高い pH の培地中もしくは培地上でクリソスポリウム属 (*Chrysosporium*) もしくは他の株を培養することを含んで成る。

【0063】

より一般的には、本発明は、中性もしくはアルカリ性の至適活性および／もしくは安定性、好ましくはアルカリ性の至適活性および／もしくは安定性を表す酵素の製造方法をさらに包含する。pH に対する好ましい範囲および至適の活性、ならびにそれを用いてこうしたものを決定するためのアッセイは、記述中の別の場所に提供されている。該酵素は、上述されたとおり、炭水化物分解酵素、プロテアーゼ、他の加水分解酵素、酸化還元酵素および転移酵素から選択されるべきであり、前記方法は、対応する酵素をコードする核酸配列で形質転換もしくはトランスフェクトされた宿主細胞を培養することを含んで成る。適しては、こうした酵素はクリソスポリウム属 (*Chrysosporium*) の酵素であることができる。このような適する一方法は、とりわけセルラーゼ、キシラナーゼ、ペクチナーゼ、リパーゼおよびプロテアーゼの製造を含んで成り、ここで、セルラーゼおよびキシラナーゼは β -1,4-結合を切断し、また、セルラーゼはエンドグルカナーゼ、セロビオヒドロラーゼおよび β -グルコシダーゼを含んで成る。本発明の方法は、こうした前述の酵素をコードする核酸を含んで成る本発明のいずれかのクリソスポリウム属 (*Chrysosporium*) の宿主を培養することを含むことができる。適しては、本発明の非組換えクリソスポリウム属 (*Chrysosporium*) 宿主の製造は、炭水化物分解酵素、加水分解酵素およびプロテアーゼの製造に向けられる。こうした場合には、酵素は適してはセルラーゼ以外である。単離方法は、国際特許出願第 WO 98/15633 号明細書に記述されるものに類似であり、そして引用することにより組み込まれる。

20

30

【0064】

本発明により製造される酵素もまた本発明により包含される。本発明の非組換えクリソスポリウム属 (*Chrysosporium*) の株から単離することができるもののクリソスポリウム属 (*Chrysosporium*) 起源の酵素もまた包含される。それらは前述の安定性、活性の特徴を表す。適しては、それらは LAS の存在下で安定である。とりわけ、前述の安定性、活性の特徴をもつ、pI 4～9.5 をもつプロテアーゼ、25～95 kD の MW をもつプロテアーゼ、4.0 と 9.5 との間の pI をもつキシラナーゼ、25 と 65 kD との間の MW をもつキシラナーゼ、3.5 と 6.5 との間の pI をもつエンドグルカナーゼ、25～55 kDa の MW をもつエンドグルカナーゼ、 β -グルコシダーゼ、4～4.5 の pI をもつ α 、 β -ガラクトシダーゼ、 β -グルコシダーゼ、45～50 kDa の MW をもつ α 、 β -ガラクトシダーゼ、pI 4～5 のセロビオヒドロラーゼ、MW 45～75 kDa のセロビオヒドロラーゼ、例えば 55 kD の MW および p

40

50

I 4.4、4.0～5.0のpIをもつポリガラクトナーゼ、60～70kDa、例えば65kDaのポリガラクトナーゼ、pI 4～5をもつエステラーゼ、ならびに95～105kDaのMWをもつエステラーゼが特許請求される。分子量(MW)はSDS-PAGEにより決定されるものである。非組換えの、すなわち天然に存在する酵素は、国際特許出願第WO 98/15633号明細書に開示されるところのセルラーゼ以外である。上に挙げられたpI値および分子量の組合せをもつ酵素もまた包含される。

【0065】

本発明はまた、本発明の突然変異体(組換え)株による非タンパク質産物の(過剰)産生にも関する。こうした非タンパク質産物は、有機酸、アミノ酸のような一次代謝物、ならびに抗生物質、例えばペニシリンおよびセファロsporin、ならびに他の治療薬のような二次代謝物を包含する。これらの産物は、目的の数種の真菌遺伝子を必要とする生化学的経路の複合の結果である。真菌の一次および二次代謝物、ならびに真菌生物体中でのこれらの代謝物を産生させるための手順は当該技術分野で公知である。一次代謝物の産生の例は、Mattey M.、The Production of Organic Acids、Current Reviews in Biotechnology、12、87-132(1992)により記述されている。二次代謝物の産生の例は、Penalvaら The Optimization of Penicillin Biosynthesis in Fungi、Trends in Biotechnology 16、483-489(1998)により記述されている。

10

【0066】

20

実施例

クリソスポリウム属(*Chrysosporium*)、トリコデルマ属(*Trichoderma*)およびトリポクラジウム ゲオデス(*Tolyposcladium geodes*)を比較した形質転換の例

2種の形質転換されないクリソスポリウム属(*Chrysosporium*)のC1株、および1種のトリコデルマ レエセイ(*Trichoderma reesei*)の参照株を、2種の培地(Gs pH6.8およびプリダム(Pridham)アガー、PA、pH6.8)上で試験した。抗生物質耐性レベルを試験するために、胞子を7日齢のPDAプレートから収集した。選択プレートを32℃でインキュベートし、そして2、4および5日後に記録した(score)。C-1株NG7C-19およびUV18-25は、フレオマイシンおよびヒグロマイシン双方に対する低い基礎の耐性レベルを明白に有したということになった。このレベルは、参照のT. レエセイ(*T. reesei*)の普遍的に使用される実験室株のものに匹敵する。従って、これら2種の標準的な真菌の選択可能なマーカーを、クリソスポリウム属(*Chrysosporium*)の株で十分に使用することができる明白な表示が存在する。他の標準的な真菌の選択可能なマーカーでの問題は期待されないはずである。

30

【0067】

Sh-ble(フレオマイシン耐性)の形質転換されたクリソスポリウム属(*Chrysosporium*)の株の選択を、50μg/mlで成功裏に実施した。これはまたT. レエセイ(*T. reesei*)に使用された選択レベルでもあり、かように、示差的選択がクリソスポリウム属(*Chrysosporium*)において容易に達成することができることを示した。同一の批評が、150μg/mlのレベルでヒグロマイシン耐性をもつ形質転換された株について妥当である。

40

【0068】

プロトプラスト形質転換技術を、最も一般に適用される真菌形質転換技術に基づいて、クリソスポリウム属(*Chrysosporium*)で使用した。1枚の90mm PDAプレートからの全部の胞子を8mlのIC1中に回収し、そして35℃かつ200rpmで15時間のインキュベーションのため、50mlのIC1培地の振とうフラスコに移した。この後、培養物を遠心分離し、ペレットをMnPで洗浄し、10mlのMnPおよび10mg/mlのケイラーゼ(*Calyase*)C₃中の溶液に戻し、そして攪拌(15

50

0 r p m) を伴い 3 5 ° C で 3 0 分間 インキュベートした。

【0069】

該溶液を濾過し、そして濾液を 3 5 0 0 r p m で 1 0 分間の遠心分離にかけた。ペレットを 1 0 m l の $MnPCa^{2+}$ で洗浄した。これを 2 5 ° C で 1 0 分間遠心分離した。その後、5 0 マイクロリットルの冷 M P C を添加した。混合物を氷上で 3 0 分間保ち、その後 2 . 5 m l の P M C を添加した。室温で 1 5 分後に、5 0 0 マイクロリットルの処理されたプロトプラストを 3 m l の MnR S o f t に混合し、そして直ちに、選択剤としてフレオマイシンもしくはヒグロマイシンを含有する MnR プレート上でプレート培養した。3 0 ° C で 5 日間のインキュベーション後に、形質転換を分析した（クローンは 4 8 時間後に見えるようになる）。形質転換効率を、1 0 マイクログラムの参照プラスミド p A N 8 - 1 ^{1 9} を使用して決定した。結果を以下の表 1 に提示する。

【0070】

【表 1】

表1:形質転換効率(10 μ g の参照プラスミド pAN8-1を使用する)

	T. reesei	NG7C-19	UV18-25
生存率	$10^6/200 \mu l$	$5 \cdot 10^6/200 \mu l$	$5 \cdot 10^6/200 \mu l$
200 μ lあたりの形質転換体	2500	10^4	10^4
10^6 個の生存可能な細胞あたりの形質転換体	2500	2000	2000

【0071】

クリソスポリウム属 (*Chrysosporium*) の形質転換体の生存率はトリコデルマ属 (*Trichoderma*) のものより優れている。これらの株の形質転換可能性は比較可能であり、そして、従って、1 回の実験で得られた形質転換体の数は、T . レエセイ (*T. reesei*) についてよりクリソスポリウム属 (*Chrysosporium*) について 4 倍より高く存する。従って、クリソスポリウム属 (*Chrysosporium*) の形質転換系は、普遍的に使用される T . レエセイ (*T. reesei*) 系に同等であるのみならず、しかしそれを凌ぎさえする。この改善は、p A N 8 - 1 より少なく形質転換が効率的であるベクターにとりわけ有用と判明する可能性がある。こうしたより少なく効率的な形質転換ベクターの例は、一般に 1 0 倍より少ない形質転換体を生じさせる非真菌タンパク質の産生のためのタンパク質キャリアベクターである。

【0072】

多数の他の形質転換および発現ベクターを、同種のクリソスポリウム属 (*Chrysosporium*) のタンパク質をコードする配列、およびまたクリソスポリウム属 (*Chrysosporium*) を用いる形質転換実験での使用のために異種タンパク質をコードする配列を用いて構築した。

【0073】

発現系の例は、キシラナーゼの読取り枠と同じ読み枠でキシラナーゼシグナル配列に連結されたクリソスポリウム属 (*Chrysosporium*) のキシラナーゼ X y 1 1 プロモーターフラグメント、次いでキシラナーゼターミネーター配列を包含する。形質転換体の選択は、選択可能なベクターを用いる共形質転換を使用することにより実施する。

【0074】

別の例は、ペニシリウム属 (*Penicillium*) のエンドグルカナーゼ 3 の読取り枠と同じ読み枠のペニシリウム属 (*Penicillium*) のエンドグルカナーゼ 3 のシグナル配列に連結されたクリソスポリウム ルクノウエンセ (*Chrysosporium lucknowense*) のセロビオヒドロラーゼプロモーター、次いでクリソス

ポリウム属 (*Chrysosporium*) のセロビオヒドロラーゼターミネーター配列である。加えて、このベクターは、選択マーカー、すなわちアセトアミダーゼ S 遺伝子 (*AmdS* 遺伝子) を含む第二の発現カセットを担持する。

【0075】

さらなる一例は、黒色アスペルギルス (*Aspergillus niger*) のグルコアミラーゼのシグナル配列に連結されたクリソスポリウム属 (*Chrysosporium*) のグリセルアルデヒド-3-リン酸脱水素酵素 1 のプロモーター、およびヒトインターロイキン 6 の読取り枠に融合されたグルコアミラーゼの読取り枠を含んで成る。加えて、このベクターは、選択マーカー、すなわち *AmdS* 遺伝子を含む第二の発現カセットを担持する。

10

【0076】

なおさらなる一例は、エンドグルカナーゼ 5 の読取り枠に連結されたアスペルギルス ニズランス (*Aspergillus nidulans*) のグリセルアルデヒド-3-リン酸脱水素酵素 A のプロモーター、次いでアスペルギルス ニズランス (*Aspergillus nidulans*) のターミネーター配列である。

クリソスポリウム属 (*Chrysosporium*) の形質転換体の異種および同種発現の例

C1 株 (NG7C-19 および／もしくは UV18-25) を、多様な異種タンパク質、すなわち、細菌タンパク質 (ストレプトアロテイクス ヒンズスタヌス (*Streptoalloteichus hindustanus*) のフレオマイシン耐性タンパク質、*Shble*)、真菌タンパク質 (トリコデルマ レエセイ (*Trichoderma reesei*) のキシラナーゼ II、XYN2) およびヒトタンパク質 (ヒトリゾチーム、HLZ) を分泌するそれらの能力について試験した。

20

トリコデルマ レエセイ (*Trichoderma reesei*) のキシラナーゼ II (XYN2) の C1 分泌。

【0077】

C1 株 UV18-25 をプラスミド pUT1064 および pUT1065 により形質転換した。

【0078】

pUT1064 は 2 種の以下の真菌発現カセットを与える。すなわち、
第一のカセットはフレオマイシン耐性の形質転換体の選択を可能にする：

30

- ・ニューロスポラ クラサ (*Neurospora crassa*) の交差経路制御遺伝子 1 (*cpc-1*) プロモーター^{1 4}

- ・ストレプトアロテイクス ヒンズスタヌス (*Streptoalloteichus hindustanus*) のフレオマイシン耐性遺伝子 *Shble*⁴

- ・アスペルギルス ニズランス (*Aspergillus nidulans*) のトリプトファン合成酵素 (*trpC*) ターミネーター⁵

第二のカセットはキシラナーゼ産生カセットである：

- ・T. レエセイ (*T. reesei*) の TR2 株の *cbh1* プロモーター^{1 5}

- ・T. レエセイ (*T. reesei*) の TR2 株の *xyn2* 遺伝子 (そのシグナル配列を包含する)^{1 6}

40

- ・T. レエセイ (*T. reesei*) の TR2 株の *cbh1* ターミネーター^{1 5}

該ベクターはまた、プラスミド pUC19 からの大腸菌 (*E. coli*) の複製起点も担持する⁶。該プラスミドの詳細された地図を図 1 に提供する。

【0079】

pUT1065 は以下の真菌発現カセットを与える：

- ・A. ニズランス (*A. nidulans*) のグリセルアルデヒド-3-リン酸脱水素酵素 (*gpdA*) プロモーター²

- ・合成の T. レエセイ (*T. reesei*) のセロビオヒドロラーゼ I (*cbh1*) シグナル配列^{1, 3}

50

- ・キャリアタンパク質¹⁰として使用されるS. ヒンズスタヌス (S. hindustanus) のフレオマイシン耐性遺伝子Shble⁴
- ・KEX-2様プロテアーゼ切断部位を特徴とするリンカーペプチド (SGERK)¹
- ・T. レエセイ (T. reesei) のTR2株のxyn2遺伝子 (シグナル配列を含まない)¹⁶
- ・A. ニズランス (A. nidulans) トリプトファン合成酵素 (trpC) ターミナー⁵

該ベクターはまた、 β -ラクタマーゼ遺伝子 (bla)、およびプラスミドpUC18からの大腸菌 (E. coli) の複製起点も担持する⁶。該プラスミドの詳細された地図を図2に提供する。

10

【0080】

C1プロトプラストを、上の例で既に記述されたと同一の手順に従って、プラスミドpUT1064もしくはpUT1065で形質転換した。プラスミドpUT1065中の融合タンパク質 (Shble::XYN2) はフレオマイシン耐性に関して機能的であり、従ってC1形質転換体の容易な選択を可能にする。さらに、フレオマイシン耐性のレベルはxyn2発現のレベルとおおまかに相関する。pUT1064において、xyn2をそれ自身のシグナル配列とともにクローン化した。

【0081】

C1形質転換体 (フレオマイシン耐性クローン) のキシラナーゼ産生を、以下のようなキシラナーゼ活性アッセイにより分析した。すなわち、一次形質転換体を、GS+フレオマイシン (5 μ g/ml) プレート (耐性確認) に、およびまたXYLANプレート (浄化領域の可視化によるキシラナーゼ活性検出¹⁷) 上に、つまようじで移した。プレートを32°Cで5日間成長させた。それぞれの妥当性が確認されたクローンをXYLANプレートにサブクロニングした。形質転換体あたり2個のサブクローンを使用して、液体培養の開始のための胞子を得るためにPDAプレートに接種した。IC1+5g/l フタル酸カリウム中の液体培養物を27°C (180 rpmで振とうする) で5日間成長させた。その後、培養物を遠心分離 (5000g、10分) した。これらのサンプルから、キシラナーゼ活性を、Millerら¹⁸によるDNS技術により測定した。

20

【0082】

【表2】

30

表2: C1(最良の生産体)中の活性のXYN2産生レベル

	培地中の活性の キシラナーゼ II 濃度	培地中の活性の キシラナーゼ II 比活性
未形質転換UV18-25	3.9 U./ml	3.8 U./mg 総タンパク質
UV18-25::1064 クローン7-1	4.7 U./ml	4.7 U./mg 総タンパク質
UV18-25::1064 クローン7-2	4.4 U./ml	4.3 U./mg 総タンパク質
UV18-25::1065 クローン1-1	29.7 U./ml	25.6 U./mg 総タンパク質
UV18-25::1065 クローン1-2	30.8 U./ml	39.4 U./mg 総タンパク質

40

【0083】

これらのデータは：

- 1) 実施例2からの点1ないし4が確認される。
- 2) C1は異種真菌タンパク質の分泌のための宿主として使用することができることを示す。

実施例に対する付録：培地

【0084】

【表3】

形質転換培地:

マンデルス(Mandels)培養基: MnP 培地:

KH ₂ PO ₄	2.0 g/l	以下を含むマンデルス(Mandels)培養基
(NH ₄) ₂ SO ₄	1.4 g/l	ペプトン 1 g/l
MgSO ₄ ·7H ₂ O	0.3 g/l	MES 2 g/l
CaCl ₂	0.3 g/l	シロ糖 100 g/l
微量元素	1.0 ml/l	pHを5に調整する

MnR

MnP+シロ糖

酵母抽出物

ブドウ糖

アガー

MnP CA²⁺:

130 g/l

2.5 g/l

2.5 g/l

15 g/l

MnP培地 +

CaCl₂ 2H₂O

pHを6.5に調整する

50 mM

10

MnR Soft:

7.5g/lのアガーのみを含むMnR

MPC:

CaCl₂

50 mM

pH 5.8

MOPS

10 mM

PEG

40%

20

選択および培養のために

GS:

ブドウ糖

10 g/l

ビオソヤーセ

5 g/l

アガー

15 g/l

[メリュー(Merieux)]

pHは6.8であるはずである

PDA:

ジャガイモD-ブドウ糖アガー

39 g/l

[ディフコ(Difco)]

pHは5.5であるはずである

MPG:

以下を含むマンデルス(Mandels)培養基

フタル酸カリウム

5 g/l

ブドウ糖

30 g/l

酵母抽出物

5 g/l

30

【0085】

50 μg/ml のフレオマイシンもしくは100～150 μg/ml のヒグロマイシンを補充された再生培地 (MnR) を使用して、形質転換体を選択する。5 μg/ml のフレオマイシンを補充されたGS培地を使用して抗生物質耐性を確認する。

【0086】

PDAは迅速な成長および良好な孢子形成のための完全な培地である。液体培地に、1/20の孢子懸濁液(5mlの0.1%トゥイーン(Tween)中の1枚の90mmのPDAプレートからの全孢子)を接種する。こうした培養物を、振とうフラスコ(200rpm)中27℃で成長させる。

C1遺伝子の単離および特徴づけ、ならびにCBH1、XYL1およびGPDの遺伝子発現配列

UV18-25のBlueSTAR遺伝子ライブラリーの構築

UV18-25の染色体DNAをSau3Aで部分的に消化し、12～15kbのフラグメントを単離し、そしてクローニングベクターBlueSTARのBamHI部位に連結した。ライゲーション混合物の20%のパッケージングは、4.6×10⁴個の独立した

40

50

クローンの遺伝子ライブラリーをもたらしした。このライブラリーを増やしそして4℃および-80℃で保存した。ライゲーション混合物の残部もまた4℃で保存した。

c b h 1、x y l 1およびg p d 1の遺伝子の単離のためのU V 1 8 - 2 5の遺伝子ライブラリーのスクリーニング

多様な遺伝子の単離のため、プローブあたり全体で 7.5×10^4 個の別個のB l u e S T A Rファージを二重で(i n d u p l o)ハイブリダイズさせた。ハイブリダイゼーションは、同種条件(65℃; 0.2×SSC)でc b h 1およびx y l 1のPCRフラグメント(国際特許出願第WO 00/20555号明細書に記述されるところの)を用いて、ならびに異種条件(53℃; 0.5×SSC)で黒色アスペルギルス(A. n i g e r)のg p d 1遺伝子を用いて実施した。陽性のシグナルの数を表3に示す。陽性クローンを再スクリーニングし、そして各クローンについて、2個の個別のファージをさらなる実験に使用した。異なるクローンのDNAを制限分析により分析して、各遺伝子から単離される異なるクローンの数を決定した(結果を表3に示す)。

10

【0087】

3種の遺伝子のそれぞれについて4~6個の異なるクローンが単離されたため、われわれは、一次遺伝子ライブラリー($4 \sim 5 \times 10^4$ 個のクローン)がU V 1 8 - 2 5のおよそ5×ゲノムを表すことを結論する。この結果から、われわれは、U V 1 8 - 2 5の完全なゲノムが 9×10^3 個のクローン中に表されることを結論する。13kbという平均のゲノム挿入物に基づけば、これは ± 120 Mbのゲノムの大きさを示すとみられ、それはアスペルギルス属(A s p e r g i l l u s)のゲノムの大きさの3倍である。

20

【0088】

プラスミド上に存在する遺伝子に特異的なプライマー(単離されたPCRフラグメントからの以前に配列決定に基づく)、ならびにp B l u e S T A Rのポリリンカー中に存在するT7およびT3プライマーを用いるPCR反応、われわれは多数のクローン中での該遺伝子の位置を決定することが可能であった。各遺伝子から、プラスミドを該遺伝子の配列決定のために使用した。

【0089】

【表4】

遺伝子	最初のスクリーニングで陽性	再スクリーニングで陽性	異なるクローン	配列決定に使用されたクローン
c b h 1	8	7	4	pCBH7
x y l 1	9	6	5	pXyl5
g p d 1	12	12	6	pGPD4

30

【0090】

表3: c b h 1遺伝子およびx y l 1遺伝子についてのU V 1 8 - 2 5のPCRフラグメント(同種条件)、ならびに黒色アスペルギルス(A. n i g e r)のg p d A遺伝子(異種条件)を用いる、U V 1 8 - 2 5の遺伝子ライブラリーの 7.5×10^4 個のファージのスクリーニング。DNAの単離および制限分析を使用して、異なるクローンの数を決定した。

40

クローン化された遺伝子の配列分析

c b h 1、x y l 1およびg p d 1遺伝子について、配列決定の結果をそれぞれ配列番号1、3および5に表す。また、該タンパク質の推定されるアミノ酸配列をこれらの配列番号2、4および6に表す。該タンパク質のいくつかの特性を表4に示す。翻訳の開始およびイントロンの位置は、同一のファミリーからの遺伝子との相同性に基づく(すなわち論文遺伝学(p a p e r g e n e t i c s))ことが言及されるべきである。

C B H 1

C B H 1のアミノ酸配列から、われわれは、該タンパク質の大きさが約63kDであるこ

50

と、および該タンパク質のC末端部分にセルロース結合ドメイン(CBD)が存在することを結論した。興味深いことに、単離された55kDの主要タンパク質中にCBDの存在の証拠は見出されなかった。しかしながら、コードされるCBH1タンパク質(配列番号1、2)中での、この55kDの主要タンパク質からの単離されたペプチドの存在は、該55kDのタンパク質がクローン化された遺伝子によりコードされることを確認する。これらの結果の可能な説明は、該55kDのタンパク質がCBDを欠くCBH1タンパク質の短縮されたバージョンであることである。

【0091】

セロビオヒドロラーゼCBH1は、MUF-セロビオシド、MUFラクトシド、FPおよびアビセルに対する、また、p-ニトロフェニルβ-グルコシド、セロビオースおよびp-ニトロフェニルラクトシドに対する活性を有する。MUFセロビオシドに対するその活性は、0.4mMの阻害定数を伴い、セロビオースにより阻害される。MUFセロビオシドに対するミカエリス定数は0.14mMであり、MUFラクトシドに対しては4mMであり、そしてCMCに対して3.6g/lであった。pHの至適は4.5から7までである。CMCに対して最大活性の50%、およびRBB-CMCに対して80%の活性がpH8で保たれる。pH5~8以内の70~80%の活性が25時間のインキュベーションの間保たれる。温度の至適は60~70℃である。CBH1は、セロビオヒドロラーゼファミリー7の1メンバーである。本発明の好ましい一態様である対応するCBHプロモーターを配列番号1に示す。

Xy11

Xy11のアミノ酸配列から、われわれは、またここにCBDが、このタンパク質中でN末端側に(すなわち、シグナル配列にもしくはそれから5アミノ酸未満の距離に直接結合されて)存在することを結論した。文献では、CBDをもつもう少しのキシラナーゼのみが既知である(フサリウム オキシスポルム (*Fusarium oxysporum*)、フミコラ グリセア (*Humicola grisea*) およびネオカリマスチクス パトリシアルム (*Neocallimastix patriciarum*))。Xy11タンパク質の推定される大きさは43kDであり、そして30kDのキシラナーゼから単離されるいくつかのペプチドはこのタンパク質から発する(配列番号3、4)。いくつかの単離されたペプチドはコードされる配列中に見出すことができなかった。これは、代替のキシラナーゼタンパク質がUV18-25中に存在することを示す可能性がある。以前の分析で、この30kDのタンパク質中でのCBDの存在の証拠は見出されなかった。また、これらの結果から、われわれは、該タンパク質のCBDがタンパク質分解により切断されることを仮定した。この仮説は今後分析されるであろう(活性、異なるC1株、すなわちC1野性型、NG7C、UV13-6、UV18-25、およびUV18-25のプロテアーゼ突然変異体中の異なるタンパク質のN末端配列および大きさの決定により)。また、酵素活性に対するCBDの存在もしくは非存在の影響を、より詳細に分析する。多様なC1宿主中での完全長の遺伝子の過剰発現を考慮してもよい。

【0092】

セルロース結合ドメイン(CBD)の存在は、この酵素の特定の一特徴である。N末端のCBDを有する唯一の他の既知のファミリー10の解糖酵素(キシラナーゼ)は、フサリウム オキシスポルム (*Fusarium oxysporum*) XylFである。クリソスポリウム ルクノウェンセ (*Chrysosporium lucknowense*) のXy11タンパク質のCBDは、配列: WGQCG GQGWT GPTTC VSGAV CQFVN DWYSQ CV (配列番号4のアミノ酸22-53)を有する。この配列は、米国特許第5,763,254号明細書(ノボ(Novo))に記述されるCBDのコンセンサス配列に従わない。米国特許第5,763,254号明細書のこのコンセンサス配列は: W/Y-G/A-Q-C-G G-Q/I/N-G/N-W/F/Y-S/T/N/Q G-P/A/C-T/R/K-T/C/N-C X-X-G/P-S/T/F/L/A/- -T/K C-V/T/R/E/K-K/Q/A-Q/I-N Q/D/A-W/F/Y-Y-Y/S/H/A-Q C-L/I/Q/V/Tであり、こ

ここでW/YはWもしくはYなどのいずれかを意味し、Xはいかなるアミノ酸も意味し、そして－は非存在を意味する。最も縮重のコンセンサス配列を伴う4つの差異がXy11中に存在し、それは上の配列7中で下線をつけられる。本発明は、従って、このコンセンサスCBDと異なるN末端CBDを有しかつフサリウム オキシスポルム (*Fusarium oxysporum*) キシラナーゼ以外のキシラナーゼに関する。より具体的には、本発明のキシラナーゼは、上の配列7と最低55%、とりわけ最低65%、好ましくは最低75%の配列の同一性を有するCBDを有する。好ましくは、CBDは、位置23のアミノ酸Phe、TyrおよびTrpの1種、もしくは4種のアミノ酸、位置20のVal、位置22のGln、位置23のPheおよび位置24のValの最低1種を含有する。好ましい配列は、Cys-Xaa-Phe、Xaa-Phe-Val、Cys-Xaa-Phe-Val、Cys-Gln-Phe、Val-Cys-Xaa-Phe、Gln-Phe-Val、Gln-Trp-Val、Gln-Tyr-Val、Val-Cys-Gln、Val-Cys-Gln-Phe、およびVal-Cys-Xaa-Phe-Valを含んで成り、ここでXaaはいずれかのアミノ酸、または好ましくはVal、Thr、Arg、Glu、GlnもしくはLys、または最も好ましくはGlnもしくはGluである。

10

【0093】

該キシラナーゼはMUFセロビアーゼ活性を所有せず、そして従って真のキシラナーゼである。それは5から8の広範なpH範囲内で高い活性を所有し、pH9~10で最大活性の65%を維持し；それはキシラナーゼFファミリーの1メンバーである。本発明の好ましい一態様である対応するキシラナーゼプロモーターを配列番号3に示す。カバのキシランに対するミカエリス定数は、30kDのキシラナーゼについて4.2g/lである。温度の至適は高く、そして該キシラナーゼについて70℃に等しい。

20

Gpd1

gpd1遺伝子のC末端部分のDNA配列は決定されていない。この遺伝子のプロモーター配列は本発明の好ましい一態様であり、そして配列番号5に描かれる。

【0094】

4種のクリソスポリウム属 (*Chrysosporium*) の遺伝子の発現レベルをノーザン分析により研究した。多様な株のC. ルクノウェンセ (*C. lucknowense*) を、セルロースおよび乳糖とともにファーマディア (pharmedia) を含有する豊富培地 (培地1)、もしくはファーマディアおよびブドウ糖を含有する豊富培地 (培地2) 中、33℃で成長させた。48時間後に、菌糸体を収穫しかつRNAを単離した。RNAを、4種の異なるプローブ、すなわちEG5、EG6、Xy11およびCBHとハイブリダイズさせた。曝露後に、ノーザンブロットを細長い形に切り、そしてブロット上のmRNAの量に対する対照としてリボソームL3に対するプローブと再度ハイブリダイズさせた。大部分の株は、培地1中でCBHに対する非常に高い応答、およびXy11に対する高い応答を示し；培地2中では、株の半分が全部の遺伝子について高い応答を示し、また、他の半分が低い応答を示した。発現の強さの順序は、これらのデータからCBH>Xy11>EG5>EG6と推定された。

30

【0095】

C. ルクノウェンセ (*C. lucknowense*) のタンパク質Xy11は、ジェンバンク (Genbank) データベース (DATABASE) 中のその最も近い相同物に67%同一 (72%相同) である (表4)。その関係するフミコラ グリセア (*Humicola grisea*) 相同物に対するCBH1タンパク質の強い相同性 (74%同一/82%相同) は注目すべきである。全部の場合で、最も近い相同物は、フサリウム属 (*Fusarium*)、フミコラ属 (*Humicola*) もしくは他の核菌類 (*Sordariomycetous*) の真菌から発する (表4) 一方、クリソスポリウム属 (*Chrysosporium*) はNCBIのデータベースによれば不整子囊菌類 (*Eurotiomycetous*) の真菌に属する (表4) こともまた注目される。従って、本発明はまた、不整子囊菌類の真菌由来のグリカン分解 (glycanolytic) 酵素、とりわ

40

50

けCBDを含んで成るセロビオヒドロラーゼおよびキシラーゼにも関する。

【0096】

【表5】

	グリコ シターゼ ファミリー	C1から 単離する	アミノ酸の 数	イントロン	備考	関係する配列 (%同一性/%相同性)
CBH1	7	70kD 55kD	526 (63kD)	1	CBD	<i>Humicola grisea</i> (74 / 82) (CBH1: P15828) <i>Fusarium oxysporum</i> (58 / 68) (CBH: P46238) <i>Neurospora crassa</i> (60 / 69) (CBH1: P38676)
XYL1	10	30kD	333 (43kD)	3	CBD	<i>Fusarium oxysporum</i> (67 / 72) (XylF: P46239) <i>Penicillium simplissicum</i> (63 / 72) (XylF: P56588) <i>Aspergillus aculeatus</i> (61 / 70) (XylF: O59859)
GPD1	-	-	不完全	2+?	-	<i>Podospira anserina</i> (85/89) (GPD: P32637) <i>Neurospora crassa</i> 80/86) (GPD: U67457) <i>Cryphonectria parasitica</i> 80/85) (GPD: P19089)

表4: 本発明のCBH1,XYL1,およびGPDの構造および比較データ

【0097】

参考文献（これらの内容は組み込まれる）

【0098】

【表6】

1. Calmels T.P., Martin F., Durand H., and Tiraby G. (1991) *Proteolytic events in the processing of secreted proteins in fungi*. J Biotechnol 17(1): p. 51-66.
2. Punt P.J., Dingemanse M.A., Jacobs-Meijnsing B.J., Pouwels P.H., and van den Hondel C.A. (1988) *Isolation and characterization of the glyceraldehyde-3-phosphate dehydrogenase gene of Aspergillus nidulans*. Gene 69(1): p. 49-57.
3. Shoemaker S., Schweickart V., Ladner M., Gelfand D., Kwok S., Myambo K., and Innis M. (1983) *Molecular cloning of exo-cellobiohydrolase I derived from Trichoderma reesei strain L27*. Bio/Technology Oct.:691-696.
4. Drocourt D., Calmels T., Reynes J.P., Baron M., and Tiraby G. (1990) *Cassettes of the Streptoalloteichus hindustanus ble gene for transformation of lower and higher eukaryotes*

【0099】

【表7】

- to phleomycin resistance. *Nucleic Acids Res* **18**(13): p. 4009.
5. Mullaney E.J., Hamer J.E., Roberti K.A., Yelton M.M., and Timberlake W.E. (1985) *Primary structure of the trpC gene from Aspergillus nidulans*. *Mol Gen Genet* **199**(1): p. 37-45.
 6. Yanisch-Perron C., Vieira J., and Messing J. (1987) *Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors*. *Gene* **33**:103-119.
 7. Durand H., Baron M., Calmels T., and Tiraby G. (1988) *Classical and molecular genetics applied to Trichoderma reesei for the selection of improved cellulolytic industrial strains*, in *Biochemistry and genetics of cellulose degradation*, J.P. Aubert, Editor. Academic Press. p. 135-151.
 8. Lowry O.H., Rosebrough N.J., Farr A.L., and Randall R.J. (1951) *Protein measurements with the folin phenol reagent*. *J. Biol. Chem.* **193**: 265-275. 10
 9. Parriche M., Bousson J.C., Baron M., and Tiraby G. *Development of heterologous protein secretion systems in filamentous fungi*. in *3rd European Conference on Fungal Genetics*. 1996. Münster, Germany.
 10. Baron M., Tiraby G., Calmels T., Parriche M., and Durand H. (1992) *Efficient secretion of human lysozyme fused to the Sh ble phleomycin resistance protein by the fungus Tolypocladium geodes*. *J Biotechnol* **24**(3): p. 253-266.
 11. Jeenes D.J., Marczinke B., MacKenzie D.A., and Archer D.B. (1993) *A truncated glucoamylase gene fusion for heterologous protein secretion from Aspergillus niger*. *FEMS Microbiol. Lett.* **107**(2-3): p. 267-271. 20
 12. Stone P.J., Makoff A.J., Parish J.H., and Radford A. (1993) *Cloning and sequence-analysis of the glucoamylase gene of neurospora-crassa*. *Current Genetics* **24**(3): p. 205-211.
 13. Mörsky P. (1983) *Turbidimetric determination of lysozyme with Micrococcus lysodeikticus cells: Reexamination of reaction conditions*. *Analytical Biochem.* **128**:77-85.
 14. Paluh J.L., Orbach M.J., Legerton T.L., and Yanofsky C. (1988) *The cross-pathway control gene of Neurospora crassa, cpc-1, encodes a protein similar to GCN4 of yeast and the DNA-binding domain of the oncogene v-jun-encoded protein*. *Proc. Natl. Acad. Sci. USA* **85**(11): p. 3728-32.
 15. Nakari T., Onnela M.L., Ilmen M., Nevalainen K., and Penttilä M. (1994) *Fungal promoters active in the presence of glucose*, WO 94/04673, Alko. 30
 16. Torronen A., Mach R.L., Messner R., Gonzalez R., Kalkkinen N., Harkki A., and Kubicek C.P. (1992) *The two major xylanases from Trichoderma reesei: characterization of both enzymes and genes*. *Biotechnology (N Y)* **10**(11): p. 1461-5.
 17. Farkas V. (1985) *Novel media for detection of microbial producers of cellulase and xylanase*. *FEMS Microbiol. Letters* **28**:137-140.
 18. Miller G.L. (1959) *Use of dinitrosalicylic acid reagent for determination of reducing sugar*. *Anal. Chem.* **31**:426-428.
 19. Punt P.J., Mattern I.E., van den Hondel C.A.M.J.J. (1988) *A vector for Aspergillus transformation conferring phleomycin resistance*. *Fungal Genetics Newsletter* **35**, 25-30. 40

【 0 1 0 0 】

【 配列表 】

【 0 1 0 1 】

【 表 8 】

配列番号1:

プロモーターおよびターミネーター配列を含む完全なクリソスポリウム(*Chrysosporium*) CBH1遺伝子のDNA配列およびアミノ酸。プロモーター配列(1-1779)、ターミネーター配列(3427-4451)およびイントロン配列(2179-2256)は小文字で示されている。

```

aaggatatccgatttggggaaacgtcgatgaaagtattgcaaaagtgcgagaggttgcgcaa 60
ctaactcgctgccgaagaagctgcggaagaaagagaacaccgaaagtgaataacgttac 120
ggatgtcctgacctcaaagtggaaaccagcccttctctatttgggaaagcggttg 180
cccttgaatgcgctgcactgtggcacgactaccagtgatcgggaggagcaaacctaccctg 240
gtccgttcccttggtggggcgccactagggcccaacttaggggtgatcggagggtcgatgccgc 300
ggctcctcggttggtctgggctcttctcatttcccggtttgcacccccgttgcaacctgctg 360
atcgcccgccaacgcgatgaggttgccgccagaccgacaatcaccgcgggtgcattccc 420
aagtatatggaagatggcaccaggtaccgggttttgcggtccagtcggttggtgccaat 480
ttgggagtttttgagcctcaagatctggggaaatcgacctcaacttccatacaagttaaa 540
gtcgcacacacggcgagttccacgaagagacacattttttctgaaggcctctctcccg 600
cacatcagaaaccaccaaataccaagactgcagaagccgggtaagtgggccaccgggac 660
tacactaaaatgcggggagaagcgagatccgttgcaagggaaggatggggtgtgctgc 720
ggcttctcctcgctctcggtgcgcttttgccttgaatctagtgtacaccagggttaggctccg 780
aaggagtatctacggcagcgctgttcgtgctgcgttgagagtcagggcgagacgagcag 840
gcgacaggagcctgcacccggcacttcggatcgcatctgcgcggagcgtcaaatacgctc 900
ttctgcggtcatcagagagcatcgtaaccaagggttcttcgcgaggcgccctgggcttc 960
gcagagtcgcactcggcggaacgccttccgtgtcaccctgataacctggctgcgcgcgcc 1020
agactcctccaatgaggtgtgtggttgccctgcgcgacccttcagcaaccttaategett 1080
ccatgcacggctccacgtcctcgaacgatgcctcagtcggtgcgcggccgtggcaacc 1140
ataacgtgacatcgccgcccagcctactagccgctatcgaccggttaggcttgtcaccgc 1200
agcgcccatcttccatcgggcctctactctlgatccacctcaccacccgcaagcactagcg 1260
agcctcaccagagtgcgaagcgacacgaccgcgttgcccttcgtccttgactatctcca 1320
gacctcttgccatcttgcgagcgccgcccctttttttctctccctccctgcgcgcaggt 1380
cggtggccccagtcccgagatggcattgctcggttgtccatgacgacccatcattcgatg 1440
gctgactggcacactcgtcttgtttgagcatcgacggcccgcgcccggtctccacggta 1500
cggaacctcggtgttacagtacctctcgtaatgatacccaacacccggggcgagcgtggg 1560
agggcgcgcttcccgagaagccgggaaggcggtggccgggtgaccttgtgacttggcg 1620
atggatgcggccatggagaatgtccgtccgaagcgacgcgacaattagcctggctaccat 1680

```

10

20

30

【0102】

【表9】

40

cgatataaaattgggtgattcccagctcttgatgggctgtcttctgctggcagccctcg 1740
 tcttcagatcaagcaactgtgtgctgatcctcttccgccATGTACGCCAAGTTCGCGACC 1800
 M Y A K F A T
 CTGCCCGCCCTTGTGGCTGGCGCCGCTGCTCAGAAGCCTGCACTCTGACCGCTGAGAAC 1860
 L A A L V A G A A A Q N A C T L T A E N
 CACCCCTCGCTGACGTGGTCCAAGTGCAGCTCTGGCGGCAGCTGCACCAGCGTCCAGGGT 1920
 H P S L T W S K C T S G G S C T S V Q G
 TCCATCACCATCGACGCCAACTGGCGGTGGACTCACCGGACCGATAGCGCCACCAACTGC 1980
 S I T I D A N W R W T H R I D S A T N C
 TACGAGGCCAACAAGTGGGATACTTCGTAAGTGCAGCGATGGTCTTCTTGGCGCTCCAAG 2040
 Y E G N K W D T S Y C S D G P S C A S K
 TGCTGCATCGACGGCGCTGACTACTCGAGCACCTATGGCATCACACGAGCGCTAACTCC 2100
 C C I D G A D Y S S T Y G I T T S G N S
 CTGAACCTCAAGTTCGTACCAAGGGCCAGTACTCGACCAACATCGGCTCGCGTACCTAC 2160
 L N L K F V T K G Q Y S T N I G S R T Y
 CTGATGGAGAGCGACACCAAGTACCAGAgtaagttcctctcgacccggcgccggggaga 2220
 L M E S D T K Y Q M
 tgatggcgcccagcccgtgacgcgaatgacacaGTGTTCCAGCTCCTCGGCAACGAGTT 2280
 F Q L L G N E F
 CACCTTCGATGTCGACGTCTCCAACCTCGGCTGCGGCCCTCAATGGCGCCCTCTACTTCGT 2340
 T F D V D V S N L G C G L N G A L Y F V
 GTCCATGGATGCCGATGGTGGCATGTCCAAGTACTCGGGCAACAAGGCAGGTGCCAAGTA 2400
 S M D A D G G M S K Y S G N K A G A K Y
 CGGTACCGGCTACTGTGATTCTCAGTGCCCCCGCGACCTCAAGTTCATCAACGGCGAGGC 2460
 G T G Y C D S Q C P R D L K F I N G E A
 CAACGTAGAGAAGTGGCAGAGCTCGACCAACGATGCCAACGCCGGCACGGGCAAGTACGG 2520
 N V E N W Q S S T N D A N A G T G K Y G
 CAGCTGCTGCTCCGAGATGGACGTCTGGGAGGCCAACAACATGGCCGCCGCTTCACTCC 2580
 S C C S E M D V W E A N N M A A A F T P
 CCACCCCTTGACCGGTGATCGGCCAGTCGCGCTGCGAGGGCGACTCGTGCGGCGGTACCTA 2640
 H P C W V I G Q S R C E G D S C G G T Y
 CAGCACCGACCGCTATGCCGGCATCTGCGACCCCGACGGATGCGACTTCAACTCGTACCG 2700
 S T D R Y A G I C D P D G C D F N S Y R
 CCAGGGCAACAAGACCTTCTACGGCAAGGGCATGACGGTCGACACGACCAAGAAGATCAC 2760
 Q G N K T F Y G K G M T V D T T K K I T
 GGTGTCACCCAGTTCTCAAGAACTCGGCCGGCGAGCTCTCCGAGATCAAGCGGTTCTA 2820
 V V T Q F L K N S A G E L S E I K R F Y
 CGTCCAGAACGGCAAGGTATCCCCAACTCCGAGTCCACCATCCGGGGCGTCGAGGGCAA 2880
 V Q N G K V I P N S E S T I P G V E G N
 CTCCATCACCCAGGACTGGTSCGACCGCCAGAAGGCCGCTTCGGCGACGTGACCGACTT 2940
 S I T Q D W C D R Q K A A F G D V T D ?

10

20

30

【0103】

【表10】

NCAGGACAAGGGCGGCATGGTCCAGATGGGCAAGGCCCTCGCGGGGGCCATGGTCCTCGT 3000
 Q D K G G M V Q M G K A L A G P M V L V
 CATGTCCATCTGGGACGACACGCCGTCAACATGCTCTGGCTCGACTCCACCTGGCCCAT 3060
 M S I W D D H A V N M L W L D S T W P I
 CGACGGCGCGGCAAGCCGGGCGCGAGCGCGGTGCCTGCCCCACCACTCGGGCGTCCC 3120
 D G A G K P G A E R G A C P T T S G V P
 CGCTGAGGTGCGAGGCGGAGGCCCACTCCAACGTATCTTCTCCAACATCCGCTTCGG 3180
 A E V E A E A P N S N V I F S N I R F G
 CCCCATCGGCTCCACCGTCTCCGGCCTGCCCGACGGCGGCGAGCGGCAACCCCAACCGCC 3240
 P I G S T V S G L P D G G S G N P N P P
 CGTCAGCTCGTCCACCCCGGTCCCCTCCTCGTCCACCACATCCTCCGGTTCCTCCGGCCC 3300
 V S S S T P V P S S S T T S S G S S G P
 GACTGGCGGCACGGGTGTGCGTAAGCACTATGAGCAATGCGGAGGAATCGGGTTCAGTGG 3360
 T G G T G V A K H Y E Q C G G I G F T G
 CCCTACCCAGTGCAGAGCCCCCTACACTTGCACCAAGCTGAATGACTGGTACTCGCAGTG 3420
 P T Q C E S P Y T C T K L N D W Y S Q C
 CCTGTAAacgaacctctctgaaggagggttctgagacacgcgcgattcttctgtatatagt 3480
 L *
 tttatcttttcaactctggagtgtctcgctccaccagtagacataaacctttttttcacgtaa 3540
 caaaatggcttcttttcagaccatgtgaaccatcttgatgccttgacctcttcagttctc 3600
 actttaacgtagtttcgcgttagtctgtatgtcccagttgcatgtagttgagataaaatacc 3660
 cctggaagtgggtctgggctttgtgggacggagccctcttctgtggtctggagagccc 3720
 gctctctaccgcctaccttcttaaccacgtacactactcacacattgctgaactgaccca 3780
 tcataccgtactttatcctgttaattcgtgggtgctgtcgactattctatttgcataatg 3840
 gagagcacattcatcggcgcagggatacacggtttatggaccccaagagtgtaggacta 3900
 ttattagtaatatattatgctcttaggcgccttaacttcaacaggcgagcactactaate 3960
 aacttttggttagacccaattacaaacgaccatacgtgcgggaaattttgggattccgtcc 4020
 gctctccccaaccaagctagaagaggcaacgaacagccaatcccggtgctaattaaatta 4080
 tatgggtcatttttttaaaaaattttttcttccattttctctctcgctttttttttc 4140
 gcatcgtagttgatcaaagtccaagtcaagcgagctatttgtgctatagctcggtggtc 4200
 taatcagtagcagcttagagaggctgtaaagggtatgataccacagcagtagtattcgcgctata 4260
 agcggcactcctagactaattgttacggtctacagaagtaggtaataaaagcggttaattg 4320
 ttctaaatactagaggcacttagagaagctatctaaatatattgacctagcttatta 4380
 tccctattagtaagttagttagctctaacctatagatagccaaatgctataataggtacc 4440
 agggttcaaaa 4451

10

20

30

【 0 1 0 4 】

【 表 1 1 】

配列番号2:

完全なクリソスポリウム(Chrysosporium)CBH1タンパク質のアミノ酸。推定シグナルペプチド(1-19)はイタリック活字体で示されており、セルロース結合ドメイン(496-526)は下線を引いた肉太活字体で示されている。

```

MYAKFATLAA LVAGAAQNA CTLTAENHPS LTWSKCTSGG SCTSVQGSIT 50
IDANNRWTHR TDSATNCYEG NKWDTSYCSD GPSCASKCCI DGADYSSTYG 100
ITTSGNLSNL KFVTKGQYST NIGSRTYLME SDTKYQMFQL LGNEFTFDVD 150
VSNLGCGLNG ALYFVSMDDAD GGMSKYSNGK AGAKYGTGYC DSQCPRDLKF 200
INGEANVENW QSSTNDANAG TGKYGSCCSE MDVWEANNMA AAFTHPC?V 250
IGQSRCEGDS CGGTYSTDRY AGICDPDGGD FNSYRQGNKT FYGKGMTVDT 300
TKKITVVTQF LKNSAGELSE IKRFYVQNGK VIPNSESTIP GVEGNSITQD 350
WCDRQKAAFG DVID?QDKGG MVQMGKALAG PMVLVMSIWD DHAVNMLWLD 400
STWPIDGAGK PGAERGACPT TSGVPAEVEA EAPNSNVIFS NIRFGPIGST 450
VSLGPDGSG NPNPPVSSST PVPSSSTSS GSSGPTGGTG VAKHYEQCGG 500
IGFTGPTQCE SPYTCTKLND WYSQCL * 526

```

10

配列番号3:

プロモーターおよびターミネーター配列を包含する完全なクリソスポリウム(Chrysosporium)Xyl1遺伝子のDNA配列およびアミノ酸。プロモーター配列(1-969)、ターミネーター配列(2428-3030(3028))およびイントロン配列(1043-1116,1181-1332(1331),1596(1595)-1674(1672))は小文字で示されている。

```

tcacacacttggcggtttggatgtactaatattacacgctcgtttgcnnagcggaggtctgtg 60
tcacatccggtggggtcggtgctccagacgacgcttcgggcccgatcctgaattcggaag 120
gaaacgggttcggctaatacaggtcctctaaaatataacgaagcactacagagggagttcct 180
cagaggacatcgatatcaaccgaagaacgaagcgccgaaaggactgatcaaacaggagta 240
ggtagggtgtgtgtagtacctaactttccataacctgacataaaaatcatcatggtgcttc 300
agacctgtttgatgaggcgagggcgaggcgccgcatgtatttttggttccttcctctcttt 360
tgttagtatatctnagggttccatcgtaaaatggaatcttcagctctactagtaattag 420
aacaatagttctgtatgtcgtgcgccaagctttttcagatgactgccaaaaaccatcatg 480
ggtatggacaaaagcagtaatacggagtcacaacgcccattttccttcattgatttccgtc 540
aaccggagaggtcgaggaggactccggccacatgtgatgcgaagaagtacatggcgcca 600
tggttctaacctcttatagtcgaaaatgcgaggaggccagcgaagccaagcccgggaac 660
cgttcttgatcgtgtttcagtaattgtttcgctaaacattctatccgattcgcgataggtg 720
cggctgccacgaagggttgatatccttaaagctttggttaagtacggagtaacggaatggaa 780

```

20

30

【 0 1 0 5 】

【 表 1 2 】

40

acgcgcgcagtcctggttccatcggtatcctccgcatgctccgcaaaaaaagaaaacc	840	
cggttatgtttacaaaggatataagagacaagatgcaccaccgcccccttcccatctgc	900	
cggttgcacacgtcgccgtcgactgcttgctcgcttctactctgcagcctctttcagaga	960	
ccatcaaacATGCGTACTCTTACGTTCGTGCTGGCAGCCGCCCCGGTGGCTGTGCTTGCC	1020	
M R T L T F V L A A A P V A V L A		
CAATCTCTCTGTGGGGCCAGTgtatglaattgccttactcggaataatagtcaccactag	1080	
Q S P L W G Q C		
agggacttaagctcactacttctgttttcacaatagGCGGCGGTCAAGGCTGGACAGGTC	1140	
G G Q G W T G		
CCACGACCTGCGTTTCTGGCGCAGTATGCCAATTCGTCAAGtcagtaactgcttttatt	1200	10
P T T C V S G A V C Q F V N		
tcttttctctctgggattacgatttctgttttgcacttagcttggttctgcatttcattgt	1260	
tgtattgttctctttttgtgtgtgagaggttttattaccacctaagggcatttgctaac	1320	
aaatctctccacagTGACTGGTACTCCCAATGCGTGCCCGGATCGAGCAACCCTCTACGGG	1380	
D W Y S Q C V P G S S N P P T G		
CACCACCAGCAGCACCCTGGAAGCACCCCGGCTCCTACTGGCGGCGGCGGAGCGGAAC	1440	
T T S S T T G S T P A P T G G G S G T		
CGGCCTCCACGACAAATTCAAGGCCAAGGGCAAGCTCTACTTCGGAACCGAGATCGATCA	1500	
G L H D K F K A K G K L Y F G T E I D H		
CTACCATCTCAACAACAATGCCTTGACCAACATTGTCAAGAAAGACTTTGGTCAAGTCAC	1560	20
Y H L N N N A L T N I V K K D F G Q V T		
TCACGAGAACAGCTTGAAGTGGGATGCTACTGAGCGtgagtgacctctctctctctctccc	1620	
H E N S L K W D A T E P		
gacaataatagataattacgagccggttcgagggctgacattgcgcgattctagCGAGCC	1680	
S R		
GCAATCAATTCAACTTTGCCAACGCCGACGCGGTTGTCAACTTTGCCAGGCCAACGGCA	1740	
N Q F N F A N A D A V V N F A Q A N G K		
AGCTCATCCGCGGCCACACCTCTCTGGCACTCTCAGCTGCCGCAAGTGGGTGCAGAACA	1800	
L I R G H T L L W H S Q L P Q W V Q N I		
TCAACGACCGCAACACCTTGACCCAGGTCATCGAGAACCACGTACACACCTTGTCACTC	1860	
N D R N T L T Q V I E N H V T T L V T R		
GCTACAAGGGCAAGATCCTCCACTGGGACGTCGTTAACGAGATCTTTGCCGAGGACGGCT	1920	30
Y K G K I L H W D V V N E I F A E D G S		
CGCTCCGCGACAGCGTCTTCAGCCGCGTCTCGGCGAGGACTTTGTCCGCATCGCCTTCC	1980	
L R D S V F S R V L G E D F V G I A F R		
GCGCCGCCCGCGCCGCGATCCCAACGCCAAGCTCTACATCAACGACTACAACCTCGACA	2040	
A A R A A D P N A K L Y I N D Y N L D I		
TTGCCAACTACGCCAAGGTGACCCGGGGCATGGTCGAGAAGGTCAACAAGTGGATCGCCC	2100	
A N Y A K V T R G M V E K V N K W I A Q		

【0106】

【表13】

AGGGCATCCCGATCGACGGCATCGGCACCCAGTGCCACCTGGCCGGGCGGCGGGTGGG 2160
 G I P I D G I G T Q C H L A G P G G W N
 ACACGGCGCGCGCGCTCCCGACGCCCTCAAGGCCCTCGCCGCGGCCAACGTCAAGGAGA 2220
 T A A G V P D A L K A L A A A N V K E I
 TCGCCATCACCGAGCTCGACATCGCCGGCGCCTCGGCCAACGACTACCTCACCGTCATGA 2280
 A I T E L D I A G A S A N D Y L T V M N
 ACGCCTGCCTCCAGGTCTCCAAGTGCCTCGGCATCACCGTCTGGGGCGTCTCTGACAAGG 2340
 A C L Q V S K C V G I T V W G V S D K D
 ACAGCTGGAGSTCGAGCAGCAACCCGCTCCTCTCGACAGCAACTACCAGCCAAAGGCGG 2400
 S W R S S S N P L L F D S N Y Q P K A A
 CATACAATGCTCTGATTAATGCCTTGTAAGaggagggtatattatttttagaggcaatgaa 2460
 Y N A L I N A L *
 gctaggaggaaaagaggggaagtgaggtaattagctaggacaggcaaatctagcagcaatt 2520
 ataagtcaacactatataaaaaatttcctataatggcttggtgcttcggtgtgcaaaaaaa 2580
 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaactcaaaaacaaaaatgatccaacatgatt 2640
 cgaaatggcgaccttgcaaatgcacacctcagataataccactatacaatacaccttaaa 2700
 tggcacctaaatccatttctgtctgcggtcatagacggggcttaagaagcctgggatgcagg 2760
 tctcgatgcaaggggttacgtcagtgatgatagagtgaaccatgctgtctgggtaat 2820
 tctccactttccctcccttacgactcttcgggtgtgcctctctagaaagtcgactcctg 2880
 gcgcctcagatcgccctttggctctgttcgggtacaatgacgtccgctgggtttcttccaaa 2940
 gaccaggtattttctcccgctggcaacaaagaataccaaatacctatatcgaaaccgtagtct 3000
 tctgataattagatgtctctcaaggcgcgg 3030

10

20

配列番号4

完全なクリソスリウム(*Chrysosporium*)XylIIタンパク質のアミノ酸配列。シグナルペプチド
 (1-20)はイタリック活字体で示されており、セルロース結合ドメイン(22-53)は下線を引いた
 肉太活字体で示されている。

1 *MRTLTFVLAA* *APVAVLAQSP* **LWGQCGGQGW** **TGPTTCVSGA** **VCQFVNDWYS**
 51 **QCV**PGSSNPP TGTTSSTGS TPAPTGGGGS GTGLHDKFKA KGKLYFGTEI
 101 DHYHLNNNAL TNIVKKDFGQ VTENSLKWDA TEPsrNQFNF ANADAVVNFA
 151 QANGKLIRGH TLLWHSQLPQ WVQNINDRNT LTQVIENHVT TLVTRYKGKI
 201 LHWQVVNEIF AEDGSLRDSV FSRVLGEDFV GIAFRAARAA DPNAKLYIND
 251 YNLDIANYAK VTRGMVEKVN KWIAQGIPID GIGTQCHLAG PGGWNTAAGV
 301 PDALKALAAA NVKEIAITEL DIAGASANDY LTVMNACLQV SKCVGITVWG
 351 VSDKDSWRSS SNPLLFDSNY QPKAAYNALI NAL*

30

【0107】

【表14】

配列番号5

プロモーター配列を包含する部分クリソスホリウム(*Chrysosporium*)GPD1遺伝子のDNA配列。
 プロモーター配列(1-1555)およびイントロン配列(1682-1781)は小文字で示されている。
 遺伝子の3'末端は欠失している。

```

tgagcagcaatgagcagcaatgagcattcctgggccaccgagtgctgagtgccagtacgga 60
gtatcgtacttcgtaccggggtttgatttggtgaacggtgcttttcacctctcgatgccc 120
aaatcgggtctaagctgagtttgatcaaatatgtgactccaacatcgcccccttcggcaa 180
accccgctcgacacgtgtgtcatccttccattgcaagcgatcactcgaggcggtgacgat 240
gaacgagatttttgccggaccgattcgcggtatagcggcagccgaccagccctaccac 300
actgatggccgtgtcactagtgtatgtctccagaaccgcaagcatacactgggcaatgct 360
tggtatgcagttgaggcagctttatgtttccatacccttccacttcggctcggggactcg 420
gcggggctcgcggaagtttgacggcagccgctcgggccttaggcgagattaccgtggtt 480
ggcccagtttttagccgttcccgctcgttttccaccggaccatgattttcgtgaaccatt 540
caatcccgaaagcgcatttccgacgttaaggagttacctccgctgccagaattcatgatc 600
gtggccggctcaaggcagcgtggcggggcatccgtgtcaagctcccaggaggaggtgcgc 660
gatttcaaatccggggccaaaacaggccaagactggctggccaaaaaaggagcgtagacg 720
gcccgggacatcggaagtcagctcgagccacccaaaaccggctccgatctactcgcttac 780
tgtggtagttcaggtacttttgagtagtaaaaacgctacggcagggccggggggttcccc 840
ggtgacggaggtgcctctgcggtggcgaacatcccacgcactctcgagctacggtgacac 900
ctcgtgtcctgttggtcttgcaatgctggggcggcaggaatgcgtcgcgctcctcccg 960
ccaagacctaaaacagacagcgccgcaaagtcgctcactagcaccgcgaaacgaagatgc 1020
cccacctcaacgcaatctgtgatgcaagcaattgggaaggctcaccacacctcagcgagg 1080
ggctcaaccatttttattatcagctcatgccaccacaacatgactgttttctttccttgc 1140
tcatccacatttgacaaaaatcgtcgattaatatctttccatacaggccgtccgcgctc 1200
tgataaccacataaaagtctcttcagtcacacagctcaaagctccctcatccctccaggta 1260
agcagccaaaagagctcccccacggaccccgactgcctcatccgcctgtatcggaactg 1320
cgcgacccagcagagaatcccaaacccttgctgcttgctgcccgggtccggactgagctg 1380
caacccaagcctttaaagccttttcccttctccacgggtgtcaactctgtcctatccct 1440
ccgacatccgttgagctcaacaactccccgaaccttttaccggcgccgagctacccctc 1500
catcaaaccacccctgacagctcgctcactcacctccccacatcacagaaatcaaaATGAC 1560
M T -
TATCAAGGTCGGCATCAACGGTTTCGGCCGTATCGGCCGTATCGTCTCCGCAACTCCAT 1620
I K V G I N G F G R I G R I V F R N S I -
CGAGCACTCGGATGTCTGAGATCGTTGCCGTCAACGACCCCTTCATTGAGCCCAAGTACGC 1680

```

10

20

30

40

【0108】

【表15】

E H S - D V E I V A V N D P F I E P K Y A -
 Tgtaagtagtttttttttttcccttcctcggttctttcctggtccatcgacagtaacgagat 1740
 GatcttgcaggcggtatcgagctaaccgcgattgtcgtagacagGAGTACATGCTCAAGTAT 1800
 E Y M L K Y -
 GACTCGACCCACGGTATCTTCAACGGCACCATCGCCGTCGAGGGCAACGACCTCATTGTC 1860
 D S T H G I F N G T I A V E G N D L I V -
 AACGGCAAGAGGGTCAAGTTCTACACTGAGCGGGMCCCGCCAAACATTCCCTGGARGGAA 1920
 N G K R V K F Y T E R ? P A N I P W ? E -
 ACTGGTGCCGAGTACATMRTCGAGTACGCGGTGTGTTACCCAMCACCAGAAAGGCTAGC 1980
 T G A E Y I ? E S T G V F T ? T ? K A S -
 GCCACCTCAAGGGCGGCGCCAAAGCGCGTCATCATCTCTGCTCCCTCGGCCGATGCCCCC 2040
 A H L K G G A K R V I I S A P S A D A P -
 ATGTACGTCATGGGCGTCAACGAGAAGACCTACGACGGCAAGGCCAGGTCATCTCTAAC 2100
 M Y V M G V N E K T Y D G K A Q V I S N -
 GCCTCGTGCAACCAACTGCCTGGCTCCCTCGCCAAGGTCATCCACGACAAGTTCGGC 2160
 A S C T T N C L A P L A K V I H D K F G -
 CTCGTTGAGGCTCTCATGACCACCGTCCACTCCTACACTGCCACCCAGAAGACCGTCGAT 2220
 L V E G L M T T V H S Y T A T Q K T V D -
 GGTCCCTCTGCCAAGGACTGGCGTGGTGGCGGTGGTCTGCTCAGAACATCATCCCCAGC 2280
 G P S A K D W R G G R G A A Q N I I P S -
 AGCACTGGCGCCGCAAGGCCGTCGCGCAAGGTCATCCCTGAGCTCAACGGCAAGCTCACC 2340
 S T G A A K A V G K V I P E L N G K L T -
 GGCATGTCCCTCCGTGTCCCAACCCCAACGTTTCCGTTGTCGACCTCACCTGCCGCCTC 2400
 G M S L R V P T P N V S V V D L T C R L -
 GAGAAGGAGGCTACCTACGACGACATCAAGGCCGCCATCAAGGAGGCCGCCGCCGCCGCC 2460
 E K E A T Y D D I K A A I K E A A A G P -
 CTAAGGgtgagttatctggttccctttttttttttggagaacgacacatgctgataaa 2520
 L K G
 acccagGCATCCTCGACTACACTGAGG 2547
 I L D Y T E

10

20

配列番号6

部分クリソヘリウム(*Chrysosporium*)GPD1タンパク質のアミノ酸
 (C末端は配列を利用可能にする際に欠失している)。

30

MTIKVGINGF GRIGRIVFRN SIEHSDVEIV AVNDPFIEPK YAEYMLKYDS
 THGIFNGTIA VEGNDLIVNG KRVKFYTER? PANIPW?ETG AEYI?ESTGV
 FT?T?KASAH LKGGAKRVII SAPSADAPMY VMGVNEKTYD GKAQVISNAS
 CTTNCLAPLA KVIHDKFGLV EGLMTTVHSY TATQKTVDPG SAKDWRGGRG
 AAQNIIPSS? GAAKAVGKVI PELNGKLTGM SLRVPTPNVS VVDLTCRLEK
 EATYDDIKAA IKEAAAGPLK GILDYTE

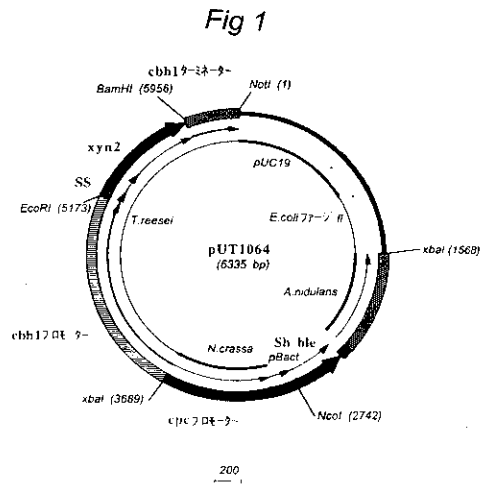
40

【図面の簡単な説明】

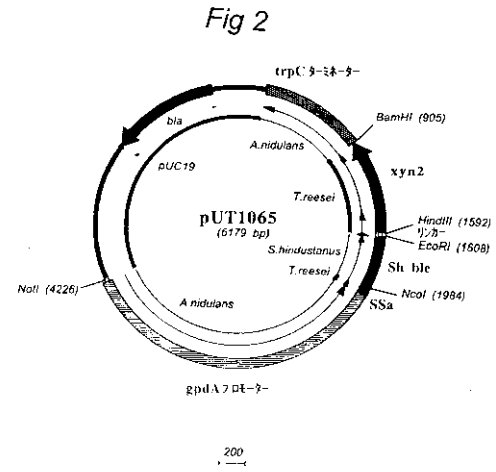
【図1】 発明にかかる、pUT1064の地図である。

【図2】 発明にかかる、pUT1065の地図である。

【 図 1 】



【 図 2 】



【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number
WO 01/79507 A2(54) International Patent Classification: C12N 15/56,
15/53, 1/15, C12Q 1/68, C12P 19/14, C12N 9/24, 9/42,
9/02, 15/80(74) Agent: JORRITSMA, Ruurd, Nederlandsch Octrooi-
bureau, Scheveningseweg 52, P.O. Box 29720, NL-2502 LS
The Hague (NL).

(21) International Application Number: PCT/NL01/00301

(22) International Filing Date: 17 April 2001 (17.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 00201343.1 13 April 2000 (13.04.2000) EP

(71) Applicant and
(72) Inventor: EMALFARR, Mark, Aaron [US/US]; 193
Spyglass Court, Jupiter, FL 33477 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): PUNT, Peter, Jan
[NL/NL]; Boekdrukkersgilde 1, NL-3994 XT Houten
[NL]; VAN ZELLEN, Cornelia, Maria, Johannes [NL/NL];
Houti Druaifan 24, NL-3451 EP Vleuten De Meern
[NL].(80) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EP, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).Published:
without international search report and to be republished
upon receipt of that reportFor two-letter codes and other abbreviations, refer to the "Guide-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/79507 A2

(54) Title: NOVEL EXPRESSION-REGULATING SEQUENCES AND EXPRESSION PRODUCTS IN THE FIELD OF HLA-
MENTOIS FUNGI(57) Abstract: The invention pertains to novel proteins corresponding to *Chrysosporium* glycerol hydrolases of families 7 and 10,
exhibiting a minimum amino acid identity of 70 and 75 %, respectively, with the amino acid sequence of SEQ ID No's 2 and 4,
and to a protein corresponding to a *Chrysosporium* glyceroldehyde phosphate dehydrogenase, exhibiting at least 86 % amino acid
identity with the partial amino acid sequence of SEQ ID No. 6. The invention further relates to nucleic acid sequences encoding
these proteins, and especially to promoter sequences regulating the expression of the corresponding genes. The preferred host for
expressing these genes is a fungus, especially a *Chrysosporium* strain.

WO 01/79507

-1-

PCT/NL01/00301

NOVEL EXPRESSION-REGULATING SEQUENCES AND EXPRESSION PRODUCTS IN THE FIELD OF FILAMENTOUS FUNGI

Field of the invention

5 The subject invention relates to novel enzymes derived from filamentous fungi, especially from strains of the genus *Chrysosporium*, and to coding sequences and expression-regulating sequences for these enzymes. The present invention is an extension on the invention disclosed in WO 00/20555 (PCT/NL99/00618), filed October 6, 1999, not published prior to the present priority date. The latter disclosure is herein incorporated by reference.

10

Background to the invention.

A number of hosts for gene expression and methods of transformation have been disclosed in the prior art. Bacteria are often mentioned e.g. *Escherichia coli*. *E. coli* is however a micro-organism incapable of secretion of a number of proteins or polypeptides and as such is undesirable
15 as host cell for production of protein or polypeptide at the industrial level. An additional disadvantage for *E. coli*, which is valid also for bacteria in general, is that prokaryotes cannot provide additional modifications required for numerous eukaryotic proteins or polypeptides to be produced in an active form. Glycosylation of proteins and proper folding of proteins are examples of processing required to ensure an active protein or polypeptide is produced. To ensure such
20 processing one can sometimes use mammalian cells; however, the disadvantage of such cells is that they are often difficult to maintain and require expensive media. Such transformation systems are therefore not practical for production of proteins or polypeptides at the industrial level. They may be cost efficient for highly priced pharmaceutical compounds requiring relatively low amounts, but certainly not for industrial enzymes.

25 A number of fungal expression systems have been developed e.g. *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus nidulans*, *Trichoderma reesei*. A number of others have been suggested but for various reasons have not found wide-spread acceptance or use. In general terms the ideal host must fulfil a large number of criteria:

- The ideal host must be readily fermented using inexpensive medium.
- 30 - The ideal host should use the medium efficiently.
- The ideal host must produce the polypeptide or protein in high yield, i.e. must exhibit high protein to biomass ratio.
- The ideal host should be capable of efficient secretion of the protein or polypeptide.
- The ideal host must enable ease of isolation and purification of the desired protein or polypeptide.
- 35 - The ideal host must process the desired protein or polypeptide such that it is produced in an

WO 01/79507

-2-

PCT/NL01/00301

active form not requiring additional activation or modification steps.

- The ideal host should be readily transformed.
- The ideal host should allow a wide range of expression regulatory elements to be used thus ensuring ease of application and versatility.
- 5 - The ideal host should allow use of easily selectable markers that are cheap to use.
- The ideal host should produce stable transformants.
- The ideal host should allow cultivation under conditions not detrimental to the protein or polypeptide product, e.g. low viscosity, low shear.

WO 96/02563 and US patents 5,602,004, 5,604,129 and 5,695,985 to Novo Nordisk describe the drawbacks of *Aspergillus* and *Trichoderma* systems and suggest that cultivation conditions for other fungi may be more suited to large scale protein production. The only examples provided for any transformed cultures are those of *Myceliophthora thermophila*, *Acremonium alabamense*, *Thielavia terrestris* and *Sporotrichum cellulosophilum* strains. The *Sporotrichum* strain is reported to lyse and produce green pigment under fermentation conditions not leading to such results for the other strains. A non-sporulating mutant of *Thielavia terrestris* is described as being the organism of choice by virtue of its morphology. However it is also stated that the protoplasting efficiency of *Thielavia* and *Acremonium* (whereby the *Acremonium* strain used was the imperfect state of the *Thielavia* strain used) is low and that hygromycin was not useful as a selection marker. A large number of others are suggested as being potentially useful by virtue of their morphology but no transformation thereof is described. The suggested strains are *Corynascus*, *Thermoascus*, *Chaetomium*, *Ctenomyces*, *Scytalidium* and *Talaromyces*. The transformed hosts are mentioned as only producing low levels of the introduced *Humicola* xylanase with *Thielavia* producing the lowest amount; however, the information is ambiguous and could actually infer *Thielavia* was the best embodiment. The nomenclature of this reference is based on the ATCC names of Industrial Fungi of 1994. Thus it is apparent that no high degree of heterologous expression was achieved and in fact no positive correlation could be derived between the postulated morphology and the degree of expression. If any correlation could be made, it was more likely to be negative. According to the 1996 ATCC fungal classification *Sporotrichum thermophilum* ATCC 20493 is a *Myceliophthora thermophila* strain. Currently the strain is still identified as *Myceliophthora thermophila*. The unpredictability of the art is apparent from these recent disclosures.

WO 97/26330 of Novo Nordisk suggests a method of obtaining mutants of filamentous fungal parent cells having an improved property for production of heterologous polypeptide. The method comprises first finding a specific altered morphology followed by assessing whether a transformant produces more heterologous polypeptide than the parent. The method is illustrated only for strains of *Fusarium A3/5* and *Aspergillus oryzae*. The method is suggested to be applicable for *Aspergillus*, *Trichoderma*, *Thielavia*, *Fusarium*, *Neurospora*, *Acremonium*, *Tolylocadium*,

WO 01/79507

-3-

PCT/NL01/00301

Humicola, *Sclerotium*, *Myceliophthora* or *Mucor*. As stated above, the unpredictability in the art and also the unpredictability of the method of the cited application do not provide a generally applicable teaching with a reasonable expectation of success.

In WO 00/20555, we have described an alternative fungal expression system with the simplicity of use of the above-mentioned *Aspergilli* and *Trichoderma* fulfilling the above requirements. The new system provides the additional advantages that transformation rates are higher than those for the frequently used *Trichoderma reesei* system. In addition the culture conditions offer the additional bonus of being advantageous for the polypeptide product.

10 *Detailed description of the invention.*

We now describe a number of industrially interesting enzymes derived from *Chrysosporium* strains, together with full sequence information. We also describe novel promoter systems derived from *Chrysosporium* strains and useful for expressing homologous and heterologous genes.

The present invention is in particular concerned with glycosyl hydrolases of the families 7 (e.g. cellobiohydrolases) and 10 (e.g. xylanases), and glyceraldehyde phosphate dehydrogenases, as identified by their amino acid sequence, as well as peptides derived from these enzymatic proteins, and with nucleic acid sequences encoding these peptides and proteins, as well as, in particular, with regulating sequences related to these genes.

In particular, the present invention pertains to isolated or recombinant enzymic proteins or active parts thereof of the three classes referred to above, including mutants thereof having at least a certain degree of sequence identity as specified in the further disclosure and in the claims, as well as nucleic acid sequences encoding these proteins or parts thereof, and/or nucleic acid sequences regulating their expression. These enzymes are especially: (1) a glycosyl hydrolase of family 7 (cellobiohydrolase, CBH1) having at least 75%, preferably at least 80% or even at least 85% amino acid identity with the sequence of SEQ ID No 2; (2) a glycosyl hydrolase of family 10 (endo-xylanase XYL1) having at least 70%, preferably at least 75% or even at least 80% amino acid identity with the sequence of SEQ ID No 4; and (3) a glyceraldehyde phosphate dehydrogenase (GPD1) having at least 86%, preferably at least 90% or even at least 93% amino acid identity with the sequence of SEQ ID No 6. Polypeptides and nucleic acid sequences encoding these polypeptides, having at least 20, preferably at least 30 contiguous amino acids of SEQ ID No's 2, 4 and 6 are also a preferred part of the invention. The corresponding nucleotide sequences are depicted in SEQ ID No's 1 (cbh1), 3 (xy1) and 5 (gpd1), respectively.

The recombinant enzymes may comprise essentially the complete protein, or a truncated protein having at least part of the enzymatic activity. Such truncated part may be the catalytic domain, or at least about 75% of the amino acids thereof. By way of example, the catalytic domain of the CBH1 according to the invention comprises the amino acids 20-495 of the amino acid

WO 01/79507

-4-

PCT/NL01/00301

sequence of SEQ ID No. 2, and the catalytic domain of the XYL1 according to the invention comprises the aminoacids 54-384 of the aminoacid sequence of SEQ ID No. 4. The catalytic domain may or may not be combined with a signal sequence originating from another protein and/or with a carbohydrate-binding domain from another enzymic protein. Alternatively, the cellulose-binding domain of the enzymes of the invention (CBH1 and XYL1) may be fused to catalytic domains of other enzymic proteins.

The nucleic acid sequences according to the invention may be complete protein-encoding regions or oligonucleotides or, preferentially, expression-regulating sequences. Oligonucleotides may be used also as probes for identifying genes corresponding to, but not identical to the genes of SEQ ID No.'s 1, 3 and 5; these genes, when fulfilling the percentage identity criteria defined herein, as well as encoding and non-encoding parts thereof and their expression products are also part of the invention. Oligonucleotides are preferably 15-75, most preferably 20-50 nucleotides in length.

The invention also pertains to expression systems (cassettes) comprising either an expression-regulating region (including a promoter) of any of the three protein classes fused to a gene encoding another protein of interest, or an encoding region of any of these proteins fused to another expression regulating region, or both the expression-regulating region and the protein-encoding region of these novel proteins. The expression-regulating region comprises at least 60%, preferably at least 70%, more preferably at least 75% or even 80% of the 5'-non-coding region of SEQ ID No.'s 1, 3 and 5, and/or at least 20, especially at least 40 contiguous nucleotides from these 5' non-coding regions. Terminating sequences similarly derived from the 3' non-coding regions of the genes of the invention are also useful in expressing cassettes, whether combined with homologous or heterologous genes.

The polynucleotides and oligonucleotides of the invention can have the minimum sequence identity with the corresponding sequences of SEQ ID NO's 1, 3 or 5, or, alternatively hybridise under stringent conditions with these given sequences. Stringent hybridisation conditions are those as understood in the art, e.g. hybridisation in 6 x SSC (20xSSC per 1000 ml : 175.3 g NaCl, 107.1 g sodium citrate.5H₂O, pH 7.0), 0.1% SDS, 0.05% sodium pyrophosphate, 5 * Denhardt's solution and 20 µg/ml denatured herring sperm DNA at 56°C for 18-24 hrs followed by two 30 min. washes in 5 x SSC, 0.1 % SDS at 56°C and two 30 min. washes in 2 x SSC, 0.1% SSC at 56°C.

These expression systems may be contained in a *Chrysosporium* host, such as a *Chrysosporium lucknowense* host, or in another non-fungal or, preferably, fungal host. Examples of other fungal hosts are other *Chrysosporium* species or strains, *Fusarium* species, *Aspergillus* species etc. Such host may be advantageously a host that does not itself, intrinsically or as a result of the culture conditions, produce a protein corresponding to the protein of interest, so as to simplify the recovery of the protein of interest.

WO 01/79507

-5-

PCT/NL01/00301

Where reference is made in this specification and in the appending claims to "polypeptides" or "peptides" or "polypeptides of interest" or "peptides of interest" as the products of the expression system of the invention, this term also comprise proteins, i.e. polypeptides having a particular function and/or secondary and/or tertiary structure. Where reference is made to a percentage amino acid identity, such identity relates to a complete protein or to a specific part defined by initial and final amino acid number, as determined by the conventionally used BLAST algorithm.

In the fungal expression system described in WO 00/20555, the pH of the culture medium can be neutral or alkaline thus no longer subjecting the produced protein or polypeptide to aggressive and potentially inactivating acid pH. It is also possible to culture at acid pH such as pH 4 for cases where the protein or polypeptide is better suited to an acidic environment. Suitably culture can occur at a pH between 4.0-10.0. A preference however exists for neutral to alkaline pH as the host strain exhibits better growth at such pH, e.g. between 6 and 9. Growth at alkaline pH which can be from pH 8 up and can even be as high as 10 is also a good alternative for some cases. Also the cultivation temperature of such host strains is advantageous to the stability of some types of produced polypeptide. The cultivation temperature is suitably at a temperature of 23-43°C. Clearly such conditions are of particular interest for production of mammalian polypeptides. The selected temperature will depend on cost effectiveness of the cultivation and sensitivity of the polypeptide or cultivation strain.

It has also been ascertained that the biomass to viscosity relation and the amount of protein produced is exceedingly favourable for the *Chrysosporium* host. Comparisons have been carried out with *Trichoderma longibrachiatum* (formerly also known as *Trichoderma reesei*) and with *Aspergillus niger*. *Trichoderma longibrachiatum* gave 2.5-5 g/l biomass, *Aspergillus niger* gave 5-10 g/l biomass and the *Chrysosporium* host gave 0.5-1 g/l biomass under their respective optimised conditions. This thus offers 5-10 fold improvement over the commercially used strains. The subject invention is directed at expression systems comprising a nucleic acid sequence encoding a heterologous protein or polypeptide, said nucleic acid sequence being operably linked to an expression regulating region described below and optionally a secretion signal encoding sequence and/or a carrier protein encoding sequence. Preferably a recombinant strain according to the invention will secrete the polypeptide of interest. This will avoid the necessity of disrupting the cell in order to isolate the polypeptide of interest and also minimise the risk of degradation of the expressed product by other components of the host cell.

Chrysosporium can be defined by morphology consistent with that disclosed in Barnett and Hunter 1972, *Illustrated Genera of Imperfect Fungi*, 3rd Edition of Burgess Publishing Company. Other sources providing details concerning classification of fungi of the genus *Chrysosporium* are known e.g. Sutton Classification (Van Oorschot, C.A.N. (1980) "A revision of *Chrysosporium* and allied genera" in *Studies in Mycology* No. 20 of the CBS in Baarn, The Netherlands p1-36). CBS is

WO 01/79507

-6-

PCT/NL01/00301

one of the depository institutes of the Budapest Treaty. According to these teachings the genus *Chrysosporium* falls within the family *Moniliaceae* which belongs to the order *Hyphomycetales*. The following strains are defined as *Chrysosporium* but the definition of *Chrysosporium* is not limited to these strains: *C. bobyoides*, *C. carmichaelii*, *C. crassitunicatum*, *C. europae*, *C. evolceanui*, *C. farinicola*, *C. fastidium*, *C. filiforme*, *C. georgiae*, *C. globiferum*, *C. globiferum* var. *articulatum*, *C. globiferum* var. *niveum*, *C. hirundo*, *C. hispanicum*, *C. holmii*, *C. indicum*, *C. inops*, *C. keratinophilum*, *C. kreiselii*, *C. kuzrovianum*, *C. lignorum*, *C. lobatum*, *C. lucknowense*, *C. lucknowense* Garg 27K, *C. medium*, *C. medium* var. *spissescens*, *C. nephiticum*, *C. merdarium*, *C. merdarium* var. *roseum*, *C. minor*, *C. pannicola*, *C. parvum*, *C. parvum* var. *crescens*, *C. pilosum*, *C. pseudomerdarium*, *C. pyriformis*, *C. queenslandicum*, *C. sigleri*, *C. sulfureum*, *C. synchronum*, *C. tropicum*, *C. undulatum*, *C. vollenarense*, *C. vespertilium*, *C. zonatum*.

C. lucknowense forms one of the species of *Chrysosporium* that have raised particular interest as it has provided a natural high producer of cellulase proteins (WO 98/15633 and related US 5,811,381). The characteristics of this *Chrysosporium lucknowense* are:

Colonies attain 55 mm diameter on Sabouraud glucose agar in 14 days, are cream-coloured, felty and fluffy; dense and 3-5 mm high; margins are defined, regular, and fimbriate; reverse pale yellow to cream-coloured. Hyphae are hyaline, smooth- and thin-walled, little branched. Aerial hyphae are mostly fertile and closely septate, about 1-3.5 μ m wide. Submerged hyphae are infertile, about 1-4.5 μ m wide, with the thinner hyphae often being contorted. Conidia are terminal and lateral, mostly sessile or on short, frequently conical protrusions or short side branches. Conidia are solitary but in close proximity to one another, 1-4 conidia developing on one hyphal cell, subhyaline, fairly thin- and smooth-walled, mostly subglobose, also clavate obovoid, 1-celled, 2.5-11 x 1.5-6 μ m, with broad basal scars (1-2 μ m). Intercalary conidia are absent. Chlamydospores are absent. ATCC 44006, CBS 251.72, CBS 143.77 and CBS 272.77 are examples of *Chrysosporium lucknowense* strains and other examples are provided in WO 98/15633 (US 5,811,381).

A further strain was isolated from this species with an even higher production capacity for cellulases. This strain is called C1 by its internal notation and was deposited with the International Depository of the All Russian Collection of micro-organisms of the Russian Academy of Sciences Bakrushina Street 8, Moscow, Russia 113184 on August 29, 1996, as a deposit according to the Budapest Treaty and was assigned Accession Number VKM F-3500D. It is called *Chrysosporium lucknowense* Garg 27K. The characteristics of the C1 strain are as follows:

Colonies grow to about 55-66 mm diameter in size on potato-dextrose agar in about 7 days; are white-cream-coloured, felty, 2-3 μ m high at the centre; margins are defined, regular, fimbriate; reverse pale, cream-coloured. Hyphae are hyaline, smooth- and thin-walled, little branched. Aerial hyphae are fertile, septate, 2-3 mm wide. Submerged hyphae are infertile. Conidia are terminal and lateral; sessile or on short side branches; absent; solitary, but in close proximity to one another.

WO 01/79507

-7-

PCT/NL01/00301

hyaline, thin- and smooth-walled, subglobose, clavate or obovoid, 1-celled, 4-10 µm. Chlamydo-spores are absent. Intersepta conidia are absent.

The method of isolation of the C1 strain is described in WO 98/15633, US 5,811,381, and US 6,015,707. Also included within the definition of *Chrysosporium* are strains derived from
5 *Chrysosporium* predecessors including those that have mutated somewhat either naturally or by induced mutagenesis. Mutants of *Chrysosporium* can be obtained by induced mutagenesis, especially by a combination of irradiation and chemical mutagenesis.

For example strain C1 was mutagenised by subjecting it to ultraviolet light to generate strain UV13-6. This strain was subsequently further mutated with N-methyl-N'-nitro-N-nitrosoguanidine
10 to generate strain NG7C-19. The latter strain in turn was subjected to mutation by ultraviolet light resulting in strain UV18-25. During this mutation process the morphological characteristics have varied somewhat in culture in liquid or on plates as well as under the microscope. With each successive mutagenesis the cultures showed less of the fluffy and felt appearance on plates that are described as being characteristic of *Chrysosporium*, until the colonies attained a flat and matted
15 appearance. A brown pigment observed with the wild type strain in some media was also less prevalent in mutant strains. In liquid culture the mutant UV18-25 was noticeably less viscous than the wild type strain C1 and the mutants UV13-6 and NG7C-19. While all strains maintained the gross microscopic characteristics of *Chrysosporium*, the mycelia became narrower with each successive mutation and with UV18-25 distinct fragmentation of the mycelia could be observed.
20 This mycelial fragmentation is likely to be the cause of the lower viscosity associated with cultures of UV18-25. The ability of the strains to sporulate decreased with each mutagenic step. The above illustrates that for a strain to belong to the genus *Chrysosporium* there is some leeway from the above morphological definition. At each mutation step production of cellulase and extracellular proteins has in addition also increased, while several mutations resulted in decrease of protease expression. Criteria with which fungal taxonomy can be determined are available from CBS,
25 VKMF and ATCC for example. The strains internally designated as *Chrysosporium* strain C1, strain UV13-6, strain NG7C-19 and strain UV18-25, have been deposited in accordance with the Budapest Treaty with the All Russian Collection (VKM) depository institute in Moscow. Wild type C1 strain was deposited with number VKM F-3500 D, deposit date 29-08-1996, C1 UV13-6 mutant
30 was deposited as VKM F-3632 D (02-09-1998), C1 NG7C-19 mutant was deposited as VKM F-3633 D (02-09-1998) and C1 UV18-25 mutant was deposited as VKM F-3631 D (02-09-1998).

It is preferable to use non-toxic *Chrysosporium* strains of which a number are known in the art as this will reduce risks to the environment upon large scale production and simplify production procedures with the concomitant reduction in costs.

35 An expression-regulating region is a DNA sequence recognised by the host *Chrysosporium* strain for expression. It comprises a promoter sequence operably linked to a nucleic acid sequence

WO 01/79507

-8-

PCT/NL01/00301

encoding the polypeptide to be expressed. The promoter is linked such that the positioning vis-à-vis the initiation codon of the sequence to be expressed allows expression. The promoter sequence can be constitutive or inducible. Any expression regulating sequence or combination thereof capable of permitting expression of a polypeptide from a *Chrysosporium* strain is envisaged. The expression
 5 regulating sequence is suitably a fungal expression-regulating region e.g. an ascomycete regulating region. Suitably the fungal expression regulating region is a regulating region from any of the following genera of fungi: *Aspergillus*, *Trichoderma*, *Chrysosporium*, *Hansenula*, *Mucor*, *Pichia*, *Neurospora*, *Tolypocladium*, *Rhizomucor*, *Fusarium*, *Penicillium*, *Saccharomyces*, *Talaromyces* or alternative sexual forms thereof like *Emmericella*, *Hypocrea* e.g. the cellobiohydrolase promoter
 10 from *Trichoderma*, glucoamylase promoter from *Aspergillus*, glyceraldehyde phosphate dehydrogenase promoter from *Aspergillus*, alcohol dehydrogenase A and alcohol dehydrogenase R promoter of *Aspergillus*, TAKA amylase promoter from *Aspergillus*, phosphoglycerate and cross-pathway control promoters of *Neurospora*, aspartic proteinase promoter of *Rhizomucor miehei*, lipase promoter of *Rhizomucor miehei* and beta-galactosidase promoter of *Penicillium canescens*.
 15 An expression regulating sequence from the same genus as the host strain is extremely suitable, as it is most likely to be specifically adapted to the specific host. Thus preferably the expression regulating sequence is one from a *Chrysosporium* strain.

Preferably an expression-regulating region enabling high expression in the selected host is applied. This is preferably an expression-regulating region derived from *Chrysosporium* according
 20 to the invention. It can also be a high expression-regulating region derived from a heterologous host, such as are well known in the art. Specific examples of proteins known to be expressed in large quantities and thus providing suitable expression regulating sequences for the invention are without being limited thereto hydrophobin, protease, amylase, xylanase, pectinase, esterase, beta-galactosidase, cellulase (e.g. endo-glucanase, cellobiohydrolase) and polygalacturonase. The high
 25 production has been ascertained in both solid state and submerged fermentation conditions. Assays for assessing the presence or production of such proteins are well known in the art. The catalogues of Sigma and Megazyme for example provide numerous examples. Megazyme is located at Bray Business Park, Bray, County Wicklow in Ireland. Sigma Aldrich has many affiliates world wide e.g. USA P.O. Box 14508 St. Louis Missouri. For cellulase we refer to commercially available
 30 assays such as CMCase assays, endovisco-metric assays, Avicelase assays, beta-glucanase assays, RBBCMCase assays. Cellazyme C assays. Xylanase assays are also commercially available (e.g. DNS and Megazyme). Alternatives are well known to a person skilled in the art and can be found from general literature concerning the subject and such information is considered incorporated herein by reference. By way of example we refer to "Methods in Enzymology" Volume 1, 1955
 35 right through to volumes 297-299 of 1998. Suitably a *Chrysosporium* promoter sequence is applied to ensure good recognition thereof by the host.

WO 01/79507

-9-

PCT/NL01/00301

We have found that heterologous expression-regulating sequences work as efficiently in *Chrysosporium* as native *Chrysosporium* sequences. This allows well known constructs and vectors to be used in transformation of *Chrysosporium* as well as offering numerous other possibilities for constructing vectors enabling good rates of expression in this novel expression and secretion host.

5 For example standard *Aspergillus* transformation techniques can be used as described for example by Christiansen et al in Bio/Technol. 6:1419-1422 (1988). Other documents providing details of *Aspergillus* transformation vectors, e.g. US patents 4,816,405, 5,198,345, 5,503,991, 5,364,770 and 5,578,463, EP-B-215,594 (also for *Trichoderma*) and their contents are incorporated by reference. As extremely high expression rates for cellulase have been ascertained for *Chrysosporium* strains, 10 the expression regulating regions of such proteins are particularly preferred. We refer for specific examples to the previously mentioned deposited *Chrysosporium* strains.

A nucleic acid construct comprising a nucleic acid expression regulatory region from *Chrysosporium*, preferably from *Chrysosporium lucknowense* or a derivative thereof forms a preferred embodiment of the invention, as does the mutant *Chrysosporium* strain comprising such 15 operably linked to a gene encoding a polypeptide to be expressed. Such a nucleic acid construct will be an expression regulatory region from *Chrysosporium* associated with cellulase or xylanase expression, preferably cellobiohydrolase expression, or glyceraldehyde phosphate dehydrogenase expression, as detailed below. The nucleic acid sequence according to the invention can suitably be obtained from a *Chrysosporium* strain, such strain being defined elsewhere in the description. The 20 manner in which promoter sequences can be determined are numerous and well known in the art. Nuclease deletion experiments of the region upstream of the ATG codon at the beginning of the relevant gene will provide such sequence. Also for example analysis of consensus sequences can lead to finding a gene of interest. Using hybridisation and amplification techniques one skilled in the art can readily arrive at the corresponding promoter sequences.

25 The promoter sequences of C1 endoglucanases were identified in this manner, by cloning the corresponding genes. Preferred promoters according to the invention are the 55 kDa cellobiohydrolase (CBH1) promoter, the 30 kDa xylanase (Xyl1) promoters, and the glyceraldehyde phosphate dehydrogenase promoter, as the enzymes are expressed at high level by their own promoters. The corresponding promoter sequences are identified in a straightforward manner by 30 cloning as described in WO 00/20555, using the sequence information given in SEQ ID No. 1 (for CBH1) and SEQ ID No. 3 (for Xyl1), respectively. The promoters of the carbohydrate-degrading enzymes of *Chrysosporium*, especially C1 promoters, can advantageously be used for expressing desired polypeptides in a host organism, especially a fungal or other microbial host organism. Promoter sequences having at least 65%, preferably at least 70%, most preferably at least 75% 35 nucleotide sequence identity with the sequence given in SEQ ID No's 1, 3 and 5, or with the sequences found for other *Chrysosporium* genes, are part of the present invention.

WO 01/78507

-10-

PCT/NL01/00301

For particular embodiments of the recombinant strain and the nucleic acid sequence according to the invention we also refer to the examples. We also refer for the recombinant strains to prior art describing high expression promoter sequences in particular those providing high expression in fungi e.g. such as are disclosed for *Aspergillus* and *Trichoderma*. The prior art provides a number of expression regulating regions for use in *Aspergillus* e.g. US 5,252,726 of Novo and US 5,705,358 of Unilever. The contents of such prior art are hereby incorporated by reference.

The hydrophobin gene is a fungal gene that is highly expressed. It is thus suggested that the promoter sequence of a hydrophobin gene, preferably from *Chrysosporium*, may be suitably applied as expression regulating sequence in a suitable embodiment of the invention. *Trichoderma reesei* and *Trichoderma harzianum* gene sequences for hydrophobin have been disclosed for example in the prior art as well as a gene sequence for *Aspergillus fumigatus* and *Aspergillus nidulans* and the relevant sequence information is hereby incorporated by reference (Munoz et al. *Curr. Genet.* 1997, 32(3):225-230; Nakari-Setälä T. et al. *Eur. J. Biochem.* 1996 15:235 (1-2):248-255; M. Parta et al. *Infect. Immun.* 1994 62 (10): 4389-4395 and Stringer M.A. et al. *Mol. Microbiol.* 1995 16(1):33-44). Using this sequence information a person skilled in the art can obtain the expression regulating sequences of *Chrysosporium* hydrophobin genes without undue experimentation following standard techniques as suggested already above. A recombinant *Chrysosporium* strain according to the invention can comprise a hydrophobin-regulating region operably linked to the sequence encoding the polypeptide of interest.

An expression regulating sequence can also additionally comprise an enhancer or silencer. These are also well known in the prior art and are usually located some distance away from the promoter. The expression regulating sequences can also comprise promoters with activator binding sites and repressor binding sites. In some cases such sites may also be modified to eliminate this type of regulation. Filamentous fungal promoters in which *creA* sites are present have been described. Such *creA* sites can be mutated to ensure the glucose repression normally resulting from the presence of the non-mutated sites is eliminated. Gist-Brocades' WO 94/13820 illustrates this principle. Use of such a promoter enables production of the polypeptide encoded by the nucleic acid sequence regulated by the promoter in the presence of glucose. The same principle is also apparent from WO 97/09438. These promoters can be used either with or without their *creA* sites. Mutants in which the *creA* sites have been mutated can be used as expression regulating sequences in a recombinant strain according to the invention and the nucleic acid sequence it regulates can then be expressed in the presence of glucose. Such *Chrysosporium* promoters ensure derepression in an analogous manner to that illustrated in WO 97/09438. The identity of *creA* sites is known from the prior art. Alternatively, it is possible to apply a promoter with *CreA* binding sites that have not been mutated in a host strain with a mutation elsewhere in the repression system e.g. in the

WO 01/79507

-11-

PCT/NL01/00301

creA gene itself, so that the strain can, notwithstanding the presence of creA binding sites, produce the protein or polypeptide in the presence of glucose.

Terminator sequences are also expression-regulating sequences and these are operably linked to the 3' terminus of the sequence to be expressed. Any fungal terminator is likely to be functional in the host *Chrysosporium* strain according to the invention. Examples are *A. nidulans* trpC terminator (1), *A. niger* alpha-glucosidase terminator (2), *A. niger* glucoamylase terminator (3), *Mucor miehei* carboxyl protease terminator (US 5,578,463) and the *Trichoderma reesei* cellobiohydrolase terminator. Naturally *Chrysosporium* terminator sequences will function in *Chrysosporium* and are suitable e.g. CBH1 terminator.

A suitable recombinant *Chrysosporium* strain to be used according to the invention has the nucleic acid sequence to be expressed operably linked to a sequence encoding the amino acid sequence defined as signal sequence. A signal sequence is an amino acid sequence which when operably linked to the amino acid sequence of the expressed polypeptide allows secretion thereof from the host fungus. Such a signal sequence may be one normally associated with the heterologous polypeptide or may be one native to the host. It can also be foreign to both host and the polypeptide. The nucleic acid sequence encoding the signal sequence must be positioned in frame to permit translation of the signal sequence and the heterologous polypeptide. Any signal sequence capable of permitting secretion of a polypeptide from a *Chrysosporium* strain is envisaged. Such a signal sequence is suitably a fungal signal sequence, preferably an ascomycete signal sequence.

Suitable examples of signal sequences can be derived from yeasts in general or any of the following specific genera of fungi: *Aspergillus*, *Trichoderma*, *Chrysosporium*, *Pichia*, *Neurospora*, *Rhizomucor*, *Hansenula*, *Humicola*, *Mucor*, *Tolypocladium*, *Fusarium*, *Penicillium*, *Saccharomyces*, *Talaromyces* or alternative sexual forms thereof like *Emericella*, *Hypocrea*. Signal sequences that are particularly useful are often natively associated with the following proteins a cellobiohydrolase, an endoglucanase, a beta-galactosidase, a xylanase, a pectinase, an esterase, a hydrophobin, a protease or an amylase. Examples include amylase or glucoamylase of *Aspergillus* or *Humicola* (4), TAKA amylase of *Aspergillus oryzae*, alpha-amylase of *Aspergillus niger*, carboxyl peptidase of *Mucor* (US 5,578,463), a lipase or proteinase from *Rhizomucor miehei*, cellobiohydrolase of *Trichoderma* (5), beta-galactosidase of *Penicillium canescens* and alpha mating factor of *Saccharomyces*.

Alternatively the signal sequence can be from an amylase or subtilisin gene of a strain of *Bacillus*. A signal sequence from the same genus as the host strain is extremely suitable as it is most likely to be specifically adapted to the specific host thus preferably the signal sequence is a signal sequence of *Chrysosporium*, especially of *Chrysosporium* strain C1, strain UV13-6, strain NG7C-19 and strain UV18-25, referred to above. Signal sequences from filamentous fungi, yeast

WO 01/79507

-12-

PCT/NL01/00301

and bacteria are useful. Signal sequences of non-fungal origin are also considered useful, particularly bacterial, plant and mammalian.

A recombinant host to be used according to any of the embodiments of the invention can further comprise a selectable marker. Such a selectable marker will permit easy selection of transformed or transfected cells. A selectable marker often encodes a gene product providing a specific type of resistance foreign to the non-transformed strain. This can be resistance to heavy metals, antibiotics and biocides in general. Prototrophy is also a useful selectable marker of the non-antibiotic variety. Non-antibiotic selectable markers can be preferred where the protein or polypeptide of interest is to be used in food or pharmaceuticals with a view to speedier or less complicated regulatory approval of such a product. Very often the GRAS indication is used for such markers. A number of such markers are available to the person skilled in the art. The FDA e.g. provides a list of such. Most commonly used are selectable markers selected from the group conferring resistance to a drug or relieving a nutritional defect e.g. the group comprising *amdS* (acetamidase), *hph* (hygromycin phosphotransferase), *pyrG* (orotidine-5'-phosphate decarboxylase), *trpC* (anthranilate synthase), *argB* (ornithine carbamoyltransferase), *sC* (sulphate adenylyltransferase), *bar* (phosphinothricin acetyltransferase), glufosinate resistance, *niaD* (nitrate reductase), a bleomycin resistance gene, more specifically *Sh ble*, sulfonylurea resistance e.g. acetolactate synthase mutation *ilvI*. Selection can also be carried out by virtue of cotransformation where the selection marker is on a separate vector or where the selection marker is on the same nucleic acid fragment as the polypeptide-encoding sequence for the polypeptide of interest.

As used herein the term heterologous polypeptide is a protein or polypeptide not normally expressed and secreted by the *Chrysosporium* host strain used for expression according to the invention. The polypeptide can be of plant or animal (vertebrate or invertebrate) origin e.g. mammalian, fish, insect, or micro-organism origin, with the proviso it does not occur in the host strain. A mammal can include a human. A micro-organism comprises viruses, bacteria, archaeobacteria and fungi i.e. filamentous fungi and yeasts. Bergey's Manual for Bacterial Determinology provides adequate lists of bacteria and archaeobacteria. For pharmaceutical purposes quite often a preference will exist for human proteins thus a recombinant host according to the invention forming a preferred embodiment will be a host wherein the polypeptide is of human origin. For purposes such as food production suitably the heterologous polypeptide will be of animal, plant or algal origin. Such embodiments are therefore also considered suitable examples of the invention. Alternative embodiments that are useful also include a heterologous polypeptide of any of bacterial, yeast, viral, archaeobacterial and fungal origin. Fungal origin is most preferred.

A suitable embodiment of the invention will comprise a heterologous nucleic acid sequence with adapted codon usage. Such a sequence encodes the native amino acid sequence of the host from which it is derived, but has a different nucleic acid sequence, i.e. a nucleic acid sequence in

WO 01/79507

-13-

PCT/NL01/00301

which certain codons have been replaced by other codons encoding the same amino acid but which are more readily used by the host strain being used for expression. This can lead to better expression of the heterologous nucleic acid sequence. This is common practice to a person skilled in the art. This adapted codon usage can be carried out on the basis of known codon usage of fungal vis-à-vis non-fungal codon usage. It can also be even more specifically adapted to codon usage of *Chrysosporium* itself. The similarities are such that codon usage as observed in *Trichoderma*, *Humicola* and *Aspergillus* should enable exchange of sequences of such organisms without adaptation of codon usage. Details are available to the skilled person concerning the codon usage of these fungi and are incorporated herein by reference.

The invention is not restricted to the above-mentioned recombinant *Chrysosporium* strains, but also covers a recombinant *Chrysosporium* strain comprising a nucleic acid sequence encoding a homologous protein for a *Chrysosporium* strain, said nucleic acid sequence being operably linked to an expression-regulating region and said recombinant strain expressing more of said protein than the corresponding non-recombinant strain under the same conditions. In the case of homologous polypeptide of interest such is preferably a neutral or alkaline enzyme like a hydrolase, a protease or a carbohydrate degrading enzyme as already described elsewhere. The polypeptide may also be acidic. Preferably the recombinant strain will express the polypeptide in greater amounts than the non-recombinant strain. All comments mentioned vis-à-vis the heterologous polypeptide are also valid (*mutatis mutandis*) for the homologous polypeptide cellulase.

Thus the invention also covers genetically engineered microbial strains wherein the sequence that is introduced can be of *Chrysosporium* origin. Such a strain can, however, be distinguished from natively occurring strains by virtue of for example heterologous sequences being present in the nucleic acid sequence used to transform or transfect the *Chrysosporium*, by virtue of the fact that multiple copies of the sequence encoding the polypeptide of interest are present or by virtue of the fact that these are expressed in an amount exceeding that of the non-engineered strain under identical conditions or by virtue of the fact that expression occurs under normally non-expressing conditions. The latter can be the case if an inducible promoter regulates the sequence of interest contrary to the non-recombinant situation or if another factor induces the expression than is the case in the non-engineered strain. The invention is directed at strains derived through engineering either using classical genetic technologies or genetic engineering methodologies.

The expression systems and host strains containing them according to the invention can comprise a nucleic acid sequence encoding a heterologous protein selected from carbohydrate-degrading enzymes (cellulases, xylanases, mannanases, mannosidases, pectinases, amylases, e.g. glucoamylases, α -amylases, α - and β -galactosidases, α - and β -glucosidases, β -glucanases, chitinases, chitinases), proteases (endoproteases, amino-proteases, amino- and carboxy-peptidases, keratinases), other hydrolases (lipases, esterases, phytases), oxidoreductases (catalases, glucose-

WO 01/79507

-14-

PCT/NL01/00301

oxidases) and transferases (transglycosylases, transglutaminases, isomerases and invertases).

The most interesting products to be produced according to invention are cellulases, xylanases, pectinases, lipases and proteases, wherein cellulases and xylanases cleave beta-1,4-bonds, and cellulases comprise endoglucanases, cellobiohydrolases and beta-glucosidases. These
 5 proteins are extremely useful in various industrial processes known in the art. Specifically for cellulases we refer e.g. to WO 98/15633 describing cellobiohydrolases and endoglucanases of use. The contents of said application are hereby incorporated by reference.

A recombinant according to the invention may have a nucleic acid sequence encoding the polypeptide of interest encodes a polypeptide that is inactivated or unstable at acid pH i.e. pH
 10 below 6, even below pH 5.5, more suitably even below pH 5 and even as low as or lower than pH 4. This is a particularly interesting embodiment, as the generally disclosed fungal expression systems are not cultured under conditions that are neutral to alkaline, but are cultured at acidic pH. Thus the system according to the invention provides a safe fungal expression system for proteins or polypeptides that are susceptible to being inactivated or are unstable at acid pH.

Quite specifically a recombinant strain as defined in any of the embodiments according to the invention, wherein the nucleic acid sequence encoding the polypeptide of interest encodes a protein or polypeptide exhibiting optimal activity and/or stability at a pH above 5, preferably at neutral or alkaline pH (i.e. above 7) and/or at a pH higher than 6, is considered a preferred
 15 embodiment of the invention. More than 50%, more than 70% and even more than 90% of optimal activities at such pH values are anticipated as being particularly useful embodiments. A polypeptide expressed under the cultivation conditions does not necessarily have to be active at the cultivation conditions, in fact it can be advantageous for it to be cultured under conditions under which it is inactive as its active form could be detrimental to the host. What is however required is for the protein or polypeptide to be stable under the cultivation conditions. The stability can be
 20 thermal stability. It can also be stability against specific compositions or chemicals, such as are present for example in compositions or processes of production or application of the polypeptide or protein of interest. LAS in detergent compositions comprising cellulases or lipases, etc. is an example of a chemical often detrimental to proteins. The time periods of use in applications can vary from short to long exposure so stability can be over a varying length of time varying per
 25 application. The skilled person will be able to ascertain the correct conditions on a case by case basis. One can use a number of commercially available assays to determine the optimal activities of the various enzymatic products. The catalogues of Sigma and Megazyme for example show such. Specific examples of tests are mentioned elsewhere in the description. The manufacturers provide guidance on the application.

35 A *Chrysosporium* strain can be suitably used to transform or transfect with the sequence of interest to be expressed and such a strain exhibits a relatively low biomass. We have found that

WO 01/79507

-15-

PCT/NL01/00301

Chrysosporium strains having a biomass two to five times lower than that of *Trichoderma reesei* when cultured to a viscosity of 200-600 cP at the end of fermentation and exhibiting a biomass of 10 to 20 times lower than that of *Aspergillus niger* when cultured to a viscosity of 1500-2000 cP under corresponding conditions, i.e. their respective optimal cultivation conditions can provide a high level of expression. This level of expression far exceeds that of the two commercial reference strains at a much lower biomass and at much lower viscosity. This means that the yield of expression of such *Chrysosporium* strains will be appreciably higher than from *Aspergillus niger* and *Trichoderma reesei*. Such a transformed or transfected *Chrysosporium* strain forms a suitable embodiment of the invention.

We find a biomass of 0,5-1,0 g/l for *Chrysosporium* strain C1(18-25) as opposed to 2,5-5,0 g/l for *Trichoderma reesei* and 5-10 g/l of *Aspergillus niger* under the above described conditions. In the Examples we provide details of this process.

In a suitable embodiment a recombinant *Chrysosporium* strain produces protein or polypeptide in at least the amount equivalent to the production in moles per liter of cellulase by the strain UV13-6 or C-19, and most preferably at least equivalent to or higher than that of the strain UV18-25 under the corresponding or identical conditions, i.e. their respective optimal cultivation conditions.

We have also found that expression and secretion rates are exceedingly high when using a *Chrysosporium* strain exhibiting the mycelial morphology of strain UV18-25 i.e. fragmented short mycelia. Thus a recombinant strain according to the invention will preferably exhibit such morphology. The invention however also covers non-recombinant strains or otherwise engineered strains of *Chrysosporium* exhibiting this novel and inventive characteristic. Also covered by the invention is a recombinant *Chrysosporium* strain in any of the embodiments described according to the invention further exhibiting reduced sporulation in comparison to C1, preferably below that of strain UV13-6, preferably below that of NG7C-19, preferably below that of UV18-25 under equivalent fermenter conditions. Also covered by the invention is a recombinant *Chrysosporium* strain in any of the embodiments described according to the invention further exhibiting at least the amount of protein production ratio to biomass in comparison to C1, preferably in comparison to that of any of strains UV13-6, NG7C-19 and UV18-25 under equivalent fermenter conditions. The invention however also covers non-recombinant strains or otherwise engineered strains of *Chrysosporium* exhibiting this novel and inventive characteristic as such or in combination with any of the other embodiments.

Another attractive embodiment of the invention also covers a recombinant *Chrysosporium* strain exhibiting a viscosity below that of strain NG7C-19, preferably below that of UV18-25 under corresponding or identical fermenter conditions. The invention however also covers non-recombinant strains or otherwise engineered strains of *Chrysosporium* exhibiting this novel and

WO 01/79507

-16-

PCT/NL01/00301

inventive characteristic as such or in combination with any of the other embodiments. We have determined that the viscosity of a culture of UV18-25 is below 10 cP opposed to that of *Trichoderma reesei* being of the order 200-600 cP, with that of *Aspergillus niger* being of the order 1500-2000 cP under their respective optimal culture conditions at the end of fermentation. The process used for such determination is provided in the examples.

Viscosity can be assessed in many cases by visual monitoring. The fluidity of the substance can vary to such a large extent that it can be nearly solid, sauce-like or liquid. Viscosity can also readily be ascertained by Brookfield rotational viscometry, use of kinematic viscosity tubes, falling ball viscometer or cup type viscometer. The yields from such a low viscosity culture are higher than from the commercial known higher viscosity cultures per time unit and per cell.

The processing of such low viscosity cultures according to the invention is advantageous in particular when the cultures are scaled up. The subject *Chrysosporium* strains with the low viscosity perform very well in cultures as large as up to 150,000 liter cultures. Thus any culture size up to 150,000 litres provides a useful embodiment of the invention. Any other conventional size of fermentation should be carried out well with the strains according to the invention. The reasoning behind this is that problems can arise in large scale production with the formation of aggregates that have mycelia that are too dense and/or are unevenly distributed. The media as a result cannot be effectively utilised during the culture thus leading to an inefficient production process in particular in large scale fermentations i.e. over 150,000 liters. Aeration and mixing become problematic leading to oxygen and nutrient starvation and thus reduced concentration of productive biomass and reduced yield of polypeptide during the culture and/or can result in longer fermentation times. In addition high viscosity and high shear are not desirable in commercial fermentation processes and in current commercial processes they are the production limiting factors. All these negative aspects can be overcome by the *Chrysosporium* host according to the invention which exhibits much better characteristics than *Trichoderma reesei*, *Aspergillus niger* and *Aspergillus oryzae* that are commercially used in this respect i.e. exhibits better protein production levels and viscosity properties and biomass figures.

A *Chrysosporium* strain according to any of the above-mentioned embodiments of the invention, said strain further exhibiting production of one or more of the fungal enzymes selected from the carbohydrate-degrading enzymes, proteases, other hydrolases, oxidoreductase, and transferases mentioned above, is considered a particularly useful embodiment of the invention. The most interesting products are specifically cellulases, xylanases, pectinases, lipases and proteases. Also useful as embodiment of the invention however is a *Chrysosporium* strain exhibiting production of one or more fungal enzymes that exhibit neutral or alkaline optimal stability and/or activity, preferably alkaline optimal stability and/or activity, said enzyme being selected from carbohydrate-degrading enzymes, hydrolases and proteases, preferably hydrolases and

WO 01/79507

-17-

PCT/NL01/00301

carbohydrate-degrading enzymes. In the case of non-recombinant *Chrysosporium*, such enzymes are suitably other than cellulase as disclosed in WO 98/15633. Enzymes of particular interest are xylanases, proteases, esterases, alpha-galactosidases, beta-galactosidases, beta-glucanases and pectinases. The enzymes are not limited to the aforementioned. The comments vis-à-vis stability and activity elsewhere in the description are valid here also.

The invention also covers a method of producing a polypeptide of interest, said method comprising culturing a host strain (e.g. fungal such as of the genera *Chrysosporium*, *Aspergillus*, *Trichoderma*, *Hansenula*, *Mucor*, *Pichia*, *Neurospora*, *Tolyposcladium*, *Rhizomucor*, *Fusarium*, *Penicillium* or bacterial or other microbial) in any of the embodiments according to the invention under conditions permitting expression and preferably secretion of the polypeptide and recovering the subsequently produced polypeptide of interest.

Where protein or polypeptide is mentioned, variants and mutants e.g. substitution, insertion or deletion mutants of naturally occurring proteins are intended to be included that exhibit the activity of the non-mutant. The same is valid vis-à-vis the corresponding nucleic acid sequences. Processes such as gene shuffling, protein engineering and directed evolution site directed mutagenesis and random mutagenesis are processes through which such polypeptides, variants or mutants can be obtained. US 5,223,409, US 5,780,279 and US 5,770,356 provide teaching of directed evolution. Using this process a library of randomly mutated gene sequences created for example by gene shuffling via error prone PCR occurs in any cell type. Each gene has a secretion region and an immobilising region attached to it such that the resulting protein is secreted and stays fixed to the host surface. Subsequently conditions are created that necessitate the biological activity of the particular polypeptide. This occurs for a number of cycles ultimately leading to a final gene with the desired characteristics. In other words a speeded up directed process of evolution. US 5,763,192 also describes a process for obtaining DNA, RNA, peptides, polypeptides or protein by way of synthetic polynucleotide coupling stochastically generated sequences, introduction thereof into a host followed by selection of the host cell with the corresponding predetermined characteristic.

Another application of the method of the present invention is in the process of "directed evolution", wherein novel protein-encoding DNA sequences are generated, the encoded proteins are expressed in a host cell, and those sequences encoding proteins having a desired characteristic are mutated and expressed again. The process is repeated for a number of cycles until a protein with the desired characteristics is obtained. Gene shuffling, protein engineering, error-prone PCR, site-directed mutagenesis, and combinatorial and random mutagenesis are examples of processes through which novel DNA sequences encoding exogenous proteins can be generated. U.S. patents 5,223,409, 5,780,279 and 5,770,356 provide teaching of directed evolution. See also Kuchner and Arnold, *Trends in Biotechnology*, 15:523-530 (1997); Schmidt-Dannert and Arnold, *Trends in*

WO 01/79507

-18-

PCT/NL01/00391

Biotech., 17:135-136 (1999); Arnold and Volkov, *Curr. Opin. Chem. Biol.*, 3:54-59 (1999); Zhao et al., *Manual of Industrial Microbiology and Biotechnology*, 2nd Ed., (Demain and Davies, eds.) pp. 597-604, ASM Press, Washington DC, 1999; Arnold and Wintrode, *Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis, and Bioseparation*, (Flickinger and Drew, eds.) pp. 971-987, John Wiley & Sons, New York, 1999; and Minshall and Stemmer, *Curr. Opin. Chem. Biol.* 3:284-290.

An application of combinatorial mutagenesis is disclosed in Liu et al., *Biochemistry*, 1998 37:10006-10015. US 5,763,192 describes a process for obtaining novel protein-encoding DNA sequences by stochastically generating synthetic sequences, introducing them into a host, and selecting host cells with the desired characteristic. Methods for effecting artificial gene recombination (DNA shuffling) include random priming recombination (Z. Shao, et al., *Nucleic Acids Res.*, 26:681-683 (1998)), the staggered extension process (H. Zhao et al., *Nature Biotech.*, 16:258-262 (1998)), and heteroduplex recombination (A. Volkov et al., *Nucleic Acids Res.*, 27:e18 (1999)). Error-prone PCR is yet another approach (Song and Rhee, *Appl. Environ. Microbiol.* 66:890-894 (2000)).

There are two widely-practised methods of carrying out the selection step in a directed evolution process. In one method, the protein activity of interest is somehow made essential to the survival of the host cells. For example, if the activity desired is a cellulase active at pH 8, a cellulase gene could be mutated and introduced into the host cells. The transformants are grown with cellulose as the sole carbon source, and the pH raised gradually until only a few survivors remain. The mutated cellulase gene from the survivors, which presumably encodes a cellulase active at relatively high pH, is subjected to another round of mutation, and the process is repeated until transformants that can grow on cellulose at pH 8 are obtained. Thermostable variants of enzymes can likewise be evolved, by cycles of gene mutation and high-temperature culturing of host cells (Liao et al., *Proc. Natl. Acad. Sci. USA* 83:576-580 (1986); Giver et al., *Proc. Natl. Acad. Sci. USA* 95:12809-12813 (1998)).

An alternative to the massively parallel "survival of the fittest" approach is serial screening. In this approach, individual transformants are screened by traditional methods, such as observation of cleared or coloured zones around colonies growing on indicator media, colorimetric or fluorometric enzyme assays, immunoassays, binding assays, etc. See for example Joo et al., *Nature* 399:670-673 (1999), where a cytochrome P450 monooxygenase not requiring NADH as a cofactor was evolved by cycles of mutation and screening; May et al., *Nature Biotech.* 18:317-320 (2000), where a hydantoinase of reversed stereoselectivity was evolved in a similar fashion; and Miyazaki et al., *J. Mol. Biol.* 297:1015-1026 (2000), where a thermostable subtilisin was evolved.

Standard cloning and protein or polypeptide isolation techniques can be used to arrive at the required sequence information. Parts of known sequences can be used as probes to isolate other

homologues in other genera and strains. The nucleic acid sequence encoding a particular enzyme activity can be used to screen a *Chrysosporium* library for example. A person skilled in the art will realise which hybridisation conditions are appropriate. Conventional methods for nucleic acid hybridisation construction of libraries and cloning techniques are described in Sambrook et al (Eds) (1989) In "Molecular Cloning. A Laboratory Manual" Cold Spring Harbor, Press Plainview, New York, and Ausubel et al (Eds) "Current Protocols in Molecular Biology" (1987) John Wiley and Sons, New York. The relevant information can also be derived from later handbooks and patents, as well as from various commercially available kits in the field.

In an alternative embodiment, said method comprises culturing a strain according to the invention under conditions permitting expression and preferably secretion of the protein or polypeptide or precursor thereof and recovering the subsequently produced polypeptide and optionally subjecting the precursor to additional isolation and purification steps to obtain the polypeptide of interest. Such a method may suitably comprise a cleavage step of the precursor into the polypeptide or precursor of interest. The cleavage step can be cleavage with a Kex-2 like protease, any basic amino acid paired protease or Kex-2 for example when a protease cleavage site links a well secreted protein carrier and the polypeptide of interest. A person skilled in the art can readily find Kex-2-like protease sequences as consensus sequence details for such are available and a number of alternatives have already been disclosed e.g. furin.

Suitably in a method for production of the polypeptide according to any of the embodiments of the invention the cultivation occurs at pH higher than 5, preferably 5-10, more preferably 6-9. Suitably in such a method the cultivation occurs at a temperature between 25-43 °C, preferably 30-40°C. The strain used in the method according to the invention is quite suitably a recombinant *Chrysosporium* strain or other fungal or non-fungal strain. The method according to the invention in such a case can further be preceded by the step of production of a recombinant strain according to the invention. The selection of the appropriate conditions will depend on the nature of the polypeptide to be expressed and such selection lies well within the realm of normal activity of a person skilled in the art.

The method of production of a recombinant strain according to the invention is also part of the subject invention. The method comprises stably introducing a nucleic acid sequence encoding a heterologous or homologous polypeptide into a suitable host strain, said nucleic acid sequence being operably linked to an expression regulating region, said introduction occurring in a manner known per se for transforming filamentous fungi. As stated above numerous references heretofore are available and a small selection has been cited. The information provided is sufficient to enable the skilled person to carry out the method without undue burden. The method comprises introduction of a nucleic acid sequence comprising any of the nucleic acid elements described in the various embodiments of the recombinant strain according to the invention as such or in combination.

WO 01/79507

-20-

PCT/NL01/00301

By way of example the introduction can occur using the protoplast transformation method. The method is described in the examples. Alternative protoplast or spheroplast transformation methods are known and can be used as have been described in the prior art for other filamentous fungi. Details of such methods can be found in many of the cited references and are thus
 5 incorporated by reference. A method according to the invention suitably comprises using a non-recombinant strain as starting material for introduction of the desired sequence encoding the polypeptide of interest.

The subject invention also covers a method of producing *Chrysosporium* enzyme, said method comprising culturing a *Chrysosporium* or other strain in or on a cultivation medium at pH
 10 higher than 5, preferably 5-10, more preferably 6-9, suitably 6-7.5, 7.5-9 as examples of neutral and alkaline pH ranges.

More in general the invention further covers a method of producing enzymes exhibiting neutral or alkaline optimal activity and/or stability, preferably alkaline optimal activity and/or stability. The preferred ranges vis-à-vis pH and optimal activity as well as assays with which to
 15 determine such have been provided elsewhere in the description. The enzyme should be selected from carbohydrate-degrading enzymes, proteases, other hydrolases, oxidoreductases, and transferases, as described above, said method comprising cultivating a host cell transformed or transfected with the corresponding enzyme-encoding nucleic acid sequence. Suitably such an enzyme will be a *Chrysosporium* enzyme. A suitable method such as this comprises production
 20 specifically of cellulase, xylanase, pectinase, lipase and protease, wherein cellulase and xylanase cleave β -1,4-bonds and cellulase comprises endoglucanase, cellobiohydrolase and β -glucosidase. The method according to the invention can comprise cultivating any *Chrysosporium* host according to the invention comprising nucleic acid encoding such aforementioned enzymes. Suitably the production of non-recombinant *Chrysosporium* hosts according to the invention is directed at
 25 production of carbohydrate degrading enzymes, hydrolases and proteases. In such a case the enzyme is suitably other than a cellulase. Methods of isolating are analogous to those described in WO 98/15633 and are incorporated by reference.

The enzymes produced according to the invention are also covered by the invention. Enzymes of *Chrysosporium* origin as can be isolated from non-recombinant *Chrysosporium* strains
 30 according to the invention are also covered. They exhibit the aforementioned stability, activity characteristics. Suitably they are stable in the presence of LAS. In particular proteases with pI 4-9.5, proteases with a MW of 25-95 kDa, xylanases with pI between 4.0 and 9.5, xylanases with MW between 25 and 65 kDa, endoglucanases with a pI between 3.5 and 6.5, endoglucanases with MW of 25-55 kDa, β -glucosidases, α , β -galactosidases with a pI of 4-4.5, β -glucosidases, α , β -galactosidases
 35 with a MW of 45-50 kDa, cellobiohydrolases of pI 4-5, cellobiohydrolases of MW 45-75 kDa, e.g. a MW of 55 kDa and pI 4.4, polygalacturonases, with a pI of 4.0-5.0 polygalacturonase of 60-70

WO 01/79507

-21-

PCT/NL01/00301

kDa, e.g. 65-kDa, esterases with a pI 4-5, and esterases with a MW of 95-105 kDa with the aforementioned stability, activity characteristics are claimed. The molecular weights (MW) are those determined by SDS-PAGE. The non-recombinant i.e. natively occurring enzyme is other than cellulase as disclosed in WO 98/15633. Enzymes with combinations of the pI values and molecular weights mentioned above are also covered.

The invention is also concerned with the (over)production of non-protein products by the mutant (recombinant) strains of the invention. Such non-protein products include primary metabolites such as organic acids, amino acids, and secondary such as antibiotics, e.g. penicillins and cephalosporins and other therapeutics. These products are the result of combinations of biochemical pathways, involving several fungal genes of interest. Fungal primary and secondary metabolites and procedures for producing these metabolites in fungal organisms are well known in the art. Examples of the production of primary metabolites have been described by Matthey M., The Production of Organic Acids, *Current Reviews in Biotechnology*, 12, 87-132 (1992). Examples of the production of secondary metabolites have been described by Penalva et al. The Optimization of Penicillin Biosynthesis in Fungi, *Trends in Biotechnology* 16, 483-489 (1998).

EXAMPLES

EXAMPLES OF TRANSFORMATION COMPARING *CHRYSO Sporium*, *TRICHODERMA* AND *TOLYPOCLADIUM GEODES*

Two untransformed *Chrysosporium* C1 strains and one *Trichoderma reesei* reference strain were tested on two media (Gs pH 6.8 and Pridham agar, PA, pH 6.8). To test the antibiotic resistance level spores were collected from 7 day old PDA plates. Selective plates were incubated at 32°C and scored after 2,4 and 5 days. It followed that the C-1 strains NG7C-19 and UV18-25 clearly have a low basal resistance level both to phleomycin and hygromycin. This level is comparable to that for a reference *T. reesei* commonly used laboratory strain. Thus there is clear indication these two standard fungal selectable markers can be used well in *Chrysosporium* strains. Problems with other standard fungal selectable markers should not be expected.

Selection of Sh-ble (phleomycin-resistance) transformed *Chrysosporium* strains was successfully carried out at 50 µg/ml. This was also the selection level used for *T. reesei* thus showing that differential selection can be easily achieved in *Chrysosporium*. The same comments are valid for transformed strains with hygromycin resistance at a level of 150 µg/ml.

The protoplast transformation technique was used on *Chrysosporium* based on the most generally applied fungal transformation technology. All spores from one 90mm PDA plate were recovered in 8ml IC1 and transferred into a shake flask of 50ml IC1 medium for incubation for 15 hours at 35°C and 200 rpm. After this the culture was centrifuged, the pellet was washed in MnP, brought back into solution in 10ml MnP and 10mg/ml Caylase C, and incubated for 30 minutes at

WO 01/79507

-22-

PCT/NL01/00301

35°C with agitation (150 rpm).

The solution was filtered and the filtrate was subjected to centrifugation for 10 minutes at 3500 rpm. The pellet was washed with 10 ml MuPCa^{2+} . This was centrifuged for 10 minutes at 25°C. Then 50 microlitres of cold MPC was added. The mixture was kept on ice for 30 minutes whereupon 2.5 ml PMC was added. After 15 minutes at room temperature 500 microlitres of the treated protoplasts were mixed to 3 ml of MnR Soil and immediately plated out on a MnR plate containing phleomycin or hygromycin as selection agent. After incubation for five days at 30°C transformants were analysed (clones become visible after 48 hours). Transformation efficiency was determined using 10 microgrammes of reference plasmid pAN8-1¹⁹. The results are presented in the following Table 1.

Table 1: Transformation efficiency (using 10 µg of reference plasmid pAN8-1)

	<i>T. reesei</i>	NG7C-19	UV18-25
Viability	$10^6/200 \mu\text{l}$	$5 \cdot 10^6/200 \mu\text{l}$	$5 \cdot 10^6/200 \mu\text{l}$
Transformants Per 200 µl	2500	10^1	10^1
Transformants per 10^6 viable cells	2500	2000	2000

The *Chrysosporium* transformant's viability is superior to that of *Trichoderma*. The transformability of the strains is comparable and thus the number of transformants obtained in one experiment lies 4 times higher for *Chrysosporium* than for *T. reesei*. Thus the *Chrysosporium* transformation system not only equals the commonly used *T. reesei* system, but even outperforms it. This improvement can prove especially useful for vectors that are less transformation efficient than pAN8-1. Examples of such less efficient transformation vectors are protein carrier vectors for production of non-fungal proteins which generally yield 10 times fewer transformants.

A number of other transformation and expression vectors were constructed with homologous *Chrysosporium* protein encoding sequences and also with heterologous protein encoding sequences for use in transformation experiments with *Chrysosporium*.

Examples of expression systems include a *Chrysosporium* xylanase XylII promoter fragment linked to a xylanase signal sequence in frame with a xylanase open reading frame followed by a xylanase terminator sequence. Transformant selection is carried out by using cotransformation with a selectable vector.

Another example is a *Chrysosporium lucknowense* cellobiohydrolase promoter linked to *Penicillium* endoglucanase 3 signal sequence in frame with the *Penicillium* endoglucanase 3 open reading frame followed by the *Chrysosporium* cellobiohydrolase terminator sequence. In addition

WO 01/79507

-23-

PCT/NL01/00301

this vector carries a second expression cassette with a selection marker i.e. the acetamidase S gene (AmdS gene).

A further example comprises *Chrysosporium* glyceraldehyde-3-phosphate dehydrogenase I promoter linked to the *Aspergillus niger* glucoamylase signal sequence and the glucoamylase open reading frame fused to the human Interleukine 6 open reading frame. In addition this vector carries a second expression cassette with a selection marker i.e. the AmdS gene.

A still further example is a *Aspergillus nidulans* glyceraldehyde-3-phosphate dehydrogenase A promoter linked to the endoglucanase 5 open reading frame followed by a *Aspergillus nidulans* terminator sequence.

10

EXAMPLES OF HETEROLOGOUS AND HOMOLOGOUS EXPRESSION OF CHRYSPORIUM TRANSFORMANTS

C1 strains (NG7C-19 and/or UV18-25) have been tested for their ability to secrete various heterologous proteins: a bacterial protein (*Streptococcus hindustanus* phleomycin-resistance protein, Sh ble), a fungal protein (*Trichoderma reesei* xylanase II, XYN2) and a human protein (the human lysozyme, HLZ).

C1 secretion of *Trichoderma reesei* xylanase II (XYN2).

C1 strain UV18-25 has been transformed by the plasmids pUT1064 and pUT1065.

pUT1064 presents the two following fungal expression cassettes:

20 The first cassette allows the selection of phleomycin-resistant transformants:

- *Neurospora crassa* cross-pathway control gene 1 (*cpc-1*) promoter ¹⁴

- *Streptococcus hindustanus* phleomycin-resistance gene *Sh ble* ⁴

- *Aspergillus nidulans* tryptophan-synthase (*trpC*) terminator ⁵

The second cassette is the xylanase production cassette:

25 - *T. reesei* strain TR2 *cbh1* promoter ¹⁵

- *T. reesei* strain TR2 *xyn2* gene (including its signal sequence) ¹⁶

- *T. reesei* strain TR2 *cbh1* terminator ¹⁸

The vector also carries an *E. coli* replication origin from plasmid pUC19 ⁶. The plasmid detailed map is provided in figure 1.

30 pUT1065 presents the following fungal expression cassette:

- *A. nidulans* glyceraldehyde-3-phosphate dehydrogenase (*gpd4*) promoter ²

- A synthetic *T. reesei* cellobiohydrolase I (*cbh1*) signal sequence ^{3,5}

- *S. hindustanus* phleomycin-resistance gene *Sh ble* ⁴ used as carrier-protein ¹⁰

- A linker peptide (SGERK) featuring a KEX2-like protease cleavage site ¹

35 - *T. reesei* strain TR2 *xyn2* gene (without signal sequence) ¹⁸

WO 01/79507

-24-

PCT/NL01/00301

- *A. nidulans* tryptophan-synthase (*trpC*) terminator⁵

The vector also carries the beta-lactamase gene (*bla*) and an *E. coli* replication origin from plasmid pUC18⁶. The plasmid detailed map is provided in figure 2.

C1 protoplasts were transformed with plasmid pUT1064 or pUT1065 following the same procedure already described in the above example. The fusion protein in plasmid pUT1065 (Sh ble :: XYN2) is functional with respect to the phleomycin resistance thus allowing easy selection of the C1 transformants. Moreover, the level of phleomycin resistance correlates roughly with the level of *xyn2* expression. In pUT1064, *xyn2* was cloned with its own signal sequence.

The xylanase production of C1 transformants (phleomycin-resistant clones) was analysed by xylanase activity assay as follows: Primary transformants were toothpicked to GS+phleomycin (5µg/ml) plates (resistance verification) and also on XYLAN plates (xylanase activity detection by clearing zone visualisation¹⁷). Plates were grown for 5 days at 32°C. Each validated clone was subcloned onto XYLAN plates. Two subclones per transformant were used to inoculate PDA plates in order to get spores for liquid culture initiation. The liquid cultures in IC1+ 5g/l KPhthalate were grown 5 days at 27°C (shaking 180 rpm). Then, the cultures were centrifuged (5000g, 10min.). From these samples, xylanase activity was measured by DNS technique according to Miller et al.¹⁸

Table 2: Active XYN2 production levels in C1 (best producers)

	Active xylanase II concentration in culture media	Xylanase II specific activity in culture media
Untransformed UV18-25	3.9 U./ml	3.8 U./mg total prot.
UV18-25::1064 clone 7-1	4.7 U./ml	4.7 U./mg total prot.
UV18-25::1064 clone 7-2	4.4 U./ml	4.3 U./mg total prot.
UV18-25::1065 clone 1-1	29.7 U./ml	25.6 U./mg total prot.
UV18-25::1065 clone 1-2	39.8 U./ml	39.4 U./mg total prot.

These data show that:

- 1) Points 1 to 4 from example 2 are confirmed.
- 2) C1 can be used as host for the secretion of a heterologous fungal protein.

Appendix to the Examples: Media

Transformation media:

Mandels Base:	MnP Medium :	
KH ₂ PO ₄	2.0 g/l	Mandels Base with
(NH ₄) ₂ SO ₄	1.4 g/l	Peptone 1 g/l
MgSO ₄ ·7H ₂ O	0.3 g/l	MES 2 g/l

WO 01/79507

-25-

PCT/NL01/00301

	CaCl ₂	0.3 g/l	Sucrose 100 g/l	
	Oligoelements	1.0 ml/l	Adjust pH to 5	
	MnR	MnP CA²⁺ :		
5	MnP+sucrose	130 g/l	MnP Medium +	
	Yeast extract	2.5 g/l	CaCl ₂ 2H ₂ O	50 mM
	Glucose	2.5 g/l	Adjust pH to 6.5	
	Agar	15 g/l		
10	MnR Soft :	MnR with only 7.5 g/l of agar.		
	MPC :			
	CaCl ₂	50 mM	pH 5.8	
	MOPS	10 mM		
15	PEG	40%		
<u>For selection and culture</u>				
	GS :			
	Glucose	10 g/l		
20	Biosoyase	5 g/l	[Merieux]	
	Agar	15 g/l	pH should be 6.8	
	PDA :			
	Potato Dextrose Agar	39 g/l	[Difco]	
25			pH should be 5.5	
	MPG :			
	Mandels Base with			
	K.Phtalate	5 g/l		
	Glucose	30 g/l		
30	Yeast extract	5 g/l		

The regeneration media (MnR) supplemented with 50 µg/ml phleomycin or 100-150 µg/ml hygromycin is used to select transformants. GS medium, supplemented with 5 µg/ml phleomycin is used to confirm antibiotic resistance.

35 PDA is a complete medium for fast growth and good sporulation. Liquid media are inoculated with 1/20th of spore suspension (all spores from one 90mm PDA plate in 5ml 0.1% Tween). Such cultures are grown at 27°C in shake flasks (200 rpm).

40 ISOLATION AND CHARACTERISATION OF *CI* GENES AND GENE EXPRESSION SEQUENCES OF *CBHI*, *XYLI*, AND *GPD*

Construction of a *BlueSTAR* gene library of *UV18-25*

Chromosomal DNA of *UV18-25* was partially digested with *Sau3A*, fragments of 12-15 kb were

WO 01/79507

-26-

PCT/NL01/00391

isolated and ligated in a *Bam*HI site of cloning vector BlueSTAR. Packaging of 20% of the ligation mixture resulted in a gene library of 4.6×10^4 independent clones. This library was multiplied and stored at 4 °C and -80°C. The rest of the ligation mixture was also stored at 4 °C.

*Screening the gene library of UV18-25 for isolation of the genes for *cbh1*, *xyl1* and *gpd1**

- 5 For the isolation of the different genes, in total $\pm 7.5 \times 10^4$ individual BlueSTAR phages per probe were hybridised in duplo. Hybridisation was carried out with the PCR fragments of *cbh1*, and *xyl1* (as described in WO 00/20555) at homologous conditions (65°C; 0.2xSSC) and with the *gpd1* gene of *A. niger* at heterologous conditions (53°C; 0.5xSSC). The number of positive signals is given in Table 3. The positive clones were rescreened and for each clone two individual phages were used for further experiments. DNA of the different clones was analysed by restriction analysis to determine the number of different clones isolated from each gene (results are given in Table 3).

- 10 As for each of the 3 genes, 4-6 different clones were isolated, we conclude that the primary gene library ($\pm 4.5 \times 10^4$ clones) represents about 5x genome of UV18-25. From this result we conclude that the complete genome of UV18-25 is represented in 9×10^3 clones. Based on an average genomic insert of 13 kb, this would indicate a genome size of ± 120 Mb, which is 3 times the size of the *Aspergillus* genome.

- 15 PCR reactions with specific primers for the gene present on the plasmid (based on previous sequence determination from the isolated PCR fragments) and the T7 and T3 primer present in the polylinker of pBlueSTAR we were able to determine the location of the genes in a number of clones. From each gene a plasmid was used for sequence determination of the gene.

Gene	Positive in first screening	positive in rescreening	different clones	clone used for sequencing
<i>cbh1</i>	8	7	4	pCBH7
<i>xyl1</i>	9	6	5	pXyl5
<i>gpd1</i>	12	12	6	pGPD4

- 25 **Table 3:** Screening of 7.5×10^4 phages of the gene library of UV18-25 with PCR fragments of UV18-25 for the *cbh1* gene and the *xyl1* gene (homologous conditions) and with the *gpdA* gene of *A. niger* (heterologous conditions). DNA isolation and restriction analysis was used to determine the number of different clones.

Sequence analysis of the cloned genes

- 30 For the *cbh1*, *xyl1*, and the *gpd1* gene, the results of the sequence determination are represented in SEQ ID No's 1, 3 and 5 respectively. Also the deduced amino acid sequences of the proteins are represented in these SEQ ID No's 2, 4 and 6. Some properties of the proteins are given in Table 4. It should be mentioned that the position of the start of the translation and the introns is based on homology with genes from the same family (i.e. paper genetics).

WO 01/79507

-27-

PCT/NL01/00301

CBH1

From the amino acid sequences of CBH1, we concluded that the protein is about 63 kD in size and that a cellulose-binding domain (CBD) is present at the C-terminal part of the protein. Interestingly, no evidence was found for the presence of a CBD in the isolated 55 kD major protein. However, the presence of the isolated peptides from this 55 kD major protein in the encoded CBH1 protein (SEQ ID No.1, 2), confirms that the 55 kD protein is encoded by the cloned gene. A possible explanation of these results is that the 55 kD protein is a truncated version of the CBH1 protein lacking the CBD.

The cellobiohydrolase CBH1 has activity against MUF-cellobioside, MUF lactoside, FP and avicel, also against p-nitrophenyl β -glucoside, cellobiose and p-nitrophenyl lactoside. Its activity toward MUF cellobioside is inhibited by cellobiose with inhibition constant of 0.4 mM. The Michaelis constant toward MUF cellobioside was 0.14 mM, toward MUF lactoside was 4 mM and toward CMC was 3.6 g/l. The pH optimum is from 4.5 to 7. 50% of maximum activity toward CMC and 80% activity toward RBB-CMC is kept at pH 8. 70-80% activity within pH 5-8 is kept during 25 hours of incubation. The temperature optimum is 60-70°C. CBH1 is a member of the cellobiohydrolase family 7. The corresponding CBH1 promoter, which is a preferred embodiment of the invention, is indicated in SEQ ID No. 1.

Xyl1

From the amino acid sequences of Xyl1 we conclude that also here a CBD is present, in this protein at the N-terminal side (i.e. directly attached to or at less than 5 amino acids distance from the signal sequence). In the literature only few more xylanases with a CBD are known (*Fusarium oxysporum*, *Hemicola grisea* and *Neocallimastix patriciarum*). The estimated size of the Xyl1 protein is 43 kD and several peptides isolated from a 30 kD xylanase originate from this protein (SEQ ID No. 3, 4). Several isolated peptides could not be found in the encoded sequence. This could indicate that alternative xylanase proteins are present in UV18-25. In previous analyses, no evidence was found for the presence of CBD in this 30 kD protein. Also from these results we hypothesized that the CBD of the protein is cleaved off by proteolysis. This hypothesis will be analysed further (by determination of activities, N-terminal sequences and sizes of the different proteins in the different C1 strains: C1 wild type, NG7C, UV13-6, UV18-25 and protease' mutants of UV18-25). Also the effect of the presence or absence of the CBD on enzymatic activities is analysed in more detail. Overexpression of the full length genes in various C1 hosts may be considered.

The presence of a cellulose-binding domain (CBD) is a particular feature of this enzyme. The only other known family 10 glycolytic enzyme (xylanase) having an N-terminal CBD is the *Fusarium oxysporum* Xyl1. The CBD of the *Chrysosporium lucknowense* Xyl1 protein has the sequence: WGQCG GQGWT GPITC VSGAY CQEVN DWYSQ CV (amino acids 22-53 of SEQ ID No. 4). This sequence does not comply to the CBD consensus sequence described in US 5,763,254 (Novo).

WO 01/79507

-28-

PCT/NL01/00301

This consensus sequence of US 5,763,254 is: W/Y-G/A-Q-C-G G-Q/I/N-G/N-W/F/Y-S/T/N/Q G-P/A/C-T/R/K-T/C/N-C X-X-G/P-S/T/F/L/A/- -T/K C-V/T/R/E/K-K/Q/A-Q/I-N Q/D/A-W/F/Y-Y-Y/S/H/A-Q C-L/I/Q/V/T, wherein W/Y means either W or Y etc., X means any amino acid, and - means absent. Four differences with the most degenerate consensus sequence are present in Xyl1, which are underlined in sequence 7 above. The invention thus pertains to xylanases having an N-terminal CBD different from this consensus CBD and other than the *Fusarium oxysporum* xylanase. More particularly the xylanase of the invention has a CBD having at least 55%, especially at least 65%, preferably at least 75% sequence identity with the sequence 7 above. Preferably the CBD contains one of the amino acids Phe, Tyr and Trp at position 23, or at least one of the four amino acids Val at position 20, Gln at position 22, Phe at position 23, and Val at position 24. Preferred sequences comprise Cys-Xaa-Phe, Xaa-Phe-Val, Cys-Xaa-Phe-Val, Cys-Gln-Phe, Val-Cys-Xaa-Phe, Gln-Phe-Val, Gln-Trp-Val, Gln-Tyr-Val, Val-Cys-Gln, Val-Cys-Gln-Phe, and Val-Cys-Xaa-Phe-Val, wherein Xaa is any amino acid or preferably Val, Thr, Arg, Glu, Gln or Lys, or most preferably Gln or Glu.

The xylanase does not possess MUF cellobiase activity and is thus a true xylanase. It possesses high activity within a broad pH range from 5-8 maintaining 65% of maximum activity at pH 9-10; it is a member of the xylanase F family. The corresponding xylanase promoter, which is a preferred embodiment of the invention, is shown in SEQ ID No. 3. The Michaelis constant towards birch xylan is 4.2 g/l for the 30kD xylanase. Temperature optimum was high and equal to 70°C for the xylanase.

20 Gpd1

The DNA sequence of the C-terminal part of the *gpd1* gene is not determined. The promoter sequences of this gene is a preferred embodiment of the present invention and is depicted in SEQ ID No. 5.

The expression level of four *Chrysosporium* genes was studied by Northern analysis. Various strains of *C. lucknowense* were grown in rich medium containing pharmedia with cellulose and lactose (medium 1) or rich medium containing pharmedia and glucose (medium 2) at 33°C. After 48 h, mycelium was harvested and RNA was isolated. The RNA was hybridised with 4 different probes: EG5, EG6, Xyl1 and CBH. After exposure, the Northern blots were stripped and hybridised again with a probe for ribosomal L3 as a control for the amount of mRNA on the blot. Most strains showed very high response for CBH1 and high response for Xyl1 in medium 1; in medium 2, half of the strain showed high response for all genes, and the other half showed low response. The order of expression strength was deduced from these data as CBH > Xyl1 > EG5 > EG6.

The protein Xyl1 of *C. lucknowense* is 67% identical (72% homologous) to its closest homologue in the Genbank DATABASE (Table 4). The strong homology of the CBH1 protein to its related *Humicola grisea* homologue (74% identical / 82% homologous) is noteworthy. It is also noted that in

WO 01/79507

-29-

PCT/NL01/00301

all cases the closest homologues originate from *Fusarium*, *Humicola* or other Pyrenomycetous (Sordariomycetous) fungi (Table 4), whereas *Chrysosporium* belongs to the Plectomycetous (Eurotiomycetous) fungi according to the NCBI database (Table 4). Thus the invention also pertains to glycanolytic enzymes, especially cellobiohydrolases and xylanases comprising a CBD, derived from

5 plectomycetous fungi.

	glycosidase family	isolated from Cl	number of amino acids	introns	remarks	related sequences (%identity / %homology)
CBH1	7	70kD 55kD	526 (63kD)	1	CBD	<i>Humicola grisea</i> (74 / 82) (CBH1: P15828) <i>Fusarium oxysporum</i> (58 / 68) (CBH: P46238) <i>Neurospora crassa</i> (60 / 69) (CBH1: P38676)
XYL1	10	30kD	333 (43kD)	3	CBD	<i>Fusarium oxysporum</i> (67 / 72) (XylF: P46239) <i>Penicillium simplicissimum</i> (63 / 72) (XylF: P56588) <i>Aspergillus aculeatus</i> (61 / 70) (XylF: Q59859)
GPD1	-	-	incomplete	2+?	-	<i>Podospira uniseriata</i> (85/89) (GPD: P32637) <i>Neurospora crassa</i> 80/86) (GPD: U67457) <i>Cryphonectria parasitica</i> 80/85) (GPD: P19089)

Table 4: Structural and comparative data of CBH1, Xyl1, and GPD1 of the invention.

10 Description of the figures

Figure 1 is a pUT1064 map

Figure 2 is a pUT1065 map

References (the contents hereof are incorporated)

1. Calmels T.P., Martin F., Durand H., and Tiraby G. (1991) *Proteolytic events in the processing of secreted proteins in fungi*. J Biotechnol 17(1): p. 51-66.
2. Punt P.J., Dingemans M.A., Jacobs-Meijsing B.J., Pouwels P.I.L., and van den Hondel C.A. (1988) *Isolation and characterization of the glyceraldehyde-3-phosphate dehydrogenase gene of Aspergillus nidulans*. Gene 69(1): p. 49-57.
3. Shoemaker S., Schweickart V., Ladner M., Gelfand D., Kwok S., Myambo K., and Innis M. (1983) *Molecular cloning of exo-cellobiohydrolase I derived from Trichoderma reesei strain L27*. Bio/Technology Oct.:691-696.
4. Drocourt D., Calmels T., Reynes J.P., Baron M., and Tiraby G. (1990) *Cassettes of the Streptococcus hindustanus ble gene for transformation of lower and higher eukaryotes*

WO 01/79507

-30-

PCT/NL01/00301

- to phleomycin resistance. *Nucleic Acids Res* 18(13): p. 4009.
5. Mullancy E.J., Hamer J.E., Roberti K.A., Yelton M.M., and Timberlake W.E. (1985) *Primary structure of the trpC gene from Aspergillus nidulans*. *Mol Gen Genet* 199(1): p. 37-45.
 6. Yanisch-Perron C., Viera J., and Messing J. (1987) *Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors*. *Gene* 33:103-119.
 7. Durand H., Baron M., Calmels T., and Tiraby G. (1988) *Classical and molecular genetics applied to Trichoderma reesei for the selection of improved cellulolytic industrial strains*, in *Biochemistry and genetics of cellulose degradation*, J.P. Aubert, Editor. Academic Press. p. 135-151.
 8. Lowry O.H., Rosebrough N.J., Farr A.L., and Randall R.J. (1951) *Protein measurements with the folin phenol reagent*. *J. Biol. Chem.* 193: 265-275.
 9. Parriche M., Bousson J.C., Baron M., and Tiraby G. *Development of heterologous protein secretion systems in filamentous fungi*, in *3rd European Conference on Fungal Genetics*. 1996. Münster, Germany.
 10. Baron M., Tiraby G., Calmels T., Parriche M., and Durand H. (1992) *Efficient secretion of human lysozyme fused to the Sh ble phleomycin resistance protein by the fungus Tolypocladium geodes*. *J Biotechnol* 24(3): p. 253-266.
 11. Jeenes D.J., Marczinke B., MacKenzie D.A., and Archer D.B. (1993) *A truncated glucosylase gene fusion for heterologous protein secretion from Aspergillus niger*. *FEMS Microbiol. Lett.* 107(2-3): p. 267-271.
 12. Stone P.J., Mekoff A.J., Parish J.H., and Radford A. (1993) *Cloning and sequence-analysis of the glucosylase gene of neurospora-crassa*. *Current Genetics* 24(3): p. 205-211.
 13. Mörsky P. (1983) *Turbidimetric determination of lysozyme with Micrococcus lysodeikticus cells: Reexamination of reaction conditions*. *Analytical Biochem.* 128:77-85.
 14. Paluh J.L., Orbach M.J., Legerton T.L., and Yanofsky C. (1988) *The cross-pathway control gene of Neurospora crassa, cpc-1, encodes a protein similar to GCN4 of yeast and the DNA-binding domain of the oncogene v-jun-encoded protein*. *Proc. Natl. Acad. Sci. USA* 85(11): p. 3728-32.
 15. Nakari T., Onnela M.L., Ilmen M., Nevalainen K., and Penttilä M. (1994) *Fungal promoters active in the presence of glucose*, WO 94/04673, Alko.
 16. Torronen A., Mach R.L., Messner R., Gonzalez R., Kalkkinen N., Harkki A., and Kubicek C.P. (1992) *The two major xylanases from Trichoderma reesei: characterization of both enzymes and genes*. *Biotechnology (N Y)* 10(11): p. 1461-5.
 17. Farkas V. (1985) *Novel media for detection of microbial producers of cellulase and xylanase*. *FEMS Microbiol. Letters* 28:137-140.
 18. Miller G.L. (1959) *Use of dinitrosalicylic acid reagent for determination of reducing sugar*. *Anal. Chem.* 31:426-428.
 19. Punt P.J., Mattem I.E., van den Hondel C.A.M.J.J. (1988) *A vector for Aspergillus transformation conferring phleomycin resistance*. *Fungal Genetics Newsletter* 35, 25-30.

WO 01/79567

-31-

PCT/NL01/00301

SEQ ID No. 1:

DNA sequence and amino acid of complete *Chrysosporium CBH1* gene including promoter and terminator sequences. Promoter sequence (1-1779), terminator sequence (3427-4451) and intron sequence (2179-2256) are given in small case.

```
aaggtatccgatttggggaacgtcgatgaaagtattgcaaaagtgcgagagttgcgcaa 60
ctaaactcgctgcgaagaagctgcggaagaagagacacccgaaagtggaaataacgttac 120
ggatgccttgacctcaaaagttgaaaccagcccttctgtctatttgggaaagcggttg 180
cccttgaatgcgctgcactgtgacacgactaccagtgatcgaggagcaaacctaccctg 240
gtcctgttcttgggtggggcactagggccaaacttaggggtgatcgagggtcgatcgccg 300
ggctcctgttgggtctgggtctcttctcatttcccggtttgcaccccccgttgacactgtc 360
atcgcccccgaacgcgatgaggttgcgcccagacgacacatcacccgggtgcattccc 420
aagtatatggaagtgaccagggtacccgggttttgcgtcccagtcgttgggtgccaat 480
ttgggagtttttgagcctcaagatctggggaatcgacctcaacttccatataagttaaa 540
gtcgacacacggcgagttccacgaagagacacattttttctgaaggcctctctccccg 600
cacatcagaacacacaaataccaagactgcagaagccgggttaagtgggccaccgggac 660
tacactaaaatgcggggaagagcgagatccgttcgcaagggaagggtgggtgtgtctgc 720
ggctttctcgcgtctcgtgcgccttttgcttgatctagtgtacacagggtaggctccg 780
aaggagtatctacggcagcgctgttctgtctgcgttgagagtcaggcgagacgagcag 840
gcgacaggagcctcgcacgggaccttcggatcgatttgcgcgagcgctcaaatacgctc 900
ttctgcggtcatcagagagcatcgtgaaccaagggttcttcgcgaggcggtctgggttc 960
gcagagtcgcactcggcggaacgccttccgtgtcaccctcgataacctggctgcgcgcgc 1020
agactcctccaatgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
ccatcgacgggtccacgtctctcgaaacgatgccttcagtcctgcggcggttggaacc 1140
ataacgtgacatcgccgcccagcctactagccgctatcgccgggttaggcttgtcaccgc 1200
agcgcctattctcctcggggtctactctgatccacctcaccacccgcaagcactagcg 1260
agcctcaccagagtgcaagcgacacgacccgcttggcccttcgtccttgactatctcca 1320
gacctcttgccatcttgcgaagcgcccccctttttttctcctccctgcggcgaggt 1380
cgggtggcccgagtcggagatggcattgtctcgttgtccatgacgacccatcattcgatg 1440
gctgactggcacactcgtcttgtttgagcatcgacgcccgcggcccgctctccacggtg 1500
cggaaacctgttgtacagtaacctctcgttaataatgatacccaacccggggccgagcgctggg 1560
agggcgggcgttcccgagaagccgggaagcggtggtggcggtgacctttgtgacttggtg 1620
atgcatcgggccatggagaatgtctctcgaagcgacgacgaatagcctggctaccat 1680
```

PCT/NI 91/00301

cgatataaattgggtgattccagcgtcttgtagggcggtgtctctcgctggcagccctcg	1740
tcttcagatcaagcaactgtgtgctgatcctcttcgcgcATGTACGCCAAGTTCGCGACC	1800
M Y A K F A T	
CTGCGCGCGCTTGTGGCTGGCGCGCGTGTCTCAGAAACGGCTGCACCTTCGACCCTGAGAAC	1860
L A A L V A G A A A Q N A C T T L T A E N	
CACCCCTCGCTGACGTGGTCCAAAGTCAGCTGTGCGCGCAGCTGCACCAGCGCTCAGGGT	1920
H P S L T W S K C T S T G S G C T S V Q	
TCCATCACCATCGAGCGCAACTGCGCGGTGGACTCACCGSACGATGACGGCCACCAACTGC	1980
S I T I D A N W R W T H R T D S A T N C	
TACGAGGGCAACAGTGGGATACTTCGTACTGCAGCAGTGGTCTCTTCTTGCGCTCCGAAG	2040
Y E G N K W D T S Y C S D G P S C A S K	
TGCTGCATCGACGGCGGCTACTCTGCAGCACTATGGCATCCACGAGCGGTAACCTCC	2100
C C I D G C A D Y S S T Y G I T T T S G N S	
CTGAACCTCAAGTTCGTCCACAAAGGGCCAGTACTCGACCAACATCGGCTCGCGCTACCTAC	2160
L N L K F V T K G Q Y S T N I G S R T Y	
CTGATGGAGAGCGACACCAAGTACAGAGtaagttctctctcgccacggcgcccgcgggaga	2220
L M E S D T K Y Q M	
tgtatggcgccagcccgctgacgcgaatgacacacGTGTTCACGCTCTCGGCAACGAGTT	2280
F Q L L G N E F	
CACCTTCGATGTGCGAGTCTCCAAACCTCGGCTCGGGCTCAATGGCGCCCTCTACTTCGT	2340
T F D V D V D S N L G C G L N G A L Y F V	
GTCCATCGSAGTGGCGGATGGTGCACTGCCGCAACGAGCGAGGTSCCAAGTA	2400
S M D A D G G M S K Y S G N K A G A K Y	
CGGTACCGGCTACTGTGATTCTCAGTGGCCCGCGCACTCAAGTTCATCAACGGCGGAGGC	2460
G T G C C D S Q C P R D L K F I N G E A	
CACCGTAGAGAACTGGCAGAGCTCGACCAACGATGCACACCGCGGCAAGGCAACGG	2520
N V E N W Q S T G C N D A N A G T G K Y G	
CAGCTGCTGCTCCGAGATGGAGCTCTGGAGGSCCAACAACATGGCCCGCGCTTCACTCC	2580
S C C C S E M D V W E A N K M A A A A F T P	
CCACCCTTGCACCGTATCGGCCAGTCGCGTGCAGGGCGGACTCGTGCGGCGGTACCTA	2640
H P C W V I G C S R C E G D S C G G T Y	
CAGCACCAGCCCGTATCGCGGCATCTGCGACCCCGAGCGGATGCGACTTCAACTCGTACCG	2700
S T D R Y A G I C D F D G C D F N S Y R	
CCAGGCGCACAGACCTTCTACGGCAAGGCACTGACGGTGCACACGACCAAGATGCAC	2760
Q G N K T F Y Y G K A G G E L S E I K R F Y	
GGTGTGCACCCAGTTCCTCAAGAGACTCGGCCGCGGCGAGCTCTCCGAGATCAAGCGGTTCTA	2820
V V T Q F L K N S A G E L S E I K R F Y	
CGTCCAGACGACGAGGTCTCCCAACTCCGAGTCACCATCCCGGGCTCGGAGGCAAA	2880
V Q N G K V I P N S A E S T I P G V E G N	
CTCCATCACCACGAGGATGGTGCAGCCGACGCGCGCGCTTCGGCGAGCTGACCGACTT	2940
S I T Q D W C D R Q K A A F G D V T D ?	

WO 01/79507

-33-

PCT/NL01/00301

NCAGGACAAGGGCGGGCATGGTCCAGATGGGCAAGGCCCTCGCGGGGGCCCATGGTCTCGT 3000
 Q D K G G M V Q M G X A L A G P M V L V
 CATGTCCATCTGGGACGACCGCCGTCAACATGCTCTGGCTCGACTCCACCTGGGCCAT 3060
 M S I W D D H A V N M L W L D S T W P I
 CGACGGCGCGGGCAAGCCGGGGCGCGAGCGCGGTGCCTGCCCCACCACCTCGGGCGTCC 3120
 D G A G K P G A E R G A C P T T S G V F
 CGCTGAGGTCGAGGCGGAGGCCCACTCCAACTGTCATCTTCTCCAACTCCGCTTCGG 3180
 A E V E A E A P N S N V I F S N I R F G
 CCCCATGGGTCCACCGTCTCCGGCTGCCCGACGGCGGCGGCAACCCCAACCCGCC 3240
 P I G S T V S G L P D G G S G N P N P F
 CGTCAGCTCGTCCACCCCGTCCCTCTCGTCCACCACATCTCCGGTTCCTCCGGGCC 3300
 V S S S T P V P S S S T T S S G S S G P
 GACTGGCGGACGGGTGTGCTAAGCACTATGAGCAATGCGGAGGAATCGGGTTCAGTGG 3360
 T G G T G V A K H Y E Q C G G I G F T G
 CCGTACCGCAGTGGGAGAGGCCCTACACTTGCACCAAGCTGAATGACTGGTACTCGCAGT 3420
 P T Q C E S P Y T C T K L N D W Y S Q C
 CCTGTAAAGcaaccctctctgaaggaggttctgagacacgcgcgaattctctgtatatagt 3480
 L
 tttatttttcaactcttgagtgcttctgctccaccagtagacataaacctttttttcacgtaa 3540
 caaaatggcttcttttcagacacatgtgaaccatcttgatgccttgacctcttcagttctc 3600
 actttaacgttagttcgcgtagtctgtatgtccagttgcatgtagttgagataaatacc 3660
 cctggaagtgggtctgggcctttgtgggacggagccctctttctgtggtctggagagccc 3720
 gctctctaccgcctaccttcttaccacagtacactactcacacattgctgaactgaccca 3780
 tcataccgtactttatcctgttaattcgtgggtgtgtgcgactattctatttgcctcaaatg 3840
 gagagcacattcatcggcgcaggatacacgggtttatggacccaagagtgtaaggacta 3900
 ttattagtaatatattatagcctctaggcgccttaacttcaacaggcgagcactactaatc 3960
 aactttttggtagacccaattacaaacgacacacgtgcccggaaattttgggattccgtcc 4020
 gctctccccaaacgaagctagaagaggcaacgaacagccaatccgggtgctaattaaatta 4080
 tatgggttcatttttttaaaaaaattttttcttccattttctctctcgtttttctttttc 4140
 gcatcgtagttgatcaagtcgaagtcgaagcgagctatttgtgctatagctcggcggcta 4200
 taatcagtagcgttagagaggtctgaaggtatgataccacagcagctattcgcgctata 4260
 agcggcactcctagactaattgttcgggtctacagaagtaggtatataaagcgttaattg 4320
 ttctaaatactagaggcacttagagaagctatctaatatctattgacctagcttatta 4380
 tccctattagtaagttagttagctctaacctctatagtagcacaatgctataataggtacc 4440
 aggggttcaaaa 4451

WO 01/79587

-34-

PCT/NL01/00301

SEQ ID No: 2:

Amino acid of complete *Chrysosporium* CBH1 protein. The putative signal peptide (1-19) is shown in italic letters and the cellulose binding domain (496-526) is shown in bold underlined letters.

```

MYAKFATLAA LVAGAAQNA CTLTAENHPS LIWSKCTSGG SCTSVQGSIT 50
IDANRWTHRH TDSATNCEYEG NKWDTSYCSG GPSCASKCCI DGADYSSTYG 100
ITTSGLSLNL KEVTKGQYST NIGSRITYLME SDTKYQMFQL LGNEFTFDVD 150
VSNLGCGLMG ALYFVSMADG GMSKYSYGNK AGAKYGTGYC DSQCPRDLKF 200
INGEANVENW QSSTNDANAG TGKYGCCSE MDVWEANNMA AAFTHPC?V 250
IGQSRCEGDS CGGTYSTDRY AGICDPDGCD FNSYRQGNK FYGKGMTVDT 300
TKKITVVVQT LKNSAGELSE IKRFYVQNGK VIPNSESTIP GVEGNSITQD 350
WCDRQKAATG DVID?QDRGG MVQMGKALAG PMVLVMSIWD DIAVNMLWLD 400
STWPIDGAGK PGAERGACPT TSGVPAEVEA EAPNSNVIFS NIRFGPIGST 450
VSGLPDGGSG NPNPPVSSST PVPSSSTSS GSSGPTGGTG VAKHYQCQG 500
IGFTGPTQCE SFYTCTKLND WYSQCL * 526

```

SEQ ID No. 3

DNA sequence sequence of complete *Chrysosporium* Xyl1 gene including promoter and terminator sequences. Promoter sequence (1-969), terminator sequence (2428-3030(3028)) and intron sequences (1043-1116, 1181-1332(1331), 1596(1595)-1674(1672)) are given in small case.

```

tcacaaacttggcggttttgatgtactaataattacacgtcgtttgcnagcggagtcgtg 60
tcacatccggtgggttcgggtgctccagacgacgcttcgggcccgtatcctgaattcgggaag 120
gaaacggttcgggtaatacaggtcctctaaaaataaacgaagcactacagagggagttcct 180
cagaggacatcgtatcaaccgaagaacgaagcgcggaaggaactgatcaaaacaggagta 240
ggtagggatgtgtgagtacctaactttccatacctgacataaaatcatcatggtgcttc 300
agacctgtttgatgagcgaggcgaggcgagccgcattgtatthttcgttccttccttcttl 360
tgttagtatatctnagggttcctcgttaaatggaatctccagctctactagtaattag 420
aacaatagtctctgctgtcgtgcgcaagcttttcagatgactgccaaaaaccatcatg 480
ggataggacaaaagcagtaatcggagtcacacgcgcgcatthttccttcattgattccgtc 540
aacgggagaggtcggaggaggaactccggccacatgtgatgcqaaagagtacatggcgcca 600
tgggttcaacctcttatagttctgaaaatgcgcggagggccagcgaagccaaagccgggaac 660
cgttcctgtcctgggttcagtatgtttcgcctaaacatctctacgattccgcalaggtg 720
cggcgcaccacgaaggttgtatccctaaagctttggtaagtcggagtaacgaatggaa 780

```

WO 01/79507

-35-

PCT/NL01/00301

```

acgcgcgcagtcctgggtccatcggtatcctccgcagtcctccgcaaaaaaagaaaacc 840
cggtgatgtttacaaggataataagagacaagatgcaccacccccccttcccatctgc 900
cggttgcaccaagtcgcgcgactgcttgcgcgttccctacctgcagccctctttcagaga 960
ccatcaaacATGCGTACTCTTACGTTGCTGCTGCGCAGCCGCCCGGTGGCTGCTTSCC 1020
      M R T L T F V L A A A P V A V L A
CAATCTCCTCTGTGGGGCAGTgtatgtatgtccttactcggaaaatagtaccactag 1080
      Q S P L W G Q C
agggaacttaagctcactacttctgttttcacaatagGCGGCGGTCAAGGCTGGACAGGTC 1140
      G G Q G W T G
CCACGACCTGCGTTTCGGCGCAGTATGCCAATTCGTCAAGtcagtaactgcttttatt 1200
      P T T C V S G A V C Q F V N
tcttttctctctgggattacgatttctgttgcacttagcttgcctctctgctttcattgt 1260
tgtattgttctcttttctgtgtgagaggttttattaccacctaaaggccatttgcataac 1320
aaatctcccccagTGACTGCTACTCCCAATGCGTGCCGGATCGAGCAACCTCCTACGGG 1380
      D W Y S Q C V P G S S N P P T G
CACCCACGACGACCACTGGAAGCACCCCGGCTCCTACTGCGGCGGCGGCGAGCGGAAC 1440
      T T S S T T G S T P A P T G G G G S G T
CGGCTCCACGACAAATTCAGGCCAAGGGCAAGCTCTACTTCGGAACGAGATCGATCA 1500
      G L H D R K F K A K G K L Y F G T E I D H
CTACCATCTCAACAACATGCCTTGACCAACATTGTCAGAAAGACTTTGGTCAAGTCAC 1560
      Y H L N N N A L T N I V K K D F G Q V T
TCACGAGAACAGCTTGAAGTGGGATGCTACTGAGCGtgagtgaccttctctctctctccc 1620
      H E N S L K W D A T E F
gacaataatagataattacgagccgggttcgaggctgacattgcgcgattcttagCGAGCC 1680
      S R
GCAATCAATTCACCTTTGGCAACGCCGACCGCGGTGTCACCTTTGCCAGGCCAACGGCA 1740
      N Q F N F A N A D A V V N F A Q A N G K
AGCTCATCCGCGGCCACACCTCCTCTGCGACTCTCAGCTGCGCGCAGTGGGTSCAGAACA 1800
      L I R G H T L L W H S Q L P Q W V Q N I
TCAACGACCGCAACACCTTGACCCAGGTGATCGAGAACCACTCACACCTTGTCACTC 1860
      N D R N T L T Q V I E N H V T T L V T R
GCTACAGGGGCAAGATCCTCCACTGGGACGTCGTTAAGGAGATCTTTGCCGAGSACGGCT 1920
      Y K G K I L H W D V V N E I F A E D G S
CGCTCCGCGACAGCGTCTTCAGCCGCGTCTCGGCGAGGACITTTGTCGGCATCSCCTTCC 1980
      L R D S V F S R V L G E D F V G I A F R
GCGCGCGCGCGCGCGATCCCAACGCCAAGCTCTACATCAACGACTACAACTCGACA 2040
      A A R A A D P N A K L Y I N D Y N I D I
TTGCCAACTACGCCAAGTGACCCGGGGCATGGTCGAGAAGGTCAACAAGTGGATCGCCC 2100
      A N Y A K V C R G M V E K V N K W I A Q

```

WO 01/79507

-36-

PCT/NL01/00301

```

AGGGCATCCCGATCGACGGCATCGGCACCCAGTGCCACCTGGCCGGGCCCGGGGGTGGG 2160
G I P I D G I G T Q C H L A G P G G W N
ACACGGCCCGCGCGCTCCCGACGCCCTCAAGGCCCTCGCCGCCCAACGTCAAGGAGA 2220
T A A G V P D A L K A L A A A N V K E I
TGCCCATCACCGAGCTCGACATCGCCGGCGCTCCGCCAACGACTACCTCACCGTCATGA 2280
A I T E L D I A G A S A N D Y L T V M N
ACGCCCTGCTCCAGGTCTCCAAAGTGGCTGGCATCAGGCTCGGGGGTCTCTGACAAGG 2340
A C L Q V S K C V G I T V W G V S D K D
ACAGCTGGAGGTCGAGCAGCAACCGCTCCTCTTCGACAGCACTACCAAGCCAAAGGCGG 2400
S W R S S S N P L L F D S N Y Q P K A A
CATACCAATGCTCTGATTAAATGCTTGTAAgaggaggtatattatttttagaggcaatgaa 2460
Y N A L I N A L *
gctaggaggaaagaggggaagtgaggtatattagctaggacaggcaaatctagcagcaatt 2520
ataagtcaacactatataaaatattcctataatggcttgtgcttcggtgtgcaaaaaaa 2580
aaaaaaaaaaaaaaaaaaaaaaaaaactcaaaaacaaaatgatccaacatgatt 2640
cgaaatggcgaccttgcaaatgcacacctcagataataccactatacaatacaccttaaa 2700
tggcacctaaatccatttctgctcggtcatagacggggttaagaagcctgggatgcagg 2760
tgtcgatgcaagggttacgtcagtgatgatgatgatgaacatgctgtctgggtaat 2820
tctccactttccctcccttacgactcttcgggtgtgctctctagaagtcgactcctg 2880
gcgcctcagatcgccctttggctctgttcggtacaatgacgtccgctggtttcttccaaa 2940
gaccagggtatttctccctgggcaacaaagaataccaaatacctatatcgaaacgtagtct 3000
tctgataattagatgtctctcaaggcgagg 3030

```

SEQ ID No. 4

Amino acid sequence of complete *Chrysosporium Xyl1* protein. The signal peptide (1-26) is shown in italic letters and the cellulose binding domain (22-53) is shown in bold underlined letters.

```

1  MRLITFVLAA APVAVLAQSP LWGQCGGQGN TGPTTCVSGA VQGFVNDWYS
51  QCVPGSSNPP TGTTSSTTGS TPAPTGGGGS GTGLHDXFKA KGRLYFGTEJ
101 DHYHLNNNAL TNIVKDFGQ VTENSLKWD TEPSRNQFNK ANADAVVNEA
151 QANGKLIRGH TLLMHSQLPQ WVQNINDRNT LTQVIENHVT TLVTRYKGI
201 LHWDVVNEIF AEDGSLRDSV FSRVLGEDFV GIAFRAARAA DFNAKLYIND
251 YNLDIAKYAK VTRGMVEKVN KWIAQGIPID GIGTQCHLAG PGGWNTAAGV
301 PDALKALAAA NVKEIAITEL DIAGASANDY LTVMNACLQV SKCVGITVWG
351 VSDKDSWRSS SNPLLFDSDNY QPKAAYNALI NAL*

```

WO 01/79507

-37-

PCT/NL01/00301

SEQ ID No: 5

DNA sequence of partial *Chrysosporium* GPD1 gene including promoter sequences. Promoter sequence (1-1555) and intron sequence (1682-1781) are given in small case. The 3' end of the gene is lacking.

```
tgagcagcaatgagcagcaatgagcattcctgggccaccgagctcgtgagtgccagtagcga 60
gtatcgtactctcgtacccgggtttgatttggtgacggtgcttttcacctctcgatgcccg 120
aaatcgggtctaaagctgagtttgatcaaatatgtgactccaacatcgcccccctcggcaa 180
accccgctcgacacgtgtgtcatccttcattgcaagcgatcactcgaggcggtgacgat 240
gaacgagatttttgccgggaccgatttcgggatatagcggcagccgaccagccctaccac 300
actgatggcgtgtcactagtgtatgtctccagaaccgcaagcatacactgggcaatgct 360
tggtatgcagttagggcagctttatgtttccatacccttcaccttcgggtcggggactcg 420
gcggggtcgggaagtttgacggcagccgctggggccttagggcagattacgtggtgt 480
ggccagttttagccgttcocgctcgtttctaccggaccatgattttcgtgaacattg 540
caatcccgaaagcccatctccgacgttaaggagttacctocgctgccagaattcatgac 600
gtggcgggtcaaggcagcgtggtgggggcatccgtgtcaagctccaggaggaggtgcgc 660
gatttcaaatccgggccaacacagcccaagactggctggccaaaaaaggagcgtagacg 720
gcccgggacatcgggacgtcagctcgcagccacccaaaaaccggtccgatctactcgcttc 780
tgtggtagttcagggtacttttgagtagtaaaaaacgctacggcagggccggggggttccc 840
ggtgacggaggtgcctctgcgggtggcgaacatccacgcactctcgcgctacggtgacac 900
ctcgtgtcctgttggtctttgcaatgctggggcggcaggaaatgcgtcgcgctcctcccg 960
ccaagacctaaaacagacagcgcgcgcaaatgcgtcactagcaccgcgaaacggaatgc 1020
cccacctcaacgcaatctgtgatgcaagcaattgggaaggctcaccaccctcagcgagg 1080
ggctcaaccatttttattatcagctcatgcccacacaacatgactgtttttcttcttgc 1140
tcatcccatatttgacaaaaatcgtcgttaattctcttccatacaggccgtccgcgtc 1200
tgataaccacataaaaagctctcttcagtcacacgctcaaaagctccctcctccctcaggta 1260
agcagccaaaagctcccccacggaccccgcactgctcctccgcctgtatcggaacctg 1320
cgcgacccagcagagaaatcccaaacctttgctgcttgctgcccgggtccggactgagctg 1380
caaccccaagctcttaaaaagcttttcccttctccacgggtgtcaactctgtcttatccct 1440
ccgacatccgttgagctcaacaaactcccgaaaccttttaccocgcgcgagctacccctc 1500
catcaaacccacctgacagctcgtcactcacctccccacatcacagaaatcaaaATGAC 1560
M T -
TATCAAGGTCGGCATCAACGGTTTCGGCCGTATCGGCCGTATCGTCTTCGGCAACTCCAT 1620
I K V G I N G F G R I G R I V F R N S I -
CGAGCACZCGGATGTCGAGATCGTTGCCGTCACGACCCCTTCATTCAGGCCAAGTACGC 1680
```

WO 01/79507

-38-

PCT/NL01/00301

E H S - D V E I V A V N D P F I E F K Y A -
 Tgtaagtagtatttttttcttccctcgcggtttcttctgttccatcgacagtagagat 1740
 GatcttgcaggcggtatcggaagctaacccgcatgttcgtacagGAGTACATGCTCAAGTAT 1800
 E Y M L K Y -
 GACTCGACCCACGGTATCTTCAACGGCACCATCGCCGTCGAGGGCAACGACCTCATTGTC 1860
 D S T H G I F N G T I A V E G N D L I V -
 AACGGCAAGAGGGTCAAGTTCTACACTGAGCGGGMCCCCCAACATTCCCTGGARGGAA 1920
 N G K R V K F Y T E R ? P A N I P W ? E -
 ACTGGTCCCGAGTACATMRTCGAGTCGACCGGTGTCTTCAACMCACCSAGAAGGCTAGC 1980
 T G A E Y I ? E S T G V F T ? T ? K A S -
 GCCCACCTCAAGGGGGGGCCCAAGCGGTCTCATCTCTCTCCCTCGGCGGATGCCCCC 2040
 A H L K G G A K R V I I S A P S A D A P -
 ATGTAGTCTATGGGCGTCAACGAGAAGCCTACGACGGCAAGGCCAGGTCTCTTAAC 2100
 M Y V M G V N E K T Y D G K A Q V I S N -
 GCCTCGTGCACCACTGCTGGTCCCTCGCCAAAGGTCTCCACGACAGTTCGGC 2160
 A S C T T N C L A P L A K V I H D K F G -
 CTCGTTGAGGGTCTCATGACCAACCGTCCACTCTCACTGCCACCCAGAAGCCGTCGAT 2220
 L V E G L M T T V H S Y T A T Q K T V D -
 GGTCCCTCTGCCAAGGACTGGCGTGGTGGCCGTGGTGTCTCAGAACATCATCCCCAGC 2280
 G P S A K D W R G G R G A A Q N I I P S -
 AGCACTGGCGCCGCAAGCCGTCGGCAAGGTCTCCCTGAGCTCAACGGCAAGCTCACC 2340
 S T G A A K A V G K V I P E L N G K L T -
 GGCATGTCCCTCCGTGTCCTCCACCCCAACGTTTCCGTGTGCGACCTCACCTGCCGCCTC 2400
 G M S L R V P T P N V S V V D L T C R L ~
 GAGAGGGAGGCTACCTAGCAGACATCAAGGCCGCCATCAAGGAGGCGCGCGCGGCCCC 2460
 E K E A T Y D D I K A A I K E A A A G P -
 CTCAGGGgtgagttatctgttcttttttttttttttggagaacgacacatgctgataaa 2520
 L K G
 acccagGCATCCTCGACTACACTGAGG 2547
 I L D Y T E

SEQ ID No. 6

Amino acid of partial *Chrysosporium* GPD1 protein (the C-terminus is lacking in the sequence available).

MTKVGINGF GRIGRIVFRN SIEHSDVEIV AVNDPFIEPK YAEYMLKYDS
 THGIENGITIA VEGNOLIVNG KRVKFYTER? PANIPW?ETG AEYI?ESTGV
 FT?T?KASAH LKGGAKRVII SAPSADAPMY VMGVNEXTYD GKAQVLSNAS
 CTTNCLAPLA KVIHOKFGLV EGLMTTVHSY TATQKTVDGP SAKDWREGRG
 AAQNIIPSSST GAAKAVGKVI PELNGXLTGM SLRVFT?NVS VVDLTCRLEK
 EATYDDIKAA IXEAAAGPLK GILDYTE

WO 01/79507

-39-

PCT/NL01/00301

CLAIMS

1. A protein corresponding to a *Chrysosporium* glycosyl hydrolase family 7, exhibiting at least 75% aminoacid identity [as determined by BLAST algorithm] with the amino acid sequence of SEQ ID No. 2 or a part thereof having at least 20 contiguous amino acids which are identical to the corresponding part of the amino acid sequence 1-246 or 394-526 of SEQ ID No. 2.
2. A protein corresponding to a *Chrysosporium* glycosyl hydrolase family 10, exhibiting at least 70% aminoacid identity [as determined by BLAST algorithm] with the amino acid sequence of SEQ ID No. 4 or a part thereof having at least 20 contiguous amino acids which are identical to the corresponding part of the amino acid sequence 1-133 or 252-383 of SEQ ID No. 4.
3. A protein corresponding to a *Chrysosporium* glycosyl hydrolase family 10, exhibiting at least 65% aminoacid identity [as determined by BLAST algorithm] with the amino acid sequence of SEQ ID No. 4 or a part thereof having at least 20 contiguous amino acids which are identical to the corresponding part of the amino acid sequence 1-383 of SEQ ID No. 4.
4. A protein corresponding to a glycosyl hydrolase family 10 and comprising a cellulose-binding domain having at least 55% aminoacid identity with the amino acid sequence 22-53 and/or having at least 20 contiguous amino acids from amino acid sequence 22-53 of SEQ ID No. 4, and/or having one of the aminoacids F, W and Y at relative position 44.
5. A protein corresponding to a *Chrysosporium* glyceraldehyde phosphate dehydrogenase, exhibiting at least 86% amino acid identity [as determined by BLAST algorithm] with the partial amino acid sequence of SEQ ID No. 6 or a part thereof having at least 20 contiguous amino acids which are identical to the corresponding part of the amino acid sequence 1-277 of SEQ ID No. 6.
6. A process of hydrolysing β -glucosidic bonds, comprising the use of an enzyme according to claim 1.

WO 01/79507

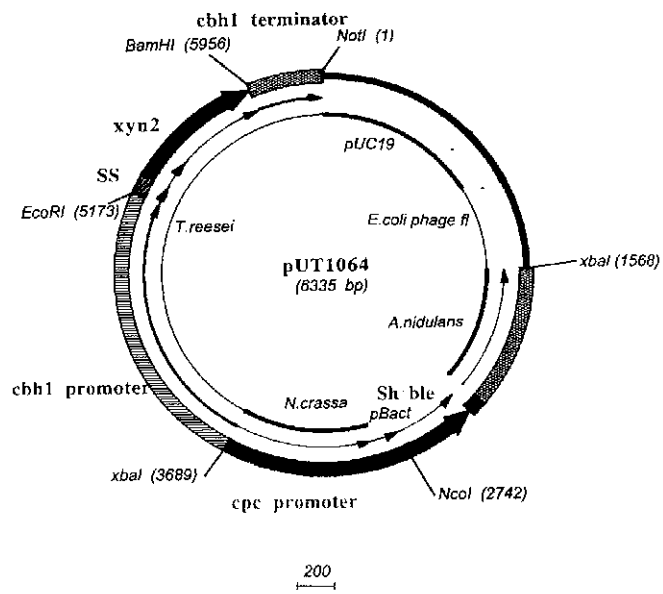
-40-

PCT/NL01/00301

7. A process of hydrolysing β -xylosidic bonds, comprising the use of an enzyme according to claim 2, 3 or 4.
8. A nucleic acid sequence encoding a protein according to any one of claims 1-5.
9. A nucleic acid sequence comprising at least 70% of the nucleotides contained in the 5'-non-coding region of the nucleic acid sequence of any one of SEQ ID No's 1, 3 and 5.
10. A nucleic acid construct comprising a nucleic acid expression-regulatory region derived from *Chrysosporium*, contained in the 5'-noncoding region of the nucleic acid sequence of any one of SEQ ID No's 1, 3 and 5, operationally linked to a nucleic acid sequence encoding a polypeptide of interest.
11. A recombinant microbial strain, preferably a fungal strain, containing a nucleic acid sequence according to any one of claims 8-10, and capable of expressing the polypeptide encoded by the coding nucleic acid sequence.
12. A process of producing a polypeptide using a construct according to claim 10 or a microbial strain according to claim 11.
13. An oligonucleotide probe comprising at least 15 contiguous nucleotides of the nucleic acid sequence of any one of SEQ ID No's 1, 3 and 5, or its complement.
14. An oligonucleotide probe according to claim 13, wherein the probe is 20-50 nucleotides in length.
15. An oligonucleotide probe according to claim 13 or 14, wherein the probe is labelled with a detectable label.

WO 01/79507

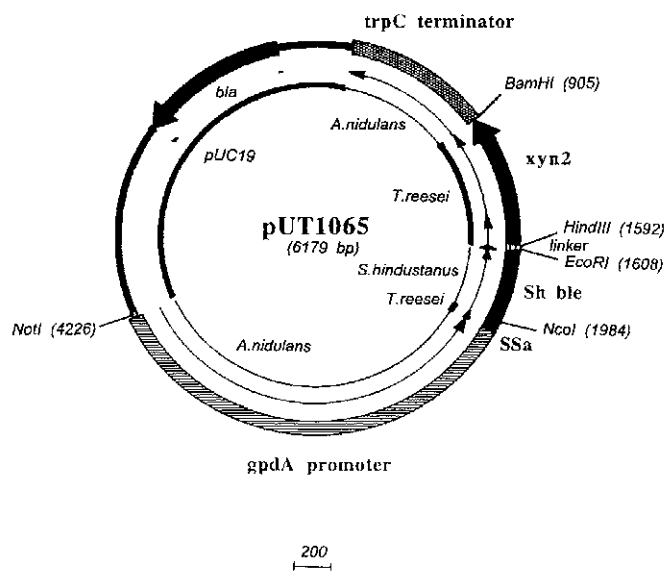
PCT/NL01/00301

Fig 1

WO 01/79507

PCT/NL01/00301

Fig 2



【国際公開パンフレット（コレクトバージョン）】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number
WO 01/79507 A3(51) International Patent Classification: C12N 15/56,
15/53, 1/15, C12Q 1/68, C12P 19/14, C12N 9/24, 9/42,
9/02, 15/89(74) Agent: JORRITSMA, Ruurd; Nederlandsch Octrooi-
bureau, Scheveningseweg 82, P.O. Box 29720, NL-2502 LS
The Hague (NL).

(21) International Application Number: PCT/NL01/00301

(81) Designated States (national): AF, AG, AI, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GE, GR, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LT, LK,
LR, LS, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RD, RU, SD, SE, SG, SI, SK, SL,
TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW.

(22) International Filing Date: 17 April 2001 (17.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 00201343.3 13 April 2000 (13.04.2000) EP

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW); Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM); European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR); OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).(71) Applicant and
(72) Inventor: EMALFARD, Mark, Aaron [US/US]; 493
Spuyss Court, Jupiter, FL 33477 (US).Published:
with international search report(72) Inventors; and
(75) Inventors/Applicants (for US only): PUNZI, Peter, Jan
[NL/NL]; Broekdrukkersgilde 1, NL-3994 XT Houten
[NL]; VAN ZELIJ, Cornelia, Maria, Johanna [NL/NL];
Hem, Dunantlaan 24, NL-3451 EP Vleuten De Meern
[NL].(88) Date of publication of the international search report:
7 February 2002For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/79507 A3

(54) Title: EXPRESSION-REGULATING SEQUENCES AND EXPRESSION PRODUCTS IN THE FIELD OF FILAMENTOUS
FUNGI *CHRYOSPORUM*(57) Abstract: The invention pertains to novel proteins corresponding to *Chrysosporium* glycosyl hydrolases of families 7 and 10,
exhibiting a minimum amino acid identity of 70 and 75 %, respectively, with the amino acid sequence of SEQ ID No's 2 and 4,
and to a protein corresponding to a *Chrysosporium* glyceraldehyde phosphate dehydrogenase, exhibiting at least 80 % amino acid
identity with the partial amino acid sequence of SEQ ID No. 6. The invention further relates to nucleic acid sequences encoding
these proteins, and especially to promoter sequences regulating the expression of the corresponding genes. The preferred host for
expressing these genes is a fungus, especially a *Chrysosporium* strain.

【国際調査報告】

INTERNATIONAL SEARCH REPORT					International Application No. PCT/NL 01/00301
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/56 C12N15/53 C12N1/35 C1201/68 C12P19/14 C12N9/24 C12N9/42 C12N9/02 C12N15/80					
According to International Patent Classification (IPC) or to both national classification and IPC					
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C12P					
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched					
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, BIOSIS, WPI Data, SEQUENCE SEARCH, EMBL, CHEM ABS Data, PAJ					
C. DOCUMENTS CONSIDERED TO BE RELEVANT					
Category *	Citation of document, with indication, where appropriate, of the relevant passages				Relevant to claim No.
A	WO 98 15633 A (CHERNOGLAZOV VLADIMIR MIKHAYLO ;OKOUNEV OLEG NICOLAEVICH (RU); SIN) 16 April 1998 (1998-04-16) cited in the application claims; examples 12-16 --- -/-				1, 6, 8-15
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.					
* Special categories of cited documents : "A" document determining the general state of the art which is not considered to be of particular relevance "T" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claims or which is cited to establish the publication date of another claim or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "I" later document published after the international filing date or priority date and not in conflict with the application but cited to substantiate the principle or theory underlying the invention "X" documents of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" documents of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "M" documents member of the same patent family					
Date of the actual completion of the international search			Date of making of the international search report		
25 October 2001			09.11.2001		
Name and mailing address of the ISA European Patent Office, P.O. Box 5516 Palantinus 2 NL-2200 LV The Hague Tel: (+31-70) 340-2040, Tx: 31 651 epo nl Fax: (+31-70) 340-3016			Authorized officer Le Cornec, N		

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

 Int. Appl. No.
PCT/NL 01/00301

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE EMBL HGC6H1A 'Online! accession number D63515; 21 August 1995 (1995-08-21) S. TAKASHIMA ET AL: "Cloning, sequencing, and expression of the cellulase genes of humicola grisea var. thermoidea" XP002147095 * 78,8% identity in 529 amino acid overlap with sequence ID no.2 and 81% identity in 1668nt with sequence ID no.1 (1765-3428:305-1958) * abstract -> S. TAKASHIMA ET AL: "Cloning, sequencing and expression of the cellulases genes of Humicola grisea var. thermoidea" JOURNAL OF BIOTECHNOLOGY, vol. 50, no. 2, 1 October 1996 (1996-10-01), pages 137-147, XP004037051	1,6,8-15
A	HARMSSEN MARTIN C ET AL: "Sequence analysis of the glyceraldehyde-3-phosphate dehydrogenase genes from the basidiomycetes Schizophyllum commune, Phanerochaete chrysosporium and Agaricus bisporus." CURRENT GENETICS, vol. 22, no. 6, 1992, pages 447-454, XP001030835 ISSN: 0172-8083 page 449; figure 3	5,8
X	RITTER R ET AL: "SEQUENCE ANALYSIS OF THE GENE CODING FOR GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE GPD OF PODOSPORA-ANSERINA USE OF HOMOLOGOUS REGULATORY SEQUENCES TO IMPROVE TRANSFORMATION EFFICIENCY" CURRENT GENETICS, vol. 21, no. 3, 1992, pages 207-213, XP001030808 ISSN: 0172-8083 the whole document	5,8,11, 12
A	--- -/--	13-15

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/NL 01/00301

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ERIKSSON K -E ET AL: "EXTRACELLULAR ENZYME SYSTEM UTILIZED BY THE FUNGUS SPOROTRICHUM PULVERULENTUM (CHRYSOSPORIUM LIGNORUM) FOR THE BREAKDOWN OF CELLULOSE. 1. SEPARATION, PURIFICATION AND PHYSICO-CHEMICAL CHARACTERIZATION OF FIVE ENDO-1,4-BETA-GLUCANASES" EUROPEAN JOURNAL OF BIOCHEMISTRY, DE, BERLIN, vol. 51, 1975, pages 193-206, XP002055449 ISSN: 0014-2956 the whole document	1,6
X	DATABASE EMBL 'Online! BUNF-FUSOX, accession number p46239, 1 November 1995 (1995-11-01) P.O. SHEPPARD ET AL: "the use of conserved cellulase family-specific sequences to clone cellulase homologue cDNAs from Fusarium oxysporum" XP002154884 abstract -s P.O. SHEPPARD ET AL: GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, vol. 150, 1994, pages 163-164, XP000647707 ELSEVIER SCIENCE PUBLISHERS, BARKING., GB ISSN: 0378-1119	4,11-15
A	WO 97 13853 A (VISSER JACOB ;DEN BROECK HENRIETTE CATHARINA (NL); GIST BROCADES B) 17 April 1997 (1997-04-17) Sequence ID no.7,8 page 21-23 claims	2-4,7-15
X	DATABASE EMBL 'Online! S71569, accession number S71569, 28 October 1994 (1994-10-28) G.W. BLACK ET AL: "xylanase b from Neocallimastix patriciarum contains a non-catalytic 455-residue linker sequence comprised of 57 repeats of an octapeptide" XP002181149 abstract & BLACK G.W. ET AL: BIOCHEM. J., vol. 299, no. pt2, 1994, pages 381-387, -/-	4,8

INTERNATIONAL SEARCH REPORT

Int. Patent Application No.
PCT/NL 01/00301

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	(Citation of document, with indication, where appropriate, of the relevant passages)	Relevant to claim No.
X	M.C. RUIZ-ROLDAN ET AL: "Two xylanase genes of the vascular wilt pathogen <i>Fusarium oxysporum</i> are differentially expressed during infection of tomato plants" MOLECULAR AND GENERAL GENETICS, vol. 261, 1999, pages 530-536, XP002181148 SPRINGER VERLAG, BERLIN, DE ISSN: 0026-8925 page 532; figure 1B ----	4,7,8, 11-15
X	US 5 763 254 A (HJORT CARSTEN MAILAND ET AL) 9 June 1998 (1998-06-09) sequences ID no.3, 22 and 23 examples -----	4,8, 11-15
X	GB 2 289 218 A (MERCK & CO INC) 15 November 1995 (1995-11-15) page 44 sequence no.143 -----	13-15
X	DATABASE EMBL 'Online! CLAF310, accession number AF000310, 27 June 1997 (1997-06-27) M. DUFRESNE ET AL: "Collettrichum lindemuthianum glyceraldehyde-3-phosphate dehydrogenase" XP002181150 * 79.7% identity in 972 nt overlap with seq ID no.5 (1509-2474:1764-2721) * * 84% identity in 275 aa overlap with seq ID no.6 (3-277:4-277) * abstract -----	5,8
T	WO 00 20555 A (BOUSSON JEAN CHRISTOPHE; PARRICHE MARTINE (FR); PUNT PETER JAN (NL)) 13 April 2000 (2000-04-13) cited in the application page 43 -page 51; claims page 34, line 29 -page 40, line 10 -----	1-15

Form PCT/ISA/210 (Continuation of second sheet) (July 1999)

INTERNATIONAL SEARCH REPORT	International application No. PCT/NL 01/00301
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)	
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:	
1. ☐ Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:	
2. ☐ Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:	
3. ☐ Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).	
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)	
This International Searching Authority found multiple inventions in this International application, as follows:	
see additional sheet	
1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.	
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.	
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:	
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:	
Remark on Protest	
☐ The additional search fees were accompanied by the applicant's protest.	
☒ No protest accompanied the payment of additional search fees.	

International Application No. PCT/IL 01 00301

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: (1,6) all totally and (8-15) all partially
Chrysosporium glycosyl hydrolase family 7
2. Claims: (2-4, 7) all totally and (8-15) all partially
Chrysosporium glycosyl hydrolase family 10
3. Claims: 5 totally and (8-15) all partially
Chrysosporium glyceraldehyde phosphate dehydrogenase

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. No. International Application No.
PCT/NL 01/00301

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
WO 9815633	A	16-04-1998	US	5811381 A		22-09-1998
			AU	4741597 A		05-05-1998
			BR	9711876 A		24-10-2000
			CN	1237207 A		01-12-1999
			EP	0932688 A1		04-08-1999
			JP	2001506845 T		29-05-2001
			TR	9900771 T2		21-06-1999
			WO	9815633 A1		16-04-1998
			US	6015707 A		18-01-2000
WO 9713853	A	17-04-1997	AU	7294396 A		30-04-1997
			WO	9713853 A2		17-04-1997
			EP	0796328 A2		24-09-1997
			JP	10510720 T		20-10-1998
US 5763254	A	09-06-1998	US	5457046 A		10-10-1995
			AU	652153 B2		18-08-1994
			BR	9106434 A		04-05-1993
			DE	69125371 D1		30-04-1997
			DE	69125371 T2		14-08-1997
			EP	0531315 A1		17-03-1993
			FI	925080 A		09-11-1992
			GR	3023734 T3		30-09-1997
			KR	191564 B1		15-06-1999
			AT	150788 T		15-04-1997
			AU	7791091 A		27-11-1991
			CA	2082259 A1		10-11-1991
			WO	9117244 A1		14-11-1991
			DK	531315 T3		22-09-1997
			ES	2100229 T3		16-06-1997
			US	5686593 A		11-11-1997
GB 2289218	A	15-11-1995	NONE			
WO 0020555	A	13-04-2000	AU	6232699 A		26-04-2000
			BR	9914278 A		19-06-2001
			EP	1117808 A2		25-07-2001
			WO	0020555 A2		13-04-2000

Form PCT/ISA220 (patent family annex) (July 1999)

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

(72)発明者 バン・ゼイール, コルネリア・マリア・ヨハンナ

オランダ・エヌエルー 3 4 5 1 イーピー ブリユーテンデメールン・ヘンリデユナントラーン 2 4

Fターム(参考) 4B024 AA03 AA05 AA10 AA11 BA12 CA02 CA07 CA09 DA11 EA04
FA02 FA07 FA18 GA11 GA21 HA13 HA14
4B050 CC01 CC03 CC05 DD03 EE01 LL02 LL03 LL04 LL05
4B063 QA01 QQ43 QQ54 QR32 QR35 QR40 QR56 QS34 QX02
4B064 AF02 AF03 AF04 AF11 CA21 CB07 CD19 DA01 DA10 DA11
DA13
4B065 AA58X AA58Y AB01 AC14 BA02 BA10 CA31 CA32 CA41 CA43
CA55