

**(12) PATENT**  
**(19) AUSTRALIAN PATENT OFFICE**

**(11) Application No. AU 199892596 B2**  
**(10) Patent No. 739445**

(54) Title  
**Cellulose degrading enzymes of aspergillus**

(51)<sup>7</sup> International Patent Classification(s)  
**C12N 015/56 C12N 009/42**  
**A21D 008/04 C12S 003/04**  
**A23L 001/00 C12S 003/08**  
**A23L 002/00**

(21) Application No: **199892596**

(22) Application Date: **1998.07.31**

(87) WIPO No: **WO99/06574**

(30) Priority Data

(31) Number	(32) Date	(33) Country
<b>97202389</b>	<b>1997.07.31</b>	<b>EP</b>

(43) Publication Date : **1999.02.22**

(43) Publication Journal Date : **1999.04.22**

(44) Accepted Journal Date : **2001.10.11**

(71) Applicant(s)  
**DSM N.V.**

(72) Inventor(s)  
**Marcus Matheus Catharina Gielkens; Jacob Visser; Leendert Hendrik De Graaff**

(74) Agent/Attorney  
**PHILLIPS ORMONDE and FITZPATRICK,367 Collins Street,MELBOURNE VIC 3000**

(56) Related Art  
**TAKASHIMA ET AL. 1996 J OF BIOTECHNOLOGY 50:137-147**  
**US 4894338**  
**EP 0137280**

OPI DATE 22/02/99 APPLN. ID 92596/98  
AOJP DATE 22/04/99 PCT NUMBER PCT/EP98/05047



AU9892596

INTE

(51) International Patent Classification <sup>6</sup> : C12N 15/56, 9/42, C12S 3/04, 3/08, A21D 8/04, A23L 1/00, 2/00 // C12R 1/685	A1	(11) International Publication Number: <b>WO 99/06574</b>
		(43) International Publication Date: 11 February 1999 (11.02.99)

(21) International Application Number: PCT/EP98/05047

(22) International Filing Date: 31 July 1998 (31.07.98)

(30) Priority Data:  
97202389.9 31 July 1997 (31.07.97) EP

(71) Applicant (for all designated States except US):  
~~GIST BROCADES B.V. [NL/NL], Wateringsweg 1,  
P.O. Box 1, NL-2600 MA Delft (NL).~~

(72) Inventors; and

(75) Inventors/Applicants (for US only): GIELKENS, Marcus, Matheus, Catharina [NL/NL]; Pomona 176, NL-6708 CG Wageningen (NL). VISSER, Jacob [NL/NL]; Haarweg 8, NL-6709 PK Wageningen (NL). DE GRAAFF, Leendert, Hendrik [NL/NL]; Cornelis Koningstraat 8, NL 6862 CK Oosterbeek (NL).

(74) Agents: RINGELING, Patricia, Louise et al.; Gist-Brocades B.V., Patents and Trademarks Dept. (994-0760), Wateringsweg 1, P.O. Box 1, NL-2600 MA Delft (NL).

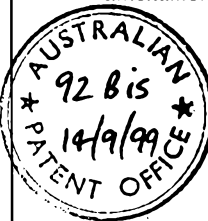
(71) DSM N.V.  
Het Overloon 1 6411 TE Heerlen  
Netherlands

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, 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), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.



(54) Title: CELLULOSE DEGRADING ENZYMES OF ASPERGILLUS

(57) Abstract

Two cellobiohydrolase polypeptides (CBHA and CBHB) derived from Aspergillus are described and can be used to degrade cellulose. Variants of these peptides are described as well as DNA encoding the peptides, vectors and host cells. The peptides can be used to produce or process food, animal feed, wood pulp, paper and textiles.

see folio 3

## CELLULOSE DEGRADING ENZYMES OF ASPERGILLUS

## FIELD OF THE INVENTION

5

The present invention relates to polypeptides which have cellulolytic activity and which can be derived from fungi; and to polynucleotides which encode the polypeptides and vectors which can express the polypeptides in a host cell. The polypeptides can be used to degrade cellulose. In particular they can be used in the production or processing of food, animal feed, wood pulp, paper and textiles.

10

## BACKGROUND OF THE INVENTION

15

Cellulose is a linear polysaccharide of glucose residues connected by beta-1,4 linkages. In nature, cellulose is usually associated with other compounds, such as hemicelluloses or lignin.

20

Three different classes of enzymatic activity have been shown to be required for the complete degradation of cellulose into glucose, viz. endoglucanases (EC 3.2.1.4), cellobiohydrolases (EC 3.2.1.91) and beta-glucosidases (EC 3.2.1.21).

25

Cellobiohydrolases attack cellulose either from the reducing or the non-reducing ends of the cellulose polymer and yield cellobiose (a glucose dimer) as the major product. Two types of cellobiohydrolase are known, cellobiohydrolase I (CBH I) and cellobiohydrolase II (CBH II). CBH I attacks cellulose from the reducing end of the cellulose polymer, and CBH II from the non-reducing end.

30

Cellobiohydrolase I and II from *Trichoderma reesei* have been described in EP-B-0137 280 and US 4,894,338, respectively. CBH I has also been identified in *Agaricus bisporus*, *Phanerochaete chrysosporium*, *Trichoderma viride* and *Humicola grisea*. Depending on both the natural origin and the previous treatment cellulose exists in many varieties differing in crystallinity, fibre composition, fibre length, fibre thickness. For fast and complete degradation of crystalline cellulose CBH and

endoglucanases work synergistically (Terri, TT., TIBTECH 15, May 1997, p. 160-167).

5 The above discussion of documents, acts, materials, devices, articles and the like is included in this specification solely for the purpose of providing a context for the present invention. It is not suggested or represented that any or all of these matters formed part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed in Australia before the priority date of each claim of this application.

10 Throughout the description and claims of the specification the word "comprise" and variations of the word, such as "comprising" and "comprises", is not intended to exclude other additives, components, integers or steps.

### SUMMARY OF THE INVENTION

15 The invention seeks to provide novel polypeptides with CBH activity which can be used to degrade cellulose.

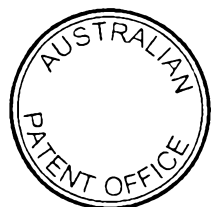
20 In one aspect the present invention provides a polypeptide which comprises the sequence of SEQ ID No: 8 (CBH A) or SEQ ID No: 10 (CBH B), or a sequence substantially homologous to either sequence, wherein the degree of homology is at least 80%.

The first aspect of the invention also provides a polypeptide which is a CBH from *Aspergillus*.

The first aspect additionally provides a polypeptide which:

- 25 (i) has CBH activity;
- (ii) has an activity of at least 50% of the maximum activity over the pH range from 3 to 5;
- (iii) has an optimum activity at a temperature which is greater than 50°C; and optionally
- (iv) does not have a cellulose binding domain or a linker peptide.

30 In a further aspect the present invention provides a polynucleotide capable of selectively hybridising (as hereinbefore defined) to SEQ ID No: 5 or SEQ ID No: 9 or a complement thereof, wherein said polynucleotide is at least 80% homologous to said SEQ ID No: 5 or SEQ ID No: 9 or their complements.



A third aspect of the invention provides a vector which comprises a nucleotide of the second aspect.

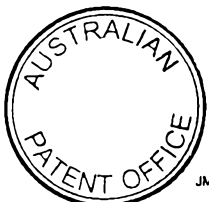
In an even further aspect the present invention provides a host cell which:

- 5 (i) comprises a polynucleotide or vector of the invention, which host cell does not naturally comprise such a polynucleotide or vector; and/or
- (ii) expresses a polypeptide of the invention, which host cell does not naturally express such a polypeptide; and/or
- 10 (iii) expresses a polypeptide of the invention at a higher level than naturally expressed by the host cell; and/or
- (iv) has been transformed or transfected by a polynucleotide or vector of the invention.

A fifth aspect of the invention provides a composition comprising a polypeptide of the invention and, for example, one or more enzyme(s).

15 A sixth aspect of the invention provides a use of the polypeptide or composition of the first or fifth aspects to degrade cellulose.

A seventh aspect provides a process of producing a host cell according to the third aspect of the invention.



## DETAILED DESCRIPTION OF THE INVENTION

5 One of the advantages of the present invention is that cellobiohydrolases may be produced which are free from endoglucanases. Since CBHs, unlike endoglucanases, do not significantly affect fibre strength they are more suitable than crude cellulase complexes for fibre modification in applications where strength is important, such as in textile and paper manufacture and finishing.

10 Another advantage is that the availability of cloned CBHs and endoglucanases (WO 97/13862) makes it possible to design tailor-made combinations of CBHs and endoglucanases for specific applications.

Yet another advantage of the invention is that the polypeptides of the invention have optimum activity at low pH, unlike known cellobiohydrolases. 15 Therefore, the polypeptides of the invention are well suited for use in industrial processes which are performed at low pH.

### Polypeptides

20 The polypeptide of the invention typically has CBH activity. Generally such an activity would be the activity defined by E.C. 3.2.1.91. The term "CBH activity" includes the ability to produce cellobiose from cellulose (as a substrate).

The CBH activity may be CBH I-like activity and/or CBH II-like activity. The cellulose which is acted on by the polypeptide of the invention may be crystalline 25 cellulose. The level of CBH activity of the polypeptide of the invention may be the same as, substantially the same as, or higher or lower than the polypeptide shown in SEQ ID No: 8 or 10.

The polypeptide of the invention may or may not comprise a cellulose binding domain. Such a domain is often linked to the catalytic domain via a linker peptide in 30 prior art CBH enzymes. Therefore if the polypeptide of the invention does not comprise a cellulose binding domain then generally it will not comprise a linker peptide either.

The polypeptide of the invention may comprise the sequence of SEQ ID No: 8 or 10 or a substantially homologous sequence. The polypeptide may be 65%

homologous to SEQ ID No: 8 or 10, preferably 80% or 90%, and more preferably at least 95% homologous thereto over the length of SEQ ID No: 8 or 10 .

The polypeptide of the invention may be derived from a fungus, such as a filamentous fungus. Preferably the (filamentous) fungus is of the species *Aspergillus oryzae*, *Aspergillus sojae*, *Aspergillus nidulans*, species from the *Aspergillus niger* Group (as defined by Raper and Fennell, *The Genus Aspergillus*, The Williams & Wilkins Company, Baltimore, pp 293-344, 1965), specifically including but not limited to *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus tubigenensis*, *Aspergillus aculeatus*, *Aspergillus foetidus*, *Aspergillus japonicus* and *Aspergillus ficuum*.

In the context of the invention the term "derived from" includes polypeptides which naturally occur in or are produced by these fungae and therefore can be obtained from these fungae, or fragments of such polypeptides. The term also includes polypeptides which are substantially homologous (such as at least 65% homologous to SEQ ID No: 8 or 10) to these naturally occurring polypeptides of the invention. Thus the invention includes polypeptides which are not naturally occurring.

The polypeptide of the invention may co-purify with an  $\alpha$ - arabinofuranosidase (see for e.g. Example 1). The polypeptide of the invention may thus co-elute with an  $\alpha$ -arabinofuranosidase. The  $\alpha$ -arabinofuranosidase may be detected using the chromogenic substrate 4- methylumbelliferyl- $\alpha$ -L-arabinofuranoside. The  $\alpha$ -arabinofuranosidase may be a fungal  $\alpha$ -arabinofuranosidase, such as one from an *Aspergillus* species, e.g. *Aspergillus niger*. The  $\alpha$ -arabinofuranosidase may be one from any of the genera or species of fungae mentioned above from which the polypeptide of the invention may be derived.

A polypeptide of the invention may comprise:

(a) a polypeptide comprising the sequence of SEQ ID No: 8 or 10;

or

(b) a polypeptide from *Aspergillus* which is a CBH or has CBH activity; or

(c) a homologue or an allelic variant of any of the polypeptides of (a) or (b) from a fungus, an *Aspergillus* species, from *Aspergillus niger* or from any of the genera or species of fungus mentioned above from which the polypeptide of the invention may be derived; or

(d) a homologue which is at least 65% homologous to (a), (b) or (c).

Methods of measuring protein homology are well known in the art and it will be understood by those of skill in the art that in the present context, homology is calculated on the basis of amino acid identity (sometimes referred to as "hard homology").

5 The polypeptide of the invention or fragments thereof may be used to identify cellobiohydrolases in other organisms. For example, they may be used for the production of antibodies. The sequence of the polypeptide of SEQ ID No: 8 or 10 and of species homologues and allelic variants can be modified to provide polypeptides of the invention.

10 The modified polypeptide may retain CBH activity.

### Polynucleotides

15 A polynucleotide of the invention may encode or comprise a sequence that encodes a polypeptide of the invention.

A polynucleotide of the invention may be capable of hybridising selectively with the coding sequence of SEQ ID No: 5 or 9 or to the sequences complementary to those coding sequences. Selective hybridisation is typically achieved using conditions of medium to high stringency (for example high stringency conditions include 0.2 x SSC at 60°C, and low stringency conditions include 2 x SSC at 60°C).

20 A nucleotide sequence capable of selectively hybridizing to the DNA coding sequence of SEQ ID No: 5 or 9 or to the sequences complementary to those coding sequences will generally be at least 65%, preferably at least 80% homologous to the coding sequence of SEQ ID No: 5 or 9 or their complements.

25 Polynucleotides of the invention, which are typically provided in isolated and/or purified form, may comprise DNA or RNA. They may also be polynucleotides which include within them synthetic or modified nucleotides. For the purposes of the present invention, it is to be understood that the polynucleotides described herein may be modified by any method available in the art.

30 Polynucleotides such as a DNA polynucleotide and primers according to the invention may be produced recombinantly, synthetically, or by any means available to those of skill in the art.

For the recombinant production of the polypeptide of the invention a DNA sequence of the invention is used for gene amplification and/or exchange of



expression signals, such as promoters, secretion signal sequences, in order to allow economic production of the polypeptide of the invention in a suitable homologous or heterologous host cell. A homologous host cell is herein defined as a host cell which is of the same species or which is a variant within the same species as the species from which the DNA sequence is derived.

Suitable host cells are preferably microorganisms like bacteria, or more preferably fungi such as yeasts or filamentous fungi. A preferred yeast host cell for the expression of a DNA sequence encoding the polypeptide of the invention is selected from the group consisting of the genera *Saccharomyces*, *Kluyveromyces*, *Hansenula*, *Pichia*, *Yarrowia*, and *Schizosaccharomyces*. More preferably a yeast host cell is selected from the group consisting of the species *Saccharomyces cerevisiae*, *Kluyveromyces lactis* (also known as *Kluyveromyces marxianus* var. *lactis*), *Hansenula polymorpha*, *Pichia pastoris*, *Yarrowia lipolytica*, and *Schizosaccharomyces pombe*.

Most preferred for the expression of a DNA sequence encoding the polypeptide of the invention are, however, filamentous fungal host cells. Preferred filamentous fungal host cells are selected from the group consisting of the genera *Aspergillus*, *Trichoderma*, *Fusarium*, *Penicillium*, *Acremonium*, *Neurospora*, *Thermoascus*, *Myceliophthora*, *Sporotrichum*, *Thielavia*, and *Talaromyces*. More preferably a filamentous fungal host cell is selected from the group consisting of the species *Aspergillus ozyae*, *Aspergillus sojae*, *Aspergillus nidulans*, species from the *Aspergillus niger* Group as defined by Raper and Fennell (1965, In: The Genus *Aspergillus*, The Williams & Wilkins Company, Baltimore, pp 293-344), specifically including but not limited to *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus tubigenensis*, *Aspergillus aculeatus*, *Aspergillus foetidus*, *Aspergillus japonicus* and *Aspergillus ficuum*, and further consisting of the species *Trichoderma reesei*, *Fusarium graminearum*, *Penicillium chrysogenum*, *Acremonium alabamense*, *Neurospora crassa*, *Myceliophthora thermophilum*, *Sporotrichum cellulophilum*, and *Thielavia terrestris*.

The expression construct can be used for transformation of the host as part of a vector carrying a selectable marker, or the expression construct is co-transformed as a separate molecule together with the vector carrying a selectable marker. Suitable selectable markers which can be used for selection of the transformed host cells are well known to the skilled person<sup>8,10</sup>. Preferred markers include but are not

limited to e.g. versatile marker genes that can be used for transformation of most filamentous fungi and yeasts such as acetamidase genes or cDNAs (the *amdS* genes or cDNAs from *A.nidulans*, *A.oryzae*, or *A.niger*), or genes providing resistance to antibiotics like G418 or hygromycin.

5 Alternatively, more specific selection markers can be used such as auxotrophic markers which require corresponding mutant host strains: e.g. *URA3* (from *S.cerevisiae* or analogous genes from other yeasts), *pyrG* (from *A.nidulans* or *A.niger*) or *argB* (from *A.nidulans* or *A.niger*). In a more preferred embodiment, the selection marker is deleted from the transformed host cell after introduction of the  
10 expression construct in accordance with the methods described in EP-A-O 635 574, so as to obtain transformed host cells capable of producing the polypeptide which are free of selection marker genes.

For most filamentous fungi and yeast, the expression construct is preferably integrated in the genome of the host cell in order to obtain stable transformants.  
15 However, for certain yeasts also suitable episomal vector systems are available into which the expression construct can be incorporated for stable and high level expression, examples thereof include vectors derived from the  $2\mu$  and pKD1 plasmids of *Saccharomyces* and *Kluyveromyces*, respectively. In case the expression constructs are integrated in the host cells genome, the constructs are either  
20 integrated at random loci in the genome, or at predetermined target loci using homologous recombination, in which case the target loci preferably comprise a highly expressed gene. A highly expressed gene is herein defined as a gene whose mRNA can make up at least 0.5% (w/w) of the total cellular mRNA, e.g. under induced conditions, or alternatively, a gene whose gene product can make up at least 1%  
25 (w/w) of the total cellular protein, or, in case of a secreted gene product, can be secreted to a level of at least 0.1 g/l. A number of examples of suitable highly expressed genes is provided herein below.

An expression construct for a given host cell will usually contain the following elements operably linked to each other in a consecutive order from the 5'-end to  
30 3'-end relative to the coding strand of the sequence encoding the polypeptide of the invention: (1) a promoter sequence capable of directing transcription of the DNA sequence encoding the polypeptide in the given host cell, (2) optionally, a signal sequence capable of directing secretion of the polypeptide from the given host cell into the culture medium, (3) the DNA sequence encoding a mature and preferably

active form of the polypeptide, and preferably also (4) a transcription termination region (terminator) capable of terminating transcription downstream of the DNA sequence encoding the polypeptide.

A variety of promoters capable of directing transcription in the host cells of the invention is available to the skilled person<sup>8,10</sup>. Preferably the promoter sequence is derived from a highly expressed gene as previously defined. Examples of preferred highly expressed genes from which promoters are preferably derived and/or which are comprised in preferred predetermined target loci for integration of expression constructs, include but are not limited to genes encoding glycolytic enzymes such as triose-phosphate isomerases (TPI), glyceraldehyde-phosphate dehydrogenases (GAPDH), phosphoglycerate kinases (PGK), pyruvate kinases (PYK), alcohol dehydrogenases (ADH), as well as genes encoding amylases, glucoamylases, xylanases, cellobiohydrolases,  $\beta$ -galactosidases, alcohol (methanol) oxidases, elongation factors and ribosomal proteins. Specific examples of suitable highly expressed genes include e.g. the *LAC4* gene from *Kluyveromyces* sp., the methanol oxidase genes (*AOX* and *MOX*) from *Hansenula* and *Pichia*, respectively, the glucoamylase (*glaA*) genes from *A.niger* and *A.awamori*, the *A.oryzae* TAKA-amylase gene, the *A.nidulans gpdA* gene and the *T.reesei* cellobiohydrolase genes.

Preferably the polypeptide is produced as a secreted protein in which case the polynucleotide sequence encoding a mature form of the polypeptide in the expression construct is operably linked to a polynucleotide sequence encoding a signal sequence. Preferably the signal sequence is native (homologous) to the DNA sequence encoding the polypeptide. Alternatively the signal sequence is foreign (heterologous) to the DNA sequence encoding the polypeptide, in which case the signal sequence is preferably endogenous to the host cell in which the DNA sequence is expressed. Examples of suitable signal sequences for yeast host cells are the signal sequences derived from yeast  $\alpha$ -factor genes. Similarly, a suitable signal sequence for filamentous fungal host cells is e.g. a signal sequence derived from a filamentous fungal (gluco)amylase gene, e.g. the *A.niger glaA* gene.

Downstream of the DNA sequence encoding the polypeptide, the expression construct preferably contains a 3' untranslated region containing one or more transcription termination sites, also referred to as a terminator. The origin of the terminator is less critical. The terminator can e.g. be native to the DNA sequence encoding the polypeptide. However, preferably a yeast terminator is used in yeast

host cells and a filamentous fungal terminator is used in filamentous fungal host cells. More preferably, the terminator is endogenous to the host cell in which the DNA sequence encoding the polypeptide is expressed.

The recombinant host cells according to the invention may be cultured using procedures known in the art. For each combination of a promoter and a host cell, culture condition are available which are conducive to the expression the DNA sequence encoding the polypeptide. After reaching the desired cell density or titre of the polypeptide the culture may be stopped and the polypeptide may be recovered using known procedures.

In addition, homologues and allelic variants of SEQ ID No: 5 or 9 may be obtained and such homologues and fragments thereof in general will be capable of selectively hybridizing to the coding sequence of SEQ ID No: 5 or 9 complements.

Selective hybridisation is typically achieved under conditions of medium to high stringency (for example high stringency conditions include 0.2 SSC at 60°C, and low stringency conditions include 2 x SSC at 60°C).

#### **Compositions containing the polypeptides**

The invention also provides a composition comprising a polypeptide of the invention and, optionally, one or more enzymes. It also provides a composition comprising a recombinant host cell of the invention.

The compositions of the invention may be in a form suitable for packaging and/or storage. In a composition which comprises the polypeptide of the invention the composition is generally of a form in which the CBH activity of the polypeptide is substantially retained. The polypeptide in the composition may be attached to or mixed with a carrier, for example the polypeptide may be immobilized, such as on a solid carrier.

In a composition which comprises host cells of the invention the composition will generally be of a form which allows some or all of the host cells to remain viable. The composition may additionally comprise nutrients for the host cell, which are provided to the host cell when it is cultured.

The composition may be in a form suitable for the process in which it will be used. The composition may be formulated in any convenient way, including as a paste, liquid, emulsion, powder, flakes, granulate, pellet or other extrudate. The

composition may be formulated for use in a process to produce human food, animal feed, wood pulp, paper or a textile. It may be stabilised in accordance with methods known in the art.

The composition may comprise additional substances which aid the cellulose degrading activity of the composition. Thus the composition may additionally  
5 comprise other enzymes. These other enzymes may be recombinant enzymes, or may have been obtained from an organism in which they occur naturally. The enzymes may have been substantially purified before addition to the composition of the invention, or they may be left substantially unpurified before being added to the  
10 composition of the invention.

Non-limiting examples of such other enzymes are endoglucanases (EC 3.2.1.4), other cellobiohydrolases (EC 3.2.1.91)  $\beta$ -glucosidases (EC 3.2.1.21), xylanases, pectinases, mannanases, phytases, alpha-amylase, proteases or various plant cell wall degrading enzymes.

The composition may comprise organisms (e.g. bacteria, fungae or yeast) which produce the above mentioned enzymes.

The composition may additionally comprise (particularly when being formulated for use in animal feed) one or more ionophores, oxidising agents, surfactants, rumen protected amino acids, enzyme enhancers or enzymes which may  
20 be produced naturally in the gastro intestinal tract of the animals to be fed.

#### **Uses of the polypeptide, polynucleotide, host cell and composition**

The invention provides the use of a polypeptide, host cell or composition of the invention to degrade cellulose. Thus the invention provides a method of  
25 degrading cellulose comprising contacting a polypeptide, host cell or composition of the invention with the cellulose (to be degraded).

In the method all or part of the cellulose may be degraded, for example from 0 to 20%, 20 to 40 %, 40 to 50%, 50 to 70% or 70 to 100% by weight may be degraded.

The cellulose which is degraded in the method of the invention may be derived  
30 from a plant or a microorganism, or may have been synthetically produced by man. The cellulose may be in crystalline form. The cellulose may be partially or wholly derivatised, for example lignified.

The method of degrading cellulose of the invention may be part of a method for producing or processing food or beverages, animal feed, pulp, paper or textiles.

Moreover, the CBHs of the present invention, have broader pH optima than CBHs of the prior art, and are particularly well suited for use in industrial processes which are performed at low pH, where the CBHs of the prior art are less or not active. Since the CBHs are also more thermostable as compared with known CBHs they can be used in a much wider range of application conditions.

In accordance with the present invention, it has been found that the CBHs produced via the present invention may be used in the baking of breads. The incorporation of small amounts of CBH to the flour imparts favourable characteristics to the dough and thus to the bread itself such as an increased loaf volume and better textural characteristics such as break and shred quality and crumb quality.

CBHs either alone or in combination with other pure or crude endoglucanase containing enzyme mixtures may also be added to animal feed compositions which are rich in cellulose. When added to feeds (including silage) for ruminants or monogastric animals (eg. poultry or swine) which feeds contain cereals such as barley, wheat, maize, rye or oats or cereal by-products such as wheat bran or maize bran, or other plant materials such as soy beans and other legumes, the enzyme(s) significantly improve the break-down of plant cell walls which leads to better utilization of the plant nutrients by the animal. As a consequence, growth rate and/or feed conversion are improved.

The CBHs of the invention may also be used for the liquefaction of plant cell wall material, e.g. in the beverage industry.

#### BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows how the activity of CBH A varies with temperature in comparison to Econase.

Figure 2 shows how the activity of CBH A varies with pH in comparison to Econase.

Figure 3 shows how the activity of CBH B varies with temperature in comparison to Econase.

Figure 4 shows how the activity of CBH B varies with pH in comparison to Econase.

The invention will now be described, by way of example, by reference to the following Examples, which are not to be construed as being limiting.

5

## EXAMPLES

10

### **Strains**

*E. coli* LE 392<sup>4</sup>:

e14:(*mcrA*) *hsdR514*, *supE44*, *supF58*, *lacY1*, or D(*lac1ZY*)6, *galK2*, *galT22*, *metB1*, *trpR55*

15

### **Example 1 Identification of an *A. niger* cellobiohydrolase**

*A. niger* N402 was grown in *Aspergillus* minimal medium (MM) (contains per litre: 6.0 g NaNO<sub>3</sub>, 1.5 g KH<sub>2</sub>PO<sub>4</sub>, 0.5 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 0.5 g KCl, pH 6.0 and 1 ml Vishniac<sup>7</sup> solution (contains per litre 10 g EDTA, 4.4 g ZnSO<sub>4</sub>·7H<sub>2</sub>O, 1.0 g MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.32 g CoCl<sub>2</sub>·6H<sub>2</sub>O, 0.32 g CuSO<sub>4</sub>·5H<sub>2</sub>O, 0.22 g (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>·4H<sub>2</sub>O, 20 1.47 g CaCl<sub>2</sub>·2H<sub>2</sub>O, 1.0 g FeSO<sub>4</sub>·7H<sub>2</sub>O, pH 4.0) supplemented with 1.5 % wheat arabinoxylan. This medium was inoculated with 1x10<sup>6</sup> spores per ml and mycelium was grown for 96 hours at 30°C and 250 rpm in an orbital New Brunswick shaker. The culture filtrate was collected after filtration of the mycelium on Myra cloth (nylon gauze) using a Büchner funnel and mild suction. The pH of the culture filtrate was adjusted to pH 6.0 with 0.1 M NaOH after which the culture filtrate was diluted by 25 the addition of 2 volumes of Millipore water.

DEAE-Sephadex A-50 was equilibrated in 50 mM sodium acetate buffer pH 5.0 and was added to the culture filtrate. After 30-60 minutes of stirring at 4°C, the DEAE-Sephadex together with the culture filtrate were passed through a funnel with 30 a glass filter holder and the DEAE-Sephadex A-50 was transferred to a column. This column was first eluted with 50 mM sodium acetate buffer pH 5.0, then with 50 mM sodium acetate buffer pH 5.0 + 0.5 M NaCl. Fractions containing alpha-arabinofuranosidase activity, detected using the chromogenic substrate 4-methylumbelliferyl-alpha-L-arabinofuranoside (detects alpha-arabinofuranosidases)

(Sigma M-9519), were pooled and desalted by dialysis against Millipore water and subsequently dialysed against 20 mM piperazine-HCl buffer pH 5.0. After dialysis the sample was loaded on a DEAE-Sepharose Fast Flow column, this column was first eluted with 3 volumes 20 mM piperazine-HCl buffer pH 5.0 and then with a linear gradient of 0.5 M NaCl in 20 mM piperazine-HCl buffer pH 5.0. Detection of the eluted protein was performed by continuous measurement of the UV absorption at 280 nm. Fractions of 10 ml were collected which were assayed for activity of alpha-arabinofuranosidase on *para*-nitrophenyl-alpha-L-arabinofuranoside (PNP-A)(Sigma N-3641) the a-arabinofuranosidase activity was found in fractions 11-27, 41-47 and 52-61. Fractions 41-47 were pooled and subsequently dialysed against 100 mM piperazine-HCl buffer pH 5.0 and the whole sample (200 ml) was loaded on a DEAE Sepharose FF column (Pharmacia) to concentrate the protein to a single peak using a 1 M NaCl gradient in 200 mM piperazine-HCl pH 5.0 buffer. Protein containing fractions were pooled (20 ml) and were first dialysed against Millipore water, then against 20 mM piperazine-HCl pH 5.0 and finally concentrated using a Speed Vac concentrator to 6 ml. This preparation was loaded on a Sephacryl S-300 column, which was eluted with 20 mM piperazine-HCl pH 5.0, 0.1 M NaCl. Fractions having activity on PNP-A were pooled, dialysed against 10 mM piperazine-HCl pH 5.0 and concentrated in a Speed Vac concentrator to 2 ml. This sample was then loaded on a Superdex 75 (Hiload column 16/60) (Pharmacia) column, which was equilibrated and eluted with 20 mM piperazine-HCl pH 5.0, 0.1 M NaCl. Fractions having a-arabinofuranosidase activity were dialysed against 10 mM Na-acetate buffer pH 3.5. The final purification was done on a Mono S cation exchange column (HR 5/5, Pharmacia). The column was equilibrated with 10 mM Na-acetate buffer pH 3.5 buffer in which the sample was loaded. Protein was eluted using 27 ml of a linear gradient 1 M NaCl in 10 mM Na-acetate buffer pH 3.5. The enzyme eluted at about 100 mM NaCl.

The purified enzyme was sent to EUROSEQUENCE (Groningen, The Netherlands). There it was applied to an SDS-PAGE gel and internal sequence analysis of the protein was performed essentially as described in reference 11.

Amino acid sequences were performed using an automated sequencer (Model 477A, Applied Biosystems) coupled to an HPLC (Model 120A ABI) for analysis of the phenylthiodantoin (PTH) amino acids. The following internal sequences were determined:



Leu-Tyr-Leu-Met-Ser-Asp-Asp-Ser-Asn-Tyr-Glu-Leu-Phe-Lys  
(SEQ ID No: 1)

Leu-Gly-Asn-Thr-Asp-Phe-Tyr-Gly-Pro-Gly-Leu-Thr-Val-Asp-Thr-Asn-Ser-Pro-Ph  
5 e -Thr-Val-Val-Thr-Gln  
(SEQ ID No: 2)

Surprisingly, the partial amino acid sequence of one of the internal fragment  
of the isolated enzyme (SEQ ID No: 1) showed a very high identity with *Agaricus*  
10 *bisporus* CBH I, i.e. 11 out of 13 residues were identical. This was the first indication  
of cellobiohydrolase genes in *Aspergillus*.

## Example 2 Isolation of cDNA clones coding for *A. niger* cellobiohydrolase

### 15 Example 2.1 Generation of probe

The *Agaricus bisporus* CBH I enzyme is encoded by the *cel2* gene (EMBL Acc.  
No. Z50094). The two following oligonucleotides were designed to PCR the coding  
region of *cel2*:

20 5'-GTC GGT ACC AAC ATG GCC G-3' (19-mer)  
(SEQ ID No: 3)

5'-ACT CAG AAA CAT TGG CTA TAG-3' (21-mer)  
(SEQ ID No: 4)

25

These oligo nucleotides, as depicted in SEQ ID NOS. 3 and 4, were used in  
PCR using plasmid DNA containing *Agaricus bisporus cel2* sequences as a template<sup>5</sup>.

For a PCR 2 ng of plasmid DNA was combined with 5  $\mu$ l 10x reaction buffer  
(200 mM Tris-HCl, pH 8.4; 500 mM KCl); 3  $\mu$ l 50 mM MgCl<sub>2</sub>; 4.0  $\mu$ l 1.25 mM of  
30 each of the four deoxynucleotide triphosphates and 50 pmol of the oligonucleotides  
in a final volume of 50  $\mu$ l. The reaction mixture was mixed and 0.5  $\mu$ l TAQ  
polymerase (5 U/ $\mu$ l) (Life Technologies) was added. The DNA was heat denatured by  
incubation for 5 minutes at 95°C followed by 25 cycles of 1 minute at 95°C, 1  
minute at 52°C and 1 minute at 72°C. After these 25 cycles the mixture was

incubated for 5 minutes at 72°C. Analysis of the reaction products revealed one discrete product of 1.5 kb, which was the expected size.

**Example 2.2 <sup>32</sup>P-labelling of 1.5 kb PCR fragment**

5 The 1.5 kb fragment obtained by PCR was isolated and labelled as described in EP-A-0 463 706, Examples 2.2 and 7.1.

**Example 2.3 Screening of the *A. niger* N400 cDNA library for cellobiohydrolase cDNA clones**

10 The construction of the *A. niger* cDNA library is described in WO 96/06935 (Example 3).

To screen the *A. niger* N400 cDNA library for *cbh* cDNA clones, 5x10<sup>3</sup> pfu per plate were plated in NZYCM topagarose containing 0.7% agarose on 85-mm-diameter NZYCM (1.5% agar) plates as described in reference 3, using *E. coli* 15 LE392 as plating bacteria.

After overnight incubation of the plates at 37°C two replicas of each were made on HybondN filters (Amersham) as described in reference 3.

The filters were prehybridized at 65°C for 2 hours in prehybridization buffer containing; 6xSSC, 0.5 % SDS, 5xDenhardt's solution, 0.01 M EDTA and 100 µg/ml 20 heat denatured herring sperm DNA (Boehringer Mannheim). After two hours prehybridization, the prehybridization buffer was replaced by hybridization buffer which was identical to the prehybridization buffer, but contained the <sup>32</sup>P labelled 1.5 kb PCR fragment containing *Agaricus bisporus cel2* sequences and prepared as described in Example 2.2. The filters were hybridized for 18 hours at a temperature 25 of 60°C.

After hybridization the filters were first washed at 60°C for 30 minutes in 4xSSC/0.5% SDS followed by two washing steps at 60°C for 30 minutes in 2xSSC/0.5% SDS. The air dried filters were taped on a sheet of Whatman 3MM paper, keying marks were made with radioactive ink and the Whatman paper and 30 filters were covered with Saran Wrap. Hybridizing plaques were identified by exposure of Kodak XAR X-ray film for 72 hours at -70°C using an intensifying screen.

About fifty positive hybridizing plaques, appearing in duplicate were found per plate. Twelve positive plaques were picked from the plate using a Pasteur pipette and

the phages were eluted from the agar plug in 1 ml of SM buffer containing 20  $\mu$ l chloroform, as described in reference 3. The phages obtained were purified by repeating the procedure described above using filter replicas from plates containing 50-100 plaques of the isolated phages.

5 After purification the phages were propagated by plating  $5 \times 10^3$  phages on NZYCM medium. After overnight incubation at 37°C confluent plates were obtained, from which the phages were eluted by adding 5 ml SM buffer and storing the plate for 2 hours at 4°C with intermittent shaking. After collection of the supernatant using a pipette, the bacteria were removed from the solution by centrifugation at  
10 4,000  $\times$  g for 10 minutes at 4°C. To the supernatant 0.3% chloroform was added and the number of pfu was determined. These phage stocks contain approximately  $10^7$  pfu/ml.

#### Example 2.4 Restriction analysis of *cbh* cDNA clones

15 The recombinant Uni-ZAP XR clones containing *cbh* cDNA were converted to Bluescript phagemids by superinfection with the filamentous helper phage ExAssist™, which is included in the ZAP™-cDNA synthesis kit from Stratagene, according to the manufacturer's instructions.

The phagemid DNA was subsequently isolated as described in reference 6.

20 The isolated DNA of eight *cbh* cDNA clones were subjected to restriction analysis using the following restriction enzymes: *EcoRI* and *XhoI*. The DNA was digested for 2 hours at 37°C in a reaction mixture composed of the following solutions; 2  $\mu$ l ( $\gg$ 1 mg) DNA solution; 2  $\mu$ l of the appropriate 10  $\times$  React buffer (Life Technologies); 10 U of each Restriction enzyme (Life Technologies) and sterile  
25 distilled water to give a final volume of 20  $\mu$ l. After addition of 4  $\mu$ l DNA loading buffer the samples were loaded on a 0.7% TAE-agarose gel. The DNA fragments were separated by electrophoresis at 80 V for 1.5 hours. The restriction analysis of the isolated cDNA clones revealed two types of plasmid; one containing inserts of 1.2 kb, the other inserts of 1.8 kb.

#### Example 2.5 Sequence analysis of *A. niger cbh* cDNA clone CBHA-C9

30 The sequence of the *A. niger cbh* cDNA clones was determined by subcloning fragments from one of the positive clones of Example 2.4 with 1.8 kb insert, called CBHA-C9, in pBluescript SK-. Sequencing reactions were conducted using the

ThermoSequenase fluorescent labelled primer cycle sequencing kit (Amersham) with universal sequencing primers. The sequencing reactions were analysed on an ALFexpress sequencer (Pharmacia). Sequence analysis was performed using the Winstar programme (LaserGene) and yielded the sequence as shown in SEQ ID No: 5. The amino acid sequence of the coding region, corresponding to nucleotides 49 to 1404, is shown in SEQ ID No: 8. The coding region encodes the mature enzyme (amino acid 18 to 452) plus its pre-sequence (amino acid 1 to 17). The protein does not contain a linker peptide, nor a cellulose binding domain.

#### Example 2.6 Screening of the *A. niger* genomic library for the *cbhA* gene

For the screening of the *A. niger* genomic library, constructed as described in reference 2, for the *cbhA* gene  $3 \times 10^3$  pfu per plate were plated in NZYCM top-agarose containing 0.7% agarose on five 85-mm-diameter NZYCM (1.5% agar) plates as described in reference 3, using *E. coli* LE392 as plating bacteria.

After overnight incubation of the plates at 37°C two replicas of each were made on HybondN filters (Amersham) as described in reference 3.

The filters were prehybridized at 68°C for two hours in prehybridization buffer containing; 6xSSC, 0.5 % SDS, 5xDenhardt's solution, 0.01 M EDTA and 100 µg/ml heat denatured herring sperm DNA (Boehringer Mannheim). After two hours prehybridization, the prehybridization buffer was replaced by hybridization buffer which was identical to the prehybridization buffer, but contained the <sup>32</sup>P labelled the 1.8 kb insert from cDNA clone CBHA-C9 and prepared as described in Example 2.2. The filters were hybridized for 18 hours at an temperature of 68°C.

After hybridization the filters were first washed at 68°C for 30 minutes in 4xSSC/0.1% SDS followed by a second wash at 68°C during 30 minutes in 2xSSC/0.1% SDS. The filters were then washed twice at 68°C for 30 minutes with 0.1xSSC/0.1% SDS. The air dried filters were taped on a sheet of Whatman 3MM paper, keying marks were made with radioactive ink and the Whatman paper and filters were covered with Saran Wrap. Hybridizing plaques were identified by exposure of Kodak XAR X-ray film for 72 h at -70°C using an intensifying screen.

Two to three positive hybridizing plaques, appearing in duplicate were found per plate. Five positive plaques were picked from the plate using a Pasteur pipette and the phages were eluted from the agar plug in 1 ml of SM buffer containing 20 µl chloroform, as described in reference 3. The phages obtained were purified by

repeating the procedure described above using filter replicas from plates containing 50-100 plaques of the isolated phages.

**Example 3 Cloning of further *A. niger cbh* cDNA**

5

**Example 3.1 Generation of probe**

The partial amino acid sequences of the internal fragments of the cellobiohydrolase which resulted from the arabinofuranosidase purification (SEQ ID No: 1 and 2), were used to design oligonucleotide mixtures. The following mixtures were derived:

10

5'-GAY GAY AGY AAY TAY GAR TTR TTY AA-3' (26-mer)  
(SEQ ID No: 6)

15

5'-GTR AAN GGR CTR TTN GTR TC-3' (20-mer)  
(SEQ ID No: 7)

20

in which R stands for an A or G; Y for a C or T and N for any base. The oligonucleotide depicted in SEQ ID No: 6 was derived from the amino acid sequence depicted in SEQ ID No: 1, as described in Example 1 from amino acid 6 (D) to amino acid 14 (K). The oligonucleotide depicted in SEQ ID No: 7 was derived from the amino acid sequence depicted in SEQ ID No: 2, as described in Example 1 from amino acid 20 (T) to amino acid 14 (D).

25

These oligo nucleotides were used in PCR using excised plasmid DNA from the *A. niger* N400 cDNA library as a template<sup>5</sup>.

30

The TaqBead™ PCR kit (Promega) was used for PCR, i.e. 5 ng of plasmid DNA was combined with 5  $\mu$ l 10\* reaction buffer, 6  $\mu$ l 25 mM MgCl<sub>2</sub>, 8  $\mu$ l 1.25 mM of each of the four deoxynucleotide triphosphates and 50 pmol of the oligonucleotides depicted in SEQ ID NO's: 6 and 7 in a final volume of 50  $\mu$ l. The reaction mixture was mixed and 0.5  $\mu$ l TAQ polymerase (5 U/ $\mu$ l) (Life Technologies) was added. The DNA was heat denatured by incubation for 5 minutes at 95°C followed by 24 cycles of 1 minute at 94°C, 1.5 minute at 48°C and 1.5 minute at 72°C, in which the annealing temperature was lowered every cycle by 0.3°C. This was followed by 10 cycles of 1 minute at 94°C, 1.5 minute 40°C and 1.5 minute

72°C. After these 10 cycles the mixture is incubated for 5 minutes at 72°C. Analysis of the reaction products revealed one discrete product of approximately 450 bp.

5           **Example 3.2 <sup>32</sup>P-labelling of 450 bp PCR fragment**

The 450 bp fragment obtained by PCR was isolated and labelled as described in EP-A-0 463 706, Examples 2.2 and 7.1.

10           **Example 3.3     Screening of the *A. niger* N400 cDNA library for further *cbh* cDNA clones**

The construction of the *A. niger* cDNA library is described in WO 96/06935 (Example 3).

To screen the *A. niger* N400 cDNA library for further *cbh* cDNA clones, 5 \* 10<sup>3</sup> pfu per plate were plated in NZYCM topagarose containing 0.7% agarose on 15 85-mm-diameter NZYCM (1.5% agar) plates as described reference 3, using *E. coli* LE392 as plating bacteria.

Hybridization was performed as described in Example 2.6, using the 450 bp PCR fragment prepared as described in Example 3.2

About ten to twenty positive hybridizing plaques, appearing in duplicate on 20 the replica filters were found per plate. Eight positive plaques were picked from the plate using a Pasteur pipette and the phages were eluted from the agar plug in 1 ml of SM buffer containing 20 µl chloroform, as described in reference 3. The phages obtained were purified by repeating the procedure described above using filter replicas from plates containing 50-100 plaques of the isolated phages.

25           After purification the phages were propagated by plating 5x10<sup>3</sup> phages on NZYCM medium. After overnight incubation at 37°C confluent plates were obtained, from which the phages were eluted by adding 5 ml SM buffer and storing the plate for 2 hours at 4°C with intermittent shaking. After collection of the supernatant using a pipette, the bacteria were removed from the solution by centrifugation at 30 4,000 x g for 10 minutes at 4°C. To the supernatant 0.3% chloroform was added and the number of pfu is determined. These phage stocks contain approximately 10<sup>7</sup> pfu/ml.

**Example 3.4 Restriction analysis of further *cbh* cDNA clones**

The recombinant Uni-ZAP XR clones containing *cbh* cDNA were converted to Bluescript phagemids by superinfection with the filamentous helper phage ExAssist™, which is included in the ZAP™-cDNA synthesis kit from Stratagene, according to the manufactures instructions.

The phagemid DNA was subsequently isolated as described in reference 6.

The isolated DNA of eight *cbh* cDNA clones were subjected to restriction analysis using the following restriction enzymes: *EcoRI* and *XhoI*. The DNA was digested for 2 hours at 37°C in a reaction mixture composed of the following solutions; 2 µl (»1 µg) DNA solution; 2 µl of the appropriate 10x React buffer (Life Technologies); 10 U of each Restriction enzyme (Life Technologies) and sterile distilled water to give a final volume of 20 µl. After addition of 4 µl DNA loading buffer the samples were loaded on a 0.7% TAE-agarose gel. The DNA fragments were separated by electrophoresis at 80 V for 1.5 hours. The restriction analysis revealed that these further *cbh* cDNA clones, called *cbhB* clones, had an insert size of approximately 1.8 kb.

**Example 3.5 Sequence analysis and description of *A. niger cbh* cDNA clone CBHB-C13**

The sequence of the *A. niger cbh* cDNA clones was determined by subcloning fragments from one of the positive clones of Example 3.4 with 1.8 kb insert, called CBHB-C13. Sequencing reactions were conducted using the ThermoSequenase fluorescent labelled primer cycle sequencing kit (Amersham) with universal sequencing primers. The sequencing reactions were analyzed on an ALFexpress sequencer (Pharmacia). Sequence analysis was performed using the Winstar programme (LaserGene) and yielded a nucleotide sequence consisting of 1781 nucleotides as shown in SEQ ID No:9, cDNA CBHB-C13. Clone CBHB-C13 contains one open reading frame from nucleotide 28 up to and including nucleotide 1635. The encoded protein consists of 536 amino acids (SEQ ID No:10, CBHB), with a calculated mass of 56,226 Da. The protein contains three domains: a catalytic domain from amino acid 22 up to and including amino acid 460; a linker peptide from amino acid 461 up to and including 500 and a cellulose binding domain (CBD) from amino acid 501 up to and including amino acid 536. After removal of the pre-sequence of 21 amino acids from the N-terminus, the mature protein consists of

515 amino acids with a calculated mass of 53,989 Da and a calculated isoelectric point of 3.9.

5

#### **Example 4 Expression of *cbhA* cDNA in *A. niger***

##### **Example 4.1 Construction of a *cbhA* expression cassette**

The cDNA clone CBHA-C9 of Example 2.5 was used to construct an expression plasmid. All digestions were conducted using 1-3  $\mu$ g DNA, 2  $\mu$ l of the appropriate 10x React buffer (Life Technologies), 10 U Restriction enzyme (Life Technologies) and sterile distilled water to give a final volume of 20  $\mu$ l. Digestions were incubated at 37°C for 1-2 hours. The digested samples were loaded on a 0.7% agarose gel in TAE buffer and the DNA fragments were subsequently separated by electrophoresis at 80 V. After electrophoresis the appropriate band was sliced out and the DNA fragment was isolated using the GeneClean™ kit (Biogel 101 Inc.), according to the manufacturers instructions. Ligations were performed in 10x ligation buffer (Promega), 50-100 ng vector DNA, a 3-fold molar excess of insert DNA, 1 U T4-Ligase resulting in a final volume of 10  $\mu$ l. After incubation overnight at 15°C, 4  $\mu$ l of the mixture was used to transform *E. coli* DH5a competent cells prepared as follows: 200  $\mu$ l of an *E. coli* DH5a overnight culture pregrown in LB medium (LB medium per 1000 ml: 10 g trypticase peptone (Life Technologies), 5 g yeast extract (Life Technologies), 10 g NaCl, 0.5 mM Tris-HCl pH 7.5). This culture was incubated in an orbital shaker at 37°C until its density corresponds to an O.D.600 of 0.15-0.2. The bacteria were then collected by centrifugation at 3000 rpm at 4°C. After discarding the supernatant the cells were kept on ice constantly. The bacterial pellet was washed in 40 ml K-MES transfer buffer (per 100 ml: 6 ml 1 M CaCl<sub>2</sub>, 0.5 ml 1 M MgCl<sub>2</sub>, 4 ml 0.5 M K-MES pH 6.0, 0.5 ml 1 M MnCl<sub>2</sub>) by resuspending these cells followed by centrifugation as described above. Finally the cells were resuspended in 10 ml K-MES/ glycerol (=K-MES transfer buffer + 15% glycerol). Aliquots (50  $\mu$ l) were either used immediately for transformation or frozen at -70°C.

*E. coli* DH5a competent cells were used in transformation experiments by combining 50  $\mu$ l of the cells suspension with 4.5  $\mu$ l of the ligation mixture. After a 30 minute incubation period on ice, the cells were incubated for 2 minutes at 42°C. Then 1 ml LB medium was added and the cells were incubated at 37°C for 1 hour.



The bacteria were then collected by centrifugation at 14,000 g for 20 seconds after discarding the supernatant the cells were resuspended in 100  $\mu$ l LB medium. The resulting bacterial suspension was plated on LB medium containing 100  $\mu$ g/ml ampicillin and 50  $\mu$ g/ml X-gal and 60  $\mu$ g/ml IPTG in case of blue/white screening.

5 A selection of 2-12 of the resulting colonies were grown overnight in LB medium containing 100  $\mu$ g/ml ampicillin. From the cultures plasmid DNA was isolated by the alkaline lysis method as described in reference 3, which was used in restriction analysis as described in Example 2.4 to select a clone harboring the desired plasmid.

10 The above described methods were used to construct the *cbhA* expression plasmid designated as pIM3006. The plasmid promH, containing the *A. niger pkiA* promoter, was digested with *Sst*I and *Nsi*I and ligated into *Sst*I-*Nsi*I digested pCBHA-C9. After linearization of the resulting plasmid, which contains the *pkiA-cbhA* fusion, using *Sst*I, this construct was partial digested using *Bgl*II. The 2.4 kb  
15 *Sst*I-*Bgl*II fragment was ligated into a *Bam*HI-*Sst*I digested plasmid carrying a *Bam*HI-*Hind*III fragment containing the terminator of the *Aspergillus nidulans trpC* gene. The *Bgl*II restriction site lies approximately 150 bp downstream the putative stop codon. The resulting construct is designated as pIM3006.

20 Plasmid DNA was isolated on a large scale from 100 ml cultures *E. coli* DH5a containing the final expression plasmid pIM3006 grown in LB medium containing 100  $\mu$ g/ml ampicillin using the Nucleobond PC-100 kit (Nagel) according to the manufacturers instructions.

#### 25 **Example 4.2 Introduction of the *cbhA* expression construct in *A. niger* by cotransformation and screening of transformants for the expression of the *A. niger cbhA* gene**

30 The plasmid pIM3006, obtained in Example 4.1 was introduced in *A. niger* by cotransformation using the *A. niger pyrA* as a selective marker on the plasmid pGW635 (Goosen *et al.*, 1989) and the plasmid pIM3006 as the cotransforming plasmid.

Protoplasts were prepared from mycelium by growing *A. niger* on minimal medium supplemented with 0.5% yeast extract, 0.2% casamino acids, 50 mM glucose, 1.5 mM leucine and 10 mM uridine for 18 hours at 30°C. The preparation of protoplasts of *A. niger* and the transformation procedure was performed as

described in reference 8. The resulting PYR<sup>+</sup> transformants were analyzed for the expression of the *A. niger cbhA* gene by SDS-PAGE.

The transformants were analyzed for the formation of the *A. niger cbhA* gene product, the CBHA protein. Twelve of these transformants were used in a growth experiment to analyze for CBHA expression. The transformants were grown in 50 ml minimal medium (per litre medium: 15 g KH<sub>2</sub>PO<sub>4</sub>, 0.5 g KCl, 0.5 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 6 g NaNO<sub>3</sub>, 1 ml Vishniac solution<sup>7</sup> ( pH 6.0) containing 4% D-fructose, 0.1% yeast extract and 1.5 mM leucine for 30 hours at 30°C. After growth, the mycelium was removed by filtration and the culture filtrate was analyzed by SDS-PAGE. Transformant *A. niger*: pIM3006-55 was the best CBHA producer (approx. 20 µg ml<sup>-1</sup>).

#### Example 4.3 Purification of the CBHA enzyme

A 1 l Erlenmeyer flask containing 300 ml medium (per litre medium: 4 g NH<sub>4</sub>Cl, 1.5 g KH<sub>2</sub>PO<sub>4</sub>, 0.5 g KCl, 0.5 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 1 ml Vishniac salts solution (contains per litre 10 g EDTA, 4.4 g ZnSO<sub>4</sub>·7H<sub>2</sub>O, 1.0 g MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.32 g CoCl<sub>2</sub>·6H<sub>2</sub>O, 0.32 g CuSO<sub>4</sub>·5H<sub>2</sub>O, 0.22 g (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>·4H<sub>2</sub>O, 1.47 g CaCl<sub>2</sub>·2H<sub>2</sub>O, 1.0 g FeSO<sub>4</sub>·7H<sub>2</sub>O, pH 4.0 (Vishniac and Santer, 1957), pH 5.0) supplemented with 0.2% ultrafiltrated yeast extract (Life Technologies), 0.5% casamino acids (Life Technologies), 10 mM leucine and 5% sucrose as C-source) was inoculated with 2x10<sup>9</sup> spores of strain *A. niger*: pIM3006-55 and grown for 6 hours at 30°C. This entire culture was then added to 1.7 litre medium in an Applikon 3 litre fermentor (BTS06) which was controlled by a Bioprocessor unit ADI 1020 (Applikon). The fermentation was temperature controlled at 30°C and the pH was allowed to drop to pH 3.0 at which it was maintained by addition of 5N NaOH. 1.9 litre culture broth was drained from the fermentor 30 hours after inoculation. The mycelia was separated from the culture broth by filtration. The fermentor was filled again with 2 litre of fresh culture medium and cultured for another 30 hours as described above. After this second run the mycelia was separated from the culture broth by filtration. The culture filtrates from the both runs were combined and piperazine was added to a final concentration of 10 mM. The pH was adjusted to pH 5.5 with 10 N NaOH. The culture filtrate was then applied on a Streamline™ anion exchange column (Pharmacia). After loading, the column was washed with 10 mM piperazine-HCl (pip-HCl) pH 5.5. The enzyme was eluted from the column by applying a salt pulse

using 10 mM pip-HCl pH 5.5 / 1 M NaCl. Fractions of 15 ml were collected and CBH activity was detected using the chromogenic substrate 4-methylumbelliferyl- $\beta$ -D-cellobioside (detects cellobio-hydrolases)(Sigma M-6018). The majority of the CBH activity was present in the fractions 2, 3 and 4. These fractions were pooled and dialysed against 10 mM pip-HCl pH 5.5.

#### Example 4.4 Influence of pH and temperature of the CBHA enzyme

Enzyme purified as described in Example 4.3 was used for characterization of the enzyme activity. A commercially available CBHI-containing *Trichoderma* enzyme preparation: Econase CEP (ALKO) was used as a reference enzyme. The assay for the determination of cellobiohydrolase activity is based on the property of the enzyme to hydrolyse p-nitrophenyl- $\beta$ -D-cellobioside.

To determine the influence of the temperature on the enzyme activity 25  $\mu$ l enzyme solution was added to 225  $\mu$ l of a solution containing 0.1% p-nitrophenyl- $\beta$ -D-cellobioside and 25 mM gluconolactone in a 50 mM sodium acetic acid buffer at pH 5.0. Following an incubation of 15 minutes at 30, 40, 50, 60 or 70°C the enzyme activity was stopped by adding 750  $\mu$ l of a 0.5 M glycine buffer pH 9.0. The concentration of p-nitrophenol formed was determined spectrophotometrically at 400 nm.

In Fig 1 it can be seen that under these assay conditions the temperature optimum of the enzyme is 60°C. Econase also has a temperature optimum at 60°C, but the remaining activity at 70°C of CBHA is 40%, whereas the Econase activity at that temperature is less than 10 % of the maximum activity.

To determine the pH optimum of CBHA the assay as described above was performed at 40°C, using 100 mM sodium citric acid buffer, set at different pH values. The activity was measured at pH-values ranging from 2.5 up to 6.5.

In Figure 2 it can be seen that the pH profile of CBHA is completely different from the pH profile of Econase. CBHA has more than 50% activity at the whole range from 2.5 to 6.5, whereas below pH 4.5 the Econase activity decrease to below 50%.

#### Example 5 Expression of *cbhB* cDNA in *A. niger*

##### Example 5.1 Construction of a *cbhB* expression cassette

The cDNA clone CBHB-C13 of Example 3.4 was used to construct an expression plasmid. To introduce a *NsiI* restriction site at the ATG codon, a PCR was conducted using the following oligo nucleotides:

5 5'-CAC AAT GCA TTC CTT CCA AAT CTA CCG-3' (27-mer)  
(SEQ ID No: 11)

5'-CAC CGT CAG CGT CCA TGG CG-3' (20-mer)  
(SEQ ID No: 12)

10

These oligo nucleotides were used in PCR using CBHB-C13 DNA as template<sup>5</sup>. For a PCR 5 ng of CBHB-C13 DNA was combined with 2.5  $\mu$ l 10x reaction buffer (200 mM Tris-HCl (pH 8.4), 500 mM KCl); 0.75  $\mu$ l 50 mM Mg<sub>2</sub>Cl<sub>2</sub>; 4  $\mu$ l dNTPs (1.25 mM each); 25 pmol of each oligo nucleotide and 0.2 U Taq polymerase (Life  
15 Technologies). The 530 bp PCR product was ligated into a pGEM-T vector (Promega) according to the manufactures instructions, resulting in plasmid pMK49. Plasmid DNA was isolated by the alkaline lysis method<sup>6</sup>. Nucleotide sequence analysis was conducted as described in Example 2.5. A 530 bp *SstI/NcoI* fragment was isolated from plasmid pMK49 and ligated into a 4.4 kb *SstI/NcoI* fragment isolated from  
20 plasmid CBHB-C13 resulting in pLIG418. All DNA manipulations were conducted as described in Example 4.1.

To introduce a *BamHI* restriction site directly downstream the stop codon, a PCR was conducted using the following oligo nucleotides:

25 5'-CCT AGG ATC CTA CAA ACA CTG CGA GTA GTA C-3' (31-mer)  
(SEQ ID No: 13)

5'-GTT AAC TCG TCG GCC TCG-3' (18-mer)  
(SEQ ID No: 14)

30

PCR was conducted as described above. The 340 bp PCR product was ligated into a pGEM-T vector (Promega), resulting in plasmid pMK51. After the sequence analysis, a 0.4 kb *SphI/BamHI* fragment was isolated from plasmid pMK51 and ligated into pUC18, which was linearized using *SphI* and *BamHI*. This resulted in

plasmid pLIG420. A 350 bp *KpnI/HpaI* fragment was isolated from plasmid pLIG420 and ligated into a 4.45 kb fragment isolated from pLIG418, resulting into pLIG423.

The plasmid promH, containing the *A. niger pkiA* promoter, was digested with *SstI* and *NsiI* to obtain a 0.75 kb fragment which was ligated into *SstI-NsiI* digested pLIG423. The resulting plasmid pLIG425, which contains the *pkiA-cbhB* fusion, was digested with *BamHI* and a 2.4 kb fragment was isolated. This fragment was ligated into a with *BamHI* linearized plasmid carrying a *BamHI-HindIII* fragment containing the terminator of the *Aspergillus nidulans trpC* gene. The resulting construct is designated as pIM3011.

Plasmid DNA was isolated on a large scale from 100 ml cultures *E. coli* DH5a containing the final expression plasmid pIM3011 grown in LB medium containing 100  $\mu$ g/ml ampicillin using the Nucleobond PC-100 kit (Nagel) according to the manufacturers instructions.

#### **Example 5.2 Introduction of the *cbhB* expression construct in *A. niger* by cotransformation and screening of transformants for the expression of the *A. niger cbhB* gene**

The plasmid pIM3011, obtained in Example 5.1 was introduced in *A. niger* by cotransformation using the *A. niger pyrA* as a selective marker on the plasmid pGW635 (Goosen *et al.*, 1989) and the plasmid pIM3011 as the cotransforming plasmid. The transformation procedure and screening of CBHB producing strains was performed as described in Example 4.2.

Transformant *A. niger*: pIM3011-34 was the best CBHB producer.

#### **Example 5.3 Purification of the CBHB enzyme**

Twelve 1 l Erlenmeyer flasks containing 250 ml medium (per litre medium: as described in Example 4.3 with 7% D-glucose in stead of 5% sucrose) were inoculated with  $2 \times 10^6$  spores per ml medium of strain *A. niger*: pIM3011-34 and grown for 30 hours at 30°C. After growth, the mycelia was separated from the culture broth by filtration and 4-methylpiperazine was added to a final concentration of 20 mM. The pH was adjusted to pH 4.5 with 10 N NaOH. The culture filtrate was then applied on a Streamline™ anion exchange column (Pharmacia). After loading, the column was washed with 20 mM 4- methylpiperazine-HCl (pip-HCl) pH 4.5. The enzyme was eluted from the column by applying a salt pulse using 20 mM

4-methylpiperazine-HCl pH 4.5 / 1 M NaCl. Fractions of 25 ml were collected and CBH activity was detected using the chromogenic substrate 4-methylumbelliferyl- $\beta$ -D-cellobioside (detects cellobiohydrolases)(Sigma M-6018). The majority of the CBH activity was present in the fractions 2, 3 and 4. Fractions  
5 containing cellobiohydrolase activity were pooled and dialyzed against 20 mM 4-methylpiperazine pH 4.5. Half of the enzyme solution was applied on a Resource30Q column (Pharmacia), washed with 20 mM 4-methylpiperazine-HCl pH 4.5, and eluted from the column using a 0-0.5 M NaCl gradient in 20 mM 4-methylpiperazine-HCl buffer pH 4.5. Fractions containing cellobiohydrolase activity  
10 on the chromogenic substrate 4-methylumbelliferyl- $\beta$ -D-cellobioside were pooled.

#### Example 5.4 Influence of pH and temperature of the CBHB enzyme

Enzyme purified as described in Example 5.3 was used for characterization of the enzyme activity. A commercially available CBHI-containing *Trichoderma* enzyme  
15 preparation: Econase CEP (ALKO) was used as a reference enzyme. The assay for the determination of cellobiohydrolase activity is based on the property of the enzyme to hydrolyse p-nitrophenyl- $\beta$ -D-cellobioside.

To determine the influence of the temperature on the enzyme activity 25  $\mu$ l enzyme solution was added to 225  $\mu$ l of a solution containing 0.1%  
20 p-nitrophenyl- $\beta$ -D-cellobioside and 25 mM gluconolactone in a 50mM sodium acetic acid buffer at pH 5.0. Following an incubation of 15 minutes at 30, 40, 50, 60 or 70°C the enzyme activity was stopped by adding 750  $\mu$ l of a 0.5 M glycine buffer pH 9.0. The concentration of p-nitrophenol formed was determined spectrophotometrically at 400 nm.

25 In Fig 3 it can be seen that under these assay conditions the temperature optimum of the enzyme is 60°C. Econase also has a temperature optimum at 60°C, but the remaining activity at 70°C of CBHB is over 70%, whereas the Econase activity at that temperature is less than 10 % of the maximum activity.

To determine the pH optimum of CBHB the assay as described above was  
30 performed at 40°C, using 100 mM sodium citric acid buffer, set at different pH values. The activity was measured at pH-values ranging from 2.5 up to 6.5.

In Figure 4 it can be seen that the pH profile of CBHB is completely different from the pH profile of Econase. CBHB has more than 80% activity at the whole range from 2.5 to 6.5, whereas below pH 4.5 the Econase activity decrease to below 50%.

**Example 6      Baking trial using *A. niger* CBH**

The *A. niger* CBH of Example 4 was tested in a full scale baking trial using the following recipe:

5

- \*      Flour 2000 g Kolibri (Meneba) & 500 g Ibis (Meneba)
- \*      1305 g water
- \*      50 g block yeast, Konings gist® (obtainable from Gist-brocades)
- \*      50 g salt
- 10      \*      12.5 g sugar
- \*      25 ppm ascorbic acid
- \*      20 ppm Fermizyme® P200 (Fungal  $\alpha$ -amylase, obtainable from Gist-brocades)
- \*      6.25 mg *A. niger* CBH from Example 4

15

20

After mixing the flour with the ingredients, three 900 gram doughs are fermented. The first fermentation for 30 and the second for 40 minutes at 30°C and after shaping a final fermentation was for 70 minutes at 35°C. The bread was baked for 30 minutes at 250°C. Thereafter the loaves were allowed to cool and loaf volume was determined by the rape seed displacement method, and other properties were determined. The baking test was carried out as a parallel test in triplicate. The result was a 10% increase in loaf volume, improved dough handling properties and crumb structure, all as compared to a reference where no CBH was added.

## REFERENCES

- 5           1. Goosen, T., Engelenburg, F. van, Debets, F., Swart, K., Bos,  
K., Briek, H. van den, (1989) Mol. Gen. Genet. 219: 282-288
2. Harmsen, J.A.M. *et al.*, (1990) Curr. Genet. 18: 161-166.
3. Maniatis T., E. F. Fritsch, J. Sambrook (1982): Molecular cloning, a  
laboratory manual; Cold Spring Harbor Laboratory, New York.
- 10          4. Murray, N.(1977) Mol. Gen. Genet. 150: 53-58
5. Saiki R.K. *et al.* (1988) Science, 239, 487-491
6. Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) In: *Molecular Cloning: a  
Laboratory Manual*, 2nd edn., Cold Spring Harbor Labatory Press, NY.
7. Visniac, W. and Santer, M. (1957), Bact. Rev. 21: 195-213
- 15          8. Goosen *et al.*, 1992, "Transformation and Gene Manipulation in filamentous  
fungi: an overview" In: Handbook of Applied Mycology" Vol. 4: "Fungal  
Biotechnology",
9. D.K. Arora, R.P. Elander and K.G. Mukerji, eds., Marcel Deker In., New  
York, pp 151-195 for filamentous fungi
- 20          10. Romanos *et al.*, 1992, Yeast 8: 423-488
11. Rosenfeld et al (Rosenfeld J, Cadevielle J, Guillemot J C, Ferrara P (1992)  
Anal. Biochem. 203,173-179).
- 25



**EDITORIAL NOTE FOR APPLICATION**

**NO. 92596/98**

**THE SEQUENCE LIST FOLLOWS WITH  
PAGE NUMBERS 1 - 9**

**THE CLAIMS FOLLOW THE SEQUENCE  
LIST ON PAGE 30**

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: Gist-Brocades
- (B) STREET: Wateringseweg 1
- (C) CITY: Delft
- (E) COUNTRY: The Netherlands
- (F) POSTAL CODE (ZIP): 2611 XT

(ii) TITLE OF INVENTION: cellulose-degrading enzymes

(iii) NUMBER OF SEQUENCES: 14

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Leu Tyr Leu Met Ser Asp Asp Ser Asn Tyr Glu Leu Phe Lys  
 1                        5                                10

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Leu Gly Asn Thr Asp Phe Tyr Gly Pro Gly Leu Thr Val Asp Thr Asn  
 1                        5                                10                                15

Ser Pro Phe Thr Val Val Thr Gln  
                                20

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GTCCGTACCA ACATGGCC

18

5

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ACTCAGAAAC ATTGGCTATA G

21

20

25

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1701 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Aspergillus niger*  
 (B) STRAIN: N402

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CGAAATCACT AAAAGGAGAC GACTAGAGTC TTATACAATC TCATTACAAT GCATCAACGT	60
GCCCTTCTCT TCTCAGCCCT GCTGACGGCT GTTCGCGCCC AGCAAGCCGG AACGCTCAGC	120
GAGGAAGTCC ATCCTTCCTT GACCTGGCAG AAATGCACCT CTGAAGGCAG CTGCACTGAA	180
CAGAGTGGCT CAGTTGTCAT TGACTCGAAC TGGCGCTGGA CCCATTCCGT CAATGACAGC	240
ACCAATTGCT ACACTGGCAA CACCTGGGAT GCAACTCTCT GCCCTGATGA TGAGACCTGT	300
GCGGCCAACT GCGCCCTGGA CGGAGCAGAC TACGAGTCCA CCTACGGTGT CACCACTGAC	360
GGTGATTCAT TGACACTGAA ATTCGTCACCT GGCTCCAATG TTGGCTCGCG GTTGTATCTA	420
ATGGACACGA GCGACGAGGG ATACCAGACG TTCAACTTGC TTGACGCAGA GTTCACTTTC	480
GACGTTGATG TGTCTAACCT CCCATGTGGG CTAAACGGCG CGTTGTACTT CACTGCAATG	540
GACGCCGATG GTGGAGTCTC AAAATACCCT GCCAATAAGG CTGGAGCCAA GTACGGAACA	600
GGATACTGTG ACTCCCAATG CCCCCGGGAC CTGAAATTCA TCGACGGACA AGCCAACGTC	660
GATGGCTGGG AACCTTCTAG CAACAATGAC AACACAGGTA TCGGCAATCA CGGTTCTTGC	720
TGCCCTGAAA TGGATATCTG GGAGGCAAAC AAGATCTCGA CCGCATTGAC ACCCCATCCT	780
TGTGACAGCA GCGAACAGAC CATGTGTGAG GGTAACGACT GCGGTGGAAC CTA CTCTCGGAT	840
GATCGCTACG GAGGAACCTG CGACCCTGAC GGCTGCGACT TCAACCCTTA TCGCATGGGC	900

70

65

50

45

40

35

30

AACGACTCTT TCTACGGTCC TGGCAAGACC ATCGACACCG GATCCAAGAT GACGGTTGTG 960  
 ACCCAGTTCA TCACTGATGG CTCTGGCTCC CTCAGCGAGA TCAAGCGTTA CTACGTGCAG 1020  
 5 AACGGAAATG TTATAGCGAA CGCTGATTC AACATCTCTG GAGTACTGG AACTCGATC 1080  
 ACAACGGACT TCTGCACTGC GCAGAAGAAG GCCTTTGGCG ACGAGGATAT ATTGCTGAG 1140  
 10 CACAATGGAC TTGCTGGAAT CAGTGATGCC ATGTCTTCCA TGGTTCAT CTGAGCTTG 1200  
 TGGGATGATT ACTATGCCAG CATGGAGTGG CTCGACAGCG ACTATCCCGA GAACGCTACC 1260  
 GCTACCGACC CAGGTGTTGC ACGCGGAACA TGCGACTCGG AATCAGGCGT CCCTGCGACA 1320  
 15 GTCGAGGGGG CGCATCCCGA TTCTTCGGTG ACCTTCTCAA ACATCAAGTT CGGTCCCATC 1380  
 AACTCGACCT TCAGCGCTTC CGCATAAGGG GAAGTGCAGG GTTCAGAGCC TCAATTACAT 1440  
 20 CCCACAAGCC GAACACAACA GGACAGGTTT CTGGAACGGA ATATGGGAGA GTTGCGGGCT 1500  
 TGTAATAGT CCGAAAAGTG GTCACTGCTT TTGTGTATCG GGCTGTCTTC CCATTTTATT 1560  
 TTCATTCCT CCAGATTGAT TGATGCAGAT CTTGTTTGT GATTCTTCA CTCGTGCTG 1620  
 25 TGACTTTTTG TACCTAATT CACATTCGTT TCTTTTCTGT TTTTGAGTCC ACTCAATCCA 1680  
 AGAGACAGTT GTTCCTTGC T 1701

30

## (2) INFORMATION FOR SEQ ID NO: 6:

35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

45 GAYGAYAGYA AYTAYGARTT RTTYAA 26

## (2) INFORMATION FOR SEQ ID NO: 7:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 20 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 55 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

60 GTRAANGGRC TRTTNGTRTC 20

## (2) INFORMATION FOR SEQ ID NO: 8:

65 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 452 amino acids  
 (B) TYPE: amino acid  
 70 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Aspergillus niger*

(B) STRAIN: N402

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

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

Ser Gly Val Thr Gly Asn Ser Ile Thr Thr Asp Phe Cys Thr Ala Gln  
 340 345 350  
 5 Lys Lys Ala Phe Gly Asp Glu Asp Ile Phe Ala Glu His Asn Gly Leu  
 355 360 365  
 Ala Gly Ile Ser Asp Ala Met Ser Ser Met Val Leu Ile Leu Ser Leu  
 370 375 380  
 10 Trp Asp Asp Tyr Tyr Ala Ser Met Glu Trp Leu Asp Ser Asp Tyr Pro  
 385 390 395 400  
 Glu Asn Ala Thr Ala Thr Asp Pro Gly Val Ala Arg Gly Thr Cys Asp  
 405 410 415  
 15 Ser Glu Ser Gly Val Pro Ala Thr Val Glu Gly Ala His Pro Asp Ser  
 420 425 430  
 20 Ser Val Thr Phe Ser Asn Ile Lys Phe Gly Pro Ile Asn Ser Thr Phe  
 435 440 445  
 Ser Ala Ser Ala  
 450  
 25

2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1767 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 30  
 (ii) MOLECULE TYPE: DNA (cDNA)  
 35  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Aspergillus niger  
 (B) STRAIN: N402  
 40  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:  
 GCAGTGAATC ACAATGTCTT CCTTCAAAT CTACCGAGCA GCTCTGCTGC TCTCTATCCT 60  
 45 TGCCACTGCC AATGCTCAGC AGGTTGGCAC CTACACCACT GAGACACATC CGTCTCTGAC 120  
 CTGGCAAACC TGCACCAGCG ATGGCAGCTG CACTACCAAC GACGGCGAGG TGGTCATCGA 180  
 50 CGCCAATTGG CGTTGGGTGC ACTCAACCTC CAGTGCCACC AACTGCTATA CTGGCAACGA 240  
 ATGGGACACC TCAATTTGTA CCGACGATGT GACCTGCGCC GCGAAGTGGC CGCTTGACGG 300  
 TGCCACTTAC GAGGCGACCT ATGGTGTGAC CACGTCCGGC AGCGAGCTGC GCCTGAAGTT 360  
 55 CGTCACTCAA GGCTCCAGCA AGAACATCGG GTCTCGTTTG TACCTCATGA GCGACGACAG 420  
 CAACTATGAG CTGTCAAGC TCCTCGGCCA GGAGTTCACC TTTGATGTCG ACGTGTCCAA 480  
 60 CCTTCCTGCG GGTCTCAACG GCGCGCTCTA CTTGCTCGCC ATGGACGCTG ACGGTGGCAC 540  
 CTCGGAGTAT TCCGGCAACA AGGCCGGTGC CAAATACGGA ACCGGCTACT GCGACTCGCA 600  
 GTGCCCTCGC GACCTGAAGT TCATCAACGG CGAAGCCAAC TGGACGGCT GGGAGCCGTC 660  
 65 CAGCAACAAC GTCAACACCG GAGTTGGCGA CCACGGCTCC TGCTGCGCCG AGATGGACGT 720  
 CTGGGAAGCG AACAGCATCT CCAACGCCTT CACCGCGCAC CCCTGCGACT CCGTCAGCCA 780  
 70 GACAATGTGC GACGGTGACT CCTGTGGTGG AACTTACAGC GCCAGCGGCG ACCGCTACAG 840

CGGCACCTGC GACCCCGACG GCTGCGACTA CAACCCCTAC CGTCTAGGCA ACACGGACTT 900  
 TTACGGCCCC GGCCTGACCG TCGACACGAA CAGCCCCTTC ACCGTCTGCA CCCAGTTCAT 960  
 5 CACCGATGAC GGCACCTCCT CCGGCACCTT GACCGAGATC AAGCGTTGT ACGTGCAGAA 1020  
 CGGGGAGGTC ATCGCCAACG GCGCCTCCAC CTA CTACTCCAGT GTCAACGGCA GCTCTATCAC 1080  
 10 CTCCGCTTTC TGTGAATCAG AGAAGACGCT GTTCGGCGAC GAGAACGTCT TCGACAAGCA 1140  
 CGGCGGTCTC GAGGGCATGG GCGAGGCTAT GGCCAAGGGC ATGGTCTTGG TCTTGAGTCT 1200  
 15 TTGGGATGAC TATGCCGCTG ACATGCTCTG GCTCGACAGC GACTACCCCG TTA ACTCGTC 1260  
 GGCCTCGACC CCTGGTGTGG CCCGCGGTAC TTGTAGCACG GACTCGGGTG TCCCGGCTAC 1320  
 CGTGGAGGCG GAGTCACCCA ATGCCTATGT CACGTA CTG AACATCAAGT TCGGGCCTAT 1380  
 20 TGGCTCGACT TACTCCAGTG GATCTTCTTC GGGGTCGGGG TCTAGCTCCA GCTCGAGTTC 1440  
 GACTACCACT AAGGCCACTT CGACGACCTT GAAGACTACC TCGACCACCA GCAGTGGAAAG 1500  
 25 CAGTTCGACA TCGGCGGCGC AGGCGTATGG ACAGTGTGGT GGACAGGGCT GGACTGGTCC 1560  
 GACCACTTGT GTGAGTGGGT ACACTTGACG GTATGAGAAT GCGTACTACT CGCAGTGTTC 1620  
 GTAGACACTA CTCGAAGGGG GATGAAGATG GCAGTGTGGG GGGGGGAATA GTTAGTGATT 1680  
 30 AGTAGTTAGC ACACGTACTA CAGTTTAGTC AAATTTAGGA AGCAATATGA CATGATCACT 1740  
 TGATGCAAAA AAAAAAAAAA AAAAAAA 1767

35

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 536 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Aspergillus niger*  
 (B) STRAIN: N402

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ser Ser Phe Gln Ile Tyr Arg Ala Ala Leu Leu Leu Ser Ile Leu  
 1 5 10 15  
 Ala Thr Ala Asn Ala Gln Gln Val Gly Thr Tyr Thr Thr Glu Thr His  
 20 25 30  
 Pro Ser Leu Thr Trp Gln Thr Cys Thr Ser Asp Gly Ser Cys Thr Thr  
 35 40 45  
 Asn Asp Gly Glu Val Val Ile Asp Ala Asn Trp Arg Trp Val His Ser  
 50 55 60  
 Thr Ser Ser Ala Thr Asn Cys Tyr Thr Gly Asn Glu Trp Asp Thr Ser  
 65 70 75 80  
 Ile Cys Thr Asp Asp Val Thr Cys Ala Ala Asn Cys Ala Leu Asp Gly  
 85 90 95

70

	Ala	Thr	Tyr	Glu	Ala	Thr	Tyr	Gly	Val	Thr	Thr	Ser	Gly	Ser	Glu	Leu
				100					105						110	
5	Arg	Leu	Asn	Phe	Val	Thr	Gln	Gly	Ser	Ser	Lys	Asn	Ile	Gly	Ser	Arg
			115					120					125			
	Leu	Tyr	Leu	Met	Ser	Asp	Asp	Ser	Asn	Tyr	Glu	Leu	Phe	Lys	Leu	Leu
		130					135					140				
10	Gly	Gln	Glu	Phe	Thr	Phe	Asp	Val	Asp	Val	Ser	Asn	Leu	Pro	Cys	Gly
	145					150					155					160
	Leu	Asn	Gly	Ala	Leu	Tyr	Phe	Val	Ala	Met	Asp	Ala	Asp	Gly	Gly	Thr
15				165						170					175	
	Ser	Glu	Tyr	Ser	Gly	Asn	Lys	Ala	Gly	Ala	Lys	Tyr	Gly	Thr	Gly	Tyr
				180					185					190		
20	Cys	Asp	Ser	Gln	Cys	Pro	Arg	Asp	Leu	Lys	Phe	Ile	Asn	Gly	Glu	Ala
			195					200					205			
	Asn	Cys	Asp	Gly	Trp	Glu	Pro	Ser	Ser	Asn	Asn	Val	Asn	Thr	Gly	Val
		210					215					220				
25	Gly	Asp	His	Gly	Ser	Cys	Cys	Ala	Glu	Met	Asp	Val	Trp	Glu	Ala	Asn
	225					230					235					240
	Ser	Ile	Ser	Asn	Ala	Phe	Thr	Ala	His	Pro	Cys	Asp	Ser	Val	Ser	Gln
30				245					250						255	
	Thr	Met	Cys	Asp	Gly	Asp	Ser	Cys	Gly	Gly	Thr	Tyr	Ser	Ala	Ser	Gly
				260					265					270		
35	Asp	Arg	Tyr	Ser	Gly	Thr	Cys	Asp	Pro	Asp	Gly	Cys	Asp	Tyr	Asn	Pro
			275					280					285			
	Tyr	Arg	Leu	Gly	Asn	Thr	Asp	Phe	Tyr	Gly	Pro	Gly	Leu	Thr	Val	Asp
		290					295					300				
40	Thr	Asn	Ser	Pro	Phe	Thr	Val	Val	Thr	Gln	Phe	Ile	Thr	Asp	Asp	Gly
	305					310					315					320
	Thr	Ser	Ser	Gly	Thr	Leu	Thr	Glu	Ile	Lys	Arg	Leu	Tyr	Val	Gln	Asn
45				325						330					335	
	Gly	Glu	Val	Ile	Ala	Asn	Gly	Ala	Ser	Thr	Tyr	Ser	Ser	Val	Asn	Gly
				340					345					350		
50	Ser	Ser	Ile	Thr	Ser	Ala	Phe	Cys	Glu	Ser	Glu	Lys	Thr	Leu	Phe	Gly
			355					360					365			
	Asp	Glu	Asn	Val	Phe	Asp	Lys	His	Gly	Gly	Leu	Glu	Gly	Met	Gly	Glu
		370					375					380				
55	Ala	Met	Ala	Lys	Gly	Met	Val	Leu	Val	Leu	Ser	Leu	Trp	Asp	Asp	Tyr
	385					390					395					400
	Ala	Ala	Asp	Met	Leu	Trp	Leu	Asp	Ser	Asp	Tyr	Pro	Val	Asn	Ser	Ser
60				405						410					415	
	Ala	Ser	Thr	Pro	Gly	Val	Ala	Arg	Gly	Thr	Cys	Ser	Thr	Asp	Ser	Gly
				420					425					430		
65	Val	Pro	Ala	Thr	Val	Glu	Ala	Glu	Ser	Pro	Asn	Ala	Tyr	Val	Thr	Tyr
			435					440					445			
	Ser	Asn	Ile	Lys	Phe	Gly	Pro	Ile	Gly	Ser	Thr	Tyr	Ser	Ser	Gly	Ser
		450					455					460				
70	Ser	Ser	Gly	Ser	Gly	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Thr	Thr	Thr	Lys



	465		470		475		480
	Ala Thr Ser Thr Thr Leu Lys Thr Thr Ser Thr Thr Ser Ser Gly Ser						
5			485		490		495
	Ser Ser Thr Ser Ala Ala Gln Ala Tyr Gly Gln Cys Gly Gly Gln Gly						
		500		505			510
10	Trp Thr Gly Pro Thr Thr Cys Val Ser Gly Tyr Thr Cys Thr Tyr Glu						
		515		520			525
	Asn Ala Tyr Tyr Ser Gln Cys Leu						
		530		535			

15

(2) INFORMATION FOR SEQ ID NO: 11:

20

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

30

CACAATGCAT TCCTTCCAAA TCTACCG

26

35

(2) INFORMATION FOR SEQ ID NO: 12:

40

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12

45

CACCGTCAGC GTCCATGGCG

50

(2) INFORMATION FOR SEQ ID NO: 13:

55

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 31 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13

60

CCTAGGATCC TACAAACACT GCGAGTAGTA C

26

65

(2) INFORMATION FOR SEQ ID NO: 14:

5

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

10

- (ii) MOLECULE TYPE: DNA (synthetic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14  
GTTAACTCGT CGGCCTCG

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A polypeptide which comprises the sequence of SEQ ID No: 8 (CBH A) or SEQ ID No: 10 (CBH B), or a sequence substantially homologous to either  
5 sequence, wherein the degree of homology is at least 80%.

2. A polypeptide from *Aspergillus* which is a CBH or has CBH activity.

3. A polypeptide according to any one of the preceding claims which has  
10 CBHI activity.

4. A polypeptide according to any of the preceding claims which is obtainable from or produced by a fungus.

15 5. A polypeptide according to claim 4 which is obtainable from or produced by *Aspergillus*.

6. A polypeptide according to claim 5 which is obtainable from or produced by *Aspergillus niger*.

20 7. A polynucleotide comprising a sequence encoding a polypeptide according to any one of the preceding claims.

8. A polynucleotide capable of selectively hybridising (as hereinbefore  
25 defined) to SEQ ID No: 5 or SEQ ID No: 9 or a complement thereof, wherein said polynucleotide is at least 80% homologous to said SEQ ID No: 5 or SEQ ID No: 9 or their complements.

9. A polynucleotide according to claim 7 or 8 which is a DNA sequence.

30 10. A vector which comprises a polynucleotide according to any one of claims 7 to 9.

11. A host cell which:



- (i) comprises a polynucleotide or vector according to any one of claims 7 to 10, which host cell does not naturally comprise such a polynucleotide or vector; and/or
- 5 (ii) expresses a polypeptide according to any one of claims 1 to 6, which host cell does not naturally express such a polypeptide; and/or
- (iii) expresses a polypeptide of the invention at a higher level than naturally expressed by the host cell; and/or
- (iv) has been transformed or transfected by a polynucleotide or vector
- 10 according to any one of claims 7 to 10.

12. A process for producing a host cell according to claim 11, the process comprising transforming or transfecting a host cell with a polynucleotide or vector according to any one of claims 7 to 10.

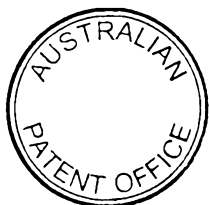
13. A method for producing a polypeptide according to any one of claims 1 to 6 comprising culturing a host cell according to claim 11 under conditions in which it expresses the polypeptide and optionally recovering the expressed polypeptide.

14. A polypeptide preparable or prepared by a process according to claim 13.

15. A composition comprising a polypeptide according to any one of claims 1 to 6 or 14 and optionally one or more enzymes.

16. A composition comprising a host cell according to claim 11 or prepared by a process according to claim 12.

17. Use of a polypeptide, host cell or composition according to any one of claims 1 to 6, 11 or 14 to 16 to degrade cellulose.



18. A method of degrading cellulose comprising contacting a polypeptide, host cell or composition according to any one of claims 1 to 6, 11, or 14 to 16 with the cellulose.

5 19. The use according to claim 17 in, or in a process for, producing food or beverages, animal feed, pulp, paper or textiles.

20. A method according to claim 18 which is part of a method for producing or processing food or beverages, animal feed, pulp, paper or textiles.

10

21. A method according to claim 20 wherein:

(i) the cellulose is contained in a material which after degradation of all or part of the cellulose within it is used to make a food or beverage for a human or animal, and/or

15

(ii) all or part of the cellulose is degraded after the food is ingested by a human or animal.

22. A bakery product or ingredient, or a baked product obtainable or obtained by a method according to claim 21.

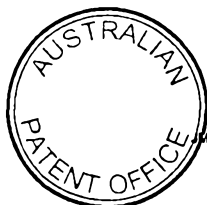
20

23. A human food or beverage, or animal feed comprising a polypeptide, host cell or composition according to any one of claims 1 to 6, 11, or 14 to 16.

25 24. The use according to claim 19 to degrade cellulose during wood pulp processing or during the manufacture of paper.

25. A polypeptide according to claim 1 or 2 substantially as hereinbefore described with reference to any of the examples.

30



26. A polynucleotide according to claim 6 or 7 substantially as hereinbefore described with reference to any of the examples.

5 DATED : 3 August, 2001

PHILLIPS ORMONDE & FITZPATRICK

Attorneys For:

DSM N.V.

10

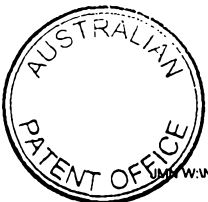
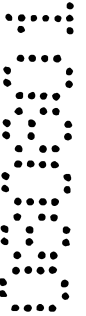
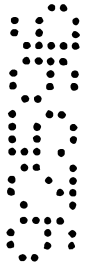


Figure 4 4/4

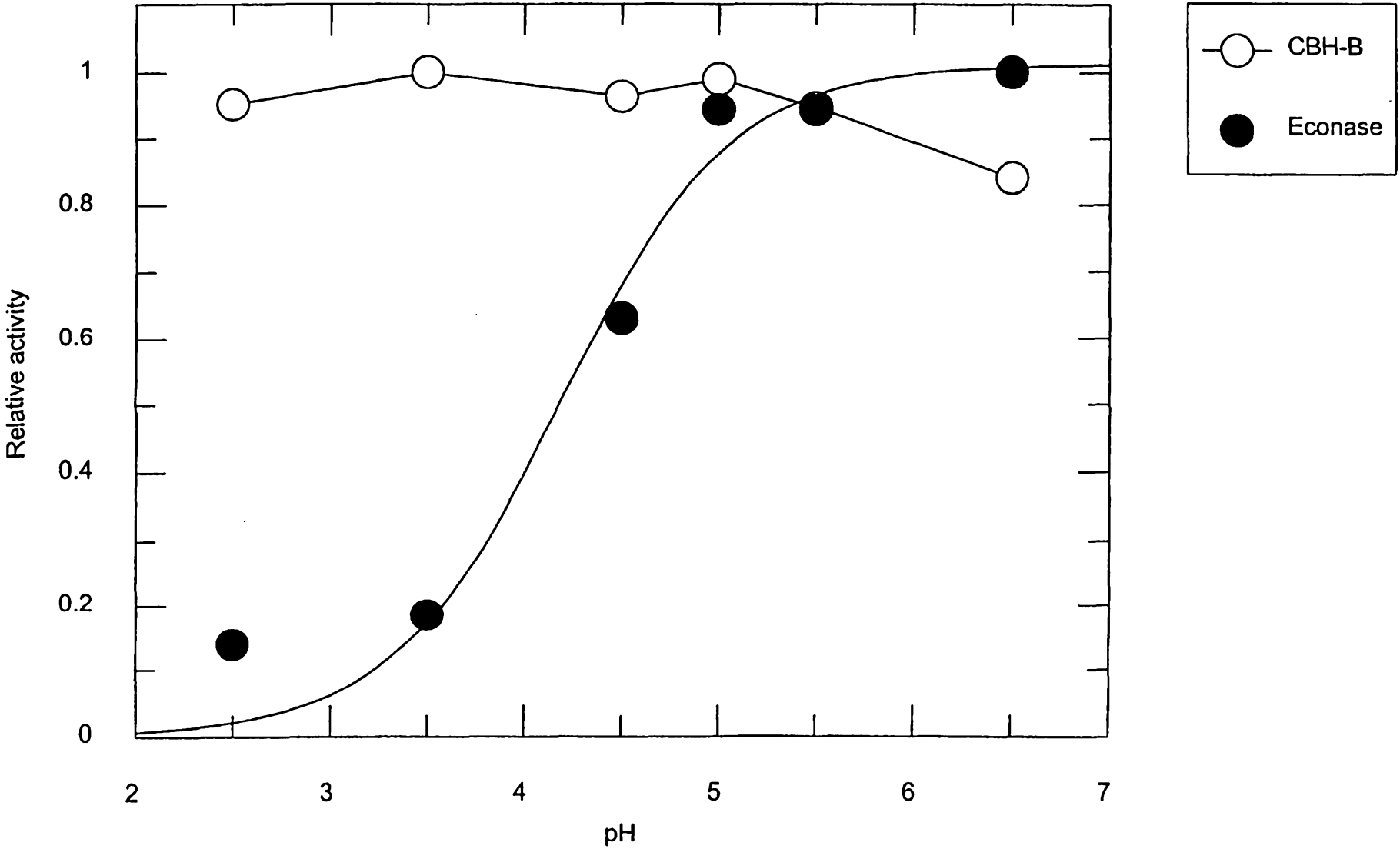


Figure 3 3/4

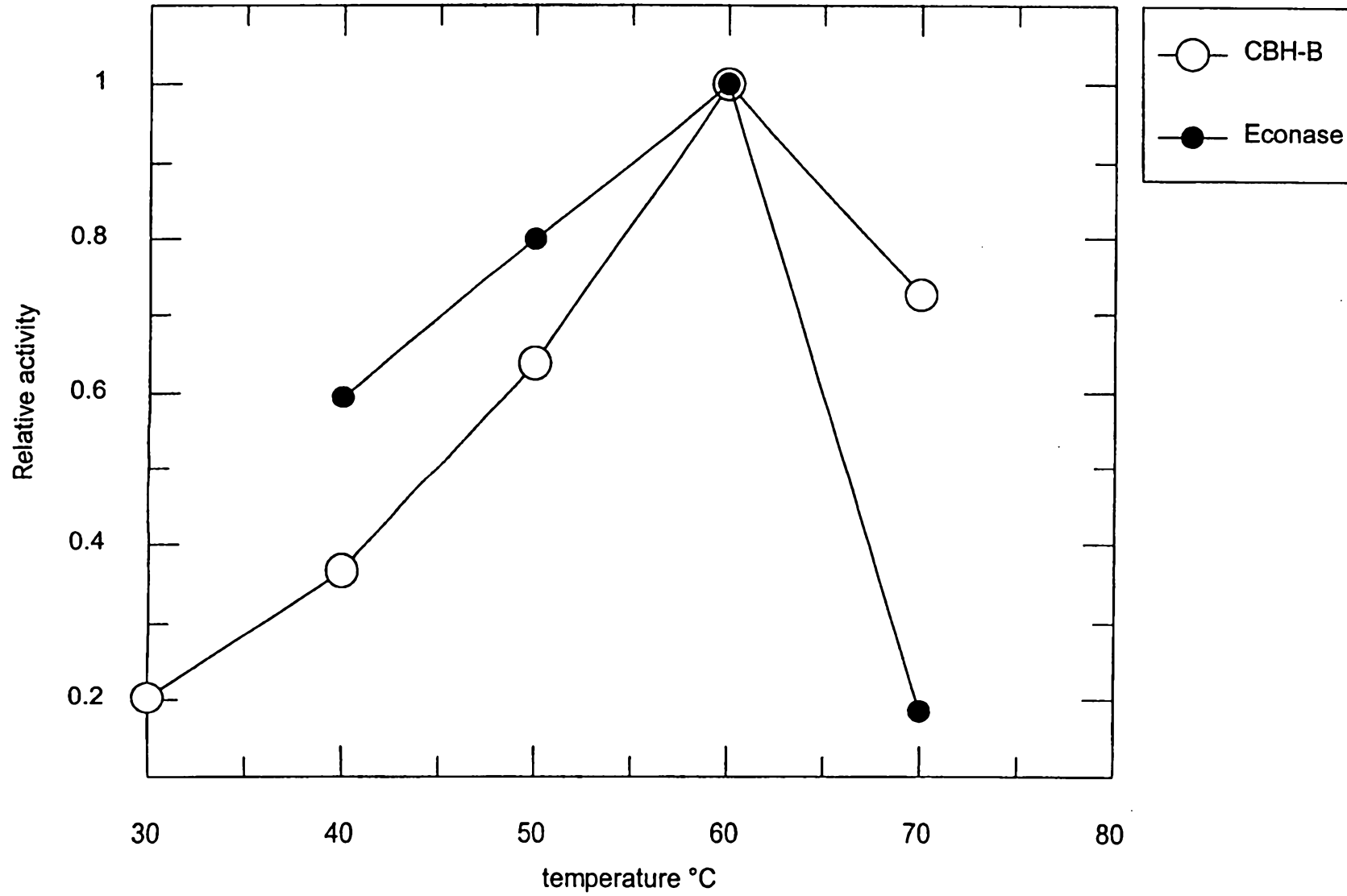




Figure 2 2/4

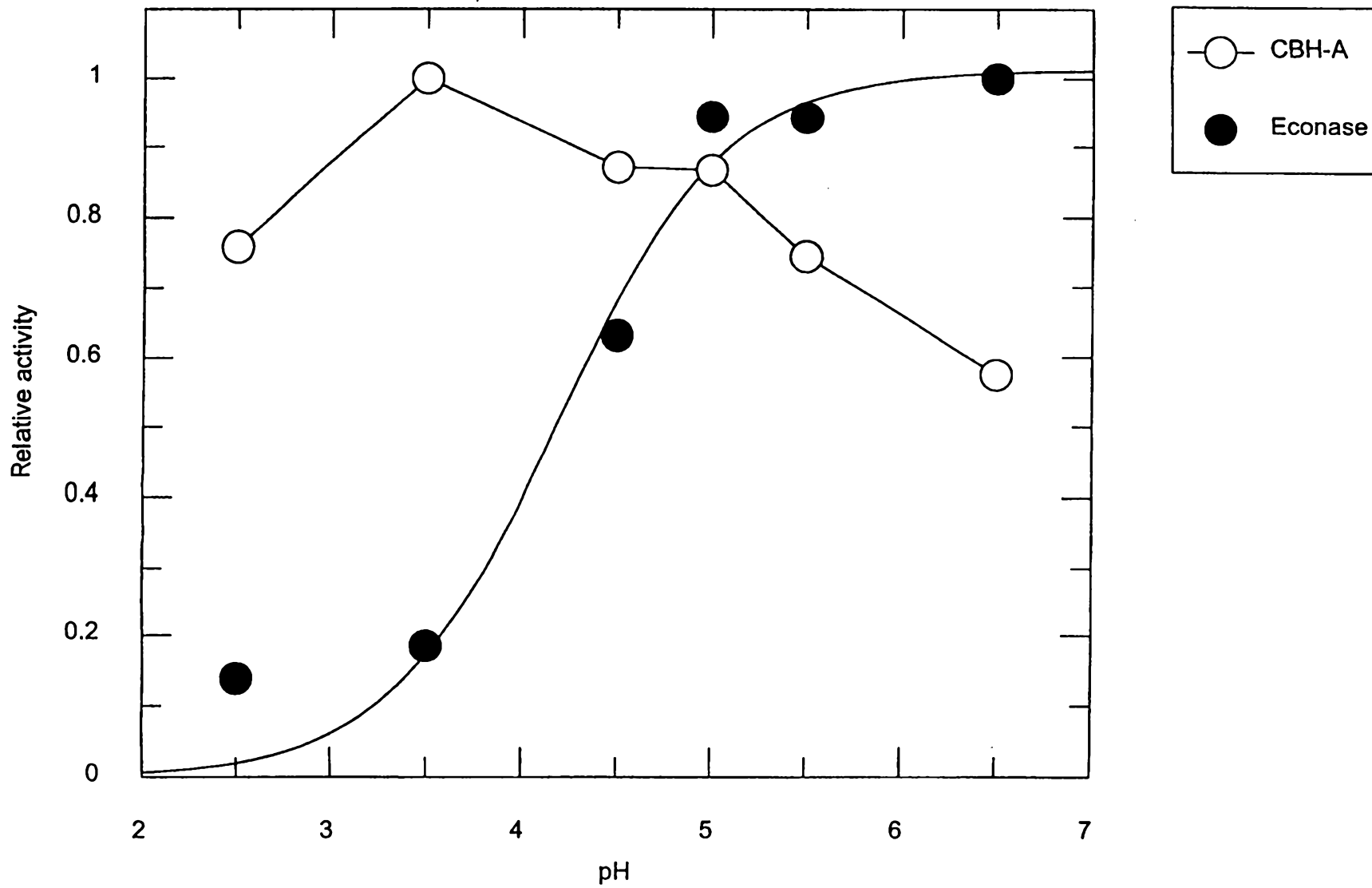


Figure 1 1/4

