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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 June 2008 (26.06.2008)

PCT

(10) International Publication Number
WO 2008/074884 A2

(51) International Patent Classification:
C12N 9/24 (2006.01) *A23K 1/165* (2006.01)

(21) International Application Number:
PCT/EP2007/064475

(22) International Filing Date:
21 December 2007 (21.12.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
06126761.3 21 December 2006 (21.12.2006) EP

(71) Applicant *(for all designated States except US)*: **VALORBEC S.E.C.** [CA/CA]; 550 Sherbrooke Street West, Suite 100, Montreal, Québec H3A 1B9 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **TSANG, Adrian** [CA/CA]; 3816 Draper Ave., Montréal, Québec H4A2P1 (CA). **POWLOWSKI, Justin** [CA/CA]; 4555 Montclair Ave., Montréal, Québec H4B2J8 (CA). **BUTLER, Gregory** [CA/CA]; 1700 René Lévesque Blvd. W., Suite 301, Montréal, Québec 3H32V1 (CA). **STORMS, Reginald** [CA/CA]; 97 Prince, Beaconsfield, Québec H9W3M8 (CA). **MOHRMAN, Lisette** [NL/NL]; van Ruysdaellaan 153, NL-2661 SH Bergschenhoek (NL). **MUTSAERS, Johanna, Henrica, Gerdina, Maria** [NL/NL]; Middenhaven 15, NL-2635 GT Den Hoorn (NL).

(74) Agents: **CHADWICK, Mark Craig** et al.; P.O. Box 1, NL-2600 MA Delft (NL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

WO 2008/074884 A2

(54) **Title:** NOVEL XYLANASE ENZYMES XYL001 AND XYL002 AND USES THEREOF

(57) **Abstract:** The present invention relates tonovel xylanasesXYL001and XYL002, a polynucleotide sequence encoding for the polypeptide according to the invention, a production process for the enzymes according to the invention and the use of the enzymes according to the invention in various industrial processes.

NOVEL XYLANASE ENZYMES XYL001 AND XYL002 AND USES THEREOF

Field of the invention

The invention relates to newly identified polynucleotide sequences comprising genes that encode novel xylanase. The enzymes may be isolated from *Gloeophyllum trabeum* or *Coprinus cinereus*. The invention features the full length coding sequence of the novel gene as well as the amino acid sequence of the full-length functional protein and functional equivalents of the gene or the amino acid sequence. The invention also relates to methods of using these proteins in industrial processes. Also included in the invention are cells transformed with a polynucleotide according to the invention suitable for producing these proteins and cells wherein a protein according to the invention is genetically modified to enhance or reduce its activity and/or level of expression.

Background of the invention

The present invention relates to a novel xylanase, suitable for use in several industrial applications, for example in food applications, such as for example cereal-based food products, in the detergent industry for removal of laundry stains; in the feed-enzyme industry such as for example increasing the digestability of nutrients; in the pulp and paper industry such as enhancing bleachability of the pulp; in the pharmaceutical industry such as antibacterial formulation; and in the bioethanol industry such as for example improving the ethanol yield.

Cereal-based food products such as pasta, noodles and bread can be prepared from a dough which is usually made from the basic ingredients (cereal) flour, water and optionally salt. Depending on the cereal-based food products, other ingredients added may be sugars, flavours etcetera. For leavened products, primarily baker's yeast is used next to chemical leavening systems such as a combination of an acid (generating compound) and bicarbonate.

In order to improve the handling properties of the dough and/or the final properties of the cereal-based food products, there is a continuous effort to develop processing aids with improved properties. Processing aids are defined herein as compounds that improve the handling properties of the dough and/or the final properties of the cereal-based food products. Dough properties that may be

improved comprise machineability, gas retaining capability, reduced stickiness, elasticity, extensibility, moldability etcetera. Properties of the cereal-based food products that may be improved comprise loaf volume, crust crispiness, crumb texture and softness, flavour-related staleness and shelf life. These dough and/or
5 baked product improving processing aids can be divided into two groups: chemical additives and enzymes.

As a result of a consumer-driven need to replace the chemical additives by more natural products, several enzymes have been developed with dough and/or cereal-based food product improving properties and which are used in all
10 possible combinations depending on the specific application conditions. Suitable enzymes include xylanase, starch degrading enzymes, oxidizing enzymes, fatty material splitting enzymes, protein degrading, modifying or crosslinking enzymes.

There is a continuous need to provide novel xylanases which can be used in the production of food products, for example cereal-based food products, and
15 other industrial applications, such as in the detergent industry, in the feed-enzyme industry, in the pulp and paper industry, in the pharmaceutical industry and in the bioethanol industry.

Object of the invention

20 It is an object of the invention to provide novel polynucleotides encoding novel xylanase. A further object is to provide naturally and recombinantly produced xylanase as well as recombinant strains producing these. Also fusion polypeptides are part of the invention as well as methods of making and using the polynucleotides and polypeptides according to the invention.

25

Summary of the invention

The invention provides for novel polynucleotides encoding novel xylanase.

In particular, the invention provides for polynucleotides having a
30 nucleotide sequence that hybridizes preferably under high stringent conditions to the complement of a sequence according to SEQ ID NO: 1 or SEQ ID NO: 3. Consequently, the invention provides nucleic acids that are at least 80%, preferably at least 85%, more preferably at least 90%, even more preferably at least 95%, 96%, 97%, 98% or 99% homologous to the sequences according to
35 SEQ ID NO: 1 or SEQ ID NO: 3.

In one embodiment the invention provides for such an isolated

polynucleotide obtainable from a brown rot fungus, in particular *Gloeophyllum* is preferred and even more preferred *Gloeophyllum trabeum*.

In another embodiment, the invention provides for an isolated polynucleotide comprising a nucleic acid sequence encoding a polypeptide with
5 an amino acid sequence as shown in SEQ ID NO: 2 or SEQ ID NO: 4 or functional equivalents of either thereof.

In a further embodiment, the invention provides an isolated polynucleotide encoding at least one functional domain of a polypeptide according to SEQ ID NO: 2 or SEQ ID NO: 4 or functional equivalents of either thereof.

10 In another embodiment the invention provides a xylanase gene according to SEQ ID NO: 1 or SEQ ID NO: 3 or variants or fragments of either thereof that are still coding for active enzyme.

The examples of activities of xylanases according to the invention are herein intended to at least cover any: 1,4(1,3;1,4)- β–D-glucan 4-
15 glucanogydrolase (EC 3.2.1.4) catalysing the endohydrolysis of 1,4-β-D-glucosidic linkages in cellulose, lichenin and cereal β-D-glucans or 1,4-β-D-xylan xylanohydrolase (EC 3.2.1.8) catalyzing the endohydrolysis of 1,4-β-D-xylosidic linkages in xylans.

The invention also relates to vectors comprising a polynucleotide
20 sequence according to the invention and primers, probes and fragments that may be used to amplify or detect the DNA according to the invention.

In a further preferred embodiment, a vector is provided wherein the polynucleotide sequence according to the invention is functionally linked with at least one regulatory sequence suitable for expression of the encoded amino acid
25 sequence in a suitable host cell, such as a filamentous fungus, for example *Aspergillus*. The invention also provides methods for preparing polynucleotides and vectors according to the invention.

The invention also relates to recombinantly produced host cells that contain heterologous or homologous polynucleotides according to the invention.

30 In another embodiment, the invention provides recombinant host cells wherein the expression of a xylanase according to the invention is significantly increased or wherein the activity of the xylanase is increased.

In another embodiment the invention provides for a recombinantly produced host cell that contains heterologous or homologous DNA according to
35 the invention and wherein the cell is capable of producing a functional xylanase

according to the invention, preferably a cell capable of over-expressing the xylanase according to the invention, for example an *Aspergillus niger* strain comprising an increased copy number of a gene according to the invention.

In yet another aspect of the invention, a purified polypeptide is provided. 5 The polypeptides according to the invention include the polypeptides encoded by the polynucleotides according to the invention. Especially preferred is a polypeptide according to SEQ ID NO: 2 or SEQ ID NO: 4 or functional equivalents of either thereof.

Fusion proteins comprising a polypeptide according to the invention are 10 also within the scope of the invention. The invention also provides methods of making the polypeptides according to the invention.

The invention also relates to the use of the xylanase according to the invention in any industrial process as described herein.

The terms XYL001 and XYL002 are used herein to refer to polypeptides 15 having the sequence of SEQ ID NO: 2 and SEQ ID NO: 4 respectively and their respective functional variants and fragments and any other amino acid sequences disclosed herein based on SEQ ID NO: 2 and SEQ ID NO: 4 (i.e. to the products encoded by the nucleic acid sequences of SEQ ID NO: 1 and SEQ ID NO: 3 and their respective functional variants and fragments and any other nucleic acid 20 sequences disclosed herein based on SEQ ID NO: 1 and SEQ ID NO: 3). However, the terms XYL001 and XYL002 may also be used to refer to nucleic acid sequences of SEQ ID NO: 1 and SEQ ID NO: 3 respectively and their respective functional variants and fragments and any other nucleic acid sequences disclosed herein based on SEQ ID NO: 1 and SEQ ID NO: 3. The 25 meaning of the term will be apparent to the skilled person depending on the context in which it is used.

Brief description of the drawings

Figure 1 shows the pH profile at 55°C and temperature profile at pH3.5 for 30 XYL001 on 1% azo-xylan. Graphs show data after subtraction of VTO (vector transformant supernatant) results.

Figure 2 shows the pH profile at 55°C and temperature profile at pH6 for XYL002 on 1% azo-xylan. Graphs show data after subtraction of VTO (vector transformant supernatant) results.

35 Figure 3 shows the specific activity for XLY001 on various substrates.

Figure 4 shows the specific activity for XLY002 on various substrates.

Figure 5 shows TLC analyses of reaction product using 4 different sources of xylan. The encircled lanes represent the products of XYL001. Standards are shown at far left.

5 Figure 6 shows TLC analyses of reaction product using 4 different sources of xylan. The encircled lanes represent the products of XYL002. Standards are shown at far left.

In the drawings, Gtra5285 refers to XYL001 and Ccin16583 refers to XYL002.

10

Detailed description of the invention

Throughout the present specification and the accompanying claims the words "comprise"and"include"and variations such as "comprises", "comprising", "includes" and "including" are to be interpreted inclusively. That is, these words 15 are intended to convey the possible inclusion of other elements or integers not specifically recited, where the context allows.

Polynucleotides

The present invention provides polynucleotides encoding a xylanase, 20 tentatively called XYL001, having an amino acid sequence according to SEQ ID NO: 2, a xylanase, tentatively called XYL002, having an amino acid sequence according to SEQ ID NO: 4 and functional equivalents of either thereof.

The sequence of the gene encoding XYL001 was determined by sequencing a cDNA clone obtained from *Gloeophyllum trabeum*. The sequence of 25 the gene encoding XYL002 was determined by sequencing a cDNA clone obtained from *Coprinus cinereus*. The invention provides polynucleotide sequences comprising the gene encoding the XYL001 xylanase or the XYL002 xylanase as well as their coding sequences. Accordingly, the invention relates to an isolated polynucleotide comprising the nucleotide sequence according to SEQ 30 ID NO: 1 or SEQ ID NO: 3 and to functional equivalents of either thereof.

In particular, the invention relates to an isolated polynucleotide hybridizable under stringent conditions, preferably under high stringent conditions, to the reverse complement of a polynucleotide according to SEQ ID NO: 1 or SEQ ID NO: 3. Advantageously, such isolated polynucleotide may be obtained from 35 brown rot fungi, in particular from *Gloeophyllacea*, such as *Gloeophyllum*, for

example *trabeum, abietinum, odoratum, sepiarium*, preferably from *Gloeophyllum trabeum*. Such an isolated polynucleotide may also be obtained from a coprinoid fungus, in particular *Coprinus, such as cinereus, calyptratus, comatus, spadiveisporus, sterquilinus, preferably from Coprinus cinereus.* More specifically,
5 the invention relates to an isolated polynucleotide having a nucleotide sequence according to SEQ ID NO: 1 or SEQ ID NO: 3.

The invention also relates to an isolated polynucleotide encoding at least one functional domain of a polypeptide according to SEQ ID NO: 2 or SEQ ID NO: 4 or to functional equivalents of either thereof.

10 As used herein, the terms “gene” and “recombinant gene” refer to nucleic acid molecules which may be isolated from chromosomal DNA, which include an open reading frame encoding a protein, e.g. a *Gloeophyllum trabeum* xylanase according to the present invention. A gene may include coding sequences, non-coding sequences, introns and regulatory sequences. Moreover, a gene refers to
15 an isolated nucleic acid molecule as defined herein.

A nucleic acid molecule of the present invention, such as a nucleic acid molecule having the nucleotide sequence of SEQ ID NO: 1 or 3 or a functional equivalent thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, using all or portion of
20 the nucleic acid sequence of SEQ ID NO: 1 or 3 as a hybridization probe, nucleic acid molecules according to the invention can be isolated using standard hybridization and cloning techniques (e. g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual.2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring
25 Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO: 1 or 3 can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence information contained in SEQ ID NO: 1 or 3.

30 A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis.

35 Furthermore, oligonucleotides corresponding to or hybridizable to

nucleotide sequences according to the invention can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO: 1 or 3. The
5 sequence of SEQ ID NO: 1 corresponds to the coding region of the *Gloeophyllum trabeum* xylanase cDNA. This cDNA comprises sequences encoding the *Gloeophyllum trabeum* XYL001 polypeptide according to SEQ ID NO: 2. The sequence of SEQ ID NO: 3 corresponds to the coding region of the *Coprinus cinereus* xylanase cDNA. This cDNA comprises sequences encoding the
10 *Coprinus cinereus* XYL002 polypeptide according to SEQ ID NO: 4.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a reverse complement of the nucleotide sequence shown in SEQ ID NO: 1 or 3 or a functional equivalent of these nucleotide sequences.

15 A nucleic acid molecule which is complementary to another nucleotide sequence is one which is sufficiently complementary to the other nucleotide sequence such that it can hybridize to the other nucleotide sequence thereby forming a stable duplex.

One aspect of the invention pertains to isolated nucleic acid molecules
20 that encode a polypeptide of the invention or a functional equivalent thereof such as a biologically active fragment or domain, as well as nucleic acid molecules sufficient for use as hybridization probes to identify nucleic acid molecules encoding a polypeptide of the invention and fragments of such nucleic acid molecules suitable for use as PCR primers for the amplification or mutation of
25 nucleic acid molecules.

An “isolated polynucleotide” or “isolated nucleic acid” is a DNA or RNA that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one on the 5’ end and one on the 3’ end) in the naturally occurring genome of the organism from which it is derived. Thus, in one
30 embodiment, an isolated nucleic acid includes some or all of the 5’ non-coding (e.g., promotor) sequences that are immediately contiguous to the coding sequence. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate
35 molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or

restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA that is part of a hybrid gene encoding an additional polypeptide that is substantially free of cellular material, viral material, or culture medium (when produced by recombinant DNA techniques), or chemical
5 precursors or other chemicals (when chemically synthesized). Moreover, an "isolated nucleic acid fragment" is a nucleic acid fragment that is not naturally occurring as a fragment and would not be found in the natural state.

As used herein, the terms "polynucleotide" or "nucleic acid molecule" are intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA
10 molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. The nucleic acid may be synthesized using oligonucleotide analogs or derivatives (e.g., inosine or phosphorothioate nucleotides). Such oligonucleotides can be used, for example,
15 to prepare nucleic acids that have altered base-pairing abilities or increased resistance to nucleases.

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a XYL001 or XYL002 nucleic acid molecule, e.g., the coding strand of a XYL001 of XYL002 nucleic acid molecule. In a further
20 embodiement, it is also provided an antisense molecule which hybridizes with at least 10 contiguous, 20 contiguous, 40 contiguous, more preferably 50 contiguous, 60 contiguous, at least 80 contiguous, more preferably 100 contiguous nucleotides or the entire length of XYL001 or XYL002 nucleic acid molecule, e.g., the coding strand of a XYL001 or XYL002 nucleic acid molecule.
25 Also included within the scope of the invention are the complement, in particular the reverse complement, strands of the nucleic acid molecules described herein.

Sequencing errors

The sequence information as provided herein should not be so narrowly
30 construed as to require inclusion of erroneously identified bases. The specific sequences disclosed herein can be readily used to isolate the complete gene from filamentous fungi, in particular *Gloeophyllum trabeum* or *Coprinus cinereus* which in turn can easily be subjected to further sequence analyses thereby identifying sequencing errors.
35 Unless otherwise indicated, all nucleotide sequences determined by

sequencing a DNA molecule herein were determined using an automated DNA sequencer and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence
5 determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other
10 approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino
15 acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

The person skilled in the art is capable of identifying such erroneously identified bases and knows how to correct for such errors.

20 **Nucleic acid fragments, probes and primers**

A nucleic acid molecule according to the invention may comprise only a portion or a fragment of the nucleic acid sequence shown in SEQ ID NO: 1 or 3, for example a fragment which can be used as a probe or primer or a fragment encoding a portion of a XYL001 or XYL002 protein. The nucleotide sequence
25 determined from the cloning of the XYL001 or XYL002 gene and cDNA allows for the generation of probes and primers designed for use in identifying and/or cloning other XYL001 or XYL002 family members, as well as XYL001 or XYL002 homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide which typically comprises a region of
30 nucleotide sequence that hybridizes preferably under highly stringent conditions to at least about 12 or 15, preferably about 18 or 20, preferably about 22 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 or more consecutive nucleotides of a nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO: 3 or of a functional equivalent of either thereof.

35 Probes based on the XYL001 or XYL002 nucleotide sequences can be

used to detect transcripts or genomic XYL001 or XYL002 sequences encoding the same or homologous proteins for instance in other organisms. In preferred embodiments, the probe further comprises a label group attached thereto, e.g., the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme cofactor. Such probes can also be used as part of a diagnostic test kit for identifying cells which express a XYL001 or XYL002 protein.

Identity & homology

The terms “homology” or “percent identity” or “percentage identity” or “% identity” are used interchangeably herein. For the purpose of this invention, it is defined here that in order to determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). Such alignment may be carried out over the full length of the sequences being compared. Alternatively, the alignment may be carried out over a shorter length, for example about 20, about 50, about 100 or more nucleic acids/bases or amino acids. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = number of identical positions/total number of positions (i.e. overlapping positions) x 100). Preferably, the two sequences are the same length.

The skilled person will be aware of the fact that several different computer programs are available to determine the homology between two sequences. For instance, a comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (J. Mol. Biol. (48): 444-453 (1970)) algorithm which has been incorporated into the GAP program in the Accelrys GCG software package (available at http://www.accelrys.com/products/gcg/), using either a Blossom 62 matrix or a

PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

In yet another embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the Accelrys GCG software package (available at http://www.accelrys.com/products/gcg/), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity two amino acid or nucleotide sequence is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989) which has been incorporated into the ALIGN program (version 2.0) (available at the ALIGN Query using sequence data of the Genestream server IGH Montpellier France http://vega.igh.cnrs.fr/bin/align-guess.cgi) using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a “query sequence” to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) J. Mol. Biol. 215:403—10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to XYL001 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to XYL001 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) Nucleic Acids Res. 25(17): 3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See the homepage of the National Center for Biotechnology Information at http://www.ncbi.nlm.nih.gov/.

Hybridization

As used herein, the term "hybridizing" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least about

50%, at least about 60%, at least about 70%, at least about 80%, more preferably at least about 85%, even more preferably at least about 90%, more preferably at least 95%, more preferably at least 98% or more preferably at least about 99% homologous to each other typically remain hybridized to each other. That is to say, such hybridizing sequences may share at least about 50%, at least about 60%, at least about 70%, at least about 80%, more preferably at least about 85%, even more preferably at least about 90%, more preferably at least 95%, more preferably at least 98% or more preferably at least about 99% sequence identity.

A preferred, non-limiting example of such hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45 °C, followed by one or more washes in 1 X SSC, 0.1% SDS at 50°C, preferably at 55°C, preferably at 60°C and even more preferably at 65°C.

Highly stringent conditions include, for example, hybridizing at 68°C in 5x SSC/5x Denhardt's solution / 1.0% SDS and washing in 0.2x SSC/0.1% SDS at room temperature. Alternatively, washing may be performed at 42°C.

The skilled artisan will know which conditions to apply for stringent and highly stringent hybridization conditions. Additional guidance regarding such conditions is readily available in the art, for example, in Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; and Ausubel et al. (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, N.Y.).

Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3' terminal poly(A) tract of mRNAs), or to a complementary stretch of T (or U) resides, would not be included in a polynucleotide of the invention used to specifically hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-standed cDNA clone).

Obtaining full length DNA from other organisms

In a typical approach, cDNA libraries constructed from other organisms, e.g. brown-rot fungi, in particular from the micro-organism family Gloeophyllacea, for example *Gloeophyllum* can be screened such as abietinum, odoratum, sepiarium or eg. a coprinoid fungus, in particular from the microorganism family *Agaricaceae*, such from the genus *Coprinus,* such as *calyptratus, comatus,*

spadiveisporus or *sterquilinus*.

For example, *Gloeophyllum* or *Agaricaceae* strains can be screened for homologous XYL001 or XYL002 polynucleotides respectively by Northern blot analysis. Upon detection of transcripts homologous to polynucleotides according to the invention, cDNA libraries can be constructed from RNA isolated from the appropriate strain, utilizing standard techniques well known to those of skill in the art. Alternatively, a total genomic DNA library can be screened using a probe hybridizable to a XYL001 or XYL002 polynucleotide according to the invention.

Homologous gene sequences can be isolated, for example, by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of nucleotide sequences as taught herein.

The template for the reaction can be cDNA obtained by reverse transcription of mRNA prepared from strains known or suspected to express a polynucleotide according to the invention. The PCR product can be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a new XYL001 or XYL002 nucleic acid sequence, or a functional equivalent thereof.

The PCR fragment can then be used to isolate a full-length cDNA clone by a variety of known methods. For example, the amplified fragment can be labeled and used to screen a bacteriophage or cosmid cDNA library. Alternatively, the labeled fragment can be used to screen a genomic library.

PCR technology also can be used to isolate full-length cDNA sequences from other organisms. For example, RNA can be isolated, following standard procedures, from an appropriate cellular or tissue source. A reverse transcription reaction can be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis.

The resulting RNA/DNA hybrid can then be "tailed" (e.g., with guanines) using a standard terminal transferase reaction, the hybrid can be digested with RNase H, and second strand synthesis can then be primed (e.g., with a poly-C primer). Thus, cDNA sequences upstream of the amplified fragment can easily be isolated. For a review of useful cloning strategies, see e.g.,Sambrook et al., supra; and Ausubel et al., supra.

Vectors

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a XYL001 or XYL002 protein or a

functional equivalent thereof. As used herein, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a “plasmid”, which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as “expression vectors”. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. The terms “plasmid” and “vector” can be used interchangeably herein as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vector includes one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, “operatively linked” is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term “regulatory sequence” is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signal). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in a certain host cell (e.g. tissue-specific regulatory sequences). It will be appreciated

by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, encoded by 5 nucleic acids as described herein (e.g. XYL001 or XYL002 proteins, mutant forms of XYL001 or XYL002 proteins, fragments, variants or functional equivalents thereof, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of XYL001 or XYL002 proteins in prokaryotic or eukaryotic cells. For 10 example, XYL001 or XYL002 proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be 15 transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression vectors useful in the present invention include chromosomal-, episomal- and virus-derived vectors e.g., vectors derived from bacterial plasmids, bacteriophage, yeast episome, yeast chromosomal elements, viruses such as 20 baculoviruses, papova viruses, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids.

The DNA insert should be operatively linked to an appropriate promoter, 25 such as the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled person. In a specific embodiment, promoters are preferred that are capable of directing a high expression level of xylanase in brown rot or coprinoid fungi. Such promoters are 30 known in the art. The expression constructs may contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be 35 translated.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host
5 cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, transduction, infection, lipofection, cationic lipidmediated transfection or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual, 2nd,ed. Cold Spring Harbor Laboratory,* Cold Spring Harbor
10 Laboratory Press, Cold Spring Harbor, NY, 1989), Davis et al., *Basic Methods in Molecular Biology* (1986) and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and
15 select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methatrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that
20 encoding XYL001 or XYL002 proteins or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g. cells that have incorporated the selectable marker gene will survive, while the other cells die).

Expression of proteins in prokaryotes is often carried out in *E. coli* with
25 vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, e.g. to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant
30 protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein.

35 As indicated, the expression vectors will preferably contain selectable

markers. Such markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracyline or ampicillin resistance for culturing in *E. coli* and other bacteria. Representative examples of appropriate host include bacterial cells, such as *E. coli*, *Streptomyces Salmonella typhimurium* and certain
5 *Bacillus* species; fungal cells such as *Aspergillus* species, for example *A. niger, A. oryzae* and *A. nidulans*, yeast cells such as *Kluyveromyces*, for example *K. lactis* and/or *Pichia,* for example *P. pastoris*; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO, COS and *Bowes melanoma*; and plant cells. Appropriate culture mediums and conditions for the above-described
10 host cells are known in the art.

Vectors preferred for use in bacteria are for example disclosed in WO-A1-2004/074468, which are hereby enclosed by reference. Other suitable vectors will be readily apparent to the skilled artisan.

Known bacterial promotors suitable for use in the present invention
15 include the promoters disclosed in WO-A1-2004/074468, which are hereby enclosed by reference.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually
20 about from 10 to 300 bp that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

25 For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signal may be incorporated into the expressed polypeptide. The signals may be endogenous to the polypeptide or they may be heterologous signals.

30 The XYL001 polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification
35 or during subsequent handling and storage. Also, peptide moieties may be added

to the polypeptide to facilitate purification.

Polypeptides according to the invention

The invention provides an isolated polypeptide having the amino acid sequence according to SEQ ID NO: 2 or 4, and an amino acid sequence obtainable by expressing the polynucleotide of SEQ ID NO: 1 or 3 in an appropriate host. Also, a peptide or polypeptide comprising a functional equivalent of the above polypeptides is comprised within the present invention. The above polypeptides are collectively comprised in the term “polypeptides according to the invention”

The terms "peptide" and "oligopeptide" are considered synonymous (as is commonly recognized) and each term can be used interchangeably as the context required to indicate a chain of at least two amino acids coupled by peptidyl linkages. The word "polypeptide" is used herein for chains containing more than seven amino acid residues. All oligopeptide and polypeptide formulas or sequences herein are written from left to right and in the direction from amino terminus to carboxyl terminus. The one-letter code of amino acids used herein is commonly known in the art and can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual, 2nd,ed. Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989)

By "isolated" polypeptide or protein is intended a polypeptide or protein removed from its native environment. For example, recombinantly produced polypeptides and proteins expressed in host cells are considered isolated for the purpose of the invention as are native or recombinant polypeptides which have been substantially purified by any suitable technique such as, for example, the single-step purification method disclosed in Smith and Johnson, Gene 67:31-40 (1988).

The XYL001 or XYL002 xylanase according to the invention can be recovered and purified from recombinant cell cultures by methods known in the art. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending

upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

5

Protein fragments

The invention also features biologically active fragments of the polypeptides according to the invention.

Biologically active fragments of a polypeptide of the invention include 10 polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the XYL001 or 2 protein (e.g., the amino acid sequence of SEQ ID NO: 2 or 4 respectively), which include fewer amino acids than the full length protein but which exhibit at least one biological activity of the corresponding full-length protein. Typically, biologically active fragments comprise 15 a domain or motif with at least one activity of the XYL001 or XYL002 protein. A biologically active fragment of a protein of the invention can be a polypeptide which is, for example, about 10, about 25, about 50 or about 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and 20 evaluated for one or more of the biological activities of the native form of a polypeptide of the invention.

The invention also features nucleic acid fragments which encode the above biologically active fragments of the XYL001 or XYL002 protein.

25 **Fusion proteins**

The proteins of the present invention or functional equivalents thereof, e.g., biologically active portions thereof, can be operatively linked to a non-XYL001 or XYL002 polypeptide (e.g., heterologous amino acid sequences) to form fusion proteins. A "non-XYL001 polypeptide" or "non-XYL002 polypeptide" 30 refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the XYL001 or XYL002 protein repsectively. Such “non-XYL001 or non-XYL002 polypeptides” can be derived from the same or a different organism. Within a XYL001 or XYL002 fusion protein the XYL001 or XYL002 polypeptide can correspond to all or a biologically active 35 fragment of a XYL001 or XYL002 protein. In a preferred embodiment, a XYL001

or XYL002 fusion protein comprises at least two biologically active portions of a XYL001 or XYL002 protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the XYL001 or XYL002 polypeptide and the non-XYL001 or XYL002 polypeptide are fused in-frame to each other. The non-
5 XYL001 or non-XYL002 polypeptide can be fused to the N-terminus or C-terminus of the XYL001 or XYL002 polypeptide.

For example, in one embodiment, the fusion protein is a GST-XYL001 or GST-XYL002 fusion protein in which the XYL001 or XYL002 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the
10 purification of recombinant XYL001 or XYL002. In another embodiment, the fusion protein is a XYL001 or XYL002 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian and yeast host cells), expression and/or secretion of XYL001 or XYL002 can be increased through use of a hetereologous signal sequence.

15 In another example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (*Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley & Sons, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase
20 (Stratagene; La Jolla, California). In yet another example, useful prokarytic heterologous signal sequences include the phoA secretory signal (Sambrook et al., *supra*) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey).

A signal sequence can be used to facilitate secretion and isolation of a
25 protein or polypeptide of the invention. Signal sequences are typically characterized by a core of hydrophobic amino acids, which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal peptides contain processing sites that allow cleavage of the signal sequence from the mature proteins as they pass through the secretory pathway.
30 The signal sequence directs secretion of the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by known methods. Alternatively, the signal sequence can be linked to the protein of interest using a sequence, which
35 facilitates purification, such as with a GST domain. Thus, for instance, the

sequence encoding the polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide, which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, such as the tag provided in a pQE
5 vector (Qiagen, Inc.), among others, many of which are commercially available. As described in *Gentz et al, Proc. Natl. Acad. Sci. USA 86*:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. The HA tag is another peptide useful for purification which corresponds to an epitope derived of influenza hemaglutinin protein, which has been described by
10 Wilson *et al., Cell 37*:767 (1984), for instance.

Preferably, a XYL001 or XYL002 fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing
15 blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR
20 amplification of gene fragments can be carried out using anchor primers, which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are
25 commercially available that already encode a fusion moiety (e.g, a GST polypeptide). A XYL001- or XYL002-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the XYL001 or XYL002 protein.

30 **Functional equivalents**

The terms “functional equivalent(s)” and “functional variant(s)” are used interchangeably herein. Functional equivalents of XYL001 or XYL002 DNA are isolated DNA fragments that encode a polypeptide that exhibits a particular function or property of the XYL001 *Gloeophyllum trabeum* xylanase or the
35 XYL002 *Coprinus cinereus* xylanase as defined herein. A functional equivalent of

a XYL001 or XYL002 polypeptide according to the invention is a polypeptide that exhibits at least one function of a *Gloeophyllum trabeum or of a Coprinus cinereus* xylanase as defined herein. Functional equivalents therefore also encompass biologically active fragments.

5 Functional protein or polypeptide equivalents may contain only conservative substitutions of one or more amino acids of SEQ ID NO: 2 or 4 or substitutions, insertions or deletions of non-essential amino acids. Accordingly, a non-essential amino acid is a residue that can be altered in SEQ ID NO: 2 or 4 without substantially altering the biological function. For example, amino acid 10 residues that are conserved among the XYL001 or XYL002 proteins of the present invention are predicted to be particularly unamenable to alteration. Furthermore, amino acids conserved among the XYL001 or XYL002 proteins according to the present invention and other xylanases are not likely to be amenable to alteration.

The term “conservative substitution” is intended to indicate a substitution 15 in which the amino acid residue is replaced with an amino acid residue having a similar side chain. These families are known in the art and include amino acids with basic side chains (e.g.lysine, arginine and hystidine), acidic side chains (e.g. aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagines, glutamine, serine, threonine, tyrosine, cysteine), non-polar side 20 chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine tryptophan, histidine).

Functional nucleic acid equivalents may typically contain silent mutations 25 or mutations that do not alter the biological function of encoded polypeptide. Accordingly, the invention provides nucleic acid molecules encoding XYL001 or XYL002 proteins that contain changes in amino acid residues that are not essential for a particular biological activity. Such XYL001 or XYL002 proteins differ in amino acid sequence from SEQ ID NO: 2 or 4 yet retain at least one biological 30 activity thereof. In one embodiment an isolated nucleic acid molecule of the invention comprises a nucleotide sequence encoding a protein, wherein the protein comprises a substantially homologous amino acid sequence of at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 72%, at least abot 75%, at least about 80%, at least 35 about 85%, at least about 90%, at least about 95%, at least about 96%, at least

about 97%, at least about 98%, at least about 99% or more homologous to the amino acid sequence shown in SEQ ID NO: 2 or 4.

The invention thus provides a polypeptide which shares at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 72%, at least abot 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% or more sequence identity with the amino acid sequence shown in SEQ ID NO: 2 or 4.

For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J.U. *et al.*, Science 247:1306-1310 (1990) and the references cited therein. As the authors state, these studies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which changes are likely to be permissive at a certain position of the protein.

An isolated nucleic acid molecule encoding a XYL001 or a XYL002 protein homologous to the protein according to SEQ ID NO: 2 or 4 respectively can be created by introducing one or more nucleotide substitutions, additions or deletions into the coding nucleotide sequences according to SEQ ID NO: 1 or 3 such that one or more amino acid substitutions, deletions or insertions are introduced into the encoded protein. Such mutations may be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis.

The term “functional equivalents” also encompasses orthologues of the *Gloeophyllum trabeum* XYL001 protein and of the *Coprinus cinereus* XYL002 protein. Orthologues of the *Gloeophyllum trabeum* XYL001 protein or of the *Coprinus cinereus* XYL002 protein are proteins that can be isolated from other strains or species and possess a similar or identical biological activity. Such orthologues can readily be identified as comprising an amino acid sequence that is substantially homologous to SEQ ID NO: 2 or 4 respectively.

As defined herein, the term “substantially homologous” refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., with similar side chain) amino acids or nucleotides to a second amino acid or nucleotide sequence such that the first and the second amino acid or nucleotide sequences have a common domain. For example, amino acid or nucleotide sequences which contain a common domain

having at least about 50%, such as at least about 55%, for example at least about 60%, at least about 65%, at least about 70%, at least about 72%, preferably at least about 75%, more preferably at least about 80%, even more preferably at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identity or more are defined herein as sufficiently identical.

Also, nucleic acids encoding other XYL001 or XYL002 family members, which thus have a nucleotide sequence that differs from SEQ ID NO: 1 or 3, are within the scope of the invention. Moreover, nucleic acids encoding XYL001 or XYL002 proteins from different species which can have a nucleotide sequence which differs from SEQ ID NO: 1 are within the scope of the invention.

Nucleic acid molecules corresponding to variants (e.g. natural allelic variants) and homologues of the XYL001 or XYL002 DNA of the invention can be isolated based on their homology to the XYL001 or XYL002 nucleic acids disclosed herein using the cDNAs disclosed herein or a suitable fragment thereof, as a hybridization probe according to standard hybridization techniques preferably under highly stringent hybridization conditions.

In addition to naturally occurring allelic variants of the XYL001 or XYL002 sequence, the skilled person will recognise that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO: 1 or 3 thereby leading to changes in the amino acid sequence of the XYL001 or XYL002 protein respectively without substantially altering the function of the XYL001 protein or XYL002 protein.

In another aspect of the invention, improved XYL001 or XYL002 proteins are provided. Improved XYL001 or XYL002 proteins are proteins wherein at least one biological activity is improved. Such proteins may be obtained by randomly introducing mutations along all or part of the XYL001 or XYL002 coding sequence, such as by saturation mutagenesis, and the resulting mutants can be expressed recombinantly and screened for biological activity. For instance, the art provides for standard assays for measuring the enzymatic activity of xylanase and thus improved proteins may easily be selected.

In a preferred embodiment the XYL001 or XYL002 protein has an amino acid sequence according to SEQ ID NO: 2. In another embodiment, the XYL001 or XYL002 polypeptide is substantially homologous to the amino acid sequence according to SEQ ID NO: 2 or 4 and retains at least one biological activity of a

polypeptide according to SEQ ID NO: 2 or 4, yet differs in amino acid sequence due to natural variation or mutagenesis as described above.

In a further preferred embodiment, the XYL001 or XYL002 protein has an amino acid sequence encoded by an isolated nucleic acid fragment capable of hybridizing to a nucleic acid according to SEQ ID NO: 1 or 3, preferably under highly stringent hybridization conditions.

Accordingly, the XYL001 or XYL002 protein is preferably a protein which comprises an amino acid sequence at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 60%, at least about 72%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99% or more homologous to (i.e. sharing identity with) the amino acid sequence shown in SEQ ID NO: 2 or 4 and retains at least one functional activity of the polypeptide according to SEQ ID NO: 2 or 4.

Functional equivalents of a protein according to the invention can also be identified e.g. by screening combinatorial libraries of mutants, e.g. truncation mutants, of the protein of the invention for xylanase activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level. A variegated library of variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g. for phage display). There are a variety of methods that can be used to produce libraries of potential variants of the polypeptides of the invention from a degenerate oligonucleotide sequence. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al. (1984) Science 198:1056; Ike et al. (1983) Nucleic Acid Res. 11:477).

In addition, libraries of fragments of the coding sequence of a polypeptide of the invention can be used to generate a variegated population of polypeptides for screening a subsequent selection of variants. For example, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of the coding sequence of interest with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can

include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal
5 fragments of various sizes of the protein of interest.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations of truncation, and for screening cDNA libraries for gene products having a selected property. The most widely used techniques, which are amenable to high through-put analysis, for screening
10 large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a
15 technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of a protein of the invention (Arkin and Yourvan (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3): 327-331).

In addition to the XYL001 or XYL002 gene sequence shown in SEQ ID
20 NO: 1 or 3, it will be apparent for the person skilled in the art that DNA sequence polymorphisms may exist within a given population, which may lead to changes in the amino acid sequence of the XYL001 or XYL002 protein. Such genetic polymorphisms may exist in cells from different populations or within a population due to natural allelic variation. Allelic variants may also include functional
25 equivalents.

Fragments of a polynucleotide according to the invention may also comprise polynucleotides not encoding functional polypeptides. Such polynucleotides may function as probes or primers for a PCR reaction.

Nucleic acids according to the invention irrespective of whether they
30 encode functional or non-functional polypeptides can be used as hybridization probes or polymerase chain reaction (PCR) primers. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having a XYL001 or XYL002 activity include, inter alia, (1) isolating the gene encoding the XYL001 or XYL002 protein, or allelic variants thereof from a cDNA library e.g.
35 from an organism other than *Gloeophyllum trabeum* or *Coprinus cinereus*; (2) in

situ hybridization (e.g. FISH) to metaphase chromosomal spreads to provide precise chromosomal location of the XYL001 or XYL002 gene as described in Verma et al., Human Chromosomes: a Manual of Basic Techniques, Pergamon Press, New York (1988); (3) Northern blot analysis for detecting expression of XYL001 mRNA in specific tissues and/or cells and 4) probes and primers that can be used as a diagnostic tool to analyse the presence of a nucleic acid hybridizable to the XYL001 probe in a given biological (e.g. tissue) sample.

Also encompassed by the invention is a method of obtaining a functional equivalent of a XYL001 or XYL002 gene. Such a method entails obtaining a labelled probe that includes an isolated nucleic acid which encodes all or a portion of the protein sequence according to SEQ ID NO: 2 or 4 or a variant thereof; screening a nucleic acid fragment library with the labelled probe under conditions that allow hybridization of the probe to nucleic acid fragments in the library, thereby forming nucleic acid duplexes, and preparing a full-length gene sequence from the nucleic acid fragments in any labelled duplex to obtain a gene related to the XYL001 or XYL002 gene.

In one embodiment, a XYL001 or XYL002 nucleic acid of the invention is at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or more homologous to (i.e. sharing identity with) a nucleic acid sequence as shown in SEQ ID NO: 1 or 3 or to the reverse complement thereof.

The XYL001 polypeptide has a pH optimum of about 3.5. Accordingly, a functional equivalent of XYL001 may have a pH optimum of from about 2 to about 5, for example from about 4 to about 5, preferably of about 3.5. A functional equivalent of XYL001 may be active at a pH of about 5 or less, of about 4.5 or less, of about 4 or of about 3.5 or less.

The XYL002 polypeptide has a pH optimum of between 6 and 6.5. Accordingly, a functional equivalent of XYL002 may have a pH optimum of from about 5 to about 8, for example example from about 5.5 to about 7, preferably of from about 6 to about 6.5. A functional equivalent of XYL002 may be active at a pH of at least about 5, of at least about 5.5, of at least about 6.0, of at least about 6.5, of at least about 7.0, of at least about 7.5 or of at least about 8.0 or higher.

The XYL001 polypeptide has a temperature optimum of about 50°C and is active at temperatures in excess of that. Accordingly, a functional equivalent of XYL001 may by active at a temperature of at least about 50°C, for example at a temperature of at least about 55°C or at a temperature of at least about 60°C or
5 higher.

The XYL002 polypeptide has a temperature optimum of about 55°C and is active at temperatures in excess of that. Accordingly, a functional equivalent of XYL001 may by active at a temperature of at least about 50°C, for example at a temperature of at least about 55°C, at least about 60°C, at least about 65°C or at
10 least about 70°C or higher.

For a functional equivalent to be “active” at a given temperature or pH, it should have at least about 40%, at least about 50%, at least about 60%, at least about 70% or at least about 80% or higher of the activity they possess at the temperature or pH optimum specified above.

15 A functional equivalent of the invention may be specific for a particular type of xylan or show a preference for a particular type of xylan. For example, a functional equivalent may have a preference for xylan derived from a particular source, for example wheat, corn, barley, malt, oats, sugar cane or sorghum (millet). Alternatively, a functional equivalent of the invention may work equally
20 well on xylans from a broad range of sources. A functional equivalent of the invention may have a preference for cleaving xylan polymers with specific types of substituents sucb as arabinose substituents, glucuronic acid substituents or a mixture thereof. Alternatively, a functional equivalent may cleave all such xylans equally well. A functional equivalent will be specific for a particular type of xylan if
25 it is capable of degrading that type of xylan whilst being substantially incapable of degrading any other type of xylan. A functional equivalent will have a preference for a particular type of xylan if it is capable of degrading that type of xylan to a greater extent that it does other types of xylan.

30 **Host cells**

In another embodiment, the invention features cells, e.g., transformed host cells or recombinant host cells, that contain a nucleic acid or vector encompassed by the invention. A "transformed cell" or “recombinant cell” is a cell into which (or into an ancestor of which) has been introduced, by means of
35 recombinant DNA techniques, a nucleic acid or vector according to the invention.

Both prokaryotic and eukaryotic cells are included, e.g., bacteria, fungi, yeast, and the like, especially preferred are cells from filamentous fungi, in particular *Gloeophyllum trabeum* or *Coprinus cinereus*. A cell of the invention is typically not a wild-type *Gloeophyllum trabeum* or *Coprinus cinereus* cell or a naturally-
5 occurring cell.

A nucleic acid (or a nucleic acid which it comprises within a vector) may be homologous or heterologous with respect to the cell into which it is introduced. In this context, a nucleic acid is homologous to a cell if the nucleic acid naturally occurs in that cell. A nucleic acid is heterologous to a cell if the nucleic acid does
10 not naturally occur in that cell. Accordingly, the invention provides a cell which comprises a heterologous or a homologous XYL001 or XYL002 sequence. A cell which comprises a homologous XYL001 sequence may comprise multiple copies of that sequence. The resulting cell may therefore overexpress the polypeptide encoded by the XYL001 polynucleotide.

15 A host cell can be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in a specific, desired fashion. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may facilitate optimal functioning of the protein.

Various host cells have characteristic and specific mechanisms for post-
20 translational processing and modification of proteins and gene products. Appropriate cell lines or host systems familiar to those of skill in the art of molecular biology and/or microbiology can be chosen to ensure the desired and correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of
25 the primary transcript, glycosylation, and phosphorylation of the gene product can be used. Such host cells are well known in the art.

Host cells also include, but are not limited to, mammalian cell lines such as CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and choroid plexus cell lines.

30 If desired, a stably transfected cell line can produce the polypeptides according to the invention. A number of vectors suitable for stable transfection of mammalian cells are available to the public, methods for constructing such cell lines are also publicly known, e.g., in Ausubel *et al.* (supra).

Use of XYL001 of XYL002 xylanase in industrial processes

The invention also relates to the use of the xylanase according to the invention in a selected number of industrial processes. Despite the long term experience obtained with these processes, the xylanase according to the invention features a number of significant advantages over the enzymes currently used. Depending on the specific application, these advantages can include aspects like lower production costs, higher specificity towards the substrate, less antigenic, less undesirable side activities, higher yields when produced in a suitable microorganism, more suitable pH and temperature ranges, better tastes of the final product as well as food grade and kosher aspects.

The present invention also relates to methods for preparing a dough or a cereal-based food product comprising incorporating into the dough an effective amount of a xylanase of the present invention. This improves one or more properties of the dough or the cereal-based food product obtained from the dough relative to a dough or a cereal-based food product in which the polypeptide is not incorporated.

The preparation of the cereal-based food product according to the invention further can comprise steps known in the art such as boiling, drying, frying, steaming or baking of the obtained dough.

The phrase “incorporating into the dough” is defined herein as adding the xylanase according to the invention to the dough, any ingredient from which the dough is to be made, and/or any mixture of dough ingredients form which the dough is to be made. In other words, the xylanase according to the invention may be added in any step of the dough preparation and may be added in one, two or more steps. The xylanase according to the invention is added to the ingredients of a dough that is kneaded and prepared to make the cereal based-food product. Preparation can include boiling, drying, frying, steaming or baking of the dough as is known in the art.

Products that are made from a dough that is boiled are for example boiled noodles, dumplings, products that are made from fried dough are for example doughnuts, beignets, fried noodles, products that are made for steamed dough are for example steamed buns and steamed noodles, examples of products made from dried dough are pasta and dried noodles and examples of products made from baked dough are bread, cookies, cake.

The term “effective amount” is defined herein as an amount of the

xylanase according to the invention that is sufficient for providing a measurable effect on at least one property of interest of the dough and/or cereal-based food product.

The term “improved property” is defined herein as any property of a dough and/or a product obtained from the dough, particularly a cereal-based food product, which is improved by the action of the xylanase according to the invention relative to a dough or product in which the xylanase according to the invention is not incorporated. The improved property may include, but is not limited to, increased strength of the dough, increased elasticity of the dough, increased stability of the dough, improved machineablity of the dough, improved proofing resistance of the dough, reduced stickiness of the dough, improved softness of the dough, improved water absorption of the dough, improved extensibility of the dough, increased volume of the cereal-based food product, reduced blistering of the cereal-based food product, improved crumb structure of the baked product, improved softness of the cereal-based food product, improved flavour of the cereal-based food product, improved anti-staling of the cereal-based food product. Improved properties related to pasta and noodle type of cereal-based products are for example improved firmness, reduced stickiness, improved cohesiveness and reduced cooking loss.

The improved property may be determined by comparison of a dough and/or a cereal-based food product prepared with and without addition of a polypeptide of the present invention in accordance with the methods of present invention are described below in the Examples. Organoleptic qualities may be evaluated using procedures well established in the baking industry, and may include, for example, the use of a panel of trained taste-testers.

The term “increased strength of the dough” is defined herein as the property of a dough that has generally more elastic properties and/or requires more work input to mould and shape.

The term “increased elasticity of the dough” is defined herein as the property of a dough which has a higher tendency to regain its original shape after being subjected to a certain physical strain.

The term “increased stability of the dough” is defined herein as the property of a dough that is less susceptible to mechanical abuse thus better maintaining its shape and volume and is evaluated by the ratio of height: width of a cross section of a loaf after normal and/or extended proof.

The term “reduced stickiness of the dough” is defined herein as the property of a dough that has less tendency to adhere to surfaces, e.g., in the dough production machinery, and is either evaluated empirically by the skilled test baker or measured by the use of a texture analyser (e.g., TAXT2) as known in the
5 art.

The term “improved softness of the dough” is defined herein as the property of a dough that it is softer in touch and is evaluated empirically by the skilled test baker.

The term “improved water absorption of the dough” is defined herein as the
10 property of a dough having a higher amount of water bound in the dough. Water absorption of dough can be measured with a Farinograph by standard methods according to the International Association of Cereal Chemistry (ICC) and the American Association of Cereal Chemistry (AACC 54-2, ICC 115).

The term “improved extensibility of the dough” is defined herein as the
15 property of a dough that can be subjected to increased strain or stretching without rupture.

The term “improved machineability of the dough” is defined herein as the property of a dough that is generally less sticky and/or more firm and/or more elastic.

20 The term “increased proofing resistance of a dough” is defined as the ability of the dough to withstand prolonged proofing times.

The term “increased volume of the baked product” is measured as the volume of a given loaf of bread determined by an automated bread volume analyser (eg. BVM-3, TexVol Instruments AB, Viken, Sweden), using ultrasound
25 or laser detection as known in the art.

The term “reduced blistering of the cereal-based food product” is defined herein as a visually determined reduction of blistering on the crust of the prepared cereal based food product.

The term “improved crumb structure of the baked product” is defined
30 herein as the property of a baked product with finer and/or thinner cell walls in the crumb and/or more uniform/homogenous distribution of cells in the crumb and is usually evaluated empirically by the skilled test baker or by digital image analysis as known in the art (eg. C-cell, Calibre Control International Ltd, Appleton, Warrington, UK).

35 The term “improved softness of the cereal-based food product” is the

opposite of “firmness” and is defined herein as the property of a baked product that is more easily compressed and is evaluated either empirically by the skilled test baker or measured by the use of a texture analyzer (e.g., TAXT2) as known in the art.

5 The term “improved flavor of the cereal-based food product” is evaluated by a trained test panel.

The term “improved anti-staling of the cereal-based food product” is defined herein as the properties of a cereal-based food product that have a reduced rate of deterioration of quality parameters, e.g., softness and/or elasticity,
10 during storage.

The term “firmness” (or “hardness”) is defined herein as the degree of resistance to the first bite and is defined as the force required penetrating pasta or noodles and is either evaluated sensorial or measured with a texture analyzer (TAXT2).

15 The term “reduced stickiness” of cooked pasta of noodle is defined herein as the property of the pasta or noodle to adhere to surfaces and is either evaluated sensorial or measured with a texture analyzer (TAXT2).

The term “cohesiveness” is defined herein as the force of internal bonds holding the pasta or noodle structure or the ability to withstand compression is
20 measured with a texture analyzer (TAXT2).

The term “cooking loss” is defined as the mass of solids lost to cooking water during boiling.

The term “dough” is defined herein as a mixture of cereal flour and other ingredients firm enough to knead or roll. Examples of cereals are wheat, rye, corn,
25 maize, barley, rice, groats, buckwheat and oat. Wheat is here and hereafter intended to encompass all known species of *Triticum* genus, for example *aestivum*, *durum* and/or *spelta*. Examples of suitable other ingredients are: the xylanase according to the present invention, additional enzymes, chemical additives and/or processing aids. The dough may be fresh, frozen, pre-pared, or
30 pre-baked. The preparation of a dough from the ingredients described above is well known in the art and comprises mixing of said ingredients and processing aids and one or more moulding and optionally fermentation steps. The preparation of frozen dough is described by Kulp and Lorenz in Frozen and Refrigerated Doughs and Batters.

35 The term “cereal-based food product” is defined herein as any product

prepared from a dough, either of a soft or a crisp character. Examples of cereal-based food products, whether of a white, light or dark type, which may be advantageously produced by the present invention are bread (in particular white, whole-meal or rye bread), typically in the form of loaves or rolls, French baguette-
5 type bread, pasta, noodles, doughnuts, bagels, cake, pita bread, tortillas, tacos, cakes, pancakes, biscuits, cookies, pie crusts, steamed bread, and crisp bread, and the like.

The term “baked product” is defined herein as any cereal-based food product prepared by baking the dough.

10 The xylanase of the present invention and/or additional enzymes to be used in the methods of the present invention may be in any form suitable for the use in question, e.g., in the form of a dry powder, agglomerated powder, or granulate, in particular a non-dusting granulate, liquid, in particular a stabilized liquid, or protected enzyme such described in WO01/11974 and WO02/26044.
15 Granulates and agglomerated powders may be prepared by conventional methods, e.g., by spraying the xylanase according to the invention onto a carrier in a fluid-bed granulator. The carrier may consist of particulate cores having a suitable particle size. The carrier may be soluble or insoluble, e.g., a salt (such as NaCl or sodium sulphate), sugar (such as sucrose or lactose), sugar alcohol (such
20 as sorbitol), starch, rice, corn grits, or soy. The xylanase according to the invention and/or additional enzymes may be contained in slow-release formulations. Methods for preparing slow-release formulations are well known in the art. Adding nutritionally acceptable stabilizers such as sugar, sugar alcohol, or another polyol, and/or lactic acid or another organic acid according to established
25 methods may for instance, stabilize liquid enzyme preparations.

The xylanase according to the invention may also be incorporated in yeast comprising compositions such as disclosed in EP-A-0619947, EP-A-0659344 and WO02/49441.

For inclusion in pre-mixes of flour it is advantageous that the polypeptide
30 according to the invention is in the form of a dry product, e.g., a non-dusting granulate, whereas for inclusion together with a liquid it is advantageously in a liquid form.

One or more additional enzymes may also be incorporated into the dough. The additional enzyme may be of any origin, including mammalian and
35 plant, and preferably of microbial (bacterial, yeast or fungal) origin and may be

obtained by techniques conventionally used in the art. Enzymes may conviently be produced in microorganisms. Microbial enzymes are available from a variety of sources; Bacillus species are a common source of bacterial enzymes, whereas fungal enzymes are commonly produced in Aspergillus species.

5 Suitable additional enzymes include starch degrading enzymes, xylanases, oxidizing enzymes, fatty material splitting enzymes, or protein-degrading, modifying or crosslinking enzymes.

Starch degrading enzymes are for instance endo-acting enzymes such as alpha-amylase, maltogenic amylase, pullulanase or other debranching enzymes and 10 exo-acting enzymes that cleave off glucose (amyloglucosidase), maltose (beta-amylase), maltotriose, maltotetraose and higher oligosaccharides.

Suitable xylanases are for instance xylanases, pentosanases, hemicellulase, arabinofuranosidase, glucanase, cellulase, cellobiohydrolase, beta-glucosidase, and others.

15 Oxidizing enzymes are for instance glucose oxidase, hexose oxidase, pyranose oxidase, sulfhydryl oxidase, lipoxygenase, laccase, polyphenol oxidases and others.

Fatty material splitting enzymes are for instance triacylglycerol lipases, phospholipases (such as A_1, A_2, B, C and D) and galactolipases.

20 Protein degrading, modifying or crosslinking enzymes are for instance endo-acting proteases (serine proteases, metalloproteases, aspartyl proteases, thiol proteases), exo-acting peptidases that cleave off one amino acid, or dipeptide, tripeptide etceteras from the N-terminal (aminopeptidases) or C-terminal (carboxypeptidases) ends of the polypeptide chain, asparagines or 25 glutamine deamidating enzymes such as deamidase and peptidoglutaminase or crosslinking enzymes such as transglutaminase.

In a preferred embodiment, the additional enzyme may be an amylase, such as an alpha-amylase (can be useful for providing sugars fermentable by yeast and retarding staling) or beta-amylase, cyclodextrin glucanotransferase, 30 peptidase, in particular, an exopeptidase (can be useful in flavour enhancement), transglutaminase, lipase (can be useful for the modification of lipids present in the dough or dough constituents), phospholipase, cellulase, hemicellulase, in particular a pentosanase such as xylanase (can be useful for the partial hydrolysis of pentosans which increases the extensibility of the dough), protease (can be 35 useful for gluten weakening in particular when using hard wheat flour), protein

disulfide isomerase, e.g., a protein disulfide isomerase as disclosed in WO 95/00636, glycosyltransferase, peroxidase (can be useful for improving the dough consistency), laccase, or oxidase, e.g., an glucose oxidase, hexose oxidase, aldose oxidase, pyranose oxidase, lipoxygenase or L-amino acid oxidase (can be useful in improving dough consistency).

When one or more additional enzyme activities are to be added in accordance with the methods of the present invention, these activities may be added separately or together with the polypeptide according to the invention, optionally as constituent(s) of the bread-improving and/or dough-improving composition. The other enzyme activities may be any of the enzymes described above and may be dosed in accordance with established baking practices.

Examples of suitable chemical additives with improving properties comprise oxidising agents such as dehydro-ascorbic acid, bromate and azodicarbonate, reducing agents such as L-cysteine and glutathione, emulsifiers acting as dough conditioners such as diacetyl tartaric esters of mono/diglycerides (DATEM), sodium stearoyl lactylate (SSL) or calcium stearoyl lactylate (CSL), or acting as crumb softeners such as glycerol monostearate (GMS) etceteras, fatty materials such as triglycerides (fat) or lecithin and others. The chemical additives are in most cases added in powder form.

Also, processing aid compositions which are tailored to specific baking applications, may be composed of a dedicated mixture of chemical additives and enzyme.

The present invention also relates to methods for preparing a baked product comprising baking a dough obtained by a method of the present invention to produce a baked product. The baking of the dough to produce a baked product may be performed using methods well known in the art.

The present invention also relates to doughs and cereal-based food products, respectively, produced by the methods of the present invention.

The present invention further relates to a pre-mix, e.g., in the form of a flour composition, for dough and/or cereal-based food products made from dough, in which the pre-mix comprises a polypeptide of the present invention. The term “pre-mix” is defined herein to be understood in its conventional meaning, i.e., as a mix of baking agents, generally including flour, which may be used not only in industrial bread-baking plants/facilities, but also in retail bakeries. The pre-mix may be prepared by mixing the polypeptide or a bread-improving and/or dough-

improving composition of the invention comprising the polypeptide with a suitable carrier such as flour, starch, a sugar, or a salt. The pre-mix may contain other dough-improving and/or bread-improving additives, e.g., any of the additives, including enzymes, mentioned above.

5 The present invention further relates to baking additives in the form of a granulate or agglomerated powder, which comprise a polypeptide of the present invention. The baking additive preferably has a narrow particle size distribution with more than 95% (by weight) of the particles in the range from 25 to 500 μm.

In dough and bread making the present invention may be used in 10 combination with the processing aids defined hereinbefore such as the chemical processing aids like oxidants (e.g. dehydro-ascorbic acid), reducing agents (e.g. L-cysteine), oxidoreductases (e.g. glucose oxidase) and/or other enzymes such as polysaccharide modifying enzymes (e.g. α–amylase, hemicellulase, branching enzymes, etc.) and/or protein modifying enzymes (endoprotease, exoprotease, 15 branching enzymes, etc.).

In addition to the use of the xylanases according to the present invention in food applications, the present invention also relates to the use of the XYL001 or XYL002 according to the present invention in other industrial applications.

The enzymes of the present invention may be used in any application for 20 hydrolyzing non-starch polysaccharide (NSP). For example, one application is in the detergent industry for removal from laundry of carbohydrate-based stains.

The textile industry uses various enzymes to improve the properties of its products. Such improvement relates to softness, quality of the finish, "stone-wash look" of denim, etc. Enzymes are used in detergents in order to improve its 25 efficacy to remove most types of dirt. Enzymes have been used in textile processing since the early part of this century to remove starch-based sizing, but only in the past decade has serious attention been given to using enzymes for a wide range of textile applications. Enzymes are expected to have an even greater impact on effluent quality as more fibre preparation, pre-treatment and value-30 added finishing processes convert to biotreatment. In addition, enzymes are very effective catalysts even under mild conditions and do not require the high energy input often associated with chemical processes. The use of the xylanase of the present invention finds utility in the detergent industry for removal from laundry of carbohydrate-based stains. Xylanase can reduce the amount of bleaching 35 chemicals to obtain a given brightness in pre-bleaching of the lignin. It is

suggested that non-starch carbohydrate enzymes depolymerises xylan blocks and increases accessibility or helps liberation of residual lignin by releasing xylan-chromophore fragments. In addition to brownstock prior to bleaching, the xylanase of the present invention can save on bleaching chemicals. The enzymes hydrolyze surface xylans and are able to break linkages between hemicellulose and lignin.

Feed enzymes have an important role to play in current farming systems. They can increase the digestibility of nutrients, leading to greater efficiency in the production of animal products such as meat and eggs. At the same time they can play a role in minimizing the environmental impact of increased animal production. Non-starch polysaccharides (NSP) can increase the viscosity of the digesta which can, in turn, decrease nutrient availability and animal performance.

Endoxylanases and phytases are the best-known feed-enzyme products. Phytase enzymes hydrolyse phytic acid and release inorganic phosphate, thereby avoiding the need to add inorganic phosphates to the diet and reducing phosphorus excretion. The use of xylanases of the present invention can improve phosphorus utilization as well as cation minerals and protein during animal digesta.

Adding specific nutrients to feed improves animal digestion and thereby reduces feed costs. A lot of feed additives are being currently used and new concepts are continuously developed. Use of specific enzymes like non-starch carbohydrate degrading enzymes could breakdown the fibre releasing energy as well as increasing the protein digestibility due to better accessibility of the protein when the fibre gets broken down. In this way the feed cost could come down as well as the protein levels in the feed also could be reduced.

Non-starch polysaccharides (NSPs) are also present in virtually all feed ingredients of plant origin. NSPs are poorly utilized and can, when solubilized, exert adverse effects on digestion. Exogenous enzymes can contribute to a better utilization of these NSPs and as a consequence reduce any antinutritional effects. The xylanases of the present invention can be used for this purpose in cereal-based diets for poultry and, to a lesser extent, for pigs and other species.

The xylanases of the present invention can be used for prebleaching of kraft pulp. Xylanases have been found to be most effective for that purpose. Xylanases attract increasing scientific and commercial attention due to applications in the pulp and paper industry for removal of hemicellulose from

dissolving pulps or for enhancement of the bleachability of pulp and, thus, reduction of the use of environmentally harmful bleaching chemicals. A similar application of xylanases for pulp prebleaching is an already well-established technology and has greatly stimulated research on hemicellulases in the past
5 decade.

The xylanase of the present invention can be used to pre-bleach the lignin to reduce the amount of bleaching chemicals to obtain a given brightness. It is suggested that xylanases depolymerise xylan blocks and increases accessibility or helps liberation of residual lignin by releasing xylan-chromophore fragments. In
10 addition to brownstock prior to bleaching, xylanases of the present invention can save on bleaching chemicals. The enzymes hydrolyze surface xylans and are able to break linkages between hemicellulose and lignin.

In addition, xylanases of the present invention can also be used in antibacterial formulation as well as in pharmaceutical products such as throat
15 lozenges, toothpastes, and mouthwash.

Other industrial applications wherein the xylanase of the present invention finds utility is for example metabolizing xylose to improve the ethanol yield from cellulosic material. The cellulose molecules are composed of long chains of glucose molecules. In the hydrolysis process, these chains are broken
20 down to "free" the sugar, before it is fermented for alcohol production. Cellulose chains can be broken into glucose molecules. Ethanol, today, is produced mostly from sugars or starches, obtained from fruits and grains. In contrast, cellulosic ethanol is obtained from cellulose, the main component of wood, straw and much of the plants. Sources of biomass for ethanol production comprise agricultural
25 residues (such as leftover crop materials from stalks, leaves, and husks of corn plants), forestry wastes (such as chips and sawdust from lumber mills, dead trees, and tree branches), energy crops (such as dedicated fast-growing trees and grasses such as switch grass), municipal solid waste (such as household garbage and paper products), food processing and other industrial wastes (such as black
30 liquor, paper manufacturing by-product, etc.). Biodiesel fuel produced from vegetable oil or other waste oil resources is advantageous in terms of sustainable resource supply, less emissions, and biodegradability. Cellulose and hemicellulose are the major components and are tightly connected and intertwined; this hampers biomass degradation by pure cellulases or pure
35 hemicellulases. The xylanase of the present invention may be used to efficiently

degrade plant cell walls to fermentable sugars for the production of energy-yielding compounds such as ethanol since a synergistic effect between xylanase and cellulase is known in the art. Cellulosic materials typically contain, in addition to cellulose, other polysaccharides, including hemicellulose. When hydrolysed, 5 hemicellulose breaks down into mostly five-carbon sugars such as xylose.

XYL001 or XYL002 xylanase according to the invention may also be used as set out above for ethanol, but for the preparation of any other fermentation product, for example lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, malic acid, fumaric acid, an amino acid, 1,3-propane-diol, 10 ethylene, glycerol, butanol, a β-lactam antibiotic or a cephalosporin.

XYL001 or XYL002 xylanase may conviently be produced in microorganisms. In the above processes, it may be advantageous to use a xylanase that is obtained by a recombinant DNA technique. Such recombinant enzymes have a number of advantages over their traditionally purified 15 counterparts. Recombinant enzymes may be produced at a low cost price, high yield, free from contaminating agents like bacteria or viruses but also free from bacterial toxins or contaminating other enzyme activities.

The invention is hereinafter illustrated by the following non-limiting examples.

20

EXAMPLES

Fermentation of the organism

25

Materials & Methods

In general, for each species, starter mycelium was grown in rich medium (either mycological broth or yeast malt broth (last case is indicated with YM in the growth conditions table)) and then washed with water. The starter was then used 30 to inoculate different liquid media or solid substrate and the resulting mycelium was used for RNA extraction and library construction.

Following are the medium recipes and the solid substrates with a referenced source (if available) as well as a table listing the media variations, since in some cases the basic recipes of the referenced source have been altered 35 depending on the species grown. This is then followed by a summary of the

specific species as grown in the examples.

A. **Mycological broth**

(source: Michelle Ricard, PAPRICAN)

5 Per liter: 10g soytone, 40g D-glucose, 1ml Trace Element solution, Double-distilled water

Adjust pH to 5.0 with hydrochloric acid (HCl) and bring volume to to 1L with double-distilled water.

Trace Element Solution contains 2mM Iron(II) sulphate heptahydrate ($FeSO_4 \cdot 7H_2O$), 1 mM Copper (II) sulphate pentahydrate ($CuSO_4 \cdot 5H_2O$), 5 mM Zinc sulphate heptahydrate ($ZnSO_4 \cdot 7H_2O$), 10 mM Manganese sulphate monohydrate ($MnSO_4 \cdot H_2O$), 5 mM Cobalt(II) chloride hexahydrate ($CoCl_2 \cdot 6H_2O$), 0.5 mM Ammonium molybdate tetrahydrate ($(NH_4)_6Mo_7O_{24} \cdot 4H_2O$), and 95 mM Hydrochloric acid (HCl)dissolved in double-distilled water.

B. **Yeast-Malt broth (YM)**

(Reference: ATCC medium No. 200)

Per liter: 3g yeast extract, 3g malt extract, 5g peptone, 10g D-glucose, Double-distilled water to 1L.

C. **Trametes Defined Medium (TDM)**

(Reference: I. D. Reid and M. G. Piace. *Effect of Residual lignin type and amount on biological bleaching of kraft pulp by Trametes versicolor*. Applied Environmental Microbiology **60**: 1395-1400, 1994.)

Per liter: 10 g D-glucose, 0.75 g L-Asparagine monohydrate, 0.68 g Potassium phosphate monobasic (KH_2PO_4), 0.25 g Magnesium sulphate heptahydrate ($Mg\ SO_4 \cdot 7H_2O$), 15 mg Calcium chloride dihydrate ($CaCl_2 \cdot 2H_2O$), 100 µg Thiamine hydrochloride, 1 ml Trace Element solution, 0.5 g Tween 80, Double distilled water

Adjust pH to 5.5 with 3M potassium hydroxide and bring volume to 1 L with double-distilled water.

Table 1. Variations of TDM media used for library construction

Variation	Description
TDM-1	Medium was prepared as in basic recipe described above.
TDM-2	Quantity of asparagine monohydrate was reduced to 0.15g.
TDM-3	Manganese sulphate monohydrate was omitted from the medium.
TDM-4	The quantity of manganese sulphate monohydrate was raised to 0.2mM final concentration in the medium.
TDM-5	The quantity of copper (II) sulphate pentahydrate was raised to 20µM.
TDM-6	Glucose was replaced with 10g per liter of cellulose (**Solka-Floc, 200FCC**)
TDM-7	Glucose was replaced with 10g per liter of xylan from birchwood (Sigma Cat. # **X-0502**)
TDM-8	Glucose was replaced with 10g per liter of wheat bran[1].
TDM-9	Glucose was replaced with 10g per liter of citrus pectin (Sigma Cat. # **P-9135).**
TDM-10	Tween80 was omitted from the medium.
TDM-11	The double-distilled water was replaced with whitewater[2] collected from peroxide bleaching (which occurs during the manufacture of fine paper).
TDM-12	The double-distilled water was replaced with whitewater[2] collected from newsprint manufacture.
TDM-13	Glucose was replaced with 5g per liter of ground hardwood kraft pulp[3].
TDM-14	The medium's pH was raised to 7.5.
TDM-15	The strain was incubated at 5°C above its optimum growth temperature.
TDM-16	The strain was incubated at 10°C below its optimum growth temperature.
TDM-17	One half of the double-distilled water was replaced with whitewater from newsprint manufacture. Glucose was omitted.
TDM-18	Potassium phosphate monobasic was replaced with 5mM phytic acid from rice (Sigma Cat. # **P3168).**
TDM-19	Asparagine monohydrate was increased to 4g per liter.
TDM-20	Asparagine monohydrate was increased to 4g per liter and glucose was replaced with 2% fructose.
TDM-21	Asparagine monohydrate was increased to 4g per liter; 100ml of double-distilled water was replaced with 100ml kerosene[4]. Glucose was omitted.
TDM-22	Asparagine monohydrate was increased to 4g per liter; 100ml of double-distilled water was replaced with 100ml heaxdecane (Sigma cat. # H0255). Glucose was omitted.
TDM-23	Asparagine monohydrate was increased to 4g per liter; one half of the double-distilled water was replaced with 25% whitewater from newsprint manufacture plus 25% white water from peroxide bleaching. Glucose was omitted.

TDM-24	Asparagine monohydrate was increased to 4g per liter and the quantity of manganese sulphate monohydrate was raised to 0.2mM final concentration in the medium.
TDM-25	Asparagine monohydrate was increased to 4g per liter and manganese sulphate monohydrate was omitted from the medium.
TDM-26	Asparagine monohydrate was increased to 4g per liter; and potassium phosphate monobasic was replaced with 5mM phytic acid from rice (Sigma Cat. # **P3168).**
TDM-27	Glucose was replaced with 10g per liter of olive oil (Sigma cat. # **O1514**)
TDM-28	One half of the double-distilled water was replaced with whitewater from peroxide bleaching. Glucose was omitted.
TDM-29	Glucose was replaced with 10g per liter of tallow.
TDM-30	Glucose was replaced with 10g per liter of yellow grease.
TDM-31	Glucose was replaced with with 10g per liter of defined lipid (Sigma cat. # L0288).
TDM-32	Glucose was replaced with with 50g per liter of D-xylose.
TDM-33	Glucose was replaced with with 20g per liter of glycerol and 20ml per liter of ethanol.
TDM-34	Glucose was reduced to 1g per liter and 10g per liter of bran was added.
TDM-35	Glucose was reduced to 1g per liter and 10g per liter of pectin (Sigma Cat. # **P-9135)** was added.
TDM-36	Glucose was replaced with 10g per liter of biodiesel.
TDM-37	Glucose was replaced with 10g per liter of soy feedstock.
TDM-38	Glucose was replaced with 10g per liter of locust bean gum (Sigma cat # **G0753**).
TDM-39	One half of double-distilled water was replaced with a 1:1 ratio of whitewater from newsprint manufacture and white water from peroxide bleaching. Glucose was omitted.
TDM-40	The medium's pH was raised to 8.5.
TDM-41	One half of double-distilled water was replaced with whitewater from peroxide bleaching; plus yeast extract was added to 1g per liter. Glucose was omitted.
TDM-42	Glucose was replaced with 5g per liter of yellow grease and 5g per liter of soy feedstock
TDM-43	Glucose was replaced with 20g per liter of fructose.
TDM-44	Glucose was replaced with 10g per liter of cellulose (**Solka-Floc, 200FCC**) plus 1g per liter of sophorose.
TDM-45	The medium's pH was raised to 8.84.

[1] We used food grade wheat bran sourced from the supermarket.
[2] All whitewaters were sourced from Quebec paper mills by PAPRICAN on our behalf.
[3] Hardwood kraft pulp was sourced from Quebec paper mills by PAPRICAN
5 on our behalf
[4] Kerosene was sourced from a general hardware store.

D. **Asparagine Salts Medium (AS):**

(Reference: R. Ikeda, T. Sugita, E. Jacobson, and T. Shinoda. *Laccase and Melanization in Clinically Important Cryptococcus Species Other Than Cryptococcus neoformans* Journal of Clinical Microbiology 40: 1214-1218,
5 2002)

Per liter: 3.0 g D-glucose, 1.0 g L-Asparagine monohydrate, 3.0 g KH_2PO_4, 0.5 g Mg $SO_4 \cdot 7H_2O$, 1 mg Thiamine

Table 2.Variations of TDM media used for library construction

Variation	Description
AS-1	Medium was prepared as in basic recipe described above.
AS-2	Glucose was replaced with 10g per liter of pectin.
AS-3	One half of double-distilled water was replaced with a 1:1 ratio of whitewater from newsprint manufacture and white water from peroxide bleaching. Glucose was omitted.

10

E. **Solid substrates used:**

SS-1 5 g Wheat Bran

SS-2 5g Wheat bran plus 5ml defined lipid.

SS-3 5g Oat bran (food grade, sourced from supermarket)

15

The *Gloeophyllum trabeum* strain was grown according to the methods described above under the following growth conditions: TDM-1, -2, -3, -4, -5, -6, -7, -8, 9, -10, -13, -14, -15, -39; YM, whereby the following optimal growth temperature was used: 25°C.

20

The *Coprinus cinereus* strain was grown according to the methods described above under the following growth conditions: TDM-1, -2, -3, -5, -6, -7, -8, 9, -10, -14, -31, -39, and -40, whereby the following optimal growth temperature was used: 37°C.

25

Building the cDNA library and annotation

Total RNA was isolated from fungal cells or mycelia when the growth cultures had reached the late log phase. For fungi with yeast-like morphology, the cells were harvested by centrifugation at 3000 rpm at 4°C and washed by
30 centrifugation with ice-cold water. After aspirating the remaining water, the cell pellets were frozen in liquid nitrogen and stored at -80°C. For the filamentous

fungi, the mycelia were collected by filtration through Miracloth and washed with water by filtration. The mycelia were padded dry using paper towels, and frozen in liquid nitrogen and stored at -80°C. To extract total RNA, the frozen mycelia or cells were ground to a fine powder in liquid nitrogen using pestle and mortar. Approximately 1-1.5 gram of frozen fungal powder was dissolved in 10 ml of TRIzol® reagent and RNA was extracted according to the manufacturer's protocol (Invitrogen Life Sciences, Catalog #15596-018). Following extraction, the RNA was dissolved at 1-1.5 mg/ml of DEPC-treated water.

The PolyATtract® mRNA Isolation Systems (Promega, Catalog #Z5300) was used to isolate poly(A)+RNA. In general, equal amounts of total RNA extracted from up to ten culture conditions were pooled. One milligram of total RNA was used for isolation of poly(A)+RNA according to the protocol provided by the manufacturer. The purified poly(A)+RNA was dissolved at 200-500 µg/ml of DEPC-treated water.

Five micrograms of poly(A)+RNA were used for the construction of cDNA library. Double-stranded cDNA was synthesized using the ZAP-cDNA® Synthesis Kit (Stratagene, Catalog #200400) according to the manufacturer's protocol with the following modifications. An anchored oligo(dT) linker-primer was used in the first-strand synthesis reaction to force the primer to anneal to the beginning of the poly(A) tail of the mRNA. The anchored oligo(dT) linker-primer has the sequence: 5'-GAGAGAGAGAGAGAGAGAGAACTAGTCTCGAGTTTTTTTTTTTTTTTTTTVN-3' (SEQ ID NO: 5) where V is A, C, or G and N is A, C, G, or T. A second modification was made by adding trehalose at a final concentration of 0.6M and betaine at a final concentration of 2M in the buffer of the first-strand synthesis reaction to promote full-length synthesis. Following synthesis and size fractionation, fractions of double-stranded cDNA with sizes longer than 600 bp were pooled. The pooled cDNA was cloned directionally into the plasmid vector BlueScript KS+® (Stratagene) or a modified BlueScript KS+ vector that contained Gateway® (Invitrogen) recombination sites. The cDNA library was transformed into *E. coli* strain XL10-Gold ultracompetent cells (Stratagene, Catalog #Z00315) for propagation.

Bacterial cells carrying cDNA clones were grown on LB agar containing the antibiotic Ampicillin for selection of plasmid-borne bacteria and X-gal and IPTG to use the blue/white system to screen for the presence cDNA inserts. The white bacterial colonies, those carrying cDNA inserts, were transferred by a

colony-picking robot to 384-well MTP for replication and storage. Clones that were to be analyzed by sequencing were transferred to 96-well deep blocks using liquid-handling robots. The bacteria were cultured at 37°C with shaking at 150 rpm. After 24 hours of growth, plasmid DNA from the cDNA clones was prepared by alkaline lysis and sequenced from the 5' end using ABI 3730xl DNA analyzers (Applied Biosystems). The chromatograms obtained following single-pass sequencing of the cDNA clones were processed using Phred (available at http://www.phrap.org) to assign sequence quality values, Lucy as described in Chou and Holmes (2001, *Bioinformatics,* 17(12) 1093-1104) to remove vector and low quality sequences, and Phrap (available at http://www.phrap.org/) to assemble overlapping sequences derived from the same gene into contigs. The protein-coding regions of the sequences were predicted using OrfPredictor (available at https://fungalgenome.concordia.ca/tools/OrfPredictor.html). Sequence similarity searches against the NCBI non-redundant database were performed with BLASTX as described in Altschul *et al.*, (1997) (Nucleic Acids Res. 25(17): 3389-3402). We used TargetIdentifier (available at https://fungalgenome.concordia.ca/tools/TargetIdentifier.html) to assess if the cDNA clones contain intact open-reading frames and to assign putative function.

Proteins targeted to the extracellular space by the classical secretory pathway possess an N-terminal signal peptide, composed of a central hydrophobic core surrounded by N- and C- terminal hydrophilic regions. We used Phobius (available at http://phobius.cgb.ki.se) and SignalP version 3 (available at http://www.cbs.dtu.dk/services/SignalP) to recognize the presence of signal peptides encoded by the cDNA clones. The tools TargetP (available at http://www.cbs.dtu.dk/services/TargetP) and Big-PI Fungal Predictor (available at http://mendel.imp.ac.at/gpi/fungi_server.html) were used to remove sequences that encode proteins which are targeted to the mitochondria or bound to the cell wall. Finally, sequences predicted to encode soluble secreted protein by these automated tools were analyzed manually. Clones that comprise full-length cDNAs which are predicted to encode soluble secreted proteins were sequenced completely.

Placing the selected cDNA into the vector for further transformation

Hybrid recombination-primer oligonucleotides were designed based on the sequence information of the completely sequenced cDNA clones. Two

sequence-specific hybrid recombination-primers were synthesized for each cDNA clone: the 5' recombination-primer comprises the Gateway (Invitrogen) recombination sequence and 20-25 nucleotides corresponding to the start of the protein translation region, and the 3' recombination-primer contains the Gateway
5 recombination sequence and 20-25 nucleotides corresponding to the end of the protein translation region. Pairs of 5' and 3' recombination-primers were used to amplify by PCR the entire protein-coding regions of the cDNAs. The PCR products were cloned by in vitro recombination first into the Entry Vector (Invitrogen) and then into the destination vector pGBFIN-GTW. The integrity of the
10 plasmids was verified by restriction enzyme analysis followed by resolution in E-gel 96 wells DNA gels (Invitrogen). The resulting plasmids containing cDNA cloned in the pGBFIN-GTW vector were used for transformation into *A. niger* for protein production.

15 **Transformation of *Aspergillus niger* in Micro Titer Plate**

Preparation of *A. niger* protoplasts for MTP transformation is performed using a previously published protocol for *A. niger* transformation (for reference, see WO199932617 and WO199846772) resulting in a protoplasts suspension of 1 x 10^8 protoplasts/ml in STC. Next, the protoplasts thus obtained are transformed
20 using either undigested or linearized donor DNA. Transformation and subsequent selection of transformants is performed as shown in the Table 3, listing the individual steps of the *A. niger* MTP transformation protocol. Transformants are plated in MTP's containing SRM (for reference, see WO199932617 and WO199846772) supplemented with 150 µg/ml phleomycin and grown for 6-7 days
25 at 30°C. The resulting transformants are transferred from the 1st selection plate to 2nd selection plates (PDA supplemented with 150 µg/ml phleomycin; for reference, see WO199932617 and WO199846772) and subsequently grown for 5 to 7 days at 30°C. Copies of the 2nd selection plates are made on PDA in MTP (preferably with phleomycin) using a 96-pin gridder.

30

Table 3: A. niger transformation and selection steps in MTP.

Step	**Volume/Remarks**
First Selection	
MTP plate	400 µl in Wide well MTP (96)
+ DNA	5 µl (manual)
+ Protoplasts/ATA/20%PEG[1]	35 µl

10 min. RT	
+ 30% PEG	90 µl
20 min. RT	
+ sorbitol 1.2 M	250 µl
Total volume	380 µl
MTP centrifuge	Spin 5 min., 1500 rpm
Aspirate remaining liquid	
+ sorbitol	20 µl
Resuspend and transfer to MTP containing SRM agar with phleomycin[2]	20 µl
Grow 5-7 days at 30°C	
Second Selection	
Transfer spores with 10% glycerol/ 0.05% Triton-X100 on PDA agar with phleomycin[2]	100 µl 2 x 30 µl
Grow 6-7 days at 30°C	

[1] Mix (for 8 MTPs); 2 ml ATA, 10 ml 20% PEG, 10 ml STC, 10 ml 1x10^8/ml protoplasts

[2] 150 µg/ml phleomycin

Using the procedure described above various DNA concentrations were used to test the efficiency of the transformation procedure. For this purpose, undigested DNA of an integrative plasmid (pGBFIN; WO199932617) was used to transform *A. niger* protoplasts. The optimal DNA concentration for MTP transformation was in the range of 0,5 – 1,0 µg/µl DNA, yielding between 50 and 100 transformants per well.

Cultivation and analysis of *Aspergillus niger* in Micro Titer Plate

Copies with fresh spores made on PDA in MTP are used for inoculation. Spores are suspended in 100 µl of STIPT medium. Two micro titer plates containing 170 µl STIPT medium are mixed with 30 µl of the spore suspension. The micro titer plates are grown for 5 to 6 days with shaking in an MTP incubator (34°C, 550 rpm, 80% humidity). Supernatants are harvested after pushing down the mycelium and the supernatants of the two inoculated MTPs are pooled. Secretion analysis is carried out on the supernatants using E-page 96 wells protein gels (Invitrogen), according to the Manufacturer catalogue, in search for visible protein bands on E-page gel.

Activity screening measurements for XYL001

Two xylanase activity assays were performed for initial screening of XYL001. The first was the hydrolysis of *p*-nitrophenylxylopyranoside (*p*NP-

xylopyranoside) and the second was measurement of release of reducing sugar from birchwood xylan.

5 *1. Assay for hydrolysis of pNP-xylanopyranoside:*

Mix on ice:

- 25 µL enzyme sample or blank (MTP supernatant (sample) or MTP medium (blank), centrifuged and diluted 1/10 with 10 mM citrate, pH 5.0)
- 25 µL 2 mM *p*-nitrophenylxylopyranoside (Sigma) in 0.1 M citrate, pH 5.0

10 Then:

Incubated for 15 min at 50°C. Cooled rapidly on ice. Added 50 µL 0.1 M sodium carbonate and mixed. Read 80 mL in a microtitre plate at 410 nm as soon as possible after adding sodium carbonate. Assays were generally performed in duplicate and raw data were corrected for the absorbance of a blank (MTP 15 medium diluted into 10 mM citrate pH 5.0). Absorbance values were converted to µmol using a standard curve prepared with different amounts of *p*-nitrophenol (in 25 µL) carried through the procedure. A positive control using 25µl of 270µg/ml beta-xylosidase (*A. niger*) SIGMA-X3501, gave an absorbance value above the blank when carried through the assay. One Unit of activity is defined as the 20 amount of enzyme releasing 1 µmol *p*-nitrophenol per min under these assay conditions.

Results

The culture supernatant from the strain transformed with the empty 25 plasmid (pGBFIN) gave apparent activities of 10 ± 4 and 4 ± 1 millUnits/mL for two independent samples.

The culture supernatant from the strain transformed with plasmid harbouring the the XYL001 gene gave an activity of 84 ± 10 milliUnits/mL. This result is significantly above the background level for the strain, and the target was 30 therefore classified as having xylosidase activity.

2. Birchwood Xylan Hydrolysis Assay

(Reference: Miller G L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem. 31:426-428)

35 This assay measures the release of reducing sugar from birchwood xylan

using 3,5-dinitrosalicylic acid (DNSA). Particulate material was first removed from MTP supernatants by centrifugation. Reducing sugars were then removed from the the MTP supernatants by repeatedly concentrating small aliquots with Microcon ultrafiltration devices (10 kDa cutoff), and diluting with 10 mM citrate buffer, pH 5: the final volume was adjusted to that of the original aliquot. Generally, a 2500-fold dilution of the original culture medium was achieved using this procedure. For use in the assay, additional dilutions of the sample were made using 10 mM citrate buffer, pH 5.0, to achieve absorbance values that fell within the standard curve.

Sample: (1/10 or greater)-diluted sample in 10mM citrate buffer pH 5.0, as described above.

Blank: (1/10 or greater)-diluted MTP-medium in 10mM citrate buffer pH 5.0 instead of sample.

Positive Control: 25 µl of 50µg/mL Xylanase (*Trichoderma viride*) SIGMA-X3876

Substrate: birchwood xylan (0.8%) in 0.1 M citrate buffer, pH 5.0 or pH 4.0

Standard: dilutions of xylose in 10 mM citrate buffer, pH 5

DNSA stock solution (100 mL):

-30g of Sodium Potassium Tartrate dissolved in 50mL of dH2O.

-1g of DNSA dissolved in 20mL of 2N NaOH

-Mix the two solutions and complete to 100ml with dH2O.

-Low heat and stirring to dissolve.

<u>Procedure</u>

25 µl-Sample (or Standard or Positive Control or Blank) were mixed with an equal volume of substrate solution, see above, on ice. Reaction mixtures were then incubated at 50°C or 70°C for 15 min, after which they were placed on ice. After 5 min, 50 µl DNSA solution (as above) were added and the mixture was incubated for 15 min at 95-98°C. After cooling on ice for 5 min, 80 µl was transferred to a 96 plate well for reading at 540 nm.

Duplicate results are expressed as U/mL (umol/min/mL), by comparing absorbance values in assay mixtures with a standard curve prepared using different amouts of xylose. Activities were all corrected for the absorbance values of the blank.

Results

The activities for two independent samples of culture supernatant (diluted 10-fold) from cells transformed with the vector (pGBFIN) only were 2.1 ± 0.01 and 1.2 ± 0.02 U/mL, at pH 5.0 and 50°C.

The supernatant from the XYL001 gene containing strain had to be diluted more than 1000 times to be within the limits of the standard curve. The measured activity was 400 ± 14 U/mL at 70°C and pH 4.0. When assayed at pH 5.0, the activity of the target was 79% of the activity at pH 4.0; when assayed at 50°C, the activity of the target was 65% that at 70°C. Applying these corrections to allow direct comparison with the empty vector transformant assays at pH 5.0 and 50°C, we can estimate that the target activity level was 98-170 times that of the empty vector control. The target was therefore classified as an active xylanase (XYL001).

Activity screening measurement for XYL002

A xylanase spot test assay was used for initial screening of XYL002, using agar plates containing oat spelt xylan.

Plate preparation

- 1 g oat spelt xylan (SIGMA-X-0627)
- 1.5 g agar

Bring to 100 ml with 0.1M citrate buffer pH 5.0. Autoclave, let cool to 50-60°C and pour into large Petri dishes (245/245mm, Fisher, 07-200-134). Dry plates.

Sample preparation

Culture supernatant diluted in 10 mM citrate buffer pH 5.0 so that vector-only control activity was barely visible on plate after development. (about 1/320 required but validated for each set of sample plates). Positive control (C+) was 30µg/ml and 100µg/ml of commercial xylanase (*Trichoderma viride*, SIGMA-X3876)

Procedure

3 µl of sample (or C+ or vector-only supernatant) was spotted on plate and dried. Incubation 16h at 40°C Pour enough (30-50 ml) Congo Red (0.5% w/v) to

cover plate surface. Shake gently and wait 15-30 min. Discard excess Congo-red solution and wash plate with 0.5M NaCl solution (2-3 changes). Zones of carbohydrate hydrolysis around spots de-colourize with washing, leaving a yellow region against a red background.

5

Result

Compared to the negative control (supernatant from vector-only transformants), and supernatants from 87 other transformants chosen at random from the cDNA library and grown on the same 96-well MTP plate, XYL002 was the

10 only sample, together with positive controls (C+), to show a clearing zone. The target was therefore classified as an active xylanase (XYL002).

Enzyme properties examples

15 **pH and temperature optima with azo-xylan substrate**

In order to determine the optimum pH for the enzymes, assays were performed using birchwood azo-xylan (Megazyme). Samples (25 ul) of a 3750x (for XYL001) and 100x (for XYL002) dilution of crude xylanase (supernatant) in appropriate buffer were incubated with 25 ul of 1% (w/v) azo-xylan in distilled

20 water at 55°C. After 15 min, 137.5 ul of 95% ethanol were added to stop the reaction, as well as 5 ul of citrate (1 M, pH 4.5). After centrifugation at 1000 x g, Absorbance of the supernatant was read at 590 nm. For the XYL001 gene product, 100 mM Citric acid-sodium phosphate buffers were used to generate a pH range of pH 3–7. The temperature optimum was determined using azo-xylan

25 as above using 100 mM Citric acid-sodium phosphate buffer pH 3.5.

For the XYL002 gene product, 100 mM Citric acid-sodium phosphate or Tris-HCl buffers were used to generate a pH range of pH 4–8. The pH of Tris-HCl buffers at 55°C was calculated by measuring the pH at lab temperature (22.5^{o}C) and subtracting 0.028/^{o}C difference.

30 The temperature optimum was determined using azo-xylan as above using 100 mM Citric acid-sodium phosphate buffer pH 6.

Fig. 1 shows the pH profile at 55°C and temperature profile at pH 3.5 for XYL001 on 1% azo-xylan. Graphs show data after subtraction of VTO (vector transformant supernatant) results.

35 Fig. 2 shows the pH profile at 55°C and temperature profile at pH 6 for

XYL002 on 1% azo-xylan. Graphs show data after subtraction of VTO (vector transformant supernatant) results.

5 **Substrate specificity**

Xylanase activity on different substrates was assayed by measuring the production of reducing-sugar ends from birchwood xylan (Sigma), beechwood xylan (Sigma), Oat Spelt xylan (Sigma), wheat arabinoxylan (Megazyme), xyloglucan (Megazyme) and carboxymethyl cellulose (CMC) (Sigma) with 10 bicinchoninic acid (BCA) (Sigma).

Samples (40 ul) of a 3200x dilution of crude XYL001 (supernatant) in citric acid-sodium phosphate (100 mM, pH 3.5) were incubated with 40 ul of 1% (w/v) xylan (or CMC or xyloglucan) in distilled water at 50°C. After 15 min, 100 ul of BCA reagent were added to 10 ul of the reaction mixture and 90 ul of distilled 15 water and the samples were incubated at 80°C for 40 min.

Samples (40 ul) of a 640x dilution of crude XYL002 (supernatant) in sodium phosphate (100 mM, pH 6) were incubated with 40 ul of 1% (w/v) xylan in distilled water at 55°C. After 15 min, 100 ul of BCA reagent were added to 10 ul of the reaction mixture and 90 ul of distilled water and the samples were incubated at 20 80°C for 40 min.

Reducing sugar production was followed by measuring the absorbance at 562 nm (Grishutin *et al.* Biochim Biophys Acta. 2004 Nov 1;1674(3):268-81). Using a standard curve generated with D-xylose, absorbance was converted into moles of reducing sugars produced. One unit of enzyme activity was defined as 1 25 umol/min of xylose released at 50°C. Specific activity was determined following protein assay (Biorad).

Figures 3 and 4 show the specific activity of XLY001 and XYL002 respectively on various substrates.

30 **Analysis of xylan degradation products by thin layer chromatography**

The products from hydrolysis of birchwood, beechwood, oatspelt xylan (Sigma) and wheat arabinoxylan (Megazyme) was analysed.

For XYL001, hydrolysis was carried out in reaction volumes of 100 µl at pH 3.5 at 50°C. The final enzyme dilution was 1500x, the final substrate concentration 35 was 0.4% and the incubations were 24 hours.

For XYL002, hydrolysis was carried out in reaction volumes of 100 µl at pH 6 at 55°C. The final enzyme dilution was 80x, the final substrate concentration was 0.4% and the incubations were 24 hours.

Samples (10 µl) of each reaction mixture were analyzed by silica gel thin layer chromatography (TLC) (Analtech, Silica gel G, 20 cm x 20 cm, 250 microns) along with molecular size standards consisting of xylose (X_1) xylobiose (X_2), xylotriose (X_3), xylostetraose (X_4) and xylopentaose (X_5) (Megazyme), using chloroform/acetic acid/H_2O (6:7:1, by vol.) as the solvent system. Reaction products were visualized by spraying a sulfuric acid/ethanol (5:95, v/v) solution on the plate followed by baking at 110°C for 10 min (Blanco *et al*. Microbiology. 1999 Aug;145 (Pt 8):2163-70).

Figures 5 and 6 show TLC analyses of reaction product using 4 different sources of xylan using XLY001 and XYL002 respectively. The encircled lanes represent the products of XYL001 and XYL002. Standards are shown at far left.

Application examples

Xylanase activity assay

Xylanase is determined on 1% wheat arabinoxylan at pH 6.0, T= 30°C, and an incubation time of 30 min. The xylose reducing residues are measured with a reducing sugar assay. One NBXU is defined as the amount of enzyme that can produce 0.5 mg of xylose reducing residues in the incubation mixture under the assay conditions.

Protein determination

Protein was determined using the Coomassie blue G-250 based Protein Assay Reagent (Product No. 1856209, PIERCE Biotechnology) according to the instructions by the manufacturer.

The xylanase sample used in baking experiments 1 and 2 (XYL001) had a protein content of 0.78 mg/ml and a xylanase activity of 640 NBXU/ml. The sample of the second xylanase (XYL002) used in baking experiment 3 had an activity of 1420 NBXU/ml and protein content of 0.19 mg/ml.

Baking experiments 1 – mini batard test

Mini batards baking test is used as a first application-screening step for

non-starch carbohydrate degrading enzymes. All tests are at least done in duplicate. The XYL001 enzyme of the invention was tested at 320 NBXU/kg flour.

Recipe

Ingredients	
Flour Kolibri (Meneba)	180 g
Flour Ibis (Meneba)	20
Fresh yeast (Koningsgist)	4.6 g
Water 59%	118 g
Salt 2%	4 g
Ascorbic acid	68 ppm
Fungal amylase Bakezyme P500	3 ppm

5

Process

Mixing	Pin mixer 6 min15 sec
Scaling	2 x 150 g
First proof	25 min, 25°C
Moulding	Bertrand moulder state 16
Final proof	90 min, 32°C, 85% RH
Baking	20 min at 240/235°C, 0.2l steam

During processing dough quality is evaluated by the baker. After cooling down bread volumes are determined by an automated bread volume analyzer
10 (BVM-3, TexVol Instruments). After baking the crumb structure is evaluated visually.

Results

The dough obtained with the xylanase XYL001 has improved extensibility
15 properties and has an improved softness. The xylanase of the invention showed a volume increase of 11% compared to a reference without xylanase. The resulting bread has a very fine crumb structure compared to the reference without xylanase.

Baking experiments 2 – batard baking test

Enzymes are tested in a baking test on batards. All tests are at least done in duplicate.

5

Recipe

Ingredients	
Flour Kolibi (Meneba)	1800 g
Flour Ibis (Meneba)	200
Fresh yeast (Koningsgist)	47 g
Water 57%	1140 g
Salt 2%	40 g
Ascorbic acid	68 ppm
Fungal amylase Bakezyme P500	3 ppm

Process

Mixing	Diosna, 2' first speed, 105 Wh second speed
Scaling	350 gram dough pieces
First proof	15 min, 32°C, 90% RH
Second proof	15 min, 32°C, 90% RH
Moulding	Stick moulder 4.5 12
Final proof	90 min, 32°C, 90% RH
Baking	25 min at 245/240°C, 0.2l steam

10 During the process the baker judges the quality of the dough (stickiness, firmness, extensibility). After cooling down to room temperature the loaf volume is determined with an automated bread volume analyzer (BVM-3, TexVol Instruments).

Results

	No xylanase (reference)	160 NBXU/kg	320 NBXU/kg	960 NBXU/kg
Volume %	100	115	122	131
Dough firmness	0	0	+	+
Dough development	0	+	+	++
Dough extensibility	0	+	+	+
Crumb structure	Open	Slightly fine	Fine	Very fine and regular

0 means reference or the same as the reference

+ means better than reference

5 ++ means much better than reference

Baking experiment 3 – mini batard test

The XYL002 xylanase was tested in the mini batard test system as described in baking experiment 1 at a dosage of 7100 NBXU/kg flour. The dosage 10 of 7100 NBXU/kg flour corresponds to only 0.95 mg protein/kg flour, due to the high specific activity of this xylanase. (Note that for the xylanase described in baking experiments 1 and 2, 320 NBXU/kg flour correspond to 0.39 mg protein/kg flour.)

15 **<u>Results</u>**

The dough obtained with XYL002 had improved extensibility properties and an improved softness. The XYL002 enzyme showed a volume increase of 9 % compared to a reference without xylanase. The resulting bread had a fine crumb structure compared to the reference without xylanase.

20

CLAIMS

1. An isolated polynucleotide which: (a) is hybridizable to a polynucleotide being the complement of SEQ ID NO: 1 or 3; or (b) has at least about 50% sequence identity with SEQ ID NO: 1 or 3.

2. An isolated polynucleotide according to claim 1 hybridizable under high stringency conditions to a polynucleotide being the complement of SEQ ID NO: 1 or 3.

3. An isolated polynucleotide according to claim 1 or 2 obtainable from a brown rot fungus or a coprinoid fungus.

4. An isolated polynucleotide according to claim 3 obtainable from *Gloeophyllum trabeum* or *Coprinus cinereus*.

5. An isolated polynucleotide encoding a polypeptide comprising an amino acid sequence according to SEQ ID NO: 2 or 4 or functional equivalents of either thereof.

6. An isolated polynucleotide encoding at least one functional domain of a polypeptide according to SEQ ID NO: 2 or 4 or functional equivalents of either thereof.

7. An isolated polynucleotide comprising a nucleotide sequence according to SEQ ID NO: 1 or 3 or functional equivalents thereof.

8. An isolated polynucleotide according to SEQ ID NO: 1 or 3.

9. A vector comprising a polynucleotide sequence according to any one of claims 1 to 8.

10. A vector according to claim 9 wherein said polynucleotide sequence according to any one of claims 1 to 8 is operatively linked with regulatory sequences suitable for expression of said polynucleotide sequence in a suitable host cell.

11. A vector according to claim 10 wherein said suitable host cell is a filamentous fungus.

12. A method for manufacturing a polynucleotide according to any one of claims 1 to 8 or a vector according to any one of claims 9 to 11 comprising the steps of culturing a host cell transformed with said polynucleotide or said vector and isolating said polynucleotide or said vector from said host cell.

5 13. An isolated polypeptide according to SEQ ID NO: 2 or 4 or functional equivalents of either thereof.

14. An isolated polypeptide according to claim 13 which has at least about 50% sequence identity with SEQ ID NO: 2 or 4.

15. An isolated polypeptide according to claim 13 or 14 obtainable from
10 *Gloeophyllum trabeum* or *Coprinus cinereus*.

16. An isolated polypeptide obtainable by expressing a polynucleotide according to any one of claims 1 to 8 or a vector according to any one of claims 9 to 11 in an appropriate host cell.

17. A xylanase comprising a functional domain of a polypeptide according to
15 any one of claims 13 to 16.

18. A xylanase according to claim 17 which is recombinant.

19. A method for manufacturing a polypeptide according to any one of claims 13 to 18 comprising the steps of transforming a suitable host cell with an isolated polynucleotide according to any one of claims 1 to 8 or a vector
20 according to any one of claims 9 to 11, culturing said cell under conditions allowing expression of said polynucleotide and optionally purifying the encoded polypeptide from said cell or culture medium.

20. A recombinant host cell comprising a polynucleotide according to any one of claims 1 to 8 or a vector according to any one of claims 9 to 11.

25 21. A recombinant host cell expressing a polypeptide according to any one of claims 13 to 18.

22. Use of a polypeptide according to claims 13 to 18 in the preparation of a food product.

23. Use according to claim 22, wherein the food product is a bakery product.

24. Process for the preparation of a food product which method comprises incorporating a polypeptide according to any one of claims 13 to 18 during
5 preparation of the said food product.

25. Process for the preparation of a dough comprising the step of adding a polypeptide according to any one of claims 13 to 18.

10 26. Use of a polypeptide according to any one of claims 13 to 18 in the preparation of a detergent.

27. Use of a polypeptide according to any one of claims 13 to 18 for in the preparation of an animal feed.

28. The use of claim 27 wherein the animal feed is cereal-based feed.

15 29. Use of a polypeptide according to any one of claims 13 to 18 for prebleaching kraft pulp.

30. Use of a polypeptide according to any one of claims 13 to 18 for prebleaching lignin.

31. Use of a polypeptide according to any one of claims 13 to 18 for producing a
20 fermentation product.

32. Use of a polypeptide according to any one of claims 13 to 18 for producing ethanol.

33. The use of any one of claims 31 or 32 further comprising adding a cellulase.

Fig.1

pH profile for Gtra5285

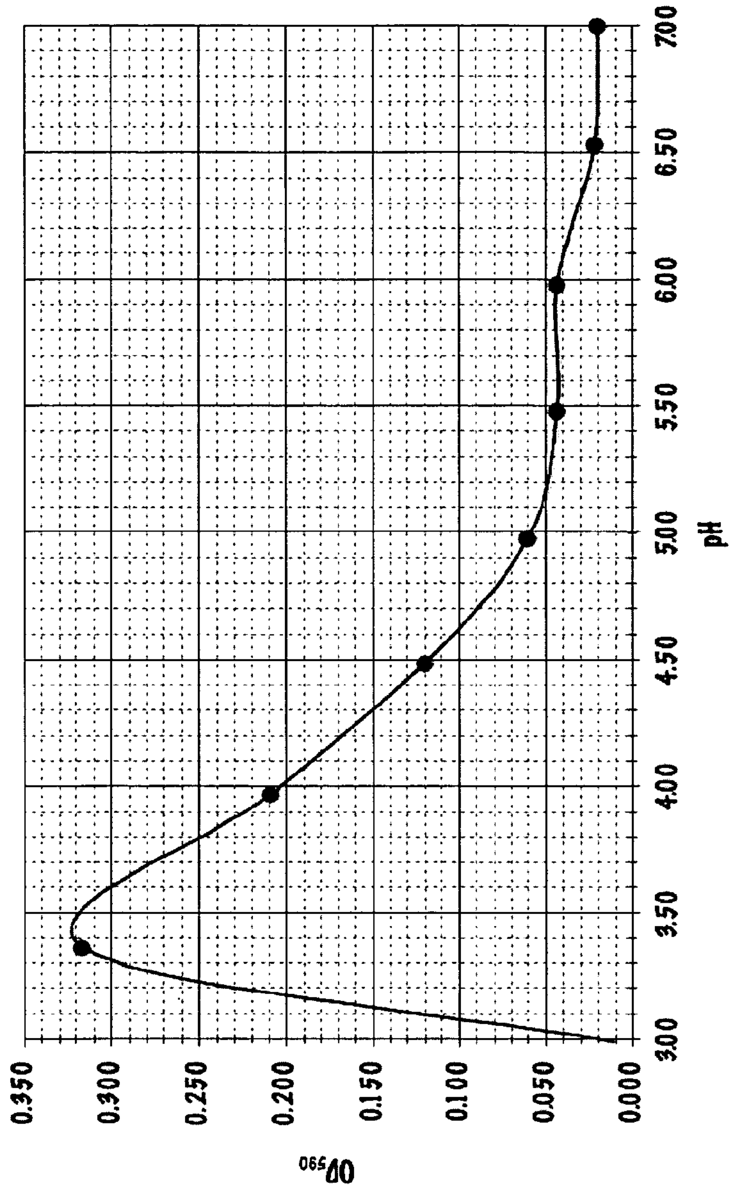


Fig.1 continued

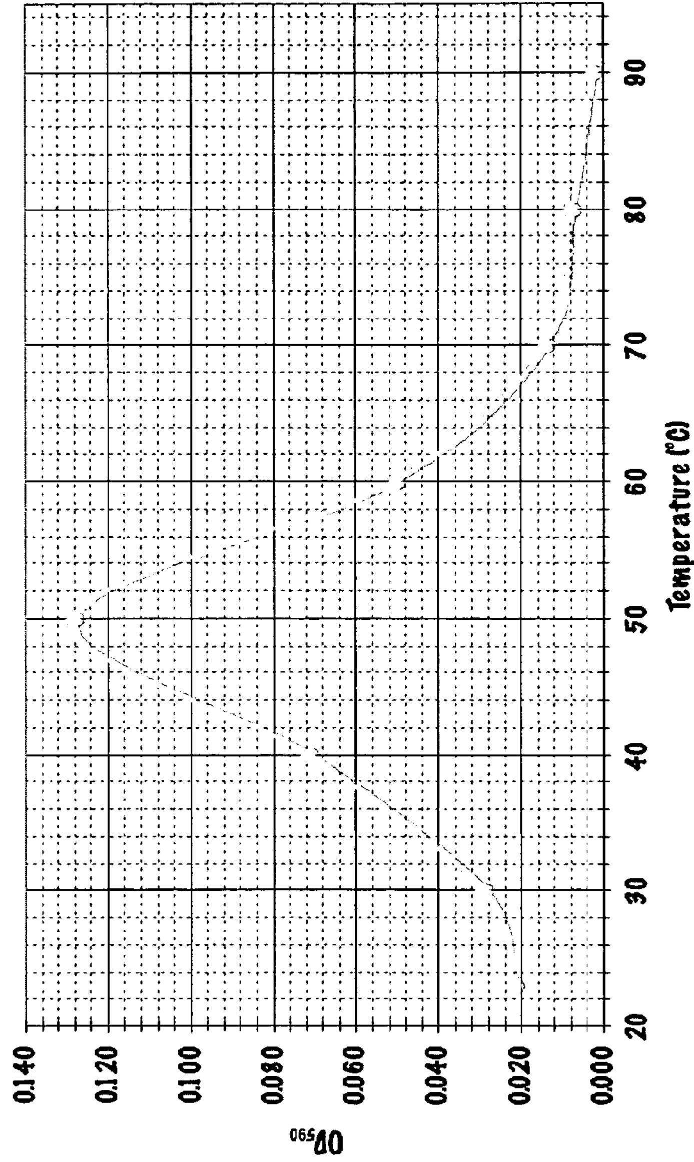


Fig.2

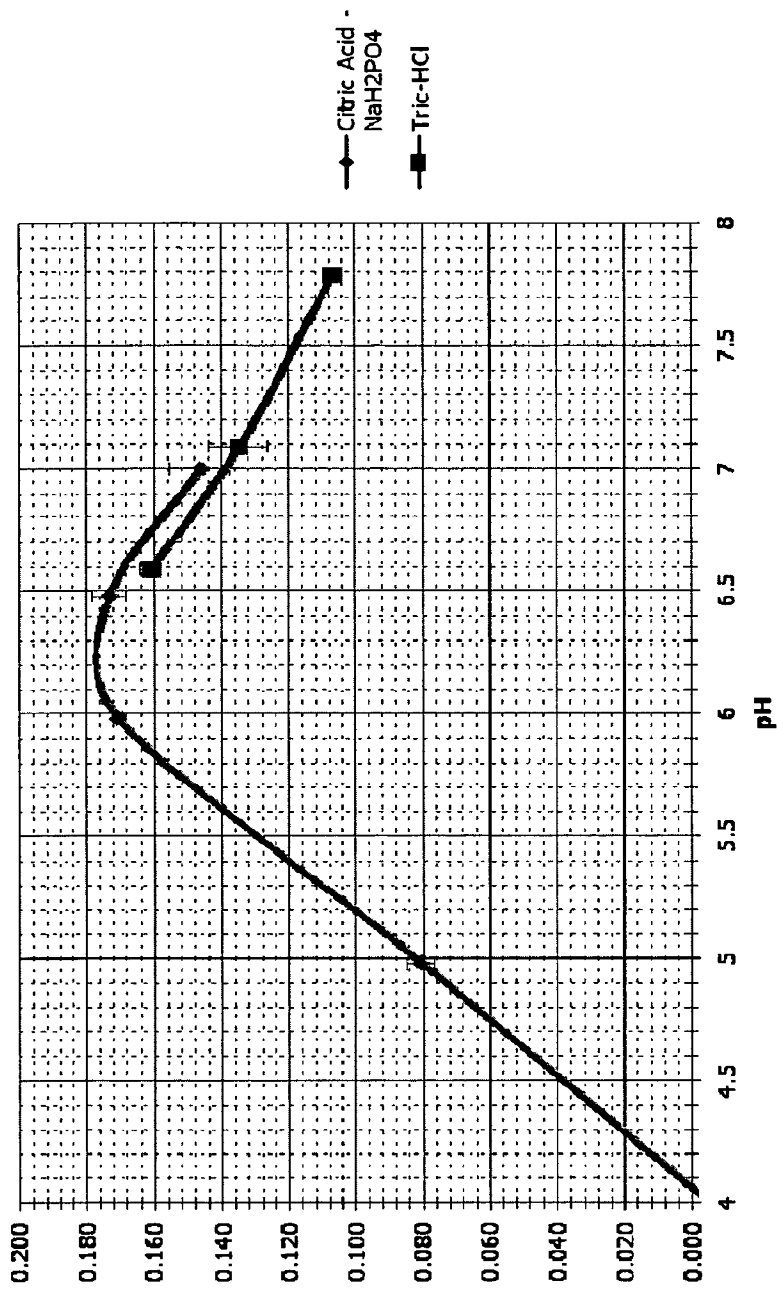


Fig.2 continued

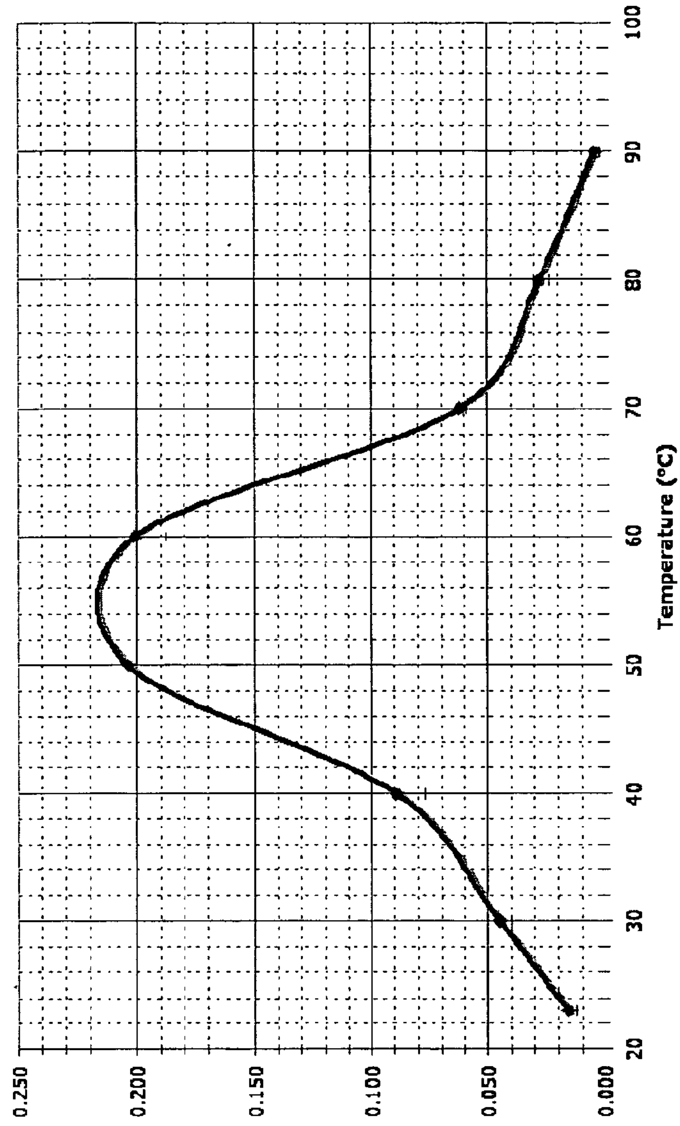


Fig.3

Substrate specificity for Gtra5285

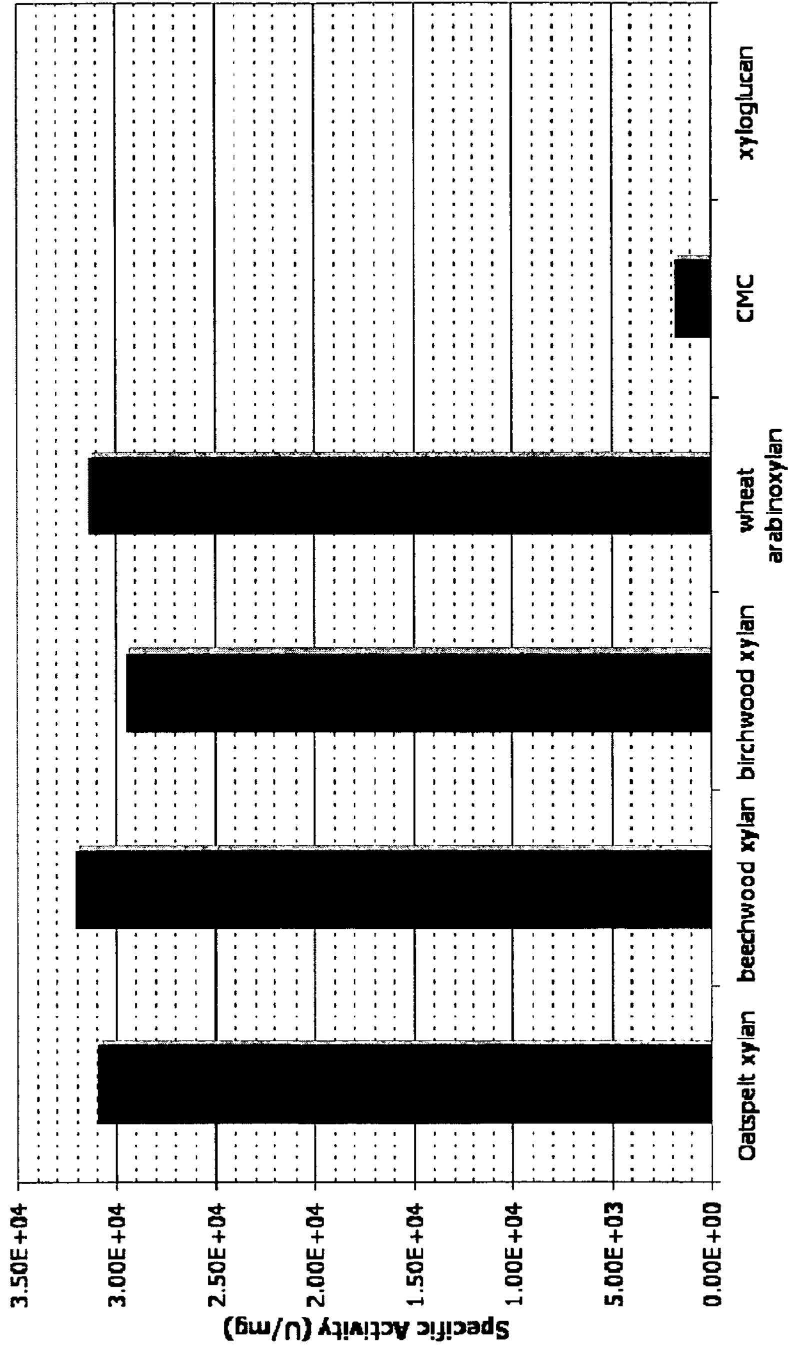
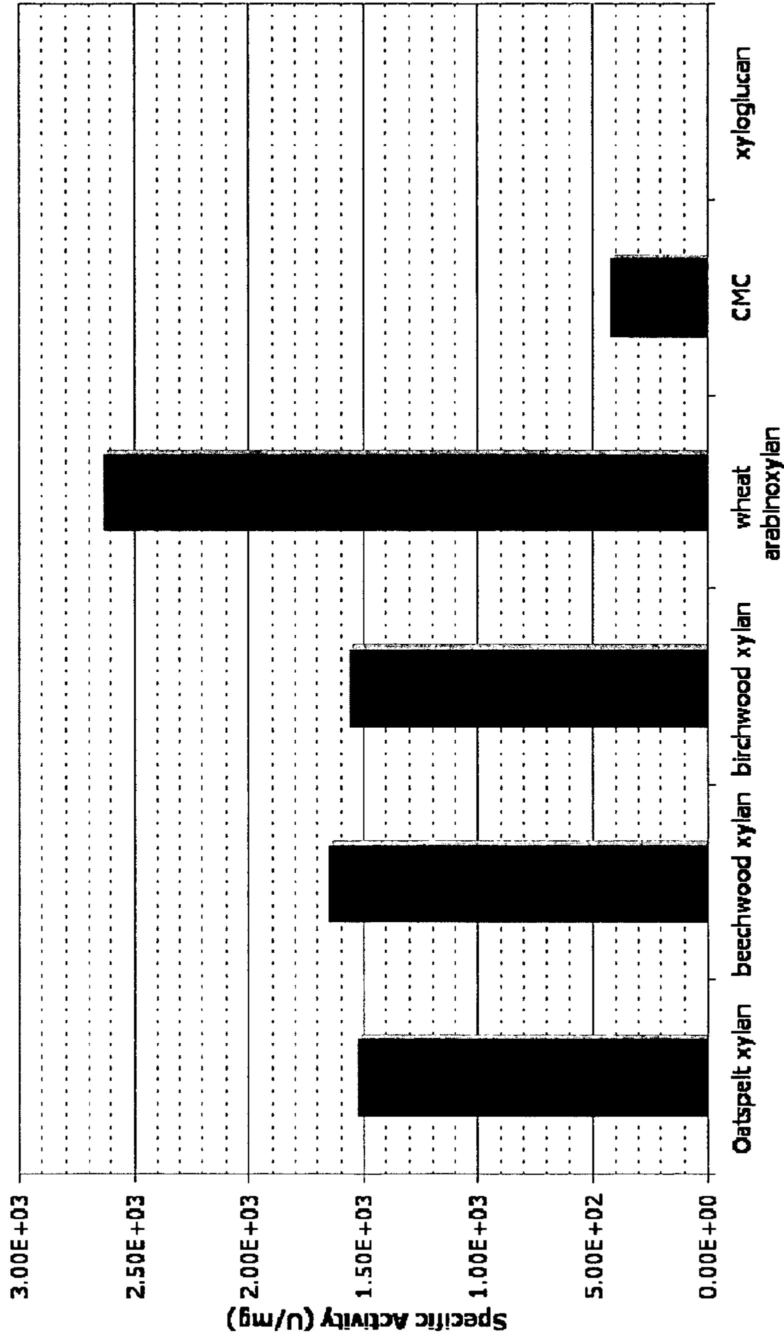
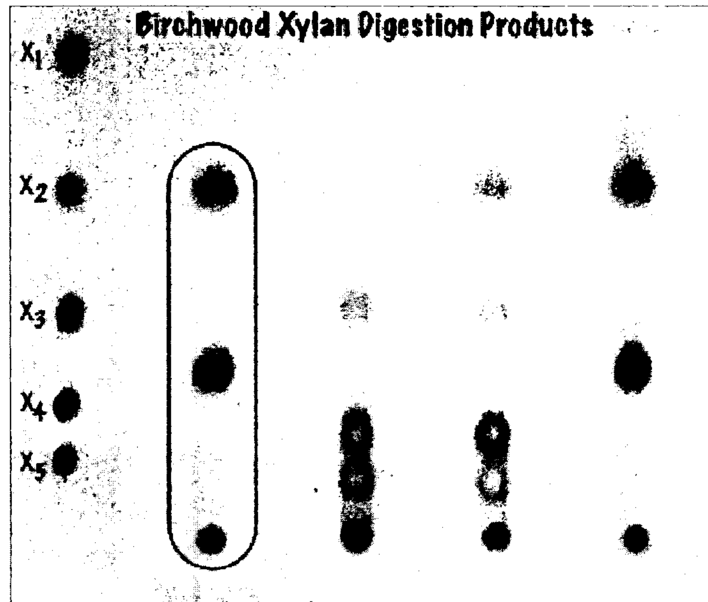


Fig.4





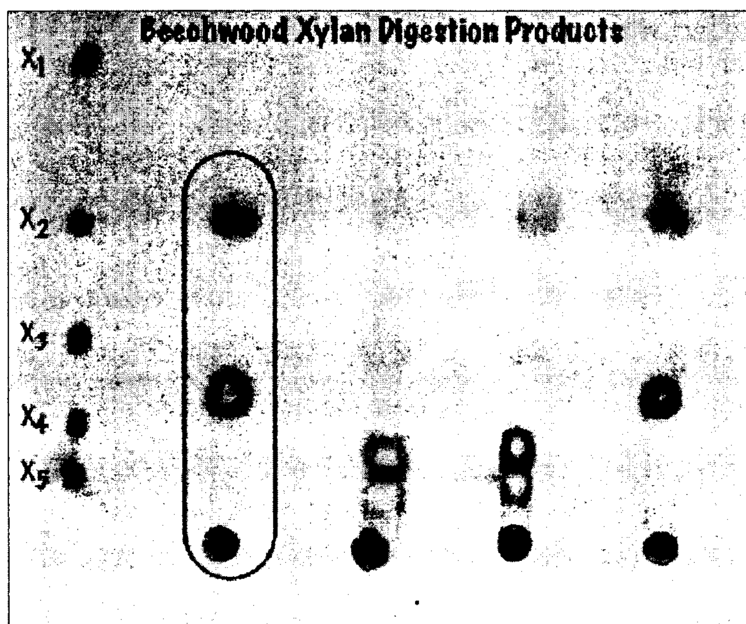
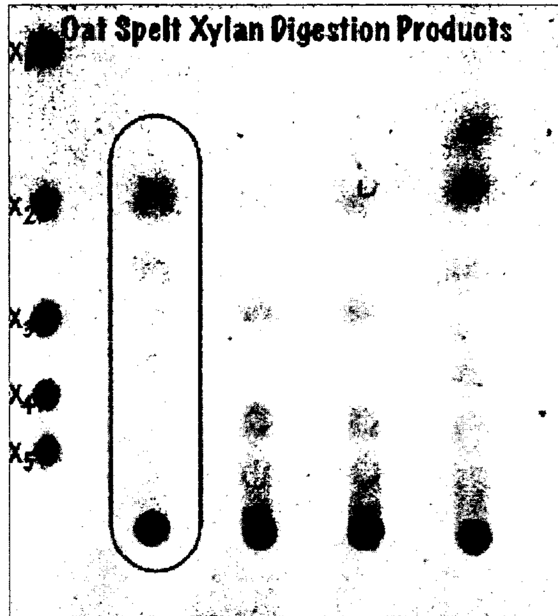


Fig. 5



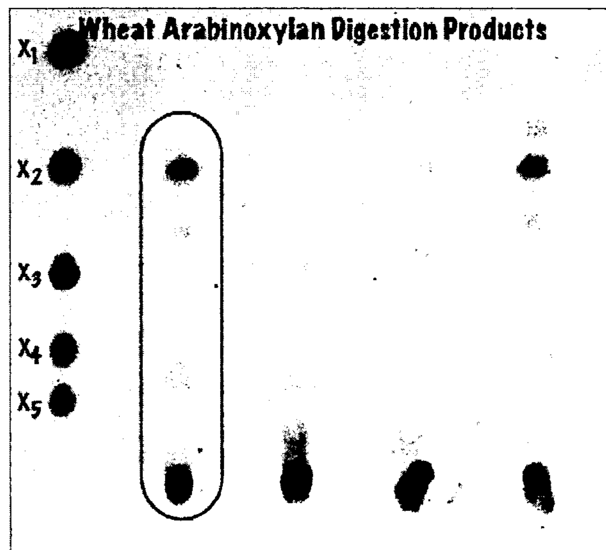
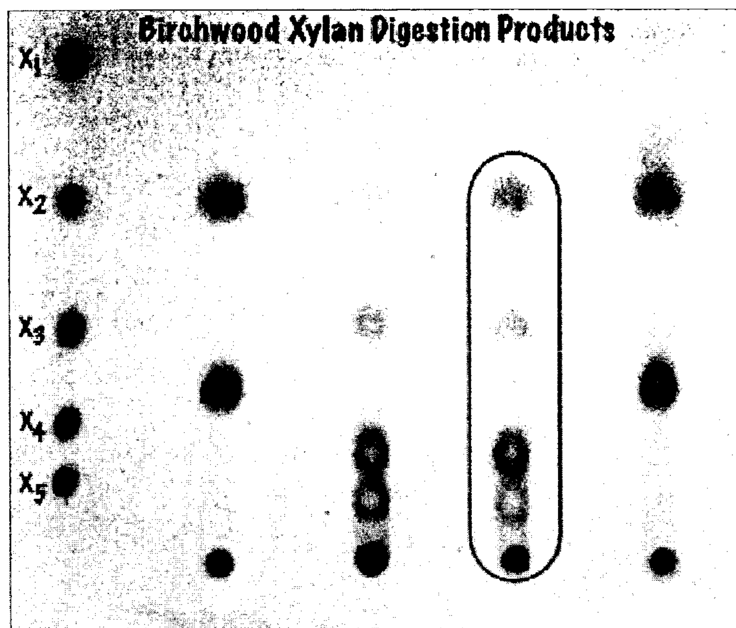


Fig.5 continued



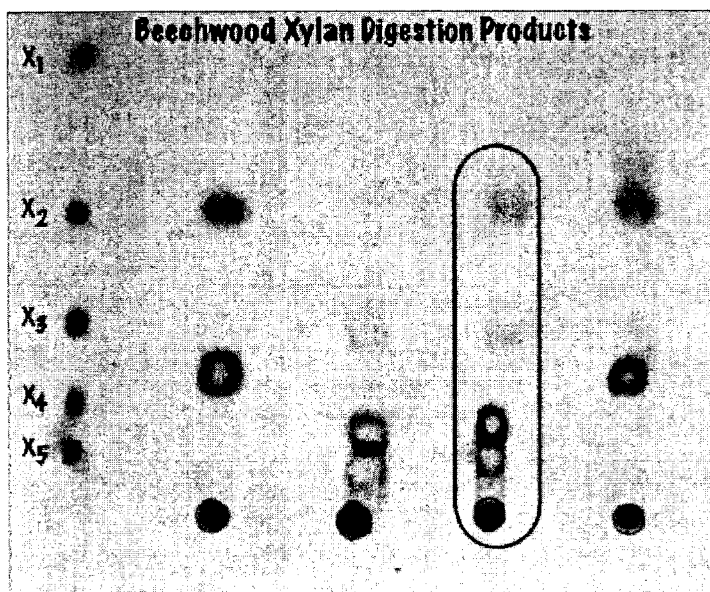
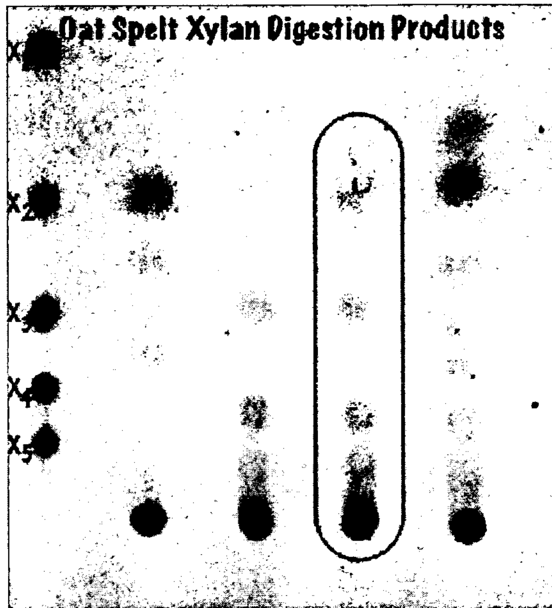


Fig. 6

10/10



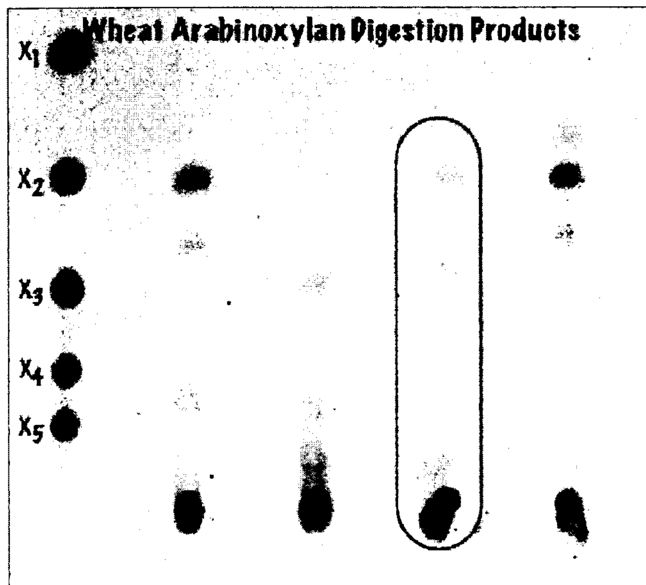


Fig.6 continued