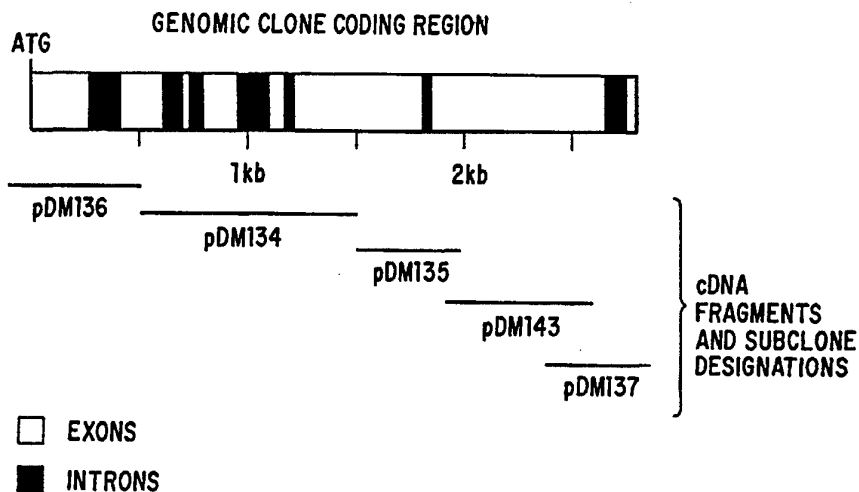




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁶ : C12N 15/53, 15/80, 9/08, 1/15 // (C12N 1/15, C12R 1:66)</p>	<p>A1</p>	<p>(11) International Publication Number: WO 96/34962 (43) International Publication Date: 7 November 1996 (07.11.96)</p>
<p>(21) International Application Number: PCT/US96/06096 (22) International Filing Date: 1 May 1996 (01.05.96) (30) Priority Data: 08/435,925 5 May 1995 (05.05.95) US (71) Applicant: NOVO NORDISK BIOTECH, INC. [US/US]; 1445 Drew Avenue, Davis, CA 95616-4880 (US). (72) Inventor: MOYER, Donna; 5219 Cowell Boulevard, Davis, CA 95616 (US). (74) Agents: ZELSON, Steve, T. et al.; Novo Nordisk of North America, Inc., Suite 6400, 405 Lexington Avenue, New York, NY 10174 (US).</p>		<p>(81) Designated States: AL, AM, AU, BB, BG, BR, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>

(54) Title: SCYTALIDIUM CATALASE GENE



(57) Abstract

The present invention relates to a nucleic acid construct containing a nucleic acid sequence encoding a *Scytalidium* catalase enzyme, as well as vectors and host cells comprising the construct.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

SCYTALIDIUM CATALASE GENE

Field of the invention

10 The present invention relates to a nucleic acid construct comprising a nucleic acid sequence encoding a *Scytalidium* catalase gene.

Background of the invention

15 The enzyme known as catalase (EC 1.11.1.6) catalyzes the decomposition of hydrogen peroxide into water and molecular oxygen. In the cell, hydrogen peroxide is produced as a by-product of aerobic metabolism and has the potential to damage a variety of macromolecules including DNA(Brawn and Fridovich, Arch. Biochem. Biophys. 206: 414-419, 1981). Catalase is induced as a defense against H₂O₂-mediated damage.

20 The enzyme has been proposed for many commercial uses. In general terms, it can be used in any situation in which it is desired to remove residual hydrogen peroxide from a system to which hydrogen peroxide has been added, e.g., for pasteurization or bleaching. One such example is the use of catalase in the textile industry for the removal of hydrogen peroxide from fabric which is bleached by an alkaline hydrogen peroxide treatment before dyeing. A similar application is its use in pulp bleaching. Catalase can also be used
25 in the removal of hydrogen peroxide from contact lenses after hydrogen peroxide disinfection.

 A number of different types of catalases have been isolated and identified, from animal, plant and microbial sources. In particular a number of filamentous fungal catalases have been found. Those which have been characterized consist of four polypeptide subunits, each having a molecular weight of 80,000 to 97,000 and contain one heme prosthetic group
30 per subunit. For example, catalase has been characterized from *Penicillium*(Vainshtein et al., J. Mol. Biol. 188:63-72, 1986), from *Neurospora*(Jacob and Orme-Johnson, Biochemistry 18:2967-2975, 1979), from *Acremonium* and *Thermoascus*(JP 5153975) and from *Aspergillus* (Fowler et al., Mol. Microbiol.9:989-998,1993). The genes encoding certain of these catalases have also been isolated, and recombinant expression achieved (WO 93/17721, Wo 93/18166;
35 JP 3103182; JP 1086879; JP 63017693); An extremely stable catalase, which retains activity at higher temperature and pH than other known catalases, has been isolated from strains of

Scytalidium and *Humicola*(WO 92/17571). These properties make the *Scytalidium/Humicola* catalases particularly effective in the removal of residual peroxide in textile applications. However, recombinant production of this enzyme has not heretofore been accomplished. The present invention provides the gene encoding the *Scytalidium* catalase and a method for recombinant expression of same.

10

Brief Description of the Figures

Figure 1 illustrates the PCR primers used to generate a 1.1 kb catalase probe. Upper case letters correspond to catalase sequence, and lower case letters correspond to restriction sites introduced by PCR(SEQ. ID NOS. 3 and 4)

15

Figure 2 illustrates the relative positions of the identified *Scytalidium thermophilum* cDNA subclones.

20

Figure 3 illustrates the restriction map of a *S. thermophilum* genomic catalase clone.

Figure 4 illustrates the *S. thermophilum* genomic clone nucleic acid and amino acid sequence(SEQ ID NOS.1 and 2)

25

Figure 5 illustrates the homology of isolated catalase peptides to the translated DNA sequence.

Figure 6 illustrates an expression construct, pDM153 for the *S. thermophilum* catalase.

30

Figure 7 illustrates the 6.3 kb *Scytalidium* genomic fragment containing the catalase gene, which is subcloned into pUC118, now called plasmid pDM125.

Summary of the Invention

35

The present invention provides a nucleic acid construct comprising a nucleic acid sequence encoding a *Scytalidium* catalase. The invention also provides a method for

production of a recombinant *Scytalidium* catalase which comprises culturing, for a time and under conditions conducive to expression of a catalase, a recombinant host cell containing a nucleic acid construct comprising a nucleic acid sequence encoding a *Scytalidium* catalase.

Detailed Description of the Invention

10 The catalase produced by strains of the genus *Scytalidium* and the genus *Humicola* have been characterized in copending US Serial No. 08/117,201, the contents of which are incorporated herein by reference. Briefly, the enzyme is one which retains at least 75% residual activity after 20 minutes at 70°C, at pH 9-10.5. Organisms of the *Humicola-Scytalidium* complex are well known, and have been described by D.H. Ellis, Trans. Br. Mycol. Soc. 78(1): 129-139, 1982. All members of this complex have been assigned to two species: *Humicola insolens*(Cooney and Emerson) and *Scytalidium thermophilum*(Cooney and Emerson)Austwick. The definition and taxonomy of the genus *Scytalidium* is described by Pesante, Annali Sper. Agr. N.S. 11:Suppl.:CCLXI-CCLXV, 1957, and by M.B. Ellis, Dematiaceous Hyphomycetes, Commonwealth Mycological Institute, Kew, Surrey, England, p.28, 1971. As used herein throughout the specification and claims, the term "*Scytalidium* catalase" gene or nucleic acid sequence is intended to encompass a catalase(and sequence encoding same) of either of the two species of the complex. It also is intended to encompass catalases of isolates which have previously been assigned to other taxonomic categories, but which are properly assigned to the *Humicola-Scytalidium* complex as defined herein. For example, the complex includes the thermophilic hyphomycetes previously classified as *Humicola grisea* var.*thermoidea* Cooney & Emerson, *H. insolens* Cooney & Emerson, and *Torula thermophila* Cooney Emerson, described in D.G. Cooney and R. Emerson, Thermophilic Fungi. An account of their biology, activities and classification, San Francisco, Freeman, 1964. Strains of the members of the complex are widely publicly available at recognized depositories, e.g., in the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland(ATCC); Centraal Bureau voor Schimmelcultures, Osterstraat 1, 3740 AG Baarn, Netherlands(CBS); University of Alberta Microfungus Collection(UAMH) and CAB International Mycological Institute (IMI). Examples of available strains from which useful genetic material can be obtained are ATCC 28085, ATCC 48409, CBS 671.88(*Scytalidium thermophilum*) and UAMH 2925, IMI 158747 and ATCC 34627(*Humicola insolens*).

As described below, to isolate a *Scytalidium* catalase gene, genomic DNA of *S. thermophilum* is used as a template in a PCR reaction with two degenerate primers based on partial peptide sequences obtained from an isolated *S. thermophilum* catalase protein. A 1.1 kb band is obtained and cloned into a plasmid to produce pDM117. A 1.1 kb HindIII/NotI fragment is isolated from this plasmid, labelled and used to screen genomic libraries.

Screening of a *S. thermophilum* genomic library yields strongly hybridizing bands. A genomic EMBL4 library is then prepared and DNA preparations made from putative positive clones. These are probed with the 1.1 kb fragment and a hybridizing 6.3 kb EcoRI fragment is identified(Figure 3); this fragment is subsequently shown to contain the entire coding region of the catalase gene as well as promoter and untranslated sequence. Screening of a cDNA library using PCR with exact match primers yields cDNA subclones which confirm the genomic sequence.

The catalase is encoded by 2791 base pairs, with 7 introns(Figure 4). The predicted amino acid sequence of 717 residues appears to encode a prepro-catalase with a 19 amino acid signal sequence, and a 17 amino acid proregion. The predicted amino acid sequence also correlates well with the amino acid sequence determined from peptides derived from sequencing of the native protein. When compared with known catalase gene sequences, the DNA sequence is closest to that of *A. niger*, showing a 61% and 57% homology at the nucleotide and amino acid levels, respectively.

The gene is PCR modified at both the 5' and 3' ends to introduce restriction sites. Three catalase fragments are ligated together and inserted between an *A. oryzae* TAKA amylase/TPI(triose phosphate isomerase fusion promoter(as described in WO 94/23022) and the *A. niger* amyloglucosidase(AMG) terminator in a plasmid(Figure 6). This plasmid is cotransformed with an amdS selection plasmid into an *A. oryzae* host cell, as described in WO 91/17243, and grown on plates for three days. Several transformants show significantly more catalase activity than untransformed controls. The best producers are cultured in shake flasks and grown for 5 days. The best producer of this group makes 12,800 CIU/ml. Yield of catalase can be improved by culturing the host cell in a medium containing heme or a heme containing material, as described in copending US Serial No. 08/284,566, the contents of which are incorporated herein by reference.

According to the invention, a *Scytalidium* gene encoding a catalase can be obtained by methods described above, or any alternative methods known in the art, using the information provided herein. The gene can be expressed, in active form, using an expression vector. A useful expression vector contains an element that permits stable integration of the vector into the host cell genome or autonomous replication of the vector in a host cell independent of the genome of the host cell, and preferably one or more phenotypic markers which permit easy selection of transformed host cells. The expression vector may also include control sequences encoding a promoter, ribosome binding site, translation initiation signal, and, optionally, a repressor gene or various activator genes. To permit the secretion of the expressed protein, nucleotides encoding a signal sequence may be inserted prior to the coding sequence of the gene. For expression under the direction of control sequences, a catalase gene to be used according to the invention is operably linked to the control sequences in the proper reading frame.

The expression vector carrying the DNA construct of the invention may be any vector which may conveniently be subjected to recombinant DNA procedures, and the choice of vector will typically depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e. a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g. a plasmid, or an extrachromosomal element, minichromosome or an artificial chromosome. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

In the vector, the DNA sequence should be operably connected to a suitable promoter sequence. The promoter may be any DNA sequence which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell. Examples of suitable promoters for directing the transcription of the DNA construct of the invention, especially in a bacterial host, are the promoter of the *lac* operon of *E.coli*, the *Streptomyces coelicolor* agarase gene *dagA* promoters, the promoters of the *Bacillus licheniformis* α -amylase gene (*amyL*), the promoters of the *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), the promoters of the *Bacillus amyloliquefaciens* α -amylase (*amyQ*), or the promoters of the *Bacillus subtilis* *xylA* and *xylB* genes. In a yeast host, a useful promoter is the *eno-1* promoter. For transcription

in a fungal host, examples of useful promoters are those derived from the gene encoding *A. oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *A. niger* neutral α -amylase, *A. niger* acid stable α -amylase, *A. niger* or *A. awamori* glucoamylase (glaA), *Rhizomucor miehei* lipase, *A. oryzae* alkaline protease, *A. oryzae* triose phosphate isomerase or *A. nidulans* acetamidase. Preferred are the TAKA-amylase, TAKA-amylase-TPI fusion, and glaA promoters.

The expression vector of the invention may also comprise a suitable transcription terminator and, in eukaryotes, polyadenylation sequences operably connected to the DNA sequence encoding the catalase of the invention. Termination and polyadenylation sequences may suitably be derived from the same or different sources as the promoter. The vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. Examples of such sequences are the origins of replication of plasmids pUC19, pACYC177, pUB110, pE194, pAMB1 and pIJ702.

The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, such as the *dal* genes from *B. subtilis* or *B. licheniformis*, or one which confers antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracycline resistance. Examples of *Aspergillus* selection markers include *amdS*, *pyrG*, *argB*, *niaD*, *sC*, and *hygB* a marker giving rise to hygromycin resistance. Preferred for use in an *Aspergillus* host cell are the *amdS* and *pyrG* markers of *A. nidulans* or *A. oryzae*. A frequently used mammalian marker is the dihydrofolate reductase (DHFR) gene. Furthermore, selection may be accomplished by co-transformation, e.g. as described in WO 91/17243.

It is generally preferred that the expression gives rise to a product that is extracellular. The catalases of the present invention may thus comprise a preregion permitting secretion of the expressed protein into the culture medium. If desirable, this preregion may be native to the catalase of the invention or substituted with a different preregion or signal sequence, conveniently accomplished by substitution of the DNA sequences encoding the respective preregions. For example, the preregion may be derived from a glucoamylase or an amylase gene from an *Aspergillus* species, an amylase gene from a *Bacillus* species, a lipase or proteinase gene from *Rhizomucor miehei*, the gene for the α -factor from

Saccharomyces cerevisiae or the calf preprochymosin gene. Particularly preferred, when the host is a fungal cell, is the preregion for *A. oryzae* TAKA amylase, *A. niger* neutral amylase, the maltogenic amylase form *Bacillus* NCIB 11837, *B. stearothermophilus* α -amylase, or *Bacillus licheniformis* subtilisin. An effective signal sequence is the *A. oryzae* TAKA amylase signal, the *Rhizomucor miehei* aspartic proteinase signal and the *Rhizomucor miehei* lipase signal.

10

The procedures used to ligate the DNA construct of the invention, the promoter, terminator and other elements, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al. Molecular Cloning, 1989).

15

The cell of the invention either comprising a DNA construct or an expression vector of the invention as defined above is advantageously used as a host cell in the recombinant production of a enzyme of the invention. The cell may be transformed with the DNA construct of the invention, conveniently by integrating the DNA construct in the host chromosome. This integration is generally considered to be an advantage as the DNA sequence is more likely to be stably maintained in the cell. Integration of the DNA constructs into the host chromosome may be performed according to conventional methods, e.g. by homologous or heterologous recombination. Alternatively, the cell may be transformed with an expression vector as described above in connection with the different types of host cells.

20

25

The host cell may be selected from prokaryotic cells, such as bacterial cells. Examples of suitable bacteria are gram positive bacteria such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus*, *Bacillus megaterium*, *Bacillus thuringiensis*, or *Streptomyces lividans* or *Streptomyces murinus*, or gram negative bacteria such as *E.coli*. The transformation of the bacteria may for instance be effected by protoplast transformation or by using competent cells in a manner known *per se*.

30

35

The host cell may also be a eukaryote, such as mammalian cells, insect cells, plant cells or preferably fungal cells, including yeast and filamentous fungi. For example, useful mammalian cells include CHO or COS cells. A yeast host cell may be selected from a species of *Saccharomyces*, *Pichia* or *Schizosaccharomyces*, e.g. *Saccharomyces cerevisiae*.

Useful filamentous fungi may be selected from a species of *Aspergillus*, e.g. *Aspergillus oryzae* or *Aspergillus niger*. Alternatively, a strain of a *Fusarium* species, e.g. *F. oxysporum*, or *F. graminearum*, can be used as a host cell. Fungal cells may be transformed by a process involving protoplast formation and transformation of the protoplasts followed by regeneration of the cell wall in a manner known *per se*. A suitable procedure for transformation of *Aspergillus* host cells is described in EP 238 023. A suitable method of transforming *Fusarium* species is described by Malardier et al., 1989.

The present invention thus provides a method of producing a recombinant catalase of the invention, which method comprises cultivating a host cell as described above under conditions conducive to the production of the enzyme and recovering the enzyme from the cells and/or culture medium. The medium used to cultivate the cells may be any conventional medium suitable for growing the host cell in question and obtaining expression of the catalase of the invention. Suitable media are available from commercial suppliers or may be prepared according to published formulae (e.g. in catalogues of the American Type Culture Collection).

The resulting enzyme may be recovered from the medium by conventional procedures including separating the cells from the medium by centrifugation or filtration, precipitating the proteinaceous components of the supernatant or filtrate by means of a salt, e.g. ammonium sulphate, followed by purification by a variety of chromatographic procedures, e.g. ion exchange chromatography, gel filtration chromatography, affinity chromatography, or the like. Preferably, the isolated protein is about 90% pure as determined by SDS-PAGE, purity being most important in food, juice or detergent applications.

In a particularly preferred embodiment, the expression of catalase is achieved in a fungal host cell, such as *Aspergillus*. As described in detail in the following examples, the catalase gene is ligated into a plasmid containing the *Aspergillus oryzae* TAKA α -amylase/TPI promoter, and the *Aspergillus nidulans amdS* selectable marker. Alternatively, the *amdS* may be on a separate plasmid and used in co-transformation. The plasmid (or plasmids) is used to transform an *Aspergillus* species host cell, such as *A. oryzae* or *A. niger* in accordance with methods described in Yelton et al. (PNAS USA 81: 1470-1474, 1984).

Those skilled in the art will recognize that the invention is not limited to use of the nucleic acid fragments specifically disclosed herein, for example, in Figure 4. It will also be apparent that the invention encompasses those nucleotide sequences that encode the

same amino acid sequences as depicted in Figure 4, but which differ from those specifically depicted nucleotide sequences by virtue of the degeneracy of the genetic code. Also, reference to Figure 4, in the specification and the claims will be understood to encompass both the genomic sequence depicted therein as well as the corresponding cDNA and RNA sequences, and the phrases "nucleic acid construct" and "nucleic acid sequences" as used herein will be understood to encompass all such variations. "nucleic acid construct" shall generally be understood to mean a nucleic acid molecule, either single- or double-stranded, which may be isolated in partial form from a naturally occurring gene or made synthetically from a naturally occurring gene template, or which has been modified to contain segments of nucleotides which are combined and juxtaposed in a manner which would not otherwise exist in nature.

In addition, the invention also encompasses other *Scytalidium* catalases, including alternate forms of catalase which may be found in *S. thermophilum* and as well as catalases which may be found in other fungi which are synonyms or fall within the definition of *Scytalidium thermophilum* as defined by Straatsma and Samson, 1993, *supra*. These include *S. indonesiacum*, *Torula thermophila*, *Humicola brevis* var. *thermoidea*, *Humicola brevispora*, *H. grisea* var. *thermoidea*, *Humicola insolens*, and *Humicola lanuginosa* (also known as *Thermomyces lanuginosus*). The invention also provides the means for isolation of catalase genes from other species of *Scytalidium*, such as *S. acidophilum*, *S. album*, *S. aurantiacum*, *S. circinatum*, *S. flaveobrunneum*, *S. hyalinum*, *S. lignicola*, and *S. uredinicum*. Identification and isolation of catalase genes from sources other than those specifically exemplified herein can be achieved by utilization of the methodology described in the present examples, with publicly available *Scytalidium* strains. Alternately, the sequence disclosed herein can be used to design primers and/or probes useful in isolating catalase genes by standard PCR or Southern hybridization techniques, using the same publicly available strains. Examples of such publicly available strains include, from the American Type Culture Collection, ATCC 16463, 28085, 36346, 48409, 66938(*S. thermophilum*); 24569(*S. acidophilum*); 16675(*S. album*); 22477(*S. aurantiacum*); 66463(*S. circinatum*); 13212(*S. flaveobrunneum*); 52297(*S. fulvum*); 38906(*S. hyalinum*); 46858(*S. indonesiacum*); 18984(*S. indonesiacum*); 32382(*S. uredinaolum*); from the International Mycological Institute(IMI; United Kingdom), IMI 243 118(*S. thermophilum*); from Centraalbureau voor Schimmelcultures(CBS; Netherlands) CBS 183.81, 671.88(*S. thermophilum*) 367.72(*S.*

acidophilum); 372.65(*S. album*);374.65(*S. aurantiacum*);654.89(*S. circinatum*);244.59 (*S. flavo-brunneum*);145.78(*S. hyalinum*);259.81(*S. indonesiacum*);233.57(*S. lignicola*);171.40(*S. terminale*); 616.84(*S. muscorum*); from Deutsche Sammlung von Mikroorganismenn und Zellkulturen (DSMZ; Germany) DSM 2842(*S. thermophilum*); DSM 2695 (*S. lignicola*). The invention also encompasses any variant nucleotide sequence, and the protein encoded thereby, which protein retains at least about an 80%, preferably about 85%, and most preferably at least about 90-95% homology with the amino acid sequence depicted in Figure 1, and which qualitatively retains the catalase activity of the sequence described herein. Useful variants within the categories defined above include, for example, ones in which conservative amino acid substitutions have been made, which substitutions do not significantly affect the activity of the protein. By conservative substitution is meant that amino acids of the same class may be substituted by any other of that class. For example, the nonpolar aliphatic residues Ala, Val, Leu, and Ile may be interchanged, as may be the basic residues Lys and Arg, or the acidic residues Asp and Glu. Similarly, Ser and Thr are conservative substitutions for each other, as are Asn and Gln. It will be apparent to the skilled artisan that such substitutions can be made outside the regions critical to the function of the molecule and still result in an active enzyme. Retention of the desired activity can readily be determined by conducting a standard titanium color method, such as is described in the present examples.

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

EXAMPLES

I. MATERIALS AND METHODS

A. Generation of PCR probes

Two degenerate PCR primers are designed based upon partial peptide sequence derived from the native *S. thermophilum* protein. The forward primer(PmeST5) corresponds to the N-terminal peptide sequence. The reverse primer(St3.2not) corresponds to the region from amino acids 265 to 271(Figure 1). 200 pmoles of each degenerate primer and 120 ng *S. thermophilum* genomic DNA are added per 100 μ l reaction. The Taq polymerase and PCR buffer are obtained from Boehringer Mannheim. A linear amplification is performed by adding only primer St3.2not, denaturing the DNA 3 minutes at 95°C, then running 10 cycles of [95°C, 30 seconds/48°C, 1 minute/72°C, 1 minute]. Primer PmeST5 is then added and the

following PCR reaction is run: 95°C, 3 minutes then 25 cycles of [95°C, 30 seconds/37°C, 40 seconds/72°C, 1 minute]. The resulting 1.1 kb band is isolated on a 1% agarose gel in TAE(0.04M Tris-acetate, 0.001M EDTA, pH 8.0) and cloned into EcoRV/NotI digested pBluescript SK-(Stratagene Cloning Systems, LaJolla,CA) to create pDM117. A 1.1 kb HindIII/NotI catalase fragment is gel purified from pDM117 and labelled in a PCR reaction containing digoxigenin-labeled deoxyuridine-triphosphate(dUTP) using the primers described above. Reaction conditions are: 95°C, 3 minutes, then 35 cycles of [95°C, 30 seconds/48°C, 1 minute/72°C, 1 minute]. This digoxigenin(DIG)labelled probe is used to screen genomic libraries.

B. Genomic library preparation

2. DNA Libraries and Identification of Catalase Clones

Genomic DNA libraries are constructed in the bacteriophage cloning vector λ -EMBL4(J.A.Sorge, in Vectors, A Survey of Molecular Cloning Vectors and Their Uses, Rodriguez et al., eds, pp.43-60, Butterworths, Boston, 1988). Briefly, total cellular DNA is partially digested with *Sau3A* and size-fractionated on low-melting point agarose gels. DNA fragments migrating between 9kb and 23 kb are excised and eluted from the gel using β -agarase (New England Biolabs, Beverly MA). The eluted DNA fragments are ligated with *Bam*HI-cleaved and dephosphorylated λ -EMBL4 vector arms, and the ligation mixtures are packaged using commercial packaging extracts (Stratagene, LaJolla, CA). The packaged DNA libraries are plated and amplified on *Escherichia coli* K802 cells.

C. Genomic library screening

Lambda phage are plated with *E. coli* K802 cells onto LB plates with NZY top agarose. Plaque lifts are made to nylon membranes(Hybond N, Amersham) using standard techniques (Sambrook et al., Molecular Cloning. A Laboratory Manual. 2nd ed. Cold Spring Harbor, 1989). DNA is bound to membranes by UV crosslinking. Filters are hybridized with the 1.1 kb DIG labelled probe described above. Hybridization and detection of catalase clones are performed using techniques described in the Boehringer Mannheim Genius™ System User's Guide. Hybridizations are performed at 65°C in 5XSSC, 0.1% L-lauroylsarcosine, 0.02% SDS, 1% blocking reagent for nucleic acid hybridization (Boehringer Mannheim). The concentration of DIG labelled probe used is 5 ng/ml hybridization solution. Hybridizing DNA is immunodetected with an alkaline phosphatase-conjugated anti-digoxigenin antibody and visualized with Lumiphos 530, a chemiluminescent

substrate(Boehringer Mannheim). DNA preparations are made from putative positive lambda clones using the Qiagen Lambda Midi kit(QIAGEN, Inc.).

D. RNA and cDNA Preparation

10 RNA and cDNA libraries are prepared from *S. thermophilum* strain A1065(ATCC 28085). Total RNA is prepared by extraction of ground frozen mycelia with guanidium thiocyanate followed by ultracentrifugation through a 5.7M CsCl cushion. Poly(A)+ RNA isolated by oligo(dT)-cellulose affinity chromatography is used to generate cDNA by the RNase H method using the hairpin modification method(Sambrook et al., Molecular Cloning: A Laboratory Manual, 1989; U.S. Serial No. 08/398,489, the contents of which are incorporated herein by reference). The cDNA is ligated to BstXI linkers and size
15 selected by agarose gel electrophoresis with a cutoff at 0.7 kb. The resulting cDNA is ligated into yeast expression plasmid pYES2(Invitrogen Corporation) and cloned into *E. coli* strain DH12S. The resulting library is divided into three pools with approximately 50,000 clones per pool.

E. cDNA Fragment Cloning

20 Three 125 ml shakeflasks containing 25 ml 2xYT(16 g/l bacto-tryptone, 10 g/l bacto-yeast extract, 5 g/l NaCl, pH 7.0), 50 µg/ml carbenicillin are inoculated with a loop full of frozen pYES2 library, one flask for each pool. Cultures are grown overnight at 37°C. Plasmid preps are made from 20 ml of culture using Qiagen tip 100 and standard Qiagen techniques. DNA is dissolved initially in 100 µl TE(10 mM Tris, pH 8.0, 1 mM EDTA, pH
25 8.0). DNA concentration is determined by absorbance at 260 nm and samples are diluted to a final concentration of 0.5 mg/ml. 100 µl PCR reactions using 250 ng DNA as template and 56 pmoles of gene specific primers(forward primer, 5' CCGCGGCGTCCTTTCTGA 3'; reverse primer, 5' GATCTTTGTCGGGTCCAG 3') are prepared. The PCR reaction buffer and Taq polymerase are obtained from Boehringer Mannheim. As a positive control a
30 reaction using 190 ng genomic *S. thermophilum* DNA is prepared. All reactions are amplified by denaturing at 95°C for 3 minutes followed by 30 cycles of [95°C, 30 seconds; 55°C, 1 minute; 72°C, 1 minute]. Subsequent PCR reactions used 250 ng pYES2 pool 5 DNA as template and the reaction conditions described above. Several combinations of catalase and pYES2 vector primers are used to generate five cDNA fragments which cover the entire
35 coding region of the gene(Figure 2). These fragments are subcloned into vector pCRII using Invitrogen's TA cloning kit.

F. Sequence Analysis

In general, nucleotide sequences are determined using Taq polymerase cycle sequencing with fluorescent labelled nucleotides(Giesecke et al., J. Virol. Methods 38: 47-60, 1992). Occasionally, reactions using bacteriophage T7 DNA polymerase(PRISM™ T7 Terminator Kit, Applied Biosystems) are employed to obtain clearer sequence in G/C rich regions. Reactions are analyzed using an Applied Biosystems automatic DNA sequencer (Model 363/A, version 1.2.0).

G. Transformation of *A. oryzae* host cells

Four wells of a 24 well tissue culture plate are filled with 500 µl YEG (5 g/l yeast extract, 20 g/l sucrose) + 1M sucrose medium. Each well is inoculated with 2-5 X 10⁶ conidia and the plate is incubated 16 hours at 37°C without shaking. The media is removed with a 1 ml pipet tip and 500 µl Novozym™ solution [OM buffer(1.2 M MgSO₄ buffered to pH 5.8 with 1M Na₂HPO₄), 5 mg/ml Novozym™ 234, 1 mg/ml BSA] is added. The fungal mat is dispersed by pipetting repeatedly with a cut-off 1 ml pipet tip. The plate is incubated at 34°C, 90 rpm until protoplasting is complete (approximately 1 hour). The protoplasts are transferred to a microfuge tube and washed three times with 1 ml STC(1.2 M sorbitol, 10 mM CaCl₂, 10 mM Tris-HCl, pH 7.5). The protoplasts are suspended at a concentration of 5 X 10⁷-1 X 10⁸/ml in STC:SPTC:DMSO(8:2:0.1).[STC, as above; SPTC:40% PEG 4000, 0.8M sorbitol, 50 mM CaCl₂, 50mM Tris-HCl, pH 8.0]

5 µg of plasmids pDM153(catalase expression plasmid) and pToC90(amdS selection) are added to 100 µl protoplasts in a Falcon tube(17 X 100mm size). 1 ml of SPTC is added. Samples are incubated 20 minutes at room temperature. Approximately 370 µl protoplast/SPTC mixture is added to 15 ml of overlay agarose that had been tempered to 50°C. Each transformation aliquot is plated onto a 150 mm plate of selective media that contains 1M sucrose. Acetamidase (amdS) positive transformants are selected on COVE minimal medium with acetamide as the sole nitrogen source. Transformants are incubated at 37°C.

H. Expression construct

The 5' end of the gene is modified using PCR to introduce a HindIII site directly upstream of the translational start. At the 3' end of the gene, an XbaI site is introduced directly after the termination codon using PCR. Three catalase genomic DNA fragments (HindIII/ApaI; ApaI/SalI; SalI/XbaI; are ligated together and inserted between the

A. oryzae TAKA amylase/TPI(triose phosphate isomerase)fusion promoter and the *A. niger* amyloglucosidase(AMG) terminator in plasmid pMT1560(described in WO 94/23022). The resulting expression construct, called pDM153, is illustrated in Figure 6.

I. Expression of catalase in *A. oryzae*

Transformants are transferred to a 24 well tissue culture plate containing 1 ml COVE agar per well and incubated at 37°C. Conidia are suspended in 0.5 ml 0.01% Tween 20. Twenty-four well plates containing 1 ml M400Da medium per well are inoculated with 15 µl conidia suspension. M400Da medium contains maltodextrin 50.0g/l; MgSO₄·7H₂O, 2.0 g/l; KH₂PO₄, 2.0 g/l; citric acid, 4.0 g/l; yeast extract, 8.0 g/l, urea, 2.0 g/l; trace metal solution I, 0.5ml/l[trace metal solution: ZnSO₄·7H₂O 14.3 g/l; CuSO₄·5H₂O, 2.5 g/l; NiCl₂·6H₂O, 0.5 g/l; FeSO₄·7H₂O, 13.8 g/l; MnSO₄·H₂O, 8.5 g/l; citric acid, 3.0 g/l]; pH adjusted to 6.0 with 5N NaOH. These plates are incubated at 37° C, 100 rpm for 3 days in a lexan culture box to maintain the humidity level and minimize loss of culture volume. Shakeflask cultures are prepared by adding 150 µl spore purified conidia stock to 25 ml M400Da medium in 125 ml polypropylene flasks. Cultures are incubated 5 days at 37°C, 100 rpm. supernatants are recovered and assayed for catalase activity.

J. Catalase assays

Catalase mediated degradation of hydrogen peroxide is measured by a modification of the titanium color method of Patti and Bonet-Maury(Bull. Soc. Chem. Biol. 35: 1177, 1953). The titanium reagent is made by mixing 1 g TiO₂ and 10 g K₂SO₄, digesting on a mantle heater for 2-3 hours at 180-220°C in 150 ml of concentrated H₂SO₄, cooling, and then diluting to 1.5 l with deionized water. Stock 16 mM H₂O₂ solution is prepared in 10 mM Na phosphate buffer, pH 7.0. The assay consists of mixing 25 µl of sample and 25 µl of the H₂O₂ solution in microtiter wells (96 well plate) and incubating for 5 minutes at room temperature. 200 µl of the titanium reagent is added and the absorbance at 405 nm is read using a Thermomax microplate reader from Molecular Devices. Standard *A. niger* catalase(Sigma) is diluted to concentrations ranging from 0.2 CIU/ml to 3.0 CIU/ml in 10 mM Na phosphate buffer, pH 7.0. One CIU causes decomposition of one µmole of H₂O₂ per minute under the conditions described above.

Catalase is also analyzable by UV. Supernatants from fungal cultures are centrifuged in a microfuge at maximum speed for 5 minutes. One ml of phosphate buffer(50 mM KPO₄ buffer, pH 7.0) is transferred to a reference cuvette. The catalase standard is an

5 *Aspergillus niger* catalase. Reaction solution is prepared by adding 0.09 ml 30% H₂O₂ stock solution(SIGMA) to 50 ml phosphate buffer, the absorbance read at 240nm, and the absorbance adjusted to 0.5 to 0.55. Catalase-containing supernatants are diluted with phosphate buffer to yield an approximate concentration of 100 to 300 CIU/ml. 25 µl of the diluted supernatant is added to 975 µl of reaction solution, mixed and ΔA_{240nm} is recorded.
10 Readings are taken, on a Shimadzu dual beam spectrophotometer (Shimadzu Scientific Instruments, Inc., Columbia, MD) for at least 15 seconds at one second intervals, and the $\Delta A_{240}/\text{minute}$ is calculated. The activity in CIU/ml is calculated according to the following formula:
15 $\text{CIU/ml} = \Delta A_{240}/\text{minute} \times 23 \times \text{dilution factor(s)}$. One CIU is defined as the amount of catalase which will decompose one µmole of H₂O₂ per minute in 50 mM KPO₄ buffer(pH 7.0) at 25°C, at a starting H₂O₂ concentration of 11.5 to 12.0 mM.

II. RESULTS AND DISCUSSION

A. Library Screening

20 (1). cDNA. DNA preparations from the pYES2 library is screened by PCR for the presence of the catalase gene. Exact match primers(forward primer: 5' CCGCGGCGTCCTTTCTGA 3'; reverse primer: 5' GATCTTTGTCGGGTCCAG 3')yields a 0.6 kb catalase fragment from the pYES2 library DNA. Primers which are used to sequence genomic clones are used as PCR primers to generate five small fragments of cDNA from the
25 pYES2 library. Figure 2 shows the cDNA subclones obtained. After subcloning into a TA vector, these fragments are sequenced using M13 and catalase specific primers. The cDNA sequence obtained verifies the genomic sequence and locations of introns.

(2). Genomic DNA. Approximately 40,000 phage are plated from the *S. thermophilum* A 1065 genomic EMBL4 library. DNA preparations are made from six out
30 of seven putative positive lambda clones using the Qiagen Lambda Midi Kit(QUIAGEN, Inc.). These preps are digested with restriction enzymes and subjected to Southern blot analysis. The blot is probed with the 1.1 kb DIG labelled catalase fragment using the hybridization conditions described above. Three independent clones are identified. A hybridizing 6.3 kb EcoRI fragment includes the entire 2.8 kb coding region of the catalase gene, 2.1 kb of
35 upstream (promoter) sequence and approximately 1.4 kb of 3' untranslated sequence. This EcoRI fragment is subcloned into pUC118 to create plasmid pDM125.

B. DNA Sequence Analysis

5 The cloned *S. thermophilum* catalase is encoded by 2791 base pairs containing 7 introns of length 131, 79, 76, 153, 53, 54, and 94 base pairs. The translated sequence is shown in Figure 4. The predicted amino acid polypeptide of the prepro-catalase is 717 residues. Prepro-catalase consists of a putative 19 amino acid signal sequence, based on the rules of von Heijne(Nucl. Acids Res. 14: 4683-4690, 1986). Amino acid sequence analysis performed on the native protein indicates that the N-terminus of the mature protein begins at amino acid 37, showing that the protein also contains a putative 17 amino acid proregion. The catalase contains three potential N-linked glycosylation sites located in the mature region of the polypeptide. The predicted molecular weight of the mature catalase (without glycosylation) is 75 kD. A comparison of the DNA translated amino acid sequence to the peptides derived from protein sequencing is shown in Figure 5. The correlations are good, and any differences observed may be due to ambiguities in protein sequencing.

10 When compared with known catalase sequences, the DNA sequence of the *S. thermophilum* catalase cloned from strain A1065 shows the closest homology with the *A. niger* catR catalase; the nucleic acid sequence is 61% identical to the DNA sequence of the *A. niger* catalase R gene. The two catalases are 57% identical on the amino acid level. However, the organization of these two genes is quite different. The *S. thermophilum* catalase coding region is interrupted by seven introns, whereas the *A. niger* catR gene contains only four introns. The first two introns are in similar positions in the two genes, but none of the remaining introns are conserved.

C. Expression in *A. oryzae*

20 An *A. oryzae* strain is cotransformed with catalase construct pDM153 and plasmid pToC90(amdS selection) as described above. Thirty-six transformants are grown in M400Da medium in 24 well plates for 3 days. Supernatants are collected and assayed using the titanium color method. Several of the transformants exhibit significantly more catalase activity than the untransformed control culture. The highest expresser makes approximately 1200 CIU/ml. The four best producers are spore purified twice. Shake flasks containing M400Da medium are inoculated and grown five days at 37°C, 100 rpm. Samples are taken at 4 and 5 days and assayed using the titanium color method. The best producer in shake flasks makes 12,800 CIU/ml.

The best producer is used to express the *Scytalidium* catalase in a fermentation medium, using a fed batch process in a 3 liter laboratory fermentor. The fermentation media have the following compositions:

10	Tanks:	Nutriose 725*	30g/l (autoclaved separately)
		(NH ₄) ₂ HPO ₄	5g/l
		yeast extract	5g/l
		MgSO ₄ ·7H ₂ O	2g/l
		KH ₂ PO ₄	2g/l
		citric acid·H ₂ O	4g/l
15		K ₂ SO ₄	3g/l
		CaCl ₂ ·2H ₂ O	2g/l
		trace metal sol.**	0.5 ml/l
		pluronic	1.0 ml/l
		hemoglobin	20g/l
20			

*maltose syrup (Roquette Corporation, Gurnee, IL)

**ZnSO₄·7H₂O, 14.3g/l; CuSO₄·5H₂O, 2.5g/l; NiCl₂·6H₂O, 0.5 g/l; FeSO₄·7H₂O, 13.8 g/l; MnSO₄·H₂O, 8.5 g/l; citric acid, 3.0 g/l

Hemoglobin is dissolved separately in 0.1 N NaOH. Salts are added while maintaining the medium at pH 10.5. The volume is brought to 1.8 l + 5% with tap water, and autoclaved 60 minutes at 125°C. The pH is adjusted to 7.6 with H₃PO₄.

30	Feed:	Nutriose 725	400 g/kg
		citric acid·H ₂ O	1 g/kg
		yeast extract	5 g/l(autoclaved separately)
		urea	10 g/kg
		pluronic	1 ml/kg

The mixture is adjusted to 1.8 kg + 5% with tap water and the pH adjusted to 4.5 before autoclaving.

To produce the inoculum, one ml of spore suspension (approximately 10⁷ spores) of the strain described above is added to 100 ml of MY50 medium (50 g/l

5 maltodextrin, 2g/l MgSO₄·7H₂O, 10g/l KH₂PO₄, 2g/l K₂SO₄, 2g/l citric acid, 10 g/l yeast
extract, 0.5 ml trace metals, 2.0 g/l urea) in a 500 ml polypropylene flask and the culture
grown for 24 hours at 34°C, at 200 rpm. 60 ml of shakeflask culture is added to 1.8 l
medium in the 3 liter fermentor, and fermentation carried out at 34°C pH 7.6(controlled with
NaOH, H₃PO₄), DO>20%(controlled by agitation of 800-1200 rpm) with aeration of 1 vvm.
10 The feed is started when DO has dropped, at approximately 20 hours. The feed is started at
8g/tank/hour, and raised to 12g/tank/hour dependent on DO. Under these conditions, yield of
catalase increases significantly relative to the yield obtained in shake flasks.

Deposit of Biological Materials

15

The following biological material has been deposited under the terms of the
Budapest Treaty with the Agricultural Research Service Patent Culture Collection, Northern
Regional Research Center, 1815 University Street, Peoria, Illinois, 61604 and given the
following accession number:

20

<u>Deposit</u>	<u>Accession Number</u>
<i>E. coli</i> containing pDM125 (a 6.3 kb EcoRI genomic fragment cloned into a pUC118 25 plasmid)	NRRL B-21426

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT:
 - (A) Novo Nordisk Biotech, Inc.
 - (B) STREET: 1445 Drew Avenue
 - (C) CITY: Davis
 - (D) STATE: California
 - (E) COUNTRY: US
 - (F) ZIP: 95616-4880
 - (G) TELEPHONE: (916) 757-8100
 - (H) TELEFAX: (916) 758-0317
- (ii) TITLE OF INVENTION: SCYTALIDUM CATALASE GENE
- (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Novo Nordisk of North America, Inc.
 - (B) STREET: 405 Lexington Avenue, 64th Floor
 - (C) CITY: New York
 - (D) STATE: New York
 - (E) COUNTRY: USA
 - (F) ZIP: 10174-6401
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: to be assigned
 - (B) FILING DATE: 01-MAY-1996
 - (C) CLASSIFICATION:
- (vi) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/435,925
 - (B) FILING DATE: 05-MAY-1995
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Agris, Cheryl H.
 - (B) REGISTRATION NUMBER: 34,086
 - (C) REFERENCE/DOCKET NUMBER: 4429.204-WO
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 212-867-0123
 - (B) TELEFAX: 212-878-9655

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2794 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ix) FEATURE:
 - (A) NAME/KEY: intron
 - (B) LOCATION: 283..413
- (ix) FEATURE:

(A) NAME/KEY: intron
 (B) LOCATION: 618..696

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION: 718..793

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION: 956..1108

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION: 1166..1218

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION: 1789..1842

(ix) FEATURE:
 (A) NAME/KEY: intron
 (B) LOCATION: 2671..2764

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: join(1..282, 414..617, 697..717, 794..955, 1109
 ..1165, 1219..1788, 1843..2670, 2765..2791)

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

ATG AAC AGA GTC ACG AAT CTC CTC GCC TGG GCC GGC GCG ATA GGG CTC	48
Met Asn Arg Val Thr Asn Leu Leu Ala Trp Ala Gly Ala Ile Gly Leu	
1 5 10 15	
GCC CAA GCA ACA TGT CCC TTT GCG GAC CCT GCC GCT CTG TAT AGT CGT	96
Ala Gln Ala Thr Cys Pro Phe Ala Asp Pro Ala Ala Leu Tyr Ser Arg	
20 25 30	
CAA GAT ACT ACC AGC GGC CAG TCG CCA CTT GCA GCA TAC GAG GTG GAT	144
Gln Asp Thr Thr Ser Gly Gln Ser Pro Leu Ala Ala Tyr Glu Val Asp	
35 40 45	
GAC AGC ACC GGA TAC CTG ACC TCC GAT GTT GGC GGG CCC ATT CAG GAC	192
Asp Ser Thr Gly Tyr Leu Thr Ser Asp Val Gly Gly Pro Ile Gln Asp	
50 55 60	
CAG ACC AGC CTC AAG GCA GGC ATC CGG GGT CCG ACC CTT CTT GAG GAC	240
Gln Thr Ser Leu Lys Ala Gly Ile Arg Gly Pro Thr Leu Leu Glu Asp	
65 70 75 80	
TTT ATG TTC CGC CAG AAG ATC CAG CAC TTC GAC CAT GAA CGG	282
Phe Met Phe Arg Gln Lys Ile Gln His Phe Asp His Glu Arg	
85 90	
GTAAGGACAT AATGCTCACA CGAGCGGCTG CGTGCCCACC TATTTCCGAG ACATTGGGCT	342
GGCTGGCTGG CTGTGACTGC TTGAGTTTGG GGACATACGG AGTACCTTAC TGACGCGCTG	402
AACCACTCCA G GTT CCC GAA AGG GCG GTC CAT GCT CGA GGC GCT GGA GCA	452
Val Pro Glu Arg Ala Val His Ala Arg Gly Ala Gly Ala	
95 100 105	
CAC GGG ACC TTC ACG AGT TAC GCC GAC TGG AGT AAC ATC ACC GCG GCG	500
His Gly Thr Phe Thr Ser Tyr Ala Asp Trp Ser Asn Ile Thr Ala Ala	
110 115 120	
TCC TTT CTG AAC GCC ACT GGA AAG CAG ACG CCG GTG TTT GTC CGG TTC	548

Ser Phe Leu Asn Ala Thr Gly Lys Gln Thr Pro Val Phe Val Arg Phe
 125 130 135

TCG ACC GTT GCT GGG TCT CGA GGG AGC GCA GAC ACG GCG AGA GAC GTT 596
 Ser Thr Val Ala Gly Ser Arg Gly Ser Ala Asp Thr Ala Arg Asp Val
 140 145 150 155

CAT GGT TTC GCG ACG CGG TTT GTAAGTTTTG TTGTGTTTCA TTCGTTCCGG 647
 His Gly Phe Ala Thr Arg Phe
 160

TCTGTAGAGG AGGGTTAGGA TATGAGCTAA CGTGTGTGTG TGTGTGAAG TAC ACT 702
 Tyr Thr

GAT GAA GGC AAC TTT GTACGTCCCA CGCATGGTCC TCAATTCTCT TATCTGGCAG 757
 Asp Glu Gly Asn Phe
 165

CCATGTGGTC ATTGTCGACG TTGCTAACTT GCGTAG GAT ATC GTC GGA AAC AAC 811
 Asp Ile Val Gly Asn Asn
 170 175

ATC CCG GTA TTC TTC ATT CAA GAT GCA ATC CAG TTC CCT GAC CTT ATC 859
 Ile Pro Val Phe Phe Ile Gln Asp Ala Ile Gln Phe Pro Asp Leu Ile
 180 185 190

CAC TCG GTC AAG CCG CGT CCC GAC AAC GAG ATT CCC CAA GCG GCG ACG 907
 His Ser Val Lys Pro Arg Pro Asp Asn Glu Ile Pro Gln Ala Ala Thr
 195 200 205

GCT CAT GAT TCA GCT TGG GAC TTC TTC AGC CAG CAG CCA AGC ACC ATG 955
 Ala His Asp Ser Ala Trp Asp Phe Ser Gln Gln Pro Ser Thr Met
 210 215 220

GTAAGCAATG GACCAAGGAG CCGCACCTGG GGTGACATGC CAGGGAGTAC ACAAGGCGTT 1015

CCGATGACCC TCGTGTGACC AAGGCAGTAC AACACTCCAC GGAGGACTCG AAGAGATTCCG 1075

GCAATATGGA ACACAGAACT GACAGGATGG TAG CAC ACG TTG TTC TGG GCC ATG 1129
 His Thr Leu Phe Trp Ala Met
 225 230

TCC GGC CAC GGA ATC CCT CGC AGC TAC CGC CAT ATG GTACGTTTGC 1175
 Ser Gly His Gly Ile Pro Arg Ser Tyr Arg His Met
 235 240

CTGGCTGAGA TGACCGTGAA TCCATTCTA ACCTCAAGCC CAG GAT GGC TTC GGC 1230
 Asp Gly Phe Gly
 245

GTC CAC ACG TTC CGG TTT GTC AAA GAT GAC GGC TCG TCC AAG TTG ATC 1278
 Val His Thr Phe Arg Phe Val Lys Asp Asp Gly Ser Ser Lys Leu Ile
 250 255 260

AAG TGG CAT TTC AAG TCA CGC CAG GGA AAG GCG AGT CTA GTC TGG GAA 1326
 Lys Trp His Phe Lys Ser Arg Gln Gly Lys Ala Ser Leu Val Trp Glu
 265 270 275

GAG GCG CAG GTT CTT TCT GGC AAG AAT GCC GAC TTC CAC CGT CAG GAC 1374
 Glu Ala Gln Val Leu Ser Gly Lys Asn Ala Asp Phe His Arg Gln Asp
 280 285 290

CTC TGG GAT GCT ATT GAG TCC GGG AAC GGA CCA GAA TGG GAT GTC TGC 1422
 Leu Trp Asp Ala Ile Glu Ser Gly Asn Gly Pro Glu Trp Asp Val Cys
 295 300 305 310

GTC Val	CAG Gln	ATT Ile	GTC Val	GAT Asp	GAG Glu	TCC Ser	CAG Gln	GCG Ala	CAA Gln	GCC Ala	TTT Phe	GGC Gly	TTC Phe	GAC Asp	TTG Leu	1470
				315						320		325				
CTG Leu	GAC Asp	CCG Pro	ACA Thr	AAG Lys	ATC Ile	ATC Ile	CCC Pro	GAG Glu	GAG Glu	TAC Tyr	GCC Ala	CCC Pro	TTG Leu	ACG Thr	AAG Lys	1518
				330						335		340				
CTG Leu	GGC Gly	CTC Leu	TTG Leu	AAG Lys	CTG Leu	GAT Asp	CGC Arg	AAT Asn	CCG Pro	ACC Thr	AAC Asn	TAC Tyr	TTC Phe	GCC Ala	GAG Glu	1566
				345						350		355				
ACG Thr	GAG Glu	CAG Gln	GTC Val	ATG Met	TTC Phe	CAA Gln	CCC Pro	GGT Gly	CAT His	ATC Ile	GTC Val	CGC Arg	GGC Gly	ATC Ile	GAC Asp	1614
				360						370		375				
TTC Phe	ACG Thr	GAG Glu	GAT Asp	CCC Pro	CTG Leu	CTA Leu	CAG Gln	GGA Gly	CGC Arg	CTC Leu	TTT Phe	TCG Ser	TAC Tyr	CTT Leu	GAC Asp	1662
				375						380		385		390		
ACG Thr	CAG Gln	CTG Leu	AAC Asn	CGG Arg	AAT Asn	GGC Gly	GGG Gly	CCC Pro	AAC Asn	TTT Phe	GAG Glu	CAG Gln	CTG Leu	CCC Pro	ATC Ile	1710
				395						400		405				
AAC Asn	ATG Met	CCG Pro	CGG Arg	GTG Val	CCG Pro	ATT Ile	CAC His	AAC Asn	AAT Asn	AAT Asn	CGC Arg	GAC Asp	GGC Gly	GCC Ala	GGC Gly	1758
				410						415		420				
CAG Gln	ATG Met	TTC Phe	ATC Ile	CAC His	AGG Arg	AAC Asn	AAG Lys	TAT Tyr	CCT Pro	GTAAGTGCCT			CTTTTGCTC			1808
				425						430						
GATCGTTGTG GTGCCGGCTT GCTGACAGAC GCAG TAC ACT CCC AAC ACC CTG																1860
Tyr Thr Pro Asn Thr Leu																
435																
AAC Asn	AGT Ser	GGT Gly	TAT Tyr	CCG Pro	CGG Arg	CAA Gln	GCC Ala	AAC Asn	CAA Gln	AAT Asn	GCC Ala	GGA Gly	CGC Arg	GGA Gly	TTC Phe	1908
				440						445		450				
TTC Phe	ACA Thr	GCG Ala	CCT Pro	GGC Gly	CGT Arg	ACC Thr	GCC Ala	AGC Ser	GGT Gly	GCC Ala	CTC Leu	GTC Val	CGT Arg	GAG Glu	GTG Val	1956
				455						460		465		470		
TCG Ser	CCA Pro	ACA Thr	TTC Phe	AAC Asn	GAC Asp	CAC His	TGG Trp	TCG Ser	CAG Gln	CCC Pro	CGT Arg	CTC Leu	TTC Phe	TTC Phe	AAC Asn	2004
				475						480		485				
TCC Ser	CTC Leu	ACT Thr	CCC Pro	GTC Val	GAA Glu	CAA Gln	CAG Gln	TTC Phe	CTC Leu	GTC Val	AAC Asn	GCC Ala	ATG Met	CGC Arg	TTC Phe	2052
				490						495		500				
GAA Glu	ATC Ile	AGC Ser	CTT Leu	GTG Val	AAG Lys	TCG Ser	GAA Glu	GAA Glu	GTC Val	AAG Lys	AAG Lys	AAC Asn	GTG Val	CTC Leu	ACC Thr	2100
				505						510		515				
CAG Gln	CTC Leu	AAC Asn	CGC Arg	GTC Val	AGC Ser	CAT His	GAC Asp	GTG Val	GCC Ala	GTG Val	CGC Arg	GTG Val	GCC Ala	GCC Ala	GCT Ala	2148
				520						525		530				
ATC Ile	GGC Gly	CTC Leu	GGC Gly	GCG Ala	CCC Pro	GAC Asp	GCG Ala	GAC Asp	GAC Asp	ACA Thr	TAC Tyr	TAC Tyr	CAC His	AAC Asn	AAC Asn	2196
				535						540		545		550		
AAG Lys	ACG Thr	GCT Ala	GGC Gly	GTC Val	TCA Ser	ATC Ile	GTT Val	GGA Gly	AGC Ser	GGG Gly	CCC Pro	TTG Leu	CCT Pro	ACC Thr	ATC Ile	2244
				555						560		565				

Asp Ser Thr Gly Tyr Leu Thr Ser Asp Val Gly Gly Pro Ile Gln Asp
 50 55 60
 Gln Thr Ser Leu Lys Ala Gly Ile Arg Gly Pro Thr Leu Leu Glu Asp
 65 70 75 80
 Phe Met Phe Arg Gln Lys Ile Gln His Phe Asp His Glu Arg Val Pro
 85 90 95
 Glu Arg Ala Val His Ala Arg Gly Ala Gly Ala His Gly Thr Phe Thr
 100 105 110
 Ser Tyr Ala Asp Trp Ser Asn Ile Thr Ala Ala Ser Phe Leu Asn Ala
 115 120 125
 Thr Gly Lys Gln Thr Pro Val Phe Val Arg Phe Ser Thr Val Ala Gly
 130 135 140
 Ser Arg Gly Ser Ala Asp Thr Ala Arg Asp Val His Gly Phe Ala Thr
 145 150 155 160
 Arg Phe Tyr Thr Asp Glu Gly Asn Phe Asp Ile Val Gly Asn Asn Ile
 165 170 175
 Pro Val Phe Phe Ile Gln Asp Ala Ile Gln Phe Pro Asp Leu Ile His
 180 185 190
 Ser Val Lys Pro Arg Pro Asp Asn Glu Ile Pro Gln Ala Ala Thr Ala
 195 200 205
 His Asp Ser Ala Trp Asp Phe Phe Ser Gln Gln Pro Ser Thr Met His
 210 215 220
 Thr Leu Phe Trp Ala Met Ser Gly His Gly Ile Pro Arg Ser Tyr Arg
 225 230 235 240
 His Met Asp Gly Phe Gly Val His Thr Phe Arg Phe Val Lys Asp Asp
 245 250 255
 Gly Ser Ser Lys Leu Ile Lys Trp His Phe Lys Ser Arg Gln Gly Lys
 260 265 270
 Ala Ser Leu Val Trp Glu Glu Ala Gln Val Leu Ser Gly Lys Asn Ala
 275 280 285
 Asp Phe His Arg Gln Asp Leu Trp Asp Ala Ile Glu Ser Gly Asn Gly
 290 295 300
 Pro Glu Trp Asp Val Cys Val Gln Ile Val Asp Glu Ser Gln Ala Gln
 305 310 315 320
 Ala Phe Gly Phe Asp Leu Leu Asp Pro Thr Lys Ile Ile Pro Glu Glu
 325 330 335
 Tyr Ala Pro Leu Thr Lys Leu Gly Leu Leu Lys Leu Asp Arg Asn Pro
 340 345 350
 Thr Asn Tyr Phe Ala Glu Thr Glu Gln Val Met Phe Gln Pro Gly His
 355 360 365
 Ile Val Arg Gly Ile Asp Phe Thr Glu Asp Pro Leu Leu Gln Gly Arg
 370 375 380
 Leu Phe Ser Tyr Leu Asp Thr Gln Leu Asn Arg Asn Gly Gly Pro Asn
 385 390 395 400
 Phe Glu Gln Leu Pro Ile Asn Met Pro Arg Val Pro Ile His Asn Asn

				405					410					415		
Asn	Arg	Asp	Gly	Ala	Gly	Gln	Met	Phe	Ile	His	Arg	Asn	Lys	Tyr	Pro	
			420					425					430			
Tyr	Thr	Pro	Asn	Thr	Leu	Asn	Ser	Gly	Tyr	Pro	Arg	Gln	Ala	Asn	Gln	
		435					440					445				
Asn	Ala	Gly	Arg	Gly	Phe	Phe	Thr	Ala	Pro	Gly	Arg	Thr	Ala	Ser	Gly	
	450					455					460					
Ala	Leu	Val	Arg	Glu	Val	Ser	Pro	Thr	Phe	Asn	Asp	His	Trp	Ser	Gln	
465					470					475					480	
Pro	Arg	Leu	Phe	Phe	Asn	Ser	Leu	Thr	Pro	Val	Glu	Gln	Gln	Phe	Leu	
				485					490					495		
Val	Asn	Ala	Met	Arg	Phe	Glu	Ile	Ser	Leu	Val	Lys	Ser	Glu	Glu	Val	
			500					505					510			
Lys	Lys	Asn	Val	Leu	Thr	Gln	Leu	Asn	Arg	Val	Ser	His	Asp	Val	Ala	
		515					520					525				
Val	Arg	Val	Ala	Ala	Ala	Ile	Gly	Leu	Gly	Ala	Pro	Asp	Ala	Asp	Asp	
	530					535					540					
Thr	Tyr	Tyr	His	Asn	Asn	Lys	Thr	Ala	Gly	Val	Ser	Ile	Val	Gly	Ser	
545					550					555					560	
Gly	Pro	Leu	Pro	Thr	Ile	Lys	Thr	Leu	Arg	Val	Gly	Ile	Leu	Ala	Thr	
				565					570					575		
Thr	Ser	Glu	Ser	Ser	Ala	Leu	Asp	Gln	Ala	Ala	Gln	Leu	Arg	Thr	Arg	
			580					585					590			
Leu	Glu	Lys	Asp	Gly	Leu	Val	Val	Thr	Val	Val	Ala	Glu	Thr	Leu	Arg	
		595					600					605				
Glu	Gly	Val	Asp	Gln	Thr	Tyr	Ser	Thr	Ala	Asp	Ala	Thr	Gly	Phe	Asp	
	610					615					620					
Gly	Val	Val	Val	Val	Asp	Gly	Ala	Ala	Ala	Leu	Phe	Ala	Ser	Thr	Ala	
625					630					635					640	
Ser	Ser	Pro	Leu	Phe	Pro	Thr	Gly	Arg	Pro	Leu	Gln	Ile	Phe	Val	Asp	
				645					650					655		
Ala	Tyr	Arg	Trp	Gly	Lys	Pro	Val	Gly	Val	Cys	Gly	Gly	Lys	Ser	Ser	
			660					665					670			
Glu	Val	Leu	Asp	Ala	Ala	Asp	Val	Pro	Glu	Asp	Gly	Asp	Gly	Val	Tyr	
		675					680					685				
Ser	Glu	Glu	Ser	Val	Asp	Met	Phe	Val	Glu	Glu	Phe	Glu	Lys	Gly	Leu	
	690					695					700					
Ala	Thr	Phe	Arg	Phe	Thr	Asp	Arg	Phe	Ala	Leu	Asp	Ser				
705					710					715						

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ix) FEATURE:
 - (A) NAME/KEY: N=deoxyinosine
 - (B) LOCATION: 24

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

AGCACTAGTT TAAACTAYGA RGTNGAYGA 29

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

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

TATTGGATAT AAGCGGCCGC CCYTGNCGYT TYTTGAARTG 40

(2) INFORMATION FOR SEQ ID NO:5:

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

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

CCGCGGCGTC CTTTCTGA 18

(2) INFORMATION FOR SEQ ID NO:6:

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

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

GATCTTTGTC GGGTCCAG 18

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

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

Ser Gly Gln Ser Pro Leu Ala Ala Tyr Glu Val Asp Asp Ser Asp Gly
 1 5 10 15
 Tyr

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

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

Ser Gly Gln Ser Pro Leu Ala Ala Tyr Glu Val Asp Asp Ser Thr Gly
 1 5 10 15

Tyr

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 8 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ala Asp Trp Ser Ile Thr Ala Ala
 1 5

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 9 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

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

Ala Asp Trp Ser Asn Ile Thr Ala Ala
 1 5

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 9 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

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

Arg Asn Pro Thr Asn Tyr Phe Ala Glu
 1 5

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ala Tyr Arg Trp Gly Lys Pro Val Gly Val Xaa Gly Gly Lys Arg Arg
 1 5 10 15

Glu

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ala Tyr Arg Trp Gly Lys Pro Val Gly Val Cys Gly Gly Lys Ser Ser
 1 5 10 15

Glu

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 11 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ser Phe Ala Leu Asp Gln Ala Ala Gln Leu Arg
 1 5 10

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 8 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Arg Phe Glu Ile Ser Leu Val Lys
 1 5

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 9 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Asn Val Leu Thr Gln Leu Asn Arg Val
 1 5

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

5 What we claim is:

1. A nucleic acid construct containing a sequence encoding a *Scytalidium* catalase.
2. The construct of Claim 1 which comprises a sequence encoding a *Scytalidium*
10 *thermophilum* catalase.
3. The construct of Claim 1 which comprises a sequence encoding the amino acid sequence depicted in SEQ ID NO. 2.
- 15 4. The construct of Claim 1, which comprises the nucleic acid sequence depicted in SEQ ID NO. 1.
5. The construct of Claim 1, which comprises the nucleic acid sequence contained in NRRL B-21426.
- 20 6. A recombinant vector comprising a nucleic acid construct containing a sequence encoding a *Scytalidium* catalase.
7. The vector of Claim 6 in which the sequence is operably linked to a promoter sequence.
- 25 8. The vector of Claim 7 in which the promoter is a fungal or yeast promoter.
9. The vector of Claim 8 in which the promoter is the TAKA amylase promoter or the triose phosphate isomerase promoter of *Aspergillus oryzae*, or a fusion thereof.
- 30 10. The vector of Claim 7 which also comprises a selectable marker.
11. The vector of Claim 10 in which the selectable marker is selected from the group consisting of *amdS*, *pyrG*, *argB*, *niaD*, *sC*, and *hygB*.

35

12. The vector of Claim 11 in which the selectable marker is the amdS marker of *Aspergillus nidulans* or *Aspergillus oryzae*, or the pyrG marker of *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus awamori*, or *Aspergillus oryzae*.

10 13. The vector of Claim 9 which comprises both the TAKA amylase promoter of *Aspergillus oryzae* and the amdS or pyrG marker of *Aspergillus nidulans* or *Aspergillus oryzae*.

14. A recombinant host cell comprising a heterologous nucleic acid construct containing a nucleic acid sequence encoding a *Scytalidium* catalase.

15 15. The host cell of Claim 14 which is a fungal cell.

16. The host cell of Claim 15 which is an *Aspergillus* cell.

17. The host cell of Claim 14 in which the construct is integrated into the host cell genome.

20

18. The host cell of Claim 14 in which the construct is contained on a vector.

19. The host cell of Claim 14 which comprises a construct containing a sequence encoding the amino acid sequence depicted in SEQ ID NO. 2.

25

20. A method for obtaining a catalase enzyme which comprises culturing, in a culture medium, a host cell comprising a nucleic acid construct containing a sequence encoding a *Scytalidium* catalase enzyme, under conditions conducive to expression of the enzyme, and recovering the enzyme from the culture.

30

21. A method according to claim 20 in which the catalase is secreted into the culture medium.

35

22. A *Scytalidium* enzyme obtained by culturing a host cell comprising a nucleic acid construct containing a sequence encoding a *Scytalidium* catalase enzyme, under conditions conducive to expression of the enzyme, and recovering the enzyme from the culture.

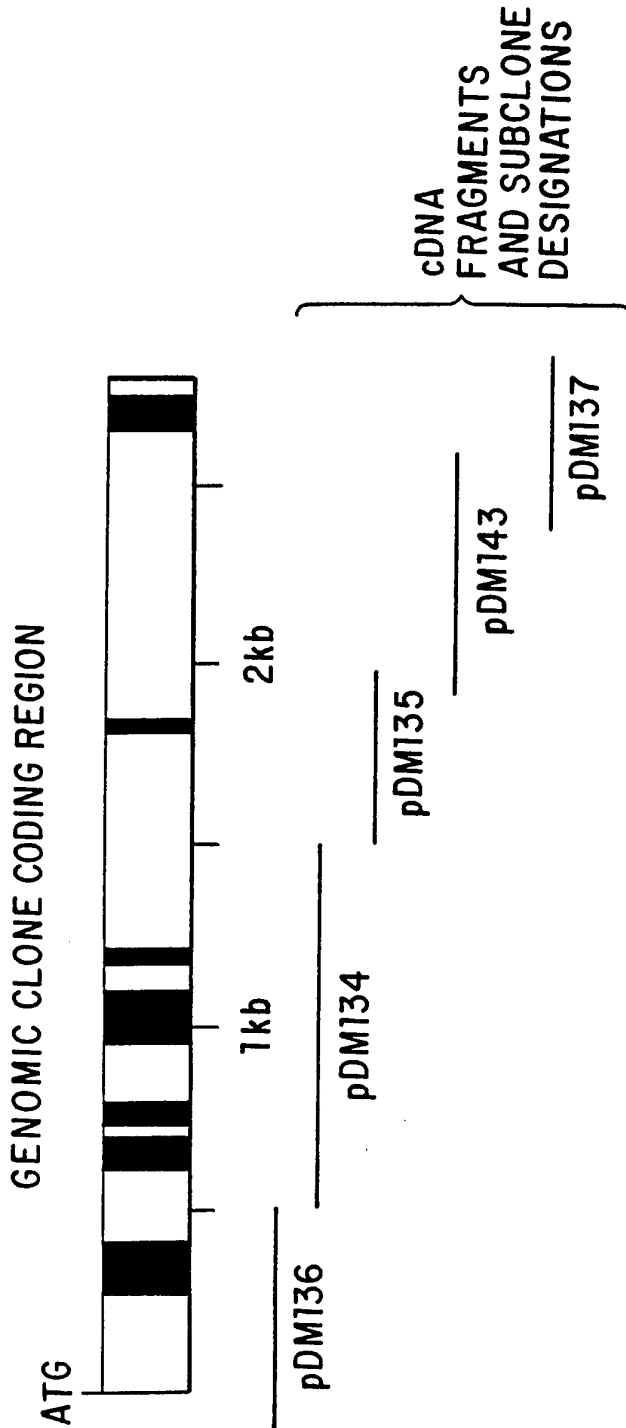


FIG. 2

3/12

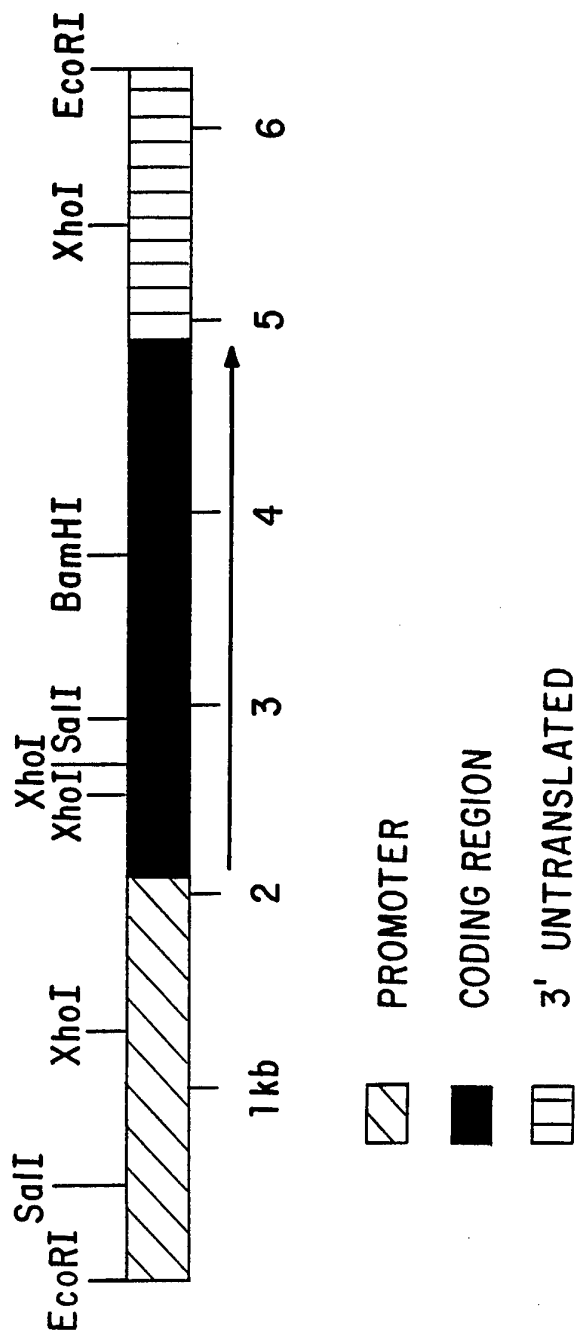


FIG. 3

4/12

	9	18	27	36	45	54														
ATG	AAC	AGA	GTC	ACG	AAT	CTC	CTC	GCC	TGG	GCC	GGC	GCG	ATA	GGG	CTC	GCC	CAA			
M	N	R	V	T	N	L	L	A	W	A	G	A	I	G	L	A	Q			
	63	72	81	90	99	108														
GCA	ACA	TGT	CCC	TTT	GCG	GAC	CCT	GCC	GCT	CTG	TAT	AGT	CGT	CAA	GAT	ACT	ACC			
A	T	C	P	F	A	D	P	A	A	L	Y	S	R	Q	D	T	T			
	117	126	135	144	153	162														
AGC	GGC	CAG	TCG	CCA	CTT	GCA	GCA	TAC	GAG	GTG	GAT	GAC	AGC	ACC	GGA	TAC	CTG			
S	G	Q	S	P	L	A	A	Y	E	V	D	D	S	T	G	Y	L			
	171	180	189	198	207	216														
ACC	TCC	GAT	GTT	GGC	GGG	CCC	ATT	CAG	GAC	CAG	ACC	AGC	CTC	AAG	GCA	GGC	ATC			
T	S	D	V	G	G	P	I	Q	D	Q	T	S	L	K	A	G	I			
	225	234	243	252	261	270														
CGG	GGT	CCG	ACC	CTT	CTT	GAG	GAC	TTT	ATG	TTC	CGC	CAG	AAG	ATC	CAG	CAC	TTC			
R	G	P	T	L	L	E	D	F	M	F	R	Q	K	I	Q	H	F			
	279	292	302	312	322	332														
GAC	CAT	GAA	CGG	GTAAGGACAT			AATGCTCACA			CGAGCGGCTG			CGTGCCCACC			TATTTCCGAG				
D	H	E	R																	
	342	352	362	372	382	392	402													
ACATTGGGCT	GGCTGGCTGG	CTGTGACTGC	TTGAGTTTGG	GGACATACGG	AGTACCTTAC	TGACGCGCTG														
	412	422	431	440	449															
AACCACTCCA	G	GTT	CCC	GAA	AGG	GCG	GTC	CAT	GCT	CGA	GGC	GCT	GGA	GCA	CAC					
		V	P	E	R	A	V	H	A	R	G	A	G	A	H					
	458	467	476	485	494	503														
GGG	ACC	TTC	ACG	AGT	TAC	GCC	GAC	TGG	AGT	AAC	ATC	ACC	GCG	GCG	TCC	TTT	CTG			
G	T	F	T	S	Y	A	D	W	S	N	I	T	A	A	S	F	L			
	512	521	530	539	548	557														
AAC	GCC	ACT	GGA	AAG	CAG	ACG	CCG	CTG	TTT	GTC	CGG	TTC	TCG	ACC	GTT	GCT	GGG			
N	X	X	G	K	Q	T	P	V	X	V	R	F	S	T	V	A	G			

FIG. 4A
SUBSTITUTE SHEET (RULE 26)

6/12

	1175	1185	1195	1205	1215	1224											
ATG	GTACGTTTGC	CTGGCTGAGA	TGACCGTGAA	TCCATTCTA	ACCTCAAGCC	CAG GAT GGC											
M						D G											
	1233	1242	1251	1260	1269	1278											
TTG	GGC	GTC	CAC	ACG	TTC	CGG	TTT	GTC	AAA	GAT	GAC	GGC	TCG	TCC	AAG	TTG	ATC
F	G	V	H	T	F	R	F	V	K	D	D	G	S	S	K	L	I
	1287	1296	1305	1314	1323	1332											
AAG	TGG	CAT	TTC	AAG	TCA	CGC	CAG	GGA	AAG	GCG	AGT	CTA	GTC	TGG	GAA	GAG	GCG
K	W	H	F	K	S	R	Q	G	K	A	S	L	V	W	E	E	A
	1341	1350	1359	1368	1377	1386											
CAG	GTT	CTT	TCT	GGC	AAG	AAT	GCC	GAC	TTC	CAC	CGT	CAG	GAC	CTC	TGG	GAT	GCT
Q	V	L	S	G	K	N	A	D	F	H	R	Q	D	L	W	D	A
	1395	1404	1413	1422	1431	1440											
ATT	GAG	TCC	GGG	AAC	GGA	CCA	GAA	TGG	GAT	GTC	TGC	GTC	CAG	ATT	GTC	GAT	GAG
I	E	S	G	N	G	P	E	W	D	V	C	V	Q	I	V	D	E
	1449	1458	1467	1476	1485	1494											
TCC	CAG	GCG	CAA	GCC	TTT	GGC	TTC	GAC	TTG	CTG	GAC	CCG	ACA	AAG	ATC	ATC	CCC
S	Q	A	Q	A	F	G	F	D	L	L	D	P	T	K	I	I	P
	1503	1512	1521	1530	1539	1548											
GAG	GAG	TAC	GCC	CCC	TTG	ACG	AAG	CTG	GGC	CTC	TTG	AAG	CTG	GAT	CGC	AAT	CCG
E	E	Y	A	P	L	T	K	L	G	L	L	K	L	D	R	N	P
	1557	1566	1575	1584	1593	1602											
ACC	AAC	TAC	TTC	GCC	GAG	ACG	GAG	CAG	GTC	ATG	TTC	CAA	CCC	GGT	CAT	ATC	GTC
T	N	Y	F	A	E	T	E	Q	V	M	F	Q	P	G	H	I	V
	1611	1620	1629	1638	1647	1656											
CGC	GGC	ATC	GAC	TTC	ACG	GAG	GAT	CCC	CTG	CTA	CAG	GGA	CGC	CTC	TTT	TCG	TAC
R	G	I	D	F	T	E	D	P	L	L	Q	G	R	L	F	S	Y

FIG. 4C
 SUBSTITUTE SHEET (RULE 26)

7/12

1665	1674	1683	1692	1701	1710
CTT GAC ACG CAG CTG AAC CGG AAT GGC GGG CCC AAC TTT GAG CAG CTG CCC ATC					
L D T Q L N R N G G P N F E Q L P I					
1719	1728	1737	1746	1755	1764
AAC ATG CCG CGG GTG CCG ATT CAC AAC AAT AAT CGC GAC GGC GCC GGC CAG ATG					
N M P R V P I H N N N R D G A G Q M					
1773	1782	1799	1809	1819	
TTC ATC CAC AGG AAC AAG TAT CCT T		GTAAGTGCCT	CTTTGCCTC	GATCGTTGTG	
F I H R N K Y P Y					
1829	1839	1848	1857	1866	1875
GTGCCGGCTT GCTGACAGAC GCAG AC ACT CCC AAC ACC CTG AAC AGT GGT TAT CCG					
		T P N T L N S G Y P			
1884	1893	1902	1911	1920	1929
CCC CAA GCC AAC CAA AAT GCC GGA CGC GGA TTC TTC ACA GCG CCT GGC CGT ACC					
R Q A N Q N A G R G F F T A P G R T					
1938	1947	1956	1965	1974	1983
GCC AGC GGT GCC CTC GTC CGT GAG GTG TCG CCA ACA TTC AAC GAC CAC TGG TCG					
A S G A L V R E V S P T F N D H W S					
1992	2001	2010	2019	2028	2037
CAG CCC CGT CTC TTC TTC AAC TCC CTC ACT CCC GTC GAA CAA CAG TTC CTC GTC					
Q P R L F F N S L T P V E Q Q F L V					
2046	2055	2064	2073	2084	2091
AAC GCC ATG CGC TTC GAA ATC AGC CTT GTG AAG TCG GAA GAA GTC AAG AAG AAC					
2100	2109	2118	2127	2136	2145
GTG CTC ACC CAG CTC AAC CGC GTC AGC CAT GAC GTG GCC GTG CGC GTG GCC GCC					
V L T Q L N R V S H D V A V R V A A					

FIG. 4D

SUBSTITUTE SHEET (RULE 26)

8/12

2154	2163	2172	2181	2190	2199
GCT ATC GGC CTC GGC GCG CCC GAC GCG GAC GAC ACA TAC TAC CAC AAC AAC AAG					
A I G L G A P D A D D T Y Y H N N K					
2208	2217	2226	2235	2244	2253
ACG GCT GGC GTC TCA ATC GTT GGA AGC GGG CCC TTG CCT ACC ATC AAG ACT CTC					
T A G V S I V G S G P L P T I K T L					
2262	2271	2280	2289	2298	2307
CGC GTC GGC ATC CTG GCT ACC ACG AGC GAG TCG AGC GCG CTG GAT CAG GCG GCC					
R V G I L A T T S E S S A L D Q A A					
2316	2325	2334	2343	2352	2361
CAG CTC CGC ACC CGT CTG GAA AAG GAC GGG CTT GTG GTC ACG GTT GTG GCT GAA					
Q L R T R L E K D G L V V T V V A E					
2370	2379	2388	2397	2406	2415
ACG CTG CGC GAG GGG GTA GAC CAG ACG TAC TCG ACG GCG GAT GCC ACG GGT TTC					
T L R E G V D Q T Y S T A D A T G F					
2424	2433	2442	2451	2460	2469
GAC GGC GTT GTT GTT GTG GAC GGG GCG GCG GCG CTG TTT GCC AGC ACC GCG TCG					
D G V V V V D G A A A A L F A S T A S					
2478	2487	2496	2505	2514	2523
TCG CCG TTG TTC CCG ACG GGC AGG CCG TTG CAG ATC TTT GTG GAC GCG TAT CCG					
S P L F P T G R P L Q I F V D A Y R					
2532	2541	2550	2559	2568	2577
TXX CGA XAX CCC GTX GGT GTG TGT GGT GGG AAG TCG AGC GAG GTG TTG GAT GCG					
W G K P V G V C G G K S S E V L D A					
2586	2595	2604	2613	2622	2631
GCG GAT GTT CCG GAA GAC GGG GAC GGG GTG TAT TCG GAG GAG TCG GTG GAC ATG					
A D V P E D G D G V Y S E E S V D M					

FIG. 4E

SUBSTITUTE SHEET (RULE 26)

9/12

	2640		2649		2658		2667		2680		2690				
	<u>TTT</u>	<u>GTG</u>	<u>GAG</u>	<u>GAG</u>	<u>TTT</u>	<u>GAG</u>	<u>AAG</u>	<u>GGG</u>	<u>TTG</u>	<u>GCT</u>	<u>ACT</u>	<u>TTC</u>	<u>AGG</u>	GTGAGTCTTG	ATGCCTTTGT
	2700		2710		2720		2730		2740		2750		2760		
	TTGTTGTGAT		GTTATTGTTT		TGTTTTGTCT		CGGACTTTGT		GAAAGAATGA		CGGACTGACG		TCTTTGGTAT		
		2770		2779		2788		2794							
CTAG	<u>TTT</u>	<u>ACC</u>	<u>GAT</u>	<u>CGG</u>	<u>TTT</u>	<u>GCT</u>	<u>CTC</u>	<u>GAC</u>	<u>TCT</u>	<u>TAG</u>					
	F	T	D	R	F	A	L	D	S						

FIG. 4F

10/12

X X
 SGQSPLAAYEVDDSDGY peptide STS2 (N-terminal sequence) (THa1, KMBr)
 ||||| ||
 SGQSPLAAYEVDDSTGY from DNA sequence
 X 40 50 X

X X
 ADWS-ITAA peptide Std1 (THa1)
 ||| |||
 ADWSNITAA from DNA sequence
 X 120 X

x X
 RNPTNYFAE peptide STd2 (THa1)
 |||||
 RNPTNYFAE from DNA sequence
 350 X

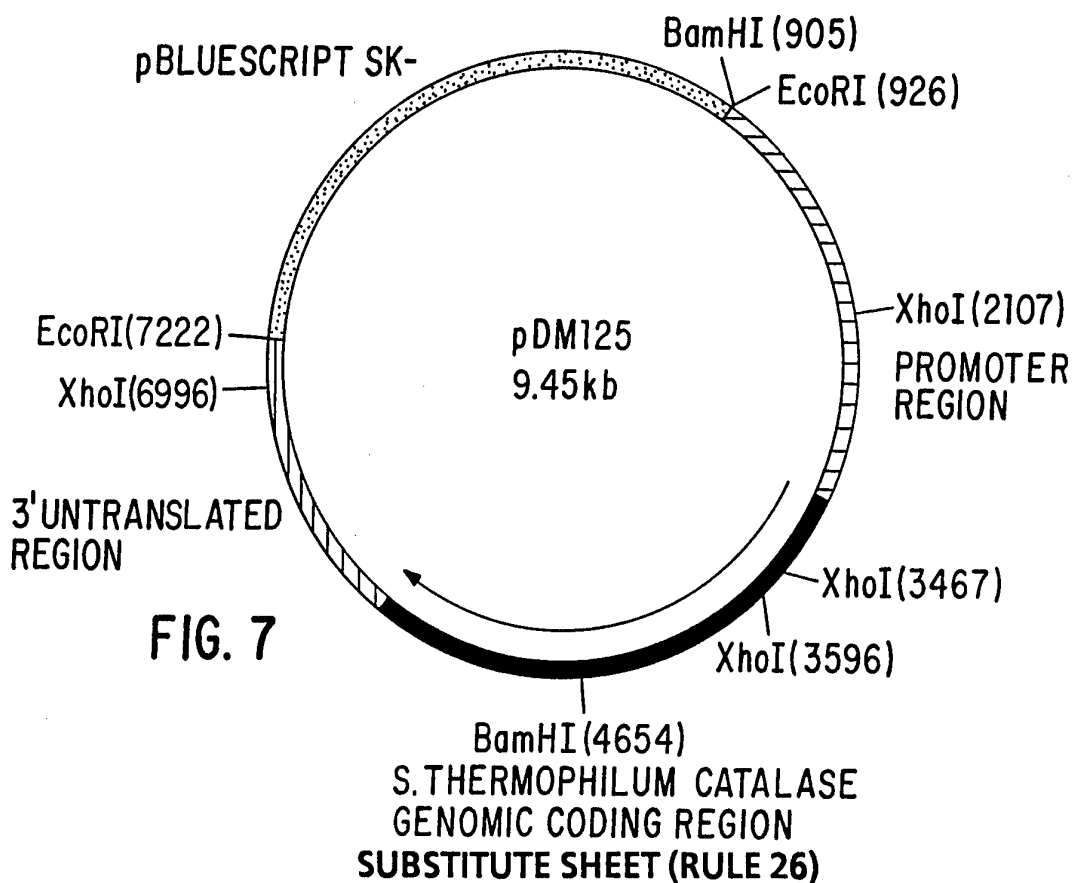
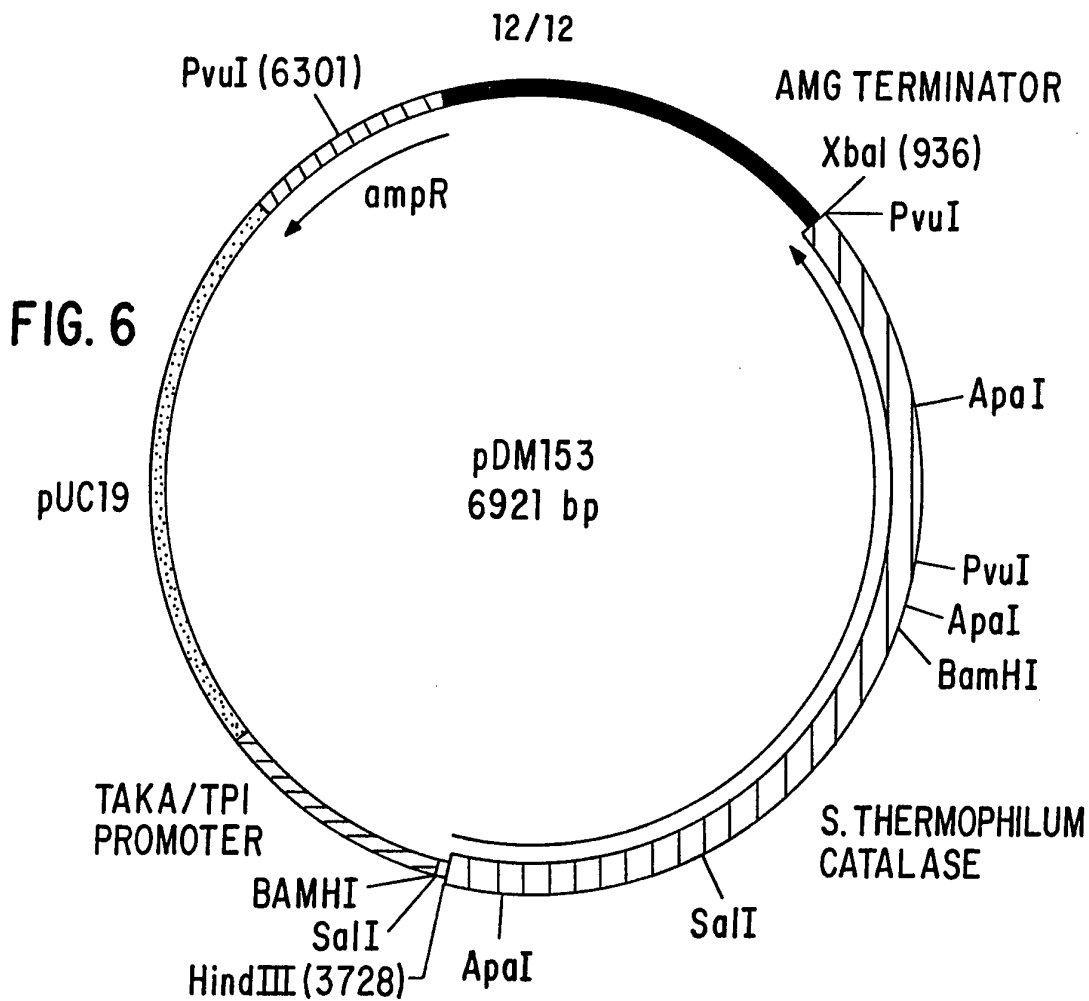
X X
 AYRWGKPVGVXGGKRRE peptide STd3 (THa1)
 ||||| || |
 AYRWGKPVGVCCKGSSE from DNA sequence
 660 670 X

X X
 SFALDQAAQLR peptide STd5 (THa1)
 | |||||
 SSALDQAAQLR from DNA sequence
 580 590

X X
 RFEISLVK peptide STS1A (KMBr)
 |||||
 RFEISLVK from DNA sequence
 x 505 X

X X
 NVLTQLNRV peptide STS1B (KMBr)
 |||||
 NVLTQLNRV from DNA sequence
 X 520 X

FIG 5A
SUBSTITUTE SHEET (RULE 26)



INTERNATIONAL SEARCH REPORT

International Application No
PCT, US 96/06096

<p>A. CLASSIFICATION OF SUBJECT MATTER IPC 6 C12N15/53 C12N15/80 C12N9/08 C12N1/15 //(C12N1/15, C12R1:66)</p>		
<p>According to International Patent Classification (IPC) or to both national classification and IPC</p>		
<p>B. FIELDS SEARCHED</p>		
<p>Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12N</p>		
<p>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched</p>		
<p>Electronic data base consulted during the international search (name of data base and, where practical, search terms used)</p>		
<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,92 17571 (NOVONORDISK AS) 15 October 1992 cited in the application see the whole document ---	20-22
A	EP,A,0 305 216 (NOVO INDUSTRI AS) 1 March 1989 see the whole document ---	1-22
P,A	WO,A,95 15391 (NOVO NORDISK BIOTECH INC) 8 June 1995 see the whole document -----	1-22
<p><input type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.</p>		
<p>* Special categories of cited documents :</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"&" document member of the same patent family</p>		
<p>Date of the actual completion of the international search 20 August 1996</p>		<p>Date of mailing of the international search report 04. 09. 96</p>
<p>Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+ 31-70) 340-3016</p>		<p>Authorized officer Hornig, H</p>

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 96/06096

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9217571	15-10-92	EP-A- 0580662	02-02-94
		JP-T- 6506347	21-07-94
EP-A-0305216	01-03-89	AT-T- 125865	15-08-95
		DE-D- 3854249	07-09-95
		DE-T- 3854249	29-02-96
		ES-T- 2076939	16-11-95
		JP-A- 1157383	20-06-89
		JP-C- 1761424	20-05-93
		JP-B- 4038394	24-06-92
WO-A-9515391	08-06-95	AU-B- 1101995	19-06-95
		CA-A- 2178007	08-06-95