



- (51) **International Patent Classification:**
C12N 9/08 (2006.01)
- (21) **International Application Number:**
PCT/EP2016/064710
- (22) **International Filing Date:**
24 June 2016 (24.06.2016)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
15174096.6 26 June 2015 (26.06.2015) EP
- (71) **Applicant:** NOVOZYMES A/S [DK/DK]; Krogshoejvej 36, 2880 Bagsvaerd (DK).
- (72) **Inventors:** VIND, Jesper; Krogshoejvej 36, 2880 Bagsvaerd (DK). KIEMER, Lars; Krogshoejvej 36, 2880 Bagsvaerd (DK). AMOURGI, Eleni; Krogshoejvej 36, 2880 Bagsvaerd (DK).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))



(54) **Title:** POLYPEPTIDES HAVING PEROXYGENASE ACTIVITY

(57) **Abstract:** The present invention relates to polypeptides having peroxygenase activity, and polynucleotides encoding the polypeptides. The invention also relates to nucleic acid constructs, vectors, and host cells comprising the polynucleotides as well as methods of producing and using the polypeptides.

POLYPEPTIDES HAVING PEROXYGENASE ACTIVITY

Reference to a Sequence Listing

This application contains a Sequence Listing in computer readable form, which is
5 incorporated herein by reference.

Background of the Invention

Field of the Invention

The present invention relates to polypeptides having peroxygenase activity, and
10 polynucleotides encoding the polypeptides. The invention also relates to nucleic acid constructs, vectors, and host cells comprising the polynucleotides as well as methods of producing and using the polypeptides.

Description of the Related Art

A method was reported for the rapid and selective spectrophotometric direct detection of
15 aromatic hydroxylation by the AaP peroxygenase (Kluge *et al.*, 2007, Appl Microbiol Biotechnol 75: 1473-1478).

WO 2008/119780 discloses several different peroxygenase polypeptides and their
encoding polynucleotides, as well as recombinant production thereof.

20 WO 2011/120938 discloses site-specific hydroxylation of aliphatic hydrocarbons using peroxygenase polypeptides.

Gröbe *et al.* "High-yield production of aromatic peroxygenase by the agaric fungus
Marasmius rotula", *AMB Express* 2011, 1:31, discloses an extracellular peroxygenase from
Marasmius rotula that was produced in liquid culture, chromatographically purified and partially
25 characterized.

The present invention provides novel polypeptides having peroxygenase activity and
polynucleotides encoding the polypeptides. The polypeptides have improved activity and
stability compared to the peroxygenase from *Marasmius rotula*.

Summary of the Invention

The present invention provides artificial polypeptides having peroxygenase activity and
polynucleotides encoding the polypeptides.

Accordingly, the present invention relates to polypeptides having peroxygenase activity
selected from the group consisting of:

35 (a) a polypeptide having at least 80% sequence identity to the mature polypeptide of SEQ ID
NO: 2;

(b) a polypeptide encoded by a polynucleotide that hybridizes under high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, or (ii) the full-length complement of (i);

(c) a polypeptide encoded by a polynucleotide having at least 80% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1;

(d) a variant of the mature polypeptide of SEQ ID NO: 2 comprising a substitution, deletion, and/or insertion at one or more (e.g., several) positions; and

(e) a fragment of the polypeptide of (a), (b), (c), or (d) that has peroxygenase activity.

The present invention also relates to polynucleotides encoding the polypeptides of the present invention; nucleic acid constructs; recombinant expression vectors; recombinant host cells comprising the polynucleotides; and methods of producing the polypeptides.

The present invention also relates to methods of using the polypeptides of the invention.

Other aspects and embodiments of the invention are apparent from the description and examples.

Definitions

Peroxygenase: The term “peroxygenase” means an enzyme exhibiting “unspecific peroxygenase” activity according to EC 1.11.2.1, that catalyzes insertion of an oxygen atom from H_2O_2 into a variety of substrates, such as nitrobenzodioxole (see also Example 1). For purposes of the present invention, peroxygenase activity is determined according to the procedure described in M. Poraj-Kobielska, M. Kinne, R. Ullrich, K. Scheibner, M. Hofrichter, “A spectrophotometric assay for the detection of fungal peroxygenases”, *Analytical Biochemistry* (2012), vol. 421, issue 1, pp. 327–329.

The peroxygenase of the present invention has at least 20%, preferably at least 40%, more preferably at least 50%, more preferably at least 60%, more preferably at least 70%, more preferably at least 80%, even more preferably at least 90%, most preferably at least 95%, and even most preferably at least 100% of the peroxygenase activity of the mature polypeptide of SEQ ID NO: 2.

Allelic variant: The term “allelic variant” means any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequences. An allelic variant of a polypeptide is a polypeptide encoded by an allelic variant of a gene.

Allelic variant: The term “allelic variant” means any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in polymorphism within populations. Gene mutations can be silent (no

change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequences. An allelic variant of a polypeptide is a polypeptide encoded by an allelic variant of a gene.

Catalytic domain: The term “catalytic domain” means the region of an enzyme containing the catalytic machinery of the enzyme.

cDNA: The term "cDNA" means a DNA molecule that can be prepared by reverse transcription from a mature, spliced, mRNA molecule obtained from a eukaryotic or prokaryotic cell. cDNA lacks intron sequences that may be present in the corresponding genomic DNA. The initial, primary RNA transcript is a precursor to mRNA that is processed through a series of steps, including splicing, before appearing as mature spliced mRNA.

Coding sequence: The term “coding sequence” means a polynucleotide, which directly specifies the amino acid sequence of a polypeptide. The boundaries of the coding sequence are generally determined by an open reading frame, which begins with a start codon such as ATG, GTG, or TTG and ends with a stop codon such as TAA, TAG, or TGA. The coding sequence may be a genomic DNA, cDNA, synthetic DNA, or a combination thereof.

Control sequences: The term “control sequences” means nucleic acid sequences necessary for expression of a polynucleotide encoding a mature polypeptide of the present invention. Each control sequence may be native (*i.e.*, from the same gene) or foreign (*i.e.*, from a different gene) to the polynucleotide encoding the polypeptide or native or foreign to each other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the polynucleotide encoding a polypeptide.

Expression: The term “expression” includes any step involved in the production of a polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

Expression vector: The term “expression vector” means a linear or circular DNA molecule that comprises a polynucleotide encoding a polypeptide and is operably linked to control sequences that provide for its expression.

Fragment: The term “fragment” means a polypeptide or a catalytic domain having one or more (*e.g.*, several) amino acids absent from the amino and/or carboxyl terminus of a mature polypeptide or domain; wherein the fragment has peroxygenase activity. In one aspect, a fragment contains at least 205 amino acid residues, at least 217 amino acid residues, or at least 229 amino acid residues.

Host cell: The term "host cell" means any cell type that is susceptible to transformation, transfection, transduction, or the like with a nucleic acid construct or expression vector comprising a polynucleotide of the present invention. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

Isolated: The term "isolated" means a substance in a form or environment that does not occur in nature. Non-limiting examples of isolated substances include (1) any non-naturally occurring substance, (2) any substance including, but not limited to, any enzyme, variant, nucleic acid, protein, peptide or cofactor, that is at least partially removed from one or more or all of the naturally occurring constituents with which it is associated in nature; (3) any substance modified by the hand of man relative to that substance found in nature; or (4) any substance modified by increasing the amount of the substance relative to other components with which it is naturally associated (*e.g.*, recombinant production in a host cell; multiple copies of a gene encoding the substance; and use of a stronger promoter than the promoter naturally associated with the gene encoding the substance). An isolated substance may be present in a fermentation broth sample; *e.g.* a host cell may be genetically modified to express the polypeptide of the invention. The fermentation broth from that host cell will comprise the isolated polypeptide.

Mature polypeptide: The term "mature polypeptide" means a polypeptide in its final form following translation and any post-translational modifications, such as N-terminal processing, C-terminal truncation, glycosylation, phosphorylation, etc. In one aspect, the mature polypeptide is amino acids 1 to 241 of SEQ ID NO: 2. It is known in the art that a host cell may produce a mixture of two or more different mature polypeptides (*i.e.*, with a different C-terminal and/or N-terminal amino acid) expressed by the same polynucleotide. It is also known in the art that different host cells process polypeptides differently, and thus, one host cell expressing a polynucleotide may produce a different mature polypeptide (*e.g.*, having a different C-terminal and/or N-terminal amino acid) as compared to another host cell expressing the same polynucleotide.

Mature polypeptide coding sequence: The term "mature polypeptide coding sequence" means a polynucleotide that encodes a mature polypeptide having peroxxygenase activity. In one aspect, the mature polypeptide coding sequence is nucleotides 1 to 723 of SEQ ID NO: 1.

Nucleic acid construct: The term "nucleic acid construct" means a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature or which is synthetic, which comprises one or more control sequences.

Operably linked: The term "operably linked" means a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of a polynucleotide such that the control sequence directs expression of the coding sequence.

Sequence identity: The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter "sequence identity".

For purposes of the present invention, the sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *J. Mol. Biol.* 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *Trends Genet.* 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

$$(\text{Identical Residues} \times 100) / (\text{Length of Alignment} - \text{Total Number of Gaps in Alignment})$$

For purposes of the present invention, the sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra*) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *supra*), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

$$(\text{Identical Deoxyribonucleotides} \times 100) / (\text{Length of Alignment} - \text{Total Number of Gaps in Alignment})$$

Stringency conditions: The term "very low stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 45°C.

The term "low stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 25% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 50°C.

The term "medium stringency conditions" means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 35% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for 15 minutes using 2X SSC, 0.2% SDS at 55°C.

The term “medium-high stringency conditions” means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 35% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed
5 three times each for 15 minutes using 2X SSC, 0.2% SDS at 60°C.

The term “high stringency conditions” means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for
10 15 minutes using 2X SSC, 0.2% SDS at 65°C.

The term “very high stringency conditions” means for probes of at least 100 nucleotides in length, prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200 micrograms/ml sheared and denatured salmon sperm DNA, and 50% formamide, following standard Southern blotting procedures for 12 to 24 hours. The carrier material is finally washed three times each for
15 15 minutes using 2X SSC, 0.2% SDS at 70°C.

Subsequence: The term “subsequence” means a polynucleotide having one or more (e.g., several) nucleotides absent from the 5' and/or 3' end of a mature polypeptide coding sequence; wherein the subsequence encodes a fragment having peroxygenase activity. In one aspect, a subsequence contains at least 615 nucleotides, at least 651 nucleotides, or at least
20 687 nucleotides.

Variant: The term “variant” means a polypeptide having peroxygenase activity comprising an alteration, i.e., a substitution, insertion, and/or deletion, at one or more (e.g., several) positions. A substitution means replacement of the amino acid occupying a position with a different amino acid; a deletion means removal of the amino acid occupying a position; and an
25 insertion means adding an amino acid adjacent to and immediately following the amino acid occupying a position.

Detailed Description of the Invention

Polypeptides Having Peroxygenase Activity

In an embodiment, the present invention relates to polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 70%, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, which have peroxygenase activity. In one aspect, the polypeptides differ by up to 10 amino acids, e.g., 1, 2,
30 3, 4, 5, 6, 7, 8, 9, or 10, from the mature polypeptide of SEQ ID NO: 2.

In a particular embodiment the invention relates to polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 80%, at least 85%, at least 90%,

at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, and wherein the polypeptide has at least at least 70% of the peroxygenase activity of the mature polypeptide of SEQ ID NO: 2.

5 In a particular embodiment the invention relates to polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, and wherein the polypeptide has at least at least 75% of the peroxygenase activity of the mature polypeptide of SEQ ID NO: 2.

10 In a particular embodiment the invention relates to polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, and wherein the polypeptide has at least at least 80% of the peroxygenase activity of the mature polypeptide of SEQ ID NO: 2.

15 In a particular embodiment the invention relates to polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, and wherein the polypeptide has at least at least 85% of the peroxygenase activity of the mature polypeptide of SEQ ID NO: 2.

20 In a particular embodiment the invention relates to polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, and wherein the polypeptide has at least at least 90% of the peroxygenase activity of the mature polypeptide of SEQ ID NO: 2.

25 In a particular embodiment the invention relates to polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, and wherein the polypeptide has at least at least 95% of the peroxygenase activity of the mature polypeptide of SEQ ID NO: 2.

30 In a particular embodiment the invention relates to polypeptides having a sequence identity to the mature polypeptide of SEQ ID NO: 2 of at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%, and wherein the polypeptide has at least at least 100% of the peroxygenase activity of the mature polypeptide of SEQ ID NO: 2.

35 In an embodiment, the polypeptide has been isolated. A polypeptide of the present invention preferably comprises or consists of the amino acid sequence of SEQ ID NO: 2 or an allelic variant thereof; or is a fragment thereof having peroxygenase activity. In another aspect,

the polypeptide comprises or consists of the mature polypeptide of SEQ ID NO: 2. In another aspect, the polypeptide comprises or consists of amino acids 1 to 241 of SEQ ID NO: 2.

In another embodiment, the present invention relates to a polypeptide having peroxygenase activity encoded by a polynucleotide that hybridizes under very low stringency conditions, low stringency conditions, medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, or (ii) the full-length complement of (i) (Sambrook *et al.*, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, New York). In an embodiment, the polypeptide has been isolated.

The polynucleotide of SEQ ID NO: 1 or a subsequence thereof, as well as the polypeptide of SEQ ID NO: 2 or a fragment thereof may be used to design nucleic acid probes to identify and clone DNA encoding polypeptides having peroxygenase activity from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic DNA or cDNA of a cell of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Such probes can be considerably shorter than the entire sequence, but should be at least 15, e.g., at least 25, at least 35, or at least 70 nucleotides in length. Preferably, the nucleic acid probe is at least 100 nucleotides in length, e.g., at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, at least 500 nucleotides, at least 600 nucleotides, at least 700 nucleotides, at least 800 nucleotides, or at least 900 nucleotides in length. Both DNA and RNA probes can be used. The probes are typically labeled for detecting the corresponding gene (for example, with ^{32}P , ^3H , ^{35}S , biotin, or avidin). Such probes are encompassed by the present invention.

A genomic DNA or cDNA library prepared from such other strains may be screened for DNA that hybridizes with the probes described above and encodes a polypeptide having peroxygenase activity. Genomic or other DNA from such other strains may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA that hybridizes with SEQ ID NO: 1 or a subsequence thereof, the carrier material is used in a Southern blot.

For purposes of the present invention, hybridization indicates that the polynucleotide hybridizes to a labeled nucleic acid probe corresponding to (i) SEQ ID NO: 1; (ii) the mature polypeptide coding sequence of SEQ ID NO: 1; (iii) the full-length complement thereof; or (iv) a subsequence thereof; under very low to very high stringency conditions. Molecules to which the nucleic acid probe hybridizes under these conditions can be detected using, for example, X-ray film or any other detection means known in the art.

In one aspect, the nucleic acid probe is nucleotides 1 to 615, nucleotides 1 to 651, nucleotides 1 to 687, or nucleotides 1 to 723 of SEQ ID NO: 1. In another aspect, the nucleic acid probe is a polynucleotide that encodes the polypeptide of SEQ ID NO: 2; the mature polypeptide thereof; or a fragment thereof. In another aspect, the nucleic acid probe is SEQ ID NO: 1.

In another embodiment, the present invention relates to a polypeptide having peroxxygenase activity encoded by a polynucleotide having a sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1 of at least 70%, e.g., at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%. In a further embodiment, the polypeptide has been isolated.

In another embodiment, the present invention relates to variants of the mature polypeptide of SEQ ID NO: 2 comprising a substitution, deletion, and/or insertion at one or more (e.g., several) positions. In an embodiment, the number of amino acid substitutions, deletions and/or insertions introduced into the mature polypeptide of SEQ ID NO: 2 is up to 10, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10. The amino acid changes may be of a minor nature, that is conservative amino acid substitutions or insertions that do not significantly affect the folding and/or activity of the protein; small deletions, typically of 1-30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to 20-25 residues; or a small extension that facilitates purification by changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain.

Examples of conservative substitutions are within the groups of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine, threonine and methionine). Amino acid substitutions that do not generally alter specific activity are known in the art and are described, for example, by H. Neurath and R.L. Hill, 1979, *In, The Proteins*, Academic Press, New York. Common substitutions are Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, and Asp/Gly.

Alternatively, the amino acid changes are of such a nature that the physico-chemical properties of the polypeptides are altered. For example, amino acid changes may improve the thermal stability of the polypeptide, alter the substrate specificity, change the pH optimum, and the like.

Essential amino acids in a polypeptide can be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, 1989, *Science* 244: 1081-1085). In the latter technique, single alanine mutations are

introduced at every residue in the molecule, and the resultant molecules are tested for peroxxygenase activity to identify amino acid residues that are critical to the activity of the molecule. See also, Hilton *et al.*, 1996, *J. Biol. Chem.* 271: 4699-4708. The active site of the enzyme or other biological interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction, or photoaffinity labeling, in conjunction with mutation of putative contact site amino acids. See, for example, de Vos *et al.*, 1992, *Science* 255: 306-312; Smith *et al.*, 1992, *J. Mol. Biol.* 224: 899-904; Wlodaver *et al.*, 1992, *FEBS Lett.* 309: 59-64. The identity of essential amino acids can also be inferred from an alignment with a related polypeptide.

Essential amino acids in the sequence of amino acids 1 to 241 of SEQ ID NO: 2 are located at positions 23, 24, and 25. Other highly conserved amino acids are located at positions 27-37 and positions 92-99.

Single or multiple amino acid substitutions, deletions, and/or insertions can be made and tested using known methods of mutagenesis, recombination, and/or shuffling, followed by a relevant screening procedure, such as those disclosed by Reidhaar-Olson and Sauer, 1988, *Science* 241: 53-57; Bowie and Sauer, 1989, *Proc. Natl. Acad. Sci. USA* 86: 2152-2156; WO 95/17413; or WO 95/22625. Other methods that can be used include error-prone PCR, phage display (*e.g.*, Lowman *et al.*, 1991, *Biochemistry* 30: 10832-10837; U.S. Patent No. 5,223,409; WO 92/06204), and region-directed mutagenesis (Derbyshire *et al.*, 1986, *Gene* 46: 145; Ner *et al.*, 1988, *DNA* 7: 127).

Mutagenesis/shuffling methods can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides expressed by host cells (Ness *et al.*, 1999, *Nature Biotechnology* 17: 893-896). Mutagenized DNA molecules that encode active polypeptides can be recovered from the host cells and rapidly sequenced using standard methods in the art. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide.

The polypeptide may be a hybrid polypeptide in which a region of one polypeptide is fused at the N-terminus or the C-terminus of a region of another polypeptide.

The polypeptide may be a fusion polypeptide or cleavable fusion polypeptide in which another polypeptide is fused at the N-terminus or the C-terminus of the polypeptide of the present invention. A fusion polypeptide is produced by fusing a polynucleotide encoding another polypeptide to a polynucleotide of the present invention. Techniques for producing fusion polypeptides are known in the art, and include ligating the coding sequences encoding the polypeptides so that they are in frame and that expression of the fusion polypeptide is under control of the same promoter(s) and terminator. Fusion polypeptides may also be constructed using intein technology in which fusion polypeptides are created post-translationally (Cooper *et al.*, 1993, *EMBO J.* 12: 2575-2583; Dawson *et al.*, 1994, *Science* 266: 776-779).

A fusion polypeptide can further comprise a cleavage site between the two polypeptides. Upon secretion of the fusion protein, the site is cleaved releasing the two polypeptides. Examples of cleavage sites include, but are not limited to, the sites disclosed in Martin *et al.*, 2003, *J. Ind. Microbiol. Biotechnol.* 3: 568-576; Svetina *et al.*, 2000, *J. Biotechnol.* 76: 245-251; Rasmussen-Wilson *et al.*, 1997, *Appl. Environ. Microbiol.* 63: 3488-3493; Ward *et al.*, 1995, *Biotechnology* 13: 498-503; and Contreras *et al.*, 1991, *Biotechnology* 9: 378-381; Eaton *et al.*, 1986, *Biochemistry* 25: 505-512; Collins-Racie *et al.*, 1995, *Biotechnology* 13: 982-987; Carter *et al.*, 1989, *Proteins: Structure, Function, and Genetics* 6: 240-248; and Stevens, 2003, *Drug Discovery World* 4: 35-48.

Sources of Polypeptides Having Peroxygenase Activity

A polypeptide having peroxygenase activity of the present invention may be obtained from microorganisms of any genus. For purposes of the present invention, the term "obtained from" as used herein in connection with a given source shall mean that the polypeptide encoded by a polynucleotide is produced by the source or by a strain in which the polynucleotide from the source has been inserted. In one aspect, the polypeptide obtained from a given source is secreted extracellularly.

It will be understood that for the aforementioned species, the invention encompasses both the perfect and imperfect states, and other taxonomic equivalents, *e.g.*, anamorphs, regardless of the species name by which they are known. Those skilled in the art will readily recognize the identity of appropriate equivalents.

Strains of these species are readily accessible to the public in a number of culture collections, such as the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau Voor Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

The polypeptide may be identified and obtained from other sources including microorganisms isolated from nature (*e.g.*, soil, composts, water, etc.) or DNA samples obtained directly from natural materials (*e.g.*, soil, composts, water, etc.) using the above-mentioned probes. Techniques for isolating microorganisms and DNA directly from natural habitats are well known in the art. A polynucleotide encoding the polypeptide may then be obtained by similarly screening a genomic DNA or cDNA library of another microorganism or mixed DNA sample. Once a polynucleotide encoding a polypeptide has been detected with the probe(s), the polynucleotide can be isolated or cloned by utilizing techniques that are known to those of ordinary skill in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

Nucleic Acid Constructs

The present invention also relates to nucleic acid constructs comprising a polynucleotide of the present invention operably linked to one or more control sequences that direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

The polynucleotide may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides utilizing recombinant DNA methods are well known in the art.

The control sequence may be a promoter, a polynucleotide that is recognized by a host cell for expression of a polynucleotide encoding a polypeptide of the present invention. The promoter contains transcriptional control sequences that mediate the expression of the polypeptide. The promoter may be any polynucleotide that shows transcriptional activity in the host cell including variant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a bacterial host cell are the promoters obtained from the *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus subtilis* *xylA* and *xylB* genes, *Bacillus thuringiensis* *cryIIIA* gene (Agaisse and Lereclus, 1994, *Molecular Microbiology* 13: 97-107), *E. coli* *lac* operon, *E. coli* *trc* promoter (Egon *et al.*, 1988, *Gene* 69: 301-315), *Streptomyces coelicolor* agarase gene (*dagA*), and prokaryotic beta-lactamase gene (Villa-Kamaroff *et al.*, 1978, *Proc. Natl. Acad. Sci. USA* 75: 3727-3731), as well as the *tac* promoter (DeBoer *et al.*, 1983, *Proc. Natl. Acad. Sci. USA* 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Gilbert *et al.*, 1980, *Scientific American* 242: 74-94; and in Sambrook *et al.*, 1989, *supra*. Examples of tandem promoters are disclosed in WO 99/43835.

Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the genes for *Aspergillus nidulans* acetamidase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Aspergillus oryzae* TKA amylase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Fusarium oxysporum* trypsin-like protease (WO 96/00787), *Fusarium venenatum* amyloglucosidase (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), *Rhizomucor miehei* lipase, *Rhizomucor*

miehei aspartic proteinase, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* xylanase III, *Trichoderma reesei* beta-xylosidase, and *Trichoderma reesei* translation elongation factor, as well as the NA2-tpi promoter (a modified promoter from an *Aspergillus* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus* triose phosphate isomerase gene; non-limiting examples include modified promoters from an *Aspergillus niger* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus nidulans* or *Aspergillus oryzae* triose phosphate isomerase gene); and variant, truncated, and hybrid promoters thereof. Other promoters are described in U.S. Patent No. 6,011,147.

In a yeast host, useful promoters are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH1, ADH2/GAP), *Saccharomyces cerevisiae* triose phosphate isomerase (TPI), *Saccharomyces cerevisiae* metallothionein (CUP1), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos *et al.*, 1992, *Yeast* 8: 423-488.

The control sequence may also be a transcription terminator, which is recognized by a host cell to terminate transcription. The terminator is operably linked to the 3'-terminus of the polynucleotide encoding the polypeptide. Any terminator that is functional in the host cell may be used in the present invention.

Preferred terminators for bacterial host cells are obtained from the genes for *Bacillus clausii* alkaline protease (*aprH*), *Bacillus licheniformis* alpha-amylase (*amyL*), and *Escherichia coli* ribosomal RNA (*rmB*).

Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* acetamidase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase, *Aspergillus oryzae* TAKA amylase, *Fusarium oxysporum* trypsin-like protease, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* xylanase III, *Trichoderma reesei* beta-xylosidase, and *Trichoderma reesei* translation elongation factor.

Preferred terminators for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1), and *Saccharomyces*

cerevisiae glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos *et al.*, 1992, *supra*.

The control sequence may also be an mRNA stabilizer region downstream of a promoter and upstream of the coding sequence of a gene which increases expression of the gene.

5 Examples of suitable mRNA stabilizer regions are obtained from a *Bacillus thuringiensis cryIIIA* gene (WO 94/25612) and a *Bacillus subtilis* SP82 gene (Hue *et al.*, 1995, *Journal of Bacteriology* 177: 3465-3471).

10 The control sequence may also be a leader, a nontranslated region of an mRNA that is important for translation by the host cell. The leader is operably linked to the 5'-terminus of the polynucleotide encoding the polypeptide. Any leader that is functional in the host cell may be used.

Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase.

15 Suitable leaders for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP).

20 The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3'-terminus of the polynucleotide and, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence that is functional in the host cell may be used.

25 Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase *Aspergillus oryzae* TAKA amylase, and *Fusarium oxysporum* trypsin-like protease.

Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, *Mol. Cellular Biol.* 15: 5983-5990.

30 The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a polypeptide and directs the polypeptide into the cell's secretory pathway. The 5'-end of the coding sequence of the polynucleotide may inherently contain a signal peptide coding sequence naturally linked in translation reading frame with the segment of the coding sequence that encodes the polypeptide. Alternatively, the 5'-end of the coding sequence may contain a signal peptide coding sequence that is foreign to the coding sequence. A foreign signal peptide coding sequence may be required where the coding
35 sequence does not naturally contain a signal peptide coding sequence. Alternatively, a foreign signal peptide coding sequence may simply replace the natural signal peptide coding sequence

in order to enhance secretion of the polypeptide. However, any signal peptide coding sequence that directs the expressed polypeptide into the secretory pathway of a host cell may be used.

Effective signal peptide coding sequences for bacterial host cells are the signal peptide coding sequences obtained from the genes for *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus* *stearothermophilus* alpha-amylase, *Bacillus stearothermophilus* neutral proteases (*nprT*, *nprS*, *nprM*), and *Bacillus subtilis* *prsA*. Further signal peptides are described by Simonen and Palva, 1993, *Microbiological Reviews* 57: 109-137.

Effective signal peptide coding sequences for filamentous fungal host cells are the signal peptide coding sequences obtained from the genes for *Aspergillus niger* neutral amylase, *Aspergillus niger* glucoamylase, *Aspergillus oryzae* TAKA amylase, *Humicola insolens* cellulase, *Humicola insolens* endoglucanase V, *Humicola lanuginosa* lipase, and *Rhizomucor miehei* aspartic proteinase.

Useful signal peptides for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding sequences are described by Romanos *et al.*, 1992, *supra*.

The control sequence may also be a propeptide coding sequence that encodes a propeptide positioned at the N-terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to an active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding sequence may be obtained from the genes for *Bacillus subtilis* alkaline protease (*aprE*), *Bacillus subtilis* neutral protease (*nprT*), *Myceliophthora thermophila* laccase (WO 95/33836), *Rhizomucor miehei* aspartic proteinase, and *Saccharomyces cerevisiae* alpha-factor.

Where both signal peptide and propeptide sequences are present, the propeptide sequence is positioned next to the N-terminus of a polypeptide and the signal peptide sequence is positioned next to the N-terminus of the propeptide sequence.

It may also be desirable to add regulatory sequences that regulate expression of the polypeptide relative to the growth of the host cell. Examples of regulatory sequences are those that cause expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory sequences in prokaryotic systems include the *lac*, *tac*, and *trp* operator systems. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the *Aspergillus niger* glucoamylase promoter, *Aspergillus oryzae* TAKA alpha-amylase promoter, and *Aspergillus oryzae* glucoamylase promoter, *Trichoderma reesei* cellobiohydrolase I promoter, and *Trichoderma reesei* cellobiohydrolase II promoter may be used. Other examples of regulatory sequences are those that allow for gene amplification. In eukaryotic systems, these regulatory sequences

include the dihydrofolate reductase gene that is amplified in the presence of methotrexate, and the metallothionein genes that are amplified with heavy metals. In these cases, the polynucleotide encoding the polypeptide would be operably linked to the regulatory sequence.

5 Expression Vectors

The present invention also relates to recombinant expression vectors comprising a polynucleotide of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleotide and control sequences may be joined together to produce a recombinant expression vector that may include one or more convenient restriction sites to
10 allow for insertion or substitution of the polynucleotide encoding the polypeptide at such sites. Alternatively, the polynucleotide may be expressed by inserting the polynucleotide or a nucleic acid construct comprising the polynucleotide into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

15 The recombinant expression vector may be any vector (e.g., a plasmid or virus) that can be conveniently subjected to recombinant DNA procedures and can bring about expression of the polynucleotide. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may be a linear or closed circular plasmid.

20 The vector may be an autonomously replicating vector, *i.e.*, a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one that, when introduced into the host cell, is integrated into the genome and replicated
25 together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids that together contain the total DNA to be introduced into the genome of the host cell, or a transposon, may be used.

The vector preferably contains one or more selectable markers that permit easy selection of transformed, transfected, transduced, or the like cells. A selectable marker is a gene the
30 product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

Examples of bacterial selectable markers are *Bacillus licheniformis* or *Bacillus subtilis* *dal* genes, or markers that confer antibiotic resistance such as ampicillin, chloramphenicol, kanamycin, neomycin, spectinomycin, or tetracycline resistance. Suitable markers for yeast host
35 cells include, but are not limited to, ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. Selectable markers for use in a filamentous fungal host cell include, but are not limited to, *adeA* (phosphoribosylaminoimidazole-succinocarboxamide synthase), *adeB* (phosphoribosyl-

aminoimidazole synthase), *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenylyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae* *amdS* and *pyrG* genes and a *Streptomyces hygroscopicus* *bar* gene. Preferred for use in a *Trichoderma* cell are *adeA*, *adeB*, *amdS*, *hph*, and *pyrG* genes.

The selectable marker may be a dual selectable marker system as described in WO 2010/039889. In one aspect, the dual selectable marker is an *hph-tk* dual selectable marker system.

The vector preferably contains an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

For integration into the host cell genome, the vector may rely on the polynucleotide's sequence encoding the polypeptide or any other element of the vector for integration into the genome by homologous or non-homologous recombination. Alternatively, the vector may contain additional polynucleotides for directing integration by homologous recombination into the genome of the host cell at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should contain a sufficient number of nucleic acids, such as 100 to 10,000 base pairs, 400 to 10,000 base pairs, and 800 to 10,000 base pairs, which have a high degree of sequence identity to the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding polynucleotides. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. The origin of replication may be any plasmid replicator mediating autonomous replication that functions in a cell. The term "origin of replication" or "plasmid replicator" means a polynucleotide that enables a plasmid or vector to replicate *in vivo*.

Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAM β 1 permitting replication in *Bacillus*.

Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6.

Examples of origins of replication useful in a filamentous fungal cell are AMA1 and ANS1 (Gems *et al.*, 1991, *Gene* 98: 61-67; Cullen *et al.*, 1987, *Nucleic Acids Res.* 15: 9163-9175; WO 00/24883). Isolation of the AMA1 gene and construction of plasmids or vectors comprising the gene can be accomplished according to the methods disclosed in WO 00/24883.

More than one copy of a polynucleotide of the present invention may be inserted into a host cell to increase production of a polypeptide. An increase in the copy number of the polynucleotide can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the polynucleotide where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the polynucleotide, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

Host Cells

The present invention also relates to recombinant host cells, comprising a polynucleotide of the present invention operably linked to one or more control sequences that direct the production of a polypeptide of the present invention. A construct or vector comprising a polynucleotide is introduced into a host cell so that the construct or vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication. The choice of a host cell will to a large extent depend upon the gene encoding the polypeptide and its source.

The host cell may be any cell useful in the recombinant production of a polypeptide of the present invention, *e.g.*, a prokaryote or a eukaryote.

The prokaryotic host cell may be any Gram-positive or Gram-negative bacterium. Gram-positive bacteria include, but are not limited to, *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, and *Streptomyces*. Gram-negative bacteria include, but are not limited to, *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, and *Ureaplasma*.

The bacterial host cell may be any *Bacillus* cell including, but not limited to, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis* cells.

The bacterial host cell may also be any *Streptococcus* cell including, but not limited to, *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, and *Streptococcus equi* subsp. *Zooepidemicus* cells.

The bacterial host cell may also be any *Streptomyces* cell including, but not limited to, *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, and *Streptomyces lividans* cells.

The introduction of DNA into a *Bacillus* cell may be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, *Mol. Gen. Genet.* 168: 111-115), competent cell transformation (see, e.g., Young and Spizizen, 1961, *J. Bacteriol.* 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, *J. Mol. Biol.* 56: 209-221), electroporation (see, e.g., Shigekawa and Dower, 1988, *Biotechniques* 6: 742-751), or conjugation (see, e.g., Koehler and Thorne, 1987, *J. Bacteriol.* 169: 5271-5278). The introduction of DNA into an *E. coli* cell may be effected by protoplast transformation (see, e.g., Hanahan, 1983, *J. Mol. Biol.* 166: 557-580) or electroporation (see, e.g., Dower *et al.*, 1988, *Nucleic Acids Res.* 16: 6127-6145). The introduction of DNA into a *Streptomyces* cell may be effected by protoplast transformation, electroporation (see, e.g., Gong *et al.*, 2004, *Folia Microbiol. (Praha)* 49: 399-405), conjugation (see, e.g., Mazodier *et al.*, 1989, *J. Bacteriol.* 171: 3583-3585), or transduction (see, e.g., Burke *et al.*, 2001, *Proc. Natl. Acad. Sci. USA* 98: 6289-6294). The introduction of DNA into a *Pseudomonas* cell may be effected by electroporation (see, e.g., Choi *et al.*, 2006, *J. Microbiol. Methods* 64: 391-397) or conjugation (see, e.g., Pinedo and Smets, 2005, *Appl. Environ. Microbiol.* 71: 51-57). The introduction of DNA into a *Streptococcus* cell may be effected by natural competence (see, e.g., Perry and Kuramitsu, 1981, *Infect. Immun.* 32: 1295-1297), protoplast transformation (see, e.g., Catt and Jollick, 1991, *Microbios* 68: 189-207), electroporation (see, e.g., Buckley *et al.*, 1999, *Appl. Environ. Microbiol.* 65: 3800-3804), or conjugation (see, e.g., Clewell, 1981, *Microbiol. Rev.* 45: 409-436). However, any method known in the art for introducing DNA into a host cell can be used.

The host cell may also be a eukaryote, such as a mammalian, insect, plant, or fungal cell.

The host cell may be a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota as well as the Oomycota and all mitosporic fungi (as defined by Hawksworth *et al.*, In, *Ainsworth and Bisby's Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK).

The fungal host cell may be a yeast cell. "Yeast" as used herein includes ascosporogenous yeast (*Endomycetales*), basidiosporogenous yeast, and yeast belonging to the *Fungi Imperfecti* (*Blastomycetes*). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in *Biology and Activities of Yeast* (Skinner, Passmore, and Davenport, editors, *Soc. App. Bacteriol. Symposium Series* No. 9, 1980).

The yeast host cell may be a *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* cell, such as a *Kluyveromyces lactis*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*,
 5 *Saccharomyces oviformis*, or *Yarrowia lipolytica* cell.

The fungal host cell may be a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth *et al.*, 1995, *supra*). The filamentous fungi are generally characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative
 10 growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

The filamentous fungal host cell may be an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Fillbasidium*,
 15 *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, or *Trichoderma* cell.

For example, the filamentous fungal host cell may be an *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*,
 20 *Aspergillus oryzae*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *Coprinus cinereus*, *Coriolus hirsutus*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*,
 25 *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*,
 30 *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* cell.

35 Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* and *Trichoderma* host cells are described

in EP 238023, Yelton *et al.*, 1984, *Proc. Natl. Acad. Sci. USA* 81: 1470-1474, and Christensen *et al.*, 1988, *Bio/Technology* 6: 1419-1422. Suitable methods for transforming *Fusarium* species are described by Malardier *et al.*, 1989, *Gene* 78: 147-156, and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, In Abelson, J.N. and
5 Simon, M.I., editors, *Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology*, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito *et al.*, 1983, *J. Bacteriol.* 153: 163; and Hinnen *et al.*, 1978, *Proc. Natl. Acad. Sci. USA* 75: 1920.

Methods of Production

10 The present invention also relates to methods of producing a polypeptide of the present invention, comprising (a) cultivating a cell, which in its wild-type form produces the polypeptide, under conditions conducive for production of the polypeptide; and optionally, (b) recovering the polypeptide.

15 The present invention also relates to methods of producing a polypeptide of the present invention, comprising (a) cultivating a recombinant host cell of the present invention under conditions conducive for production of the polypeptide; and optionally, (b) recovering the polypeptide.

20 The host cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods known in the art. For example, the cells may be cultivated by shake flask cultivation, or small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial
25 suppliers or may be prepared according to published compositions (*e.g.*, in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

30 The polypeptide may be detected using methods known in the art that are specific for the polypeptides. These detection methods include, but are not limited to, use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of the polypeptide.

35 The polypeptide may be recovered using methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, collection, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation. In one aspect, a fermentation broth comprising the polypeptide is recovered.

The polypeptide may be purified by a variety of procedures known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing), differential solubility (e.g., ammonium sulfate precipitation), SDS-PAGE, or extraction (see, e.g., *Protein Purification*, Janson and Ryden, editors, VCH Publishers, New York, 1989) to obtain substantially pure polypeptides.

In an alternative aspect, the polypeptide is not recovered, but rather a host cell of the present invention expressing the polypeptide is used as a source of the polypeptide.

Production in plants

The present invention also relates to isolated plants, e.g., a transgenic plant, plant part, or plant cell, comprising a polynucleotide of the present invention so as to express and produce a polypeptide or domain in recoverable quantities. The polypeptide or domain may be recovered from the plant or plant part. Alternatively, the plant or plant part containing the polypeptide or domain may be used as such for improving the quality of a food or feed, e.g., improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor.

The transgenic plant can be dicotyledonous (a dicot) or monocotyledonous (a monocot). Examples of monocot plants are grasses, such as meadow grass (blue grass, *Poa*), forage grass such as *Festuca*, *Lolium*, temperate grass, such as *Agrostis*, and cereals, e.g., wheat, oats, rye, barley, rice, sorghum, and maize (corn).

Examples of dicot plants are tobacco, legumes, such as lupins, potato, sugar beet, pea, bean and soybean, and cruciferous plants (family *Brassicaceae*), such as cauliflower, rape seed, and the closely related model organism *Arabidopsis thaliana*.

Examples of plant parts are stem, callus, leaves, root, fruits, seeds, and tubers as well as the individual tissues comprising these parts, e.g., epidermis, mesophyll, parenchyme, vascular tissues, meristems.

Plant cells and specific plant cell compartments, such as chloroplasts, apoplasts, mitochondria, vacuoles, peroxisomes and cytoplasm are also considered to be a plant part.

Also included within the scope of the present invention are the progeny of such plants, plant parts, and plant cells.

The transgenic plant or plant cell expressing the polypeptide or domain may be constructed in accordance with methods known in the art.

The present invention also relates to methods of producing a polypeptide or domain of the present invention comprising (a) cultivating a transgenic plant or a plant cell comprising a polynucleotide encoding the polypeptide or domain under conditions conducive for production of the polypeptide or domain; and (b) recovering the polypeptide or domain.

Removal or Reduction of Peroxygenase Activity

The present invention also relates to methods of producing a mutant of a parent cell, which comprises disrupting or deleting a polynucleotide, or a portion thereof, encoding a polypeptide of the present invention, which results in the mutant cell producing less of the polypeptide than the parent cell when cultivated under the same conditions.

The mutant cell may be constructed by reducing or eliminating expression of the polynucleotide using methods well known in the art, for example, insertions, disruptions, replacements, or deletions. The polynucleotide to be modified or inactivated may be, for example, the coding region or a part thereof essential for activity, or a regulatory element required for expression of the coding region. An example of such a regulatory or control sequence may be a promoter sequence or a functional part thereof, *i.e.*, a part that is sufficient for affecting expression of the polynucleotide. Other control sequences for possible modification include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, signal peptide sequence, transcription terminator, and transcriptional activator.

The methods of the present invention for producing an essentially peroxygenase-free product are of particular interest in the production of polypeptides, in particular proteins such as enzymes. The peroxygenase-deficient cells may also be used to express heterologous proteins of pharmaceutical interest such as hormones, growth factors, receptors, and the like.

In a further aspect, the present invention relates to a protein product essentially free from peroxygenase activity that is produced by a method of the present invention.

Fermentation Broth Formulations or Cell Compositions

The present invention also relates to a fermentation broth formulation or a cell composition comprising a polypeptide of the present invention. The fermentation broth product further comprises additional ingredients used in the fermentation process, such as, for example, cells (including, the host cells containing the gene encoding the polypeptide of the present invention which are used to produce the polypeptide of interest), cell debris, biomass, fermentation media and/or fermentation products. In some embodiments, the composition is a cell-killed whole broth containing organic acid(s), killed cells and/or cell debris, and culture medium.

The term "fermentation broth" as used herein refers to a preparation produced by cellular fermentation that undergoes no or minimal recovery and/or purification. For example, fermentation broths are produced when microbial cultures are grown to saturation, incubated under carbon-limiting conditions to allow protein synthesis (*e.g.*, expression of enzymes by host cells) and secretion into cell culture medium. The fermentation broth can contain unfractionated or fractionated contents of the fermentation materials derived at the end of the fermentation. Typically, the fermentation broth is unfractionated and comprises the spent culture medium and

cell debris present after the microbial cells (e.g., filamentous fungal cells) are removed, e.g., by centrifugation. In some embodiments, the fermentation broth contains spent cell culture medium, extracellular enzymes, and viable and/or nonviable microbial cells.

In an embodiment, the fermentation broth formulation and cell compositions comprise a first organic acid component comprising at least one 1-5 carbon organic acid and/or a salt thereof and a second organic acid component comprising at least one 6 or more carbon organic acid and/or a salt thereof. In a specific embodiment, the first organic acid component is acetic acid, formic acid, propionic acid, a salt thereof, or a mixture of two or more of the foregoing and the second organic acid component is benzoic acid, cyclohexanecarboxylic acid, 4-methylvaleric acid, phenylacetic acid, a salt thereof, or a mixture of two or more of the foregoing.

In one aspect, the composition contains an organic acid(s), and optionally further contains killed cells and/or cell debris. In one embodiment, the killed cells and/or cell debris are removed from a cell-killed whole broth to provide a composition that is free of these components.

The fermentation broth formulations or cell compositions may further comprise a preservative and/or anti-microbial (e.g., bacteriostatic) agent, including, but not limited to, sorbitol, sodium chloride, potassium sorbate, and others known in the art.

The cell-killed whole broth or composition may contain the unfractionated contents of the fermentation materials derived at the end of the fermentation. Typically, the cell-killed whole broth or composition contains the spent culture medium and cell debris present after the microbial cells (e.g., filamentous fungal cells) are grown to saturation, incubated under carbon-limiting conditions to allow protein synthesis. In some embodiments, the cell-killed whole broth or composition contains the spent cell culture medium, extracellular enzymes, and killed filamentous fungal cells. In some embodiments, the microbial cells present in the cell-killed whole broth or composition can be permeabilized and/or lysed using methods known in the art.

A whole broth or cell composition as described herein is typically a liquid, but may contain insoluble components, such as killed cells, cell debris, culture media components, and/or insoluble enzyme(s). In some embodiments, insoluble components may be removed to provide a clarified liquid composition.

The whole broth formulations and cell compositions of the present invention may be produced by a method described in WO 90/15861 or WO 2010/096673.

Compositions

The peroxygenase polypeptides of the invention may be added to and thus become a component of a detergent composition.

The detergent composition of the present invention may be formulated, for example, as a hand or machine laundry detergent composition including a laundry additive composition

suitable for pre-treatment of stained fabrics and a rinse added fabric softener composition, or be formulated as a detergent composition for use in general household hard surface cleaning operations, or be formulated for hand or machine dishwashing operations.

In a specific aspect, the present invention provides a detergent additive comprising a polypeptide of the invention as described herein.

The detergent composition may comprise one or more surfactants, which may be anionic and/or cationic and/or non-ionic and/or semi-polar and/or zwitterionic, or a mixture thereof. In a particular embodiment, the detergent composition includes a mixture of one or more nonionic surfactants and one or more anionic surfactants. The surfactant(s) is typically present at a level of from about 0.1% to 60% by weight, such as about 1% to about 40%, or about 3% to about 20%, or about 3% to about 10%. The surfactant(s) is chosen based on the desired cleaning application, and includes any conventional surfactant(s) known in the art.

When included therein the detergent will usually contain from about 1% to about 40% by weight, such as from about 5% to about 30%, including from about 5% to about 15%, or from about 20% to about 25% of an anionic surfactant. Non-limiting examples of anionic surfactants include sulfates and sulfonates, in particular, linear alkylbenzenesulfonates (LAS), isomers of LAS, branched alkylbenzenesulfonates (BABS), phenylalkanesulfonates, alpha-olefinsulfonates (AOS), olefin sulfonates, alkene sulfonates, alkane-2,3-diylbis(sulfates), hydroxyalkanesulfonates and disulfonates, alkyl sulfates (AS) such as sodium dodecyl sulfate (SDS), fatty alcohol sulfates (FAS), primary alcohol sulfates (PAS), alcohol ethersulfates (AES or AEOS or FES, also known as alcohol ethoxysulfates or fatty alcohol ether sulfates), secondary alkanesulfonates (SAS), paraffin sulfonates (PS), ester sulfonates, sulfonated fatty acid glycerol esters, alpha-sulfo fatty acid methyl esters (alpha-SFMe or SES) including methyl ester sulfonate (MES), alkyl- or alkenylsuccinic acid, dodecenyl/tetradecenyl succinic acid (DTSA), fatty acid derivatives of amino acids, diesters and monoesters of sulfo-succinic acid or soap, and combinations thereof.

When included therein the detergent will usually contain from about 0.2% to about 40% by weight of a non-ionic surfactant, for example from about 0.5% to about 30%, in particular from about 1% to about 20%, from about 3% to about 10%, such as from about 3% to about 5%, or from about 8% to about 12%. Non-limiting examples of non-ionic surfactants include alcohol ethoxylates (AE or AEO), alcohol propoxylates, propoxylated fatty alcohols (PFA), alkoxyated fatty acid alkyl esters, such as ethoxylated and/or propoxylated fatty acid alkyl esters, alkylphenol ethoxylates (APE), nonylphenol ethoxylates (NPE), alkylpolyglycosides (APG), alkoxyated amines, fatty acid monoethanolamides (FAM), fatty acid diethanolamides (FADA), ethoxylated fatty acid monoethanolamides (EFAM), propoxylated fatty acid monoethanolamide (PFAM), polyhydroxy alkyl fatty acid amides, or N-acyl N-alkyl derivatives of glucosamine

(glucamides, GA, or fatty acid glucamide, FAGA), as well as products available under the trade names SPAN and TWEEN, and combinations thereof.

The detergent composition may contain about 0-65% by weight of a detergent builder or co-builder, or a mixture thereof. In a dish wash detergent, the level of builder is typically 40-65%, particularly 50-65%. The builder and/or co-builder may particularly be a chelating agent that forms water-soluble complexes with Ca and Mg. Any builder and/or co-builder known in the art for use in laundry detergents may be utilized. Non-limiting examples of builders include zeolites, diphosphates (pyrophosphates), triphosphates such as sodium triphosphate (STP or STPP), carbonates such as sodium carbonate, soluble silicates such as sodium metasilicate, layered silicates (e.g., SKS-6 from Hoechst), ethanolamines such as 2-aminoethan-1-ol (MEA), iminodiethanol (DEA) and 2,2',2''-nitrilotriethanol (TEA), and carboxymethylinulin (CMI), and combinations thereof.

The detergent composition may contain 0-50% by weight of a bleaching system. Any bleaching system known in the art for use in laundry detergents may be utilized. Suitable bleaching system components include bleaching catalysts, photobleaches, bleach activators, sources of hydrogen peroxide such as sodium percarbonate and sodium perborates, preformed peracids and mixtures thereof. Suitable preformed peracids include, but are not limited to, peroxycarboxylic acids and salts, percarbonic acids and salts, perimidic acids and salts, peroxymonosulfuric acids and salts, for example, Oxone (R), and mixtures thereof. Non-limiting examples of bleaching systems include peroxide-based bleaching systems, which may comprise, for example, an inorganic salt, including alkali metal salts such as sodium salts of perborate (usually mono- or tetra-hydrate), percarbonate, persulfate, perphosphate, persilicate salts, in combination with a peracid-forming bleach activator. By Bleach activator is meant herein a compound which reacts with peroxygen bleach like hydrogen peroxide to form a Peracid. The peracid thus formed constitutes the activated bleach. Suitable bleach activators to be used herein include those belonging to the class of esters amides, imides or anhydrides, Suitable examples are tetracetyl ethylene diamine (TAED), sodium 3,5,5 trimethyl hexanoyloxybenzene sulphonate, diperoxy dodecanoic acid, 4-(dodecanoyloxy)benzenesulfonate (LOBS), 4-(decanoyloxy)benzenesulfonate, 4-(decanoyloxy)benzoate (DOBS), 4-(3,5,5-trimethylhexanoyloxy)benzenesulfonate (ISONOBS), tetraacetylenediamine (TAED) and 4-(nonanoyloxy)benzenesulfonate (NOBS), and/or those disclosed in WO98/17767. A particular family of bleach activators of interest was disclosed in EP624154 and particularly preferred in that family is acetyl triethyl citrate (ATC). ATC or a short chain triglyceride like Triacin has the advantage that it is environmental friendly as it eventually degrades into citric acid and alcohol. Furthermore acetyl triethyl citrate and triacetin has a good hydrolytical stability in the product upon storage and it is an efficient bleach activator. Finally ATC provides a good building capacity to the laundry additive. Alternatively, the bleaching system may comprise peroxyacids

of, for example, the amide, imide, or sulfone type. The bleaching system may also comprise peracids such as 6-(phthaloylamino)percapronic acid (PAP). The bleaching system may also include a bleach catalyst.

Other ingredients of the detergent composition, which are all well-known in art, include hydrotropes, fabric hueing agents, anti-foaming agents, soil release polymers, anti-redeposition agents etc.

The detergent additive as well as the detergent composition may comprise one or more additional enzymes such as a protease, lipase, cutinase, amylase, carbohydrase, cellulase, pectinase, mannanase, arabinase, galactanase, xylanase, oxidase, e.g., a laccase, and/or peroxidase.

The polypeptide of the present invention may be added to a detergent composition in an amount corresponding to 0.001-100 mg of protein, such as 0.01-100 mg of protein, preferably 0.005-50 mg of protein, more preferably 0.01-25 mg of protein, even more preferably 0.05-10 mg of protein, most preferably 0.05-5 mg of protein, and even most preferably 0.01-1 mg of protein per liter of wash liquor.

The polypeptide having peroxygenase activity (the peroxygenase), and optionally also a source of hydrogen peroxide, may be formulated as a liquid (e.g. aqueous), a solid, a gel, a paste or a dry product formulation. The dry product formulation may subsequently be re-hydrated to form an active liquid or semi-liquid formulation usable in the methods of the invention.

When the peroxygenase and the source of hydrogen peroxide are formulated as a dry formulation, the components may be mixed, arranged in discrete layers or packaged separately.

When other than dry form formulations are used, and even in that case, it is preferred to use a two-part formulation system having the peroxygenase separate from the source of hydrogen peroxide.

The composition of the invention may further comprise auxiliary agents such as wetting agents, thickening agents, buffer(s) for pH control, stabilisers, perfume, colourants, fillers and the like.

Useful wetting agents are surfactants, i.e. non-ionic, anionic, amphoteric or zwitterionic surfactants. Surfactants are further described above.

Methods and Uses

The peroxygenase polypeptides of the invention may be used for site specific hydroxylation in position 2 or position 3 of an aliphatic hydrocarbon. The aliphatic hydrocarbon must include a chain of at least 3 carbons, and either (one or more) end of the aliphatic hydrocarbon may be used as the starting point to determine which carbon is in position 2 or 3. The aliphatic hydrocarbon must have at least one hydrogen attached to the carbon (which is

hydroxylated) in position 2 or 3. In a preferred embodiment, the carbon in position 2 or 3, which is hydroxylated with the peroxygenase, is unsubstituted (before the hydroxylation is carried out).

Accordingly, in a first aspect, the present invention provides a method for hydroxylation in position 2 or 3 of either end (one or more ends) of a substituted or unsubstituted, linear or branched, aliphatic hydrocarbon having at least 3 carbons and having a hydrogen attached to the carbon in position 2 or 3, comprising contacting the aliphatic hydrocarbon with hydrogen peroxide and a polypeptide having peroxygenase activity of the invention.

The method of the invention may be used for a variety of purposes, like bulk chemical synthesis (biocatalysis), increasing aqueous solubility of aliphatic hydrocarbons, bioremediation, and modification of the characteristics of food products.

The method of the invention may also be used for a number of industrial processes in which said hydroxylation reactions are beneficial. An example of such use is in the manufacture of pulp and paper products where alkanes and other relevant aliphatic hydrocarbons that are present in the wood (resin) can result in deposition problems in the pulp and paper manufacturing process. These hydrophobic compounds are the precursors of the so-called pitch deposits within the pulp and paper manufacturing processes. Pitch deposition results in low quality pulp, and can cause the shutdown of pulp mill operations. Specific issues related to pulps with high extractives content include runnability problems, spots and holes in the paper, and sheet breaks. Treatment with peroxygenase can increase the solubility of said compounds and thereby mitigate problems.

Yet another use of the method of the invention is in i.e. oil or coal refineries where the peroxygenase catalyzed hydroxylation can be used to modify the solubility, viscosity and/or combustion characteristics of hydrocarbons. Specifically the treatment can lead to changes in the smoke point, the kindling point, the fire point and the boiling point of the hydrocarbons subjected to the treatment.

In the synthesis of bulk chemicals, agro chemicals (incl. pesticides), specialty chemicals and pharmaceuticals the method of the invention may obviously be relevant in terms of selectively introducing hydroxy groups in the substrates thereby affecting the solubility of the modified compound. Furthermore, the selective hydroxylation provides a site for further modification by methods known in the art of organic chemical synthesis and chemo-enzymatic synthesis.

Natural gas is extensively processed to remove higher alkanes. Hydroxylation of such higher alkanes may be used to improve water solubility, and thus facilitate removal of the higher alkanes by washing the natural gas stream. Removal may be performed at the well or during refining.

Hydroxylation of oil waste will significantly improve biodegradability and will be applicable both in connection with waste water treatment from refineries and bioremediation of contaminated ground or water

In a second aspect, the present invention provides a method for hydroxylation in position 2 or 3 of the terminal end of an acyl group of a lipid, comprising contacting the lipid with hydrogen peroxide and a polypeptide having peroxygenase activity of the invention.

Hydroxylation of the acyl group of a lipid generally improves the aqueous solubility of the lipid. Accordingly, the method of the invention may be used to remove or reduce oil or lipid containing stains, like chocolate, from laundry, by contacting the laundry with a peroxygenase and a source of hydrogen peroxide, and optionally a surfactant.

In another aspect, the methods of the invention may be used to reduce unpleasant odors from laundry by contacting the laundry with a peroxygenase and a source of hydrogen peroxide, and optionally a surfactant. The method of the invention results in reduction of the amount of butanoic acid (butyric acid) in the laundry. Butanoic acid is formed during washing of laundry when certain animal fats and plant oils are hydrolyzed, e.g. by detergent lipase, to yield free fatty acids, including butanoic acid. Butanoic acid has an extremely unpleasant odor. The peroxygenase hydroxylates the butanoic acid to 2-hydroxybutyric acid (*alpha*-hydroxybutyric acid) or 3-hydroxybutyric acid (*beta*-hydroxybutyric acid).

The present invention also provides a method for site specific introduction of a hydroxy and/or an oxo (keto) group at the second or third carbon of at least two ends of an aliphatic hydrocarbon, using a peroxygenase polypeptide of the invention, and hydrogen peroxide.

The aliphatic hydrocarbon must include a chain of at least five carbons. The second and third carbons are determined by counting the carbon atoms from any end of the aliphatic hydrocarbon.

The aliphatic hydrocarbon must have at least one hydrogen attached to a carbon which is hydroxylated by attachment of a hydroxy group; and at least two hydrogens attached to a carbon when an oxo group is introduced. In a preferred embodiment, the second or third carbon is unsubstituted before being contacted with the peroxygenase.

According to the method of the invention, the hydroxy and/or oxo groups are introduced independently of each other at the (at least) two ends of the aliphatic hydrocarbon. Thus, a hydroxy group can be introduced at one end, at the same time as an oxo group is introduced at another (the other) end - and vice versa. Two hydroxy groups, or two oxo groups, or one hydroxy group and one oxo group, cannot be introduced at the same end of the aliphatic hydrocarbon.

In the context of the present invention, "oxidation" means introduction of a hydroxy and/or an oxo group.

Accordingly, in a first aspect, the present invention provides a method for introducing a hydroxy and/or an oxo (keto) group at the second or third carbon of (at least) two ends of a substituted or unsubstituted, linear or branched, aliphatic hydrocarbon having at least five carbons and having at least one hydrogen attached to said second or third carbon, comprising
5 contacting the aliphatic hydrocarbon with hydrogen peroxide and a polypeptide having peroxygenase activity of the invention.

In a preferred embodiment, the aliphatic hydrocarbon is oxidized to (converted to) a diol, by introduction of two hydroxy groups. More preferably, the two hydroxy groups are located at each end of a linear aliphatic hydrocarbon.

10 The method of the invention may be used for a variety of purposes, like bulk chemical synthesis (biocatalysis), increasing aqueous solubility of aliphatic hydrocarbons, bioremediation, and modification of the characteristics of food products.

The method of the invention may also be used for a number of industrial processes in which said oxidation reactions are beneficial. An example of such use is in the manufacture of
15 pulp and paper products where alkanes and other relevant aliphatic hydrocarbons that are present in the wood (resin) can result in deposition problems in the pulp and paper manufacturing process. These hydrophobic compounds are the precursors of the so-called pitch deposits within the pulp and paper manufacturing processes. Pitch deposition results in low quality pulp, and can cause the shutdown of pulp mill operations. Specific issues related to
20 pulps with high extractives content include runnability problems, spots and holes in the paper, and sheet breaks. Treatment with peroxygenase can increase the solubility of said compounds and thereby mitigate problems.

Yet another use of the method of the invention is in, for example, oil or coal refineries where the peroxygenase catalyzed oxidation can be used to modify the solubility, viscosity
25 and/or combustion characteristics of hydrocarbons. Specifically the treatment can lead to changes in the smoke point, the kindling point, the fire point and the boiling point of the hydrocarbons subjected to the treatment.

In the synthesis of bulk chemicals, agro chemicals (incl. pesticides), specialty chemicals and pharmaceuticals the method of the invention may obviously be relevant in terms of
30 selectively introducing hydroxy groups in the substrates thereby affecting the solubility of the modified compound. Furthermore, the selective oxidation provides a site for further modification by methods known in the art of organic chemical synthesis and chemo-enzymatic synthesis.

Natural gas is extensively processed to remove higher alkanes. Oxidation of such higher alkanes may be used to improve water solubility, and thus facilitate removal of the higher
35 alkanes by washing the natural gas stream. Removal may be performed at the well or during refining.

Oxidation, according to the invention, of oil waste will significantly improve biodegradability and will be applicable both in connection with waste water treatment from refineries and bioremediation of contaminated ground or water

5 The methods of the invention may be carried out with an immobilized peroxygenase polypeptide of the invention.

The methods of the invention may be carried out in an aqueous solvent (reaction medium), various alcohols, ethers, other polar or non-polar solvents, or mixtures thereof. By studying the characteristics of the aliphatic hydrocarbon used in the methods of the invention, suitable examples of solvents are easily recognized by one skilled in the art. By raising or
10 lowering the pressure at which the hydroxylation/oxidation is carried out, the solvent (reaction medium) and the aliphatic hydrocarbon can be maintained in a liquid phase at the reaction temperature.

The methods according to the invention may be carried out at a temperature between 0 and 90 degrees Celsius, preferably between 5 and 80 degrees Celsius, more preferably
15 between 10 and 70 degrees Celsius, even more preferably between 15 and 60 degrees Celsius, most preferably between 20 and 50 degrees Celsius, and in particular between 20 and 40 degrees Celsius.

The methods of the invention may employ a treatment time of from 10 seconds to (at least) 24 hours, preferably from 1 minute to (at least) 12 hours, more preferably from 5 minutes
20 to (at least) 6 hours, most preferably from 5 minutes to (at least) 3 hours, and in particular from 5 minutes to (at least) 1 hour.

Diols (di-hydroxy aliphatic hydrocarbons) produced by the method of the invention may be used for producing polyurethan. Polyurethane is a polymer composed of a chain of organic units joined by carbamate (urethane) links. Polyurethane polymers are formed through step-growth
25 polymerization, by reacting a monomer (with at least two isocyanate functional groups) with another monomer (with at least two hydroxyl groups) in the presence of a catalyst.

The present invention also provides a method for introducing an oxo (keto) group at the second or third carbon of a substituted or unsubstituted, linear or branched, aliphatic
30 hydrocarbon having at least five carbons and having at least two hydrogens attached to said second or third carbon, comprising contacting the aliphatic hydrocarbon with hydrogen peroxide and a polypeptide having peroxygenase activity of the invention.

In yet another aspect, the present invention also provides a method for introducing a
35 hydroxy or an oxo group at a terminal carbon of a linear or branched aliphatic hydrocarbon having at least five carbons, which is substituted with a carboxy group, comprising contacting

the aliphatic hydrocarbon with hydrogen peroxide and a polypeptide having peroxygenase activity of the invention.

In an embodiment, the aliphatic hydrocarbon which is substituted with a carboxy group is a fatty acid; preferably butanoic acid (butyric acid), pentanoic acid (valeric acid), hexanoic acid (caproic acid), heptanoic acid (enanthic acid), octanoic acid (caprylic acid), nonanoic acid (pelargonic acid), decanoic acid (capric acid), dodecanoic acid (lauric acid), tetradecanoic acid (myristic acid), hexadecanoic acid (palmitic acid), octadecanoic acid (stearic acid), eicosanoic acid (arachidic acid), linoleic acid, linolenic acid, arachidonic acid, eicosapentaenoic acid, or docosahexaenoic acid.

In yet another aspect, the present invention also provides a method for changing (oxidizing) a primary alcohol of a linear or branched aliphatic hydrocarbon having at least five carbons to the corresponding acid, comprising contacting the alcohol of an aliphatic hydrocarbon with hydrogen peroxide and a polypeptide having peroxygenase activity of the invention.

For example, pentanol may be changed (oxidized) to pentanoic acid (valeric acid), hexanol may be changed to hexanoic acid (caproic acid), heptanol may be changed to heptanoic acid (enanthic acid), octanol may be changed to octanoic acid (caprylic acid), nonanol may be changed to nonanoic acid (pelargonic acid), decanol may be changed to decanoic acid (capric acid), dodecanol may be changed to dodecanoic acid (lauric acid), tetradecanol may be changed to tetradecanoic acid (myristic acid), hexadecanol may be changed to hexadecanoic acid (palmitic acid), octadecanol may be changed to octadecanoic acid (stearic acid), and eicosanol may be changed to eicosanoic acid (arachidic acid).

The polypeptides having peroxygenase activity of the invention (peroxygenase polypeptides or peroxygenases) are used in the methods of the invention in an amount of 0.005-50 ppm (mg/l), or 0.01-40, 0.02-30, 0.03-25, 0.04-20, 0.05-15, 0.05-10, 0.05-5, 0.05-1, 0.05-0.8, 0.05-0.6, or 0.1-0.5 ppm. The amount of enzyme refers to mg of a well-defined enzyme preparation.

In the methods of the invention, the peroxygenase may be applied alone or together with an additional enzyme. The term "an additional enzyme" means at least one additional enzyme, e.g. one, two, three, four, five, six, seven, eight, nine, ten or even more additional enzymes.

The term "applied together with" (or "used together with") means that the additional enzyme may be applied in the same, or in another step of the method of the invention. The other process step may be upstream or downstream, as compared to the step in which the peroxygenase is used.

In particular embodiments the additional enzyme is an enzyme which has protease, lipase, xylanase, cutinase, oxidoreductase, cellulase, endoglucanase, amylase, mannanase, steryl esterase, and/or cholesterol esterase activity. Examples of oxidoreductase enzymes are enzymes with laccase, and/or peroxidase activity.

5 The term "a step" of a method means at least one step, and it could be one, two, three, four, five or even more method steps. In other words the peroxygenases of the invention may be applied in at least one method step, and the additional enzyme(s) may also be applied in at least one method step, which may be the same or a different method step as compared to the step where the peroxygenase is used.

10 The term "enzyme preparation" means a product containing at least one peroxygenase. The enzyme preparation may also comprise enzymes having other enzyme activities. In addition to the enzymatic activity, such a preparation preferably contains at least one adjuvant. Examples of adjuvants are buffers, polymers, surfactants and stabilizing agents.

15 **Hydrogen peroxide**

20 The hydrogen peroxide (or source of hydrogen peroxide) required by the peroxygenase may be provided as an aqueous solution of hydrogen peroxide or a hydrogen peroxide precursor for in situ production of hydrogen peroxide. Any solid entity which liberates upon dissolution a peroxide which is useable by peroxygenase can serve as a source of hydrogen peroxide. Compounds which yield hydrogen peroxide upon dissolution in water or an appropriate aqueous based medium include but are not limited to metal peroxides, percarbonates, persulphates, perphosphates, peroxyacids, alkylperoxides, acylperoxides, peroxyesters, urea peroxide, perborates and peroxydicarboxylic acids or salts thereof.

25 Another source of hydrogen peroxide is a hydrogen peroxide generating enzyme system, such as an oxidase together with a substrate for the oxidase. Examples of combinations of oxidase and substrate comprise, but are not limited to, amino acid oxidase (see e.g. US 6,248,575) and a suitable amino acid, glucose oxidase (see e.g. WO 95/29996) and glucose, lactate oxidase and lactate, galactose oxidase (see e.g. WO 00/50606) and galactose, and aldose oxidase (see e.g. WO 99/31990) and a suitable aldose.

30 By studying EC 1.1.3._, EC 1.2.3._, EC 1.4.3._, and EC 1.5.3._ or similar classes (under the International Union of Biochemistry), other examples of such combinations of oxidases and substrates are easily recognized by one skilled in the art.

35 Hydrogen peroxide or a source of hydrogen peroxide may be added at the beginning of or during the method of the invention, e.g. as one or more separate additions of hydrogen peroxide; or continuously as fed-batch addition. Typical amounts of hydrogen peroxide correspond to levels of from 0.001 mM to 25 mM, preferably to levels of from 0.005 mM to 5 mM, and particularly to levels of from 0.01 to 1 mM hydrogen peroxide. Hydrogen peroxide may

also be used in an amount corresponding to levels of from 0.1 mM to 25 mM, preferably to levels of from 0.5 mM to 15 mM, more preferably to levels of from 1 mM to 10 mM, and most preferably to levels of from 2 mM to 8 mM hydrogen peroxide.

5 Aliphatic Hydrocarbons

The hydrocarbons, which are hydroxylated in the method of the invention, are aliphatic hydrocarbons having a chain of at least 3 carbons, and having a hydrogen attached to the carbon in position 2 or 3. Preferably, the aliphatic hydrocarbon is an alkane or an alkene; more preferably, the aliphatic hydrocarbon is an alkane, such as propane, butane, pentane, hexane, 10 heptane, octane, nonane or decane, or isomers thereof.

The aliphatic hydrocarbons are linear or branched, but not cyclic, as site specific hydroxylation is not possible with cyclic hydrocarbons. Branched hydrocarbons correspond to isomers of linear hydrocarbons.

The aliphatic hydrocarbons are substituted or unsubstituted. Preferably, the aliphatic 15 hydrocarbons are unsubstituted, such as non-activated hydrocarbons.

When the aliphatic hydrocarbons are substituted (functional groups attached), the preferred substituents are halogen, hydroxyl, carboxyl, amino, nitro, cyano, thiol, sulphonyl, formyl, acetyl, methoxy, ethoxy, phenyl, benzyl, xylyl, carbamoyl and sulfamoyl; more preferred substituents are chloro, hydroxyl, carboxyl and sulphonyl; and most preferred substituents are 20 chloro and carboxyl.

The aliphatic hydrocarbons may be substituted by up to 10 substituents, up to 8 substituents, up to 6 substituents, up to 4 substituents, up to 2 substituents, or by up to one substituent.

In a preferred embodiment, the aliphatic hydrocarbon is a fatty acid (the substituent is a 25 carboxyl group). Examples of fatty acids include, but are not limited to, butanoic acid (butyric acid), pentanoic acid (valeric acid), hexanoic acid (caproic acid), heptanoic acid (enanthic acid), octanoic acid (caprylic acid), nonanoic acid (pelargonic acid), decanoic acid (capric acid), dodecanoic acid (lauric acid), tetradecanoic acid (myristic acid), hexadecanoic acid (palmitic acid), octadecanoic acid (stearic acid), eicosanoic acid (arachidic acid), linoleic acid, linolenic 30 acid, arachidonic acid, eicosapentaenoic acid, and docosahexaenoic acid.

In a second aspect, the aliphatic hydrocarbon is an acyl group of a lipid, such as a monoglyceride, diglyceride, triglyceride, phospholipid or sphingolipid; and the hydroxylation takes place in position 2 or position 3 of the terminal end of the acyl group. The acyl group must have at least one hydrogen attached to the carbon in position 2 or 3 of the terminal end. The 35 acyl group may be saturated or unsaturated, and optionally functional groups (substituents) may be attached. Examples of acyl groups include, but are not limited to, the acyl forms of butanoic acid (butyric acid), pentanoic acid (valeric acid), hexanoic acid (caproic acid), heptanoic acid

(enantic acid), octanoic acid (caprylic acid), nonanoic acid (pelargonic acid), decanoic acid (capric acid), dodecanoic acid (lauric acid), tetradecanoic acid (myristic acid), hexadecanoic acid (palmitic acid), octadecanoic acid (stearic acid), eicosanoic acid (arachidic acid), linoleic acid, linolenic acid, arachidonic acid, eicosapentaenoic acid, and docosaheptaenoic acid.

5

The present invention is further described by the following examples that should not be construed as limiting the scope of the invention.

EXAMPLES

10 Chemicals were commercial products of at least reagent grade. Unless otherwise indicated the percentages in the Examples are weight based.

Media and Solutions

15 AMG trace metal solution: per liter 14.3 g $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 2.5 g $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.5 g NiCl_2 , 13.8 g FeSO_4 , 8.5 g MnSO_4 , 3.0 g citric acid.

MDU-2BP: per liter 45 g maltose- $1\text{H}_2\text{O}$, 7 g yeast extract, 12 g KH_2PO_4 , 1 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 2 g K_2SO_4 , 5 g Urea, 1 g NaCl, 0.5 ml AMG trace metal solution pH 5.0.

EXAMPLE 1

20 Recombinant Expression and Purification of the Peroxygenase

The artificial peroxygenase amino acid sequence of SEQ ID NO: 2 was prepared based on the *Marasmius rotula* peroxygenase and homologous peroxygenases.

A synthetic gene encoding the peroxygenase (SEQ ID NO: 1) was cloned into an expression vector using restriction site BamHI and XhoI, thus creating an expression plasmid.

25 The plasmid was transformed into an *Aspergillus oryzae* host cell.

Expression of the peroxygenase was identified/verified as band on SDS-PAGE gel electrophoresis.

The strain expressing the peroxygenase was inoculated in 5 shake flask each containing 200 ml MDU-2BP and added protoporphyrin IX (final concentration 100 mg/L). The strain was
30 grown at 30°C for 4 days at 200 rpm. The culture broth was sterile-filtered before starting the purification.

The filtered culture broth was reduced to 100-200 mL avoiding protein precipitation by using ultra-filtration. 5 mM Tris buffer pH 8 was added until 1 L, and then the volume was again reduced to 100-200 mL using ultra-filtration. This step was repeated until the conductivity of the
35 sample matched the conductivity of buffer A of ion exchange chromatography: 25 mM Tris pH 8. The volume of sample was finally reduced to 100 mL using ultra-filtration.

A Q-sepharose column was used for ion exchange chromatography. The column was equilibrated with 25 mM Tris pH 8 buffer. Flow rate was 10 mL/min. A gradient 0-100% of 25 mM Tris with 0.5 M NaCl buffer pH 8 buffer was applied. Fractions with high absorbance at 280 and 420 nm were loaded to SDS-PAGE gel and the ones with high protein concentration were collected. The purified peroxygenase was used in the following Examples.

EXAMPLE 2

Oxidation of 4-nitrobenzodioxole (NBD)

Peroxygenases from EC 1.11.2.1 oxidize 4-nitrobenzodioxole (1,2-(Methylenedioxy)-4-nitrobenzene) to 4-nitrocatechol and the produced yellow color was quantified spectrophotometrically at 425 nm ($\epsilon_{425} = 9,700 \text{ M}^{-1}\text{cm}^{-1}$).

A 10 mM stock solution of 4-nitrobenzodioxole (98% pure, 161500 Aldrich) was prepared in acetonitrile. A 0.1 mg/ml stock of purified peroxygenase (mature peroxygenase of SEQ ID NO: 2) was prepared in de-ionized water. The final reaction mixture (0.2 mL) contained 1.0 mM 4-nitrobenzodioxole, 10% acetonitrile, 50 mM phosphate buffer pH 7, 0.005 mg/mL of peroxygenase, and 0.5 mM hydrogen peroxide. The reaction was started by addition of hydrogen peroxide. A SpectraMax Plus 384 plate reader was applied (kinetics at 30°C at 425 nm) using a 96 well microtitre plate from Nunc (no. 260836). Each sample was analysed in triplicates. Blanks prepared without addition of hydrogen peroxide were subtracted.

The peroxygenase oxidized 4-nitrobenzodioxole to 4-nitrocatechol. The concentration of 4-nitrocatechol was calculated by spectrophotometric measurements (absorption) at 425 nm, using a standard curve.

Table 1. Nitrobenzodioxole oxidation to nitrocatechol.

Time (seconds)	Concentration of 4-nitrocatechol (mM)
0	0.004
60	0.009
120	0.011
300	0.014

EXAMPLE 3

Oxidation of Veratryl alcohol

Oxidation of 1 mM veratryl alcohol (3,4-dimethoxybenzyl alcohol) with 0.5 mM H_2O_2 was carried out with 20% acetonitrile and 50 mM phosphate buffer pH 7, using 0.025 mg/mL of

purified peroxygenase (mature peroxygenase of SEQ ID NO: 2) in a total reaction volume of 0.2 mL. The reaction was performed at room temperature for 5 minutes.

The peroxygenase oxidised veratryl alcohol to veratryl aldehyde. The concentration of veratryl aldehyde was calculated by spectrophotometric measurements (absorption) at 310 nm, using a standard curve.

Table 2. Veratryl alcohol oxidation to veratryl aldehyde.

Time (seconds)	Concentration of veratryl aldehyde (mM)
0	0.004
60	0.043
120	0.075
300	0.133

EXAMPLE 4

Oxidation of Naphthalene

Oxidation of 1 mM naphthalene with 0.5 mM H₂O₂ was carried out with 20% acetonitrile and 50 mM phosphate buffer pH 7, using 0.025 mg/mL of purified peroxygenase (mature peroxygenase of SEQ ID NO: 2) in a total reaction volume of 0.2 mL. Reactions were performed at room temperature for 5 minutes.

The peroxygenase oxidised naphthalene to 1-naphthol. The concentration of 1-naphthol was calculated by spectrophotometric measurements (absorption) at 303 nm, using a standard curve.

Table 3. Naphthalene oxidation to 1-naphthol.

Time (seconds)	Concentration of 1-naphthol (mM)
0	0.005
60	0.023
120	0.034
300	0.046

EXAMPLE 5**Oxidation of ABTS**

Peroxygenases oxidize ABTS (2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) in the presence of hydrogen peroxide and the produced green color is quantified

spectrophotometrically at 405 nm ($\epsilon_{405}=36,800\text{M}^{-1}\text{cm}^{-1}$).

The reaction mixture contained 0.5 mM ABTS, 50 mM phosphate buffer pH 7, 0.005 mg/mL of purified peroxygenase (mature peroxygenase of SEQ ID NO: 2), 0.5 mM hydrogen peroxide, and water ad 0.2 mL.

The reaction was started by adding the peroxygenase supernatant to the other ingredients used in the assay. A SpectraMax microtitre plate reader from Molecular Devices was applied to monitor the change in absorbance at 405 nm in a 96 well microtitre plate at room temperature. Blanks prepared without addition of enzyme were included.

The increase in absorbance was recorded over 5 minutes and the results are shown in Table 4.

Table 4. ABTS oxidation.

Time (seconds)	Absorbance at 405 nm
0	0.02
60	0.22
120	0.43
300	1.18

EXAMPLE 6**Oxidation of Myristic acid**

Oxidation of 0.1 mM myristic acid with 0.5 mM hydrogen peroxide was carried out with 20% acetonitrile and 50 mM phosphate buffer pH 6.5, using 0.01 mg/mL of purified peroxygenase (mature peroxygenase of SEQ ID NO: 2) in a total reaction volume of 1 mL. Reactions were performed at room temperature for 30 minutes.

After the enzymatic reaction, water was immediately removed in a vacuum centrifuge. The compounds were derivatized in presence of BSTFA + TMCS 99:1 and pyridine at 70°C for 30 min. Samples were analyzed by GC/MS. The GC/MS analyses were performed using 7HG-G011-02 ZB-1MS GC capillary column (30 m x 0.25 mm x 0.10 μm) from Phenomenex. The oven was heated from 120°C (1 min) to 155°C at 20°C/min and held for 1 min, then to 185°C at 5°C/min and finally to 300°C at 20°C/min.

The products formed during enzymatic reaction and qualified by the MS were mainly hydroxylated fatty acids at the (ω -1), (ω -2) and (ω) positions. Di-carboxylic acid and ketones were also identified among the products.

CLAIMS

1. A polypeptide having peroxygenase activity, selected from the group consisting of:

(a) a polypeptide having at least 80% sequence identity to the mature polypeptide of SEQ ID

NO: 2;

(b) a polypeptide encoded by a polynucleotide that hybridizes under high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, or (ii) the full-length complement of (i);

(c) a polypeptide encoded by a polynucleotide having at least 80% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1;

(d) a variant of the mature polypeptide of SEQ ID NO: 2 comprising a substitution, deletion, and/or insertion at one or more positions; and

(e) a fragment of the polypeptide of (a), (b), (c), or (d) that has peroxygenase activity.

2. The polypeptide of claim 1, having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity to the mature polypeptide of SEQ ID NO: 2.

3. The polypeptide of claim 1 or 2, which is encoded by a polynucleotide that hybridizes under low stringency conditions, low-medium stringency conditions, medium stringency conditions, medium-high stringency conditions, high stringency conditions, or very high stringency conditions with (i) the mature polypeptide coding sequence of SEQ ID NO: 1, or (ii) the full-length complement of (i).

4. The polypeptide of any of claims 1-3, which is encoded by a polynucleotide having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity to the mature polypeptide coding sequence of SEQ ID NO: 1.

5. The polypeptide of any of claims 1-4, comprising or consisting of SEQ ID NO: 2 or the mature polypeptide of SEQ ID NO: 2.

6. The polypeptide of any of claims 1-5, wherein the mature polypeptide is amino acids 1 to 241 of SEQ ID NO: 2.

7. A whole broth formulation or cell culture composition comprising the polypeptide of any of claims 1-6.

8. A polynucleotide encoding the polypeptide of any of claims 1-6.

9. A nucleic acid construct or expression vector comprising the polynucleotide of claim 8 operably linked to one or more control sequences that direct the production of the polypeptide in an expression host.

10. A recombinant host cell comprising the polynucleotide of claim 8 operably linked to one or more control sequences that direct the production of the polypeptide.

11. A method of producing a polypeptide having peroxygenase activity, comprising cultivating the host cell of claim 10 under conditions conducive for production of the polypeptide.

12. The method of claim 11, further comprising recovering the polypeptide.

13. A detergent composition, comprising a surfactant and a polypeptide according to any of claims 1-6.

14. A method for hydroxylation in position 2 or 3 of either end of a substituted or unsubstituted, linear or branched, aliphatic hydrocarbon having at least 3 carbons and having a hydrogen attached to the carbon in position 2 or 3, comprising contacting the aliphatic hydrocarbon with hydrogen peroxide and a polypeptide according to any of claims 1-6.

15. A method for hydroxylation in position 2 or 3 of the terminal end of an acyl group of a lipid, comprising contacting the lipid with hydrogen peroxide and a polypeptide according to any of claims 1-6.

16. A method for introducing a hydroxy or a keto group at the second or third carbon of at least two ends of a substituted or unsubstituted, linear or branched, aliphatic hydrocarbon having at least five carbons and having at least one hydrogen attached to said second or third carbon, comprising contacting the aliphatic hydrocarbon with hydrogen peroxide and a polypeptide according to any of claims 1-6.

17. The method of any of claims 14-16, wherein the aliphatic hydrocarbon is an alkane.

18. The method of claim 17, wherein the alkane is pentane, hexane, heptane, octane, nonane, decane, undecane, dodecane, tridecane, tetradecane, pentadecane or hexadecane, or isomers thereof.

5 19. The method of any of claims 14-18, wherein the aliphatic hydrocarbon is unsubstituted.

20. The method of any of claims 14-19, wherein the aliphatic hydrocarbon is linear.

10 21. The method of any of claims 14-20, wherein the aliphatic hydrocarbon is converted to a diol, by introduction of two hydroxy groups.

22. The method of any of claims 14-16, wherein the aliphatic hydrocarbon is a fatty acid.

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2016/064710

A. CLASSIFICATION OF SUBJECT MATTER
 INV. C12N9/08
 ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data, EMBL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2013/021061 A1 (NOVOZYMES AS [DK]; LANDVIK SARA [DK]; OESTERGAARD LARS HENRIK [DK]; KA) 14 February 2013 (2013-02-14)	1,3,7-22
A	SEQ ID NO: 2 (accession no. BAL21719) page 2, lines 5-31 page 27, lines 11-26	2,4-6
X	WO 2013/021063 A2 (NOVOZYMES AS [DK]; LANDVIK SARA [DK]; OESTERGAARD LARS HENRIK [DK]; KA) 14 February 2013 (2013-02-14) SEQ ID NO: 1 (accession no. BAL21714) page 2, lines 5-33 page 27, lines 11-30	1,3,7-22



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"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)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search

6 September 2016

Date of mailing of the international search report

16/09/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040,
 Fax: (+31-70) 340-3016

Authorized officer

Barnas, Christoph

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2016/064710

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013021061 A1	14-02-2013	CN 103827297 A EP 2742129 A1 US 2014248674 A1 US 2016244731 A1 WO 2013021061 A1	28-05-2014 18-06-2014 04-09-2014 25-08-2016 14-02-2013
WO 2013021063 A2	14-02-2013	CN 103732740 A EP 2742060 A2 US 2014234918 A1 WO 2013021063 A2	16-04-2014 18-06-2014 21-08-2014 14-02-2013