



- (51) **International Patent Classification:**
C12N 9/10 (2006.01)
- (21) **International Application Number:**
PCT/EP2017/058327
- (22) **International Filing Date:**
7 April 2017 (07.04.2017)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
16164245.9 7 April 2016 (07.04.2016) EP
- (71) **Applicant:** ETH ZURICH [CH/CH]; Raemistrasse 101 / ETH transfer, 8092 Zurich (CH).
- (72) **Inventors:** AEBI, Markus; Bellevuestr. 4, 5430 Wettlingen (CH). KÜNZLER, Markus; Im Bungert 2, 5432 Neuenhof (CH). PIEL, Jörn; Minervastrasse 26, 8032 Zürich (CH). FREEMAN, Michael; 68 Bessbrook St., Feeding Hills, Massachusetts 01030 (US). VAN DER VELDEN, Niels; Scherpenkamp 19, 3992 RB Houten (NL). KÄLIN, Noemi; Alpstrasse 15, 6020 Emmenbrücke (CH).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM,

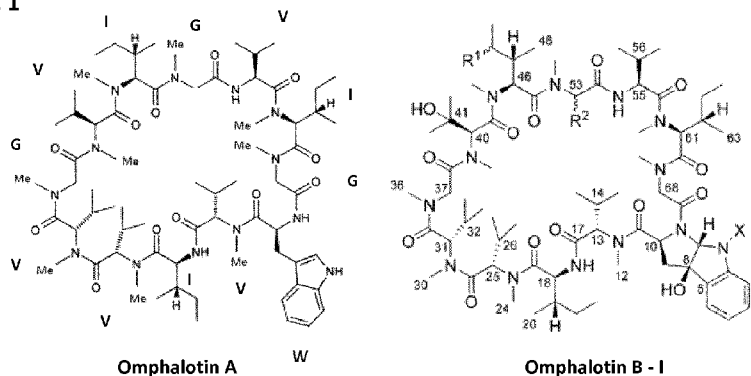
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, 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:

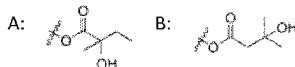
- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) **Title:** NOVEL MULTIPLY BACKBONE N-METHYL TRANSFERASES AND USES THEREOF

Fig. 1

- 1 Omphalotin B: $R^1 = A, R^2 = OH$
 2 Omphalotin C: $R^1 = A, R^2 = OAc$
 3 Omphalotin D: $R^1, R^2 = OAc$
 4 Omphalotin E: $R^1, R^2, X = H$
 5 Omphalotin F: $R^1, R^2 = H, X = OH$

- 6 Omphalotin G: $R^1 = H, R^2, X = OH$
 7 Omphalotin H: $R^1, R^2 = OAc, X = OH$
 8 Omphalotin I: $R^1 = B, R^2 = OAc, X = OH$



(57) **Abstract:** The present invention is directed to all aspects of novel methyl transferase enzymes that methylate backbone amides of (poly)peptides. The present invention also relates to nucleic acids encoding these enzymes as well as corresponding vectors and host cells comprising these. Moreover, the present invention encompasses the use of said enzymes for modifying (poly)peptides as well as corresponding methods. Also, the present invention pertains to further novel enzymes for modifying (poly)peptides derived from the omphalotin gene cluster of *O. olearius* and the homologous gene clusters from *D. bispora*, *L. edodes* and *F. mediterranea* as well as related aspects.

NOVEL MULTIPLY BACKBONE N-METHYL TRANSFERASES AND USES THEREOF

The present invention is directed to all aspects of novel methyl transferase enzymes that methylate backbone amides of (poly)peptides. The present invention also relates to nucleic acids encoding these enzymes as well as corresponding vectors and host cells comprising these. Moreover, the present invention encompasses the use of said enzymes for modifying (poly)peptides as well as corresponding methods. Also, the present invention pertains to further novel enzymes for modifying (poly)peptides derived from the omphalotin gene cluster of *O. olearius* and the homologous clusters from *D. bispora*, *L. edodes* and *F. mediterranea* as well as related aspects.

Background of the invention

Multiple backbone N-methylated (macrocyclic) peptides have interesting pharmacodynamic properties, e.g. increased bioavailability due to increased permeability for oral and intestinal epithelial membranes, increased half-life *in vivo* due to increased stability towards proteinases, reduced structural flexibility and the ability to disassemble aggregates and fibrils of their corresponding native non-methylated sequences (Chatterjee et al., (2008) *Accts Chem Res* 41:1331-1342, Chatterjee et al., (2013) *Angew Chem Int Ed* 52:254-269, Butterfield et al., (2012) *J. Mol. Biol.* 421:204-236). The chemical synthesis of such peptides is complicated and therefore very expensive (White et al., (2011) *Nat Chem Biol* 7:810-817). Multiply backbone N-methylated peptides like cyclosporin A are produced by architecturally highly complex non-ribosomal peptide synthetases (NRPSs) and N-methylated during elongation of the peptide by a built-in methyl transferase domain of the respective NRPS (Velkov et al., (2011) *Chem Biol* 18:464-475). Such a biosynthetic pathway is difficult to modify e.g. for the production of slightly altered versions, with regard e.g. to the peptide sequence or the N-methylation pattern, of the respective peptide or the production of peptide libraries because one has to reengineer the NRPS machinery which is very challenging (Thern et al., (2002) *Angew Chem Int Ed* 41:2307-2309, Lawen (2015) *Biochem Biophys Acta* 1850:2111-2120, Peel and Scribner, (2015) *Biochem Biophys Acta* 1850:2121-2144).

Large-scale production of backbone N-methylated (cyclic) peptides requires either fermentation followed by purification from the respective naturally occurring microbes (bacteria or fungi) or from recombinant bacteria carrying the respective NRPSs. The only other currently available alternative to provide such compounds is complicated chemical synthesis. The methods can also be

combined e.g. by chemical or enzymatic modification of the fermentation product (Peel and Scribner (2015) *Biochem Biophys Acta* 1850:2121-2144).

Ribosomally synthesized and post-translationally modified peptides, so-called RiPPs, also known as ribosomal peptide natural products, are a diverse class of natural products of ribosomal origin consisting of more than 20 subclasses that are produced by a variety of organisms, including prokaryotes, eukaryotes and archaea (Arnison PG et al., (2013) *Nat Prod Rep* 30:108-160). Ribosomally synthesized peptides have the advantage that they can be modified easily by simply changing the template/coding region of the peptide-encoding gene. So far, only two fungal RiPPs have been described, the ama/phallotoxins (Hallen et al., (2007) *Proc Natl Acad Sci USA* 104:19097-19101) and the ustiloxins (Umemura et al., (2014) *Fung Gen Biol* 68:23-30), and none of them are backbone N-methylated.

Omphalotin A was isolated in 1997 from the basidiomycete *Omphalotus olearius* (*O. olearius*) in order to find an active compound against the root-knot nematode *Meloidogyne incognita* (Mayer, (1997) *Nat. Prod. Lett.* 10:33-38). The omphalotins are highly backbone N-methylated, cyclic peptides consisting of 12 amino acids (Fig. 1, Atta-ur-Rahman (2015). "Studies in Natural Products Chemistry", Vol. 44 Chapter 3: 93-110). Due to the high degree of posttranslational modifications in the omphalotins (e.g. cyclization, hydroxylation's, O-acylation and backbone N-methylation) they are presently thought to belong to the class of non-ribosomally synthesized fungal peptides.

In view of the above, it is the objective technical problem underlying the present invention to provide new methylated (poly)peptides, preferably N-methylated (poly) peptides, more preferably backbone N-methylated (poly)peptides, and novel enzymatic tools for producing these as well as corresponding methods and uses of these enzymatic tools.

It is a further objective underlying the present invention to provide new enzyme tools for modifying (poly)peptides, in particular methylated (poly)peptides, preferably novel proteases, oxygenases, O-acyl transferases, and oxidoreductases with broad substrate specificity, to manufacture these enzymatic tools economically, and to provide corresponding (poly)peptide-modifying uses and methods.

The above objectives are solved in a first aspect by a preferably isolated and purified nucleic acid, comprising a nucleic acid sequence selected from the group consisting of:

- (i) a nucleic acid sequence selected from the group consisting of nucleic acid sequences listed in SEQ ID NOs: 1 to 14 and 108 to 111, preferably SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, and 108 (MT domains), more preferably SEQ ID NOs: 1 and 2;
- (ii) a nucleic acid sequence of at least 60 or 70 % identity, 80 or 90 % identity, preferably at least 95 % identity, more preferred at least 98 % identity with the nucleic acid sequence listed in SEQ ID NO 1 to 14 or 108 to 111, preferably over the whole sequence;
- (iii) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions;
- (iv) a fragment of any of the nucleic acid sequences of (i) to (iii), that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions; and
- (v) a nucleic acid sequence, wherein said nucleic acid sequence is derivable by substitution, addition and/or deletion of one of the nucleic acids of (i) to (iv) that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions,

wherein the nucleic acid sequence encodes a polypeptide having N-methyl transferase activity.

It was surprisingly found that the above nucleic acids encode the first enzymes to N-methylate the backbone of (poly)peptides / ribosomally synthesized (poly)peptides. The encoded enzymes define a new natural product (RiPP) class involved in the biosynthesis of the nematotoxin Omphalotin A.

It was found that the genome of *O. olearius* encompasses a gene containing a nucleic acid sequence that encodes the 12 amino acids of the omphalotin core peptide, followed by an additional six amino acids at the C-terminus and an N-terminal domain displaying homology to methyltransferases (hereinafter the methyl transferase and omphalotin core peptide gene is referred to as the omphalotin precursor gene, see Fig. 2). The omphalotin precursor gene is part of a gene cluster encoding a prolyl oligopeptidase (2090, SEQ ID NOs: 29, 30, 36 and 37), which cleaves and cyclizes the omphalotin peptide. In addition, two P450 monooxygenases (2085, SEQ ID NOs: 31 and 38; and 2089, SEQ ID NOs: 33, 34, 40 and 41), an O-acyl transferase (2088, SEQ ID NOs: 32 and 39) and an oxidoreductase (2091, SEQ ID NOs: 35 and 42) reside in the same gene cluster, which contribute to the further modifications found in omphalotin variants B-I (see Fig. 1 and 3). Similar gene clusters encoding methyl transferase genes were also identified in other fungi, e.g. *D. bispora*, *L. edodes* and *F. mediterranea* (Fig. 3 and 4).

The term “% (percent) identity” as known to the skilled artisan and used herein in the context of nucleic acids indicates the degree of relatedness among two or more nucleic acid molecules that is determined by agreement among the sequences. The percentage of “identity” is the result of the percentage of identical regions in two or more sequences while taking into consideration the gaps and other sequence peculiarities.

The identity of related nucleic acid molecules can be determined with the assistance of known methods. In general, special computer programs are employed that use algorithms adapted to accommodate the specific needs of this task. Preferred methods for determining identity begin with the generation of the largest degree of identity among the sequences to be compared. Preferred computer programs for determining the identity among two nucleic acid sequences comprise, but are not limited to, BLASTN (Altschul et al., (1990) J. Mol. Biol., 215:403-410) and LALIGN (Huang and Miller, (1991) Adv. Appl. Math., 12:337-357). The BLAST programs can be obtained from the National Center for Biotechnology Information (NCBI) and from other sources (BLAST handbook, Altschul et al., NCB NLM NIH Bethesda, MD 20894).

The nucleic acid molecules according to the invention may be prepared synthetically by methods well-known to the skilled person, but also may be isolated from suitable DNA libraries and other publicly available sources of nucleic acids and subsequently may optionally be mutated. The preparation of such libraries or mutations is well-known to the person skilled in the art.

In a preferred embodiment, the nucleic acid molecules of the invention are cDNA, genomic DNA, synthetic DNA, RNA or PNA, either double-stranded or single-stranded (i.e. either a sense or an anti-sense strand). The nucleic acid molecules and fragments thereof, which are encompassed within the scope of the invention, may be produced by, for example, polymerase chain reaction (PCR) or generated synthetically using DNA synthesis or by reverse transcription using mRNA from *O. olearius*, *D. bispora*, *L. edodes* or *F. mediterranea*.

In some instances, the present invention also provides novel nucleic acids encoding the (poly)-peptides of the present invention characterized in that they have the ability to hybridize to a specifically referenced nucleic acid sequence, preferably under stringent conditions. Next to common and/or standard protocols in the prior art for determining the ability to hybridize to a specifically referenced nucleic acid sequence under stringent conditions (e.g. Sambrook and Russell, (2001) Molecular cloning: A laboratory manual (3 volumes)), it is preferred to analyze and determine the ability to hybridize to a specifically referenced nucleic acid sequence under stringent conditions by

comparing the nucleotide sequences, which may be found in gene databases (e.g. <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=nucleotide> and <http://genome.jgi.doe.gov/-programs/fungi/index.jsf>) with alignment tools, such as e.g. the above-mentioned BLASTN (Altschul et al., (1990) J. Mol. Biol., 215:403-410), LALIGN alignment tools and multiple alignment tools such as e.g. CLUSTALW (Sievers F et al., (2011) Mol. Sys. Bio. 7: 539), MUSCLE (Edgar., (2004) Nucl. Acids Res. 32:1792-7) or T-COFFEE (Notredame et al., (2000) J of Mol. Bio 302 1: 205-17).

Most preferably, the ability of a nucleic acid of the present invention to hybridize to a nucleic acid, e.g. those listed in any of SEQ ID NOs 1 to 14 or 108 to 111, is confirmed in a Southern blot assay under the following conditions: 6x sodium chloride/sodium citrate (SSC) at 45°C followed by a wash in 0.2x SSC, 0.1% SDS at 65°C. Further preferred stringent conditions are selected from the group consisting of (a) 0.15 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50 °C for washing, (b) 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl and 75 mM sodium citrate at 42 °C during hybridization, and (c) hybridization in 50% formamide, 5xSSC, 5xDenhardt's solution, 50 g/mL sonicated salmon perm DNA, 0.1% SDS, and 10% dextran sulfate at 42 °C, with washes at 42 °C in 0.2xSSC and 0.1% SDS.

The nucleic acid of the present invention preferably encodes a polypeptide having N-methyl transferase activity for methylating peptidic backbone amides, preferably for multiply methylating peptidic backbone amides, more preferably for multiply methylating peptidic backbone amides of hydrophobic residues, in particular residues valine, glycine, and isoleucine (VGI). A method for confirming N-methyl transferase activity for methylating peptidic backbone amides is described further below.

The term "nucleic acid encoding a polypeptide" as used in the context of the present invention is meant to include allelic variations and redundancies in the genetic code.

In a preferred embodiment, reflecting the current US patent practice with regard to the patentability of products by nature, the nucleic acids of the present invention do not encode a naturally occurring methyl transferase, preferably do not encode a methyl transferase encoded by an omphalotin gene cluster selected from the group consisting of omphalotin gene clusters from *O. olearius*, *D. bispora*, *L. edodes* and *F. mediterranea*. More preferably, the claimed nucleic acids exclude any nucleic acid sequence selected from the group consisting of nucleic acid sequences SEQ ID NOs: 1 to 14 and 108 to 111, preferably SEQ ID NOs: 1, 3, 5, 7, 9, 11 and 13.

In a most preferred embodiment, the nucleic acid of the present invention further comprises a nucleic acid sequence encoding a (poly)peptide for N-methylation by the N-methyl transferase encoded by the nucleic acid of the invention. It is preferred that the nucleic acid of the present invention encodes a fusion protein with the methyl transferase domain at the N-terminus and the (poly)peptide for N-methylation at the C-terminus, optionally together with further amino acids at the C- and/or N-terminus and/or bridging the methyl transferase and the (poly)peptide for N-methylation. Preferred examples for fusion proteins comprising a methyl transferase domain according to the present invention and varying (poly)peptides for N-methylation are provided below. For example, the omphalotin peptide sequence (i.e. the sequence of the precursor peptide that is subject to N-methylation) can be exchanged for an omphalotin peptide sequence without follower sequence, for an omphalotin peptide sequence with an upstream TEV cleavage site or for a CycA, KenA, R1 and/or DicA peptide sequence. Specific experimental data for the aforementioned fusion proteins are provided in Example 6 and Figure 8 and 9. Said data demonstrate the robustness and versatility of the fusion proteins having methyl transferase activity according to the present invention and show that (i) the substrate for the methyl transferase can be varied largely in size and amino acid composition, (ii) a cleavage site, e.g. a TEV cleavage site, for separating the fusion proteins after methylation can be introduced into the fusion protein without hampering the methyl transferase efficacy, and (iii) that peptides can be methylated without a follower sequence.

The term “(poly)peptide for N-methylation” as used herein encompasses oligopeptides featuring 2 to 10 amino acids as well as polypeptides featuring more than 10 amino acids. The (poly)peptides for N-methylation for practicing the present invention preferably have 4 to 24, 6 to 20, or 8 to 16, more preferably 10 to 14, most preferably about 12 amino acids.

The nucleic acid of the present invention is preferably operably linked to a promoter that governs expression in suitable vectors and/or host cells producing the polypeptides of the present invention *in vitro* or *in vivo*.

Suitable promoters for operable linkage to the isolated and purified nucleic acid are known in the art. In a preferred embodiment the nucleic acid of the present invention is one that is operably linked to a promoter selected from the group consisting of the *Pichia pastoris* AOX1 promoter (see for example *Pichia* Expression Kit Instruction Manual, Invitrogen Corporation, Carlsbad, Calif.), the *Saccharomyces cerevisiae* GAL1, ADH1, GAP, ADH2, MET25, GPD, CUP1 or TEF promoter (see for example *Methods in Enzymology*, 350, 248, 2002), the Baculovirus polyhedrin p10 or ie1 promoter

(see for example Bac-to-Bac Expression Kit Handbook, Invitrogen Corporation, Carlsbad, Calif., and Novagen Insect Cell Expression Manual, Merck Chemicals Ltd., Nottingham, UK), the Lentivirus CMV, UbC, EF1 α , or MSCV promoter (see for example System Biosciences, Mountain View, CA, USA), the Adenovirus CMV promoter (see for example ViraPower Adenoviral Expression System, Life Technologies, Carlsbad, CA, USA), the *E. coli* T7, araBAD, rhaP BAD, tetA, lac, trc, tac or pL promoter (see Applied Microbiology and Biotechnology, 72, 211, 2006), the *B. subtilis*, vegI, vegII, σ_A , P_{grac}, P_{glv}, manP or P43 promoter (see Applied Microbiology and Biotechnology, 72, 211, 2006), the plant CaMV35S, ocs, nos, Adh-1, Tet promoters (see e.g. Lau and Sun, Biotechnol Adv. 2009, 27, 1015-22) or inducible promoters for mammalian cells as described in Sambrook and Russell (2001).

Preferably, the isolated and purified nucleic acid is in the form of a recombinant vector, such as an episomal or viral vector. The selection of a suitable vector and expression control sequences as well as vector construction are within the ordinary skill in the art. Preferably, the viral vector is a lentivirus vector (see for example System Biosciences, Mountain View, CA, USA), adenovirus vector (see for example ViraPower Adenoviral Expression System, Life Technologies, Carlsbad, CA, USA), baculovirus vector (see for example Bac-to-Bac Expression Kit Handbook, Invitrogen Corporation, Carlsbad, Calif.), bacterial vector (see for example Novagen, Darmstadt, Germany) or yeast vector (see for example ATCC Manassas, Virginia). Vector construction, including the operable linkage of a coding sequence with a promoter and other expression control sequences, is within the ordinary skill in the art.

Hence and in a further aspect, the present invention relates to a recombinant vector, comprising one or more nucleic acids of the invention.

A further aspect of the present invention is directed to a host cell comprising a nucleic acid and/or a vector of the invention and preferably producing polypeptides of the invention. Preferred host cells for producing the polypeptide of the invention are selected from the group consisting of yeast cells, preferably *Saccharomyces cerevisiae* (see for example Methods in Enzymology, 350, 248, 2002), *Pichia pastoris* cells (see for example *Pichia* Expression Kit Instruction Manual, Invitrogen Corporation, Carlsbad, Calif.), bacterial cells preferably *E. coli* cells (BL21(DE3), K-12 and derivatives) (see for example Applied Microbiology and Biotechnology, 72, 211, 2006) or *B. subtilis* cells (1012 wild type, 168 Marburg or WB800N) (see for example Westers et al., (2004) Mol. Cell. Res. Volume 1694, Issues 1-3 P: 299-310), plant cells, preferably *Nicotiana tabacum* or *Physcomitrella patens* (see e.g. Lau and Sun, Biotechnol Adv. 2009 May 18. [electronic publication ahead of print]), NIH-3T3

mammalian cells (see for example Sambrook and Russell, 2001) and insect cells, preferably sf9 insect cells (see for example Bac-to-Bac Expression Kit Handbook, Invitrogen Corporation, Carlsbad, Calif.).

Another important aspect of the present invention is directed to a preferably isolated and purified polypeptide selected from the group consisting of:

- (a) polypeptides having an amino acid sequence selected from the group consisting of SEQ ID NOs: 15-28 and 112 to 115, preferably SEQ ID NO: 16, 18, 20, 22, 24, 26, 28 and 112, more preferably SEQ ID NOs: 15 and 16.
 - (b) polypeptides encoded by a nucleic acid of the present invention,
 - (c) polypeptides having an amino acid sequence identity of at least 50, preferably at least 54, 60, 65, 70, 75, 80 or 85 %, preferably at least 90, 91, 92, 93, 94 or 95 % with the polypeptides of (a) and/or (b),
 - (d) a functional fragment and/or functional derivative of (a), (b) or (c),
- wherein the polypeptide has N-methyl transferase activity.

The polypeptide of the present invention preferably has N-methyl transferase activity for methylating peptidic backbone amides, preferably for multiply methylating peptidic backbone amides, more preferably multiply methylating peptide backbone amides of hydrophobic residues, in particular residues valine, glycine, and isoleucine (VGI).

The N-methyl transferase activity can be easily assayed and confirmed, e.g. by contacting the polypeptide suspected of N-methyl transferase activity with a suitable substrate (poly)peptide for N-methylation, e.g. WVIVVGIVGIVG (SEQ ID NO: 63), under suitable conditions. Identifying N-methylation, preferably in the backbone of the (poly)peptide substrate, e.g. by mass spectroscopy, NMR spectroscopy, etc. confirms enzymatic activity. For example, suitable conditions are presented in Example 6 below.

The identity of related amino acid molecules can be determined with the assistance of known methods. In general, special computer programs are employed that use algorithms adapted to accommodate the specific needs of this task. Preferred methods for determining identity begin with the generation of the largest degree of identity among the sequences to be compared. Preferred computer programs for determining the identity among two amino acid sequences comprise, but are not limited to, TBLASTN, BLASTP, BLASTX, TBLASTX (Altschul et al., (1990) J. Mol. Biol., 215, 403-410), ClustalW (Larkin MA et al., Bioinformatics, 23, 2947-2948, 2007) or PHYRE2 (Kelley LA et al., (2015) Nature Protocols 10, 845-858). The BLAST programs can be obtained from the National

Center for Biotechnology Information (NCBI) and from other sources (BLAST handbook, Altschul et al., NCB NLM NIH Bethesda, MD 20894). The ClustalW program can be obtained from <http://www.clustal.org> and the PHYRE2 program from <http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>.

The term “functional derivative” of a polypeptide of the present invention is meant to include any polypeptide or fragment thereof that has been chemically or genetically modified in its amino acid sequence, e.g. by addition, substitution and/or deletion of amino acid residue(s) and/or has been chemically modified in at least one of its atoms and/or functional chemical groups, e.g. by additions, deletions, rearrangement, oxidation, reduction, etc. as long as the derivative still has at least some methyl transferase activity to a measurable extent, e.g. of at least about 1 to 10 % methyl transferase activity of the original unmodified polypeptide of the invention, e.g. SEQ ID NOs: 15 to 28 or 112 to 115.

In this context a “functional fragment” of the invention is one that forms part of a polypeptide or derivative of the invention and still has at least some methyl transferase activity to a measurable extent, e.g. of at least about 1 to 10 % methyl transferase activity of the original unmodified polypeptide of the invention, e.g. SEQ ID NOs: 15 to 28 or 112 to 115.

The term “isolated and purified polypeptide” as used herein refers to a polypeptide or a peptide fragment which either has no naturally-occurring counterpart (e.g., a peptide mimetic), or has been separated or purified from components which naturally accompany it, e.g. in *O. olearius*, *D. bispora*, *L. edodes*, *F. mediterranea* or a fraction thereof. Preferably, a polypeptide is considered “isolated and purified” when it makes up for at least 60 % (w/w) of a dry preparation, thus being free from most naturally-occurring polypeptides and/or organic molecules with which it is naturally associated. Preferably, a polypeptide of the invention makes up for at least 80%, more preferably at least 90%, and most preferably at least 99% (w/w) of a dry preparation. More preferred are polypeptides according to the invention that make up for at least at least 80%, more preferably at least 90%, and most preferably at least 99% (w/w) of a dry polypeptide preparation. Chemically synthesized polypeptides are by nature “isolated and purified” within the above context.

An isolated polypeptide of the invention may be obtained, for example, by extraction from *O. olearius*, *D. bispora*, *L. edodes* or *F. mediterranea*; by expression of a recombinant nucleic acid encoding the polypeptide in a host, preferably a heterologous host; or by chemical synthesis. A polypeptide that is produced in a cellular system being different from the source from which it

naturally originates is “isolated and purified”, because it is separated from components which naturally accompany it. The extent of isolation and/or purity can be measured by any appropriate method, e.g. column chromatography, polyacrylamide gel electrophoresis, HPLC analysis, NMR spectroscopy, gas liquid chromatography, or mass spectrometry.

In a preferred embodiment, reflecting the current US patent practice with regard to the patentability of products of nature, the polypeptides of the present invention are not a naturally occurring methyl transferase, preferably are not a methyl transferase encoded by an omphalotin gene cluster selected from the group consisting of omphalotin gene clusters from *O. olearius*, *D. bispora*, *L. edodes* and *F. mediterranea*. More preferably the polypeptide does not consist of an amino acid sequence selected from the group consisting of amino acid sequences of SEQ ID NOs: SEQ ID NOs: 15 to 28 and 112 to 115, preferably 15, 17, 19, 21, 23, 25 and 27.

The methyl transferase of the present invention can, and preferably is present, e.g. expressed, together with the substrate (poly)peptide for N-methylation. Hence, the enzyme and its substrate form a fusion polypeptide, wherein the enzyme and its substrate may be separated by further bridging amino acids and the fusion protein may also comprise further amino acids at the C-terminus or N-terminus.

In a preferred embodiment, the present invention is also directed to a fusion polypeptide comprising a first polypeptide sequence encoding a methyl transferase according to the invention, and at least a second (poly)peptide sequence as substrate for N-methylation by the N-methyl transferase activity of the first polypeptide sequence.

Of course, the methyl transferase or a corresponding methyl transferase fusion protein of the present invention can also methylate (poly)peptide substrates without being fused to them, e.g. by contacting the enzyme and the substrate under typical physiological conditions allowing for enzyme activity.

Furthermore, in one aspect the present invention relates to antibodies, functional fragments and functional derivatives thereof or antibody-like binding proteins that specifically bind a polypeptide of the invention. As used herein, the term antibody is meant to include whole antibodies, functional fragments and functional derivatives thereof that specifically bind a polypeptide of the invention. These are routinely available by hybridoma technology (Kohler and Milstein, Nature 256, 495-497, 1975), antibody phage display (Winter et al., (1994) Annu. Rev. Immunol. 12, 433-455), ribosome display (Schaffitzel et al., (1999) J. Immunol. Methods, 231, 119-135) and iterative colony filter

screening (Giovannoni et al., (2001) *Nucleic Acids Res.* 29, E27) once the target antigen is available. Typical proteases for fragmenting antibodies into functional products are well-known. Other fragmentation techniques can be used as well as long as the resulting fragment has a specific high affinity and, preferably a dissociation constant in the micromolar to picomolar range.

A very convenient antibody fragment for targeting applications is the single-chain Fv fragment, in which a variable heavy and a variable light domain are joined together by a polypeptide linker. Other antibody fragments for identifying the polypeptide of the present invention include Fab fragments, Fab₂ fragments, miniantibodies (also called small immune proteins), tandem scFv-scFv fusions as well as scFv fusions with suitable domains (e.g. with the Fc portion of an immunoglobulin). For a review on certain antibody formats, see Holliger P, Hudson PJ.; *Engineered antibody fragments and the rise of single domains.* *Nat Biotechnol.* 2005 Sep., 23(9):1126-36).

The term “functional derivative” of an antibody for use in the present invention is meant to include any antibody or fragment thereof that has been chemically or genetically modified in its amino acid sequence, e.g. by addition, substitution and/or deletion of amino acid residue(s) and/or has been chemically modified in at least one of its atoms and/or functional chemical groups, e.g. by additions, deletions, rearrangement, oxidation, reduction, etc. as long as the derivative has substantially the same binding affinity as to its original antigen and, preferably, has a dissociation constant in the micro-, nano- or picomolar range.

In a preferred embodiment, the antibody, fragment or functional derivative thereof for use in the invention is one that is selected from the group consisting of polyclonal antibodies, monoclonal antibodies, chimeric antibodies, humanized antibodies, CDR-grafted antibodies, Fv-fragments, Fab-fragments and Fab₂-fragments and antibody-like binding proteins, e.g. affilines, anticalines and aptamers.

For a review of antibody-like binding proteins see Binz et al. on engineering binding proteins from non-immunoglobulin domains in *Nature Biotechnology*, Vol. 23, No. 10, October 2005, 1257-1268. The term “aptamer” describes nucleic acids that bind to a polypeptide with high affinity. Aptamers can be isolated from a large pool of different single-stranded RNA molecules by selection methods such as SELEX (see, e.g., Jayasena, *Clin. Chem.*, 45, p. 1628 – 1650, (1999); Klug and Famulok, *M. Mol. Biol. Rep.*, 20, p. 97 - 107 (1994); US 5,582,981). Aptamers can also be synthesized and selected in their mirror form, for example, as the L-ribonucleotide (Nolte et al., (1996) *Nat. Biotechnol.*, 14, pp.1116 – 1119; Klusmann et al., (1996) *Nat. Biotechnol.* 14, p. 1112 – 1115).

Forms isolated in this way have the advantage that they are not degraded by naturally occurring ribonucleases and, therefore, have a greater stability.

Another antibody-like binding protein and alternative to classical antibodies are the so-called “protein scaffolds”, for example, anticalines, that are based on lipocaline (Beste et al., (1999) Proc. Natl. Acad. Sci. USA, 96, p. 1898 - 1903). The natural ligand binding sites of lipocalines, for example, of the retinol-binding protein or bilin-binding protein, can be changed, for example, by employing a “combinatorial protein design” approach, and in such a way that they bind selected haptens (Skerra, (2000) Biochem. Biophys. Acta, 1482, pp. 337 – 350). For other protein scaffolds it is also known that they are alternatives for antibodies (Skerra, (2000) J. Mol. Recognit, 13, pp. 167 - 287; Hey, (2005) Trends in Biotechnology, 23, pp. 514-522).

In summary, the term functional antibody derivative is meant to include the above protein-derived alternatives for antibodies, i.e. antibody-like binding proteins, e.g. affilines, anticalines and aptamers that specifically recognize a polypeptide, fragment or derivative thereof.

A further aspect relates to a hybridoma cell line, expressing a monoclonal antibody according to the invention.

The nucleic acids, vectors, host cells, polypeptides and antibodies of the present invention have a number of new applications.

In a further aspect, the present invention is directed to the use of a vector, a host cell and/or a polypeptide according to the invention for producing an N-methylated, preferably a backbone N-methylated (poly)peptide.

In another aspect, the present invention relates to a method for producing an N-methylated (poly)peptide, preferably a backbone N-methylated (poly)peptide, more preferably a multiply backbone N-methylated (poly)peptide, most preferably a multiply N-methylated (poly)peptide backbone of hydrophobic residues (VGI) comprising the steps of:

- (a) providing a polypeptide having methyl transferase activity according to the present invention (termed hereinafter “MT”),
- (b) providing a (poly)peptide for N-methylation (termed hereinafter “substrate”),
- (c) contacting the polypeptide having methyl transferase activity (MT) of (a) and the (poly)peptide for N-methylation (substrate) of (b) under conditions that allow for N-methylation,
- (d) optionally cyclisation of the methylated peptide,

- (e) optionally further modifying the N-methylated (poly)peptide,
- (f) optionally cleaving the methylated peptide from the precursor,
- (g) optionally at least partially purifying the N-methylated (poly)peptide.

In a preferred embodiment, the MT of step (a) and/or the substrate of step (b) are provided by a vector and/or a host cell according to the present invention.

In a further preferred embodiment, the MT of step (a) and the substrate of step (b) are provided as a single fusion protein according to the present invention or as separate entities, preferably as a single fusion protein.

In yet another preferred embodiment, the conditions in step (c) allow for in vivo or in vitro N-methylation.

Preferably, the MT and the substrate are contacted intracellularly, more preferably expressed as a fusion protein or co-expressed in the same host cell, and the substrate is methylated within the host cell, preferably within *E. coli* or *S. cerevisiae*.

There are many further enzymes suited for modifying (poly)peptide substrates that can be used to further modify a substrate for use in the present invention before or after the N-methylation step.

In a very preferred embodiment the method according to the present invention is one, wherein in steps (d) and/or (e) at least one further enzyme and/or protein is added for further modifying the N-methylated (poly)peptide, preferably further enzymes and proteins selected from the group consisting of prolyl oligopeptidase (POP), monooxygenase, preferably P450 monooxygenase, O-acyl transferase, oxidoreductase, aldo/keto reductase, aspartic peptidase, and serine-type peptidase, more preferably enzymes having an amino acid sequence according to any one of SEQ ID NOs: 36 to 42, and 53 to 58 or functional fragments or functional derivatives thereof.

The method of the present invention can also be varied by modifying the substrate prior to N-methylation by the polypeptide of the invention, e.g. by the above referenced enzymes.

Preferably, in step (g) the at least partially purifying of the N-methylated (poly)peptide is accomplished using ammonium sulfate precipitation or general chromatography procedures such as anionic and cationic affinity chromatography, hydrophobic interaction chromatography or size-exclusion chromatography (e.g. see: Protein and peptide purification Technique selection guide (GE Healthcare, UK) or Sterner et al., (1997) Nat. Prod. Let. Volume 10, issue 1). The peptide can also be purified exploiting genetically introduced peptide or protein tags using affinity chromatography.

Preferred affinity chromatography methods are metal affinity chromatography, strep-tag-streptavidin affinity chromatography, glutathione-GST affinity chromatography, maltose-MBP affinity chromatography or FLAG-antiFLAG affinity chromatography (e.g. see: Affinity Chromatography Principles and Methods (GE Healthcare, UK) or Terpe et al., (2003) Appl. Microbiol. Biotechnol 60:523-533).

The omphalotin A gene cluster is the first example of a ribosomally synthesized multiply backbone N-methylated (cyclic) peptide and founds a novel class of ribosomally synthesized and post-translationally modified peptides, so-called RiPPs (Arnison PG et al., (2013) Nat Prod Rep 30:108-160). The structural gene coding for omphalotin A is unique in that the leader peptide is the MT domain that is responsible for the N-methylation of the backbone amides of the core peptide. MT is the first example of an enzyme that reliably N-methylates amides of existing peptide bonds in a ribosomally synthesized peptide or protein.

The proteins and methods of the present invention for N-methylating substrates have utility in (poly)peptide drug discovery, research and development in that they can be easily and economically used for modifying the properties of potentially medically active substrates or to disassemble aggregates and fibrils of their corresponding native non-methylated sequences. Next to designing new (poly)peptide-based drugs the methyl transferase enzymes can modify or attach a methylated peptide to known (poly)peptide-based drugs and lead to improved target-specificity, efficacy, stability, improved pharmacokinetics and/or improved oral and intestinal membrane permeability.

The substrates for use in the present invention are not limited to specific amino acid sequences and can be varied widely. All they require is the presence of at least one amide bond in a (poly)peptide or (poly)peptide-like compound, e.g. a depsipeptide.

On the nucleic acid level, preferably forming part of the methyl transferase enzyme as a fusion protein, the substrate can be easily, preferably randomly varied to provide a whole library of potentially suitable substrates that can be screened for medical utility. Hence, the nucleic acids, polypeptides, vectors and/or host cells can form part of a screening assay for identifying medically active (poly)peptides.

In addition, it is a further objective of the present invention to provide new, alternative and/or improved enzymatic tools for modifying (poly)peptides.

In a further independent aspect, the above objective technical problem is solved by the provision of new fungal enzymes and proteins, preferably selected from the group consisting of

prolyl oligopeptidase (POP), monooxygenase, preferably P450 monooxygenase, O-acyl transferase, oxidoreductase, aspartic peptidase and serine type peptidase.

The enzymes and/or proteins of the present invention are preferably derived from *O. olearius*, *D. bispora*, *L. edodes* or *F. mediterranea*. More preferably, these enzymes and/or proteins relate to a prolyl oligopeptidase (POP) derived from *O. olearius*, *D. bispora*, *L. edodes* or *F. mediterranea*.

In a first alternative aspect, the present invention relates to a preferably isolated and purified nucleic acid, comprising a nucleic acid sequence selected from the group consisting of:

- (i) a nucleic acid sequence selected from the group consisting of nucleic acid sequences listed in SEQ ID NOs: 29, 43, 45 and 47 (2090, 765750, 144537, 801057 (complete POP genes);
- (ii) a nucleic acid sequence of at least 60 or 70, 80 or 90 % identity, preferably at least 95 % identity, more preferred at least 98 % identity with a nucleic acid sequence listed in SEQ ID 29, 30, 43-48, preferably SEQ ID NOs: 30, 44, 46 and 48, more preferably SEQ ID NOs: 29 and 30, preferably over the whole sequence;
- (iii) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions;
- (iv) a fragment of any of the nucleic acid sequences of (i) to (iii), that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions; and
- (v) a nucleic acid sequence, wherein said nucleic acid sequence is derivable by substitution, addition and/or deletion of one of the nucleic acids of (i) to (iv) that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions,

wherein the nucleic acid sequence encodes a prolyl oligopeptidase (POP).

In a preferred embodiment, the nucleic acid encoding POP is a DNA, RNA or PNA, preferably DNA or PNA, more preferably DNA.

In a further preferred embodiment, the above nucleic acid encodes a polypeptide having peptidase and/or cyclase activity for polypeptides, preferably N-methylated (poly)peptides, more preferably for backbone N-methylated peptides, most preferably for multiply backbone N-methylated (poly)peptides.

In a preferred embodiment, reflecting the current US patent practice with regard to the patentability of products by nature, the nucleic acids of the present invention do not encode a naturally occurring prolyl oligopeptidase, preferably do not encode a prolyl oligopeptidase encoded by an omphalotin gene cluster selected from the group consisting of omphalotin gene clusters from

O. olearius, *D. bispora*, *L. edodes* and *F. mediterranea*. More preferably, the nucleic acid does not consist of a nucleic acid sequence selected from the group consisting of nucleic acid sequences SEQ ID NOs: 29, 30, 43-48, preferably 29, 43, 45 and 47.

In a further preferred embodiment, the present invention is directed to a recombinant vector comprising a nucleic acid encoding POP, preferably a viral or episomal vector, preferably a baculovirus vector, lentivirus vector, adenovirus vector, yeast or bacterial episomal vector.

In a further preferred embodiment, the present invention relates to a host cell comprising a nucleic acid encoding POP, or a vector of the present invention, preferably a host cell selected from the group consisting of yeast cells, preferably *Saccharomyces cerevisiae*, *Pichia pastoris* cells, *E. coli* cells, *B. subtilis* cells, plant cells, preferably *Nicotiana tabacum* or *Physcomitrella patens* cells, NIH-3T3 mammalian cells and insect cells, preferably sf9 insect cells.

In a further alternative aspect, the present invention is directed to a preferably isolated and purified polypeptide selected from the group consisting of:

- (a) polypeptides having an amino acid sequence selected from the group consisting of SEQ ID NOs: 36, 37 and 53-58, preferably SEQ ID NOs: 37, 54, 56 and 58, more preferably 36 and 37,
- (b) POP polypeptides encoded by a nucleic acid of the present invention,
- (c) polypeptides having an amino acid sequence identity of at least 50 or 60 %, at least 70 or 80 %, preferably at least 90 or 95 % with the polypeptides of (a) and/or (b),
- (d) a functional fragment and/or functional derivative of (a), (b) or (c),

wherein the polypeptide has prolyl oligopeptidase activity.

In a preferred embodiment, the POP polypeptide has peptidase and/or cyclase activity for (poly)peptides, preferably N-methylated (poly)peptides, more preferably for backbone N-methylated peptides, most preferably for multiply backbone N-methylated (poly)peptides.

In a further preferred embodiment, the POP polypeptide is not a naturally occurring prolyl oligopeptidase, preferably is not a prolyl oligopeptidase encoded by an omphalotin gene cluster selected from the group consisting of omphalotin gene clusters from *O. olearius*, *D. bispora*, *L. edodes* and *F. mediterranea*. More preferably, the polypeptide does not consist of an amino acid sequence selected from the group consisting of amino acid sequences SEQ ID NOs: 36, 37 and 53 to 58, preferably 36, 53, 55 and 57.

In a further alternative aspect, the present invention provides an antibody, preferably a monoclonal antibody, a functional fragment or functional derivative thereof or an antibody-like binding compound that specifically binds a POP polypeptide of the present invention.

In yet another alternative aspect, the present invention relates to a hybridoma cell line, expressing a monoclonal antibody that specifically binds a POP polypeptide according to the present invention.

Also, the present invention is directed to the use of a vector, a host cell and/or a POP polypeptide according to the present invention for producing a cleaved and/or cyclized (poly)peptide.

In another alternative aspect, the present invention teaches a method for producing a cleaved and/or cyclized (poly)peptide, comprising the steps of:

- (a) providing a polypeptide having prolyl oligopeptidase activity according to the present invention (termed hereinafter "POP"),
- (b) providing a (poly)peptide for peptide cleavage and/or cyclization (termed hereinafter "POP substrate"),
- (c) contacting the polypeptide having prolyl oligopeptidase activity of (a) and the (poly)peptide for peptide cleavage and/or cyclization of (b) under conditions that allow for cleavage and/or cyclization,
- (d) optionally further modifying the cleaved and/or cyclized (poly)peptide,
- (e) optionally at least partially purifying the cleaved and/or cyclized (poly)peptide.

Preferably, the POP of step (a) and/or the POP substrate of step (b) are provided by a vector and/or a host cell of the invention.

Preferably, in step (c), the conditions allow for intra- or extracellular cleavage and/or cyclization.

In a preferred embodiment the invention relates to a method, wherein in step (d) at least one further enzyme and/or protein is added for further modifying the cleaved and/or cyclized (poly)peptide, preferably further enzymes and proteins selected from the group consisting of methyl transferase (MT), preferably N-methyl transferase, more preferably backbone N-methyl transferase, monooxygenase, preferably P450 monooxygenase, O-acyl transferase, oxidoreductase, aldo/keto reductase, more preferably one or more enzymes having an amino acid sequence according to any one of SEQ ID Nos: 15 to 28, 112 to 115 and 38 to 42 or functional fragments or derivatives thereof.

In step (e), the at least partially purifying of the cleaved and/or cyclized (poly)peptide preferably comprises the same procedures identified above for at least partially purifying methyl transferase products.

In another aspect, the present invention is also directed to the (poly)peptide-modifying enzymes having nucleic acid and amino acid sequences listed in SEQ ID NOs: 31 to 35 and 38 to 42, respectively.

In a preferred embodiment, this aspect relates to a preferably isolated and purified nucleic acid, comprising a nucleic acid sequence selected from the group consisting of:

- (i) a nucleic acid sequence selected from the group consisting of nucleic acid sequences listed in SEQ ID NOs: 31 to 35;
- (ii) a nucleic acid sequence of at least 60 or 70, at least 80 or 90 % identity, preferably at least 95 % identity, more preferred at least 98 % identity with a nucleic acid sequence listed in SEQ ID NO 31 to 35, preferably over the whole sequence;
- (iii) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions;
- (iv) a fragment of any of the nucleic acid sequences of (i) to (iii), that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions; and
- (v) a nucleic acid sequence, wherein said nucleic acid sequence is derivable by substitution, addition and/or deletion of one of the nucleic acids of (i) to (iv) that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions,

wherein the nucleic acid sequence encodes a polypeptide having monooxygenase (SEQ ID NO: 31), O-acyl transferase (SEQ ID NO: 32), monooxygenase (SEQ ID NO: 33, 34), or oxidoreductase (SEQ ID NO: 35) activity.

In a second preferred embodiment, the present invention relates to an isolated and purified polypeptide selected from the group consisting of:

- (a) polypeptides having an amino acid sequence selected from the group consisting of SEQ ID NOs: 38 to 42,
- (b) polypeptides encoded by a nucleic acid of any of claims 1 to 5,
- (c) polypeptides having an amino acid sequence identity of at least 50 or 60 %, at least 70 or 80 %, preferably at least 90 or 95 % with the polypeptides of (a) and/or (b),
- (d) a functional fragment and/or functional derivative of (a), (b) or (c),

wherein the polypeptide has monooxygenase (SEQ ID NO: 38), O-acyl transferase (SEQ ID NO: 39), monooxygenase (SEQ ID NO: 40, 41), or oxidoreductase (SEQ ID NO: 42) activity.

Preferably, reflecting the current US patent practice with regard to the patentability of products by nature, the nucleic acids and amino acids of the present invention do not encode a naturally occurring (poly)peptide-modifying enzyme and/or protein, preferably do not encode an enzyme encoded by an omphalotin gene cluster selected from the group consisting of omphalotin gene clusters from *O. olearius*, *D. bispora*, *L. edodes* and *F. mediterranea*. More preferably, the nucleic acid and amino acid does not consist of a nucleic acid or amino acid sequence selected from the group consisting of nucleic acid and amino acid sequences SEQ ID NOs: 31-35 and 38-42.

The here-described novel monooxygenase (SEQ ID NO: 31), O-acyl transferase (SEQ ID NO: 32), monooxygenase (SEQ ID NO: 33, 34), and oxidoreductase (SEQ ID NO: 35) and their nucleic acid and amino acid derivatives can form part of a recombinant vector, a host cell, a fusion protein or have utility in methods as described above for the inventive MT and POP enzymes for modifying (poly)peptides, preferably for modifying methylated and/or cyclized (poly)peptide substrates.

In this regard, it is noted that the definitions provided in the context of the methyl transferase (MT) enzymes above and generally relating to the nature of nucleic acids, polypeptides and (poly)peptides, antibodies, hybridoma cells, vectors, host cells, nucleic acid and amino acid sequence identity, hybridization conditions, etc. are also valid for the prolyl oligopeptidase (POP) enzymes, the monooxygenase, O-acyl transferase, monooxygenase, and oxidoreductase enzymes. Moreover, the promoters, vectors, host cells, method parameters, etc. mentioned as preferred embodiments in the context of MT enzymes also form preferred embodiments for practicing the present invention with regard to the other enzymes.

In the following, the invention is illustrated by figures and specific examples, none of which are to be interpreted as limiting the scope of the invention as taught in the appended claims.

Figures

Fig. 1 is a schematic overview of the natural N-methylation pattern of the nematocidal cyclic peptide omphalotin A (left) and its omphalotin variants B to I (right) (adapted from Büchel et al., (1998) Tetrahedron 54:5345-5352).

Fig. 2 shows a schematic bar (top) and corresponding sequence listing (below) illustrating the amino acid sequence of the omphalotin precursor according to SEQ ID NO: 15. The amino acids of the methyl transferase domain (underlined) show homology towards the tetrapyrrole (Corrin

/Porphyrin) methylases (pfam: 00590), the amino acids in italic letters show weak homology towards the conserved domains of unknown function TrbG and YabN (PRK13885 and COG3956) and the amino acids in bold letters form omphalotin. Bold and underlined residues are involved in S-Adenosyl methionine (SAM) binding and compromise methylation when mutated to an alanine. For the protein sequence, see SEQ ID NO 15 and protein ID number: 2087 on the JGI website (<http://genome.jgi.doe.gov/Ompol1/Ompol1.home.html>)

Fig. 3A is a comparative schematic overview of the OmpA gene cluster of *O. olearius* with further fungal homologous omphalotin gene clusters in *D. bispora*. The gene cluster of *O. olearius* consists of (i) the methyltransferase (MT) encoding the omphalotin peptide (2087, SEQ ID NOs: 1 and 15), the prolyl oligopeptidase (POP) (2090, SEQ ID NO: 29 and 36), which is involved in cleavage of the omphalotin peptide from the precursor and the cyclisation reaction. In addition, two monooxygenases (2085, SEQ ID NOs: 31, 38; and 2089, SEQ ID NOs: 33, 40; and 2089Late start, SEQ ID NOs: 34, 41), an O-acyl transferase (2088, SEQ ID NOs: 32, 39), and an oxidoreductase (2091, SEQ ID NOs: 35, 42) are shown that are involved in the modifications of omphalotin variants B to I. The protein sequence of the OmpA gene cluster with the respective protein ID numbers can be found on the JGI website (<http://genome.jgi.doe.gov/Ompol1/Ompol1.home.html>). **Fig. 3B** is a tabular overview of the details of genes located in the *O. olearius* and *D. bispora* borosin gene clusters. The asterisk denotes that the protein sequences corresponding to the respective protein ID numbers can be found on the JGI website (<http://genome.jgi.doe.gov/Ompol1/Ompol1.home.html> and <http://genome.jgi.doe.gov/Denbi1/Denbi1.home.html>).

Fig. 4 is a comparative schematic overview of the OmpA gene cluster of *O. olearius* with further fungal homologous omphalotin gene clusters in *F. mediterranea*. The protein sequences corresponding to the respective protein ID numbers can be found on the JGI website (<http://genome.jgi.doe.gov/programs/fungi/index.jsf>).

Fig. 5A shows an alignment of the methyltransferase domain of the omphalotin precursor with other fungal homologues. Residues in black boxes highlight conserved residues and residues with similar properties are high-lighted in bolt. Sequence identity with the omphalotin precursor (2087) ranges from 89.42% (765759) to 54.36% (32310). Sequences of all homologues are according to SEQ ID NO: 15, 17, 19, 21, 23, 25 and 27 (aa MT only). **Fig. 5B** shows an alignment of the peptidase domain of the omphalotin prolyl oligopeptidase with the prolyl oligopeptidase of *D. bispora*. Sequence identity of the prolyl oligopeptidase from omphalotin with the peptidase from *D.*

bispora is 87.86%. Sequences according to SEQ ID NOs: 37 and 54 (aa protease domains only). Alignment created with clustal Omega (Sievers F et al., (2011) Mol. Sys. Bio. 7: 539) and visualized using ESPript 3 (Robert X et al., (2014) Nucl. Acids Res. 42(W1):W320-W324).

Fig. 6 shows SDS-PAGE gels and immuno detections for confirming the expression and solubility of the omphalotin precursor in *E. coli* BL21. **Gel A** shows a whole cell extract (WCE) of *E. coli* BL21 that was not induced (lane 1), expressing the non-tagged omphalotin precursor (lanes 2, 4 and 6), or the N-terminally HIS-tagged omphalotin precursor (lanes 3, 5 and 7), when separated on a 12% SDS PAGE gel. Lanes 2 and 3 show the WCE before centrifugation, lane 4 and 5 the WCE after low spin (LS) centrifugation (5000xg 5min) and lanes 6 and 7 after high spin (HS) centrifugation (15000xg 30min). **Gel B** shows immuno-detection of the omphalotin precursor with an anti-HIS antibody (Qiagen, Switzerland). A WCE of *E. coli* BL21 expressing an empty vector (EV) was used as a control. Lane 2 shows the WCE after low spin centrifugation and lane 3 after high spin centrifugation. *E. coli* BL21 was induced for 24hr with 0.2mM IPTG at 16°C.

Fig. 7 shows SDS-PAGE gels and immuno detections for confirming the expression and solubility of the omphalotin precursor in *S. cerevisiae*. **Gel A** shows WCEs of *S. cerevisiae* W303 MAT α (lane 1-3) or MAT α /a (lane 3-6) grown for 3 hrs on YP galactose to induce expression of the N-terminally HIS-tagged omphalotin precursor. Lanes 1 and 4 show the WCE before centrifugation, lanes 2 and 5 the WCE after high spin centrifugation (16000g for 30min) and lanes 3 and 6 the purified protein after HIS tag purification. **Gel B** shows the immuno-detection of the omphalotin precursor with an anti-HIS tag antibody on a WCE of *S. cerevisiae* W303 MAT α /a grown for 24hr on galactose at 25°C (lane 2). *S. cerevisiae* W303 MAT α /a grown on glucose was used as a control (lane 1).

Fig. 8A, 8B, 8C and 8D lists methylated fragments of different variants of the omphalotin precursor (SEQ ID NOs: 99, 101 and 102) and the ten times methylated fragment of the *D. bispora* 765750 homologue (SEQ ID NO 100) observed by MS/MS data after different expression times in *E. coli* BL21. The omphalotin precursor was expressed for 1 or 5 days in *E. coli* BL21. The purified protein was digested by trypsin and subjected to LC-MS/MS. The MS-identified trypsin fragments containing the omphalotin core peptide are listed. The top fragment shows where methylations (underlined) are expected in the peptide (bold) of the naturally isolated omphalotins. The fragments below were found experimentally, the third fragment (italic letters) was the most abundant after 1

day of expression. Underlining means that this amino acid was found to be methylated according to the MS/MS data.

Fig. 9 shows the N-methylation pattern of therapeutic peptides produced with the methyltransferase of *O. olearius*. The top line **(1)** illustrates the natural N-methylation pattern of the cyclic non-ribosomal immunosuppressant cyclosporine A, and the natural N-methylation pattern of the cyclin-dependent kinase 4 inhibitor dictyonamide A. Line **(7)** shows the natural N-methylated synthetic antimalarial R1 peptide and kendarimide which reverses P-glycoprotein-mediated multidrug resistance in tumor cells. The following bottom columns (2) to (6) and (8) to (10) show the patterns of N-methylation observed after 5 days of *in vivo E. coli* expression at 16°C when the omphalotin encoding peptide in the precursor was exchanged with the peptide sequences that resembled cyclosporin A (line (1) left, SEQ ID NO: 103), dictyanomide A (line (1) middle, SEQ ID NO: 104), R1 (line (7) left, SEQ ID NO: 105) or kendarimide, (line (7) right, SEQ ID NOs: 106) thus demonstrating general substrate applicability for methylation by the methyl transferase.

Examples

Example 1 - cDNA synthesis and cloning of the omphalotin precursor

O. olearius strain DSM3398 (DSMZ, Germany) was used for RNA extraction and cDNA synthesis. *O. olearius* mycelium was grown on cellophane disks (Celloclair, Switzerland) on 1.5% (w/v) agar plates containing yeast maltose agar (0.4% (w/v) yeast extract (Oxoid AG, England), 1% (w/v) malt extract (Oxoid AG, England), 0.4% (w/v) glucose) for 5 days at 28°C in the dark. To extract RNA, 8.2 mg of lyophilized mycelium was lysed with 8 mg of 0.5 mm glass beads in three FastPrep steps of 45 s at 4.5, 5.5 and 6.5, cooling the sample for 5 min on ice between each step. RNA was extracted with 1 ml Qiazol (Qiagen, Germany) and 0.2 ml chloroform. After a centrifugation at 12000xg and 4 °C for 15 min, RNA was recovered in the aqueous phase, washed on-column using the RNeasy Lipid Tissue Mini Kit (Qiagen, Germany) and eluted in RNase-free water. cDNA was synthesized from 1 µg of extracted RNA using the Transcriptor-first strand cDNA synthesis kit (Roche Applied Science, Germany) following the instructions of the manufacturer. The coding sequence of the omphalotin precursor was amplified from cDNA by PCR using Phusion high-fidelity DNA polymerase using standard protocols (Sambrook J, Russell D, 2001, Molecular cloning 3rd edition) using primers:

Forward primer: 5'-ATGGAGACTTCCACTCAGAC-3' (SEQ ID NO: 64)

Reverse primer: 5'-TTATCCGTGCTCATGACTG-3'. (SEQ ID NO: 65)

An A-tailing reaction was performed on the PCR product after which it was cloned into the pGEM-T-easy vector (Promega, USA) and transformed to *E. coli* DH5 α . A PCR on a plasmid containing the right cDNA sequence was performed using primers:

Forward primer: 5'-GGGGGGCATATGGAGACTTCCACTCAGAC-3' (SEQ ID NO: 66)

Forward primer: 5'-TTTTTTCATATGGAGCATCATCATCATCATCATCATCACT
TCCACTCAGACCAAAGCTGGCTCA-3' (SEQ ID NO: 67)

Reverse primer: 5'-CCCCCGCGGCCGCTTATTCCGTGCTCATGACT-3' (SEQ ID NO: 68)

after which the PCR products were cloned into pET24 (Novagen, Germany) using the restriction sites NdeI and NotI (Thermo Scientific, USA) and transformed to *E. coli* BL21 or cloned into the *S. cerevisiae* expression vector pRS426 (Sikorski R. S. et al., (1989) *Genetics* 122: 19-275) using the restriction sites SpeI and EcoRI.

Example 2 Cloning of the CycD, DicA, KenA and no follower omphalotin variants

An adapter ligation was performed to introduce a BsaI restriction site in the omphalotin precursor gene using primers:

Forward primer: 5'-CATGAGACCTAGTCATGAGCACGGAATAAGC-3' (SEQ ID NO: 69)

Reverse primer: 5'-GGCCGCTTATTCCGTGCTCATGACTAGGTCT-3' (SEQ ID NO: 70)

10 μ l of each primer (100 μ M) was added to 80 μ l of 10mM Tris pH 7.5. Primer annealing was performed with the following program: 5min at 98°C with cooling rate at 4°C/s, 30 seconds at 80°C with cooling rate 0.01°C/s and 30 seconds at 40°C with cooling rate 0.02°C/s. The annealed primers were phosphorylated and ligated into a pET24 vector containing the omphalotin precursor restricted with NcoI and NotI (Thermo Scientific, USA) at a vector-to-insert molar ratio of 1:7. To introduce the sequence encoding for CycD, DicA, KenA and to remove the follower an adapter ligation was performed with the following primers:

CycD forward primer:

5'-AGGTCTACCCACTGGTGCTGGCTGCACTGCTCGTCATCGTTGGTTCAGTCATGAGCACGGAATAAGCGGC
CGCA-3' (SEQ ID NO: 71)

CycD reverse primer: 5'-TGCGGCCGCTTATTCCGTGCTCATGACTGAACCAACGATGACGAGCAGTGCAGCC
AGCACCAGTGGGTGAGACCT-3' (SEQ ID NO: 72)

DicA forward primer:

5`-AGGTCTCACCCAGCAACGACCGTAGTAGTTGTTGTTATTGTGGGTTAAGCGGCCGCA-3` (SEQ ID NO: 73)

DicA reverse primer:

5`-TGCGGCCGCTTAACCCACAATAACAACAATACTACTACGGTCGTTGCTGGGTGAGACCT-3` (SEQ ID NO: 74)

Ken A forward primer:

5`-AGGTCTCACCCACAGTTCATTGCCGTTGTAGTTGTGGCAGTCGTGTGCTGTTTCTAAGCGGCCGCA-3`
(SEQ ID NO: 75)

Ken A reverse primer:

5`-TGCGGCCGCTTAGAAACAGCACACGACTGCCACAATAACAACGGCAATGAACTGTGGGTGAGACCT-3` (SEQ ID NO: 76)

No follower forward primer:

5`-AGGTCTCACCCATGGGTCATCGTCGTTGGTGTTATCGGTGTCATCGGATAAGCGGCCGCA-3` (SEQ ID NO: 77)

No follower reverse primer:

5`-TGCGGCCGCTTATCCGATGACACCGATAACACCAACGACGATGACCCATGGGTGAGACCT-3` (SEQ ID NO: 78)

Primer annealing was performed as described above, restricted using BsaI and NotI, purified using a Nucleospin gel and PCR cleanup kit (Macherey Nagel, Switzerland) and ligated in pET24 containing the omphalotin precursor that was restricted with BsaI.

Example 3 Cloning of OMP TEV, R1 and 765759

For replacement of the omphalotin peptide with that of R1 and to introduce a TEV cleavage site the following sequences were synthesized by Genscript (USA):

OMP TEV

5`-GGATCCCAAGGCACTCGCCGACTACAAAGCTGATCACCGCGCCTTTGCTCAATCTGTCCCCGACTTGACGCCTCAGGAGCGTGCGGCTTTGGAGCTCGGTGATTCGTGGGCTATTCGTTGCGCGATGAAGAATATGCCCTCGTCGCTCTTGACGCTGCTCGTGAATCCGGCGAAAATTGTACTTCCAAGGTTTCCCATGGGTGTCATCGTCGTTGGTGTATCGGTGTCATCGGATAAGAATTC-3` (SEQ ID NO 79)

OMP R1

5`-GGATCCCAAGGCACTCGCCGACTACAAAGCTGATCACCGCGCCTTTGCTCAATCTGTCCCCGACTTGACGC

CTCAGGAGCGTGCGGCTTTGGAGCTCGGTGATTCGTGGGCTATTCGTTGCGCGATGAAGAATATGCCCTCGT
CGCTCTTGGACGCTGCTCGTGAATCCGGCGAAGAGGCATCCCAAAACGGTTTCCCAGTCTTGCAGAATTTCT
GCCTCTGTTCAAGCAATTCGGTTCGCGGATGCACATTCTGAAATAAGCGGCCGC-3' (SEQ ID NO 80)

The synthesized sequences were restricted with BamHI and EcoRI for OMP TEV and BamHI and NotI for R1, purified using a Nucleospin gel and PCR cleanup kit (Macherey Nagel, Switzerland) and ligated into pET24 containing the omphalotin precursor.

The sequence of 765759 from *D. bispora* was codon optimized for expression in *E. coli* and synthesized by Genscript (USA) SEQ ID NO 107. The synthesized sequence was restricted with HindIII and NdeI, purified using a Nucleospin gel and PCR cleanup kit (Macherey Nagel, Switzerland) and ligated into pET24 containing the omphalotin precursor.

Example 4 - Expression and purification of the omphalotin precursor and variants from *E. coli* BL21

The omphalotin precursor, the different variants and the N-terminal HIS-tagged precursor were expressed in *E. coli* BL21 in TB medium (2.4% (w/v) yeast extract (Oxoid AG, England), 1.2% (w/v) tryptone (Oxoid AG, England), 0.4% (w/v) glycerol, 0.17 M KH₂PO₄ and 0.72 M K₂HPO₄) containing 50 µg/ml kanamycin. For protein purification of the HIS-tagged precursor cultures were grown at 37°C to an OD₆₀₀ nm between 1.5-2 after which cultures were cooled down on ice for 30min and 0.2 mM IPTG was added. Further incubation was performed at 16°C for the indicated expression times. Bacterial cell pellets were resuspended in ice-cold buffer (50mM HEPES pH8, 0.1% triton X-100, 10% glycerol) containing 20 mM Imidazole and lysed using a French press. Cell debris was removed by centrifugation at 16000 g for 30 min. For protein purification the supernatant was incubated with Ni-NTA beads (Thermo Scientific, USA) at 4 °C for 1 h and the protein was finally eluted in buffer containing 400 mM Imidazole. Purified protein was concentrated on an Amicon Ultra-4 centrifugal filter device (Millipore, USA) with a molecular weight cut off of 30 kDa, desalted using a PD-10 column (Amersham Biosciences, UK) and flash frozen with liquid nitrogen before storage at -20°C. Protein concentrations were measured using the Pierce BCA protein assay (Thermo Scientific, USA) using a standard curve derived from bovine albumin.

Example 5 - Expression and purification of the omphalotin precursor from *S. cerevisiae*

S. cerevisiae W303 MAT α and W303 MAT α /a containing the omphalotin precursor or the N-terminal HIS-tagged precursor were grown at 30°C in synthetic dropout (SD) media containing

raffinose and without histidine to the exponential phase. Cells were harvested and washed once with YP containing galactose and diluted to $OD_{600}=0.2$. Cells were grown till $OD=1$, washed with ice-cold buffer (50mM HEPES pH8, 0.1% triton X-100, 10% glycerol) and lysed using glass beads. Cell debris was removed by centrifugation at 16000xg for 30 min, Ni-NTA beads (Thermo Scientific, USA) were added and the protein was purified according to the instructions of the manufacturer. Proteins were desalted and stored as described in example 4.

Example 6 - Confirmation of backbone N-methylation of omphalotin peptide and variants by LC-MS/MS

An Amicon ultra column (30 kDa cutoff) was washed 2 times with ddH₂O. 100 µg protein and ddH₂O to a final volume of 400 µL was added. The sample was centrifuged (11000xg for 5 min) and washed with 500 µl of ddH₂O. 50 mM NH₄HCO₃ pH 8 was added and the sample was concentrated to 4 µg/µl. 25 µl was transferred to a glass vial, trypsin (Promega, USA) was added to a molar ratio of 1:70 and the sample incubated for 3 h at 37°C. 3µl of each sample was used for mass spectrometry analysis. Electrospray ionization-LC-MS data was recorded on a Thermo Scientific Q Exactive mass spectrometer equipped with a Dionex Ultimate 3000 UHPLC system using a Phenomenex Kinetex 2.6 µm C18 100 Å (150 × 4.6 mm) column heated at 50°C. Elution was performed with a linear gradient using water with 0.1% (v/v) formic acid (solvent A) and acetonitrile with 0.1% (v/v) formic acid (solvent B) at a flow rate of 0.8mL/min. The column was equilibrated with 5% solvent B for 2 min followed by a linear increase of solvent B to 85% over 15 min and a final elution step with 98% solvent B for 5 min. Mass spectra were acquired in positive ion mode with the following settings; spray voltage 3500V, capillary temperature at 268.75°C, probe heater at 437.5°C and an S-lens level at 50. Full MS was done at a resolution of 35000 (AGC target 1e6, maximum IT 100 ms, range between 500-1800 *m/z*) and data-dependent MSMS was performed at a resolution of 17500 (AGC target at 2e5, maximum IT of 300ms isolation window 1.2) using a stepped NCE of 16, 20 and 24. The inclusion list contained the mass of the trypsin fragment encoding for the omphalotin peptide (or one of the variants) and its different methylation states.

Example 7 - In vitro methylation of the omphalotin precursor and variants

Omphalotin precursor and variants were purified from *E. coli* induced for only 4 hours to obtain non-methylated protein. To 5µM of protein 1µM S-adenosyl methionine (Sigma-Aldrich, USA)

was added and samples were incubated overnight at room temperature. Reactions were performed in 50µl of 50mM HEPES pH8, 0.1% triton X-100 and 10% glycerol. Methylation of the omphalotin precursor or variants were asses by LC-MS as described in example 6.

Example 8 - In vitro methylation of omphalotin peptide with and without follower

The syntheses of peptides were performed by automated Fmoc solid-phase peptide synthesis using the following conditions. Fmoc deprotections were performed using 20% piperidine in DMF for 8 min, repeated another cycle to ensure complete removal. For each coupling, Fmoc-amino acid (4.0 equiv), HCTU (3.9 equiv), and NMM (8.0 equiv) in DMF were coupled to free amine on-resin for 45 min. After coupling, resin was treated with 20% acetic anhydride in DMF and NMM (1 equiv to acetic anhydride) for 10 min to cap any unreacted free amine. Synthesis of omphalotin A: Due to the highly hydrophobic sequence of omphalotin A, the synthesis was performed on ChemMatrix Rink amide resin using hexa-arginine tag with a base labile linker at the C-terminal. The first arginine coupling was performed manually on ChemMatrix Rink amide resin (0.1 mmole scale). Afterwards, five subsequent arginine residues and the basic-labile linker (4-hydroxymethyl-benzoic acid, HMBA) were coupled on synthesizer to form the Arg₆ tag (SEQ ID NO: 108) to enhance the solubility. The C-terminal residue (Glu¹⁸ or Gly¹²) was introduced by pre-formed symmetrical anhydride of the corresponding Fmoc-amino acid (5.0 equiv) in the presence of DMAP (0.1 equiv) for 6 hours. The rest of the sequence was introduced by standard automated Fmoc SPPS protocol resulting in the following peptides:

With follower, **1a**: WVIVGVGVG¹⁰IGSVMSTE¹⁸-HMBA-RRRRRR SEQ ID NO: 83

Without follower, **1b**: WVIVGVGVG¹⁰IG¹²-HMBA-RRRRRR SEQ ID NO: 84

Upon the completion of synthesis, the full peptide on-resin with the final Fmoc was cleaved by treatment with a mixture of 94:2:2:2 TFA/TIPS/DODT/H₂O for **1a** and 95:2.5:2.5 TFA/TIPS/H₂O for **1b** respectively for 2.5 hours. The solid support was filtered off and the filtrate was evaporated under vacuum. The residue was triturated with Et₂O and centrifuged to obtain the crude peptides. The crude peptides were purified by reverse phase HPLC (RP-HPLC) using a gradient of 40% CH₃CN to 90% CH₃CN over 30 min. The fractions containing the product was pooled and lyophilized. The final Fmoc was removed by treatment the peptides with 5% diethylamine in DMSO (2 mM) for 5 min. The solution was neutralized with equal volume of cold TFA (10% in 1:1 CH₃CN/H₂O) and the resulting

mixture was purified by RP-HPLC and lyophilized to obtain N-terminal free amine of **1a** and **1b**. For the final removal of the Arg₆ tag, **1a** and **1b** were dissolved in DMSO/H₂O (9:1) to the concentration of 10 mM and treated with one-tenth volume of 1.0 M NaOH_(aq) for 90 min. The reaction mixture was cooled in an ice bath, neutralized with TFA, purified by RP-HPLC, and lyophilized to afford the final products:

Omphalotin A (with follower): WVIVGVVGV¹⁰IGSVMSTE¹⁸ SEQ ID NO: 85

Omphalotin A (without follower): WVIVGVVGV¹⁰IG¹² SEQ ID NO: 86

5μM of omphalotin precursor, 1μM S-adenosyl methionine (Sigma-Aldrich, USA) and 1mM peptide was added to 200μl of 50mM HEPES pH8, 0.1% triton X-100 and 10% glycerol. Samples were incubated overnight at room temperature after which they were run over an Amicon ultra column (30 kDa cutoff). The supernatant was collected and methylation of the peptide was assessed by LC-MS as described in example 6.

Example 9 In vitro cleavage of omphalotin precursor by TEV protease

TEV protease (Sigma-Aldrich, USA) was added to purified omphalotin precursor containing a TEV cleavage site at a molar ratio of 1:100 and incubated overnight at 4°C. Cleavage of the precursor was assessed by LC-MS/MS as described in example 6.

Example 10 In vitro proteolysis and cyclisation of the omphalotin precursor by the prolyl oligopeptidase

Proteolysis of the peptide from the omphalotin precursor and cyclisation reactions were performed as described by Luo et al., (2014) Chemistry and Biology 21, 1610-1617. In short, 2μg of omphalotin precursor together with 10μg of POP was dissolved in 50mM Tris-HCl (pH7.5) in a reaction volume of 50μl and incubated at 37°C for 4hr. For POP activity on peptides, 1mM of omphalotin peptide with or without follower (see above) were dissolved in 50mM Tris-HCl (pH7.5) containing 1mM dithiothreitol, or 0.1% trifluoroacetic acid and incubated together with 2μg of POP. After the incubation reactions were quenched with 50μl of methanol and the samples were centrifuged for 5 min. The supernatant was analyzed by HPLC (see Luo et al., (2014) Chemistry and Biology 21:1610-1617) or LC-MS as described in example 6.

Example 11 In vivo proteolysis and cyclisation of the omphalotin precursor by the prolyl oligopeptidase

STREP-tagged POP and HIS-tagged OMP were co-expressed in *E. coli* BL21 using the pCDFDUET-1 vector (Merck Millipore, Germany) and brought to expression as described in Example 4. For analysis of proteolysis of the omphalotin precursor protein was purified as described in Example 4 and sent for analysis by whole protein LC/TOF-MS by the Functional Genomics Center (University Zürich). To assess cyclisation and cleavage the supernatant of the whole cell extract after high spin centrifugation was run over an Amicon Ultra-4 centrifugal filter device (Millipore, USA) with a molecular weight cut off of 10 kDa. The supernatant was analyzed by HPLC (see Luo et al., (2014) Chemistry and Biology 21:1610-1617) or LC-MS as described in example 6.

Claims

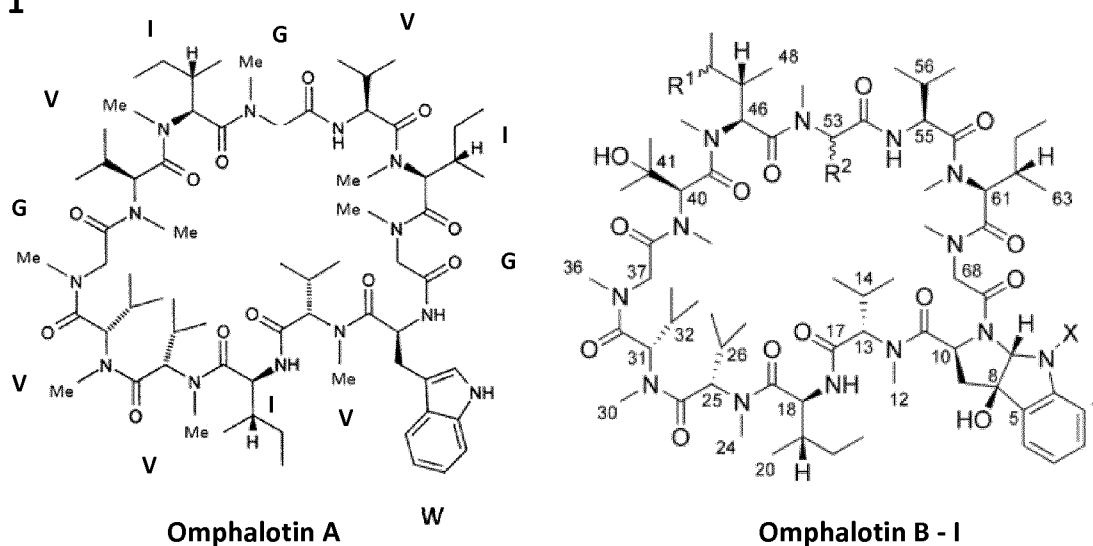
1. An isolated and purified nucleic acid, comprising a nucleic acid sequence selected from the group consisting of:
 - (i) a nucleic acid sequence selected from the group consisting of nucleic acid sequences listed in SEQ ID NOs: 1 to 14 and 108 to 111, preferably SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, and 108, more preferably SEQ ID NOs: 1 and 2;
 - (ii) a nucleic acid sequence of at least 80 or 90 % identity, preferably at least 95 % identity, more preferred at least 98 % identity with a nucleic acid sequence listed in SEQ ID NO 1 to 14 or 108 to 111, preferably over the whole sequence;
 - (iii) a nucleic acid sequence that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions;
 - (iv) a fragment of any of the nucleic acid sequences of (i) to (iii), that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions; and
 - (v) a nucleic acid sequence, wherein said nucleic acid sequence is derivable by substitution, addition and/or deletion of one of the nucleic acids of (i) to (iv) that hybridizes to a nucleic acid sequence of (i) or (ii) under stringent conditions, wherein the nucleic acid sequence encodes a polypeptide having N-methyl transferase activity.
2. The nucleic acid according to claim 1, wherein said nucleic acid is a DNA, RNA or PNA, preferably DNA or PNA, more preferably DNA.
3. The nucleic acid according to claim 1 or 2, wherein said nucleic acid encodes a polypeptide having N-methyl transferase activity for methylating peptidic backbone amides, preferably for multiply methylating peptidic backbone amides, more preferably for multiply methylating peptidic backbone amides of hydrophobic residues.
4. The nucleic acid according to any one of claims 1 to 3, wherein the nucleic acid further comprises a nucleic acid sequence encoding a (poly)peptide for N-methylation by the N-methyl transferase encoded by the nucleic acid according to any one of claims 1 to 4.

5. A recombinant vector comprising a nucleic acid of any of claims 1 to 4, preferably a viral or episomal vector, preferably a baculovirus vector, lentivirus vector, adenovirus vector, yeast or bacterial episomal vector.
6. A host cell comprising a nucleic acid according to any one claims 1 to 4, or a vector according to claim 5, preferably a host cell selected from the group consisting of yeast cells, preferably *Saccharomyces cerevisiae*, *Pichia pastoris* cells, Bacterial *E. coli* or *B. subtilis* cells, plant cells, preferably *Nicotiana tabacum* or *Physcomirella patens* cells, NIH-3T3 mammalian cells and insect cells, preferably sf9 insect cells.
7. An isolated and purified polypeptide selected from the group consisting of:
 - (a) polypeptides having an amino acid sequence selected from the group consisting of SEQ ID NOs: 15 to 28 and 112 to 115, preferably SEQ ID NOs: 16, 18, 20, 22, 24, 26, 28, and 112, more preferably SEQ ID NOs: 15 and 16,
 - (b) polypeptides encoded by a nucleic acid of any of claims 1 to 5,
 - (c) polypeptides having an amino acid sequence identity of at least 70 or 80 %, preferably at least 90 or 95 % with the polypeptides of (a) and/or (b),
 - (d) a functional fragment and/or functional derivative of (a), (b) or (c),wherein the polypeptide has N-methyl transferase activity.
8. The polypeptide according to claim 7, wherein said polypeptide has N-methyl transferase activity for methylating peptidic backbone amides, preferably for multiply methylating peptidic backbone amides, more preferably for multiply methylating peptidic backbone amides of hydrophobic residues.
9. Fusion polypeptide comprising a first polypeptide sequence according to any one of claims 7 to 8, and at least a second polypeptide sequence for N-methylation by the N-methyl transferase activity of the first polypeptide sequence.
10. An antibody, preferably a monoclonal antibody, a functional fragment or functional derivative thereof or antibody-like binding proteins that specifically binds a polypeptide of any of claims 7 to 8.

11. Use of a vector according to claim 5, a host cell according to claim 6 and/or a polypeptide according to any of claims 7 to 8 for producing an N-methylated, preferably a backbone N-methylated (poly)peptide.
12. Method for producing an N-methylated, preferably a backbone N-methylated (poly)peptide, more preferably a multiply backbone N-methylated (poly)peptide comprising the steps of:
 - (a) providing a polypeptide having methyl transferase activity according to any of claims 7 to 8,
 - (b) providing a (poly)peptide for N-methylation,
 - (c) contacting the polypeptide having methyl transferase activity of (a) and the (poly)peptide for N-methylation of (b) under conditions that allow for N-methylation,
 - (d) optionally cyclisation of the methylated peptide,
 - (e) optionally further modifying the N-methylated (poly)peptide,
 - (f) optionally cleaving the methylated peptide from the precursor, and
 - (g) optionally at least partially purifying the N-methylated (poly)peptide.
13. Method according to claim 12, wherein the polypeptide having methyl transferase activity of step (a) and/or the (poly)peptide for N-methylation of step (b) are provided by a vector according to claim 5 and/or a host cell according to claim 6.
14. Method according to claim 12 or 13, wherein the polypeptide having methyl transferase activity of step (a) and the (poly)peptide for N-methylation of step (b) are provided as a single fusion protein according to claim 9 or separately, preferably as a single fusion protein according to claim 9.
15. Method according to any of claims 12 to 14, wherein in steps (d) and/or (e) at least one further enzyme and/or protein is added for further modifying the N-methylated (poly)peptide, preferably further enzymes and proteins selected from the group consisting of prolyl oligopeptidase (POP), monooxygenase, preferably P450 monooxygenase, O-acyl transferase and oxidoreductase, more preferably one or more enzymes having an amino acid sequence according to any one of SEQ ID Nos: 36 to 42, and 53 to 58.

1 / 9

Fig. 1



- 1 Omphalotin B: $R^1 = A, R^2 = OH$
 2 Omphalotin C: $R^1 = A, R^2 = OAc$
 3 Omphalotin D: $R^1, R^2 = OAc$
 4 Omphalotin E: $R^1, R^2, X = H$
 5 Omphalotin F: $R^1, R^2 = H, X = OH$

- 6 Omphalotin G: $R^1 = H, R^2, X = OH$
 7 Omphalotin H: $R^1, R^2 = OAc, X = OH$
 8 Omphalotin I: $R^1 = B, R^2 = OAc, X = OH$

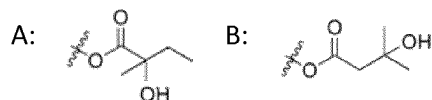
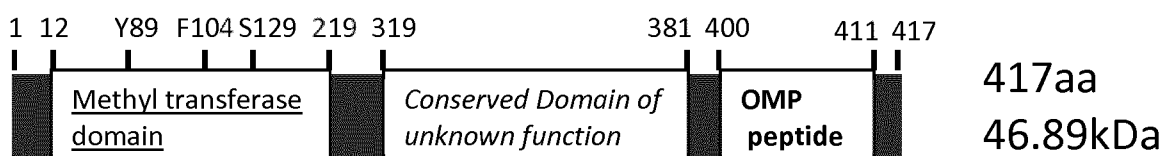
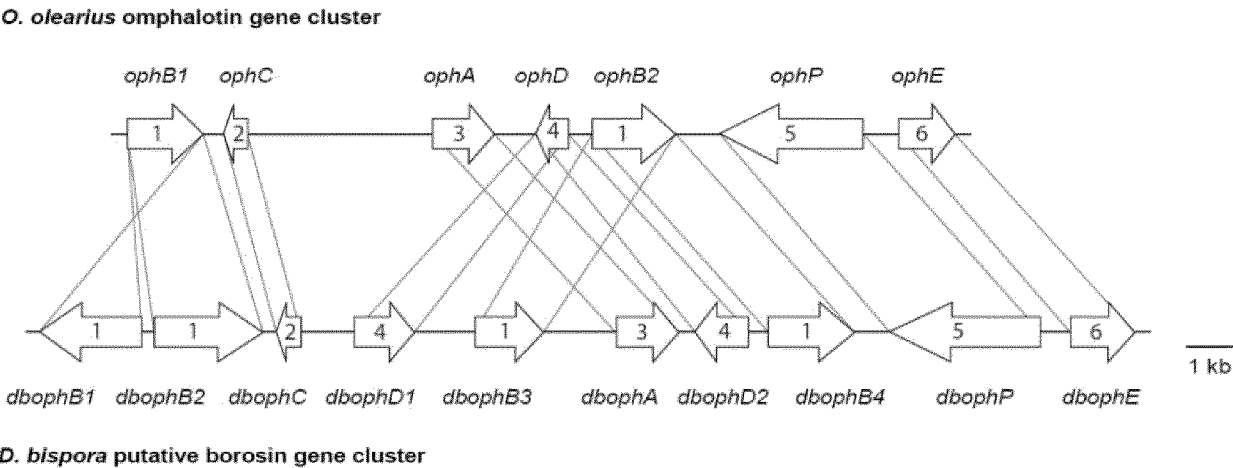


Fig. 2



METSTQTKAGSLTIVGTGIESIGQMTLQALSIEAAKVFYCVIDPATEAFILTKNKNCVDLYQYYD
 NGKSRLNTYTQMSELMVREVRKGLDVVGVFYGHPGVFVNPSHRALAIKSEGYRARMPLPGVS
 AEDCLFADLCIDPSNPGCLTYEASDFLIRDRPVSIHSHLVLFQVGCVGIADFNFTGFDNNKFGVLV
 DRLEQEYGAHPVVHYIAAMMPHQDPVTDKYTVAQLREPEIAKRGGVSTFYIPPKARKASNLD
 IIRRELLPAGQVPDKKARIYPANQWEPDVPEVEPYRPSDQAAIAQLADHAPPEQYQPLATSKA
 MSDVMTKLALDPKALADYKADHRAFAQSVPDLTPQERAALGDSWAIRCAMKNMPSSLLDA
 ARESGEEASQNGFP**WVIVGVIGVIGSVMSTE**
 (SEQ ID NO: 15)

Fig. 3A



Dendrothele bispora

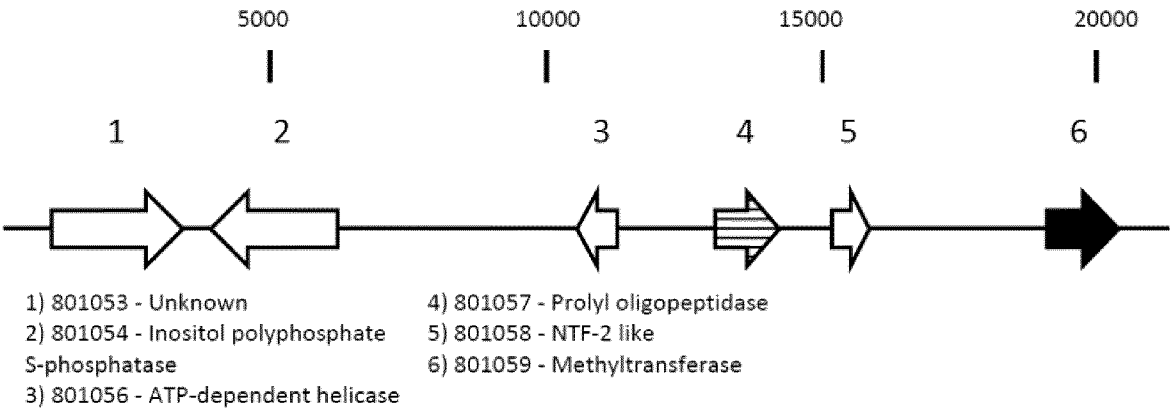
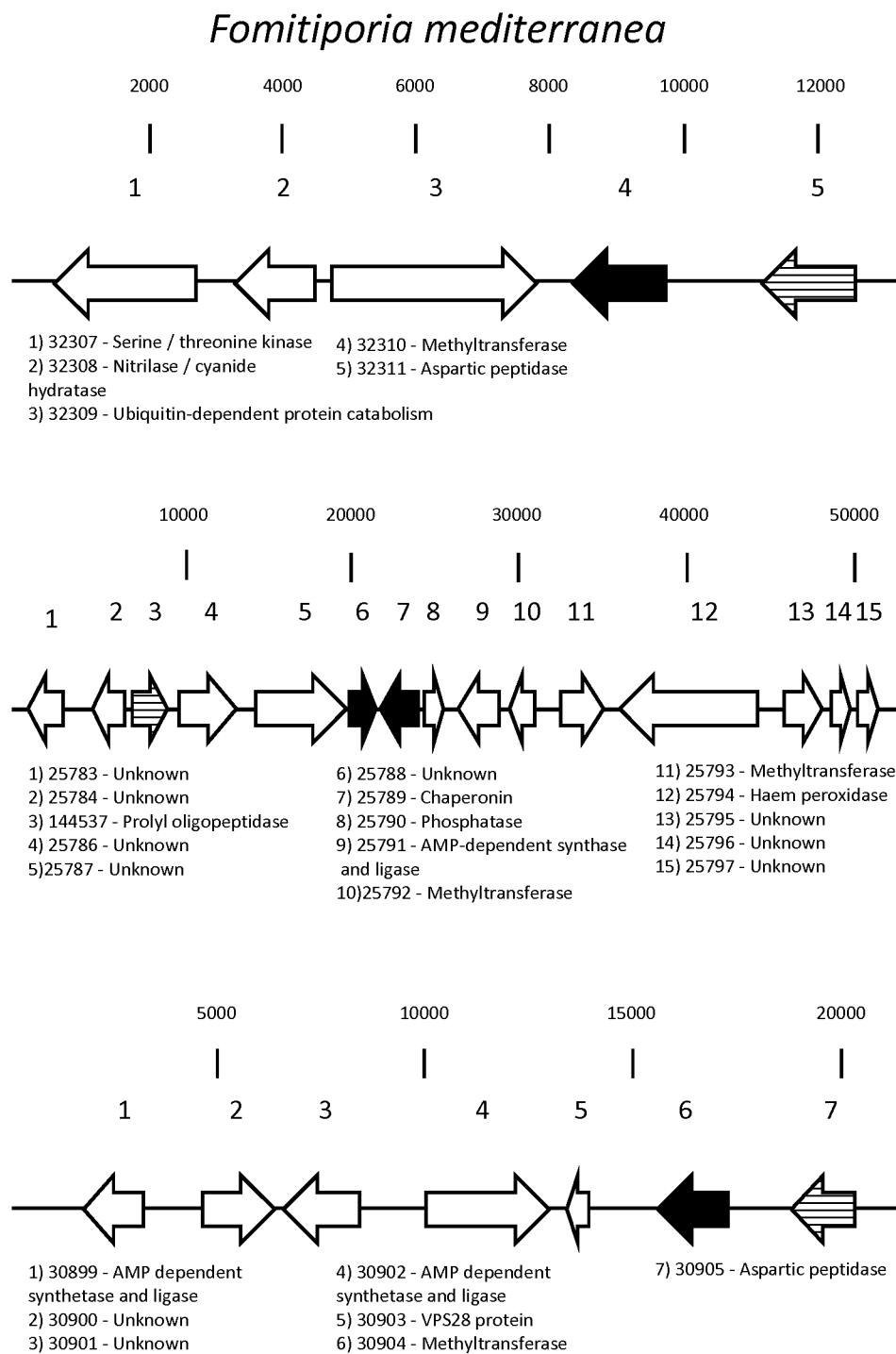


Fig. 3B

Gene	Protein ID*	Scaffold number	Gene start	Gene stop	Gene length	# of introns	Transcript length	Protein length
<i>O. olearius</i> gene cluster								
<i>ophB1</i>	2085	169	2522	4276	1161	10	1755	386
<i>ophC</i>	2086	169	4718	5286	375	3	569	124
<i>ophA</i>	2087	169	9536	10904	1369	2	1254	417
<i>ophD</i>	2088	169	11859	12616	758	1	696	231
<i>ophB2</i>	2089	169	13183	15046	1497	7	1864	498
<i>ophP</i>	2090	169	16016	19346	2172	21	3331	723
<i>ophE</i>	2091	169	20170	21484	1251	1	1315	416
<i>D. bispora ophA</i> gene cluster								
<i>dbophB1</i>	971518	621	40768	43005	2128	11	1515	504
<i>dbophB2</i>	971517	621	38089	40469	2313	11	1629	542
<i>dbophC</i>	871170	621	37262	37791	570	3	363	120
<i>dbophD1</i>	784138	621	34955	36331	1208	1	1092	363
<i>dbophB3</i>	831474	621	32392	34679	1864	7	1497	498
<i>dbophA</i>	765759	621	28937	30516	1384	2	1254	417
<i>dbophD2</i>	765756	621	27371	28885	1232	1	1113	370
<i>dbophB4</i>	765754	621	25288	27217	1854	6	1485	542
<i>dbophP</i>	765750	621	17398	24733	3500	20	1830	609
<i>dbophE</i>	871164	621	19236	20583	1278	1	1278	313

4 / 9

Fig. 4



5 / 9

Fig. 5A (SEQ ID NOs: 15, 17, 27, 23, 25, 19, 21)

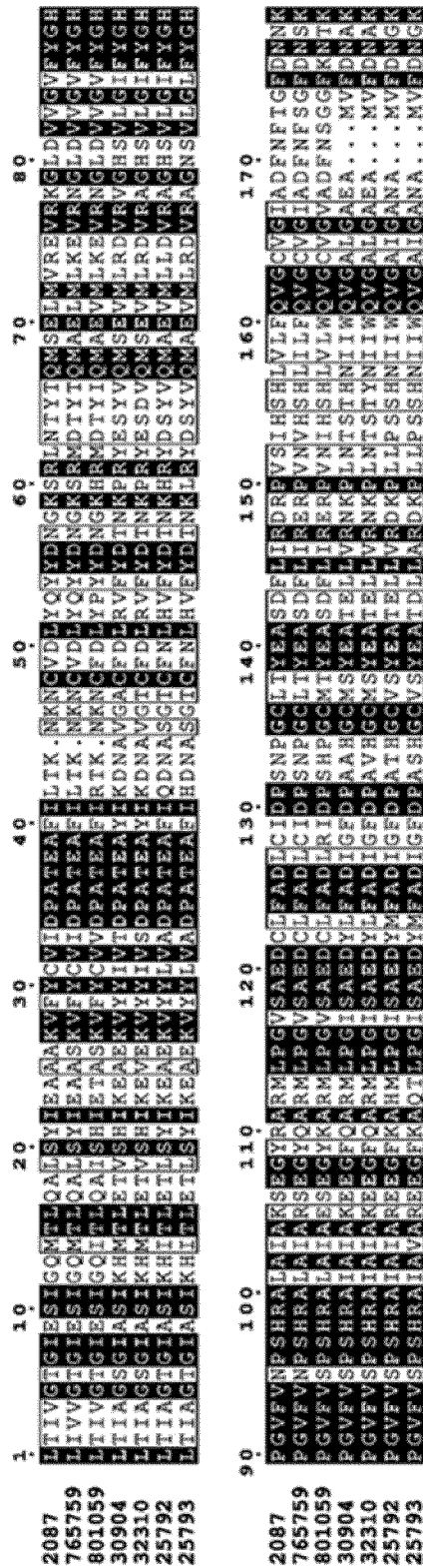


Fig. 5B (SEQ ID NOs: 37, 54)

180 190 200 % identity

2087 FGVLDRLQEYGAEPVHVHYIAAMMPHQ... 100

765759 FGVLDRLQEYGAEPVHVHYIAAMMPHQ... 89.42

801059 FGVLDRLQEYGAEPVHVHYIAAMMPHQ... 82.83

30904 FGVLDRLQEYGAEPVHVHYIAAMMPHQ... 56.41

32310 FGVLDRLQEYGAEPVHVHYIAAMMPHQ... 54.36

25792 FGVLDRLQEYGAEPVHVHYIAAMMPHQ... 55.38

25793 FGVLDRLQEYGAEPVHVHYIAAMMPHQ... 55.38

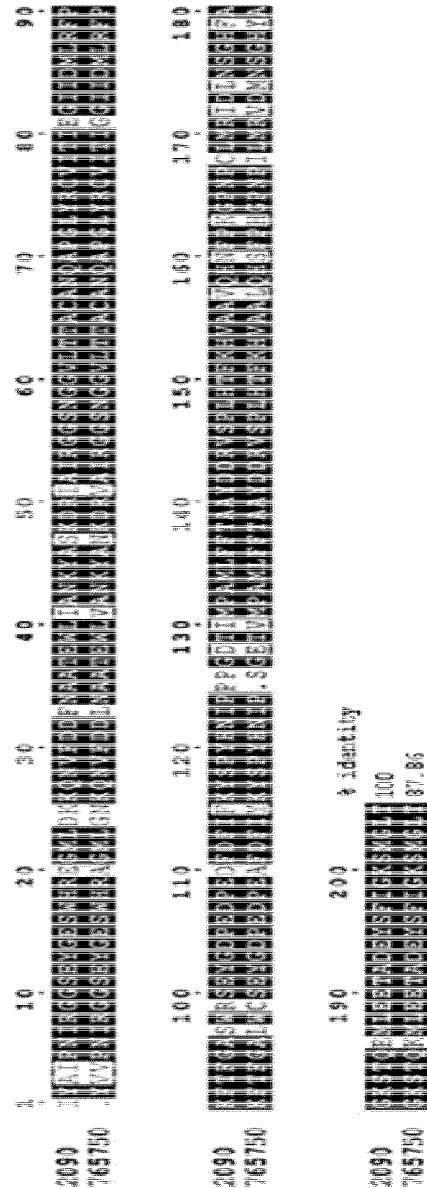


Fig. 6

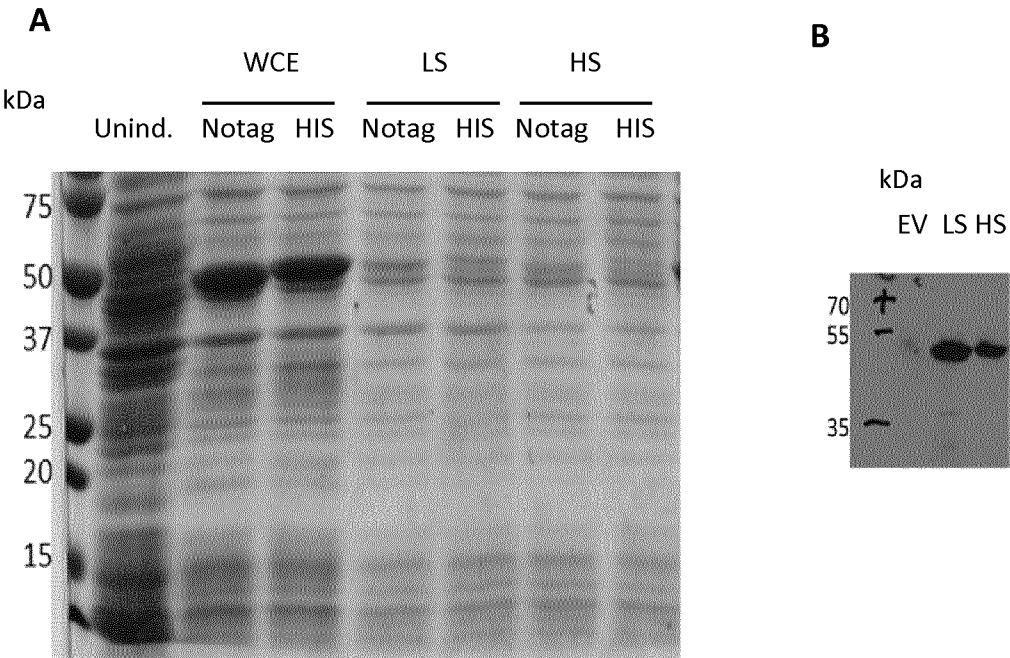
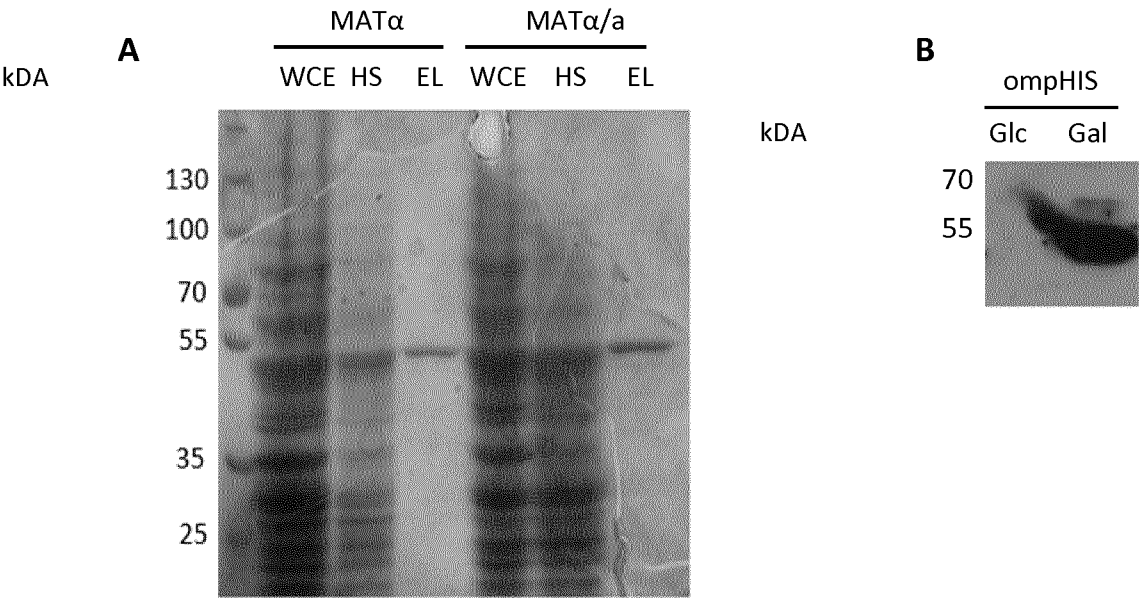


Fig. 7



7 / 9

Fig. 8A

Omp SEQ ID NO: 99

ESGEEASQNGFPWWIVVGVIGVIGSVMST

5 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: *ESGEEASQNGFPW*IVVGVIGVIGSVMST*E*

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

Minor components

5 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

1 Day + Trypsin: ESGEEASQNGFPWIVVGVIGVIGSVMST

ESGEEASQNGFPWIVVGVIGVIGSVMST

N = N-Me detectedN = expected N-Me*N* = major species
observed

Fig. 8B

D. bispora 765750 (SEQ ID NO: 100)1 Day + Trypsin: NMPSSLLEAASQSVEEASMNGFPWIVITGIVGVIGSVVSSA

8 / 9

Fig. 8C

Omp no follower (SEQ ID NO: 101)

ESGEEASQNGFPWVIVVGVIGVIG

5 Day + Trypsin: ESGEEASQNGFPWVIVVGVIGVIG

5 Day + Trypsin: *ESGEEASQNGFPW*VIVVGVIGVIG

Fig. 8D

Omp TEV cleavage site (SEQ ID NO:102)

ESGENLYFQGFPWVIVVGVIGVIG

5 Day + Trypsin: *ESGENLYFQGFPW*VIVVGVIGVIG

5 Day + Trypsin: ESGENLYFQGFPWVIVVGVIGVIG

1 Day + Trypsin: ESGENLYFQGFPWVIVVGVIGVIG

1 Day + Trypsin: ESGENLYFQGFPWVIVVGVIGVIG

1 Day + Trypsin: ESGENLYFQGFPWVIVVGVIGVIG

1 Day + Trypsin: ESGENLYFQGFPWVIVVGVIGVIG

1 Day + Trypsin: ESGENLYFQGFPWVIVVGVIGVIG

1 Day + Trypsin: ESGENLYFQGFPWVIVVGVIGVIG

1 Day + Trypsin: ESGENLYFQGFPWVIVVGVIGVIG

Minor component observed

(ESGENLYFQGFPWVIVVGVIGVIG)

N = N-Me detected

N = expected N-Me

N = major species observed

9 / 9

Fig. 9

cyclosporin A		dictyonamide A	
(1)	<div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div></div></div>		

N = N-Me detected
N = N-Me detected
 position not defined

* = D-
 R1= butenyl-methyl
 R2= L-alpha-aminobutyric acid
 Abz= anthranilic acid

R3= L-pyroMeGlu
 R4=L-Phenylalaninol
 cc= disulphide bridge
 between two cysteines

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/058327

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/10
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 C07K

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, Sequence Search, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>LYNDEL W MEINHARDT ET AL: "Genome and secretome analysis of the hemibiotrophic fungal pathogen, <i>Moniliophthora roreri</i>, which causes frosty pod rot disease of cacao: mechanisms of the biotrophic and necrotrophic phases", BMC GENOMICS, BIOMED CENTRAL LTD, LONDON, UK, vol. 15, no. 1, 27 February 2014 (2014-02-27), page 164, XP021179819, ISSN: 1471-2164, DOI: 10.1186/1471-2164-15-164 the whole document</p> <p>-/--</p>	1-15



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

17 July 2017

Date of mailing of the international search report

06/09/2017

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

Surdej, Patrick

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	<p>-& DATABASE UniProt [Online]</p> <p>22 January 2014 (2014-01-22), "SubName: Full=Uroporphyrin-iii c tetrapyrrole (Corrin porphyrin) methyltransferase {ECO:0000313 EMBL:ESK83659.1}"; XP002760295, retrieved from EBI accession no. UNIPROT:V2XW98 Database accession no. V2XW98 the whole document</p> <p>-----</p>	
X	<p>D. FLOUDAS ET AL: "The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes", SCIENCE, vol. 336, no. 6089, 29 June 2012 (2012-06-29), pages 1715-1719, XP055291730, US ISSN: 0036-8075, DOI: 10.1126/science.1222218 the whole document -& D Floudas ET AL: "tetrapyrrole methylase [Fomitiporia mediterranea MF3/22] - GenBank: EJD06538.1", NCBI - GenBank, 18 March 2015 (2015-03-18), pages 1-2, XP055291776, Retrieved from the Internet: URL:http://www.ncbi.nlm.nih.gov/protein/393221053 [retrieved on 2016-07-28] the whole document</p> <p>-----</p>	1-15
A	<p>GRUENEWALD J ET AL: "Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides", MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 70, no. 1, 1 March 2006 (2006-03-01), pages 121-146, XP002431363, ISSN: 1092-2172, DOI: 10.1128/MMBR.70.1.121-146.2006 page 138, right-hand column, paragraph 2</p> <p>-----</p>	1-15
X	<p>CA 2 418 798 A1 (ACTINODRUG PHARMACEUTICALS GMB [DE]) 5 November 2002 (2002-11-05) page 2 - page 3</p> <p>-----</p>	1-15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2017/058327

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-15(partially)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. claims: 1-15(partially)

An isolated and purified nucleic acid, comprising a nucleic acid sequence encoding a protein having an amino acid sequence selected from the group consisting of SEQ ID NOs: 15 to 18, 27-28, 112.

2. claims: 1-15(partially)

An isolated and purified nucleic acid, comprising a nucleic acid sequence encoding a protein having an amino acid sequence selected from the group consisting of SEQ ID NOs: 19-20.

3. claims: 1-15(partially)

An isolated and purified nucleic acid, comprising a nucleic acid sequence encoding a protein having an amino acid sequence selected from the group consisting of SEQ ID NOs: 21-22

4. claims: 1-15(partially)

An isolated and purified nucleic acid, comprising a nucleic acid sequence encoding a protein having an amino acid sequence selected from the group consisting of SEQ ID NOs: 23-26.

5. claims: 1-15(partially)

An isolated and purified nucleic acid, comprising a nucleic acid sequence encoding a protein having an amino acid sequence selected from the group consisting of SEQ ID NOs: 113.

6. claims: 1-15(partially)

An isolated and purified nucleic acid, comprising a nucleic acid sequence encoding a protein having an amino acid sequence selected from the group consisting of SEQ ID NOs: 114.

7. claims: 1-15(partially)

An isolated and purified nucleic acid, comprising a nucleic acid sequence encoding a protein having an amino acid

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

sequence selected from the group consisting of SEQ ID NOs:
115.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2017/058327

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
CA 2418798	A1	05-11-2002	AU 5965100 A 02-01-2001
		CA 2377592 A1 21-12-2000	
		CA 2418798 A1 05-11-2002	
		DE 10081623 D2 29-05-2002	
		DE 19928313 A1 21-12-2000	
		EP 1190066 A2 27-03-2002	
		WO 0077220 A2 21-12-2000	
