



(51) International Patent Classification:

C07K 14/195 (2006.01)

(21) International Application Number:

PCT/JP2018/038013

(22) International Filing Date:

28 September 2018 (28.09.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2017136428 16 October 2017 (16.10.2017) RU

(71) Applicant: **AJINOMOTO CO., INC.** [JP/JP]; 15-1, Kyobashi 1-chome, Chuo-ku, Tokyo, 1048315 (JP).

(72) Inventors: **SMIRNOV, Sergey, Vasilievich**; Bldg. 1, 1, 1st Dorozhny pr., Moscow 117545 (RU). **KOTLIAROVA, Veronika, Aleksandrovna**; Bldg. 1, 1, 1st Dorozhny pr., Moscow 117545 (JP). **FUJII, Hidemi**; c/o Ajinomoto Co., Inc. 1-1, Suzuki-cho, Kawasaki-ku, Kawasaki-shi, Kanagawa 2108681 (JP). **YAMAZAKI, Shunsuke**; c/o Ajinomoto Co., Inc., 1-1, Suzuki-cho, Kawasaki-ku, Kawasaki-shi, Kanagawa, 2108681 (JP).

(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, JO, 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).

(54) Title: A METHOD FOR PRODUCTION OF A PROTEIN HAVING PEPTIDYLGLYCINE ALPHA-HYDROXYLATING MONOOXYGENASE ACTIVITY

(57) Abstract: The present invention provides a method for production of a protein having peptidylglycme alpha-hydroxylating monooxygenase (PHM) activity using a coryneform bacterium having an ability to produce a protein by secretory production, which method includes steps of culturing the bacterium, which has been modified so that the nucleic acids encoding the domains CuA and CuB of the PHM are co-expressed on separate genetic constructs in the bacterium, in a culture medium to produce the protein, and collecting the produced protein from the bacterium and / or the medium. The novel technique allows for improved production of the protein having an activity of a heterologous PHM native to various species using a coryneform bacterium such as, for example, a bacterium belonging to the genera *Corynebacterium* and *Brevibacterium*.



DESCRIPTION
**A METHOD FOR PRODUCTION OF A PROTEIN HAVING
PEPTIDYLGLYCINE ALPHA-HYDROXYLATING MONOOXYGENASE
ACTIVITY**

Background of the Invention

Field of the Invention

The present invention relates to the biotechnology industry, and specifically to a method for production of a protein having peptidylglycine alpha-hydroxylating monooxygenase activity using a coryneform bacterium that is able to produce and secrete a protein. The protein produced by the method of the present invention can be used for production of amidated peptides.

Description of the Related Art

Amidation of a glycine residue at the carboxy-terminus (hereinafter, C-terminus) is required to attain biological activity of many neuronal and endocrine peptides. It is estimated that about 50% of all mammalian peptide hormones have an alpha-amidated C-terminus (Eipper B.A. *et al.*, Peptides in the nervous system, *Trends Neurosci.*, 1986, 9:463-468). The alpha-amidated peptides are produced by oxidative cleavage of precursor peptides at the C-terminal glycine residue to generate the alpha-amidated target peptides and glyoxylate (Bradbury A.F. *et al.*, Mechanism of C-terminal amide formation by the pituitary enzymes, *Nature*, 1982, 298:686-688). The oxidative cleavage (herein, also called an amidation reaction) is catalyzed by the bifunctional enzyme peptidylglycine alpha-amidating monooxygenase (PAM) in a two-step process (Kolhekar A.S. *et al.*, Peptidylglycine α -hydroxylating monooxygenase: active site residues, disulfide linkages, and a two-domain model of the catalytic core, *Biochemistry*, 1997, 36(36):10901-10909; and references therein). The first step, hydroxylation at the alpha-carbon of the C-terminal glycine residue of a precursor peptide, is catalyzed by peptidylglycine alpha-hydroxylating monooxygenase (PHM, EC 1.14.17.3). The second step, dealkylation of the intermediate peptidyl-alpha-

hydroxyglycine, is catalyzed by peptidyl-alpha-hydroxyglycine-alpha-amidating lyase (PAL, EC 4.3.2.5). As a result, the peptide amidated at the C-terminus and glyoxylate are obtained.

As alpha-amidated peptides are of therapeutic importance, methods for inexpensive and large-scale commercial production of such peptides are in demand. One of the considerable obstacles in the production of recombinant alpha-amidated peptides is the availability of the amidating enzymes. Difficulties both in manufacturing soluble and active amidating enzymes such as PAM, including catalytic subunits thereof such as PHM and PAL, and in subsequent commercial production of the objective alpha-amidated peptides, are well-recognized. Nonetheless, methods for production of PAM, PHM, and PAL, that can be used in a method for C-terminal amidation of a precursor peptide, have been reported. For example, a method for production of rat PAM using a *Streptomyces* bacterium that has been modified to co-express the rat PAM together with a precursor peptide, such as calcitonin or glucagon, extended by a glycine residue at the C-terminus, has been reported (see, for example, Hong B. *et al.*, Production of C-terminal amidated recombinant salmon calcitonin in *Streptomyces lividans*, *Appl. Biochem. Biotechnol.*, 2003, 110:113-123; Qi X. *et al.*, Expression, purification, and characterization of C-terminal amidated glucagon in *Streptomyces lividans*, *J. Microbiol. Biotechnol.*, 2008, 18:1076-1080). In another example, high levels of the catalytic core of PHM (PHMcc) from rat were produced using CHO (DG44) (*dhfr*⁻) cells (Kolhekar A.S. *et al.*, 1997). This method for the production of rat PHMcc was improved further by using bioreactors such as Cellmax 100 (B1) and Accusyst-MiniMax (B2) (Bauman A.T. *et al.*, Large scale production of the copper enzyme peptidylglycine monooxygenase using an automated bioreactor, *Protein Expr Purif.*, 2007, 51(1):34-38). In another example, the catalytic core of human PHM (hPHMcc) was produced in *Escherichia coli* as a fusion protein with glutathione-S-transferase (GST), thioredoxin (Trx), N utilization substance A (NusA), maltose binding protein (MBP), and a His6-tag (Handa S. *et al.*, Production of the catalytic core of human peptidylglycine α -hydroxylating monooxygenase (hPHMcc) in *Escherichia coli*, *Protein Expr Purif.*, 2012, 84(1):9-13). Among the fusion

proteins tested, only Trx-hPHMcc-His6 was expressed successfully to produce active and soluble hPHMcc that displayed steady-state kinetic parameters similar to that of the wild-type rat PHMcc. However, the amount of the hPHMcc thus produced was rather low.

It has been reported that the catalytic core of rat PHM (rPHMcc) can be cleaved via limited proteolysis using endoprotease Lys-C into two major fragments, referred to as the N-terminal domain and the C-terminal domain, with molecular weights of 18.5 and 16.5 kDa, respectively (Kolhekar A.S. *et al.*, 1997). The cleavage site is located to the C-terminal side of the Lys219-residue, which is present in an exposed surface region that links the two domains. Each domain of the PHM contains an independent copper binding site, referred to as CuA- and CuB-binding site, and, henceforth, they can be referred to as domain CuA and domain CuB, respectively. The proposed model for the PHM suggests that the two domains are separable and held together by a flexible hinge.

However, a method for production of a protein having peptidylglycine alpha-hydroxylating monooxygenase activity using a coryneform bacterium that is able to produce and secrete a protein, which method includes the steps of culturing a coryneform bacterium that has been modified so that the nucleic acids encoding the domain CuA and domain CuB of the protein are co-expressed separately in the bacterium, was not known.

Brief Summary of the Invention

It is one aspect of the present invention to provide a novel technique for improving production of a protein having peptidylglycine alpha-hydroxylating monooxygenase activity, also called the objective protein, using a coryneform bacterium belonging to the genus *Corynebacterium* or *Brevibacterium*. According to the presently disclosed subject matter, production of the active form of a PHM such as, for example, mammalian PHM, including human PHM and rat PHM, using a coryneform bacterium which is able to produce a protein by secretory production can be improved. As the production of a protein having peptidylglycine alpha-hydroxylating monooxygenase activity (also referred to as "an activity of PHM") can be improved by the methods as

described herein, peptides that are alpha-amidated at the C-terminus can be produced at a lower price. Therefore, it is a further aspect of the present invention to provide a method for C-terminal amidation of a precursor peptide to produce an alpha-amidated target peptide using the protein having the activity of PHM and produced by the developed novel technique as described herein.

Production of human PHM and rat PHM using a coryneform bacterium belonging to the genus *Corynebacterium* that is able to produce a protein by secretory production can be improved when the bacterium is modified in such a way that a nucleic acid encoding the domain CuA of the PHM and a nucleic acid encoding the domain CuB of the PHM are co-expressed on separate genetic constructs in the bacterium. That is, using a coryneform bacterium belonging to the genus *Corynebacterium* that is able to produce a protein by secretory production to produce a protein having an activity of PHM native to various species can be improved when the domains CuA and CuB of the protein are encoded by the nucleic acids harbored in the bacterium on separate genetic constructs such that co-expression of the nucleic acids is attained.

It is one aspect of the present invention to provide a method for production of a protein having peptidylglycine alpha-hydroxylating monooxygenase activity using a coryneform bacterium that is able to produce and secrete a protein by secretory production, wherein the method comprises:

i) culturing the coryneform bacterium in a culture medium to produce the protein having peptidylglycine alpha-hydroxylating monooxygenase activity, and

ii) collecting the protein having peptidylglycine alpha-hydroxylating monooxygenase activity from the bacterium and/or the medium;

wherein said bacterium has been modified so that a nucleic acid encoding domain CuA of a peptidylglycine alpha-hydroxylating monooxygenase and a nucleic acid encoding domain CuB of a peptidylglycine

alpha-hydroxylating monooxygenase are co-expressed separately in the bacterium.

It is a further aspect of the present invention to provide the method as described above, wherein the protein having peptidylglycine alpha-hydroxylating monooxygenase activity is native to a mammal.

It is a further aspect of the present invention to provide the method as described above, wherein said mammal is selected from the group consisting of human, horse, pig, and rat.

It is a further aspect of the present invention to provide the method as described above, wherein the nucleic acid encoding the domain CuA is a DNA selected from the group consisting of:

(A) a DNA having the nucleotide sequence shown in SEQ ID NO: 11, 15, 17, 25, 27, 29, 31, 33, 35, 37, 57, or 108,

(B) a DNA which encodes the domain CuA using any synonymous amino acid codons according to the standard genetic code table,

(C) a DNA encoding the protein having the amino acid sequence shown in SEQ ID NO: 12, 16, 18, 26, 28, 30, 32, 34, 36, 38, 58, or 109,

(D) a DNA encoding the protein having the amino acid sequence shown in SEQ ID NO: 12, 16, 18, 26, 28, 30, 32, 34, 36, 38, 58, or 109, but which includes substitution, deletion, insertion, or addition of 1 to 20 amino acid residues, and wherein said protein has peptidylglycine alpha-hydroxylating monooxygenase activity with the domain CuB.

It is a further aspect of the present invention to provide the method as described above, wherein the nucleic acid encoding the domain CuB is a DNA selected from the group consisting of:

(a) a DNA having the nucleotide sequence shown in SEQ ID NO: 13, 39, 41, 45, 49, 55, or 59,

(b) a DNA which encodes the domain CuB using any synonymous amino acid codons according to the standard genetic code table,

(c) a DNA encoding the protein having the amino acid sequence shown in SEQ ID NO: 14, 40, 42, 46, 50, 56, or 60,

(d) a DNA encoding the protein having the amino acid sequence shown in SEQ ID NO: 14, 40, 42, 46, 50, 56, or 60, but which includes substitution, deletion, insertion, or addition of 1 to 20 amino acid residues, and wherein said protein has the peptidylglycine alpha-hydroxylating monooxygenase activity with the domain CuA.

It is a further aspect of the present invention to provide the method as described above, wherein the coryneform bacterium belongs to the genus *Corynebacterium* or *Brevibacterium*.

It is a further aspect of the present invention to provide the method as described above, wherein the coryneform bacterium is *Corynebacterium glutamicum*.

It is a further aspect of the present invention to provide the method as described above, wherein the coryneform bacterium harbors a genetic construct for production of the protein having peptidylglycine alpha-hydroxylating monooxygenase activity, wherein the genetic construct comprises:

- (i) a promoter sequence that functions in the coryneform bacterium,
- (ii) a nucleic acid encoding a signal peptide that functions in the coryneform bacterium, wherein said nucleic acid encoding a signal sequence is ligated downstream from the promoter sequence, and
- (iii) the nucleic acid encoding the domain CuA, wherein said nucleic acid encoding the domain CuA is ligated downstream from the nucleic acid encoding the signal peptide;

wherein the genetic construct is co-expressed with a second genetic construct comprising the nucleic acid encoding the domain CuB.

It is a further aspect of the present invention to provide the method as described above, wherein the coryneform bacterium harbors the genetic construct for production of the protein having peptidylglycine alpha-hydroxylating monooxygenase activity, wherein the genetic construct comprises:

(i) a promoter sequence that functions in the coryneform bacterium,

(ii) a nucleic acid encoding a signal peptide that functions in the coryneform bacterium, wherein said nucleic acid encoding a signal peptide is ligated downstream from the promoter sequence, and

(iii) the nucleic acid encoding the domain CuB, wherein said nucleic acid encoding the domain CuB is ligated downstream from the nucleic acid encoding the signal peptide;

wherein the genetic construct is co-expressed with a second genetic construct comprising the nucleic acid encoding the domain CuA.

It is a further aspect of the present invention to provide the method as described above, wherein the signal peptide is a signal peptide for the Sec-mediated protein secretion system.

It is another aspect of the present invention to provide a method for C-terminal amidation of a precursor peptide to produce a target peptide, wherein the method comprises a step of hydroxylating at the alpha-carbon of the C-terminal glycine residue of the precursor peptide in the presence of the protein having peptidylglycine alpha-hydroxylating monooxygenase activity produced by the method as described herein.

It is a further aspect of the present invention to provide the method as described above, wherein the target peptide is a neuropeptide, cytokine or hormone.

It is a further aspect of the present invention to provide the method as described above, wherein the target peptide is exenatide.

The present invention is described in detail below.

Brief Description of Drawings

FIG.1 shows an alignment of amino acid sequences of PHMcc native to *Homo sapiens* (human) - hPHMcc(S39-V351), *Rattus norvegicus* (rat) - rPHMcc(S44-V356), *Equus caballus* (horse) - eqPHMcc(S38-V350), and *Sus scrofa* (pig) - ssPHMcc(S39-V351). Identical amino acid residues are shown in grey boxes. The cleavage site of endoprotease Lys-C is shown by the arrow.

FIG.2 shows the scheme for C-terminal amidation of a precursor peptide (adopted from Handa S. *et al.*, 2012).

FIGs.3(A) and (B) show the generalized structure of the constructed DNA-fragments (A) and, specifically, the structure of the DNA-fragment <cspA-Pe03G> encoding Exenatide precursor (B). Pr_{scpB} – promoter of *cspB* gene, SD – Shine-Dalgarno sequence, SP_{cspA} – signal peptide of *cspA* gene, cleavage site for signal peptidase is marked by the arrow, R1 and R2 – restriction sites of endonucleases, glycine residue (G, Gly) at the C-terminus of Exenatide precursor is underlined.

FIG.4 shows the structure of the *C. glutamicum*/*E. coli* shuttle vector pPK4.

FIG.5 shows the scheme for construction of pPK4-F1 and pPK4-FA-FB plasmids. F1, FA, and FB are sets of DNA fragments as shown in Table 2.

FIG.6 shows the data on the activity of PHM using a standard preparation of PAM: A – raw data analysis, B – normalized data analysis.

Description of Sequences

SEQ ID NO: 1	peptidylglycine alpha-amidating monooxygenase, isoform 1, human (hPAM)
SEQ ID NO: 2	peptidylglycine alpha-amidating monooxygenase, isoform 1, human (hPAM)
SEQ ID NO: 3	peptidylglycine alpha-amidating monooxygenase, isoform X3, rat (rPAM)

SEQ ID NO: 4	peptidylglycine alpha-amidating monooxygenase, isoform X3, rat (rPAM)
SEQ ID NO: 5	catalytic core of peptidylglycine alpha-hydroxylating monooxygenase, human (hPHMcc)
SEQ ID NO: 6	catalytic core of peptidylglycine alpha-hydroxylating monooxygenase, human (hPHMcc(S39-V351))
SEQ ID NO: 7	catalytic core of peptidylglycine alpha-hydroxylating monooxygenase, rat (rPHMcc)
SEQ ID NO: 8	catalytic core of peptidylglycine alpha-hydroxylating monooxygenase, rat (rPHMcc(S44-V356))
SEQ ID NO: 9	peptidylglycine alpha-amidating monooxygenase, isoform 1, rat (rPAM)
SEQ ID NO: 10	peptidylglycine alpha-amidating monooxygenase, isoform 1, rat (rPAM)
SEQ ID NO: 11	nucleotide sequence of rCuA
SEQ ID NO: 12	amino acid sequence of rCuA
SEQ ID NO: 13	nucleotide sequence of rCuB
SEQ ID NO: 14	amino acid sequence of rCuB
SEQ ID NO: 15	nucleotide sequence of hCuA(S39-E213)
SEQ ID NO: 16	amino acid sequence of hCuA(S39-E213)
SEQ ID NO: 17	nucleotide sequence of hCuA(S39-I209)
SEQ ID NO: 18	amino acid sequence of hCuA(S39-I209)
SEQ ID NO: 19	nucleotide sequence of hCuA(S39-S204)
SEQ ID NO: 20	amino acid sequence of hCuA(S39-S204)
SEQ ID NO: 21	nucleotide sequence of hCuA(S39-L201)
SEQ ID NO: 22	amino acid sequence of hCuA(S39-L201)
SEQ ID NO: 23	nucleotide sequence of hCuA(S39-L195)
SEQ ID NO: 24	amino acid sequence of hCuA(S39-L195)
SEQ ID NO: 25	nucleotide sequence of hCuA(N40-E213)
SEQ ID NO: 26	amino acid sequence of hCuA(N40-E213)

SEQ ID NO: 27	nucleotide sequence of hCuA(N40K-E213)
SEQ ID NO: 28	amino acid sequence of hCuA(N40K-E213)
SEQ ID NO: 29	nucleotide sequence of hCuA(E41-E213)
SEQ ID NO: 30	amino acid sequence of hCuA(E41-E213)
SEQ ID NO: 31	nucleotide sequence of hCuA(E41K-E213)
SEQ ID NO: 32	amino acid sequence of hCuA(E41K-E213)
SEQ ID NO: 33	nucleotide sequence of hCuA(C42-E213)
SEQ ID NO: 34	amino acid sequence of hCuA(C42-E213)
SEQ ID NO: 35	nucleotide sequence of hCuA(V49-E213)
SEQ ID NO: 36	amino acid sequence of hCuA(V49-E213)
SEQ ID NO: 37	nucleotide sequence of hCuA(D53-E213)
SEQ ID NO: 38	amino acid sequence of hCuA(D53-E213)
SEQ ID NO: 39	nucleotide sequence of hCuB(K214-V351)
SEQ ID NO: 40	amino acid sequence of hCuB(K214-V351)
SEQ ID NO: 41	nucleotide sequence of hCuB(K214-S353)
SEQ ID NO: 42	amino acid sequence of hCuB(K214-S353)
SEQ ID NO: 43	nucleotide sequence of hCuB(V215K-V351)
SEQ ID NO: 44	amino acid sequence of hCuB(V215K-V351)
SEQ ID NO: 45	nucleotide sequence of hCuB(V216K-V351)
SEQ ID NO: 46	amino acid sequence of hCuB(V216K-V351)
SEQ ID NO: 47	nucleotide sequence of hCuB(N217-V351)
SEQ ID NO: 48	amino acid sequence of hCuB(N217-V351)
SEQ ID NO: 49	nucleotide sequence of hCuB(N217K-V351)
SEQ ID NO: 50	amino acid sequence of hCuB(N217K-V351)
SEQ ID NO: 51	nucleotide sequence of hCuB(K214-M337)
SEQ ID NO: 52	amino acid sequence of hCuB(K214-M337)
SEQ ID NO: 53	nucleotide sequence of hCuB(K214-T340)

SEQ ID NO: 54	amino acid sequence of hCuB(K214-T340)
SEQ ID NO: 55	nucleotide sequence of hCuB(K214-N346)
SEQ ID NO: 56	amino acid sequence of hCuB(K214-N346)
SEQ ID NO: 57	nucleotide sequence of rCuA(S44-E218)
SEQ ID NO: 58	amino acid sequence of rCuA(S44-E218)
SEQ ID NO: 59	nucleotide sequence of rCuB(K219-V356)
SEQ ID NO: 60	amino acid sequence of rCuB(K219-V356)
SEQ ID NO: 61	nucleotide sequence of signal peptide of PS1
SEQ ID NO: 62	amino acid sequence of signal peptide of PS1
SEQ ID NO: 63	nucleotide sequence of signal peptide of PS2 (CspB)
SEQ ID NO: 64	amino acid sequence of signal peptide of PS2 (CspB)
SEQ ID NO: 65	nucleotide sequence of signal peptide of SlpA (CspA)
SEQ ID NO: 66	amino acid sequence of signal peptide of SlpA (CspA)
SEQ ID NO: 67	artificial DNA-fragment encoding Exenatide precursor
SEQ ID NO: 68	Exenatide precursor having additional C-terminal Gly residue
SEQ ID NO: 69	artificial DNA-fragment <KpnI-Pr _{cspB} -pre1pro13-hCuA(S39-E213)-XbaI>
SEQ ID NO: 70	artificial DNA-fragment <KpnI-Pr _{cspB} -pre1pro11-hCuB(K214-V351)-XbaI>
SEQ ID NO: 71	artificial DNA-fragment <KpnI-Pr _{cspB} -pre1pro11-hPHMcc(S39-V351)-XbaI>
SEQ ID NO: 72	artificial DNA-fragment <KpnI-Pr _{cspB} -pre1pro11-rPHMcc(S44-V356)-XbaI>
SEQ ID NO: 73	artificial DNA-fragment <KpnI-Pr _{cspB} -pre1pro13-hCuA(S39-E213)-BamHI>
SEQ ID NO: 74	artificial DNA-fragment <KpnI-Pr _{cspB} -pre1pro12-hCuA(S39-E213)-BamHI>
SEQ ID NO: 75	artificial DNA-fragment <KpnI-Pr _{cspB} -pre1pro11-hCuA(S39-E213)-BamHI>

SEQ ID NO: 76	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(N40-E213)- <i>Bam</i> HI>
SEQ ID NO: 77	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro14-hCuA(N40-E213)- <i>Bam</i> HI>
SEQ ID NO: 78	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(N40K-E213)- <i>Bam</i> HI>
SEQ ID NO: 79	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(E41-E213)- <i>Bam</i> HI>
SEQ ID NO: 80	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(E41K-E213)- <i>Bam</i> HI>
SEQ ID NO: 81	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(C42-E213)- <i>Bam</i> HI>
SEQ ID NO: 82	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(V49-E213)- <i>Bam</i> HI>
SEQ ID NO: 83	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(D53-E213)- <i>Bam</i> HI>
SEQ ID NO: 84	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(S39-I209)- <i>Bam</i> HI>
SEQ ID NO: 85	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(S39-S204)- <i>Bam</i> HI>
SEQ ID NO: 86	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(S39-L201)- <i>Bam</i> HI>
SEQ ID NO: 87	artificial DNA-fragment < <i>KpnI</i> -Pr _{cspB} -pre1pro11-hCuA(S39-L195)- <i>Bam</i> HI>
SEQ ID NO: 88	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-rCuA(S44-E218)- <i>Xba</i> I>
SEQ ID NO: 89	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-hCuB(K214-V351)- <i>Xba</i> I>
SEQ ID NO: 90	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-hCuB(V215K-V351)- <i>Xba</i> I>
SEQ ID NO: 91	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-hCuB(V216K-V351)- <i>Xba</i> I>
SEQ ID NO: 92	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-hCuB(N217-V351)- <i>Xba</i> I>

SEQ ID NO: 93	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-hCuB(N217K-V351)- <i>Xba</i> I>
SEQ ID NO: 94	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-hCuB(K214-M337)- <i>Xba</i> I>
SEQ ID NO: 95	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-hCuB(K214-T340)- <i>Xba</i> I>
SEQ ID NO: 96	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-hCuB(K214-N346)- <i>Xba</i> I>
SEQ ID NO: 97	artificial DNA-fragment < <i>Bam</i> HI-Pr _{cspB} -pre1pro11-hCuB(K214-S353)- <i>Xba</i> I>
SEQ ID NO: 98	artificial DNA-fragment < <i>Xba</i> I-Pr _{cspB} -pre1pro11-rCuB(K219-V356)- <i>Sal</i> I>
SEQ ID NO: 99	promoter Pr _{cspB}
SEQ ID NO: 100	nucleotide sequence of artificial N-terminal propeptide pro11
SEQ ID NO: 101	amino acid sequence of artificial N-terminal propeptide pro11
SEQ ID NO: 102	nucleotide sequence of artificial N-terminal propeptide pro12
SEQ ID NO: 103	amino acid sequence of artificial N-terminal propeptide pro12
SEQ ID NO: 104	nucleotide sequence of artificial N-terminal propeptide pro13
SEQ ID NO: 105	amino acid sequence of artificial N-terminal propeptide pro13
SEQ ID NO: 106	nucleotide sequence of artificial N-terminal propeptide pro14
SEQ ID NO: 107	amino acid sequence of artificial N-terminal propeptide pro14
SEQ ID NO: 108	nucleotide sequence of hCuA(E41-I209)
SEQ ID NO: 109	amino acid sequence of hCuA(E41-I209)
SEQ ID NO: 110	primer P1
SEQ ID NO: 111	primer P2

SEQ ID NO: 112 primer P3

SEQ ID NO: 113 primer P4

Detailed Description of the Invention

<1> Method for production of a protein having peptidylglycine alpha-hydroxylating monooxygenase activity

1. *Coryneform bacterium*

In a method as described herein, a coryneform bacterium is used as a host, which bacterium is able to produce and secrete a protein, and has been modified so that a nucleic acid encoding the domain CuA of the peptidylglycine alpha-hydroxylating monooxygenase (PHM) and a nucleic acid encoding the domain CuB of the PHM are co-expressed separately in the bacterium. The explanations given hereinafter to the coryneform bacterium can be applied *mutatis mutandis* to any coryneform bacterium strain that can be used equivalently in the methods as described herein.

In a method as described herein, the coryneform bacteria are aerobic gram-positive bacilli, and include *Corynebacterium* bacteria, *Brevibacterium* bacteria, *Microbacterium* bacteria, and so forth. The coryneform bacteria include bacteria which were previously classified into the genus *Brevibacterium*, but have been united into the genus *Corynebacterium* (Liebl W. et al., Transfer of *Brevibacterium divaricatum* DSM 20297T, "*Brevibacterium flavum*" DSM 20411, "*Brevibacterium lactofermentum*" DSM 20412 and DSM 1412, and *Corynebacterium glutamicum* and their distinction by rRNA gene restriction patterns, *Int. J. Syst. Bacteriol.*, 1991, 41:255-260). The coryneform bacteria also include bacteria which were previously classified into *Corynebacterium ammoniagenes*, but have been reclassified into *Corynebacterium stationis* by nucleotide sequence analysis of 16S rRNA and so forth (Bernard K.A. et al., Assignment of *Brevibacterium stationis* (ZoBell and Upham 1944) Breed 1953 to the genus *Corynebacterium*, as *Corynebacterium stationis* comb. nov., and emended description of the genus *Corynebacterium* to include isolates that can alkalinize citrate, *Int. J. Syst. Evol. Microbiol.*, 2010, 60:874-879). One advantage of using the coryneform

bacteria is that they inherently secrete an extremely small amount of proteins out of cells compared with fungi, yeasts, *Bacillus* bacteria, etc., which are conventionally used for production of proteins, and therefore the purification process of a heterologous protein produced by the coryneform bacterium should be simplified or eliminated. Another advantage of using coryneform bacteria is that they grow well in a simple medium containing a saccharide, ammonia, mineral salts, etc., and therefore they are excellent in view of cost of medium, culture method, culture productivity, and so forth.

Specific examples of coryneform bacteria include the following species:

Corynebacterium acetoacidophilum,
Corynebacterium acetoglutamicum,
Corynebacterium alkanolyticum,
Corynebacterium callunae,
Corynebacterium glutamicum,
Corynebacterium lilium,
Corynebacterium melassecola,
Corynebacterium thermoaminogenes (*Corynebacterium efficiens*),
Corynebacterium herculis,
Brevibacterium divaricatum,
Brevibacterium flavum,
Brevibacterium immariophilum,
Brevibacterium lactofermentum (*Corynebacterium glutamicum*),
Brevibacterium roseum,
Brevibacterium saccharolyticum,
Brevibacterium thiogenitalis,
Corynebacterium ammonia genes (*Corynebacterium stationis*),
Brevibacterium album,
Brevibacterium cerinum,
Microbacterium ammoniophilum.

Specific examples of coryneform bacteria include the following strains:

Corynebacterium acetoacidophilum ATCC 13870,

Corynebacterium acetoglutamicum ATCC 15806,
Corynebacterium alkanolyticum ATCC 21511,
Corynebacterium callunae ATCC 15991,
Corynebacterium glutamicum ATCC 13020, ATCC 13032, ATCC 13060,
ATCC 13869, FERM BP-734,
Corynebacterium lilium ATCC 15990,
Corynebacterium melassecola ATCC 17965,
Corynebacterium thermoaminogenes AJ12340 (FERM BP-1539),
Corynebacterium herculis ATCC 13868,
Brevibacterium divaricatum ATCC 14020,
Brevibacterium flavum ATCC 13826, ATCC 14067, AJ12418 (FERM BP-2205),
Brevibacterium immariophilum ATCC 14068,
Brevibacterium lactofermentum ATCC 13869,
Brevibacterium roseum ATCC 13825,
Brevibacterium saccharolyticum ATCC 14066,
Brevibacterium thiogenitalis ATCC 19240,
Corynebacterium ammoniagenes (*Corynebacterium stationis*) ATCC 6871, ATCC 6872,
Brevibacterium album ATCC 15111,
Brevibacterium cerinum ATCC 15112,
Microbacterium ammoniaphilum ATCC 15354.

These strains are available from, for example, the American Type Culture Collection (ATCC; Address: P.O. Box 1549, Manassas, VA 20108, United States of America). That is, registration numbers are assigned to the respective strains, and the strains can be ordered by using these registration numbers (refer to <http://www.atcc.org/>). The registration numbers of the strains are listed in the catalogue of the American Type Culture Collection.

In particular, the *Corynebacterium glutamicum* (*C. glutamicum*) AJ12036 strain (FERM BP-734), which was isolated as a streptomycin (Sm)-resistant mutant strain from the wild-type strain, *C. glutamicum* ATCC 13869, is predicted to have a mutation in the functional gene responsible for secretion of proteins, and has an extremely high ability to produce and

secrete heterologous proteins, as high as about 2 to 3 times in terms of accumulation amount of proteins under optimum culture conditions, compared with the parent strain such as, for example, a wild-type strain. Therefore, this strain is preferred as a host bacterium. The AJ12036 strain was originally deposited at the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (currently, the incorporated administrative agency, National Institute of Technology and Evaluation (NITE), International Patent Organism Depository (IPOD), Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, 305-8566, Japan) on March 26, 1984 as an international deposit, and assigned an accession number of FERM BP-734.

Moreover, a host strain able to produce and secrete a protein may be obtained from such parental coryneform bacterial strains as described above by using a mutagenesis method or a genetic recombination method. For example, after a parental strain is treated with ultraviolet irradiation or a chemical mutation agent such as N-methyl-N'-nitrosoguanidine, a strain with enhanced ability to produce and secrete a protein can be selected.

Furthermore, a host coryneform bacterium can be modified so that it does not produce a cell surface layer protein. If a strain that does not produce a cell surface layer protein is used as the host, purification of the heterologous protein secreted into the medium or on the cell surface layer is easy. Such modification can be carried out by introducing a mutation into the coding region of the cell surface layer protein or an expression control region thereof, on the chromosome by mutagenesis or genetic recombination. Examples of coryneform bacterium modified so to not produce a cell surface layer protein include the *C. glutamicum* YDK010 strain (WO2004029254 A1), which is a cell surface layer protein PS2 (also referred to as "CspB") deficient strain of the *C. glutamicum* AJ12036 strain (FERM BP-734). Depending on the chosen coryneform bacterial strain, it may be deficient in cell surface layer protein PS1, PS2 (CspB), and/or SlpA (also referred to as "CspA").

Yet furthermore, a host coryneform bacterium may be modified further to reduce the activity of a penicillin-binding protein. Examples of coryneform

bacterial strains modified to reduce the activity of a penicillin-binding protein include the *C. glutamicum* YDK010 Δ pbp1A strain (WO2013065772 A1), which is a penicillin-binding protein (PBP1a) deficient strain of the *C. glutamicum* YDK010 strain (WO2004029254 A1). Nonetheless, the bacterium may also be modified to reduce activity of another penicillin-binding protein such as class B HMW-PBPs and LMW-PBPs (WO2013065772 A1). An advantage of using a coryneform bacterium modified to reduce the activity of a penicillin-binding protein is that the production of the protein is improved further compared with that observed for a non-modified strain. Such a modification can be carried out by, for example, reducing expression of a gene encoding the protein, disrupting the gene encoding the protein, or using any means known to the one skilled in the art such as those that can be used to render the coryneform bacterium not able to produce a cell surface layer protein as described above.

The bacterium as described herein has the “ability to produce and secrete a protein”. The phrase “ability to produce and secrete a protein” can refer to an ability of the bacterium to secrete the protein into a medium and/or the cell surface layer such that the protein accumulates to such an extent that the protein can be collected from the medium and/or the cell surface layer, when the bacterium is cultured in the medium. The accumulation amount may be, for example, in terms of the accumulation amount in the medium, 10 μ g/L or more, 1 mg/L or more, 100 mg/L or more, or 1 g/L or more. Also, the accumulation amount, if in the cell surface layer, can be, for example, such that if the protein in the cell surface layer is collected and suspended in a liquid of the same volume as the medium, the concentration of the protein in the suspension is 10 μ g/L or more, 1 mg/L or more, or 100 mg/L or more. The protein that can be produced by the method as described herein can also be referred to as peptides and polypeptides.

The protein to be produced is not particularly limited so long as it is a heterologous protein. The phrase “heterologous protein” can refer to an exogenous protein relative to the coryneform bacterium that expresses and secretes that protein. The heterologous protein may be, for example, a protein native to a microorganism, a protein native to a plant, a protein

native to an animal, a protein native to an insect, a protein native to a virus, or even an artificially-designed protein. The protein can be a monomeric protein or a multimeric protein. A monomeric protein can refer to a protein that has one subunit. Specific examples of the “monomeric protein” can include domains CuA and CuB of a PHM native to various species. The more detailed explanations of the domains CuA and CuB are provided herein.

The multimeric protein can refer to a protein that may exist as a multimer having two or more subunits. In the multimer, the subunits may be linked by covalent bonds such as disulfide bonds, linked by non-covalent bonds such as hydrogen bonds and hydrophobic interaction, or linked by a combination of these. The multimer can include one or more intermolecular disulfide bonds. The multimer may be a homo-multimer having a single kind of subunit, or may be a hetero-multimer having two or more kinds of subunits. When the multimeric protein is a hetero-multimer, it is sufficient that at least one subunit in the hetero-multimer is a heterologous protein. That is, all the subunits may be heterologous, or only one or some of the subunits may be heterologous. The multimeric protein may be a secretory protein in nature, or may be a non-secretory protein in nature. The multimeric protein can include a single kind of protein, or two or more kinds of proteins. When the multimeric protein is a hetero-multimer, all the subunits in the hetero-multimer can be produced according to methods described herein. A specific example of the “multimeric protein” is PHM. More detailed explanations of PHM as derived from or native to various species are provided herein.

Furthermore, the protein to be produced by the method as described herein may be a protein having a pro-structure moiety (proprotein). When the protein is a proprotein, the proprotein may be processed into the mature protein by cleavage of the pro-structure moiety, or may exist as the proprotein. The cleavage of the pro-structure moiety can be attained with, for example, a protease. When a protease is used, generally, the proprotein can be cleaved at a position that is substantially the same as that of the natural protein, or can be at exactly the same position as that of the natural protein. In this way, the mature protein is identical to the natural mature protein, in

view of their activity. Therefore, generally, a protease that cleaves the proprotein at a position so that the obtained protein is the same as the naturally occurring mature protein is preferred particular example. However, the N-terminal region of the obtained mature protein may not be the same as that of the natural protein as described above. For example, depending on type, purpose of use, *etc.* of the protein to be produced, the obtained protein may have an N-terminus that is longer or shorter by one to several amino acid residues as compared with the natural protein since such a protein may have higher activity. Proteases that can be used in the present invention can include, for example, commercially available proteases such as Dispase (produced by Boehringer Mannheim) as well as proteases obtained from culture broth of a microorganism such as actinomycetes. Such proteases may be used in an unpurified state, or after purification to an appropriate purity as required.

The phrase “native to” in reference to a protein or a nucleic acid native to a particular species such as, for example, human, horse, pig and rat can refer to a protein or a nucleic acid that is native to that species. That is, a protein or a nucleic acid native to a particular species can mean the protein or the nucleic acid, respectively, that exists naturally in the species and can be isolated from that species and sequenced using means known to the one of ordinary skill in the art. Moreover, as the amino acid sequence or the nucleotide sequence of a protein or nucleic acid, respectively, isolated from a species in which the protein or nucleic acid exists, can easily be determined, the phrase “native to” in reference to a protein or a nucleic acid can also refer to a protein or a nucleic acid that can be obtained using, for example, a genetic engineering technique, including recombinant DNA technology, or a chemical synthesis method, or the like, so long as the amino acid sequence of the protein or the nucleotide sequence of the nucleic acid thus obtained is identical, accordingly, to the amino acid sequence of the protein or the nucleotide sequence of the nucleic acid that exists naturally in the species. Examples of amino acid sequences native to particular species include, but are not limited to, peptides, oligopeptides, polypeptides, including proteins, specifically enzymes, and so forth. Examples of nucleotide sequences native

to particular species include deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), and these are not limited to regulatory sequences, including promoters, attenuators, terminators, and the like, genes, intergenic sequences, sequences encoding signal peptides, pro-moieties of proteins, artificial amino acid sequences, and so forth. Specific examples of amino acid sequences and nucleotide sequences, and homologues thereof native to various species are described herein, and these examples include PHM having the amino acid sequence shown in SEQ ID NO: 6 or 8, which are native to human and rat, respectively, encoded by the corresponding genes having the nucleotide sequence shown in SEQ ID NOs: 5 and 7.

The phrase that a protein is “secreted” can mean that the protein is transported out of the bacterial cell, that is, the protein is extracellularly transported. The phrase that a protein is “secreted” can include when all the molecules of the protein are transported out of the cell and are present in the medium in free form, when all the molecules of the protein are present in the cell surface layer, or when some of the protein molecules are present in the medium and some are present in the cell surface layer.

The bacterium can be modified so that a nucleic acid encoding domain CuA of PHM and a nucleic acid encoding domain CuB of PHM (cumulatively, nucleic acids encoding domains CuA and CuB of PHM) are co-expressed separately in the bacterium such that the bacterium is able to produce and secrete a protein having an activity of PHM. The bacterium can be obtained by modifying a coryneform bacterium that is able to produce and secrete a protein so that the nucleic acids encoding domains CuA and CuB of PHM are co-expressed separately in the bacterium. Alternatively, the bacterium can also be obtained by modifying a coryneform bacterium so that the nucleic acids encoding domains CuA and CuB of PHM are co-expressed separately in the bacterium, and then imparting the ability to produce and secrete a protein to it. The modification and impartation of the ability to produce and secrete a protein and constructing the bacterium can be carried out in any arbitrary order. The bacterium as described herein may be obtained from a bacterium that can produce and secrete a protein even before it is modified so that the nucleic acids encoding domains CuA and CuB of PHM are co-expressed

separately in the bacterium. Alternatively, the bacterium as described herein may also be obtained from a bacterium that is not able to produce and secrete a protein although it may have the genetic constructs for secretory expression of a protein, but by virtue of the modification so that nucleic acids encoding domains CuA and CuB of PHM are co-expressed separately in the bacterium, the bacterium is then able to produce and secrete the protein. Virtually, any coryneform bacterium can be used in the method as described herein, provided that the bacterium has an ability to produce and secrete a protein and it is modified so that the nucleic acids encoding domains CuA and CuB of PHM are co-expressed separately in the bacterium, and the protein having an activity of PHM can be produced.

A protein, that can be a monomeric protein or a multimeric protein, which can be produced using a bacterium as described herein, is not particularly limited so long as the protein can be produced by the bacterium of the method as described herein. For example, the phrase “a monomeric protein” can include either of the domains CuA and CuB. As the PHM may be native to various species, the phrase “a monomeric protein” can include, for example, any CuA and CuB domain native to the mammalian PHM. The phrase “a multimeric protein” can include a protein PHM that includes domains CuA and CuB, and the specific examples include, but are not limited to, the PHM native to mammals such as, for example, human, monkey, rat, horse, pig, and the like. The human PHM and rat PHM are particular examples, and these proteins and their homologues are described herein.

It is known that a secretory protein is generally translated as a preprotein (also referred to as “prepeptide”) or a preproprotein (also referred to as “prepropeptide”), and then becomes a mature protein through processing. Specifically, a secretory protein is generally translated as a preprotein or preproprotein having a signal peptide or pre-moiety, which is then cleaved with a protease (generally called signal peptidase), and the secretory protein is thereby converted into a mature protein or proprotein. As for the proprotein, the pro-moiety thereof is further cleaved by a protease, and the proprotein thereby becomes a mature protein. Therefore, a signal

peptide can be used for production of a protein in the method as described herein. A preprotein and a preproprotein of a secretory protein may be collectively referred to as "secretory protein precursor". The "signal peptide", also referred to as "signal sequence", can refer to an amino acid sequence present at the N-terminus of a secretory protein precursor, and is usually not present in a natural mature protein. A secretory protein can be produced as a mature protein devoid of a signal peptide and the pro-moiety thereof, or as a proprotein devoid of a signal peptide and having the pro-moiety thereof.

The coryneform bacterium as used in the method as described herein can be able to co-express separately a nucleic acid encoding domain CuA of PHM and a nucleic acid encoding domain CuB of PHM. This can be attained by introducing the first genetic construct for production of the protein (also referred to as "the first genetic construct") and the second genetic construct for production of the protein (also referred to as "the second genetic construct") into the coryneform bacterium as mentioned above so that the bacterium harbors the genetic constructs, wherein said first genetic construct and second genetic construct each includes a nucleic acid encoding the domain CuA or CuB of the PHM, provided that when the domain CuA is present in the first genetic construct, the second genetic construct includes the domain CuB, or when the domain CuA is present in the second genetic construct, the first genetic construct includes the domain CuB. That is, the coryneform bacterium used in the method as described herein can be obtained by introducing the first genetic construct and second genetic construct such that each genetic construct includes the nucleic acid encoding the domain CuA or CuB of the PHM to render the bacterium harboring nucleic acids encoding the domains CuA and CuB. Therefore, the phrase "co-expressed separately" in relation to a nucleic acid encoding domain CuA of PHM and a nucleic acid encoding domain CuB of PHM, can mean that the nucleic acids encoding domains CuA and CuB of PHM are harbored by the bacterium on separate genetic constructs so that expression of the first genetic construct can be attained independently from expression of the second genetic construct (that is, separately), and the nucleic acids encoding the domains CuA and CuB in the genetic constructs can be

expressed in such a way when the expression of the first genetic construct occurs, the expression of the second genetic construct occurs also, or contrariwise (that is, co-expressed).

The phrases “the first genetic construct” and “the second genetic construct” can be referred to as generally “a genetic construct”, which can also be referred to as equivalently “a genetic construct for production of a protein” and “a genetic construct used for the present invention”. That is, the bacterium of the method as described herein harbors, at least, two genetic constructs for production of a protein. The first genetic construct and second genetic construct can be introduced into a coryneform bacterium in such a way that the genetic constructs are present on different nucleic acid molecules. Alternatively, the first genetic construct and second genetic construct can be introduced into a coryneform bacterium in such a way that the genetic constructs are present on one nucleic acid molecule. Specifically, for example, the genetic constructs for production of a protein may be present in a single expression vector, or present on the chromosome. Alternatively, the genetic constructs for production of the protein may be present on two or more expression vectors, or may be separately present on one or more expression vectors and the chromosome. When the genetic constructs are present on different nucleic acid molecules or are present on a sole nucleic acid molecule, the constructs can be present in the nucleic acid molecule(s) in such a way that expression of the first genetic construct does not affect expression of the second genetic construct, and contrariwise, so that co-expression of the nucleic acids encoding the domains CuA and CuB in the genetic constructs can be attained. Virtually, any way of introducing the genetic constructs into the coryneform bacterium of the method as described herein can be chosen so long as the expression of the genetic constructs can be attained in the bacterium. The phrases “expression can be attained” and “co-expression can be attained” can refer to the case when transcription from the DNA can take place such that RNA that complements the DNA as a template can be synthesized. The phrases “expression can be attained” and, specifically, “co-expression can be attained” also refer to the case when transcription from the DNA can occur such that the RNA that

complements the DNA as a template can be synthesized, and translation from the RNA can occur such that a peptide such as, for example, a protein can be produced according to the method as described herein. Therefore, the phrase “expression can be attained independently” with regard to genetic constructs can mean that the expressions of genetic constructs can be attained. That is, the phrase “expression can be attained independently” with regard to genetic constructs can mean that when expression of the first genetic construct can be attained, the expression of the second genetic construct can be attained also, and contrariwise. That is, alternatively, the phrase “expression can be attained independently” with regard to genetic constructs can mean that the co-expression of genetic constructs can be attained.

Specific examples of “a genetic construct for production of a protein” include “a genetic construct for production of a protein having peptidylglycine alpha-hydroxylating monooxygenase activity”. As the protein having an activity of PHM produced according to the method as described herein may include domains CuA and CuB, it is acceptable that “a genetic construct for production of a protein having peptidylglycine alpha-hydroxylating monooxygenase activity” can refer also, equivalently, to the first genetic construct and the second genetic construct that include the nucleic acids encoding domains CuA and CuB as explained above.

Examples of nucleic acids suitable for the expression of nucleic acids in general and co-expression of the genetic constructs as described herein, in particular, can include those nucleic acids that can function in a coryneform bacterium, that is, at least transcription from DNA can occur such that the RNA that complements the DNA as a template can be synthesized, and, optionally, translation from the RNA can take place such that a peptide such as, for example, a protein can be produced according to the method as described herein. The examples can include chromosomal DNA, plasmids, vectors, and so forth, that are native to or naturally present in a coryneform bacterium and can function in the bacterium. As the vector, a vector autonomously replicable in a cell of the coryneform bacterium can be used. The vector can be a multi-copy vector. Furthermore, the vector can include a

marker such as an antibiotic resistance gene for selection of transformants. The vector may be, for example, a vector native to a bacterial plasmid, a vector native to a yeast plasmid, a vector native to a bacteriophage, cosmid, phagemid, or the like. Specific examples of vector autonomously replicable in coryneform bacteria can include pHM1519 (Miwa K. *et al.*, *Agric. Biol. Chem.*, 1984, 48:2901-2903); pAM330 (Miwa K. *et al.*, 1984); plasmids obtained by improving these and having a drug resistance gene; plasmid pCRY30 described in Japanese Patent Laid-open (Kokai) No. 3-210184; plasmids pCRY21, pCRY2KE, pCRY2KX, pCRY31, pCRY3KE, and pCRY3KX described in Japanese Patent Laid-open (Kokai) No. 2-72876 and U.S. Patent No. 5,185,262; plasmids pCRY2 and pCRY3 described in Japanese Patent Laid-open (Kokai) No. 1-191686; pAM330 described in Japanese Patent Laid-open (Kokai) No. 58-67679; pHM1519 described in Japanese Patent Laid-open (Kokai) No. 58-77895; pAJ655, pAJ611, and pAJ1844 described in Japanese Patent Laid-open (Kokai) No. 58-192900; pCG1 described in Japanese Patent Laid-open (Kokai) No. 57-134500; pCG2 described in Japanese Patent Laid-open (Kokai) No. 58-35197; pCG4 and pCG11 described in Japanese Patent Laid-open (Kokai) No. 57-183799; pVK7 described in Japanese Patent Laid-open (Kokai) No. 10-215883; pVC7 described in Japanese Patent Laid-open (Kokai) No. 9-070291; and so forth.

Methods which can be used to introduce a nucleic acid such as, for example, a genetic construct, a gene, a vector, and the like into a coryneform bacterium can include, but are not limited to, genetic engineering methods known to persons of ordinary skill in the art, and these are not particularly limited. In the bacterium of the method as described herein, the genetic construct(s) can be present on a vector that autonomously replicates outside of the chromosome such as a plasmid, or may be incorporated into the chromosome. In addition, as described above, for constructing the bacterium of the method as described herein, modifications such as introduction of the genetic construct(s), impartation or enhancement of the ability to produce a protein by secretory production, reduction of the activity of a penicillin-binding protein, and reduction of the activity of a cell surface layer protein can be performed in an arbitrary order.

The genetic construct as described herein can be introduced into a host by using, for example, a vector that includes the genetic construct. The vector is not particularly limited so long as a vector autonomously replicable in a coryneform bacterium is chosen, and may be, for example, a vector based on a bacterial plasmid, a vector based on a yeast plasmid, a vector based on a bacteriophage cosmid or phagemid, or the like. As the vector, for example, a plasmid native to a coryneform bacterium is preferred. Specific examples of vectors autonomously replicable in coryneform bacteria are described above.

Furthermore, an artificial transposon and so forth can also be used. When a transposon is used, a protein-encoding gene is introduced into a chromosome by homologous recombination or the translocation ability of the transposon itself. Other examples of the introduction method utilizing homologous recombination can include, for example, the methods utilizing a linear DNA, a plasmid having a temperature sensitive replication origin, a plasmid capable of conjugative transfer, a suicide vector having a replication origin that does not function in the chosen host, and so forth. In addition, when a protein gene is introduced into a chromosome so that the genetic construct is present on the chromosome, either one or both of the promoter sequence and the nucleic acid encoding a signal peptide may be originally present on the host chromosome. Specifically, for example, by using a promoter sequence that is native to the host chromosome and a nucleic acid encoding a signal peptide that is also native to the host chromosome and is ligated downstream from the promoter sequence, and then replacing only the gene ligated downstream from the nucleic acid encoding the signal peptide with a gene encoding an objective protein, the genetic construct is also present on the chromosome. The more detailed explanations of the genetic construct are provided herein.

The method for introducing the genetic constructs into a coryneform bacterium is not particularly limited, and a generally used method, for example, the protoplast method (Miwa K. *et al.*, *Gene*, 1985, 39:281-286), the electroporation method (Dunican L.K. and Shivnan E., *Nat. Biotechnol.*, 1989,

7:1067-1070), the electric pulse method (JP H2-207791 A), and so forth can be used.

The copy number, presence, or absence of the gene can be measured, for example, by restricting the chromosomal DNA followed by Southern blotting using a probe based on the gene sequence, fluorescence *in situ* hybridization (FISH), and the like. The level of gene expression can be determined by measuring the amount of mRNA transcribed from the gene using various well-known methods, including Northern blotting, quantitative RT-PCR, and the like. The amount of the protein encoded by the gene can be measured by known methods including SDS-PAGE followed by immunoblotting assay (Western blotting analysis), or mass spectrometry analysis of the protein samples, and the like.

Methods for manipulation with recombinant molecules of DNA and molecular cloning such as preparation of plasmid DNA, digestion, ligation and transformation of DNA, selection of an oligonucleotide as a primer, incorporation of mutations, and the like may be ordinary methods well-known to persons of ordinary skill in the art. These methods are described, for example, in Sambrook J., Fritsch E.F. and Maniatis T., "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor Laboratory Press (1989) or Green M.R. and Sambrook J.R., "Molecular Cloning: A Laboratory Manual", 4th ed., Cold Spring Harbor Laboratory Press (2012); Bernard R. Glick, Jack J. Pasternak and Cheryl L. Patten, "Molecular Biotechnology: principles and applications of recombinant DNA", 4th ed., Washington, DC, ASM Press (2009).

To evaluate the degree of protein or DNA homology, several calculation methods can be used, such as a BLAST search, FASTA search and ClustalW method. The BLAST (Basic Local Alignment Search Tool, www.ncbi.nlm.nih.gov/BLAST/) search is the heuristic search algorithm employed by the programs blastp, blastn, blastx, megablast, tblastn, and tblastx; these programs ascribe significance to their findings using the statistical methods of Karlin S. and Altschul S.F. ("Methods for assessing the statistical significance of molecular sequence features by using general

scoring schemes" *Proc. Natl. Acad. Sci. USA*, 1990, 87:2264-2268; "Applications and statistics for multiple high-scoring segments in molecular sequences". *Proc. Natl. Acad. Sci. USA*, 1993, 90:5873-5877). The computer program BLAST calculates three parameters: score, identity and similarity. The FASTA search method is described by Pearson W.R. ("Rapid and sensitive sequence comparison with FASTP and FASTA", *Methods Enzymol.*, 1990, 183:63-98). The ClustalW method is described by Thompson J.D. *et al.* ("CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice", *Nucleic Acids Res.*, 1994, 22:4673-4680),

The bacterium can have, in addition to the properties already mentioned, other specific properties such as various nutrient requirements, drug resistance, drug sensitivity, and drug dependence.

2. *Peptidylglycine alpha-hydroxylating monooxygenase*

Hereinafter, peptidylglycine alpha-hydroxylating monooxygenase (PHM) from various species and genes encoding it are described.

PHM, along with peptidyl-alpha-hydroxylglycine-alpha-amidating lyase (PAL), is a catalytic subunit of peptidylglycine alpha-amidating monooxygenase (PAM). The individual PHM and PAL subunits are each catalytically competent and connected through a peptide linker (Kolhekar A.S. *et al.*, 1997, and references therein). The PAM from various species is known, and the examples include, but are not limited to, PAM native to mammals and insects. As PAM includes PHM, the PHM native to various mammalian species are also known, and the examples include, but are not limited to, the PHM native to human, monkey, rat, horse, pig, and so forth. The PHM native to human (also referred to as "human PHM", "hPHM") and rat (also referred to as "rat PHM", "rPHM") are particular examples.

It is known that a less-sized PHM protein having shortened N- and C-termini, also referred to as "a catalytic core of PHM" ("PHMcc"), can be obtained from the wild-type PHM (also referred to as "native PHM" or "natural PHM"), that is, the PHM not having the shortened N- and C-termini (Kolhekar

A.S. *et al.*, 1997; Handa S. *et al.*, 2012). The peptidylglycine alpha-hydroxylating monooxygenase not having the shortened N- and C-termini can also be referred to as “full-sized peptidylglycine alpha-hydroxylating monooxygenase” or, simply, “full-sized PHM”. The PHMcc, which is stable and has an activity of PHM, can be obtained from the full-sized PHM by shortening the N- and C-termini. Therefore, the phrase “peptidylglycine alpha-hydroxylating monooxygenase” (“PHM”) can be applied equivalently to the full-sized PHM and the PHM having shortened N- and C-termini (PHMcc) so long as the full-sized PHM and PHMcc have the activity of PHM. That is, the explanations given herein to “PHM” can be applied *mutatis mutandis* to “PHMcc” and, specifically, to “hPHMcc” and “rPHMcc”.

As the nucleotide sequences of genes encoding PAM are known, the nucleotide sequences of nucleic acids encoding PHM, including full-sized PHM and PHMcc, are known also. Correspondingly, as a protein is encoded by a gene, amino acid sequences of PHM, including full-sized PHM and PHMcc, are known also.

Hereinafter, more detailed explanations of nucleotide sequences of PHMcc native to human (hPHMcc) and rat (rPHMcc), and the amino acid sequences encoded thereby, are described.

The amino acid sequence of hPHMcc (SEQ ID NO: 6) has the position numbers from S39 (serine-39) to V351 (valine-351) in the amino acid sequence of PAM, isoform 1 from *Homo sapiens* (UniProt Knowledgebase: locus AMD_HUMAN, accession number: P19021; SEQ ID NO: 2), and it is encoded by the nucleotide sequence shown in SEQ ID NO: 5.

The amino acid sequence of rPHMcc (SEQ ID NO: 8) has the position numbers from S44 (serine-44) to V356 (valine-356) in the amino acid sequence of PAM, isoform X3 from *Rattus norvegicus* (NCBI Reference Sequence: XP_006245656.1; SEQ ID NO: 4), and it is encoded by the nucleotide sequence shown in SEQ ID NO: 7.

Since the nucleotide sequence of the gene encoding PHM may differ depending on the originating species, the gene encoding a PHM may be a

variant of any herein described nucleotide sequences so long as it encodes a protein having peptidylglycine alpha-hydroxylating monooxygenase activity. The variants of the hPHM and rPHM genes can include homologues of the genes. Generally, homologues of the PHM genes native to *Homo sapiens* and *Rattus norvegicus*, including homologues of the hPHMcc and rPHMcc genes from the same species, can easily be obtained from public databases by BLAST search or FASTA search using the wild-type (also referred to as “native” or “natural”) hPHM gene or rPHM gene, or their shortened aforementioned counterparts hPHMcc and rPHMcc genes, respectively, as a query sequence, and can also be obtained by PCR (polymerase chain reaction) using a chromosome of a mammalian species as a template and oligonucleotides primers prepared on the basis of a known gene sequence.

For example, amino acid homologues of the hPHM and rPHM, that are native to *Homo sapiens* (human) and *Rattus norvegicus* (rat), respectively, can be obtained from public databases by BLAST search or FASTA search. In particular, the PHMcc from other species, besides *Homo sapiens* and *Rattus norvegicus*, can be obtained using the search for homologous amino acid sequences, such as, for example, eqPHMcc from *Equus caballus* (horse) and ssPHMcc from *Sus scrofa* (pig) (FIG.1, Table 1).

A protein having an activity of PHM produced according to the method as described herein can include domains CuA and CuB as described above. The protein may have copper ions as a cofactor to have the activity of PHM, and, optionally, it can further include water molecules or other additives such as, for example, monovalent metal ions bound to the protein.

The phrase “wild-type” in reference to a gene (for example, “a wild-type gene”) or a protein (for example, “a wild-type protein”) can mean a native gene or protein, naturally expressed in or produced by a wild-type coryneform bacterium, for example, by the wild-type *C. glutamicum* ATCC 13032 or *C. glutamicum* ATCC 13869 strain. A wild-type protein can be encoded by the “wild-type gene” naturally occurring in genome of a wild-type bacterium.

The gene encoding PHM may be a gene encoding a protein having the aforementioned amino acid sequence and can include substitution, deletion,

insertion, or addition of one or several amino acid residues at one or several positions so long as the gene encodes a protein having peptidylglycine alpha-hydroxylating monooxygenase activity. In such a case, usually 70% or more, 80% or more, 90% or more of the activity is maintained as compared to the protein not having any substitution, deletion, insertion, or addition of one or several amino acid residues. Although the number of "one or several" amino acid residues may differ depending on the position in the three-dimensional structure of the protein or types of amino acid residues, specifically, it 1 to 20, 1 to 10, or 1 to 5.

The aforementioned substitution, deletion, insertion, or addition of one or several amino acid residues can be a conservative mutation that maintains the normal function of the protein. Typical examples of conservative mutations are conservative substitutions. The conservative substitution is a mutation wherein substitution takes place mutually among Phe, Trp, and Tyr, if the substitution site is an aromatic amino acid; among Leu, Ile, and Val, if it is a hydrophobic amino acid; between Gln and Asn, if it is a polar amino acid; among Lys, Arg, and His, if it is a basic amino acid; between Asp and Glu, if it is an acidic amino acid; and between Ser and Thr, if it is an amino acid having a hydroxyl group. Examples of substitutions considered as conservative substitutions include, specifically, substitution of Ser or Thr for Ala, substitution of Gln, His, or Lys for Arg, substitution of Glu, Gln, Lys, His, or Asp for Asn, substitution of Asn, Glu, or Gln for Asp, substitution of Ser or Ala for Cys, substitution of Asn, Glu, Lys, His, Asp, or Arg for Gln, substitution of Gly, Asn, Gln, Lys, or Asp for Glu, substitution of Pro for Gly, substitution of Asn, Lys, Gln, Arg, or Tyr for His, substitution of Leu, Met, Val, or Phe for Ile, substitution of Ile, Met, Val, or Phe for Leu, substitution of Asn, Glu, Gln, His, or Arg for Lys, substitution of Ile, Leu, Val, or Phe for Met, substitution of Trp, Tyr, Met, Ile, or Leu for Phe, substitution of Thr or Ala for Ser, substitution of Ser or Ala for Thr, substitution of Phe or Tyr for Trp, substitution of His, Phe, or Trp for Tyr, and substitution of Met, Ile, or Leu for Val. Further, such substitution, deletion, insertion, addition, inversion or the like of amino acid residues as mentioned above includes a naturally

occurring mutation due to an individual difference, or a difference of species from which the gene is derived (mutant or variant).

Furthermore, the gene providing such a conservative mutation as mentioned above may be a gene encoding a protein having a homology of 80% or more, 90% or more, 95% or more, 97% or more, or 99% or more, to the total encoded amino acid sequence, and having peptidylglycine alpha-hydroxylating monooxygenase activity. In addition, in this specification, "homology" may mean "identity", that is the identity of amino acid residues. The sequence identity between two sequences is calculated as the ratio of residues matching in the two sequences when aligning the two sequences so as to achieve a maximum alignment with each other.

Moreover, the gene encoding PHM may be a DNA that is able to hybridize with a probe that can be prepared from a known gene sequence, such as a sequence complementary to a part or the whole of the aforementioned nucleotide sequence, under stringent conditions, and encodes a protein having peptidylglycine alpha-hydroxylating monooxygenase activity. The "stringent conditions" can refer to conditions under which a so-called specific hybrid is formed, and a non-specific hybrid is not formed. Examples of the stringent conditions can include those under which highly homologous DNAs hybridize to each other, for example, DNAs not less than 80% homologous, not less than 90% homologous, not less than 95% homologous, not less than 97% homologous, not less than 99% homologous, hybridize to each other, and DNAs less homologous than the above do not hybridize to each other, or conditions of washing of typical Southern hybridization, i.e., conditions of washing once, or 2 or 3 times, at a salt concentration and temperature corresponding to 1 x SSC (saline-sodium citrate), 0.1% SDS (sodium dodecyl sulfate) at 60°C, 0.1 x SSC, 0.1% SDS at 60°C, or 0.1 x SSC, 0.1% SDS at 68°C.

The probe used for the aforementioned hybridization may be a part of a sequence that is complementary to the gene as described above. Such a probe can be prepared by PCR using oligonucleotides prepared on the basis of a known gene sequence as primers and a DNA fragment containing the

nucleotide sequence as a template. For example, when a DNA fragment having a length of about 300 bp is used as the probe, the washing conditions of the hybridization may be, for example, 50°C, 2 x SSC and 0.1% SDS.

Furthermore, although a naturally occurring nucleic acid having a nucleotide sequence encoding PHM can be used as it is, a nucleotide sequence encoding the PHM in which an arbitrary codon is replaced with an equivalent codon, that is, synonymous amino acid codon may also be used according to the standard genetic code table (see, e.g., Lewin B., "Genes VIII", 2004, Pearson Education, Inc., Upper Saddle River, NJ 07458). For example, the nucleotide sequence encoding a PHM may be modified so that it has optimal codons according to codon frequencies in the chosen host.

The above descriptions concerning variants of the genes and proteins can also be applied *mutatis mutandis* to arbitrary proteins such as cell surface layer proteins, penicillin-binding proteins, and monomeric and multimeric proteins described herein, including those that can be produced according to the method as described herein, and genes encoding them.

The phrase "PHM" may be used equivalently to the phrase "a protein having an activity of PHM" so long as the PHM can be the protein having the activity of PHM. The same explanation is applied to the phrase "PHMcc" and the variants of PHM and PHMcc. As explained hereinafter, a PHM or a part thereof such as, for example, PHMcc is an example of the protein having an activity of PHM. Furthermore, the phrase "a protein having an activity of PHM" can be equivalent to the phrase "a protein having PHM activity" that is abbreviated from the phrase "a protein having peptidylglycine alpha-hydroxylating monooxygenase activity".

The phrase "a peptidylglycine alpha-hydroxylating monooxygenase activity" in reference to a protein having such activity, can mean an activity wherein the protein causes catalysis of the reaction of hydroxylating at the alpha-carbon of the C-terminal glycine residue of a precursor peptide (so-called peptidylglycine) to generate peptidyl alpha-hydroxyglycine (EC 1.14.17.3; Kolhekar A.S. *et al.*, 1997). Specifically, the phrase "a peptidylglycine alpha-hydroxylating monooxygenase activity" can mean the

activity of PHM domain in PAM of catalyzing the reaction of hydroxylating at the alpha-carbon of the C-terminal glycine residue of peptidylglycine to generate peptidyl alpha-hydroxyglycine. For example, the peptidylglycine alpha-hydroxylating monooxygenase activity can mean the activity of the protein having the amino acid sequence shown in SEQ ID NO: 2, 4, 6 or 8 and homologues thereof capable of catalyzing the reaction of hydroxylating at the alpha-carbon of the C-terminal glycine residue of peptidylglycine to generate peptidyl alpha-hydroxyglycine.

The peptidylglycine alpha-hydroxylating monooxygenase activity of the protein can be determined by evaluating the rate of oxygen consumption measured using a Clarke-type oxygen electrode (Bauman A.T. *et al.*, 2007). Alternatively, an approach based on measuring a fluorescent signal from a dansyl-labeled amidated derivative of peptidyl alpha-hydroxyglycine can also be used to determine the peptidylglycine alpha-hydroxylating monooxygenase activity *in vitro* as described in Handa S. *et al.*, 2012. Other means may be used to determine the peptidylglycine alpha-hydroxylating monooxygenase activity (see, for example, Kim K.-H. and Seong B.L., Peptide amidation: Production of peptide hormones *in vivo* and *in vitro*, *Biotechnol. Bioproc. Eng.*, 2001, 6(4):244-251). The protein concentration can be determined by the Bradford protein assay using bovine serum albumin (BSA) as a standard (Bradford M.M., *Anal. Biochem.*, 1976, 72:248-254). Specifically, the PHM concentration can be determined using fractionation of proteins in SDS-PAGE followed by transferring the proteins onto Immobilon P membrane and visualization using Coomassie dye or Western blotting analysis as described in Kolhekar A.S. *et al.*, 1997.

3. Genetic construct

Hereinafter, a "genetic construct for production of a protein" will be explained in more details.

The genetic construct is not particularly limited so long as production of a protein having an activity of PHM can be attained, and it can include a promoter sequence that functions in a coryneform bacterium, a nucleic acid encoding a signal peptide that functions in the coryneform bacterium that is

ligated downstream from the promoter sequence, and a nucleic acid encoding a protein that is ligated downstream from the nucleic acid encoding the signal peptide. The nucleic acid encoding a signal peptide may be ligated downstream from the promoter sequence so that the signal peptide is expressed under the control of the promoter. The nucleic acid encoding the protein may be ligated downstream from the nucleic acid encoding the signal peptide so that the protein is expressed as a fusion protein with the signal peptide. The genetic construct can also include a control sequence (operator, terminator, *etc.*) effective for expression of the protein gene in a coryneform bacterium at such an appropriate position that it can function. Furthermore, the genetic construct can also include an artificial sequence (affinity tag, *etc.*) effective for isolation from a coryneform bacterium and/or a culture medium and purification of a secretory protein produced by the method.

The promoter is not particularly limited so long as it functions in a coryneform bacterium, and it may be native to a coryneform bacterium, or it may be a heterologous promoter. The “promoter that functions in a coryneform bacterium” can refer to a promoter that possesses promoter activity in a coryneform bacterium. Specific examples of the heterologous promoter include, for example, promoters native to *E. coli* such as *tac* promoter, *lac* promoter, *trp* promoter, and *araBAD* promoter. Among these, potent promoters such as the *tac* promoter is a particular example, and inducible promoters such as *araBAD* promoter are also particular examples.

Examples of the promoter native to a coryneform bacterium include, for example, promoters of the genes encoding the cell surface layer proteins PS1, PS2 (also referred to as “CspB”), and SlpA (also referred to as “CspA”), and promoters of various amino acid biosynthesis system genes. Specific examples of the promoters of various amino acid biosynthesis system genes include, for example, promoters of the glutamate dehydrogenase gene of the glutamic acid biosynthesis system, the glutamine synthetase gene of the glutamine synthesis system, the aspartokinase gene of the lysine biosynthesis system, the homoserine dehydrogenase gene of the threonine biosynthesis system, the acetohydroxy acid synthetase gene of the isoleucine and valine biosynthesis system, 2-isopropylmalate synthetase gene of the

leucine biosynthesis system, the glutamate kinase gene of the proline and arginine biosynthesis system, the phosphoribosyl-ATP pyrophosphorylase gene of the histidine biosynthesis system, the deoxyarabinoheptulonate phosphate (DAHP) synthetase gene of the aromatic amino acid biosynthesis systems such as those for tryptophan, tyrosine, and phenylalanine, the phosphoribosyl pyrophosphate (PRPP) amidotransferase gene of the nucleic acid biosynthesis systems such as those for inosinic acid and guanylic acid, the inosinic acid dehydrogenase gene, and the guanylic acid synthetase gene.

As the promoter, an existing promoter that is highly active may be obtained by using various reporter genes. For example, by making the -35 and -10 regions in a promoter region closer to a consensus sequence, the activity of the promoter can be enhanced (WO0018935 A1). Examples of the method for evaluating the strength of a promoter and strong promoters are described in the paper of Goldstein M.A. *et al.* (Prokaryotic promoters in biotechnology, *Biotechnol. Annu. Rev.*, 1995, 1:105-128) and so forth. Furthermore, it is known that substitution, insertion, or deletion of several nucleotides in a spacer region between the ribosome-binding site (RBS) and the start codon, especially in a sequence immediately upstream of the start codon (5'-UTR), significantly affects stability and translation efficiency of mRNA, and these sequences can also be modified.

The signal peptide is not particularly limited so long that the chosen signal peptide functions in a coryneform bacterium, and it may be native to a coryneform bacterium, or it may be a heterologous signal peptide. The "signal peptide that functions in a coryneform bacterium" can refer to a peptide that, when it is ligated to the N-terminus of an objective protein, allows the coryneform bacterium to secrete the protein. The signal peptide can be a signal peptide of a secretory protein of the host coryneform bacterium, or a signal peptide of a cell surface layer protein of the coryneform bacterium. Examples of the cell surface layer protein of coryneform bacteria include PS1 and PS2 (CspB) native to *C. glutamicum* (Japanese Patent Laid-open (Kohyo) No. 6-502548), and SlpA (CspA) native to *C. ammoniagenes* (*C. stationis*) (Japanese Patent Laid-open (Kokai) No. 10-108675). The amino acid sequence of the signal peptide of PS1 is shown in SEQ ID NO: 62, the amino

acid sequence of the signal peptide of PS2 (CspB) is shown in SEQ ID NO: 64, and the amino acid sequence of the signal peptide of SlpA (CspA) is shown in SEQ ID NO: 66. Moreover, U.S. Patent No. 4,965,197 describes that there are signal peptides for DNases native to coryneform bacteria, and such signal peptides can also be used for the present invention. A signal peptide for the Sec-mediated protein secretory system (so-called Sec-pathway) or the Tat-mediated protein secretory system (so-called Tat-pathway) can be used so long as a signal peptide that functions in a coryneform bacterium is chosen (Natale P. *et al.*, Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-distinct translocases and mechanisms, *Biochim. Biophys. Acta.*, 2008, 1778(9):1735-1756). Exemplarily, the Sec-dependent CspA and PorB signal peptides can be used (Matsuda Y. *et al.*, Double mutation of cell wall proteins CspB and PBP1a increases secretion of the antibody Fab fragment from *Corynebacterium glutamicum*, *Microb. Cell Fact.*, 2014, 13(1):56; Yim S.S. *et al.*, High-level secretory production of recombinant single-chain variable fragment (scFv) in *Corynebacterium glutamicum*, *Appl. Microbiol. Biotechnol.*, 2014, 98:273-284).

Although signal peptides tend to have common sequences among biological species, a signal peptide that exhibits a secretory function in a certain biological species does not necessarily exhibit a secretory function in another biological species. Therefore, when a heterologous signal peptide is used, a signal peptide that functions in a coryneform bacterium may be appropriately chosen. Whether a certain signal peptide functions in a coryneform bacterium can be confirmed by, for example, expressing the objective protein as a fusion protein with that signal peptide, and confirming whether the protein is secreted or not.

The signal peptide may include a part of the N-terminal amino acid sequence of the secretory protein from which the signal peptide is derived. The signal sequence is generally cleaved by a signal peptidase, when the translation product is secreted out of the cell. As a gene encoding a signal peptide, although a native or naturally occurring gene may be used as it is, it may be modified so that it has the optimal codons according to codon frequencies in the chosen host.

A nucleic acid encoding a protein, which is ligated downstream from the nucleic acid encoding the signal peptide, can mean a nucleic acid encoding domain CuA of a protein having an activity of PHM or a nucleic acid encoding domain CuB of a protein having an activity of PHM. Domain CuA encoded by the nucleic acid included in a genetic construct as described herein is not particularly limited so long as the domain CuA has the activity of PHM with the domain CuB. Similarly, the domain CuB encoded by the nucleic acid included in a genetic construct as described herein is not particularly limited so long as the domain CuB has the activity of PHM with the domain CuA. The domains CuA and CuB are not particularly limited so long as the domains can form a protein having an activity of PHM as described herein. That is, the domains CuA and CuB are not particularly limited so long as the domains can form a hetero-multimer protein that can be produced as described herein. That is, the domains CuA and CuB are not particularly limited so long as the domains can form a hetero-multimer protein having an activity of PHM. The domains CuA and CuB can each have any amino acid sequence, including any kind and any length so long as the domains can form a protein having the activity of PHM. Therefore, the domains CuA and CuB may derive from or be native to the same or different species, so long as the protein which is formed by the domains CuA and CuB has the activity of PHM. As the domains CuA and CuB are different in structure, the resulting PHM is an example of a hetero-multimer having two different subunits that is produced using a coryneform bacterium having an ability to produce a protein by secretory production. As nucleotide sequences of the nucleic acids encoding PHM are known, the nucleotide sequences of the nucleic acids encoding PHMcc are known also, including the nucleotide sequences of the nucleic acids encoding domains CuA and CuB of the PHM.

Hereinafter, more detailed explanations of nucleotide sequences of the domains CuA and CuB of the PHM native to human (hCuA and hCuB) and rat (rCuA and rCuB), and the amino acid sequences encoded thereby, are described.

It is known that domains CuA and CuB can be obtained as a result of protease treatment of PHM. For example, N-terminal fragment (herein, the

domain CuA) and C-terminal fragment (herein, the domain CuB) were obtained by treating rPHMcc with endoprotease Lys-C that cleaved PHM after the K219 (lysine-219) (Kolhekar A.S. *et al.*, 1997). As a result of the treatment, the domains rCuA and rCuB were obtained, and those amino acid sequences are described hereinafter.

The amino acid sequence of domain rCuA of rPHMcc has the position numbers from S42 (serine-42) to K219 (lysine-219) in the amino acid sequence of PAM, isoform 1 from *Rattus norvegicus* (UniProt Knowledgebase: locus AMD_RAT, accession number: P14925; SEQ ID NO: 10), and it is shown in SEQ ID NO: 12, which is encoded by the nucleotide sequence shown in SEQ ID NO: 11. The amino acid sequence of domain rCuB of rPHMcc has the position numbers from V220 (valine-220) to V356 (valine-356) in the amino acid sequence of PAM, isoform 1 from *Rattus norvegicus* (UniProt Knowledgebase: locus AMD_RAT, accession number: P14925; SEQ ID NO: 10), and it is shown in SEQ ID NO: 14, which is encoded by the nucleotide sequence shown in SEQ ID NO: 13.

Examples of the domain rCuA of rPHMcc also include, but are not limited to, the domain rCuA(S44-E218) (SEQ ID NO: 36), wherein "S44" represents the N-terminal amino acid residue serine in the position number 44 and "E218" represents the C-terminal amino acid residue glutamic acid in the position number 218 in the amino acid sequence of PAM, isoform X3 from *Rattus norvegicus* (NCBI Reference Sequence: XP_006245656.1; SEQ ID NO: 4). Examples of the domain rCuB of rPHMcc also include, but are not limited to, the domain rCuB(K219-V356) (SEQ ID NO: 37), wherein "K219" represents the N-terminal amino acid residue lysine in the position number 219 and "V356" represents the C-terminal amino acid residue valine in the position number 356 in the amino acid sequence of PAM, isoform X3 from *Rattus norvegicus* (NCBI Reference Sequence: XP_006245656.1; SEQ ID NO: 4).

Examples of the domain hCuA of hPHMcc can include, but are not limited to, the domains hCuA(Xa-Ya) having varying length and structure of the amino acid sequence, wherein "Xa" represents an N-terminal amino acid

residue and its position number and “Ya” represents a C-terminal amino acid residue and its position number in the amino acid sequence of PAM, isoform 1 from *Homo sapiens* (UniProt Knowledgebase: locus AMD_HUMAN, accession number: P19021; SEQ ID NO: 2), and these can be, for example, hCuA(Xa-Ya), wherein Xa is S39, N40, E41, C42, V49, or D53, and Ya is E213, I209, S204, L201, or L195. Moreover, an N-terminal amino acid residue (Xa) may be a substitution of the amino acid residue native to or naturally present in the PAM, isoform 1 having SEQ ID NO: 2 to another amino acid residue, such as, for example, the substitution N40K or E41K, wherein the native or naturally present amino acid residue (N, asparagine, position number 40; or E, glutamic acid, position number 41) is replaced with lysine (K). The examples of the domain hCuA of hPHMcc include hCuA(S39-E213) (SEQ ID NO: 16), hCuA(S39-I209) (SEQ ID NO: 18), hCuA(N40-E213) (SEQ ID NO: 26), hCuA(N40K-E213) (SEQ ID NO: 28), hCuA(E41-E213) (SEQ ID NO: 30), hCuA(E41K-E213) (SEQ ID NO: 32), hCuA(C42-E213) (SEQ ID NO: 34), hCuA(V49-E213) (SEQ ID NO: 36), hCuA(D53-E213) (SEQ ID NO: 38). An example of the domain hCuA of hPHMcc also can include the domain hCuA(E41-I209) (SEQ ID NO: 109), wherein “E41” represents the N-terminal amino acid residue glutamic acid in the position number 41 and “I209” represents the C-terminal amino acid residue isoleucine in the position number 209 in the amino acid sequence of PAM, isoform 1 from *Homo sapiens* (UniProt Knowledgebase: locus AMD_HUMAN, accession number: P19021; SEQ ID NO: 2).

Examples of the domain hCuB of hPHMcc can include, but are not limited to, the domains hCuB(Xb-Yb) having varying length and structure of the amino acid sequence, wherein “Xb” represents an N-terminal amino acid residue and its position number and “Yb” represents a C-terminal amino acid residue and its position number in the amino acid sequence of PAM, isoform 1 from *Homo sapiens* (UniProt Knowledgebase: locus AMD_HUMAN, accession number: P19021; SEQ ID NO: 2), and these can be, for example, hCuB(Xb-Yb), wherein Xb is N217, and Yb is V351, M337, T340, N346, S353. Moreover, an N-terminal amino acid residue (Xb) may be a substitution of the amino acid residue native to or naturally present in the PAM, isoform 1

having SEQ ID NO: 2 to another amino acid residue, such as, for example, the substitution V215K, V216K or N217K, wherein the native to or naturally present amino acid residue (V, valine, position number 215 or 216; or N, asparagine, position number 217) is replaced with lysine (K). The examples of the domain hCuB of hPHMcc include hCuB(K214-V351) (SEQ ID NO: 40), hCuB(K214-S353) (SEQ ID NO: 42), hCuB(V216K-V351) (SEQ ID NO: 46), hCuB(N217K-V351) (SEQ ID NO: 50), hCuB(K214-N346) (SEQ ID NO: 56).

An amino acid sequence of the domain CuA of the PHMcc native to human and rat as described above can be chosen and used in the method as described herein so long as the domain CuA having the chosen amino acid sequence has an activity of PHM with the domain CuB of the PHMcc. Similarly, an amino acid sequence of the domain CuB of the PHMcc native to human and rat as described above can be chosen and used in the method as described herein so long as the domain CuB having the chosen amino acid sequence has an activity of PHM with the domain CuA of the PHMcc.

4. Culturing the bacterium

A method for production of a protein having an activity of PHM using a coryneform bacterium having an ability to produce a protein by secretory production is described, which method includes the steps of culturing the bacterium in a culture medium and collecting the protein from the bacterium and/or the medium. That is, by culturing the bacterium as described above, a large amount of the protein having an activity of PHM can be obtained.

The bacterium as described herein can be cultured according to a known methods and conditions. For example, the bacterium can be cultured in a typical medium containing a carbon source, a nitrogen source, and inorganic ions. In order to obtain still higher proliferation, organic micronutrients such as vitamins and amino acids can also be added as required.

As the carbon source, carbohydrates such as glucose and sucrose, organic acids such as acetic acid, alcohols, and others can be used. As the nitrogen source, ammonia gas, aqueous ammonia, ammonium salts, and

others can be used. As the inorganic ions, calcium ions, magnesium ions, phosphate ions, potassium ions, iron ions, and so forth can be appropriately used as required. The culture can be performed within appropriate ranges of pH 5.0 to 8.5 and 15 to 37°C under aerobic conditions for 1 to 7 days. Furthermore, the culture conditions for L-amino acid production by coryneform bacteria and other conditions described for the methods for producing a protein using a signal peptide of the Sec-mediated protein secretion system (so-called Sec type) or the Tat-mediated protein secretion system (so-called Tat type) can be used (WO0123591 A1 and WO2005103278 A1). Furthermore, when an inducible promoter is used for expression of the protein, a promoter-inducing agent can be added to the medium. By culturing the bacterium as described herein under such conditions, a large amount of the objective protein can be produced in cells and efficiently secreted out of the cells.

The protein secreted in the medium can be separated and purified from the medium after the culturing by a method well-known in the art. For example, after the cells are removed by centrifugation or the like, the protein can be separated and purified by a known appropriate method such as salting out, ethanol precipitation, ultrafiltration, gel filtration chromatography, ion exchange column chromatography, affinity chromatography, medium or high pressure liquid chromatography, reverse phase chromatography, and hydrophobic chromatography, or a combination of these. Furthermore, in some instances, the culture or culture supernatant may be used as it is. The protein secreted into the cell surface layer can also be separated and purified in the same manner as when the protein is secreted in the medium, after solubilizing it by a well-known method such as raising the salt concentration and use of a surfactant. Furthermore, in some instances, the protein secreted in the cell surface layer may be used as, for example, an immobilized enzyme, without solubilizing it.

Production of the objective protein can be confirmed by performing SDS-PAGE with a sample of the culture supernatant and/or a fraction containing the cell surface layer, and confirming the molecular weights of the separated protein bands. Production of the objective protein can also be

confirmed by performing Western blotting analysis using antibodies for the culture supernatant and/or a fraction containing the cell surface layer (Molecular Cloning, Cold spring Harbor Laboratory Press, Cold Spring Harbor (USA), 2001). Production of the objective protein can also be confirmed by determining the N-terminal amino acid sequences of the constituents of the obtained objective protein using a protein sequencer (see, for example, Kolhekar A.S. *et al.*, 1997). Furthermore, production of the objective protein can also be confirmed by determining the mass of the constituents of the obtained objective protein using a mass spectrometer. Furthermore, when the objective protein is an enzyme or a protein having a certain measurable physiological activity, production of the objective protein can be confirmed by measuring the enzymatic activity or the physiological activity of the objective protein in the culture supernatant and/or a fraction containing the cell surface layer as a sample.

<2> Method for C-terminal amidation of a precursor peptide

A method for C-terminal amidation of a precursor peptide to produce a target peptide is also described herein. The method includes a step of hydroxylating at the alpha-carbon of the C-terminal glycine residue of the precursor peptide in the presence of a protein having an activity of PHM produced as described herein using a coryneform bacterium having an ability to produce a protein by secretory production as explained above, wherein the method for production of the protein includes the steps of culturing the coryneform bacterium, which has been modified so that nucleic acids encoding domains CuA and CuB of the PHM are co-expressed separately in the bacterium (FIG.2). The method for C-terminal amidation of a precursor peptide includes a further step of dealkylating the intermediate peptidyl-alpha-hydroxyglycine, which is produced at the preceding step of hydroxylating the C-terminal glycine residue, in the presence of peptidyl-alpha-hydroxyglycine-alpha-amidating lyase (PAL) to produce a target peptide that is amidated at the C-terminus and glyoxylate (Kolhekar A.S. *et al.*, 1997).

The protein having an activity of PHM used in the method for C-terminal amidation of a precursor peptide is not particularly limited, and a

protein produced by the method for production of a protein having an activity of PHM as described above can be used. As the coryneform bacterium is able to produce a protein by secretory production, the protein having an activity of PHM can be used in the method for C-terminal amidation of a precursor peptide in such a way that the protein is not separated and purified from the bacterium and/or medium in which the bacterium is cultured. That is, a culture medium, in which the bacterium is cultured and the protein having an activity of PHM is present, can be used. Also, a culture medium can be used that contains disrupted cells of the bacterium (so-called crude cells lysate) so that the protein having an activity of PHM secreted in the cell surface layer can be used. Methods of cells disruption are well-known in the art, and include, for example, mechanical disruption, liquid homogenization (including the French press), high frequency sound waves (so-called ultrasonic lysis), freeze-thaw cycles, manual grinding, and so forth. It is also acceptable that the protein having an activity of PHM is separated and purified from the bacterium and/or medium using the techniques described above such that the pure protein can be used in the method for C-terminal amidation. Cells can be disrupted before separation and purification of the protein having an activity of PHM so that the amount of the purified protein is higher as compared with that of the protein separated and purified exclusively from the medium.

As a PAL, any protein (such as enzyme) can be used so long as the protein can dealkylate peptidyl-alpha-hydroxyglycine to produce a peptide that is amidated at the C-terminus (so-called peptide amide, or "target peptide") and glyoxylate. Specifically, those proteins can be used as the PAL that are classified as peptidylamidoglycolate lyases (synonym: alpha-hydroxyglycine amidating dealkylases) according to the Enzyme Commission number (EC) of 4.3.2.5. The non-limiting examples of PAL include the PAL from human (UniProt Knowledgebase: locus AMD_HUMAN, accession number: P19021), rat (UniProt Knowledgebase: locus AMD_RAT, accession number: P14925), mouse (UniProt Knowledgebase: locus AMD_MOUSE, accession number: P97467), frog *Xenopus laevis* (UniProt Knowledgebase: locus AMDB_XENLA, accession number: P12890), fly *Drosophila*

melanogaster (UniProt Knowledgebase: locus PAL1_DROME, accession number: Q9V5E1), and so forth, that can be the PAL subunit of PAM native to the same species.

Moreover, the peptidyl-alpha-hydroxyglycine produced in the presence of a protein having an activity of PHM can be dealkylated without using the PAL. This can be attained using also chemical methods such as, for example, the treatment of peptidyl-alpha-hydroxyglycine with alkaline solution upon heating. The conditions of dealkylation can be, but are not limited to, those that are described in the forthcoming Examples, and these include treating peptidyl-alpha-hydroxyglycine with aqueous solution of alkali such as, for example sodium hydroxide or potassium hydroxide at pH above neutral, for example, pH 9 with heating at 60°C for 30 min.

Any "precursor peptide" can be used in the method for C-terminal amidation so long as the precursor peptide can be hydroxylated at the alpha-carbon of the C-terminal glycine residue of the peptide in the presence of the protein having an activity of PHM produced as described above and dealkylated in the presence of PAL or without using PAL. Specifically, the precursor peptide can have the structure shown as X-Gly, wherein "X" denotes a peptide. Thus, the precursor peptide can also be referred to as "peptidylglycine". Examples of precursor peptides can include, but are not limited to, those peptides that have the structure shown as X-Gly and from which "target peptides" can be obtained using the method for C-terminal amidation of a precursor peptide as described herein. Specifically, examples of precursor peptides can include the peptides from which target peptides having alpha-amidated C-terminus can be obtained such as, for example, neuropeptides, cytokines and hormones, specific and non-limiting examples of which include Allatostatin, Amylin, alpha-Melanocyte-stimulating hormone (α -MSH), Arginine-vasopressin, Neurokinin A, Calcitonin, Bombesin, Conotoxin M1, Corticotropin releasing factor (CRF), Dermorphin, Gastrin releasing peptide, Thyroliberin (THR), Lem-KI, Luteinizing hormone-releasing hormone (LHRH), Leucopyrokinin, Gastrin I, Pigment dispersing hormone, Melanocyte-inhibiting factor (MIF-1), Melittin, Neuropeptide Y (NPY), Neuromedin B, Oxytocin, Sarcotoxin 1A, Substance P (SP), Vasoactive

intestinal peptide (VIP), and so forth, and a peptide having the amino acid sequence shown in SEQ ID NO: 68 that corresponds to Exenatide precursor having additional glycine residue (Gly, G) at the C-terminus.

So long as a “precursor peptide” that can be used in the method for C-terminal amidation is chosen, the derivative thereof amidated at the C-terminus can easily be determined. Therefore, the non-limiting examples of the “target peptide” can include those that are described above, including Exenatide (also referred to as “bydureon”, “Ex4 peptide”, “exendin-4”).

The C-terminal amidation of a precursor peptide can be attained under conditions that are suitable for performing the reactions of hydroxylating at the alpha-carbon of the C-terminal glycine residue of the precursor peptide in the presence of a protein having an activity of PHM and dealkylating the intermediate peptidyl-alpha-hydroxyglycine in the presence of PAL or without using PAL such that the target peptide that is amidated at the C-terminus and glyoxylate can be produced, and such conditions are well-known in the art. Specifically, the conditions described in the forthcoming Examples or elsewhere (see, for example, US2006292672 A1; Kolhekar A.S. *et al.*, 1997) may be used. Cofactors such as copper ions (Cu^{2+}), divalent metal ions (for example, Zn^{2+} and Ca^{2+}), ascorbate and oxygen, and other additives can be added into or induced in the reaction medium to attain C-terminal amidation. The target peptide can be isolated from reaction medium by a known appropriate method such as, for example, ion exchange column chromatography, affinity chromatography, medium or high pressure liquid chromatography, reverse phase chromatography, and hydrophobic chromatography, or a combination of these.

Examples

The present invention is more precisely explained below with reference to the following non-limiting Examples.

Example 1. Production of a protein having PHM activity using *C. glutamicum*.

1.1. Design and chemical synthesis of DNA fragments.

Three sets of DNA fragments were constructed that are denoted as F1, FA, and FB in Table 2. These DNA fragments have the structure R1-Pr_{cspB}-pre1-(secreted polypeptide)-R2, where Pr_{cspB} denotes a promoter of the *cspB* gene and pre1 denotes a signal peptide of CspB protein encoded by the *cgR_2373* gene native to *C. glutamicum* R. Referring to the secreted polypeptide component in the above structure in Table 2, pro11, pro12, pro13, and pro14 denote artificial N-terminal amino acid sequences, which can also be called artificial N-terminal propeptides, so-called N-terminal peptides; hPHMcc and rPHMcc denote the catalytic core of PHM native to *Homo sapiens* (h) or *Rattus norvegicus* (r), respectively; hCuA(Xa-Ya) and rCuA(Xa-Ya) denote a polypeptide containing the domain CuA of PHMcc flanked by amino acid residues Xa and Ya (shown in parenthesis in the 'Secreted polypeptide' column in Table 2), and native to *Homo sapiens* (h) or *Rattus norvegicus* (r), respectively; and hCuB(Xb-Yb) and rCuB(Xb-Yb) denote a polypeptide containing domain CuB of PHMcc flanked by amino acid residues Xb and Yb (shown in the parenthesis in the 'Secreted polypeptide' column in Table 2), and native to *Homo sapiens* (h) or *Rattus norvegicus* (r), respectively. All DNA fragments have flanking restriction sites R1 (*KpnI*) and R2 (*XbaI* or *BamHI*) for cloning. The nucleotide sequence of the promoter of *cspB* gene (Pr_{cspB}) is shown in SEQ ID NO: 99. The amino acid sequence of the signal peptide pre1, also known as PS2, of CspB protein is shown in SEQ ID NO: 64, and it is encoded by the nucleotide sequence shown in SEQ ID NO: 63. The amino acid sequence of the signal peptide pre1 corresponds to the 30 amino acid residues of the CspB protein when counting from the N-terminus of the protein, native to *C. glutamicum* R (NCBI Reference Sequence: NC_009342.1; GI: 145294042). The CspB protein is encoded by the *cgR_2373* gene (NCBI Reference Sequence: NC_009342.1; gene ID 4992619; locus_tag='cgR_2373'; nucleotides position from 2608342 to 2609838, complement). The constructed DNA fragments encode the following peptides:

- 1) artificial N-terminal propeptides pro11 (SEQ ID NO: 101), pro12 (SEQ ID NO: 103), pro13 (SEQ ID NO: 105), and pro14 (SEQ ID NO: 107);

2) catalytic core of the PHM native to *Homo sapiens* (hPHMcc; UniProt Knowledgebase: locus AMD_HUMAN, accession P19021) hPHMcc(S39-V351) (SEQ ID NO: 6);

3) catalytic core of the PHM native to *Rattus norvegicus* (rPHMcc; NCBI Reference Sequence: XP_006245656.1) rPHMcc(S44-V356) (SEQ ID NO: 8);

4) domain hCuA of the hPHMcc such as hCuA(S39-E213) (SEQ ID NO: 16), hCuA(S39-I209) (SEQ ID NO: 18), hCuA(S39-S204) (SEQ ID NO: 20), hCuA(S39-L201) (SEQ ID NO: 22), hCuA(S39-L195) (SEQ ID NO: 24), hCuA(N40-E213) (SEQ ID NO: 26), hCuA(N40K-E213) (SEQ ID NO: 28), hCuA(E41-E213) (SEQ ID NO: 30), hCuA(E41K-E213) (SEQ ID NO: 32), hCuA(C42-E213) (SEQ ID NO: 34), hCuA(V49-E213) (SEQ ID NO: 36), and hCuA(D53-E213) (SEQ ID NO: 38);

5) domain hCuB of the hPHMcc such as hCuB(K214-V351) (SEQ ID NO: 40), hCuB(K214-S353) (SEQ ID NO: 42), hCuB(V215K-V351) (SEQ ID NO: 44), hCuB(V216K-V351) (SEQ ID NO: 46), hCuB(N217-V351) (SEQ ID NO: 48), hCuB(N217K-V351) (SEQ ID NO: 50), hCuB(K214-M337) (SEQ ID NO: 52), hCuB(K214-T340) (SEQ ID NO: 54), and hCuB(K214-N346) (SEQ ID NO: 56);

6) domain rCuA of the rPHMcc such as rCuA(S44-E218) (SEQ ID NO: 36); and

7) domain rCuB of the rPHMcc such as rCuB(K219-V356) (SEQ ID NO: 37).

Nucleotide sequences of the DNA-fragments were adopted for the expression in *C. glutamicum* using Gene Designer program© (2005-2011, DNA 2.0 Inc.). The DNA-fragments were synthesized chemically (ATG Service Gene; Russian Federation, St. Petersburg, www.service-gene.spb.ru). Structure of the constructed DNA-fragments is shown in FIG.3. Structure of the *C. glutamicum*/*E. coli* shuttle vector pPK4 (US6,090,597 A) is shown in FIG.4.

1.2. Construction of pPK4-F1 and pPK4-FA-FB plasmids.

The scheme for construction of pPK4-F1 and pPK4-FA-FB plasmids is shown in FIG.5. The constructed plasmids are listed in Table 3. All enzymes and protocols were obtained from Fermentas (Thermo Fisher Scientific). To construct the pPK4-F1 and pPK4-FA-FB plasmids containing the F1, FA, and FB sets of DNA-fragments (Table 2), the DNA-fragments were digested using endonucleases (restrictases R1 (*KpnI*) and R2 (*XbaI* and *BamHI*)) and ligated using T4 DNA-ligase with vector pPK4 digested with the same restrictases. Thus, the pPK4/R1-R2 vector was obtained. *Escherichia coli* JM109 strain (Promega, cat. No. P9751) was transformed by the mixture obtained after ligation. The obtained pPK4-F1 and pPK4-FA-FB plasmids were screened among plasmids isolated from 5-20 arbitrary chosen kanamycin-resistant (Kn^R) colonies. The structure of the plasmids was verified using restriction analysis and sequencing. As a result, the first set of plasmids pPK4-F1 was constructed for the expression of:

- 1) hPHMcc(S39-V351) native to *Homo sapiens* (Table 3, plasmid No.4);
- 2) rPHMcc(S44-V356) native to *Rattus norvegicus* (Table 3, plasmid No.5);
- 3) domain CuA hCuA(S39-E213) (Table 3, plasmid No.2); and
- 4) domain CuB hCuB(K214-V351) (Table 3, plasmid No.3).

As another result, the second set of plasmids pPK4-FA-FB was constructed for the expression of nucleic acids encoding polypeptides containing domains CuA and CuB (Table 3, plasmid Nos.6-29).

1.3. Introduction of pPK4-F1 and pPK4-FA-FB plasmids into *C. glutamicum* YDK110Δpbp1a strain. Culturing the strains containing pPK4-F1 and pPK4-FA-FB plasmids.

The pPK4-F1 and pPK4-FA-FB plasmids were each introduced into *C. glutamicum* YDK110Δpbp1a strain using a standard electroporation procedure (van der Rest M.E. *et al.*, A heat shock following electroporation induces highly efficient transformation of *Corynebacterium glutamicum* with xenogeneic plasmid DNA, *Appl. Microbiol. Biotechnol.*, 1999, 52:541-545).

Thus, the *C. glutamicum* YDK110Δpbp1a/pPK4-F1 and *C. glutamicum* YDK110Δpbp1a/pPK4-FA-FB strains were obtained. The *C. glutamicum* YDK110Δpbp1a strain was obtained from the parent *C. glutamicum* YDK010Δpbp1a strain (WO2013065772 A1) by restoring the wild-type *rpsL* gene (Auxiliary example). Test-tube cultivation was performed as described previously (Matsuda Y. *et al.*, 2014). The resulted culture broth was stored at -20°C.

1.4. Determination of the activity of PHM in a standard preparation of PAM.

Procedures well-known to those of ordinary skill in the art were used to determine the activity of PHM (Consalvo A.P. *et al.*, Rapid fluorimetric assay for the detection of the peptidyl alpha-amidating enzyme intermediate using high-performance liquid chromatography, *J. chromatogr.*, 1992, 607:25-29; Jones B.N. *et al.*, A fluorometric assay for peptidyl alpha-amidation activity using high-performance liquid chromatography, *Anal. biochem.*, 1988, 168:272-279). A reaction mixture with a total volume of 200 μL contained MES buffer (50 mM, pH 5.5), KCl (40 mM), KI (30 mM), CuSO₄ (5 μM), L-ascorbate (0.5 mg/mL), catalase (Sigma-Aldrich, cat. No. C3155; 5 μL), Dansyl-Tyr-Phe-Gly (Life Tein, USA; www.LifeTein.com; 35-40 mg/L), and culture broth (1-3 μL) or a standard PAM preparation (1-4 μL; Daiichi Sankyo, Japan; modified PAM from *Xenopus tropicalis*; WO 2009/005140). In order to make the standard PAM preparation, the PAM solution (9.8 mg/mL) was diluted as 1:1000 using storage buffer (100 mM Tris-HCl, pH 9.3; 0.3 M NaCl). The reaction mixture was kept at 30°C for 3 hours with periodic shaking. The reaction was stopped by adding KOH to pH 9 (11 μL of 1 M KOH per 200 μL of reaction mixture) and then kept at 60°C for 30 minutes. The obtained reaction mixture was then centrifuged (10 minutes at maximum rotation of 14 rpm). Then, 50 μL of the mixture was mixed with 450 μL of 30% (v/v) MeOH and analyzed using HPLC. The amount of Dansyl-Tyr-Phe-Gly and Dansyl-Tyr-Phe-NH₂ peptides was determined. The conditions for HPLC analysis were as follows:

Equipment: UPLC Acuity Waters (Milford, USA) / Agilent 1200 infinity series,

Column: TSK GEL DEAE-5PW 2X75,

Detection: FLR-detectors, λ_{ex} . 352 nm, λ_{em} . 525 nm,

Elution: isocratic,

Elution buffer: 100 mM sodium acetate, 44% (v/v) acetonitrile, pH 6.5, adjusted using acetic acid,

Temperature: 50 °C,

Elution rate: 1 mL/min,

Injection volume: 10 μ L or 20 μ L,

Calibration samples: 1-10 mg/L,

Time of injection: 10 min.

The concentration of added PAM and the reaction time were chosen so that linear dependence was observed between the concentration of Dansyl-Tyr-Phe-NH₂ peptide (a product of reaction) and the amount of PAM. Because hydrogen peroxide (H₂O₂) generated by PHM as a by-product can oxidize Dansyl-group and, therefore, block fluorescent signal, the cumulative amount of the substrate (Dansyl-Tyr-Phe-Gly, DNS-YFG) and product (Dansyl-Tyr-Phe-NH₂, DNS-YF-NH₂) that is measured may be erroneous and could result in overestimating the substrate consumption rate and underestimating the product synthesis rate. Therefore, the P-slope of product (p) would not be equal to the P-slope of substrate (s) (Table 4; FIG.6A, raw data analysis). Assuming that the rate of oxidation of Dansyl-labeled substrate and Dansyl-labeled product correlates linearly with the concentration of the substrate and the product, the raw experimental data were normalized by calculating the amount of substrate and product (in %) in the reaction mixtures at each sample-point. The normalization resulted in the equal absolute values (in %) for the P-slope of substrate (s) and the P-slope of the product (p) and thus allowed use of the P-slope parameter for characterization of the activity of PHM (Table 4; FIG.6B, normalized data analysis).

In the Table 4, which shows the results of determination of the activity of PHM using the standard preparation of PAM, DNS-YFG means Dansyl-Tyr-

Phe-Gly peptid, and DNS-YF-NH₂ means Dansyl-Tyr-Phe-NH₂ peptide. Normalized data were calculated as a percentage (%) of substrate (DNS-YFG) and product (DNS-YF-NH₂) that remained in the reaction mixture after incubation using the raw data. P-slope (p) refers to the increment of concentration (mg/L) or percentage (%) of the amount of product synthesized under reaction conditions per 1 μ L of the added PAM. P-slope (s) refers to the decrement of concentration (mg/L) or percentage (%) of the amount of substrate converted to the product under the same conditions. The increment and decrement were calculated using the linear regression analysis of the raw data and normalized experimental data (FIG.6). The average specific activity of PHM in the standard PAM preparation was calculated using the formula: Activity (μ mol/min/mg) = $0.234 \times (35 \times 10^{-3} \text{ g/L}) / (618 \text{ g/mol}) \times (200 \times 10^{-6} \text{ L}) / (180 \text{ min} \times 9.8 \times 10^{-6} \text{ g})$, considering the molecular weight (MW) of DNS-YFG is 618 g/mol and the average initial concentration of the substrate is 35 mg/L. The activity of PHM in the standard PAM preparation was determined as 1.5 μ mol/min/mg.

1.5. Determination of the activity of PHM in culture broth of *C. glutamicum* strains.

The data obtained from the calibration experiment (Example 1, item 1.4) showed that the standard PAM preparation having 9.8 mg/L of protein characterized with the P-slope ((s) and (p)) of 23.4 (%)/ μ L (Table 4; FIG.6B, normalized data analysis). Therefore, the activity of PHM in a sample preparation can be determined by referring to the concentration of PHM in standard PAM preparation as the standard using the formula: Cref (mg/L) = $9.8 \times [\text{P-slope}(\text{sample}) / 23.4]$ so long as the P-slope value of the sample is known. By using this approach and the conditions described in Example 1 (item 1.4), the activity of PHM was determined in the culture broths of *C. glutamicum* YDK110 Δ pbp1a/pPK4, *C. glutamicum* YDK110 Δ pbp1a/pPK4-F1, and *C. glutamicum* YDK110 Δ pbp1a/pPK4-FA-FB strains (Example 1, item 1.3). The data are shown in Table 5. As one can see from the Table 5, production of hPHMcc and rPHMcc having the peptidylglycine alpha-hydroxylating monooxygenase activity using a coryneform bacterium was

improved when the nucleic acid encoding the domain CuA of the PHM and the nucleic acid encoding the domain CuB of the PHM were co-expressed separately in the bacterium, as compared with the bacterium in which the domains CuA and CuB were not co-expressed separately (that is, as compared with the bacterium in which the nucleic acid encoding hPHMcc or rPHMcc was expressed). It follows also from the Table 5 that production of the active form of human PHM and rat PHM using the coryneform bacterium having an ability to produce a protein by secretory production and modified so that a nucleic acid encoding the domain CuA of the PHM and a nucleic acid encoding domain the CuB of the PHM are co-expressed separately in the bacterium was improved by the method as described herein.

1.6. Substrate specificity of PHM.

Substrate specificity of the PHM was confirmed in culture broth of *C. glutamicum* YDK110Δpbp1a/pPK4-FA-FB strain co-expressing pre1pro11-hCuA(S39-E213) and pre1pro11-hCuB(K214-V351) polypeptides (Example 1, item 1.3; Table 5) using artificial Dansyl-labeled peptides as substrates. The Dansyl-labeled peptides (Table 6) were obtained from Life Tein (USA, www.LifeTein.com). Standard reaction conditions were used (Example 1, item 1.3). As one can see from the Table 6, the PHM showed activity towards all of the tested precursor peptides.

Example 2. C-terminal amidation of Exenatide precursor using the PHM.

2.1. Cloning of a nucleic acid encoding Exenatide precursor.

A plasmid overexpressing a nucleic acid encoding Exenatide precursor was constructed using the general knowledge of one of ordinary skill in the art. In particular, a DNA-fragment encoding Exenatide precursor having the nucleotide sequence is shown in SEQ ID NO: 67 was synthesized (Life Technologies Japan Ltd., Japan) and used as a template for PCR-amplification using primers P1 (SEQ ID NO: 110) and P2 (SEQ ID NO: 111). As a result, the DNA-fragment <cspA-Pe03G> encoding the Exenatide precursor having glycine residue at the C-terminus (SEQ ID NO: 68; FIG.3B)

was obtained. The obtained DNA-fragment was ligated into *KpnI* and *BamHI* sites of the pPK4 vector (US6,090,597 A). Thus, the pPS-Pe03G plasmid was constructed.

2.2. Preparation of Exenatide precursor.

C. glutamicum YDK0107 strain (WO2016171224 A1) having the pPS-Pe03G plasmid was seed-cultured at 30°C for 16 hr in a tube containing 3 mL of CMDX medium (5 g of glucose, 0.4 g of magnesium sulfate heptahydrate, 10 g of trypton (Becton, Dickinson, Sparks, MD, USA), 10 g of yeast extract (Becton, Dickinson), 1 g of potassium dihydrogen phosphate, 0.01 g of iron sulfate heptahydrate, 0.01 g of manganese sulfate pentahydrate, 3 g of urea, 25 mg of kanamycin, adjusted to 1 L with water). Then, the seed culture was inoculated into a tube containing 4 mL of MM liquid media (60 g of glucose, 1 g of magnesium sulfate heptahydrate, 30 g of ammonium sulfate, 1.5 g of potassium dihydrogen phosphate, 0.01 g of iron sulfate heptahydrate, 0.0082 g of manganese sulfate pentahydrate, 0.45 mg of thiamine hydrochloride, 0.45 mg of biotin, 0.15 g of DL-methionine, 50 g of calcium carbonate, 25 mg of kanamycin, adjusted to 1 L with water and to pH 7.0), and cultured aerobically at 30°C for 72 hr. Cells were removed by centrifugation (12,000 g, 5 min). Thus, 4 mL of supernatant containing Exenatide precursor was prepared.

2.3. Purification of Exenatide precursor from culture broth.

The Exenatide precursor was purified from the culture supernatant using column chromatography (AKTA avant, GE Healthcare, Little Chalfont, UK) under the following conditions:

Column: Hi Trap Q HP 5 mL (GE Healthcare),

Column temperature: 4°C,

Mobile phase A: 20 mM Tris-HCl, pH 8.0,

Mobile phase B: 20 mM Tris-HCl, 1.0 M NaCl, pH 8.0,

Flow rate: 5.0 mL/min,

Detection: 280 nm,

Gradient: Mobile phase B, from 0% to 100% (v/v) for 10 min.

As a result, 2 mg of purified Exenatide precursor was prepared from 4 mL culture broth. The sample of the Exenatide precursor was dialyzed against 20 mM Tris-HCl, pH 8.0 and stored at 4°C.

2.4. C-terminal amidation.

The C-terminal amidation reaction was performed in two steps as follows: first, 1 mL of reaction mixture containing 0.130 mg/mL Exenatide precursor, 200 mM NaCl, 1.0 μ M CuSO₄, 0.1 mg/mL L-ascorbate, 768 U/mL catalase (Sigma-Aldrich, cat. No. C3155), PHM protein preparation (4 mL of culture broth of *C. glutamicum* YDK110 Δ pbp1a/pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(K214-V351) strain; Example 1, item 1.5), and 50 mM MES-NaOH, pH 5.5 was incubated at 30°C for 16 hr. Secondly, the reaction mixture obtained after the first step was adjusted to pH 9.0 with KOH and incubated at 60°C for 30 min.

2.5. Analysis of amidated product.

The amidated product obtained in Example 2 (item 2.4) was quantified using HPLC (GL Science, Tokyo, Japan) under the following conditions:

Column: YMC-Triart C18, 4.6x150 mm, particle diameter 5 μ m (YMC Co., Ltd.),

Column temperature: 40°C,

Mobile phase: 20 mM NaHCO₃-NaOH buffer (pH 10.3) supplemented with acetonitrile (27%, v/v),

Flow rate: 1.0 mL/min,

Detection: 220 nm.

Exenatide acetate salt (BACHEM, H-8730.10000) was used as the standard. As confirmed by the HPLC analysis, Exenatide precursor was converted into the amidated form thereof (Exenatide) with the yield of 94% (Table 7).

Auxiliary example. Construction of *C. glutamicum* YDK110 Δ pbp1a strain.

The *C. glutamicum* YDK110 Δ pbp1a strain was constructed from the parent strain *C. glutamicum* YDK010 Δ pbp1a (WO2013065772 A1) by restoring the wild-type *rpsL* gene.

The genome sequence of *C. glutamicum* ATCC 13869 and the nucleotide sequence of the *rpsL* gene encoding the 30S ribosomal protein S12 are known (GenBank accession number: AP017557, NCBI locus_tag CGBL_0105810). The K43N mutation (replacement of the amino acid residue lysine in position number 43 with an asparagine residue) in the 30S ribosomal protein S12 was determined in the *rpsL* gene of the *C. glutamicum* YDK010 Δ pbp1a strain. The K43N mutation conferred on the strain resistance to streptomycin (Sm). In order to restore this mutation to the wild-type amino acid residue, a *rpsL*-restored strain was constructed.

A DNA-fragment containing the wild-type *rpsL* gene was amplified by PCR using primers P3 (SEQ ID NO: 112) and P4 (SEQ ID NO: 113), and the chromosomal DNA of the *C. glutamicum* ATCC 13869 strain as a template. Pyrobest DNA Polymerase (Takara Bio) was used, and the PCR reaction conditions were as those that are recommended by the manufacturer. The resulting DNA-fragment (about 2.0 kbp) was inserted into the *Sma*I site of pBS5T (WO2006057450 A1) to obtain a vector pBS5T-*rpsL*(wt) for restoring the *rpsL* gene. Then, the *C. glutamicum* YDK010 Δ pbp1a strain (WO2013065772 A1) was transformed with the constructed pBS5T-*rpsL*(wt) vector. The resulting strain was selected from the obtained transformants according to the methods described in WO2005113744 A1 and WO2006057450 A1. Thus, the *C. glutamicum* YDK110 Δ pbp1a strain harboring the restored (wild-type) *rpsL* gene was obtained.

While the invention has been described in detail with reference to preferred embodiments thereof, it will be apparent to one skilled in the art that various changes can be made, and equivalents employed, without departing from the scope of the invention.

Table 1. Homology (as identity*) of PHMcc's from various species.

	horse PHMcc	pig PHMcc	human PHMcc	rat PHMcc
horse PHMcc	100	96	96	92
pig PHMcc		100	97	93
human PHMcc			100	93
rat PHMcc				100

* : in %, calculated using VectorNTI software (Invitrogen)

Table 2. Constructed DNA fragments.

DNA fragment		Structure of a DNA-fragment R1-Pr _{csfB} -pre1-(secreted polypeptide)-R2		
Set	Nucleotide sequence	Secreted polypeptide	Cloning sites	
			R1	R2
F1	SEQ ID NO: 69	pro13-hCuA(S39-E213)	<i>KpnI</i>	<i>XbaI</i>
	SEQ ID NO: 70	pro11-hCuB(K214-V351)		
	SEQ ID NO: 71	pro11-hPHMcc(S39-V351)		
	SEQ ID NO: 72	pro11-rPHMcc(S44-V356)		
FA	SEQ ID NO: 73	pro13-hCuA(S39-E213)	<i>KpnI</i>	<i>BamHI</i>
	SEQ ID NO: 74	pro12-hCuA(S39-E213)		
	SEQ ID NO: 75	pro11-hCuA(S39-E213)		
	SEQ ID NO: 76	pro11-hCuA(N40-E213)		
	SEQ ID NO: 77	pro14-hCuA(N40-E213)		
	SEQ ID NO: 78	pro11-hCuA(N40K-E213)		
	SEQ ID NO: 79	pro11-hCuA(E41-E213)		
	SEQ ID NO: 80	pro11-hCuA(E41K-E213)		
	SEQ ID NO: 81	pro11-hCuA(C42-E213)		
	SEQ ID NO: 82	pro11-hCuA(V49-E213)		
	SEQ ID NO: 83	pro11-hCuA(D53-E213)		
	SEQ ID NO: 84	pro11-hCuA(S39-I209)		
	SEQ ID NO: 85	pro11-hCuA(S39-S204)		
	SEQ ID NO: 86	pro11-hCuA(S39-L201)		
	SEQ ID NO: 87	pro11-hCuA(S39-L195)		
	SEQ ID NO: 88	pro11-rCuA(S44-E218)		
FB	SEQ ID NO: 89	pro11-hCuB (K214-V351)		
	SEQ ID NO: 90	pro11-hCuB(V215K-V351)		
	SEQ ID NO: 91	pro11-hCuB(V216K-V351)		
	SEQ ID NO: 92	pro11-hCuB(N217-V351)		
	SEQ ID NO: 93	pro11-hCuB(N217K-V351)		
	SEQ ID NO: 94	pro11-hCuB(K214-M337)		
	SEQ ID NO: 95	pro11-hCuB(K214-T340)		
	SEQ ID NO: 96	pro11-hCuB(K214-N346)		
	SEQ ID NO: 97	pro11-hCuB(K214-S353)		
	SEQ ID NO: 98	pro11-rCuB(K219-V356)		

Table 3. Constructed plasmids.

Plasmid number	Construct
Cloning vector:	
1	pPK4
Plasmid pPK4-F1:	
2	pPK4-pre1pro13-hCuA(S39-E213)
3	pPK4-pre1pro11-hCuB(K214-V351)
4	pPK4-pre1pro11-hPHMcc(S39-V351)
5	pPK4-pre1pro11-rPHMcc(S44-V356)
Plasmid pPK4-FA-FB:	
6	pPK4-pre1pro13-hCuA(S39-E213)-pre1pro11-hCuB(K214-V351)
7	pPK4-pre1pro12-hCuA(S39-E213)-pre1pro11-hCuB(K214-V351)
8	pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(K214-V351)
9	pPK4-pre1pro11-hCuA(N40-E213)-pre1pro11-hCuB(K214-V351)
10	pPK4-pre1pro14-hCuA(N40-E213)-pre1pro11-hCuB(K214-V351)
11	pPK4-pre1pro14-hCuA(N40-E213)-pre1pro11-hCuB(K214-S353)
12	pPK4-pre1pro11-hCuA(N40K-E213)-pre1pro11-hCuB(K214-V351)
13	pPK4-pre1pro11-hCuA(E41-E213)-pre1pro11-hCuB(K214-V351)
14	pPK4-pre1pro11-hCuA(E41K-E213)-pre1pro11-hCuB(K214-V351)
15	pPK4-pre1pro11-hCuA(C42-E213)-pre1pro11-hCuB(K214-V351)
16	pPK4-pre1pro11-hCuA(V49-E213)-pre1pro11-hCuB(K214-V351)
17	pPK4-pre1pro11-hCuA(D53-E213)-pre1pro11-hCuB(K214-V351)
18	pPK4-pre1pro11-hCuA(S39-I209)-pre1pro11-hCuB(K214-V351)
19	pPK4-pre1pro11-hCuA(S39-S204)-pre1pro11-hCuB(K214-V351)
20	pPK4-pre1pro11-hCuA(S39-L201)-pre1pro11-hCuB(K214-V351)
21	pPK4-pre1pro11-hCuA(S39-L195)-pre1pro11-hCuB(K214-V351)
22	pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(V215-V351)
23	pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(V216K-V351)
24	pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(N217-V351)
25	pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(N217K-V351)
26	pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(K214-M337)
27	pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(K214-T340)
28	pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(K214-N346)
29	pPK4-pre1pro11-rCuA(S44-E218)-pre1pro11-rCuB(K219-V356)

Table 4. Results of the determination of the activity of PHM using the standard preparation of PAM.

Reaction number	Added PAM, μL	A. Raw data analysis					
		DNS-YFG, mg/L	DNS-YF-NH ₂ , mg/L	Sum, mg/L	P-slope, (mg/L)/ μL	Cref, mg/L	Activity, $\mu\text{mol/min/mg}$
1	0	35	0	35	8.2 (p) -8.6 (s)	9.8	1.5
2	1	26	8	34			
3	2	14	23	37			
4	3	8	26	34			
5	4	3	32	35			
Reaction number	Added PAM, μL	B. Normalized data analysis					
		DNS-YFG, %	DNS-YF-NH ₂ , %	Sum, %	P-slope, (%)/ μL	Cref, mg/L	Activity, $\mu\text{mol/min/mg}$
1	0	100	0	100	23.4 (p) -23.4 (s)	9.8	1.5
2	1	76	24				
3	2	38	62				
4	3	24	76				
5	4	9	91				

Table 5. Results of the determination of the activity of PHM in culture broths of *C. glutamicum* YDK110Δpbb1a/pPK4, *C. glutamicum* YDK110Δpbb1a/pPK4-F1, and *C. glutamicum* YDK110Δpbb1a/pPK4-FA-FB.

Plasmid	Activity (in Cref, μg/L)
pPK4	nd*
pPK4-F1	
pPK4-pre1pro13-hCuA(S39-E213)	nd
pPK4-pre1pro11-hCuB(K214-V351)	nd
pPK4-pre1pro11-hPHMcc(S39-V351)	nd
pPK4-pre1pro11-rPHMcc(S44-V356)	nd
pPK4-FA-FB	
pPK4-pre1pro13-hCuA(S39-E213)-pre1pro11-hCuB(K214-V351)	400
pPK4-pre1pro12-hCuA(S39-E213)-pre1pro11-hCuB(K214-V351)	900
pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(K214-V351)	1800
pPK4-pre1pro11-hCuA(N40-E213)-pre1pro11-hCuB(K214-V351)	2000
pPK4-pre1pro14-hCuA(N40-E213)-pre1pro11-hCuB(K214-V351)	2000
pPK4-pre1pro14-hCuA(N40-E213)-pre1pro11-hCuB(K214-S353)	2000
pPK4-pre1pro11-hCuA(N40K-E213)-pre1pro11-hCuB(K214-V351)	1000
pPK4-pre1pro11-hCuA(E41-E213)-pre1pro11-hCuB(K214-V351)	1400
pPK4-pre1pro11-hCuA(E41K-E213)-pre1pro11-hCuB(K214-V351)	1200
pPK4-pre1pro11-hCuA(C42-E213)-pre1pro11-hCuB(K214-V351)	300
pPK4-pre1pro11-hCuA(V49-E213)-pre1pro11-hCuB(K214-V351)	200
pPK4-pre1pro11-hCuA(D53-E213)-pre1pro11-hCuB(K214-V351)	200
pPK4-pre1pro11-hCuA(S39-I209)-pre1pro11-hCuB(K214-V351)	1700
pPK4-pre1pro11-hCuA(S39-S204)-pre1pro11-hCuB(K214-V351)	nd
pPK4-pre1pro11-hCuA(S39-L201)-pre1pro11-hCuB(K214-V351)	nd
pPK4-pre1pro11-hCuA(S39-L195)-pre1pro11-hCuB(K214-V351)	nd
pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(V215K-V351)	nd
pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(V216K-V351)	700
pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(N217-V351)	nd
pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(N217K-V351)	110
pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(K214-M337)	nd
pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(K214-T340)	nd
pPK4-pre1pro11-hCuA(S39-E213)-pre1pro11-hCuB(K214-N346)	500
pPK4-pre1pro11-rCuA(S44-E218)-pre1pro11-rCuB(K219-V356)	200

* : nd - not detected

Table 6. Substrate specificity of PHM.

Substrate	P-slope, (%)/ μ L	Cref, μ g/L	Relative activity, %
Dansyl-Tyr-Phe-Gly	4.7	2000	100
Dansyl-Tyr-Gly-Gly	2.2	937	47
Dansyl-Tyr-Val-Gly	8.6	3664	183
Dansyl-Tyr-Lys-Gly	4.1	1747	87
Dansyl-Tyr-Asn-Gly	7.1	3025	151

Table 7. Conversion of Exenatide precursor to Exenatide.

	Concentration, mg/mL
Exenatide precursor	0.130
Exenatide	0.122

CLAIMS

1. A method for production of a protein having peptidylglycine alpha-hydroxylating monooxygenase activity using a coryneform bacterium that is able to produce and secrete a protein by secretory production, wherein the method comprises:

i) culturing the coryneform bacterium in a culture medium to produce the protein having peptidylglycine alpha-hydroxylating monooxygenase activity, and

ii) collecting said protein having peptidylglycine alpha-hydroxylating monooxygenase activity from the bacterium and/or the medium;

wherein said bacterium has been modified so that a nucleic acid encoding domain CuA of a peptidylglycine alpha-hydroxylating monooxygenase and a nucleic acid encoding domain CuB of a peptidylglycine alpha-hydroxylating monooxygenase are co-expressed separately in the bacterium.

2. The method according to claim 1, wherein said protein having the peptidylglycine alpha-hydroxylating monooxygenase activity is native to a mammal.

3. The method according to claim 2, wherein said mammal is selected from the group consisting of human, horse, pig, and rat.

4. The method according to any one of claims 1 to 3, wherein said nucleic acid encoding the domain CuA is a DNA selected from the group consisting of:

(A) a DNA having the nucleotide sequence shown in SEQ ID NO: 11, 15, 17, 25, 27, 29, 31, 33, 35, 37, 57, or 108,

(B) a DNA which encodes the domain CuA using any synonymous amino acid codons according to the standard genetic code table,

(C) a DNA encoding the protein having the amino acid sequence shown in SEQ ID NO: 12, 16, 18, 26, 28, 30, 32, 34, 36, 38, 58, or 109,

(D) a DNA encoding the protein having the amino acid sequence shown in SEQ ID NO: 12, 16, 18, 26, 28, 30, 32, 34, 36, 38, 58, or 109, but which includes substitution, deletion, insertion, or addition of 1 to 20 amino acid residues, and wherein said protein has peptidylglycine alpha-hydroxylating monooxygenase activity with the domain CuB.

5. The method according to any one of claims 1 to 3, wherein said nucleic acid encoding the domain CuB is a DNA selected from the group consisting of:

(a) a DNA having the nucleotide sequence shown in SEQ ID NO: 13, 39, 41, 45, 49, 55, or 59,

(b) a DNA which encodes the domain CuB using any synonymous amino acid codons according to the standard genetic code table,

(c) a DNA encoding the protein having the amino acid sequence shown in SEQ ID NO: 14, 40, 42, 46, 50, 56, or 60,

(d) a DNA encoding the protein having the amino acid sequence shown in SEQ ID NO: 14, 40, 42, 46, 50, 56, or 60, but which includes substitution, deletion, insertion, or addition of 1 to 20 amino acid residues, and wherein said protein has the peptidylglycine alpha-hydroxylating monooxygenase activity with the domain CuA.

6. The method according to any one of claims 1 to 5, wherein said coryneform bacterium belongs to the genus *Corynebacterium* or *Brevibacterium*.

7. The method according to claim 6, wherein said coryneform bacterium is *Corynebacterium glutamicum*.

8. The method according to any one of claims 1 to 7, wherein said coryneform bacterium harbors a genetic construct for production of the protein having peptidylglycine alpha-hydroxylating monooxygenase activity, wherein said genetic construct comprises:

(i) a promoter sequence that functions in the coryneform bacterium,

(ii) a nucleic acid encoding a signal peptide that functions in the coryneform bacterium, wherein said nucleic acid encoding a signal peptide is ligated downstream from the promoter sequence, and

(iii) the nucleic acid encoding the domain CuA, wherein said nucleic acid encoding the domain CuA is ligated downstream from the nucleic acid encoding the signal peptide;

wherein said genetic construct is co-expressed with a second genetic construct comprising the nucleic acid encoding the domain CuB.

9. The method according to any one of claims 1 to 7, wherein said coryneform bacterium harbors the genetic construct for production of the protein having peptidylglycine alpha-hydroxylating monooxygenase activity, wherein said genetic construct comprises:

(i) a promoter sequence that functions in the coryneform bacterium,

(ii) a nucleic acid encoding a signal peptide that functions in the coryneform bacterium, wherein said nucleic acid encoding a signal peptide is ligated downstream from the promoter sequence, and

(iii) the nucleic acid encoding the domain CuB, wherein said nucleic acid encoding the domain CuB is ligated downstream from the nucleic acid encoding the signal peptide;

wherein said genetic construct is co-expressed with a second genetic construct comprising the nucleic acid encoding the domain CuA.

10. The method according to claim 8 or 9, wherein said signal peptide is a signal peptide for the Sec-mediated protein secretion system.

11. A method for C-terminal amidation of a precursor peptide to produce a target peptide, wherein said method comprises a step of hydroxylating at the alpha-carbon of the C-terminal glycine residue of the precursor peptide in the presence of the protein having peptidylglycine alpha-hydroxylating monooxygenase activity produced by the method of any one of claims 1 to 10.

12. The method according to claim 11, wherein said target peptide is a neuropeptide, cytokine or hormone.

13. The method according to claim 12, wherein said target peptide is exenatide.

FIG. 1.

		1		50
humanPHMcc	(1)	SNECLGTTTRPVVPIDSSDFALDIRMPGVTPKQSDTYFCMSMRI	FPVDEEAF	
ratPHMcc	(1)	SNECLGTIGPVTPLDASDFALDIRMPGVTPKESD TYFCMSMRL	FPVDEEAF	
pigPHMcc	(1)	SNECLGTTTRPVIPIDSSDFALDIRMPGVTPKQSDTYFCMSMRL	FPVDEEAF	
horsePHMcc	(1)	SNECLGTTTRPVIPIDSSDFALDIRMPGVTPKQSDTYFCMSMRL	PMDEETF	
		51		100
humanPHMcc	(51)	VIDFKPRASMDTVHHMLLFGCNMPSS	TGSIWFCDEGTCTDKANILYAWAR	
ratPHMcc	(51)	VIDFKPRASMDTVHHMLLFGCNMPSS	TGSIWFCDEGTCTDKANILYAWAR	
pigPHMcc	(51)	VIDFKPRASMDTVHHMLLFGCNMPAST	GNWFCDEGTCTDKANILYAWAR	
horsePHMcc	(51)	VIDFKPRASMDTVHHMLLFGCNMPSS	TGSIWFCDEGVCTDKANILYAWAR	
		101		150
humanPHMcc	(101)	NAPPTRLPKGVGFRVGGETGSKYFVLQVHYGDI	SAFRDNHKDCSGVSLHL	
ratPHMcc	(101)	NAPPTRLPKGVGFRVGGETGSKYFVLQVHYGDI	SAFRDNHKDCSGVSVHL	
pigPHMcc	(101)	NAPPTRLPKGVGFRVGGETGSKYFVLQVHYGDI	SAFRDNHKDCSGVSLHL	
horsePHMcc	(101)	NAPPTRLPKGVGFRVGGETGSKYFVLQVHYGDI	SAFRDNHKDCSGVSLHL	
		151		200
humanPHMcc	(151)	TRLPQPLIAGMYLMSVDTVIPAGEKVVNSDI	SCHYKNYPMHVFAIRVHT	
ratPHMcc	(151)	TRVPQPLIAGMYLMSVDTVIPPEKVVNADI	SCQYKMYPMHVFAIRVHT	
pigPHMcc	(151)	TRLPQPLIAGMYLMSVDTVIPAGEKVVNSDI	SCHYKKYPMHVFAIRVHT	
horsePHMcc	(151)	TRLPQPLIAGMYLMALD TVIPAGEKVVNSDL	SCHYKKYPMHVFAIRVHT	
		201		250
humanPHMcc	(201)	HHLGKVVSGYRVRNGQWTLIGRQSPQLPQAFYPVGH	PVDVSFGDLLAARC	
ratPHMcc	(201)	HHLGKVVSGYRVRNGQWTLIGRQNPQLPQAFYPVEH	PVDVTFGDI LAARC	
pigPHMcc	(201)	HHLGKVVSGYRVRNGQWTLIGRQSPQLPQAFYPVEH	PVDVSFGDI LAARC	
horsePHMcc	(201)	HHLGKVVSGYRVRNGQWTLIGRQSPQLPQAFYPVEH	PVDVSFGDI LAARC	
		251		300
humanPHMcc	(251)	VFTGEGRTEATHIGGTSSDEMONLYIMYYMEAKH	AVSFMTCTQNVAPDMF	
ratPHMcc	(251)	VFTGEGRTEATHIGGTSSDEMONLYIMYYMEAKY	ALSFMTCTKNVAPDMF	
pigPHMcc	(251)	VFTGEGRTEATHIGGTSSDEMONLYIMYYMEAKH	AVSFMTCTQNVAPDIF	
horsePHMcc	(251)	VFTGEGRTEATHIGGTSSDEMONLYIMYYMEAKH	AVSFMTCTQNVAPENF	
		301		313
humanPHMcc	(301)	RTIPPEANIPIPV		
ratPHMcc	(301)	RTIPAEANIPIPV		
pigPHMcc	(301)	RTIPTEANIPIPV		
horsePHMcc	(301)	RTIPPEANIPIPV		

FIG.2.

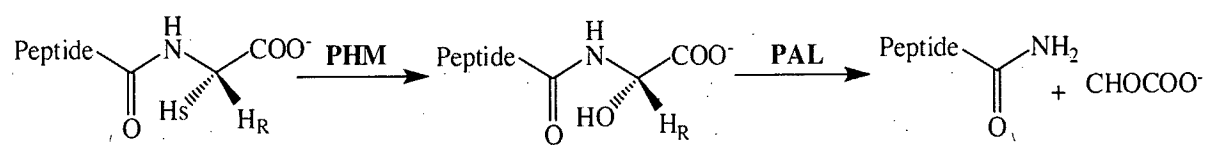


FIG.3.

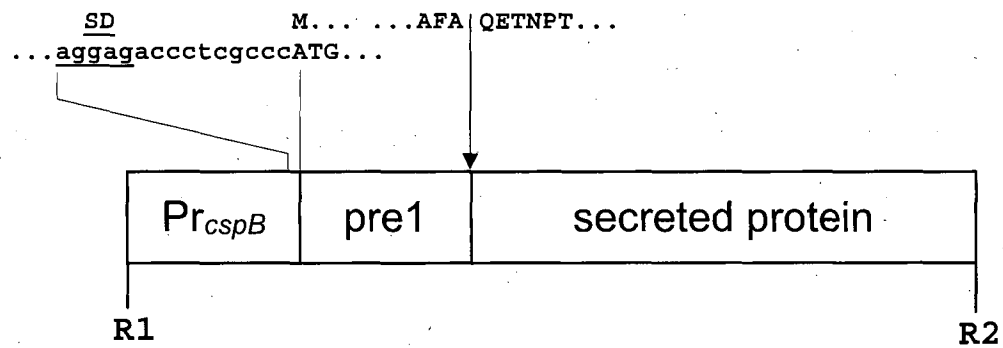
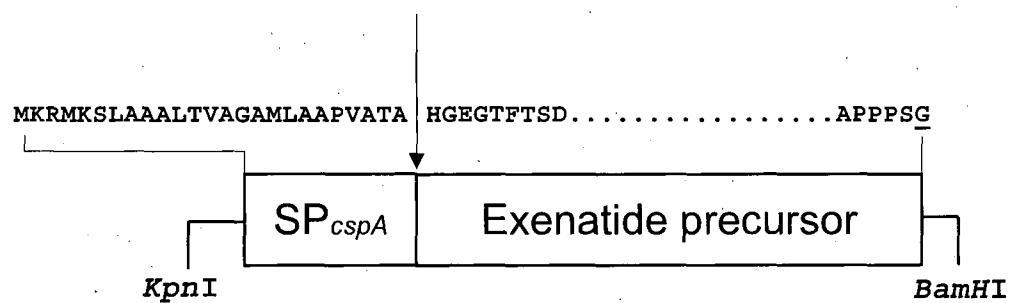
A**B**

FIG.4.

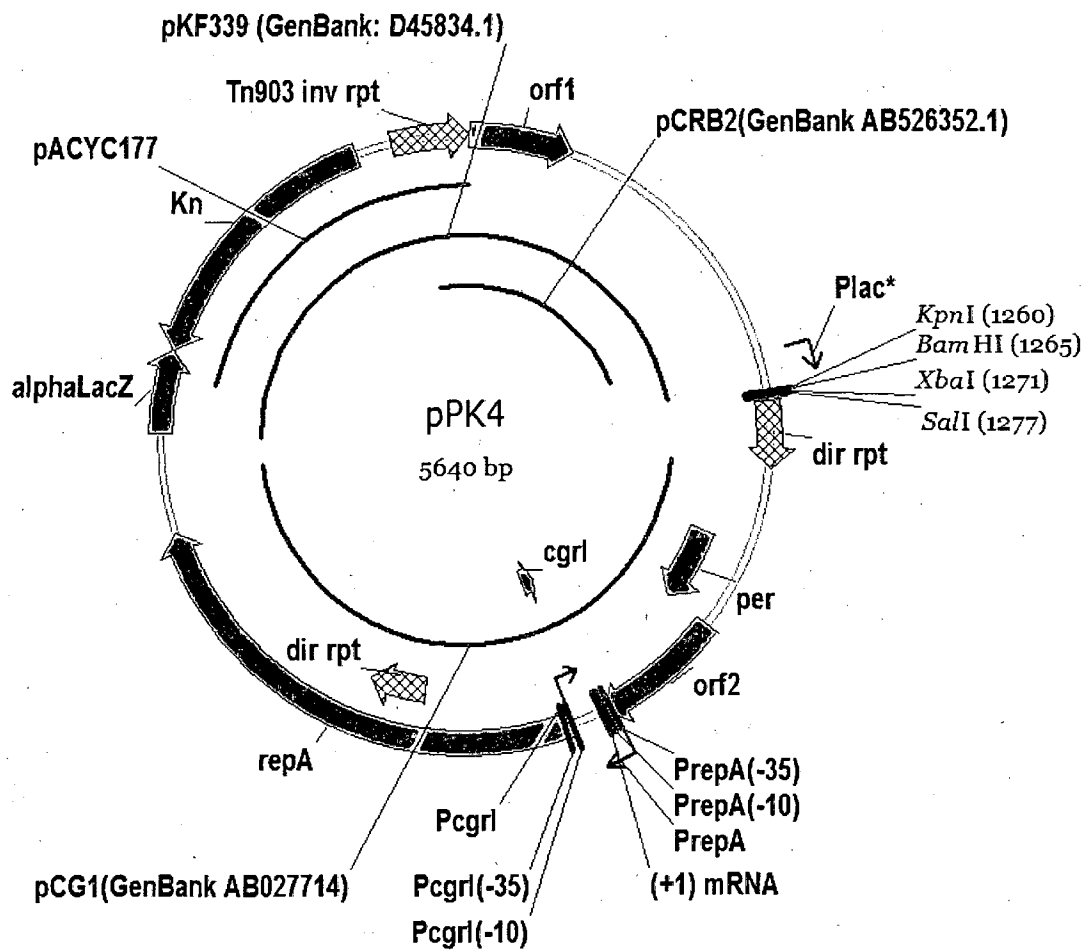


FIG.5.

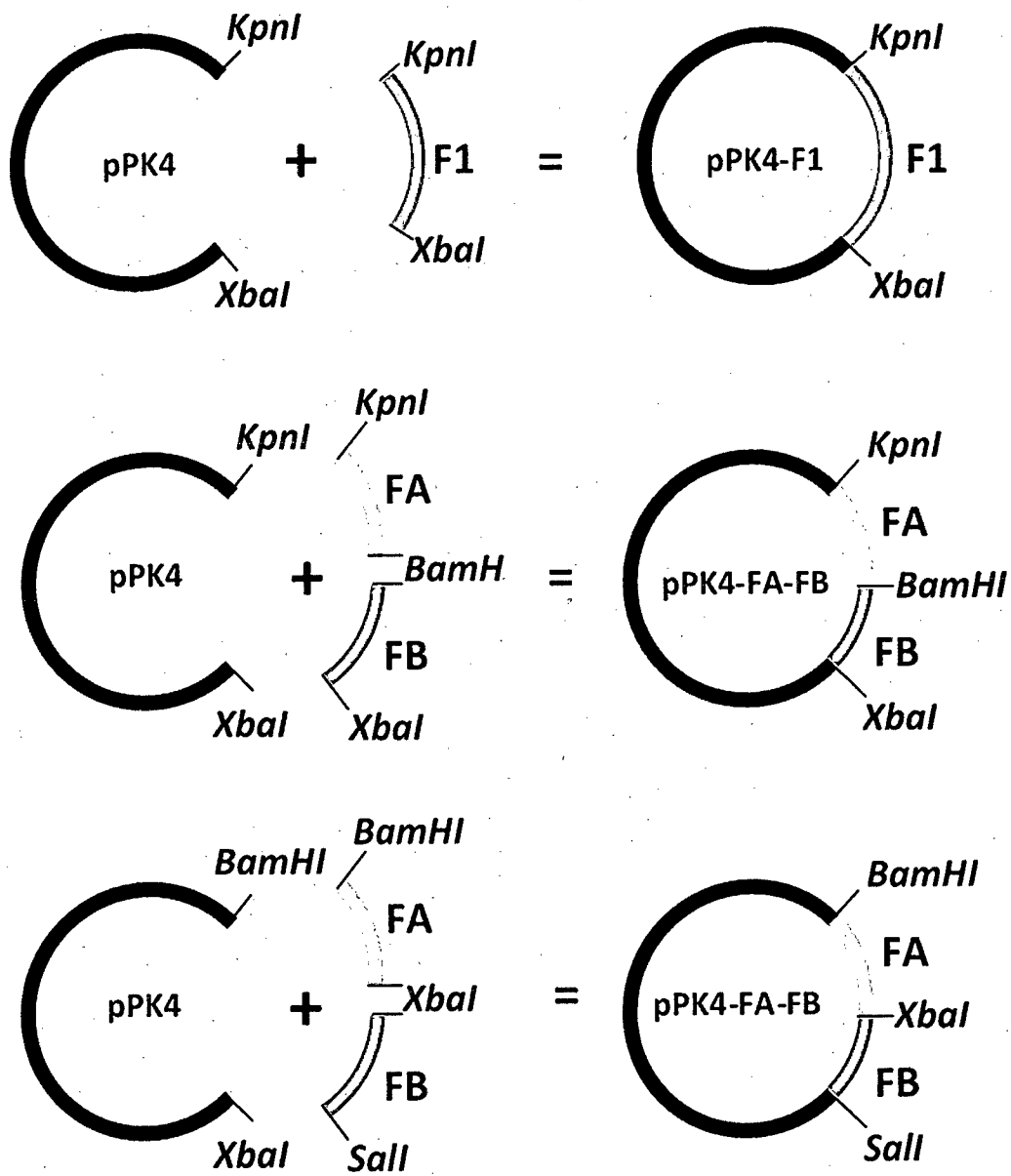
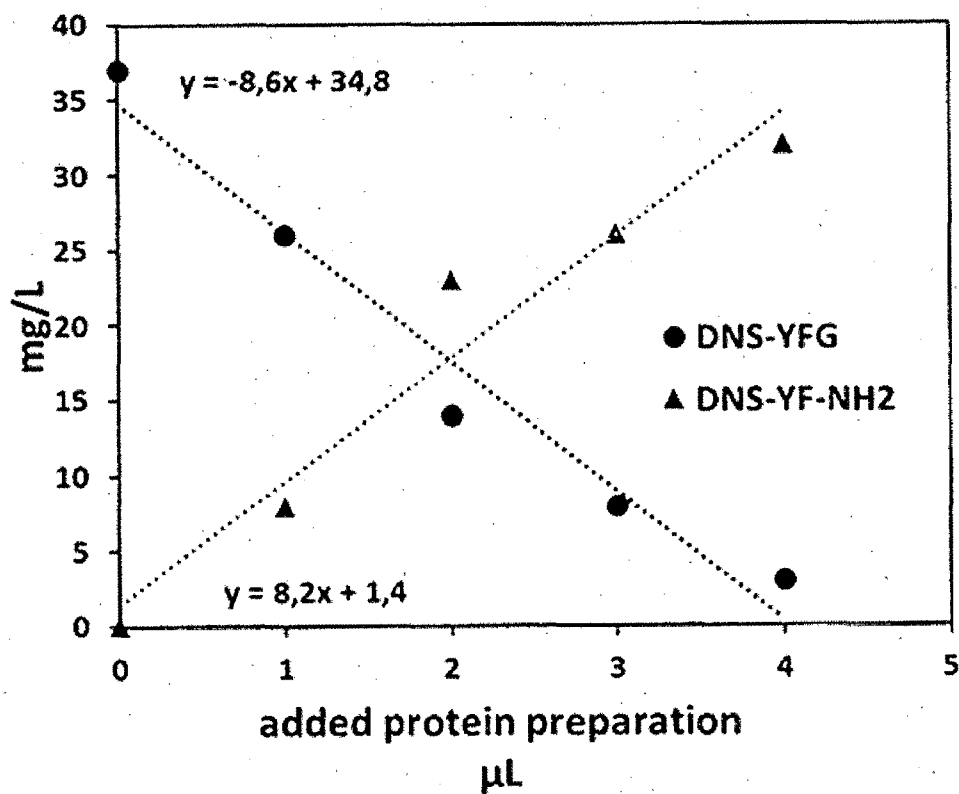
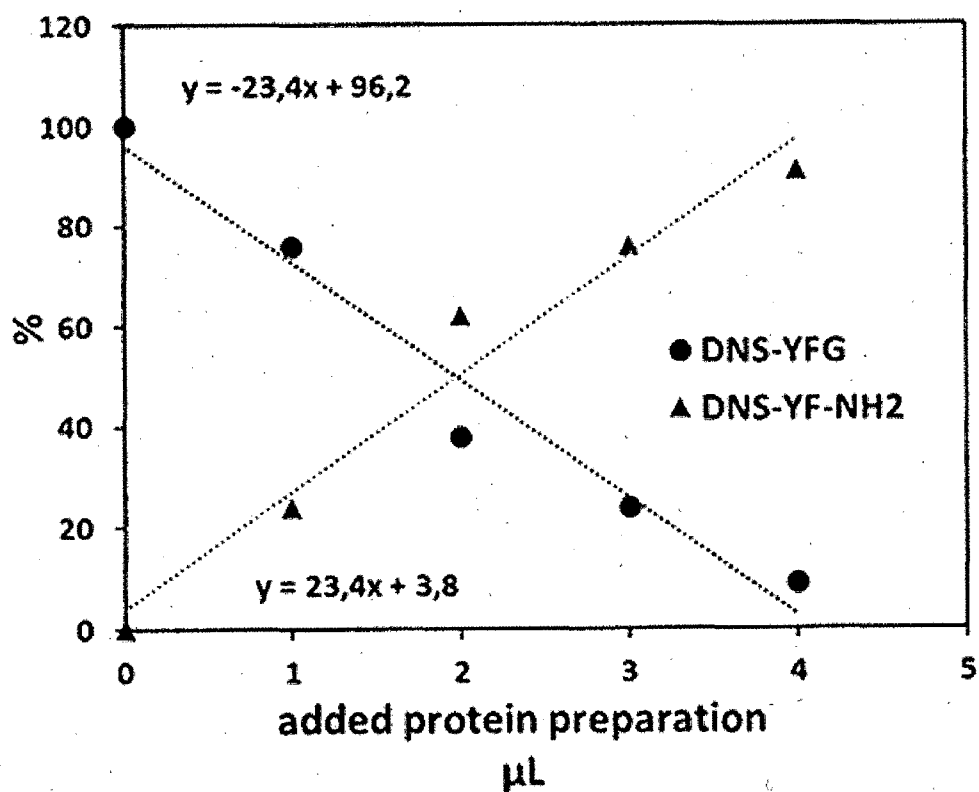


FIG.6.

A**B**

INTERNATIONAL SEARCH REPORT

International application No
PCT/JP2018/038013

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/195
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)
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, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2013/065772 A1 (AJINOMOTO KK [JP]) 10 May 2013 (2013-05-10) claims 1-13 -----	1-13
A	EP 0 448 513 A2 (JAPAT LTD [CH]) 25 September 1991 (1991-09-25) claims 1-11 -----	1-13
A	WO 2007/002532 A2 (UNIGENE LAB INC [US]; MILLER DUNCAN A [US]; MEHTA NOZER M [US]; CONSAL) 4 January 2007 (2007-01-04) claims 32-61 -----	1-13



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 December 2018

Date of mailing of the international search report

11/01/2019

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

Stoyanov, Borislav

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/JP2018/038013

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013065772 A1	10-05-2013	AU 2012333515 A1	12-06-2014
		CN 103946372 A	23-07-2014
		EP 2774982 A1	10-09-2014
		JP 6136930 B2	31-05-2017
		JP W02013065772 A1	02-04-2015
		US 2014234901 A1	21-08-2014
		WO 2013065772 A1	10-05-2013

EP 0448513 A2	25-09-1991	AU 7355991 A	03-10-1991
		CA 2038569 A1	22-09-1991
		EP 0448513 A2	25-09-1991
		IE 910934 A1	25-09-1991
		JP H04211367 A	03-08-1992
		PT 97069 A	31-10-1991

WO 2007002532 A2	04-01-2007	AU 2006261909 A1	04-01-2007
		CA 2613471 A1	04-01-2007
		CN 101287840 A	15-10-2008
		DK 1907561 T3	02-12-2013
		EP 1907561 A2	09-04-2008
		ES 2437068 T3	08-01-2014
		JP 4856177 B2	18-01-2012
		JP 5568703 B2	13-08-2014
		JP 5677918 B2	25-02-2015
		JP 2008543344 A	04-12-2008
		JP 2011139714 A	21-07-2011
		JP 2012016354 A	26-01-2012
		US 2006292672 A1	28-12-2006
		WO 2007002532 A2	04-01-2007
