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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 November 2014 (27.11.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/187960 A1

(51) International Patent Classification:
C12P 21/06 (2006.01) *C12N 15/62* (2006.01)

(21) International Application Number:
PCT/EP2014/060669

(22) International Filing Date:
23 May 2014 (23.05.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

13169173.5	24 May 2013 (24.05.2013)	EP
13171191.3	10 June 2013 (10.06.2013)	EP
61/834,100	12 June 2013 (12.06.2013)	US
13194053.8	22 November 2013 (22.11.2013)	EP
61/909,412	27 November 2013 (27.11.2013)	US

(71) Applicant: **NOVO NORDISK A/S** [DK/DK]; Novo Allé, DK-2880 Bagsværd (DK).

(72) Inventor: **SHAW, Allan Christian**; Novo Allé, DK-2880 Bagsværd (DK).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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.

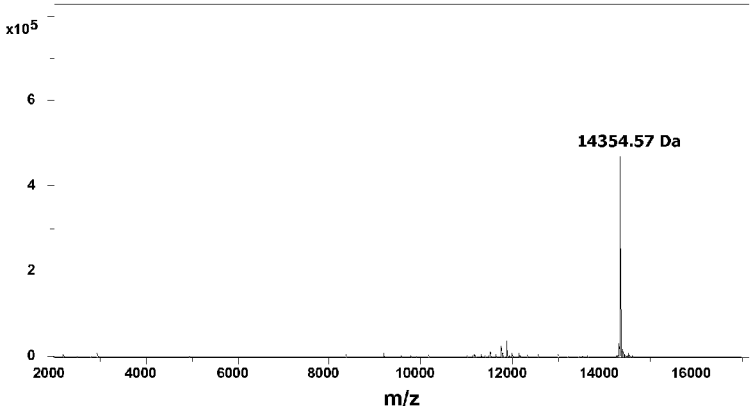
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) Title: REMOVAL OF N-TERMINAL EXTENSIONS FROM FUSION PROTEINS

Figure 1 / 7



(57) Abstract: This invention relates to an improved method for manufacturing a matured protein with its intact sequence from a fusion protein. More specifically, the present invention relates to enzymatic methods using picornaviral 3C protease and Xaa-Pro dipeptidyl aminopeptidase for the removal of an N-terminal extension from a fusion protein, and the use of these methods to produce recombinant products with the correct N-terminal sequence.

WO 2014/187960 A1

REMOVAL OF N-TERMINAL EXTENSIONS FROM FUSION PROTEINS

TECHNICAL FIELD

The present invention relates to the technical fields of protein expression and protein chemistry where a matured protein is to be released from a fusion protein.

BACKGROUND

The techniques of recombinant protein expression allow for the production of large quantities of desirable proteins which may be used for their biological activity. Such proteins are often expressed as recombinant fusion proteins in microbial host cells. The matured protein (protein of interest) is often attached to a fusion partner protein or a smaller amino acid extension in order to increase the expression level, increase the solubility, promote protein folding or to facilitate the purification and downstream processing.

Removal of the fusion partner protein from the fusion protein, to release the mature protein with native N- and C-terminus, may be pivotal for maintaining intact biological activity of the protein as well as for drug regulatory purposes.

Presently a limited number of proteases useful for removal of fusion partner proteins from fusion proteins, which leaves a native N-terminus in the released maturated target protein are available as economically sustainable enzymes for industrial use.

One such enzyme is enterokinase which, however, lacks the specificity to be generally applicable. Other such enzymes are Factor Xa, trypsin, clostripain, thrombin, TEV or rhinoviral 3C protease, all of which either lacks specificity as most proteins comprise internal secondary cleavage sites or leaves an amino acid extension in the C- or N-terminal of the mature protein.

Waugh, Protein Expr. Purif. 80:283-293 (2011) discloses an overview of enzymatic reagents for the removal of affinity tags.

US 8,137,929 B2 discloses the purification of a tagged protein comprising a target protein, a purification tag and a cleavage site for in vitro cleavage by e.g. enterokinase, Factor Xa, thrombin, TEV or HRV14 3C.

WO2008/043847 discloses a processing enzyme comprising an N-terminally attached tag derived from highly basic ribosomal proteins from thermophilic bacteria.

WO92/10576 discloses the use of fusion proteins with DPP IV cleavable extension peptide portions in medicinal preparations.

US 5427927 disclose a process for preparation of a desired protein by introducing an IgA protease recognition site in a precursor and cleaving the precursor with IgA protease. An option is mentioned for additional cleaving of a X-Pro amino terminus from the desired protein using dipeptidyl-aminopeptidase.

Xin, Protein Expr. Purif. 2002, 24, pp530-538 discloses the cloning, expression in *Escherichia coli* and application of X-prolyl dipeptidyl aminopeptidase from *Lactococcus lactis* for removal of N-terminal Pro-Pro from recombinant proteins.

In the pharmaceutical industry protein pharmaceuticals are now constituting a substantial proportion of the competitive market and efficient processes for the large scale manufacture of these protein pharmaceuticals are therefore needed. A key issue for the industrial use of fusion proteins remains the removal of the fusion protein partner from the fusion protein to liberate the intact matured protein.

Thus, there is a need for an industrial process for specifically removing a fusion partner protein without cleaving internal sites in the mature protein and without leaving any amino acid extension on the mature protein. Preferably, this process encompasses only process steps which do not cause cleavage at any internal sites in the matured protein under the applied process conditions. There is also a need for such a process which can serve this function for many different proteins at mild process conditions such that unintended chemical and physical changes to the mature protein do not occur. Hence, it is desirable that following cleavage the mature protein is directly obtained as a uniform product free from any significant amounts of mis-cleaved forms.

SUMMARY

According to a first aspect of the invention, there is provided a method for release of a correctly matured protein from a fusion protein by enzymatic treatment using a picornaviral 3C protease and a Xaa-Pro-dipeptidyl aminopeptidase (XaaProDAP).

Both picornaviral 3C protease and XaaProDAP are very specific enzymes which by the present invention are shown to exhibit complementing activities very useful for the pharmaceutical industry manufacturing protein pharmaceuticals. The combination of the two enzymes have the further advantage of XaaProDAP being easily removed from the matured protein by a simple gel-filtration process.

According to a second aspect of the invention there is provided a method for release of a matured protein from a fusion protein by enzymatic treatment using a picornaviral 3C protease and a Xaa-Pro-dipeptidyl aminopeptidase (XaaProDAP), wherein said fusion protein has the structure :

Tag sequence – Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-QGP - Protein of interest, wherein

Tag sequence is a fusion partner protein connected at its C-terminal,

Xaa1, Xaa2, Xaa3, Xaa4 and Xaa5 are independently selected from the group consisting of the naturally occurring amino acid residues, and

Protein of interest is a protein connected at its N-terminal and comprising said matured protein and optionally an N-terminal extension which can be cleaved by XaaProDAP or by picornaviral 3C protease and XaaProDAP.

In one embodiment, the XaaProDAP is able to cleave a peptide after the proline residue in a dipeptide comprising Gly-Pro or Ala-Pro.

In one embodiment, the picornaviral 3C protease natively has a Gly-Pro in the N-terminal. In another embodiment, the picornaviral 3C protease is HRV14 3C or a functional variant thereof.

In one embodiment the XaaProDAP is an E.C. 3.4.14.11 enzyme or a functional variant thereof. In another embodiment, the XaaProDAP is an enzyme having bacterial origin. In another embodiment, the XaaProDAP is an enzyme from *Lactococcus sp.* or a functional variant derived from the enzyme in *Lactococcus sp.*

According to a third aspect of the invention there is provided a fusion enzyme comprising the catalytic domains of a picornaviral 3C protease and a XaaProDAP or functional variants thereof.

In one embodiment the fusion enzyme comprises the full length of a picornaviral 3C protease and the full length XaaProDAP or functional variant thereof.

In one embodiment the fusion enzyme comprises a combination of either a catalytic domain or full-length picornaviral 3C protease with either a catalytic domain or full-length XaaProDAP or functional variants thereof.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1. Mass spectrum of intact fusion protein RL27_HRV14_PYY(3-36) control with no enzyme added. X-axis: Mass over charge ratio (m/z) in Da. Y-axis: Relative intensity.

Figure 2. Mass spectrum of Reaction 1 where RL27_HRV14_PYY(3-36) is incubated with XaaProDAP and HRV14 protease for 1 hour at 37 °C using 1:1:40 molar ratio (XaaProDAP:HRV14:fusion protein). X-axis: Mass over charge ratio (m/z) in Da. Y-axis: Relative intensity.

Figure 3. Mass spectrum of Reaction 2 where RL27_HRV14_PYY(3-36) is incubated with XaaProDAP and HRV14 protease for 3 hours at 37 °C using a 1:1:40 molar ratio. X-axis: Mass over charge ratio (m/z) in Da. Y-axis: Relative intensity.

Figure 4. Mass spectrum of Reaction 3 where RL27_HRV14_PYY(3-36) is incubated with XaaProDAP and HRV14 protease for 1 hour at 37 °C using a 1:1:20 molar ratio. X-axis: Mass over charge ratio (m/z) in Da. Y-axis: Relative intensity.

Figure 5. Mass spectrum of Reaction 4 where RL27_HRV14_PYY(3-36) is incubated with XaaProDAP and HRV14 protease for 3 hours at 37 °C using a 1:1:20 molar ratio. X-axis: Mass over charge ratio (m/z) in Da. Y-axis: Relative intensity.

Figure 6. Mass spectrum of Reaction 5 where RL27_HRV14_PYY(3-36) is incubated with XaaProDAP and HRV14 protease for 1 hour at 37 °C using a 1:1:10 molar ratio. X-axis: Mass over charge ratio (m/z) in Da. Y-axis: Relative intensity.

Figure 7. Mass spectrum of Reaction 6 where RL27_HRV14_PYY(3-36) is incubated with XaaProDAP and HRV14 protease for 3 hours at 37 °C using a 1:1:10 molar ratio. X-axis: Mass over charge ratio (m/z) in Da. Y-axis: Relative intensity.

DESCRIPTION

According to a first aspect of the invention, there is provided a method for release of a matured protein from a fusion protein by enzymatic treatment using a picornaviral 3C protease and a Xaa-Pro-dipeptidyl aminopeptidase (XaaProDAP).

The method of the invention provides a number of advantages over previously described methods for release of a matured protein from a fusion protein. For example, it has been surprisingly found that a very specific hydrolysis of the fusion protein can be obtained such that the mature protein is released with the correct native N-terminal amino acid in the absence or with a minimum level of related impurities and in high yields. The presence of any related impurities, i.e. proteins resembling the mature protein by having limited differences in chemical structure, is clearly undesirable as they are difficult and thus expensive to remove in a manufacturing process.

Certain embodiments of the invention have the further advantage of allowing release of the matured protein from the fusion protein in a single process step. Additional embodiments have the advantage of allowing release of the matured protein from the fusion protein at reactions conditions having low temperatures.

“Xaa-Pro dipeptidyl aminopeptidase” (“XaaProDAP”) as used herein is intended to mean an enzyme having dipeptidase activity specific for Xaa-Pro dipeptides, i.e. the scissile bond connecting the C-terminal of the Xaa-Pro dipeptide with the N-terminal of a peptide or

protein. XaaProDAP´s are classified according to the international union of Biochemistry and molecular Biology Enzyme (IUBMB) Enzyme Nomenclature as the enzymes EC 3.4.14.11 from the peptidase family S15 and as the enzymes EC 3.4.14.5 from the peptidase family S9B. Non-limiting examples of XaaProDAP are dipeptidyl-peptidase IV (DPP-IV) from mammals. Other non-limiting examples of XaaProDAP are Xaa-Prolyl dipeptidyl aminopeptidase from bacteria such as *Lactococcus lactis, Streptococcus thermophilus, Lactobacillus delbrueckii* and *Streptococcus suis*. Xaa-Prolyl dipeptidyl aminopeptidase from *Lactococcus lactis subsp. cremoris CNCM I-1631* has the sequence :

MRFNHFSIVDKNFDEQLAELDQLGFRWSVFWDEKKILKDFLIQSPTDMTVLQANTELDVIEFL
KSSIELDWEIFWNITLQLLDFVPNFDFEIGKATEFAKKLNLPQRDVEMTTETIISAFYYLLCSRR
KSGMILVEHWVSEGLLPLDNHYHFFNDKSLATFDSSLLEREVVWVESPVDTEQKGKNDLIKI
QIIRPKSTEKLPVVITASPYHLGINEKANDLALHEMNVDLEKKDSHKIHVQGKLPQKRPSETK
ELPIVDKAPYRFTHGWTYSLNDYFLTRGFASIYVAGVGTRGSNGFQTSGDYQQIYSMTAVID
WLNGRTRAYTSRKKTHEIKATWANGKVAMTGKSYLGTMAYGAATTGVDGLEVILAEAGISS
WYNYYRENGLVRSPGGFPGEDLDVLAALTYSRNLDGADYLKGNDEYEKRLAEMTTALDRK
SGDYNQFWHDRNYLINSDQVRADVLIVHGLQDWNVTPEQAYNFWQALPEGHAKHAFLHR
GAHIYMNSWQSIDFSETINAYFSAKLLDRDLNLNLPPVILQENSKEQVWSAVSKFGGDDQLK
LPLGKTAVSFAQFDNHYDDESFKKYSKDFNVFKKDLFENKANEAVIDLELPSELTINGPIELEI
RLKLNDSKGLLSAQILDFGPKKRLEDKARVKDFKVLDRGRNFMLDDLVELPLVESPYQLVTK
GFTNLQNKDLLTVSDLKADEWFTLKFELQPTIYHLEKADKLRVILYSTDFEHTVRDNRKVTYE
IDLSQSKLIIPIESVKK (SEQ ID NO: 1)

Xaa-Prolyl dipeptidyl aminopeptidase from *Streptococcus suis* has the sequence :

MRFNQFSFIKKETSVYLQELDTLGFQLIPDASSKTNLETFVRKCHFLTANTDFALSNMIAEWD
TDLLTFFQSDRELTDQIFYQVAFQLLGFVPGMDYTDVMDFVEKSNFPIVYGDIIDNLYQLLNT
RTKSGNTLIDQLVSDDLIPEDNHYHFFNGKSMATFSTKNLIREVVYVETPVDTAGTGQTDIVK
LSILRPHFDGKIPAVITNSPYHQGVNDVASDKALHKMEGELAEKQVGTIQVKQASITKLDLDQ
RNLPVSPATEKLGHITSYSLNDYFLARGFASLHVSGVGTLGSTGYMTSGDYQQVEGYKAVI
DWLNGRTKAYTDHTRSLEVKADWANGKVATTGLSYLGTMSNALATTGVDGLEVIIAEAGISS
WYDYYRENGLVTSPGGYPGEDLDSLTALTYSKSLQAGDFLRNKAAYEKGLAAERAALDRTS
GDYNQYWHDRNYLLHADRVKCEVVFTHGSQDWNVKPIHVWNMFHALPSHIKKHLFFHNGA
HVYMNNWQSIDFRESMNALLSQKLLGYENNYQLPTVIWQDNSGEQTWTTLDTFGGENETV
LPLGTGSQTVANQYTQEDFERYGKSYSAFHQDLYAGKANQISIELPVTEGLLLNGQVTLKLR
VASSVAKGLLSAQLLDKGNKKRLAPIPAPKARLSLDNGRYHAQENLVELPYVEMPQRLVTK
GFMNLQNRTDLMTVEEVVPGQWMNLTWKLQPTIYQLKKGDVLELILYTTDFECTVRDNSQ
WQIHLDLSQSQLILPH (SEQ ID NO:18)

The XaaProDAP may be an enzyme naturally occurring in e.g. bacteria or mammals, but it may also be a functional variant of such an enzyme. A non-limiting example of a functional variant is an analogue, an extended or a truncated version of a naturally occurring XaaProDAP which functional variant retain dipeptidase activity specific for Xaa-Pro dipeptides.

The 3C proteases (or Protein 3C, Picornian 3C or Picornaviral 3C) are a group of cysteine proteases with a serine proteinase-like fold that are responsible for generating mature viral proteins from a precursor polyprotein in vira from the Picornaviridae family. "Picornaviral 3C protease" as used herein is intended to mean a protease originating from the *Picorna viridae* including functional variants thereof, which protease cleave the peptide bond between a P1-P1´ Gln-Gly pair where the scissile bond connects Gln and Gly (where P1 and P1´ according to commonly used notation denote the first amino acids on the N-terminal and C-terminal sides of the scissile bond, respectively). Several picornaviral 3C proteases, have an additional preference for Pro in P2' where P2' denote the second amino acid on the C-terminal side of the scissile bond. Enzymes with this substrate specificity are typically isolated from virus of the genus enterovirus, which currently comprises Coxsackie,Echovirus, Enterovirus 71, Poliovirus and Rhinovirus. Non-limiting examples of such picornian 3C proteases are Human Rhino Virus 14 3C (HRV14 3C) protease having the sequence GPNTEFALSLLRKNIMTITTSKGEFTGLGIHDRVCVIPTHAQPGDDVLVNGQKIRVKDKYKLV DPENINLELTVLTLDRNEKFRDIRGFISEDLEGVDATLVVHSNNFTNTILEVGPVTMAGLINLS STPTNRMIRYDYATKTGQCGGVLCATGKIFGIHVGGNGRQGFSAQLKKQYFVEKQ (SEQ ID NO: 2), Enterovirus 71 3C protease, Coxsackievirus A16 3C protease, Coxsackievirus B3 3C protease (CVB3 3C) having the sequence GPAFEFAVAMMKRNSSTVKTEYGEFTMLGIYDRWAVLPRHAKPGPTILMNDQEVGVLDAKE LVDKDGTNLELTLLKLNRNEKFRDIRGFLAKEEVEVNEAVLAINTSKFPNMYIPVGQVTEYGF LNLGGTPTKRMLMYNFPTRAGQCGGVLMSTGKVLGIHVGGNGHQGFSAALLKHYFNDEQ (SEQ ID NO: 19), cowpea mosaic comovirus-type picornian 3C and Human Poliovirus 3C protease. These 3C proteases are able to release a protein with Gly-Pro in the N-terminal from a large fusion protein and can often be identified by having a Gly-Pro naturally occurring in their own native N-terminal. According to the present invention the picornaviral 3C protease may be an enzyme naturally occurring in the *Picorna viridae*, but it may also be a functional variant of such an enzyme. A non-limiting example of a functional variant is an

analogue, an extended or a truncated version of a naturally occurring picornaviral 3C protease which functional variant retain substrate specificity for the Gln-Gly pair.

Picornaviral 3C protease is commercially available as e.g. His-tagged HRV14 3C (Sino Biological Inc), GST-tagged HRV14 3C (Speed Biosystems) and as dual-tagged (GST plus His) HRV14 3C (Thermo Scientific Pierce).

In one embodiment the picornaviral 3C protease cleaves the peptide bond between a P1-P1' Gln-Ala pair where the scissile bond connects Gln and Ala. In another embodiment the picornaviral 3C protease cleaves the peptide bond between a P1-P1' Gln-Ser pair where the scissile bond connects Gln and Ser. In another embodiment the picornaviral 3C protease natively has a proline as the second amino acid residue from the N-terminal, or the picornaviral 3C protease is an analog of a picornaviral 3C protease which natively has a proline as the amino acid residue in position 2 from the N-terminal. In another embodiment the picornaviral 3C protease natively has a Gly-Pro in the N-terminal.

"Matured protein" as used herein is intended to mean a protein or a polypeptides of interest, or an extended version thereof which extended version can be cleaved at its N-terminus by XaaProDAP. The matured protein is often present as a fusion protein during its manufacture, such as a protein comprising a tag sequence, an optional linker sequence, a picornaviral 3C protease site in addition to the matured protein. Non-limiting examples of a mature protein is glucagon, PYY, GLP-1(7-37), Arg34-GLP1(7-37), Arg34-GLP-1(9-37) and Arg34-GLP-1(11-37). Using the commonly used single letter abbreviation of amino acid residues, for instance, Arg34-GLP-1(7-37) is K34R-GLP-1(7-37).

"Fusion protein" as used herein is intended to mean a hybrid protein expressed by a nucleic acid molecule comprising nucleotide sequences encoding at least two proteins. For example, a fusion protein can comprise a tag protein fused with a protein having an activity of pharmaceutical interest. Fusion proteins are often used for improving recombinant expression of therapeutic proteins as well as for improved recovery and purification of such proteins from cell cultures and the like. Fusion proteins may also comprise synthetic sequences, e.g. a linker sequence.

"Fusion partner protein" as used herein is intended to mean a protein which is part of a fusion protein, i.e. one of the at least two proteins encompassed by the fusion protein. Non-limiting examples of fusion partner proteins are tag proteins and solubilisation domains such as His6-tags, Maltose-binding protein, Thioredoxin, etc.

"Fusion enzyme" as used herein is intended to mean a fusion protein comprising at least two proteins which are both enzymes (in the sense that the two proteins have backbone sequences that are covalently connected).

"Tag protein" as used herein is intended to mean a protein which is attached to another protein in order to facilitate or improve the manufacture of said other protein, e.g. facilitating or improving the recombinant expression, recovery and/or purification of said other protein. Non-limiting examples of tag proteins are His6-tags, Glutathione S-transferase (GST), Maltose-binding Protein (MBP), *Staphylococcus aureus* protein A, biotinylated peptides and highly basic proteins from thermophilic bacteria as described in WO2006/108826 and WO2008/043847.

"Tag sequence" as used herein is intended to mean a sequence comprising a protein. A tag sequence may optionally also comprise an additional sequence, e.g. a linker sequence. Protein tags are peptide sequences genetically grafted onto a recombinant protein,which may be removable by chemical agents or by enzymatic means, such as proteolysis. Tags are attached to proteins for various purposes

"Linker" as used herein is intended to mean an amino acid sequence which is typically used to facilitate the function, folding or expression of fusion proteins. It is known in the art that two proteins present in the form of a fusion enzyme may interact with the enzyme activities of each other, an interaction that can often be eliminated or reduced by the insertion of a linker between the two enzyme sequences.

"Analogues" as used herein is intended to mean proteins which are derived from another protein by means of substitution, deletion and/or addition of one or more amino acid residues from the protein. Non-limiting example of analogues of GLP-1(7-37) are K34R-GLP-1(7-37) where residue 34 has been substituted by an arginine residue and K34R-GLP-1(9-37) where residue 34 has been substituted with an arginine residue and amino acid residues 7-8 have been deleted (using the common numbering of amino acid residues for GLP-1 peptides).

"Functional variant" as used herein is intended to mean a chemical variant of a certain protein which has an altered sequence of amino acids but retains substantially the same function as the original protein. Non-limiting examples of functional variants are e.g. extended proteins, truncated proteins, fusion proteins and analogues. Non-limiting examples of functional variants of HRV14 3C are e.g. His6 tagged HRV14 3C and GST-tagged HRV14 3C. Non-limiting functional variants of GLP-1(7-37) are K34R-GLP-1(7-37).

In one embodiment, a function variant of a protein comprises from 1-2 amino acid substitutions, deletions or additions as compared said protein. In another embodiment, a functional variant comprises from 1-5 amino acid substitutions, deletions or additions as compared to said protein. In another embodiment, a functional variant comprises from 1-15 amino acid substitutions, deletions or additions relative to the corresponding naturally

occurring protein or naturally occurring sub-sequence of a protein. In yet another embodiment a functional variant of a protein analogue in the overlapping sequence comprises no more than 10 amino acids different from the sequence of the native protein. . In yet another embodiment a functional variant of a protein analogue in the overlapping sequence comprises no more than 5 amino acids different from the sequence of the native protein.

A “Solubilisation domain” as used herein is intended to mean a protein which is part of a fusion protein and which is to render said fusion protein more soluble than the fusion partner protein itself under certain conditions. Non-limiting examples of solubilisation domains are DsbC (Thiol:disulfide interchange protein), RL9 (Ribosomal Protein L9) as described in WO2008/043847, MBP (Maltose-binding Protein), NusA (Transcription termination/antitermination protein) and Trx(Thioredoxin).

The term “enzymatic treatment” as used herein is intended to mean a contacting of a substrate protein with an enzyme which catalyses at least one reaction involving said substrate protein. One common enzymatic treatment is the contacting of a fusion protein with an enzyme having proteolytic activity in order to separate two proteins being constituents of the fusion protein.

Phosphate buffered saline (abbreviated PBS) is a buffer solution commonly used and typically is a water-based salt solution containing sodium phosphate, sodium chloride and, in some formulations, potassium chloride and potassium phosphate. A typical PBS buffer used for enzymatic reactions in the present invention is (8.05 mM Na2HPO4x2H2O, 1,96 mM KH2PO4, 140 mM NaCl, pH 7.4).

According to a second aspect of the invention there is provided a method for release of a matured protein from a fusion protein by enzymatic treatment using a picornaviral 3C protease and a Xaa-Pro-dipeptidyl aminopeptidase (XaaProDAP), wherein said fusion protein has the structure :

Tag sequence – Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-QGP - Protein of interest,

wherein

Tag sequence is a fusion partner protein connected at its C-terminal,

Xaa1, Xaa2, Xaa3, Xaa4 and Xaa5 are independently selected from the group consisting of the naturally occurring amino acid residues, and

Protein of interest is a protein connected at its N-terminal and comprising said matured protein and optionally an N-terminal extension which can be cleaved by XaaProDAP or by picornaviral 3C protease and XaaProDAP.

In one embodiment the protein of interest (or the matured protein) is GLP-1, Glucagon, PYY , amylin or a functional variant thereof.

In one embodiment Xaa2-Xaa3-Xaa4-Xaa5 is EVLF. In another embodiment Xaa1 is selected from L and S. In another embodiment Xaa1-Xaa2-Xaa3-Xaa4-Xaa5 is selected from LETLF (SEQ ID NO:16) and LFAQT (SEQ ID NO:17).

In another embodiment of the method the fusion protein has the structure :

Tag sequence – SSSGGSEVLFQGP – Protein of interest.

The method of the present invention form a very low amount of product related impurities for most proteins. However, particularly low amounts of product related impurities may arise when a) the amino acid residue at position 2 from the N-terminal of said matured protein is different from proline, b) the amino acid residue at position 2 from the N-terminal of said matured protein is different from alanine, and/or 3) the N-terminal amino acid residue of said matured protein is different from proline.

In an embodiment of the method according to the present invention said enzymatic treatment is done after the fusion protein has been isolated from an expression host. There are a number of different ways of carrying out the enzymatic treatment, all of which are suitable for liberating the matured protein. Firstly, the enzymatic treatment may be done by simultaneously contacting said fusion protein with said picornaviral 3C protease and said XaaProDAP. Secondly, the enzymatic treatment may be done by first contacting said fusion protein with said picornaviral 3C protease to separate a tag from the fusion protein, and then contacting said de-tagged protein with said XaaProDAP to form said matured protein. In one embodiment said picornaviral 3C protease may be inactivated or separated from the de-tagged protein prior to contacting said de-tagged protein with said XaaProDAP to form said matured protein. In another embodiment the XaaProDAP may be added to a reaction mixture containing the picornaviral 3C protease, fusion protein and de-tagged protein once the majority of the fusion protein has been cleaved by picornaviral 3C protease into the de-tagged protein.

The enzymatic treatment to liberate the matured protein may be done at a pH in the range from about 6 to about 9 , such as in the range from about 7 to about 8.5, in the range from about 7.5 to about 8.5, in the range from about 8 to about 9, or in the range from about 6 to about 7.

Suitable temperatures for the enzymatic treatment are in the range from about 0 °C to about 50 °C, in the range from 30 °C or 37 °C, in the range from about 0 °C to about 15 °C, in the range from about 0 °C to about 10 °C, in the range from about 2 °C to about 10 °C, in

the range from about 5 °C to about 15 °C, in the range from about 0 °C to about 5 °C, or in the range from about 2 °C to about 8 °C.

In an embodiment of the inventino the enxymatic treatment is done at a pH in the range from about pH 7.5 to about pH 8.5 and at a temperature in the range from about 4 °C to about 10 °C.

Suitable reaction times for the enzymatic treatment are in the range for about 1 minute to about 3 hours, for about 3 hours to about 24 hours, for about 3 hours to about 16 hours, for about 6 hours to about 24 hours, or for about 10 hours to about 16 hours.

Suitable buffers for the enzymatic treatment are various buffers exhibiting buffer capacity at the desired pH for the enzymatic treatment. In one embodiment the enzymatic treatment is done in the presence of a buffer selected from PBS, TRIS and HEPES.

In a further embodiment of the method of the present invention the picornaviral 3C protease and thed XaaProDAP are co-expressed in an expression host.

In a further embodiment of the method of the present invention the picornaviral 3C protease and the XaaProDAP are expressed as a fusion enzyme in an expression host.

In a further embodiment of the method of the present invention the fusion protein is co-expressed in an expression host together with the picornaviral 3C protease and the XaaProDAP.

Manufacture of picornaviral 3C protease and XaaProDAP.

The picornaviral 3C protease, the XaaProDAP as well as the fusion protease may be produced by means of recombinant nucleic acid techniques. In general, a cloned wild-type picornaviral 3C protease or XaaProDAP nucleic acid sequence is modified to encode the desired protein. This modification can also include the in-frame fusion of the nucleic acid sequences encoding two or more proteins to be expressed as a fusion protein. Such a fusion protein can be the fusion protease (picornaviral 3C protease and XaaProDAP) as well as the fusion protease, picornaviral 3C protease or XaaProDAP fused to e.g. a His-tag or a solubilization domain (DsbC, RL9, MBP, NusA or Trx). This modified sequence is then inserted into an expression vector, which is in turn transformed or transfected into expression host cells.

The nucleic acid construct encoding the picornaviral 3C protease, XaaProDAP or fusion protease may suitably be of genomic, cDNA or synthetic origin. Amino acid sequence alterations are accomplished by modification of the genetic code by well known techniques.

The DNA sequences encoding the picornaviral 3C protease, XaaProDAP or fusion protease are usually inserted into a recombinant vector which may be any vector, which may

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

The vector is preferably an expression vector in which the DNA sequence encoding the picornaviral 3C protease, XaaProDAP or fusion protease is operably linked to additional segments required for transcription of the DNA. The term, “operably linked” indicates that the segments are arranged so that they function in concert for their intended purposes, e.g. transcription initiates in a promoter and proceeds through the DNA sequence coding for the polypeptide until it terminates within a terminator.

Thus, expression vectors for use in expressing picornaviral 3C protease, XaaProDAP or fusion protease will comprise a promoter capable of initiating and directing the transcription of a cloned gene or cDNA. The promoter may be any DNA sequence, which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell.

Additionally, expression vectors for use of expression of picornaviral 3C protease, XaaProDAP or fusion protease will also comprise a terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3’ terminus of the nucleic acid sequence encoding the polypeptide. Any terminator which is functional in the host cell of choice may be used in the present invention.

Expression of picornaviral 3C protease, XaaProDAP or fusion protease can be aimed for either intracellular expression in the cytosol of the host cell or be directed into the secretory pathway for extracellular expression into the growth medium.

Intracellular expression is the default pathway and requires an expression vector with a DNA sequence comprising a promoter followed by the DNA sequence encoding the picornaviral 3C protease, XaaProDAP or fusion protease polypeptide followed by a terminator.

To direct the picornaviral 3C protease, XaaProDAP or fusion protease into the secretory pathway of the host cells, a secretory signal sequence (also known as signal peptide or a pre sequence) is needed as an N-terminal extension of the picornaviral 3C protease, XaaProDAP or fusion protease. A DNA sequence encoding the signal peptide is joined to the 5’ end of the DNA sequence encoding the picornaviral 3C protease,

XaaProDAP or fusion protease in the correct reading frame. The signal peptide may be that normally associated with the protein or may be from a gene encoding another secreted protein.

The procedures used to ligate the DNA sequences coding for the picornaviral 3C protease, XaaProDAP or fusion protease, the promoter, the terminator and/or secretory signal sequence, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989).

The host cell into which the DNA sequences encoding picornaviral 3C protease, XaaProDAP or fusion protease is introduced may be any cell that is capable of expressing picornaviral 3C protease, XaaProDAP or fusion protease either intracellularly or extracellulaly. If posttranslational modifications are needed, suitable host cells include yeast, fungi, insects and higher eukaryotic cells such as mammalian cells.

The picornaviral 3C protease and the XaaProDAP may be co-expressed in an expression host. In one embodiment the picornaviral 3C protease and the XaaProDAP may be expressed as a fusion enzyme in an expression host. In another embodiment said fusion protein is co-expressed in an expression host together with said picornaviral 3C protease and said XaaProDAP, thus allowing the manufacture of the 3C protease and the XaaProDAP simulataneously with the enzymatic treatment of the fusion protein to liberate the matured protein.

Bacterial expression

Examples of suitable promoters for directing the transcription of the nucleic acid constructs in a bacterial host cell are, for expression in *E. coli*, the promoters obtained from the *lac* operon, the *trp* operon and hybrids thereof *trc* and *tac*, all from *E. coli* (DeBoer *et al.*, 1983, *Proceedings of the National Academy of Sciences USA* 80: 21-25). Other even stronger promoters for use in *E. coli* are the bacteriophage promoters from T7 and T5 phages. The T7 promoter requires the presence of the T7 polymerase in the *E. coli* host (Studier and Moffatt, *J. Mol. Biol.* 189, 113, (1986)). All these promoters are regulated by induction with IPTG, lactose or tryptophan to initiate transcription at strategic points in the bacterial growth period. *E. coli* also has strong promoters for continuous expression, eg. the synthetic promoter used to express hGH in Dalbøge et al, 1987, Biotechnology 5, 161-164.

For the expression in *Bacillus,* the promoters from *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*),

Bacillus licheniformis penicillinase gene (*penP*), *Bacillus subtilis xylA* and *xylB* genes are suitable examples. Further promoters are described in "Useful proteins from recombinant bacteria" in *Scientific American*, 1980, 242: 74-94; and in Sambrook *et al.*, 1989, *supra*.

Effective signal peptide coding regions for bacterial host cells are, for *E. coli*, the signal peptides obtained from the genes *DegP, OmpA, OmpF, OmpT, PhoA* and Enterotoxin STII, all from *E. coli.* For *Bacillus* the signal peptide regions obtained from *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* neutral proteases (*nprT*, *nprS*, *nprM*) and *Bacillus subtilis prsA*. Further signal peptides are described by Simonen and Palva, 1993, *Microbiological Reviews* 57: 109-137. For both *E. coli* and *Bacillus*, signal peptides can be created *de novo* according to the rules outlined in the algorithm SignalP (Nielsen et al, 1997, Protein Eng. 10, 1-6., Emanuelsen et al, 2007, Nature Protocols 2, 953-971). The signal sequences are adapted to the given context and checked for SignalP score.

Examples of strong terminators for transcription are the aspartase *aspA* as in the Thiofusion Expression System, the T7 gene *10* terminator in the pET vectors (Studier et al) and the terminators of the ribosomal RNA genes *rrnA, rrnD*.

Examples of preferred expression hosts are *E. coli* K12 W3110, *E. coli* K12 with a trace of B, MC1061 and *E. coli* B BL21 DE3, harbouring the T7 polymerase by lysogenization with bacteriophage λ. These hosts are selectable with antibiotics when transformed with plasmids for expression. For antibiotics free selection the preferred host is eg. *E. coli* B BL21 DE3 3xKO with deletion of the 2 D,L-alanine racemase genes *Δalr, ΔdadX*, and deletion of the Group II capsular gene cluster *Δ (kpsM-kpsF)*, specific for *E. coli* B and often associated with pathogenic behaviour. The deletion of the Group II gene cluster brings *E. coli* B BL21 DE3 3xKO into the same safety category as *E. coli* K12. Selection is based on non-requirement of D-alanine provided by the alr gene inserted in the expression plasmid instead of the AmpR gene.

Expression of the fusion protein may be done in a similar way by bacterial expression as described above for the picornaviral 3C protease, XaaProDAP or fusion protease.

Once picornaviral 3C protease, XaaProDAP or a fusion protease has been expressed in a host organism it may be recoved and purified to the required purity by conventional techniques. Non-limiting examples of such conventional recovery and purification techniques are centrifugation, solubilization, filtration, precipitation, ion-exchange

chromatography, immobilized metal affinity chromatography (IMAC), RP-HPLC, gel-filtration and freeze drying.

Examples of recombinant expression and purification of picornaviral 3C protease may be found in e.g. Cordingley et al., J. Virol. 1989, 63, pp5037-5045, Birch et al., Protein Expr Purif., 1995, 6, pp609-618 and WO2008/043847.

Examples of microbial expression and purification of XaaProDAP from *Lactococcus lactis* may be found in e.g. Chich et al, Anal. Biochem, 1995, 224, pp 245-249 and Xin et al., Protein Expr. Purif. 2002, 24, pp530-538

The invention is further described by the following non-limiting embodiments:

1. Method for release of a matured protein from a fusion protein by enzymatic treatment using a picornaviral 3C protease and a Xaa-Pro-dipeptidyl aminopeptidase (XaaProDAP).
2. The method according to embodiment 1 wherein said picornaviral 3C protease natively has a Gly-Pro in the N-terminal.
3. The method according to any of embodiments 1-2 wherein said picornaviral 3C protease is HRV14 3C or a functional variant thereof.
4. The method according to any of embodiments 1-2 wherein said picornaviral 3C protease is CVB3 3C or a functional variant thereof.
5. The method according to any of embodiments 1-2 wherein said picornaviral 3C protease originates from a virus selected from Enterovirus, Coxsackievirus, Cowpea mosaic comovirus, Rhinovirus and Poliovirus.
6. The method according to any of embodiments 1-5 wherein said XaaProDAP is an EC 3.4.14.11 enzyme or a functional variant thereof.
7. The method according to any of embodiments 1-6 wherein said XaaProDAP is an enzyme from *Lactococcus sp.* or a functional analogue derived from the enzyme from *Lactococcus sp.*
8. The method according to any of embodiments 1-6 wherein said XaaProDAP is an enzyme from *Streptococcus suis* or a functional analogue thereof.
9. The method according to any of embodiments 1-5 wherein said XaaProDAP is an EC 3.4.14.5 enzyme or a functional variant thereof.
10. The method according to any of embodiments 3-4 and 6-8, wherein said functional analogue in the overlapping sequence comprises no more than 10 amino acids different from the sequence of the native protein.
11. The method according to any of embodiments 1-10 wherein said fusion protein has the structure :

Tag sequence – Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-QGP - Protein of interest, wherein

Tag sequence is a fusion partner protein connected at its C-terminal,

Xaa1, Xaa2, Xaa3, Xaa4 and Xaa5 are independently selected from the group consisting of the naturally occurring amino acid residues, and

Protein of interest is a protein connected at its N-terminal and comprising said matured protein and optionally an N-terminal extension which can be cleaved by XaaProDAP or by picornaviral 3C protease and XaaProDAP.

12. The method according to embodiment 11 wherein Xaa2-Xaa3-Xaa4-Xaa5 is EVLF.
13. The method according to embodiment 12 wherein Xaa1 is selected from L and S.
14. The method according to any of embodiments 12-13 wherein said fusion protein has the structure :

Tag sequence – SSSGGSEVLFQGP – Protein of interest.

15. The method according to any of embodiments 11-12 wherein Xaa1-Xaa2-Xaa3-Xaa4-Xaa5 is selected from LETLF (SEQ ID NO:16) and LFAQT (SEQ ID NO:17).
16. The method according to any of embodiments 1-15 wherein the amino acid residue at position 2 from the N-terminal of said matured protein is different from proline.
17. The method according to any of embodiments 1-15 wherein the amino acid residue at position 2 from the N-terminal of said matured protein is different from alanine.
18. The method according to any of embodiments 1-17 wherein the N-terminal amino acid residue of said matured protein is different from proline.
19. The method according to any of embodiments 1-18 wherein said enzymatic treatment is done after the fusion protein has been isolated from an expression host.
20. The method according to any of embodiments 1-19 wherein said enzymatic treatment is done by simultaneously contacting said fusion protein with said picornaviral 3C protease and said XaaProDAP.
21. The method according to any of embodiments 1-19 wherein said enzymatic treatment is done by first contacting said fusion protein with said picornaviral 3C protease to separate a tag from the fusion protein, and then contacting said de-tagged protein with said XaaProDAP to form said matured protein.
22. The method according to embodiment 21 wherein said picornaviral 3C protease is inactivated or separated from the de-tagged protein prior to contacting said de-tagged protein with said XaaProDAP to form said matured protein.
23. The method according to embodiment 21 wherein said XaaProDAP is added to a reaction mixture containing the picornaviral 3C protease, fusion protein and de-tagged protein

once the majority of the fusion protein has been cleaved by picornaviral 3C protease into the de-tagged protein.

24. The method according to any of embodiments 1-23 wherein said enzymatic treatment is done at a pH in the range from about 6 to about 9.
25. The method according to any of embodiments 1-24 wherein said enzymatic treatment is done at a pH in the range from about 7 to about 8.5.
26. The method according to any of embodiments 1-25 wherein said enzymatic treatment is done at a pH in the range from about 7.5 to about 8.5.
27. The method according to any of embodiments 1-24 wherein said enzymatic treatment is done at a pH in the range from about 8 to about 9.
28. The method according to any of embodiments 1-24 wherein said enzymatic treatment is done at a pH in the range from about 6 to about 7.
29. The method according to any of embodiments 1-28 wherein said enzymatic treatment is done at a temperature in the range from about 0 °C to about 50 °C, such as 30 °C or 37 °C.
30. The method according to any of embodiments 1-29 wherein said enzymatic treatment is done at a temperature in the range from about 0 °C to about 15 °C.
31. The method according to any of embodiments 1-30 wherein said enzymatic treatment is done at a temperature in the range from about 0 °C to about 10 °C.
32. The method according to any of embodiments 1-31 wherein said enzymatic treatment is done at a temperature in the range from about 2 °C to about 10 °C.
33. The method according to any of embodiments 1-30 wherein said enzymatic treatment is done at a temperature in the range from about 5 °C to about 15 °C.
34. The method according to any of embodiments 1-31 wherein said enzymatic treatment is done at a temperature in the range from about 0 °C to about 5 °C.
35. The method according to any of embodiments 1-32 wherein said enzymatic treatment is done at a temperature in the range from about 2 °C to about 8 °C.
36. The method according to any of embodiments 1-26 wherein said enzymatic treatment is done at a pH in the range from about pH 7.5 to about pH 8.5 and at a temperature in the range from about 4 °C to about 10 °C.
37. The method according to any of clams 1-36 wherein said enzymatic treatment is done for about 1 minute to about 3 hours.
38. The method according to any of embodiments 1-36 wherein said enzymatic treatment is done for about 3 hours to about 24 hours.

39. The method according to embodiment 38 wherein said enzymatic treatment is done for about 3 hours to about 16 hours.
40. The method according to embodiment 38 wherein said enzymatic treatment is done for about 6 hours to about 24 hours.
41. The method according to any of embodiments 1-40 wherein said enzymatic treatment is done in the presence of a buffer selected from PBS (phosphate buffered saline, TRIS(tris(hydroxymethyl)aminomethane) and HEPES(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid).
42. The method according to any of embodiments 38 or 41 wherein said enzymatic treatment is done for about 10 hours to about 16 hours.
43. The method according to any of embodiments 1-42 wherein said picornaviral 3C protease and said XaaProDAP are co-expressed in an expression host.
44. The method according to any of embodiments 1-42 wherein said picornaviral 3C protease and said XaaProDAP are expressed as a fusion enzyme in an expression host.
45. The method according to any of embodiments 1-20 wherein said fusion protein is co-expressed in an expression host together with said picornaviral 3C protease and said XaaProDAP.
46. The method according to any of embodiments 43-45 wherein said expression host is *E. coli*.
47. The method according to any of embodiments 1-46 wherein said matured protein is GLP-1, Glucagon, PYY, amylin or a functional variant thereof.
48. The method according to embodiment 47 wherein said functional variant is an analogue.
49. The method according to embodiment 48 wherein said analogue comprises no more than 10 amino acid differences as compared to the native matured protein.
50. The method according to embodiment 49 wherein said analogues comprise no more than 5 amino acid differences as compared to the native matured protein.
51. The method according to any of embodiments 1-20, 24-42 or 44-50 wherein said treatment using a picornaviral 3C protease and a XaaProDAP is done by using a fusion enzyme comprising a picornaviral 3C protease and a XaaProDAP.
52. Fusion enzyme comprising the catalytic domains of a picornaviral 3C protease and a XaaProDAP.
53. The fusion enzyme according to embodiment 52 which comprises HRV14 3C.
54. The fusion enzyme according to any of embodiments 52-53 wherein said XaaProDAP is an EC 3.4.14.11 enzyme.

55. The fusion enzyme according to any of embodiments 52-54 wherein said XaaProDAP is an enzyme from *Lactococcus sp* or a functional analogue derived from the enzyme from *Lactococcus sp*.
56. The fusion enzyme according to any of embodiments 52-53 wherein said XaaProDAP is an EC 3.4.14.5 enzyme
57. The fusion enzyme according to any of embodiments 52-56 comprising an amino acid residue linker between said picornaviral 3C protease or the catalytic domain thereof and said XaaProDAP or the catalytic domain thereof, where said linker has a length of from 2 to 100 amino acid residues.
58. The fusion enzyme according to any of embodiments 52-57 wherein a fusion partner is attached to the N-terminal of said fusion enzyme.
59. Recombinant expression of the fusion enzyme according to any of embodiments 52-58.
60. Recombinant expression according to embodiment 59, wherein the host cell is *E. coli.*

EXAMPLES

Example 1

Plasmid constructs and expression of XaaProDAP variants

The pET system was used for expression of enzymes as this system provides a powerful means of expressing proteins in *E.coli*. In pET vectors, target genes are cloned under control of strong bacteriophage T7 transcription and translation signals, and expression is induced by providing a source of T7 RNA polymerase in the host cell.

E.coli expression plasmids (pET22b, Novagen) encoding Xaa-Prolyl Dipeptidyl aminopeptidase (XaaProDAP, full length) from *Lactococcus lactis subsp. cremoris* CNCM I-1631 (SEQ ID NO:1) with different N-terminal extensions/fusion partners were prepared.

One plasmid comprised a His6 tag in the N-terminal with the sequence MHHHHHH (SEQ ID NO 3). In five other plasmids different fusions partners were designed to comprise a His6 tag either in the N-or C-terminal intervened with a small GSGSG linker to promote solubilisation and expression yields. The sequences were as follows:

Trx with the sequence

MHHHHHHGSGSGSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA (SEQ ID NO:4)

RL9 with the sequence

MKVILLRDVPKIGKKGEIKEVSDGYARNYLIPRGFAKEYTEGLERAIKHEKEIEKRKKEREREE

SEKILKELKKRTHVVKVKAGEGGKIFGAVTAATVAEEISKTTGLKLDKRWFKLDKPIKELGEY SLEVSLPGGVKDTIKIRVEREEGSGSGHHHHHH (SEQ ID NO: 5)

DsbC with the sequence

MDDAAIQQTLAKMGIKSSDIQPAPVAGMKTVLTNSGVLYITDDGKHIIQGPMYDVSGTAPVN VTNKMLLKQLNALEKEMIVYKAPQEKHVITVFTDITCGYCHKLHEQMADYNALGITVRYLAFP RQGLDSDAEKEMKAIWCAKDKNKAFDDVMAGKSVAPASCDVDIADHYVLGVQLGVSGTPA VVLSNGTLVPGYQPPKEMKEFLDEHQKMTSGKGSGSGHHHHHH (SEQ ID NO: 6),

NusA with the sequence

MNKEILAVVEAVSNEKALPREKIFEALESALATATKKKYEQEIDVRVQIDRKSGDFDTFRRWL VVDEVTQPTKEITLEAARYEDESLNLGDYVEDQIESVTFDRITTQTAKQVIVQKVREAERAMV VDQFREHEGEIITGVVKKVNRDNISLDLGNNAEAVILREDMLPRENFRPGDRVRGVLYSVRP EARGAQLFVTRSKPEMLIELFRIEVPEIGEEVIEIKAAARDPGSRAKIAVKTNDKRIDPVGACV GMRGARVQAVSTELGGERIDIVLWDDNPAQFVINAMAPADVASIVVDEDKHTMDIAVEAGNL AQAIGRNGQNVRLASQLSGWELNVMTVDDLQAKHQAEAHAAIDTFTKYLDIDEDFATVLVEE GFSTLEELAYVPMKELLEIEGLDEPTVEALRERAKNALATIAQAQEESLGDNKPADDLLNLEG VDRDLAFKLAARGVCTLEDLAEQGIDDLADIEGLTDEKAGALIMAARNICWFGDEAGSGSGH HHHHH (SEQ ID NO: 7).

MBP with the sequence

MHHHHHHGSGSGKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQ VAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEA LSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGK YDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDT SKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPL GAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEAL KDAQTRITK (SEQ ID NO: 8)

The gene fragments encoding full-length of XaaProDAP (SEQ ID NO:1), His6 tags, fusion partners and linkers were codon-optimized for expression in *E.coli* and prepared by gene synthesis (GenScript). A nucleotide sequence encoding a flexible Gly-Ser-rich linker, comprising a Hepatitis A Virus 3C protease (HAV) cleavage site with the sequence GGSSGSGSELRTQSGS (SEQ ID NO:9) was introduced adjacent to the N-terminal amino acid of the XaaProDAP sequence to allow enzymatic separation of the fusion partner from the XaaProDAP enzyme if needed. This linker region also contained an XhoI site and BamHI site, which allows introduction of another N-terminal extension to the enzyme. For cloning purposes the XaaProDAP sequence was prepared with small 3'-end overhang encoding

GSS, which includes a SacI restriction site that allows introduction C-terminal extensions to the XaaProDAP enzyme.

The plasmid constructs specified in Table 1 were generated by DNA synthesis and subcloning into pET22b vectors using standard cloning technologies known to those of ordinary skill in the art (obtained from GenScript):

Table 1. Plasmid constructs using pET22b vectors

Product name	Calculated Molecular weight (Dalton)	His6 / fusion partner	HAV linker	XaaProDAP enzyme	Linker
His-LLXaaProDAP	90362.97 Da	SEQ ID NO: 3	SEQ ID NO: 9	SEQ ID NO: 1	GSS
His-Trx-LLXaaProDAP	102365.83 Da	SEQ ID NO: 4	SEQ ID NO: 9	SEQ ID NO: 1	GSS
RL9-His-LLXaaProDAP	107621.14 Da	SEQ ID: NO:5	SEQ ID NO:9	SEQ ID NO: 1	GSS
DsbC-His-LLXaaProDAP	114178.60 Da	SEQ ID NO: 6	SEQ ID NO:9	SEQ ID NO: 1	GSS
NusA-His-LLXaaProDAP	145430.53 Da	SEQ ID NO: 7	SEQ ID NO:9	SEQ ID NO: 1	GSS
His-MBP-LLXaaProDAP	131397.99 Da	SEQ ID NO: 8	SEQ ID NO:9	SEQ ID NO: 1	GSS

Evaluation of the XaaProDAP variants by small scale expression and purification

Expression plasmids were transformed into *E.coli* BL21(DE3) and expressed in small scale.

E.coli BL21(DE3) were transformed with plasmid using a procedure based on Heat Shock at 42°C according to the manufacturer. Transformed cells were plated onto LB agar plates and incubate overnight at 37ºC with 10 mg/L ampicillin. Overnight Terrific broth (TB) culture with 0.5% glucose and 50 mg/L carbenicillin of each transformant was prepared at 30ºC and shaking at 700 rpm using a Glas-Col shaker (Glas-Col). 20 µL of overnight culture of each transformant was used to inoculate 0.95 µL of TB medium with 50 mg/L carbenicillin in 96 Deep-Well plates (2 ml) and transformants were propagated overnight at 700 rpm. Expression cultures were incubated at 37ºC until OD600 of 1.5 was reached. The cultures

were then cooled to 20°C and protein induction was carried out overnight using 0.3 mM IPTG. Pellets containing expressed protein were harvested by centrifugation at 1800xG.

Purification screen: Small scale purification using IMAC resin was performed to evaluate the combined expression and purification potential of each construct. In short, 250µL of lysis buffer (50 mM NaPO4, 300 mM NaCl, 10 mM Imidiazole, 10mg/ml Lysozyme, 250U/µL Benzoase and 10% DDM) was added to each pellet and the cells were lyzed using freeze/thaw cycles. Debris was removed by centrifugation, the supernatant was filtered (0.45 µm) and transferred onto 1.2 µm filter plates containing Ni2+-loaded Sepharose Fast Flow (prepared from washing 30 µL of a 50% slurry in 20% EtOH) (GE Healthcare). The supernatant was incubated for 20 min by shaking at 400 rpm with resin to bind the protein and the solute was removed by gentle centrifugation at 100xg for 1 min. The resin was washed with 50mM sodium phosphate, 300 mM NaCl, 30 mM Imidiazole, pH 7.5 by gentle mixing and the resin was dried by centrifugation. To elute the protein, 40µL of elution buffer (50 mM sodium phosphate, 300 mM NaCl, 300 mM Imidazole) was added to the resin, incubated for 10 min by shaking at 400 rpm and the eluate containing partly purified enzymes was collected.

SDS-PAGE analysis of pellets containing partly purified enzyme was performed. Whole lysates of pellets from the expression of 6 XaaProDAP variants were analysed by SDS-PAGE. 5 µL of protein was mixed with 15 µL of SDS-PAGE reducing sample buffer and separated on NuPage 4-12% Bis-Tris polyacrylamide gels. Except for MBP tagged XaaProDAP all constructs gave rise to clear protein bands of the expected size following Coomasie Blue staining of the gels. The eluted protein was also evaluated using SDS-PAGE. In agreement with analysis of the whole lysates, His6, RL9, Trx, DsbC and NusA tagged variants of XaaProDAP, gave rise to a clear protein band of the expected size following Coomassie staining of gels, showing that the *L. lactis* XaaProDAP enzyme could be expressed as soluble protein with a range of different fusion partners and captured using IMAC.

Variants of XaaProDAP, with different fusions partners were initially tested for activity by estimation of color reaction using a small synthetic peptide H-Gly-Pro-p-Nitroanilide (L-1880, Bachem) as substrates for the enzymes. All enzymes exept the MBP showed activity based on a clear color reaction compared to controls without enzyme, thus indicating that large N-terminal fusion partners are not prohibitive for the activity of L. lactis XaaProDAP.

Example 2

Scaling up expression of His-Trx-LLXaaProDAP

In order to prepare a larger amount of XaaProDAP for further testing His-Trx-LLXaaProDAP was selected for expression and purification in larger scale.

BL21(DE3) transformants (from a glycerol stock) harbouring the pET22b plasmid encoding His-Trx-XaaProDAP was propagated overnight in 50 ml of Terrific Broth medium containing 50 mg/L Carbenicillin and 0.5% glucose by shaking at 37°C with 100 rpm (Multitron Standard shaker, 50mm amplitude, Infors HT). The following day, 7.5 ml overnight culture was used to inoculate 750 ml of TB medium with 50 mg/L Carbenicillin in a 2L shaker flask and the culture was subsequently incubated at 37°C with 100 rpm. When OD600 of ~1.5 was reached, the culture was cooled to 20°C for 30 min., before 0.3 mM IPTG was added to induce the protein. The induction was carried out overnight at 20°C at 100 rpm, and cells were harvested by centrifugation at 4000 x g for 10 minutes. Pelleted cells were frozen until usage.

Purification of His-Trx-XaaProDAP

In order to obtain purified protein for further analysis, two consecutive purification steps were conducted.

12 g of cell pellets were suspended in 100 ml lysis buffer containing 50mM sodium phosphate pH 7.5 and 3 µL benzonase). The cells were disrupted in a cell homogenizer at 1.4 kBar for one cycle and cell debris was spun down at 18.000g for 20 min. The supernatant was then sterile filtrated (0.22 µm).The purification of His-Trx-XaaProDAP was done using an ÄKTAExpress for two consecutive purification steps. In the capture step, enzyme from the 100 ml of sample application was purified on a 1 ml HisTrap crude column (GE Healthcare) with a flow rate of 0.8 ml/min using the following buffers :

BufferA: 50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole pH 7.5

Buffer B: 50 mM sodium phosphate, 300 mM NaCl, 300 mM imidazole pH 7.5

Buffer C: : 50 mM sodium phosphate, 300 mM NaCl, 30 mM imidazole pH 7.5

The column was initially equilibrated for 10 column volumes of buffer A. After loading of the application, unbound protein was removed by washing using 7 column volumes of buffer C. A step elution from 0-100 % buffer B for 5 column volumes was used to elute the XaaProDAP protease from the column.

SDS-PAGE was used to analyse the recoved fractions and showed a clear band of the expected size corresponding to His-Trx-XaaProDAP. The purest fractions were pooled and loaded onto a HiLoad S200 16/600 (GE-Healthcare) gel filtration column. And size

separation was performed with a flow rate of 1.2 ml/min using 1X PBS buffer (phosphate buffered saline, pH 7.4). Collected fraction of the predominant peak were analysed by SDS-PAGE and a clear band of the expected size was observed. The purity was estimated to be at least 90% as judged by SDS-PAGE and HPLC analysis. Fractions containing the most enzyme were pooled and the concentration was measured to be 3.5 mg/ml using UV280 measurements (NanoDrop). The His-Trx-XaaProDAP enzyme was active by observation of a clear color change following incubation at 37°C with H-Gly-Pro-p-Nitroanilide synthetic substrate (L-1880, Bachem) compared to a control without enzyme. The enzyme was aliqouted and frozen until used.

Example 3

Plasmid constructs and expression of model fusion proteins containing basic tag.

In order to demonstrate the value of combining XaaProDAP and HRV14 3C for removal of N-terminal tags, three different model fusion proteins were prepared to be used as protein substrates. A basic tag comprising Ribosomal Protein L27 from *T. maritima*, previously described in WO2008/043847 was used as a fusion partner and has the sequence MAHKKSGGVAKNGRDSLPKYLGVKVGDGQIVKAGNILVRQRGTRFYPGKNVGMGRDFTLF ALKDGRVKFETKNNKKYVSVYEE (SEQ ID NO: 10). The fusion proteins were designed so that the RL27 fusion partner can be removed by HRV14 3C enzyme and the remaining GP sequence can be removed by XaaProDAP.

A flexible linker containing a HRV14 cleavage site was used to link the basic tag to the model peptide sequences and had the sequence SSSGGSEVLFQGP (SEQ ID NO: 11). The model peptide sequences used were human pancreatic peptide YY (PYY(3-36)), Glucagon and Glucagon-like peptide-1 (7-37,K34R) (GLP-1(7-37,K34R)) having the following sequences:

PYY(3-36): IKPEAPGEDASPEELNRYYASLRHYLNLVTRQRY (SEQ ID NO: 12)

Glucagon: HSQGTFTSDYSKYLDSRRAQDFVQWLMNT (SEQ ID NO: 13)

GLP-1(7-37, K34R): HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRG (SEQ ID NO: 14)

E.coli expression plasmids (pET22b, Novagen) were prepared such that they encoded the three fusion proteins as specified in Table 2.

Table 2. Fusion proteins encoded by plasmid constructs using pET22b vectors.

Product name	Calculated Molecular mass (without Met)	RL27	HRV14 linker	Peptide
RL27_HRV14_PYY(3-36)	14354.5 Da	SEQ ID NO: 10	SEQ ID NO: 11	SEQ ID NO: 12
RL27_HRV14_Glucagon	13787.1 Da	SEQ ID NO: 10	SEQ ID NO: 11	SEQ ID NO: 13
RL27_HRV14_GLP-1(7-37,K34R)	13688.1 Da	SEQ ID: NO: 10	SEQ ID NO:11	SEQ ID NO: 14

Gene fragments codon-optimized for *E.coli* and spanning the entire fusion proteins were made by gene synthesis and ligated into the cloning site of the pET22b vector using standard cloning techniques (obtained from GenScript).

Expression of model fusion proteins

Expression of RL27_HRV14_PYY(3-36) was done essentially as described for XaaProDAP in Example 2. In short, expression of RL27_HRV14_Glucagon and RL27_HRV14_GLP-1(7-37,K34R) was done as follows: *E.coli* BL21(DE3) was transformed with the plasmid and plated on LB agar plates containing 100 mg/L ampicillin and overnight cultures were dissolved in 10 ml LB medium and used to inoculate 750 ml LB with 50mg/ml Carbenicillin in shaker flasks. Shaker flasks were incubated at 100 rpm at 37°C. When OD600 of 0.4 was reached protein expression was induced by adding 0.3 mM IPTG and cells were harvested by centrifugation following 3 hours incubation at 37°C.

Purification of model fusion proteins

In short, capture of the fusion proteins from supernatants resulting from cell disruption (essentially as described in Example 2)was done by cation exchange chromatography as previously described (WO2008/043847) using a SP FF HiTrap 5 ml (GE Healthcare) column on an AKTA Express at a flowrate of 4 ml/min and the following buffers:
Buffer A: 50mM sodium phosphate, pH 7.0
Buffer B: 50mM sodium phosphate, 1000mM NaCl, pH 7.0

In short, after sample loading, and a washing step, the fusion proteins were eluted from the columns using Buffer B. To increase purity of following capture, the proteins were purified by gel filtration essentially as described in Example 2, but using a S75 16/600

column (GE-Healthcare) for the separation. The purified proteins were evaluated by SDS-PAGE analysis and the correct intact mass was verified by LC-MS. UV280 was used to determine the concentration of the fusion proteins.

Example 4

In order to provide proof of concept, that a combined XaaProDAP activity and picornaviral 3C protease activity can be used to release a maturated protein from a fusion protein, while maintaining the native N-terminal of interest, the two enzymes were tested on the fusion proteins from Example 3.

The purified His-Trx-LLXaaProDAP was prepared as described in Example 2. A purified variant of the HRV14 protease was used for the experiments, which had an N-terminal basic tag (ribosomal L9 from *T. maritima* described previously in WO2008/043847). This protease was named RL9-HRV14 3C and was prepared essentially as described previously, using a SP FF for capture and gel filtration for polishing. The RL9-HRV14 3C protease was identical to the one described in WO2008/043847 and had the following sequence:

MKVILLRDVPKIGKKGEIKEVSDGYARNYLIPRGFAKEYTEGLERAIKHEKEIEKRKKEREREE
SEKILKELKKRTHVVKVKAGEGGKIFGAVTAATVAEEISKTTGLKLDKRWFKLDKPIKELGEY
SLEVSLPGGVKDTIKIRVEREESSSGSSGSSGSSGPNTEFALSLLRKNIMTITTSKGEFTGLGI
HDRVCVIPTHAQPGDDVLVNGQKIRVKDKYKLVDPENINLELTVLTLDRNEKFRDIRGFISED
LEGVDATLVVHSNNFTNTILEVGPVTMAGLINLSSTPTNRMIRYDYATKTGQCGGVLCATGKI
FGIHVGGNGRQGFSAQLKKQYFVEKQ (SEQ ID NO: 15)

LC-MS analysis of enzymatic digests

The combined action of His-Trx-LLXaaProDAP and RL9-HRV14 3C protease was evaluated by mass spectrometry using a MaXis Impact Ultra high resolution time-of-flight (UHR-TOF) mass spectrometer (Bruker Daltonics) equipped with a Dionex UltiMate3000™ liquid chromatometer (Dionex) allowing Diode array measuments at UV215 nm with general settings according to the instructions of the manufacturer. Enzymatic digests were separated on a Waters BEH300 C18 Reversed phase 1.0 X 100 mm column with 1.7µm pore size using a column temperature of 45°C and a flow rate of 0.2 ml/min. The solvents used were are follows

Solvent A: 0.1% formic acid in H2O

Solvent B: 99,9% MeCN, 0.1% formic acid(v/v)

Liquid Chromatography was performed with the following gradient to separate the enzyme digests

Time (min)	%A	%B
0	90	10
2	90	10
10	10	90
11	10	90
12	90	10
13	90	10
14	50	50

The recorded mass spectra were deconvoluted and analysed using the Bruker Compass dataanalysis version 4.1 software (Bruker Daltonics) with adjusted settings for mass ranges (2000-17000 Da) and resolutions (>20.000) according to manufactorers instructions. The UV215 nm chromatogram and total ion count (TIC) chromatograms were evaluated in parallel, to ensure that there was agreement between MS data obtained and UV215 nm traces of the peptides. The experimental determined masses indicated in the following examples refers to the most abundant mass, e.g. the mass of the molecule with the most highly represented isotope distribution, based on the natural abundance of the isotopes of the protein detected. In the following, the mass spectrometry data was obtained with a mass accuracy lower than 100 ppm.

Enzymatic reactions with RL27_HRV14_PYY(3-36) as a fusion protein model substrate

The concentration RL27-HRV14-PYY(3-36) was adjusted to a concentration of 1 mg/ml. Enzymatic reaction were setup in reaction volumes of 100 µl using 1x PBS, pH 7.4 as enzyme reaction buffer. His-Trx-LLXaaProDAP: RL9-HRV14 3C: RL27-HRV14-PYY(3-36) substrate was setup in molar ratios of 1:1:40 or 1:1:20, respectively, and the reactions were carried out at 37°C for 1 or 3 hours (as depicted in Table 3): The enzymatic reactions were stopped by addition >0.5 M AcOH prior to LC-MS analysis

Table 3: Enzymatic reactions using co-treatment of His-Trx-LLXaaProDAP and RL9-HRV14 3C with RL27_HRV14_PYY(3-36) from Example 3. Experimentally determined predominant peaks detected in deconvoluted mass spectra of reactions 1-6 are indicated.

Reaction number	Molar ratio	Incubation (hours)	Predominant detected peaks	Determined molecular masses (Dalton)	Calculated mass (Dalton)	Corresponds to
Reaction 1	1:1:40	1	Peak #1	4050.08	4050.1	PYY(3-36) (SEQ ID NO:12)
			Peak #2	10168.45	10168.4	RL27
			Peak #3	4204.15	4204.1	GP-PYY(3-36)
			Peak #4	14354.59	14354.5	RL27-HRV14-PYY(3-36)
Reaction 2	1:1:40	3	Peak #1	4050.09	4050.1	PYY(3-36) (SEQ ID NO:12)
			Peak#2	10167.48	10168.4	RL27
			Peak#3	4204.16	4204.1	GP-PYY(3-36)
Reaction 3	1:1:20	1	Peak #1	4050.08	4050.1	PYY(3-36) (SEQ ID NO:12)
			Peak #2	10168.41	10168.4	RL27
			Peak#3	4204.14	4204.1	GP-PYY(3-36)
Reaction 4	1:1:20	3	Peak #1	4050.09	4050.1	PYY(3-36) (SEQ ID NO:12)
			Peak#2	10168.41	10168.4	RL27
Reaction 5	1:1:10	1	Peak#1	4050.08	4050.1	PYY(3-36) (SEQ ID NO: 12)
		3	Peak#2	10168.44	10168.4	RL27
Reaction 6	1:1:10	1	Peak#1	4050.09	4050.1	PYY(3-36) (SEQ ID NO: 12)
		3	Peak#2	10167.46	10168.4	RL27

Results with RL27_HRV14_PYY(3-36) as fusion protein model substrate.

Analysis of deconvoluted mass spectra showed that the RL27_HRV14_PYY(3-36) fusion protein (control without enzyme) had a mass of 14354.57 Da. This was in agreement with the calculated mass(14354.5 Da) for the fusion protein without the Initiator Methionine (Fig. 1). Analysis of Reaction 1 (Fig. 2) showed two highly predominant mass peaks. Peak #1 corresponded exactly to mature PYY(3-36) and Peak#2 corresponded to the released RL27 tag. In addition, Peak 3# was observed with ~30 % of the relative intensity of Peak#1 and

corresponded to GP-PYY(3-36) and Peak 4# corresponded to the intact RL27-HRV14-PYY(3-36) fusion protein (Fig. 2). This indicates, incomplete cleavage with both His-Trx-XaaProDAP and the RL9-HRV14 3C protease following one hour incubation at these reaction conditions. Reaction 2 shows that 3 hours of incubation completely removes the full length protein and leaves the released mature PYY(3-36) and the RL27 tag (Fig. 3). The GP-PYY(3-36) peak was still present but with reduced intensity (Fig. 3). At 1:1:20 molar ratio and 1 hour incubation time (Reaction 3) no full length protein could be observed , PYY(3-36) was highly predominant and only very little GP-PYY(3-36) was observed (Fig. 4). Following 3 hours at the same molar ratio (Reaction 4) only fully mature PYY(3-36) was detected, indicating that this cleavage process was optimized (Fig. 5). A doubling in the amount of XaaProDAP compared to Reaction 4, did not cause any unspecific XaaProDAP cleavage as indicated by Reaction 5 and 6 (Fig. 6 & 7). Analysis of the corresponding UV215 nm chromatogram supported the MS analysis. Combined analysis showed that the PYY(3-36) UV215 nm peaks had a retention time of ~6.4 min, the released tag at ~5.9 min. and the intact fusion protein at ~6 min and confirmed, that no fusion protein was left following 1:1:40 at 3 hours.

The data demonstrates that the combined action of XaaProDAP and HRV14 activity can specifically release a fully maturated PYY(3-36) peptide with the correct N-terminal amino acid from a fusion protein and proofs the value of using these two enzymes for processing of fusion proteins.

Example 5

Enzymatic reactions with RL27_HRV14_Glucagon as fusion protein model substrate.

The concentrations of model fusion proteins was adjusted to a concentration of 0.5 mg/ml (in PBS buffer pH 7.4). Three enzymatic reaction were setup in reaction volumes of 55 micro litre using PBS, pH 7.4 as reaction buffer. His-Trx-LLXaaProDAP: RL9-HRV14 3C: RL27_HRV14_Glucagon was setup in molar ratios of either 1:100:2000 or 1:500:10000 as depicted in Table 4. The enzymatic reactions were carried out for 3 hours at 37°C and stopped by addition of >0.5 M AcOH and LC-MS analysis was conducted as described in Example 4.

Results with RL27_HRV14_Glucagon as fusion protein model substrate.

Analysis of deconvoluted mass spectra showed that the RL27_HRV14_Glucagon (control without enzyme) had a molecular mass of 13787.01 Da in agreement with the calculated mass of 13787.1 Da for the fusion protein without a N-terminal Methionine. Two

highly predominant mass peaks were detected upon analysis of Reaction 7. One corresponded to mature Glucagon (#Peak 1) and one of the released RL27 tag (#Peak2). No peaks corresponding to full-length fusion protein was observed indicating complete cleavage of the fusion protein at this reaction condition. One additional peak was observed that had a molecular mass of 3275.5 Da corresponding to Glucagon without the His-Ser N-terminal dipeptide, indicating unspecific cleavage of XaaProDAP, but this peak only had an intensity less than ~3% of the intensity observed for the predominant Glucagon peak.

The same highly predominant peaks were observed for Reaction 8, for which the His-Trx-XaaProDAP was diluted even further compared to the RL9-HRV14 3C protease and fusion protein substrate. In this reaction Peak#3 with the mass 3635.66 Da corresponding to Glucagon with an Gly-Pro N-terminal extension was observed and had an intensity of ~10% of the intensity observed for the predominant Glucagon peak(Peak#2). This indicates that XaaProDAP did not completely remove the GP part of the RL9-relased peptide using the experimental condition in Reaction 8.

The results clearly demonstrated that combined treatment of a HRV14 3C protease and XaaProDAP can successfully process the RL27_HRV14-Glucagon fusion protein to yield fully maturated Glucagon with a correct N-terminal amino acid. The person skilled in the art, knows that the reaction can be optimized to a level which introduces very low contaminant in the form of unspecific degradation products.

Table 4: Enzymatic reactions using co-treatment of His-Trx-LLXaaProDAp and RL9-HRV14 3C with RL27_HRV14_Glucagon from Example 3. Experimentally determined predominant peaks detected in deconvoluted mass spectra of reaction 7 and 8 are indicated.

Reaction number	Molar ratio	Temp. Degrees Celcius	Predominant peaks	Determined molecular masses (Dalton)	Corresponds to
Reaction 7	1:100:2000	37	Peak #1	3481.64	Glucagon (SEQ ID NO:13)
			Peak #2	10167.48	RL27
Reaction 8	1:500:10000	37	Peak #1	3481.64	Glucagon (SEQ ID NO:13)
			Peak#2	10168.45	RL27
			Peak#3	3635.66	GP-Glucagon

Example 6

Enzymatic reactions with RL27_HRV14_GLP-1(7-37,K34R) as fusion protein model substrate.

The concentrations of model fusion proteins was adjusted to a concentration of 0.5 mg/ml (in PBS buffer pH 7.4). The following enzymatic reaction were setup in reaction volumes of 55 micro litre using PBS, pH 7.4 as reaction buffer.

Molar ratios which were evaluated was either 1:100:2000 or 1:500:10000 of His-Trx-LLXaaProDAP: RL9-HRV14 3C: GLP-1(7-37, K34R) as depicted in Table 5. The enzymatic reactions were carried out for 3 hours at 37°C and stopped by addition of >0.5 M AcOH and LC-MS analysis was conducted as described in Example 4.

Table 5: Enzymatic reactions using co-treatment of His-Trx-LLXaaProDAp and RL9-HRV14 3C with RL27_HRV14_GLP-1(7-37,K34R) from Example 3. Experimentally determined predominant peaks detected in deconvoluted mass spectra of reactions 9 and 10 are indicated.

Reaction number	Molar ratio	Temp. oC	Predominant detected peaks	Determined molecular Mass (Dalton)	Corresponds to
Reaction 9	1:100:2000	37	Peak #1	3382.68	GLP-1(7-37,K34R) (SEQ ID NO:14)
		37	Peak #2	10168.41	RL27
Reaction 10	1:500:10000	37	Peak #1	3382.67	GLP-1(7-37, K34R) (SEQ ID NO:14)
		37	Peak#2	10168.41	RL27
		37	Peak#3	3537.76	GP-GLP1(7-37, K34R)

Results with RL27_HRV14_GLP-1(K34R, 7-37) as fusion protein model substrate.

Analysis of deconvoluted mass spectra showed that the RL27_HRV14_GLP-1(7-37, K34R) (control, fusion protein without enzyme) had a mass of 13687.21 Da in agreement with the calculated mass for the fusion protein without N-terminal Methionine (13688.1 Da). Two highly predominant peaks were detected for reaction 9, which correspond to released RL27 tag (Peak #2) and mature GLP-1(7-37,K34R) (Peak#1), respectively. For reaction 10 an additional predominant peak with a mass of 3537.76 Da (Peak #3) and an intensity approximately equal to GLP-1(7-37, K34R) peak#1 was also clearly detected and this mass corresponded to a GP extended version of GLP-1(7-37, K34R). This indicates that the

XaaProDAP amount used in Reaction 10 is not sufficient to obtain full processing. However, as the 3537.76 Da peak was not observed in reaction 9, the our data proofs that a reaction using combined actions of HRV14 and XaaProDAP can be optimized to deliver correctly processed and maturated GLP-1 analogs with the correct N-terminal amino acid and with a minimum amount of unwanted cleavage products.

Example 7

Testing different reaction temperatures for combined action of HRV14 and XaaProDAP

One advantage of the picornaviral 3C proteases is that they are active at low temperature, which is beneficial for industrial process. It was therefore evaluated whether XaaProDAP in combination with HRV14 could work at lower temperatures. The concentrations of model fusion proteins were adjusted to 0.5 mg/ml (in PBS buffer pH 7.4). The following enzymatic reaction was setup in reaction volumes of 30 micro litre using PBS, pH 7.4 as reaction buffer.

His-Trx-LLXaaProDAP: RL9-HRV14 3C: RL27_HRV14_Glucagon

His-Trx-LLXaaProDAP: RL9-HRV14:RL27_HRV14:GLP-1(7-37,K34R)

Molar ratios of either 1:1:20 or 1:100:2000 was evaluated at either 4°C,10°C or 30°C as described in Table 6. The enzymatic reactions were incubated overnight and stopped by addition of >0.5 M AcOH and LC-MS analysis was conducted as described in Example 4.

Results with RL27_HRV14_Glucagon as fusion protein model substrate.

The analysis of deconvoluted mass spectra from reaction 11 to 16 showed that all reactions did not contain any remaining fusion protein and the released RL27 tag was observed in all spectra. This indicates that the HRV14 3C protease completely removed the RL27 tag in these conditions. In reaction 11 two predominant peaks corresponding to Glucagon (Peak#1) and Glucagon(5-29)(Peak#2) was observed indicating that mature Glucagon could be released, but also, that the XaaProDAP enzyme can remove additionally two dipeptides. Using the same molar ratio, at higher temperature resulted in more of the degraded Glucagon product due to unspecific cleavage. However, upon 100-fold less XaaProDAP added (Reaction 12), the degraded Glucagon product was removed and only the intact mature Glucagon peak(Peak#1) was observed. Increase of the temperature to 10°C, resulted in the same profile (Reaction 14), whereas 30°C incubation resulted in a small additional fragment corresponding to Glucagon(3-29) (Reaction 16). All together, these data shows that the enzymatic reaction comprising XaaProDAP and HRV14 3C protease can be

optimized to release mature correctly processed Glucagon from a fusion protein, with minimum contamination from unspecific fragments even at low temperatures.

Results with RL27_HRV14_GLP-1(7-37,K34R) as fusion protein model substrate.

As for Glucagon, the analysis of deconvoluted mass spectra from reaction 17 to 22 indicated that HRV14 3C protease completely removed the RL27 tag. In reaction 17, the clearly predominant peak observed corresponded to mature GLP-1(7-37,K34R)(Peak#2). However, GLP-1(9-37,K34R) (Peak#1) derived from unspecific cleavage of XaaProDAP was also observed, but with significant lower intensity on the mass spectra. As expected, Reaction 19 and 21 shows, that increase of temperature increases the amount of this degradation product when using the 1:1:20 molar ratio. However, analysis of Reaction 18, shows that a 100-fold reduction of XaaProDAP compared to substrate completely removes the unspecific GLP-1(9-37,K34R) fragment at 4°C. However, a smaller fragment corresponding to GP-GLP-1(7-37,K34R) was observed. Analysis of reaction 20 and 22, shows that a combination of 1:100:2000 molar ratio and 10°C overnight resulted in efficient reduction of both unwanted peaks to very low levels. This shows that the ratio of XaaProDAP/HRV14 3C and the reaction temperature are parameters that can easily be optimized to allow successful processing and limit the amount of contaminants to an acceptable minimum even at low temperatures.

Table 6: Enzymatic reactions using co-treatment of His-Trx-LLXaaProDAp and RL9-HRV14 3C with RL27_HRV14_Glucagon or RL27_HRV14_GLP-1(7-37,K34R) from Example 3 at different temperatures and molar ratios. Experimentally determined predominant peaks detected in deconvoluted mass spectra of reaction 11 to 22

Reaction number	Molar ratio	Temp. oC	Predominant detected peaks	Determined Molecular mass (Dalton)	Corresponds to
Reaction 11	1:1:20	4	Peak#1	3482.67	Glucagon (SEQ ID NO: 13)
			Peak#2	3072.49	Glucagon (5-29)
			Peak#3	10167.54	RL27
Reaction 12	1:100:2000	4	Peak#1	3481.66	Glucagon (SEQ ID NO: 13)
			Peak#2	10167.54	RL27
Reaction 13	1:1:20	10	Peak#1	3482.64	Glucagon (SEQ ID NO: 13)
			Peak#2	3072.47	Glucagon (5-29)
			Peak#3	10167.46	RL27

Reaction 14	1:100:2000	10	Peak#1	3481.64	Glucagon (SEQ ID NO: 13)
			Peak#2	10167.46	RL27
Reaction 15	1:1:20	30	Peak#1	3072.47	Glucagon(5-29)
			Peak#2	10168.48	RL27
Reaction 16	1:100:2000	30	Peak#1	3257.55	Glucagon(3-29)
			Peak#2	3481.65	Glucagon (SEQ ID NO: 13)
			Peak#3	10168.5	RL27
Reaction 17	1:1:20	4	Peak#1	3174.61	GLP-1(9-37,K34R)
			Peak#2	3382.71	GLP-1(7-37,K34R) (SEQ ID NO: 14)
			Peak#3	10168.53	RL27
Reaction 18	1:100:2000	4	Peak#1	3382.72	GLP-1(7-37,K34R) (SEQ ID NO:14)
			Peak#2	3537.79	GP-GLP-1(K34R,7-37)
			Peak#3	10167.47	RL27
Reaction 19	1:1:20	10	Peak#1	3174.61	GLP-1(9-37,K34R)
			Peak#2	3382.71	GLP-1(7-37,K34R) (SEQ ID NO:14)
			Peak#3	10167.48	RL27
Reaction 20	1:100:2000	10	Peak#1	3382.69	GLP-1(7-37,K34R) (SEQ ID NO: 14)
			Peak#2	10168.45	RL27
Reaction 21	1:1:20	30	Peak#1	3174.61	GLP-1(9-37,K34R)
			Peak#2	3382.7	GLP-1(7-37,K34R) (SEQ ID NO:14)
			Peak#3	10168.5	RL27
Reaction 22	1:100:2000	30	Peak#1	3382.7	GLP-1(7-37,K34R) (SEQ ID NO:14)
			Peak#2	10168.47	RL27 (SEQ ID NO: 10)

While certain features of the invention have been illustrated and described herein, many modifications, substitutions, changes, and equivalents will now occur to those of ordinary skill in the art. It is, therefore, to be understood that the appended claims are intended to cover all such modifications and changes as fall within the true spirit of the invention.

CLAIMS

1. Method for release of a matured protein from a fusion protein by enzymatic treatment using a picornaviral 3C protease and a Xaa-Pro-dipeptidyl aminopeptidase (XaaProDAP).
2. The method according to claim 1 wherein said picornaviral 3C protease natively has a Gly-Pro in the N-terminal.
3. The method according to any of claims 1-2 wherein said picornaviral 3C protease is HRV14 3C or a functional variant thereof.
4. The method according to any of claims 1-2 wherein said picornaviral 3C protease is CVB3 3C or a functional variant thereof.
5. The method according to any of claims 1-4 wherein said XaaProDAP is an EC 3.4.14.11 enzyme or a functional variant thereof.
6. The method according to any of claims 1-5 wherein said XaaProDAP is an enzyme from *Lactococcus sp.* or a functional analogue derived from an enzyme from *Lactococcus sp.*
7. The method according to any of claims 1-5 wherein said XaaProDAP is an enzyme from *Streptococcus suis* or a functional analogue derived from an enzyme from *Streptococcus suis.*
8. The method according to any of claims 1-7 wherein said fusion protein has the structure :

   Tag sequence – Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-QGP - Protein of interest,

   wherein

   Tag sequence is a fusion partner protein connected at its C-terminal,

   Xaa1, Xaa2, Xaa3, Xaa4 and Xaa5 are independently selected from the group consisting of the naturally occurring amino acid residues, and

   Protein of interest is a protein connected at its N-terminal and comprising said matured protein and optionally an N-terminal extension which can be cleaved by XaaProDAP or by picornaviral 3C protease and XaaProDAP.
9. The method according to claim 8, wherein Xaa2-Xaa3-Xaa4-Xaa5 is EVLF, ETLF or FAQT.
10. The method according to any of claims 1-9 wherein said enzymatic treatment is done by simultaneously contacting said fusion protein with said picornaviral 3C protease and said XaaProDAP.
11. The method according to any of claims 1-10 wherein said XaaProDAP is added to a reaction mixture containing the picornaviral 3C protease, fusion protein and de-

tagged protein once the majority of the fusion protein has been cleaved by picornaviral 3C protease into the de-tagged protein.

12. The method according to any of claims 1-11 wherein said enzymatic treatment is done at a temperature in the range from about 0 °C to about 15 °C.
13. The method according to any of claims 1-12 wherein said enzymatic treatment is done at a pH in the range from about pH 7.5 to about pH 8.5 and at a temperature in the range from about 4 °C to about 10 °C.
14. The method according to any of claims 1-10 or 12-13 wherein said treatment using a picornaviral 3C protease and a XaaProDAP is done by using a fusion enzyme comprising a picornaviral 3C protease and a XaaProDAP.
15. The method according to any of claims 1-14 wherein said matured protein is GLP-1, Glucagon, PYY(3-36), amylin or a functional variant thereof.

Figure 1 / 7

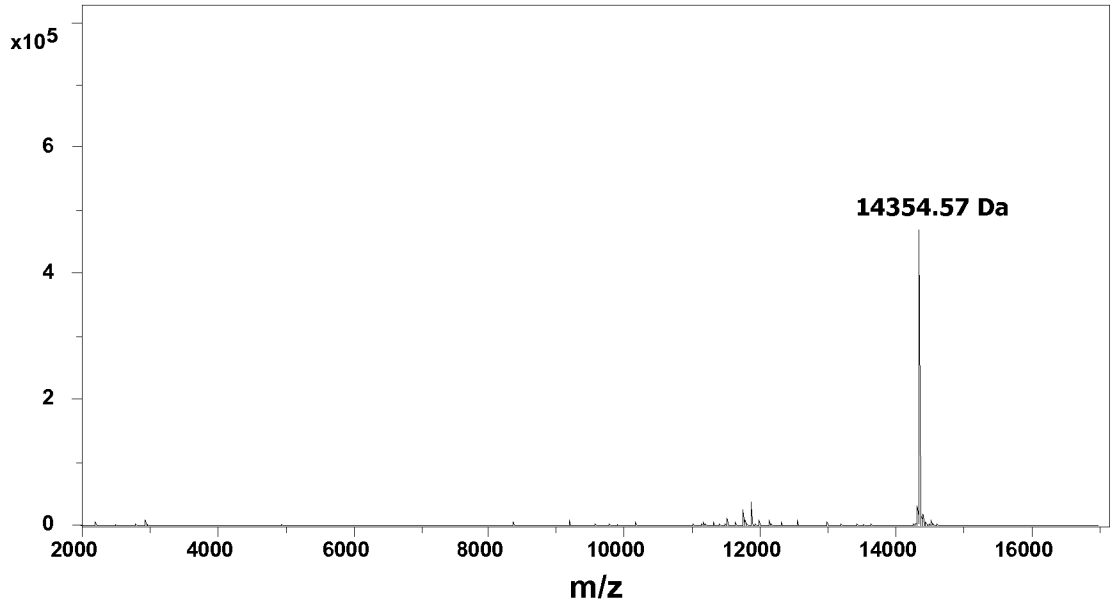

x10^5
6
4
2
0
14354.57 Da
2000
4000
6000
8000
10000
12000
14000
16000
m/z

Figure 2 / 7

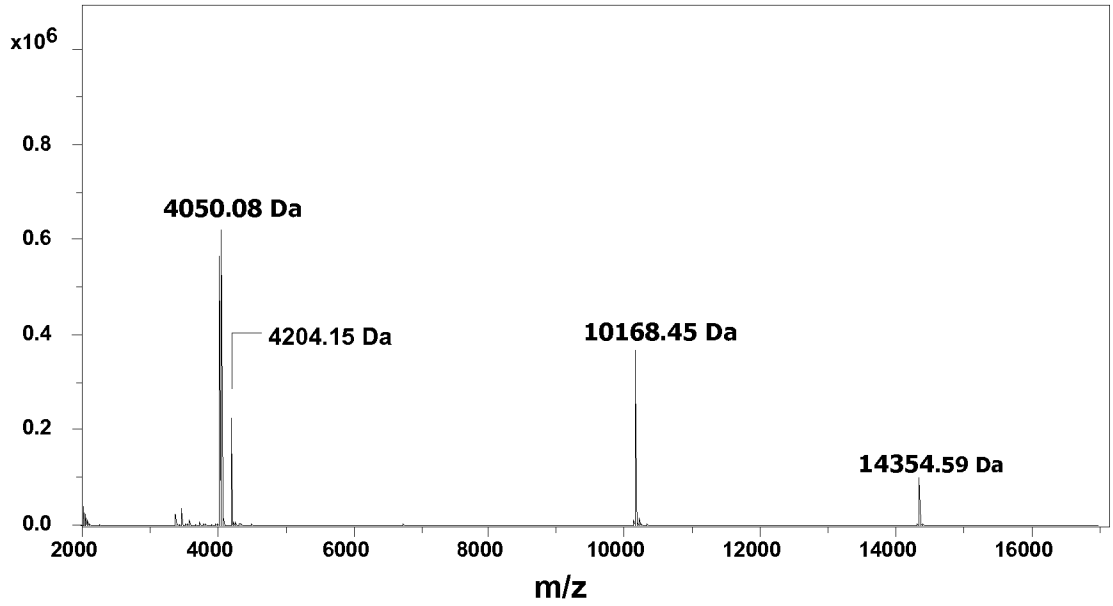


Figure 3 / 7

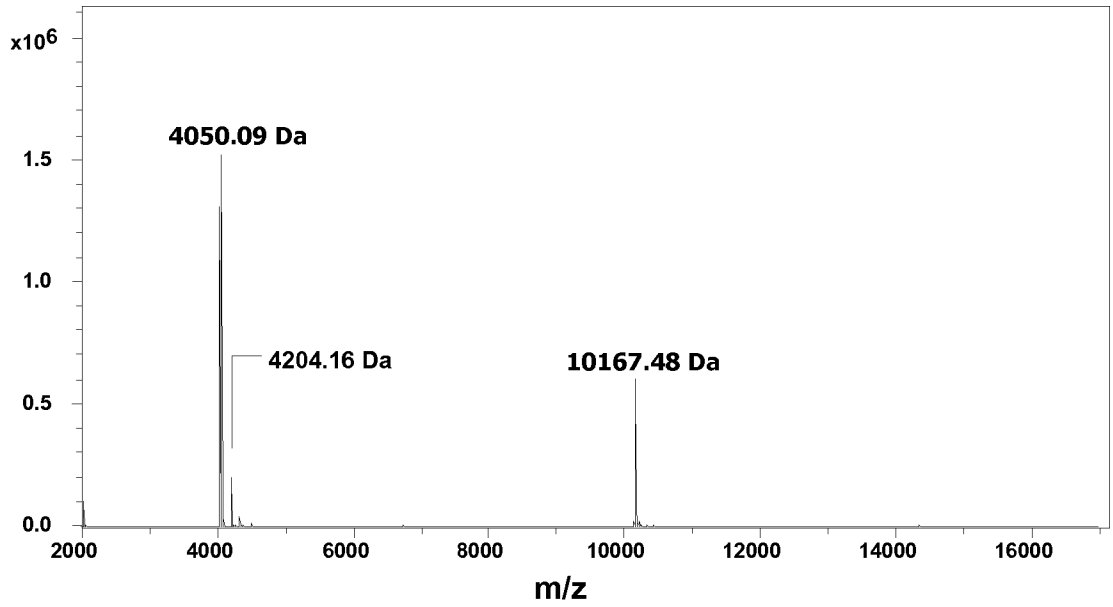


Figure 4 / 7

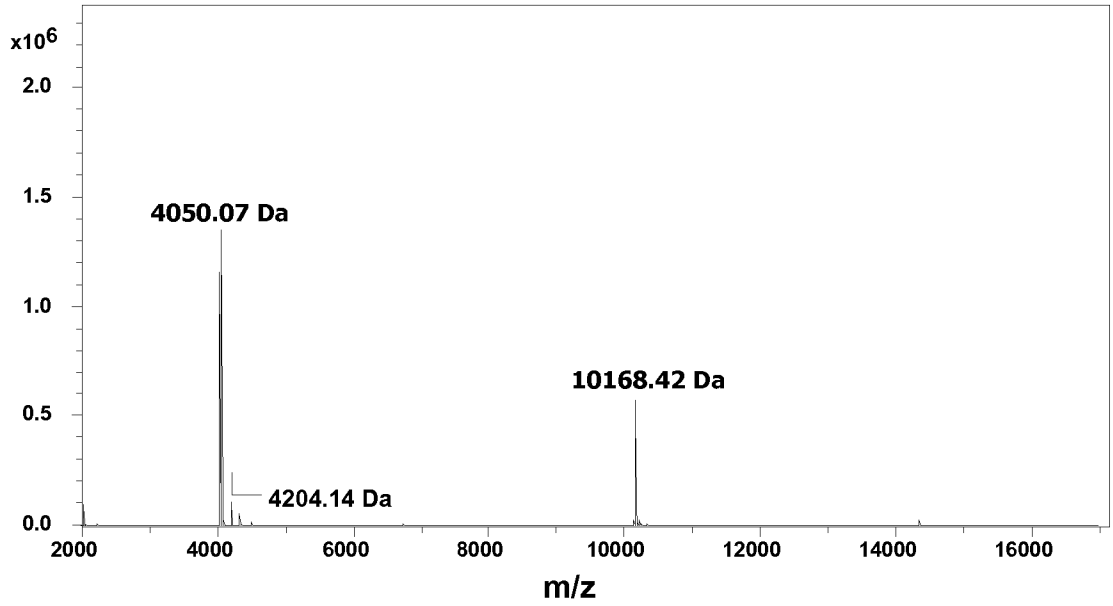


Figure 5 / 7

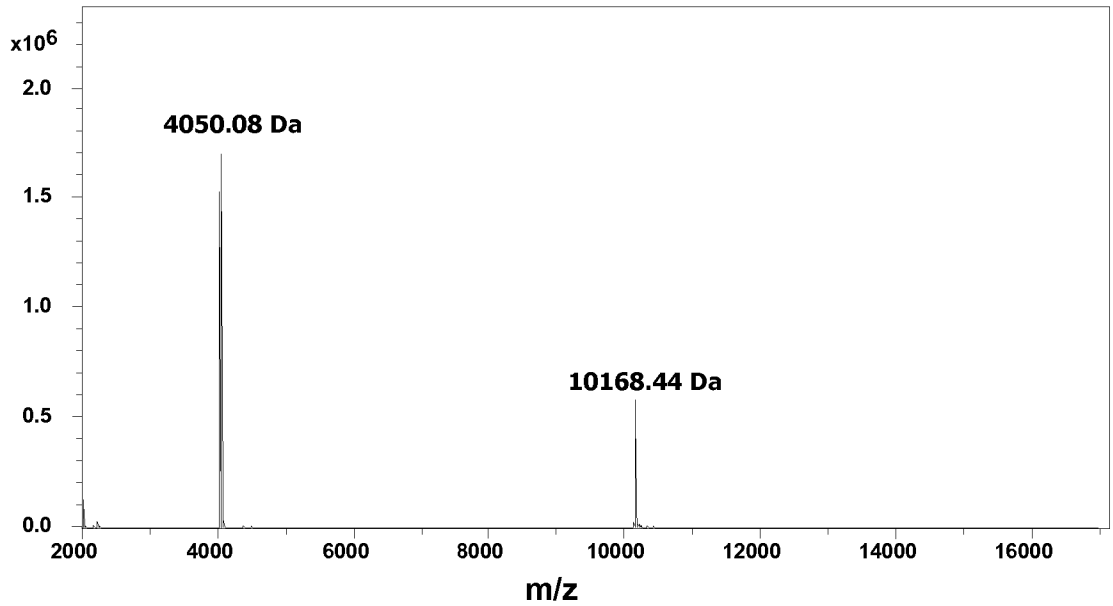


Figure 6 / 7

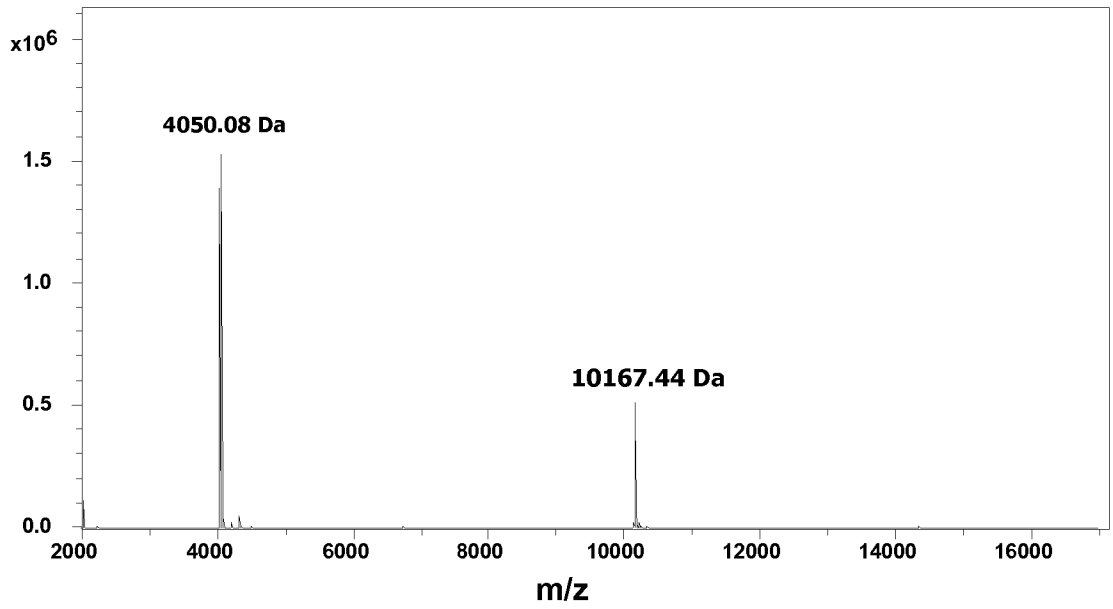
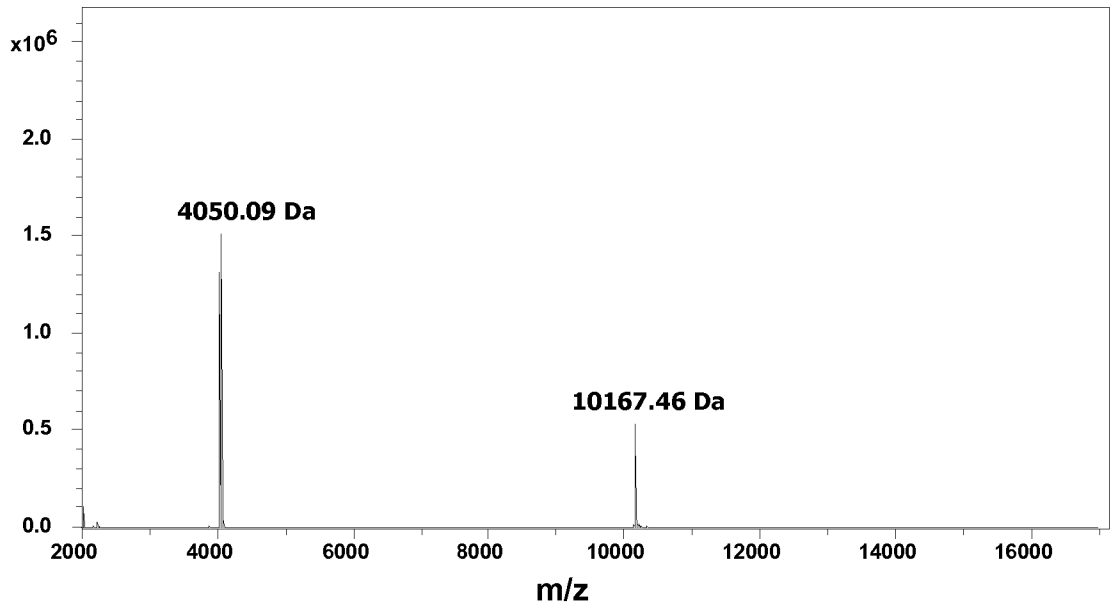


Figure 7 / 7



INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2014/060669

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. (means)

   ☐ on paper

   ☒ in electronic form

   b. (time)

   ☐ in the international application as filed

   ☒ together with the international application in electronic form

   ☐ subsequently to this Authority for the purpose of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/060669

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12P21/06 C12N15/62
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)
C12P C12N C07K

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data, BIOSIS, Sequence Search, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 91/11520 A1 (MAX PLANCK GESELLSCHAFT [DE]) 8 August 1991 (1991-08-08) page 21, paragraph 1 - page 22, paragraph 1 claims 21,22 ----- -/--	1-15

[X] Further documents are listed in the continuation of Box C.

[X] 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	Date of mailing of the international search report
25 August 2014	29/08/2014

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Schwachtgen, J

4

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DAVID S WAUGH: "An overview of enzymatic reagents for the removal of affinity tags", PROTEIN EXPRESSION AND PURIFICATION, vol. 80, no. 2, 19 August 2011 (2011-08-19), pages 283-293, XP028312708, ISSN: 1046-5928, DOI: 10.1016/J.PEP.2011.08.005 [retrieved on 2011-08-19] cited in the application page 287, column 1, paragraph 3 - page 288, column 1, paragraph 1	1-15
Y	US 4 769 326 A (RUTTER WILLIAM J [US]) 6 September 1988 (1988-09-06) column 15, line 30 - line 36; claim 7	1-15
Y	XIN M ET AL: "An X-Prolyl Dipeptidyl Aminopeptidase from Lactococcus lactis: Cloning, Expression in Escherichia coli, and Application for Removal of N-Terminal Pro-Pro from Recombinant Proteins", PROTEIN EXPRESSION AND PURIFICATION, ACADEMIC PRESS, SAN DIEGO, CA, vol. 24, no. 3, 1 April 2002 (2002-04-01), pages 530-538, XP004445164, ISSN: 1046-5928, DOI: 10.1006/PREP.2001.1579 cited in the application abstract	6
Y	M.-C. JOBIN ET AL: "Cloning, Purification, and Enzymatic Properties of Dipeptidyl Peptidase IV from the Swine Pathogen Streptococcus suis", JOURNAL OF BACTERIOLOGY, vol. 187, no. 2, 15 January 2005 (2005-01-15), pages 795-799, XP055134888, ISSN: 0021-9193, DOI: 10.1128/JB.187.2.795-799.2005 abstract	7
Y	C.-C. LEE ET AL: "Structural Basis of Inhibition Specificities of 3C and 3C-like Proteases by Zinc-coordinating and Peptidomimetic Compounds", JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 284, no. 12, 20 March 2009 (2009-03-20), pages 7646-7655, XP055134941, ISSN: 0021-9258, DOI: 10.1074/jbc.M807947200 abstract	7

-/--

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KREIL ET AL: "Processing of precursors by dipeptidylaminopeptidases: a case of molecular ticketing", TRENDS IN BIOCHEMICAL SCIENCES, ELSEVIER, HAYWARDS, GB, vol. 15, no. 1, 1 January 1990 (1990-01-01), pages 23-26, XP025442469, ISSN: 0968-0004, DOI: 10.1016/0968-0004(90)90126-V [retrieved on 1990-01-01] table 1 -----	1-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2014/060669

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9111520 A1	08-08-1991	AT 124456 T	15-07-1995
		AT 219518 T	15-07-2002
		AU 638309 B2	24-06-1993
		CA 2074943 A1	04-08-1991
		CS 9100241 A2	13-08-1991
		DK 0513073 T3	30-10-1995
		EP 0513073 A1	19-11-1992
		EP 0602688 A1	22-06-1994
		ES 2076521 T3	01-11-1995
		ES 2177536 T3	16-12-2002
		FI 923463 A	31-07-1992
		HU T63195 A	28-07-1993
		HU 217103 B	29-11-1999
		IE 910358 A1	14-08-1991
		JP 2766621 B2	18-06-1998
		JP H0789952 B2	04-10-1995
		JP H05501360 A	18-03-1993
		JP H07278195 A	24-10-1995
		LV 10309 A	20-10-1994
		NO 923010 A	14-09-1992
		NZ 236819 A	27-07-1993
		PT 96658 A	31-10-1991
		US 5427927 A	27-06-1995
		WO 9111520 A1	08-08-1991
US 4769326 A	06-09-1988	NONE	