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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 June 2013 (27.06.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/092855 Al

(51) **International Patent Classification:**
C12N 9/64 (2006.01) ***C12N 15/62*** (2006.01)

(21) **International Application Number:**
PCT/EP2012/076372

(22) **International Filing Date:**
20 December 2012 (20. 12.2012)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
PCT/CN201 1/002169
23 December 201 1 (23. 12.201 1) CN

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

(72) **Inventors: WÖLDICKE, Helle Fabricius;** Novo Nordisk A/S, Novo Alle, DK-2880 Bagsvaerd (DK). **ZHANG, Xujia;** Novo Nordisk R/D Center China Internal, 1st Floor, Tower B no. 29, Life Science Park Road, Beijing, Changping District 102206 (CN). **LIU, Yun;** Novo Nordisk R/D Center China Internal, 1st Floor, Tower B no. 29, Life Science Park Road, Beijing, Changping 102206 (CN). **WEI-WEI, Tong;** Novo Nordisk R/D Center China Internal, 1st Floor, Tower B no. 29, Life Science Park Road, Changping District, Beijing, Changping 102206 (CN).

(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, IS, JP, KE, KG, KM, 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, 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, ML, MR, NE, SN, TD, TG).

Published:

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

— *with sequence listing part of description (Rule 5.2(a))*

WO 2013/092855 A1

(54) Title: MODIFIED ENTEROKMASE LIGHT CHAIN

(57) Abstract: The present invention is related to novel mammalian enterokinase analogues such as mammalian enterokinase light chain analogues and methods of making such. Also described herein is a method for cleaving proteins having an enterokinase cleavage site.

MODIFIED ENTEROKINASE LIGHT CHAIN

TECHNICAL FIELD

The present invention is related to novel mammalian enterokinase analogues, methods of making such and the use of said mammalian enterokinase analogues for cleaving proteins having an enterokinase cleavage site.

INCORPORATION-BY-REFERENCE OF THE SEQUENCE LISTING

The Sequence Listing, entitled "SEQUENCE LISTING", is 10 kb, was created on 17-Dec-2012 and is incorporated herein by reference.

BACKGROUND

The serine protease enterokinase (in short enterokinase or EK), also known as enteropeptidase, is a heterodimeric glycoprotein, a mammalian enzyme catalyzing the conversion of trypsinogen into active trypsin. Enterokinase has preference for the substrate sequence Asp-Asp-Asp-Asp-Lys ($(Asp)_4$-Lys, DDDDK), where it selectively cleaves after lysine. Enterokinase isolated from bovine duodenal mucosa exhibits a molecular weight (MW) of 150,000 and a carbohydrate content of 35 percent. The enzyme is comprised of a heavy chain (MW-1 15,000) and a disulfide-linked light chain (MW~35,000) (Liepnieks et al., J. Biol. Chem., 254(5): 1677-1 683 (1979)). The function of the heavy chain is to anchor the enzyme to the mucosal membrane. The light chain acts as the catalytic subunit.

In E.coli many mammalian proteins are expressed as fusion proteins, which have to be cleaved to release the mature, active protein. For that purpose a processing enzyme is needed, preferably one which cleaves directly at the junction leaving no extra amino acids on the product. Enterokinase is such an enzyme, and much effort has been made to establish a recombinant process to obtain enterokinase or enterokinase analogues in E.coli. However, the results so far have been rather poor: Available commercial products are expensive and of low specific activity, due to inefficient renaturation of precipitated EK or inefficient secretion of soluble EK.

A process in E.coli aiming at a soluble EK product leads to a mixture of soluble and insoluble protein, requiring 2 routes of purification, expensive affinity columns and low yields altogether. In order to get a uniform product, the EK has to be produced as insoluble material in inclusion bodies. They are easy to isolate but challenging to renature in satisfactory yields, due to possible aggregation of the protein.

An object of the invention is to obtain a mammalian enterokinase analogue with improved properties.

SUMMARY

The present invention is related to mammalian enterokinase analogues mutated in appropriate sites. One or more substitutions of an enterokinase analogue of the invention may e.g. be from hydrophobic to hydrophilic, charged amino acids relative to the amino acids in the parent (wild type) mammalian enterokinase.

In one aspect of the invention, a bovine enterokinase light chain analogue is obtained which comprises at least one substitution in position 134 and/or 135 from hydrophobic to a hydrophilic charged amino acid(s). In one aspect, the bovine enterokinase light chain analogue according to the invention further comprises a substitution in position 112.

The invention is also related to a method for obtaining improved solubility in a renaturation process of an enterokinase light chain analogue. In one aspect, the method comprises the step of mutating one or more hydrophobic amino acids of wild type bovine enterokinase light chain to hydrophilic amino acids and optionally mutating other amino acids of wild type bovine enterokinase light chain, wherein the hydrophobic amino acids subject to mutation are present on the surface of folded wild type bovine enterokinase light chain.

In one aspect, the invention provides an improved production process for obtaining mammalian enterokinase analogues. Also or alternatively, in a second aspect, the invention provides an improved production process resulting in improved production yield.

In one aspect of the invention, the method for production of a bovine enterokinase light chain analogue comprises the steps:

a) culturing the host cells in a growth medium comprising inducer, wherein the host cells comprise a polynucleotide sequence encoding the amino acid sequence of the enterokinase light chain analogue;

b) recovering the cells with enterokinase light chain analogue in inclusion bodies

c) solubilizing and refolding the enterokinase light chain analogue; and

d) purifying the enterokinase light chain analogue.

In one aspect, the invention provides a method for recombinantly producing a peptide or protein in a bacterial or yeast host cell. In one aspect the method comprises:

a) expressing in yeast or bacteria a fusion protein comprising the peptide or protein to be produced;

b) cleaving the fusion protein with a bovine enterokinase light chain analogue according to any one of aspects 1-9; and

c) isolating the produced peptide or protein.

The invention may also solve further problems that will be apparent from the disclosure of the exemplary embodiments.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1: Dependence of both Trx-EK$_L$ (A) and Trx-EK$_{LM}$ (B) expression upon induction time. **M:** Marker; **BI:** Before Induction; **I2, 13, I4** and **I6** represent induction time (hr) by IPTG, respectively; 15% gel; Fermentation defined medium (FDM) used.

Figure 2: Flowchart for EK purification

Figure 3: Figure 3: % refolding yield (fig. 3A) and the amount of purified EK$_L$ and EK$_{LM}$ in 1L refolding buffer (mg, fig. 3B) as a function of the Trx-linker-EK$_L$ and Trx-linker-EK|_M concentration during refolding. A/Δ: Trx- linker-EK$_L$, 1mg/ml inclusion body (IB); **»/o:** Trx- linker-EK$_{LM}$, 6mg/ml IB; ◆/◇: Trx- linker-EK$_{LM}$, 4mg/ml IB. 1.3g cell pellets of Trx-linker-EK|_M or Trx-linker-EK|_ were lysed and inclusion bodies were solublized to different concentrations, i.e. 1mg/ml for Trx-linker-EK$_L$, 4mg/ml or 6mg/ml for Trx-linker-EK$_{LM}$ in buffer containing 20mM Tris, 8M urea, pH8.0, 20mM DTT. After dilution to the concentrations as indicated in the refolding buffer containing 20mM Tris, 1M Urea, 1mM GSSG, 3mM GSH, pH 8.3 and incubation at 20°C for 24hrs, the EK$_{LM}$/ EK$_L$ was subjected to purification by Q HP chromatography as described in Experiments.

Figure 4: The refolding yield of Trx-EK$_L$ increases with incubation time. 1.3g cell pellets of Trx-EK$_L$ were lysed and inclusion bodies were solublized to 1.6mg/ml in buffer containing 20mM Tris, 8M urea, pH8.0, 20mM DTT. After 100 fold dilution in the refolding buffer containing 20mM Tris, 1M Urea, 1mM GSSG, 3mM GSH, pH 8.3 and incubated at 20°C for 24hrs or 48hrs, respectively, the enzyme activity was assayed as described in Experiments.

Figure 5: Dependence of the refolding yield upon urea concentration. 1.3g cell pellets of Trx-EK$_L$ were lysed and inclusion bodies were solubilized to 1.6mg/ml in buffer containing 20mM Tris, 8M urea, pH8.0, 20mM DTT. After 100 fold dilution in the refolding buffer containing 20mM Tris, 1mM GSSG, 3mM GSH, pH 8.3 and OmM, 0.5mM, 1mM, 1.5mM or 2mM urea, respectively, and incubated at 20°C for 24hrs, the enzyme activity was assayed as described in Experiments.

Figure 6: Dependence of the refolding yield with redox GSSG/GSH ratio. 1.3g cell pellets of Trx-EK$_L$ were lysed and inclusion bodies were solublized to 1.6mg/ml in buffer

containing 20mM Tris, 8M urea, pH8.0, 20mM DTT. After 100 fold dilution in the refolding buffer containing 20mM Tris, 1M Urea, pH 8.3 and GSSG/GSH as indicated, and incubated at 20°C for 24hrs, the enzyme activity was assayed as described in Experiments.

Figure 7: Purification of EK_LM by Q HP chromatography. **(A):** A chromatogram. EK_LM was eluted by sodium gradient, as shown in P2. The fractions containing EK enzymatic activity were indicated. **(B):** SDS-PAGE of EK_LM at each step under reduced conditions. **EK_{LM}:** High purity EK_{LM} (>90%) obtained from further purification of P2 by Hydrophobic Interaction Chromatography; **M:** Marker, **BI:** Before Induction, **Total:** Total lysates; **Sup:** Supernatant after lysis of cells; **IB:** Inclusion bodies subjected to refolding and purification; **App:** Samples applied to Q HP column after refolding and auto-activation; **P1, P2** and **P3** represent the pooled fractions of each peak indicated in Fig 7A. **(C):** Enzymatic activity. Δ: P 1. 1ul of sample added to 100ul of reaction buffer; ·: P2. After 5 fold dilution of P2, 1ul of diluted sample added to 100ul of reaction buffer ; **o:** P3. 1ul of sample added to 100ul of reaction buffer ; ■: Blank. 1ul of buffer (20mM Tris, pH 8.0) added to 100ul of reaction buffer. 1.3g cell pellets of Trx-EK_LM were lysed and inclusion bodies were solubilized to 4mg/ml in buffer containing 20mM Tris, 8M urea, pH8.0, 20mM DTT. After 80 fold dilution into refolding buffer containing 20mM Tris, 1M Urea, 1mM GSSG, 3mM GSH, pH 8.3 and incubated at 20°C for 24hrs, the EK_LM was subjected to purification by Q HP chromatography as described in Experiments.

Figure 8: Similar specific enzymatic activity between EK_L and EK_LM- 25EU of purified EK_L and EK_LM was loaded on SDS-PAGE.

Figure 9: EK_LM is stable for at least 3 month at -80°C or 4°C. The purified EK_LM as described in Experiments was aliquoted and stored at -80°C or 4°C. After 3 month, 5μg of EK|_M from each temperature was loaded on SDS-PAGE under reduced and non-reduced condition, and compared with freshly purified EK_LM (Fresh).

Figure 10: Comparison of amino acid sequences trxEK_LM (SEQ ID No: 9) and trx-linker-EK|_M (SEQ ID No: 8). In trx-linker-EK_LM the spacer between trx and EKLM is 37 amino acids longer than in trxEK_LM.

Figure 11: The refolding efficiency of Trx-linker-EK_LM increases with PEG1000 or cyclodextrin added into the refolding buffer. The inclusion body was solubilized into 7.3 mg/ml and diluted with the ratio of 1 to 20 into the refolding buffer. The final concentration of PEG1000 and cyclodextrin in the refolding buffer is 1% and 1.5% respectively.

DESCRIPTION

The present invention is related to mammalian enterokinase analogues mutated in appropriate sites. One or more substitutions of an enterokinase analogue of the invention may e.g. be from hydrophobic to hydrophilic, charged amino acids relative to the amino acids in the parent (wild type) mammalian enterokinase. In one aspect, one or more substitutions of a mammalian enterokinase analogue of the invention is from hydrophobic to hydrophilic, charged amino acids relative to the amino acids in wild type bovine enterokinase. In one aspect, the hydrophobic amino acids subject to mutation are present on the surface of folded wild type mammalian enterokinase light chain such as folded wild type bovine enterokinase light chain.

The wild type bovine enterokinase light chain generally exhibits good activity in the presence of various detergents and denaturants over a wide pH range (4.5-9.5) and temperature range (4-45 °C). Therefore, the enterokinase light chain as a powerful tool has been used in biotechnology for the in vitro cleavage of fusion proteins.

However, the complicated production processes and low production yield extracted from animals, such as porcine and bovine, has set a limitation to EK application in biotechnology. Recently, recombinant enterokinase light chain in *E.coli* has been obtained by secretion of active enterokinase light chain or by intracellular accumulation of inclusion bodies of inactive enterokinase light chain, refolding and activation. Moreover, it has been demonstrated that substitution of Cys1 12 to Ala of bovine enterokinase light chain enhanced the enzymatic activity, presumably due to facilitated refolding. Cys1 12 links the light chain to the heavy chain in the holoenzyme and is not an essential part of the light chain.

In one aspect of the invention the mammalian enterokinase analogue is a mammalian enterokinase light chain analogue such as a bovine enterokinase light chain analogue. In one aspect of the invention the mammalian enterokinase analogue is a bovine enterokinase light chain analogue. In one aspect according to the invention the bovine light chain analogue comprises substitution(s) in position 134 and/or position 135. In one aspect the bovine enterokinase light chain analogue comprises substitutions in positions 112, 134 and/or 135. In one aspect, the bovine enterokinase light chain analogue comprises at least two substitutions. In one aspect, the bovine enterokinase light chain analogue comprises at least three substitutions. In one aspect the bovine enterokinase light chain analogue comprises substitutions in positions 112, 134 and 135. In one aspect the bovine enterokinase light chain analogue comprises the substitutions C112A, L134K and I135K.

Novel bovine enterokinase light chain analogues of the invention include those having the primary structural conformation (i. e., amino acid sequence) of the light chain of

wild type bovine enterokinase. The light chain of wild type bovine enterokinase has the sequence substantially as set forth in SEQ ID NO:1 .

```
  1  IVGGSDSREG AWPWVVALYF DDQQVCGASL VSRDWLVSAA HCVYGRNMEP
 51  SKWKAVLGLH MASNLTSPQI ETRLIDQIVI NPHYNKRRKN NDIAMMHLEM
101  KVNYTDYIQP ICLPEENQVF PPGRICSIAG WGALIYQGST ADVLQEADVP
151  LLSNEKCQQQ MPEYNITENM VCAGYEAGGV DSCQGDSGGP LMCQENNRWL
201  LAGVTSFGYQ CALPNRPGVY ARVPRFTEWI QSFLH
```

SEQ ID NO: 1

According to an aspect bovine enterokinase light chain analogues of the invention have enterokinase protease activity. Antibodies to such proteases are also available.

The bovine enterokinase light chain analogue described by the present invention, maintains enterokinase wild type protease activity for use as a restriction proteases to specifically cleave fusion proteins.

The term "bovine enterokinase" as used herein means the bovine enterokinase enzyme whose structure and properties are well-known. Mammalian enterokinases are carbohydrate containing heterodimers with a heavy chain of 650-800 amino acids and a catalytic light chain of around 235 amino acids and an overall homology of 75-80% (Liepniecks et al., J. Biol. Chem. 254 , 1677 (1979), Matsushima et al., J.Biol. Chem. 269 (31), 19976 (1994), Kitamoto et al., Biochemistry 34, 4562 (1995) for bovine, porcine and human enterokinase, respectively). Further studies of the catalytic light chains are reported in LaVallie et al., J. Biol. Chem. 268 (31), 2331 1-17 (1993) on the bovine EK and in Matsushima et al., J. Biochem. 125, 947, (1999) on the porcine EK.

The term "bovine enterokinase light chain" as used herein means the light chain of bovine enterokinase having 4 disulphide bridges. The bovine enterokinase light chain is e.g. described in LaVallie et al, above.

When used herein the term "surface" in connection with amino acids present on the surface of folded wild type bovine enterokinase light chain means amino acids identified as present on the surface of the folded wild type bovine enterokinase light chain on a 3D structure as e.g. described in Mod Base P 98072.

"An enterokinase light chain" according to the invention is herein to be understood as bovine enterokinase light chain or an enterokinase light chain from another species such as porcine or human enterokinase light chain.

The term "enterokinase light chain peptide" as used herein means a peptide which is either bovine enterokinase light chain or an analog or a derivative thereof with enterokinase activity.

As used herein, enterokinase activity means the capability of cleaving peptide or protein substrates at a specific site; for protein substrates, this is generally following the sequence $(Asp)_4$-Lys, or a similar sequence such as those described in Light et al., Anal. Biochem. 106: 199(1 980); (a cluster of negatively charged amino acids followed by a positively charged amino acid). Typically, such activity is measured by activation of trypsinogen by cleaving the N-terminal propeptide (containing $(Asp)_4$-Lys) with the enterokinase or enterokinase analogue and subsequently assaying the amount of active trypsin generated using tosyl-arginine-methylester (TAME). Alternatively, enterokinase activity can be measured directly by incubating the enzyme with the peptide substrate Gly $(Asp)_4$-Lys-ss-naphthylamide and measuring the increase in fluorescence (excitation at 337 nm, emission at 420 nm) generated by cleavage and release of the ss-NA (ss-naphthylamide) moiety. See, e.g., Grant et al., Biochem. Biophys. Acta. 567:207(1979). Bovine enterokinase is also active on some trypsin substrates like TAME and BAEE (benzyl-arginine-ethyl-ester).

The term "wild type enterokinase light chain" as used herein is intended to mean an enterokinase light chain before any substitutions according to the invention have been applied thereto.

The term "enterokinase light chain analogue" or "bovine enterokinase light chain analogue" as used herein means a modified bovine enterokinase light chain wherein one or more amino acid residues of the enterokinase light chain have been substituted by other amino acid residues and/or wherein one or more amino acid residues have been deleted from the enterokinase light chain and/or wherein one or more amino acid residues have been added and/or inserted to the enterokinase light chain.

In one embodiment an enterokinase light chain analogue comprises less than 10 amino acid modifications (substitutions, deletions, additions (including insertions) and any combination thereof) relative to bovine enterokinase light chain, alternatively less than 9, 8, 7, 6, 5, 4, 3or 2 modifications relative to bovine enterokinase light chain. In one aspect an enterokinase light chain analogue comprises 5 amino acid modifications, in one aspect 4 amino acid modifications, in one aspect 3 amino acid modifications, in one aspect 2 amino acid modifications and in one aspect 1 amino acid modification relative to bovine enterokinase light chain.

Modifications in the enterokinase molecule light chain are denoted stating the position and the one or three letter code for the amino acid residue substituting the native amino acid residue. Using the one letter codes for amino acids, terms like 134K and 135K designates that the amino acid in position 134 and 135, respectively, is K. Using the three letter codes for amino acids, the corresponding expressions are 134Lys and 135Lys, respectively. Thus, e.g., 112Ala,134Lys,135Lys bovine enterokinase light chain is an analogue of bovine enterokinase light chain where the amino acid in position 112 is substituted with alanine, the amino acid in position 134 is substituted with lysine and the amino acid in position 135 is substituted with lysine.

Herein, the term "amino acid residue" is an amino acid from which, formally, a hydroxy group has been removed from a carboxy group and/or from which, formally, a hydrogen atom has been removed from an amino group.

Examples of bovine enterokinase light chain analogues are such wherein Leu in position 134 is substituted with Lys or another charged amino acid, at position 135 where lie is substituted with Lys or another charged amino acid. Furthermore, Cys in position 112 may be substituted with a number of amino acids including Ala and Ser.

Further examples of bovine enterokinase light chain analogues according to the invention include, without limitation: 134Lys bovine enterokinase light chain; 135Lys bovine enterokinase light chain; 134Lys,135Lys bovine enterokinase light chain; 112Ala,134Lys,135Lys bovine enterokinase light chain; 112Ala,134Lys bovine enterokinase light chain; 112Ala,135Lys bovine enterokinase light chain and any such combinations including substitutions with other charged amino acids.

In one aspect a bovine enterokinase light chain analogue is obtained which has improved solubility in a renaturation process relative to natural bovine enterokinase light chain. In one aspect a bovine enterokinase light chain analogue according to the invention has one or more surface oriented hydrophobic amino acids which have been mutated to hydrophilic, charged amino acids wherein improved solubility in a renaturation process relative to natural bovine enterokinase light chain is obtained. In one aspect surface oriented hydrophobic amino acids for substitution to hydrophilic charged amino acids are selected after aligning the bovine enterokinase light chain with other serine proteases and scanning the solvent-accessable surfaces through a computational 3D model of enterokinase.

The method for refolding a bovine enterokinase light chain analogue according to the invention is known to the person skilled in the art. For example, refolding may be carried out by denaturation in urea, followed by oxidative refolding in glutathione or another re-dox environment.

In one aspect a buffer (refolding buffer) is used during the refolding process. In one aspect of the invention, the refolding buffer comprises urea. In one aspect, the refolding buffer comprises between 0M and 2M urea. In one aspect, the refolding buffer comprises between 0.5M and 2M urea, between 0M and 1.5M urea or between 0.5M and 1.5M urea. In one aspect, the refolding buffer comprises about 1M urea.

The initial concentration of inclusion body may affect the refolding yield. In one aspect of the invention, the concentration of inclusion body is between 1 and 4 mg/ml .

In one aspect of the invention, the thioredoxin (Trx) tag is removed during refolding, i.e. during dilution and incubation under refolding conditions. It has thus been found that refolding and activation may be obtained without addition of an activation enzyme. In one aspect of the invention, the linker connecting the trx tag and the bovine enterokinase light chain analogue of the invention is removed by autocleavage. It has thus by the inventors surprisingly been found that the linker connecting the trx tag and the bovine enterokinase light chain analogue of the invention facilitates the refolding.

In one aspect, less aggregation during the renaturation process of a bovine enterokinase light chain analogue according to the invention is obtained relative to the aggregation obtained during the renaturation process of wild type EK. In one aspect, a bovine enterokinase light chain analogue according to the invention has the substitutions L134K and I135K, where the bovine enterokinase light chain analogue is more soluble during the renaturation process relative to wild type EK. In one aspect, a bovine enterokinase light chain analogue according to the invention further has the substitution C112A. It is believed by the inventors that by mutating the lone cysteine in position 112, which in wild type EK heterodimer is involved in the disulfide binding from the light chain to the heavy chain, formation of the 4 disulfide bridges in the EK light chain may be facilitated.

In one aspect, a bovine enterokinase light chain analogue of the invention has full enterokinase activity compared to wild type bovine enterokinase. In one aspect, a bovine enterokinase light chain analogue of the invention has a substantially equivalent functional or biological activity as wild type bovine enterokinase. For example, a bovine enterokinase light chain analogue has substantially equivalent functional or biological activities (i.e., is a functional equivalent) of the polypeptide having the amino acid sequence set forth as SEQ ID NO: 1 (e.g., has a substantially equivalent enteropeptidase activities).

Nucleic acid forms encoding enterokinase light chain analogues of the present invention are also within the scope of the invention. Nucleic acids according to the invention include genomic DNA (gDNA), complementary DNA (cDNA), synthetic DNA prepared by chemical synthesis as well as DNA with deletions or substitutions, allelic variants and

sequences that hybridize thereto under stringent conditions as long as they encode enterokinase light chain analogues of the present invention.

In one embodiment a nucleic acid is provided wherein said nucleic acid comprises a polynucleotide sequence, and wherein said nucleic acid encodes a mammalian enterokinase light chain analogue such as a bovine enterokinase light chain analogue according to the invention. In one embodiment, the nucleic acid is operably linked to an inducible promoter. In one embodiment, a recombinant vector is provided which comprises the nucleic acid operably linked to the inducible promoter. In one embodiment, the inducible promoter is selected from a group consisiting of AraB, T7, trp, lac, tac.

A further embodiment of the invention provides a host cell comprising the recombinant vector comprising the polynucleotide sequence coding for the amino acid sequence of a mammalian enterokinase light chain analogue such as a bovine enterokinase light chain analogue according to the invention.

A further aspect of the invention provides the host cell comprising the recombinant vector comprising the polynucleotide sequence coding for the amino acid sequence encoding a mammalian enterokinase light chain analogue such as a bovine enterokinase light chain analogue according to the invention. In one embodiment, the host cell is selected from a group consisting of *E.coli, B.subtilis, S.saccaromyces* and *A.oryzae.*

The production of polypeptides, e.g., enterokinase light chain, is well known in the art. The bovine enterokinase light chain analogue may for instance be produced by classical peptide synthesis, e.g., solid phase peptide synthesis using t-Boc or Fmoc chemistry or other well established techniques, see, e.g., Greene and Wuts, "Protective Groups in Organic Synthesis", John Wiley & Sons, 1999. The bovine enterokinase light chain analogue may also be produced by a method which comprises culturing a host cell containing a DNA sequence encoding the analogue and capable of expressing the bovine enterokinase light chain analogue in a suitable nutrient medium under conditions permitting the expression of the bovine enterokinase light chain analogue. Several recombinant methods may be used in the production of bovine enterokinase light chain and bovine enterokinase light chain analogues. Examples of methods which may be used in the production of enterokinase in microorganisms such as, e.g., *Escherichia coli* and *Saccharomyces cerevisiae* are, e.g., disclosed in WO 94/16083.

Typically, the bovine enterokinase light chain analogue is produced by expressing a DNA sequence encoding the bovine enterokinase light chain analogue in question or a precursor thereof in a suitable host cell by well known technique as disclosed in e.g. WO 94/16083

The bovine enterokinase light chain analogues of the invention may be recovered from the cell culture medium or from the cells. The bovine enterokinase light chain analogues of the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing (IEF), differential solubility (e.g., ammonium sulfate precipitation), or extraction (see, e.g., Protein Purification, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

In one aspect, the bovine enterokinase light chain analogues of the present invention are purified using anion exchange chromatography. In a further aspect, the anion exchange chromatography is followed by hydrophobic interaction chromatography. In one aspect, the bovine enterokinase light chain analogues of the present invention are purified using Q HP anion exchange chromatography. In a further aspect, the Q HP anion exchange chromatography is followed by Phenyl FF hydrophobic interaction chromatography.

In one aspect of the present invention an improved process for production of a mammalian enterokinase light chain analogue such as a bovine enterokinase light chain analogue is provided, wherein said method comprises the steps:

a) culturing the host cells in a growth medium comprising inducer, wherein the host cells comprise a polynucleotide sequence encoding the amino acid sequence of the enterokinase light chain analogue;

b) recovering the cells with enterokinase light chain analogue in inclusion bodies

c) solubilizing and refolding the enterokinase light chain analogue; and

d) purifying the enterokinase light chain analogue.

The invention provides a new recombinant process for production of mammalian enterokinase light chain analogue such as a bovine enterokinase light chain analogue in E.coli in a very efficient and economic way.

The expression of a bovine enterokinase light chain analogue according to the invention may e.g. be localized in the inclusion bodies of *E. coli* or in the secreted material of yeast. In one embodiment expression of enterokinase is localized in the inclusion bodies of *E. coli.*

Various strains of E. coli are useful as host cells for the production of non-glycosylated, homogeneous enterokinase activity are also well-known in the art. A non-exclusive list of such strains includes E.coli B BL21 DE3, E.coli K12 W31 10, MC1061 , DH1,

K803, HB101 , JM101 and other K12 like strains. Alternatively, other bacterial species may be used, including B. subtilis, various strains of Pseudomonas, other bacilli and the like.

Many strains of yeast cells, known to those skilled in the art, are also available as host cells for expression of the enterokinase activity of the present invention. Yeast cells are especially useful as a host for pre/pro fusion to mature enterokinase. When expressed using a suitable yeast vector, the fusion is secreted by virtue of a signal peptide.

When the bovine enterokinase light chain analogue of this invention is expressed in bacterial cells, it may be expressed intracellular^ usually as inclusion bodies, or it may be secreted from bacterial cells in active form if a secretory signal is included. Where necessary or desired, as when reduced bioactivity is observed, the enterokinase activity may be obtained by conventional methods such as solubilization of protein in urea or guanidine HCI, followed by dilution to reduce the concentration of these reagents and treatment with oxidizing agents such as dithiothreitol or ss-mercapto ethanol to enhance refolding.

In one embodiment, the bovine enterokinase light chain analogues according to the invention are enzymatically active proteases which cleave specifically after a $(Asp)_4$-Lys (DDDDK) sequence in various numbers of fused protein products between affinity tag and the mature protein. In one embodiment, the bovine enterokinase light chain analogues according to the invention have retained enzymatic activity

In one aspect of the invention, a process for preparing a bovine enterokinase light chain analogue in E. coli cells is obtained, wherein the E. coli cells are transformed with a plasmid carrying the bovine enterokinase light chain analogue gene and an inducible promoter by fermentation involving batch and fed batch stages and isolation and purification of the expressed protein from the cultures.

In one aspect of the invention, a refolding process for a bovine enterokinase light chain analogue according to the invention is obtained, wherein the expression of the enterokinase light chain analogue is in the form of inclusion bodies in recombinant *E. coli.* In one embodiment denaturation followed by refolding in a redox system is used.

The enterokinase light chain analogues of the invention may be used in a method for cleaving proteins having an enterokinase cleavage site, and especially fusion proteins having such a cleavage site engineered into their sequence. The amounts needed are readily determined empirically by one skilled in the art.

The term "fusion protein" as used herein is meant to refer to a protein created through genetic engineering from two or more proteins or peptides. As used herein, a fusion protein can refer to a protein in which a Asp-Asp-Asp-Asp-Lys (D4K) sequence has been intentionally introduced for specific cleavage. Generally, cleavage of the fusion protein

generates two polypeptides. A fusion protein according to the invention can be a recombinant fusion protein. In particular embodiments, a fusion protein can be generated, for example, from the addition of a vector-derived residue peptide at one terminus, for example the N-terminus, in addition to the amino acid sequence of the wild type protein of interest. In this way, for example, a recombinant fusion protein can be constructed to have Asp-Asp-Asp-Lys (D4K) cleavage site in the vector upstream joined to the protein of interest.

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

The term "protease" is intended to include any polypeptide/s, alone or in combination with other polypeptides, that break peptide bonds between amino acids of proteins.

The term "proteolytic activity" is meant to refer to the cleavage activity of a substrate by an enzyme. In particular embodiments, the term refers to the enzymatic cleavage by enteropeptidases. In exemplary embodiments, the term is meant to refer to the specific activity of a bovine enterokinase light chain analogue of the invention for Asp-Asp-Asp-Asp-Lys cleavage sites. "Non-specific proteolytic activity" is meant to refer to cleavage activity that is not directed to a specific cleavage site. "Specific proteolytic activity" is meant to refer to cleavage activity that is directed to a specific cleavage site.

Indeed, as described herein, a bovine enterokinase light chain analogue according to the invention is superior for cleavage of fusion proteins when compared to the bovine-derived two-chain form.

As another aspect of the invention, the enterokinase light chain analogue of the invention is incorporated as one of the fusion protein partners to yet another protein. As such, with the addition of a minimal amount of exogenous enterokinase activity to the reaction vessel the fusion protein results in the release of additional enterokinase activity which in turn can catalyze many more proteolytic cleavages of fusion proteins. In this way, large amounts of enterokinase activity can be produced from a fusion protein in an autocatalytic manner.

Another particular aspect of the invention teaches a method for cleavage of a protein containing an Asp-Asp-Asp-Asp-Lys cleavage site using any of the bovine enterokinase light chain analogues of the invention described herein, the method comprising contacting the protein with any of the bovine enterokinase light chain analogues of the

invention, and wherein the contacting of the protein with the bovine enterokinase light chain analogue results in specific cleavage.

In one embodiment, the protein is a fusion protein. In another embodiment, the fusion protein is a recombinant fusion protein. In a further embodiment, the protein is bacterially produced. In a more particular embodiment, the protein is a synthetic protein.

In a further aspect, the invention teaches a method for the preparation of recombinant protein using any of the bovine enterokinase light chain analogues according to the invention as described herein, the method comprising providing a recombinant fusion protein containing a Asp-Asp-Asp-Asp-Lys cleavage site, and contacting the fusion protein with any of the bovine enterokinase light chain analogues of the invention, wherein contacting the recombinant fusion protein with the bovine enterokinase light chain analogue results in Asp-Asp-Asp-Asp-Lys specific cleavage and preparation of recombinant protein.

The following is a non-limiting list of aspects according to the invention:

1. A bovine enterokinase light chain analogue comprising at least one substitution in position 134 and/or 135 from hydrophobic to a hydrophilic charged amino acid(s).

2. The bovine enterokinase light chain analogue according to aspect 1, wherein both positions 134 and 135 have substitutions from a hydrophobic to a hydrophilic charged amino acid.

3. The bovine enterokinase light chain analogue according to aspect 1 or 2, further comprising a substitution in position 112.

4. The bovine enterokinase light chain analogue according to aspect 3, wherein the amino acid in position 112 is selected from the group consisting of: alanine, serine and glycine.

5. The bovine enterokinase light chain analogue according to aspect 3, wherein the amino acid in position 112 is alanine.

6. The bovine enterokinase light chain analogue according to any one of the previous aspects, wherein the hydrophilic charged amino acid(s) are one or more amino acids selected from the group consisting of: lysine, arginine, glutamic acid and aspartic acid.

7. The bovine enterokinase light chain analogue according to any one of the previous aspects, wherein the hydrophilic charged amino acid(s) are lysine.

8. The bovine enterokinase light chain analogue according to any one of the previous aspects, comprising the substitutions C112A, L134K and I135K.

9. The bovine enterokinase light chain analogue according to any one of the previous aspects, wherein the enterokinase light chain to be mutated is SEQ ID NO:1 .

10. A method for obtaining improved solubility in a renaturation process of an enterokinase light chain analogue comprising the step of mutating one or more hydrophobic amino acids of wild type bovine enterokinase light chain to hydrophilic amino acids and optionally mutating other amino acids of wild type bovine enterokinase light chain, wherein the hydrophobic amino acids subject to mutation are present on the surface of folded wild type bovine enterokinase light chain.

11. A method according to aspect 10, wherein the hydrophobic amino acid(s) to be mutated are selected from the group consisting of: I, V, L, M, W, F, A

12. A method according to aspect 10, wherein the hydrophobic amino acid(s) to be mutated are selected from the group consisting of: Leucin and isoleucin.

13. A method according to any one of aspects 10-12, wherein the hydrophilic amino acid(s) are selected from the group consisting of: Lysine, arginine, glutamic acid and aspartic acid.

14. A method according to aspect 13, wherein the hydrophilic amino acid(s) are lysine.

15. A method according to any one of aspects 10-14, wherein the hydrophobic amino acid(s) to be mutated are in one or more positions selected from the group consisting of: position 11-14 (amino acids AWPW), position 78-80 (amino acids I V I) and position 133-136 (amino acids A L I Y).

16. A method according to aspect 15, wherein the hydrophobic amino acid(s) to be mutated are in positions 134 and/or 135.

17. A method for production of a bovine enterokinase light chain analogue, wherein said method comprises the steps:

a) culturing the host cells in a growth medium comprising inducer, wherein the host cells comprise a polynucleotide sequence encoding the amino acid sequence of the enterokinase light chain analogue;

b) recovering the cells with enterokinase light chain analogue in inclusion bodies

c) solubilizing and refolding the enterokinase light chain analogue; and

d) purifying the enterokinase light chain analogue.

18. A method for production of a bovine enterokinase light chain analogue according to aspect 17, wherein a refolding buffer is used during the refolding process.

19. A method for production of a bovine enterokinase light chain analogue according to aspect 17 or 18, wherein the refolding buffer comprises urea.

20. A method for production of a bovine enterokinase light chain analogue according to any one of aspects 17-19, wherein the refolding buffer comprises about 1M urea.

21. A method for production of a bovine enterokinase light chain analogue according to any one of aspects 19-20, wherein the refolding buffer further comprises low molecular weight polyethylene glycol (low-PEG).

22. A method for production of a bovine enterokinase light chain analogue according to any one of aspects 19-21, wherein the refolding buffer further comprises PEG1000 such as 1% PEG1000.

23. A method for production of a bovine enterokinase light chain analogue according to any one of aspects 19-22, wherein the refolding buffer further comprises hydroxypropyl-β-cyclodextrin such as 1.5% hydroxypropyl -β-cyclodextrin.

24. A method for production of a bovine enterokinase light chain analogue according to any one of aspects 17-23, wherein the concentration of inclusion body is between 1 and 4 mg/ml.

25. A method for production of a bovine enterokinase light chain analogue according to any one of aspects 17-24, wherein the host cell is *E.coli.*

26. A method for production of a bovine enterokinase light chain analogue according to any one of aspects 17-25, wherein the bovine enterokinase light chain analogue is an analogue according to any one of aspects 1-9.

27. A method for recombinantly producing a peptide or protein in a bacterial or yeast host cell, comprising

a) expressing in yeast or bacteria a fusion protein comprising the peptide or protein to be produced;

b) cleaving the fusion protein with a bovine enterokinase light chain analogue according to any one of aspects 1-9; and

c) isolating the produced peptide or protein.

28. A method for recombinantly producing a peptide or protein according to aspect 27, wherein the fusion protein expressed in step a) further comprises an Asp-Asp-Asp-Asp-Lys cleavage site.

29. A method for recombinantly producing a peptide or protein according to aspect 28, wherein step b) results in Asp-Asp-Asp-Asp-Lys specific cleavage.

30. A method for recombinantly producing a peptide or protein according to any one of aspects 27-29, wherein the host cell is *E. coli.*

31. A method for recombinantly producing a peptide or protein according to any one of aspects 27-30, wherein the peptide or protein to be produced is a GLP-1 peptide.

All references, including publications, patent applications, and patents, cited herein are hereby incorporated by reference in their entirety and to the same extent as if each

reference were individually and specifically indicated to be incorporated by reference and were set forth in its entirety herein (to the maximum extent permitted by law).

All headings and sub-headings are used herein for convenience only and should not be construed as limiting the invention in any way.

The use of any and all examples, or exemplary language (e.g., "such as") provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

The citation and incorporation of patent documents herein is done for convenience only and does not reflect any view of the validity, patentability, and/or enforceability of such patent documents.

This invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law.

EXAMPLES

Herein a production process for making bovine enterokinase light chain analogues has been developed. The bovine enterokinase light chain analogues were fused to thioredoxin tag expressed as inclusion bodies in *E.coli.* After refolding and auto-activation, the active enterokinase light chain analogue was purified by Q HP anion exchange chromatography. Moreover, it was found that triple substitutions (C112A, L134K and I135K) of bovine enterokinase light chain (EK_{LM}), which improved the surface hydrophilic properties, increased the refolding yield 4 fold without loosing activity. The yield of purified enterokinase light chain analogue was 800mg/L from a culture of 4g/L, and the specific activity was determined as 5000 ± 10 EU/mg. Thus, our enterokinase light chain analogue production process provides a valuable tool for processing therapeutic fusion proteins and other fusion proteins.

Abbreviations :

EK: Enterokinase

EK_L: Bovine Enterokinase light chain with C112A mutation

EK_{LM} (alternatively herein named EK_M or EK_{LM}(C1 12A, L134K, I135K)): Bovine Enterokinase light chain with mutations in positions 112 to Ala, 134 to Lys and 135 to Lys.

TrxEK|_M : EK_LMfused with N-terminal Thioredoxin tag with a linker of 12AA

Trx-Linker-EK|_M : Effused with N-terminal Thioredoxin tag with a longer linker of 49AA

Trx-Linker-EK|_ : EK_Lfused with N-terminal Thioredoxin tag with a longer linker of 49AA

IPTG: Isopropyl β-D-l-thiogalactopyranoside

Tris: Tris(hydroxymethyl)aminomethane

DTT: Dithiothreitol

GSSG: Glutathione disulfide

GSH: Glutathione

FDM: Fermentation defined medium

Trx: Thioredoxin

LC-MS: Liquid chromatography-mass spectrometry

SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis

BL21 : Ecoli strain E.coli B BL21 DE3

PCR reaction: Polymerase chain reaction

Low-PEG: Low molecular weight polyethylene glycol such as polyethylene glycols with a molecular weight up to 1000

PEG1000: Polyethylene Glycol 1000, a polyethylene glycol with approximate molecular weight 1000.

Example 1. Plasmid construction of Trx-linker-EK_L and Trx-linker-EK_{LM}

The DNA sequence encoding the catalytic subunit of bovine enterokinase was amplified with the following primers:

5'-ggcggtaccgacgacgacgacaagattgtcggaggaagtgac-3' **SEQ ID NO: 2**

5'- ggcgaattcctaatgtagaaaactttgtatccactctgtgaacc-3' **SEQ ID NO: 3**

These two primers contained Kpn I and EcoR I restriction enzyme sites, respectively. The target fragment was introduced into pET32a (Novagen) from Kpnl and EcoRI site. Routine PCR reaction was performed using *Pfu* DNA Polymerase from Stratagene. The sequence of plasmid pET32a-EK_L was confirmed by sequencing. Three

substitution sites, i.e. C112A, L134K, I135K were introduced by using QuikChange® XL Site-Directed Mutagenesis Kit from Stratagene with the primers:

C112AF 5'-acacagattatatacagcctat tgcgttaccagaagaaaatcaag-3' **SEQ ID NO: 4**

C112AR 5'-cttgattttcttctggtaacgcaataggctgtatataatctgtgt-3' **SEQ ID NO: 5**

L134KJ135KF 5'-ctattgctggctgggggcaaagaaatatcaaggttctactgcagacg-3' **SEQ ID NO: 6**

L134K,I135KR5'-cgtctgcagtagaaccttgatatttctttcccc ccagccagcaatag-3' **SEQ ID NO: 7**

Amino acid Sequence of Trx-linker-EK$_{LM}$:

MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQ NPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSS GLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDK ***IVGGSDSREGAWPWVVALYFDDQQ VCGASLVSRDWLVSAAHCVYGRNMEPSKWKA VLGLHMASNLTSPQIETRLIDQIVINPHY NKRRKNNDIAMMHLEMKVNYTDYIQPIALPEENQVFPPGRICSIAGWGAKKYQGSTADVLQ EADVPLLSNEKCQQQMPEYNITENMVCAGYEA GGVDSCQGDSGGPLMCQENNRWLLA G VTSFGYQCALPNRPGVYARVPRFTEWIQSFLH***

SEQ ID NO: 8

Underlined: Trx; Regular: linker; ***Bold italic: EK$_{LM}$***

Amino acid Sequence of TRX-EKLM:

MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQ NPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGGTDDDDK ***IyGG SDSREGAWPWWALYFDDQQVCGASLVSRDWLVSAAHCVYGRNMEPSKWKA VLGLHM ASNLTSPQIETRLIDQIVINPHYNKRRKNNDIAMMHLEMKVNYTDYIQPIALPEENQVFPPGR ICSIAGWGAKKYQGSTADVLQEADVPLLSNEKCQQQMPEYNITENMVCAGYEAGGVDSC QGDSGGPLMCQENNRWLLAGVTSFGYQCALPNRPGVYARVPRFTEWIQSFLH***

SEQ ID NO: 9

Underlined: Trx; Regular: linker; ***Bold italic: EK$_{LM}$***

Example 2. Fermentation and expression of Trx-linker-EK$_L$ and Trx-linker-EK$_{LM}$

Cells from a glycerol stock were inoculated on an EC1 plate grown overnight at 37°C, and washed with 0.9% sodium chloride (NaCl) to suspend the cells. The culture was allowed to grow in a fermentor containing fermentation defined medium (FDM) at 37°C for 16

hrs, and induced with 1.0mM IPTG at an OD600 of 150, and then grown for 6 hours at 37°C before harvesting by centrifugation.

Both Trx-linker-EK$_L$ and Trx-linker-EK$_{LM}$ in *E.coli* BL21 were expressed in fed-batch fermentation. As shown in Fig.1, no apparent leaky expression judged by SDS-PAGE was observed before IPTG induction. An induced band just above 43kD on SDS-PAGE by IPTG appeared, and it was confirmed by LC-MS that this band represented the target protein. Moreover, the expression level of the target protein was dependent upon the induction time. 4hrs or 6hrs of induction for Trx-linker-EK$_L$ and Trx-linker-EK$_{LM}$ by using fermentation defined medium (FDM), respectively gave acceptable expression, and ~4g/L of the target proteins was achieved.

Example 3. Refolding, auto-catalytic activation and purification

Cells from fermentation were resuspended in lysis buffer (1:10, w/w) containing 20mM Tris, pH 8.0, and lysed by French press. Inclusion bodies were sedimented at 20,000g for 1hr at 4°C, and then washed once by using lysis buffer. The inclusion bodies were solublized to 3.2mg/ml in buffer containing 20mM Tris, 8M urea, pH8.0, 20mM DTT and incubated at 4°C for 3hrs. After centrifugation at 20,000g for 30min, the solublized EK (i.e. Trx-linker-EK$_L$ and/or Trx-linker-EKuvi) was diluted 80 fold into refolding buffer containing 20mM Tris, 1M Urea, 1mM GSSG, 3mM GSH, pH 8.3 and incubated at 20°C for 24hrs.

During dilution and incubation of the refolding procedure, auto-catalytic cleavage occurred, and liberated fully active enzyme without thioredoxin (Trx) tag. Finally, the enzyme was purified by Q HP anion exchange chromatography.

The process scheme is shown in Fig.2. The inclusion bodies were solublized in the buffer containing 5-8M urea and 10-20mM DTT. It should be noted that the inclusion body concentration affected the refolding yield. It was found that the refolding yield of 4mg/ml Trx-linker-EKLM was 2 fold higher than that of 6mg/ml Trx-linker-EK$_{LM}$ (Fig.3A).

The refolding occurred by dilution. The amount of purified enzyme from a fixed volume was also dependant upon the Trx linker EK concentration in the refolding buffer, and reached a maximum when Trx-linker-EK$_{LM}$ concentration was 120μg/ml.

The auto-catalytic activation occurred concomitantly with the refolding process. The active EK was liberated from Trx-linker-EK by the escape active EK, which specifically cleaved Trx tag off at DDDDK recognition site just before the mature EK. The refolding and auto-catalytic activation process seemed optimal at48hrs (Fig.4). Considering the inhibition of EK by urea, it was found that the refolding yield was largely reduced if above 2M urea in refolding buffer. Our result showed that 1M urea in refolding buffer was optimal (Fig.5).

The refolding yield was dependent upon the redox system. GSSG/GSH in the ratio 1:3 was found optimal and better than Cystine/cystein (Fig. 6).

The active EK after refolding and auto-activation was purified and concentrated by one step anion exchange chromatographic purification (QHP column, Fig.7A). It was found that Trx tag was in P1, EK_LM was mainly in P2 together with the impurity of Trx tag, and P3 contained trace amount of EK_LM, which is confirmed by the activity assay shown in Fig. 7C. It should be noted that high purity EK_LM (>90%) was obtained by further purification of P2 using hydrophobic interaction chromatography (HIC) (Fig.7B). Moreover, the enzymatic activity of each fraction was also assayed (Fig.7C), and pooled. For Trx-linker-EK_L, the refolding yield was rather low beyond 40μg/ml of Trx-linker-EK_L during the refolding process (4.4% at 40μg/ml), which made this process practically difficult. In other words, a huge holding tank is required to produce large amount of EK (~1,000g).

The low refolding yield could be due to protein aggregation caused by protein hydrophobic interactions. After surface hydrophobicity mapping of EK_L based on its 3D structure, it was found that the ^{133}ALIY is one of the most hydrophobic patches on the surface. Therefore, EK_LM with 3 substitutions (C112A, L134K and I135K) was constructed and subjected to study. By using the exact same process, EK_LM greatly improved the refolding yield, especially when EK_LM concentration in refolding buffer was beyond 40μg/ml, which is the bottle neck for the large scale production of EK_L (Fig.3A). As shown in Fig.3A, at 40μg/ml of Trx-linker-EK_LM concentration in the refolding buffer, the refolding yield of Trx-linker-EK|_M (17%) was 4 fold higher than that of Trx-linker-EK_L (4.4%). Moreover, ~16mg of active EK_LM could be purified from 1L refolding tank in which the EK_LM concentration is 120Mg/ml.

The specific enzymatic activity between EK_L and EK_LM was compared as in Fig. 8. The triple substitutions of EK_LM had no apparent effect on enzyme activity, which was evidenced by the fact that EK_L and EK_LM have similar bands on SDS-PAGE if loaded the same activity. Moreover, EK_LM was quite stable if stored in buffer containing 20mM Tris, 200mM NaCl at -80°C or 4°C. No apparent degradation and decrease of activity were observed up to 3 months (Fig. 9).

Example 4. Enzyme assays

The enzymatic activity was measured directly using a fluorogenic substrate, GDDDDK-Beta-naphthylamide. The reaction was started with addition of 1ul sample into each well of Fluorescent 96 well plate containing 100ul of reaction buffer. After mixing for 10

seconds, the fluorescence was measured with Fluostar OPTIMA (excitation at 340nM and emission at 420nM). The enzyme activity was defined by arbitrary unit (EU), which derived from slope * 60/30,000, where the slope represented linear range.

Example 5. Linker region

Two EK|_M amino acid sequences connected to trx were produced where the linker region differed, trxEK $_{L}$M and trx-linker-EK $_{L}$M (see figure 10). In trx-linker-EK $_{L}$M the spacer between trx and EKLM is 37 amino acids longer than in trxEK $_{L}$M.

TrxEK$_{LM}$

Cell disruption and IBs solubilization

7.41 g TrxEK $_{LM}$ cell pellet was resuspended in 100ml of lysis buffer (20mM Tris, pH 8.0), and the cells were disrupted by using a homogenizer under a pressure of 30,000psi. After the supernatant was discarded, the IBs weighed 3.53g. The isolated IBs were resuspended in 70ml of solublization buffer (20mM Tris, 8M urea, pH8.0, 20mM DTT (freshly added)) and incubated at 4°C for 4hrs. The solublized samples were clarified by centrifugation.

Refolding of TrxEK $_{L}$M

16ml of IBs solution was diluted into 500ml refolding buffer (20mM Tris, 1mM GSSG, 3mM GSH, 1M Urea, pH 8.0) and stirred at 20°C for 54hrs. The concentration of protein during refolding is 60µg/ml.

Purification of TrxEK $_{LM}$

Column:	Q HP column
Sample buffer:	20mM Tris, 1mM GSSG, 3mM GSH, 0.62mM DTT, 1M Urea, pH 8.0
Buffers:	Buffer A: 20mM Tris, pH 8.0
	Buffer B: 20mM Tris, 0.5M NaCl, pH 8.0
Procedure:	10 CV 100% A
	Application at 10ml/min
	5 CV 100% A
	7 CV 0% B-70%B
	1 CV 70% B-100%B
	1.5CV 100%B
Column volume:	28 ml

Speed: 10ml/min

The elution fractions with highest enzyme activity were combined resulting in a pool volume of 30ml and total enzyme activity of 14,100EU. The protein amount was 2.82mg.

Trx-linker-EKLM

Cell disruption and IBs solubilization

66.9g Trx-linkerEK_LMcell pellet was resuspended in 1000ml of lysis buffer (20mM Tris, pH 8.0), and the cells were disrupted by using a homogenizer under a pressure of 30,000psi. After the supernatant was discarded, the IBs weighed 22g and were washed by 1000ml of 20mM Tris, pH 8.0 once. After wash, the IBs solution was divided into 6 bottles for centrifugation. After the supernatant was discarded, 41ml of solublization buffer (20mM Tris, 8M urea, pH8.0, 20mM DTT (freshly added)) was added into one bottle and incubated at 4°C for 3hrs. The solublized IBs were clarified by centrifugation and the final volume was 43ml.

Refolding of Trx-linker-EK_{LM}

9ml of IBs solution was diluted into 500ml of refolding buffer (20mM Tris, 1M Urea, 1mM GSSG, 3mM GSH, pH 8.0) and stirred at 20°C for 18hrs. The concentration of protein during refolding was 60µg/ml.

Purification of Trx-linker-EK_{LM}

Column: Q HP column

Sample buffer: 20mM Tris, 1M Urea, 1mM GSSG, 3mM GSH, 0.296mM DTT, pH 8.0

Buffers: Buffer A1: 20mM Tris, 1M Urea, pH 8.0

Buffer A2: 20mM Tris, pH 8.0

Buffer B: 20mM Tris, 0.5M NaCl, pH 8.0

Procedure: 10 CV 100% A1

Application at 10ml/min

5 CV 100% A1

5 CV 100% A2

7 CV 0% B-70%B(100% A2-30%A2)

1 CV 70% B-100%B(30% A2-0%A2)

1.5CV 100%B

Column volume: 28 ml

Speed: 10ml/min

The enzyme activity of elution fractions 18-23 is higher than the other fractions through activity test. The elution fractions with highest enzyme activity were combined resulting in a pool volume of 30ml and total enzyme activity of 24,900EU. The protein amount was 4.98mg.

Result:

2.82mg of EK_{LM} protein was produced from 0.5L of refolding solution when using $TrxEK_{LM}$ when the protein concentration was 60 μg/ml during refolding, whereas 4.98mg of EK protein was produced from Trx-linker-EK_{LM} version under the same conditions. Thus, the fusion protein with longer linker showed 76% higher of refolding efficiency than the fusion protein with shorter linker.

Example 6: Components optimization of the refolding buffer

Several different additives, including detergents, cyclodextrins, amino acids , PEG (polyethylene glycol) and sugars, were combined into the current refolding buffer (20mM Tris, 1M Urea, 1mM GSSG, 3mM GSH, pH 8.3) individually to test their capacity to improve the refolding efficiency of Trx-linker-EK_{LM}. The refolding process was performed as described in Example 3 with small modifications. Briefly, the inclusion bodies were solubilized to 7.3mg/ml in the buffer containing 20mM Tris, 8M urea, pH8.0, 20mM DTT, and then the solubilized Trx-linker-EK_{LM} was added into the optimized refolding buffer containing certain additive by 20-fold dilution. The mixture was incubated at 4°C for 20hrs and the amount of correctly refolded Trx-linker-EK_{LM} was quantified by protease activity assay as described in Example 4.

Both low-PEG (eg.PEGI 000, 1%) and hydroxypropyl -β-cyclodextrin (1.5%) exhibited strong capacity to enhance the refolding efficiency of Trx-linker-EK_{LM}, with 57.9% and 106.2% increase, respectively, to that from urea-only refolding buffer (as shown in figure 11). These two additives have no obvious impact on the maturation of EK_{LM} and the following purification process.

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. A bovine enterokinase light chain analogue comprising at least one substitution in position 134 and/or 135 from hydrophobic to a hydrophilic charged amino acid(s).

2. The bovine enterokinase light chain analogue according to claim 1, further comprising a substitution in position 112.

3. The bovine enterokinase light chain analogue according to any one of the previous claims, wherein the hydrophilic charged amino acid(s) are one or more amino acids selected from the group consisting of: lysine, arginine, glutamic acid and aspartic acid.

4. The bovine enterokinase light chain analogue according to any one of the previous claims, wherein the enterokinase light chain to be mutated is SEQ ID NO:1 .

5. A method for obtaining improved solubility in a renaturation process of an enterokinase light chain analogue comprising the step of mutating one or more hydrophobic amino acids of wild type bovine enterokinase light chain to hydrophilic amino acids and optionally mutating other amino acids of wild type bovine enterokinase light chain, wherein the hydrophobic amino acids subject to mutation are present on the surface of folded wild type bovine enterokinase light chain.

6. A method according to claim 5, wherein the hydrophobic amino acid(s) to be mutated are selected from the group consisting of: I, V, L, M, W, F, A

7. A method according to any one of claims 5-6, wherein the hydrophilic amino acid(s) are selected from the group consisting of: Lysine, arginine, glutamic acid and aspartic acid.

8. A method according to any one of claims 5-7, wherein the hydrophobic amino acid(s) to be mutated are in one or more positions selected from the group consisting of: position 11-14 (amino acids AWPW), position 78-80 (amino acids I V I) and position 133-136 (amino acids A L I Y).

9. A method according to claim 8, wherein the hydrophobic amino acid(s) to be mutated are in positions 134 and/or 135.

10. A method for production of a bovine enterokinase light chain analogue, wherein said method comprises the steps:

a) culturing the host cells in a growth medium comprising inducer, wherein the host cells comprise a polynucleotide sequence encoding the amino acid sequence of the enterokinase light chain analogue;

b) recovering the cells with enterokinase light chain analogue in inclusion bodies

c) solubilizing and refolding the enterokinase light chain analogue; and

d) purifying the enterokinase light chain analogue.

11. A method for production of a bovine enterokinase light chain analogue according to claim 10, wherein the bovine enterokinase light chain analogue is an analogue according to any one of claims 1-4.

12. A method for recombinantly producing a peptide or protein in a bacterial or yeast host cell, comprising

a) expressing in yeast or bacteria a fusion protein comprising the peptide or protein to be produced;

b) cleaving the fusion protein with a bovine enterokinase light chain analogue according to any one of claims 1-4; and

c) isolating the produced peptide or protein.

13. A method for recombinantly producing a peptide or protein according to claim 12, wherein the fusion protein expressed in step a) further comprises an Asp-Asp-Asp-Asp-Lys cleavage site.

14. A method for recombinantly producing a peptide or protein according to any one of claims 12-13, wherein the host cell is *E. coli.*

15. A method for recombinantly producing a peptide or protein according to any one of claims 12-14, wherein the peptide or protein to be produced is a GLP-1 peptide.

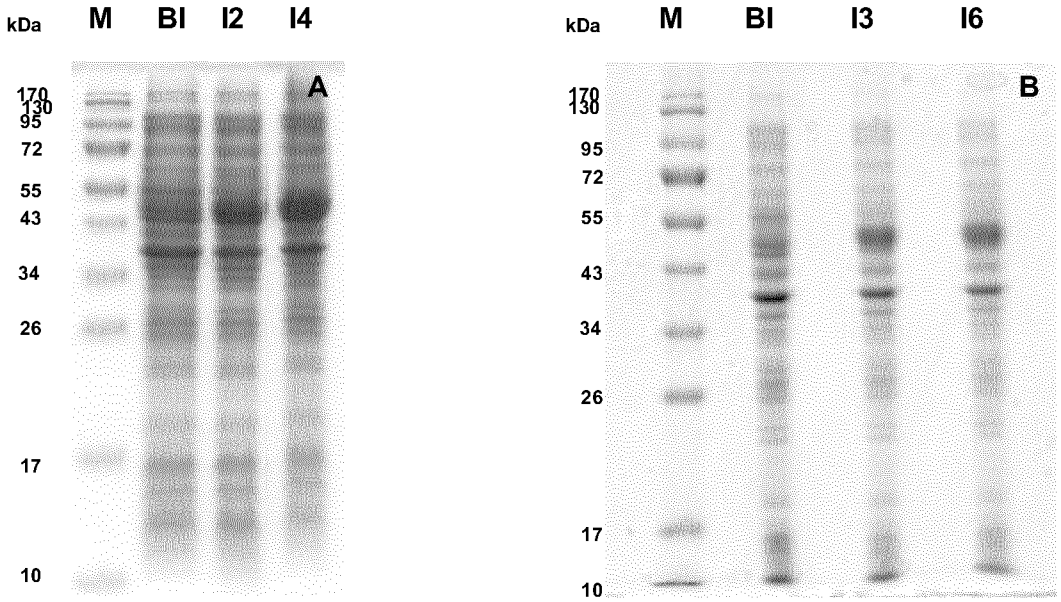


Figure 1

Fig. 1/11

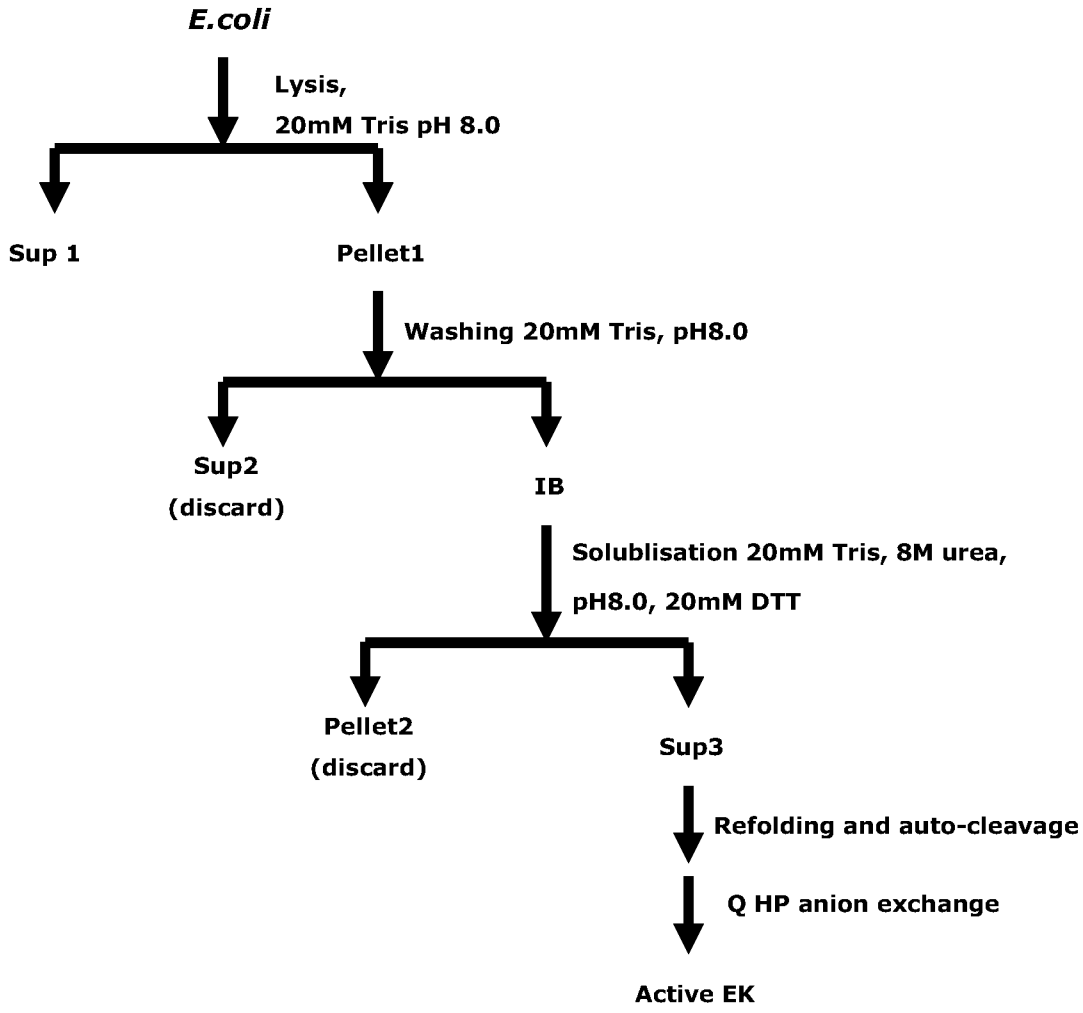
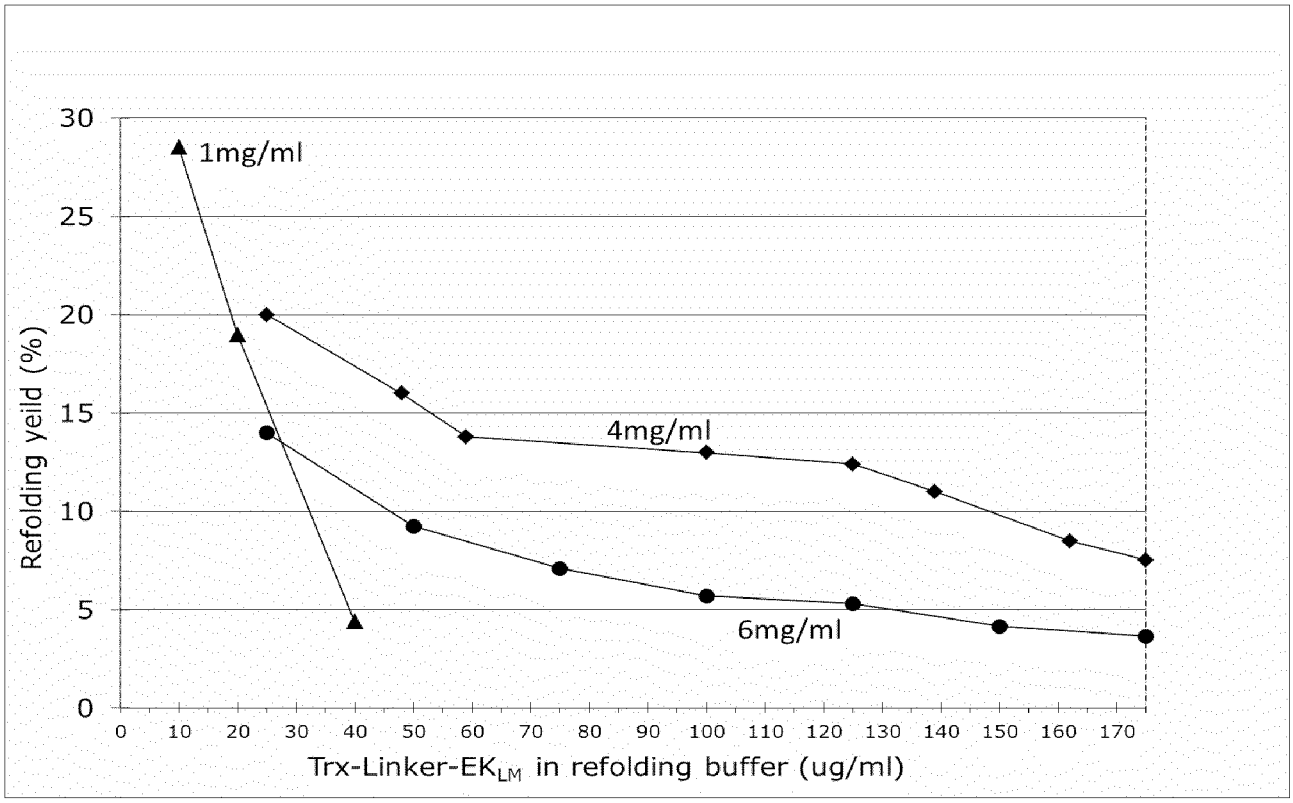


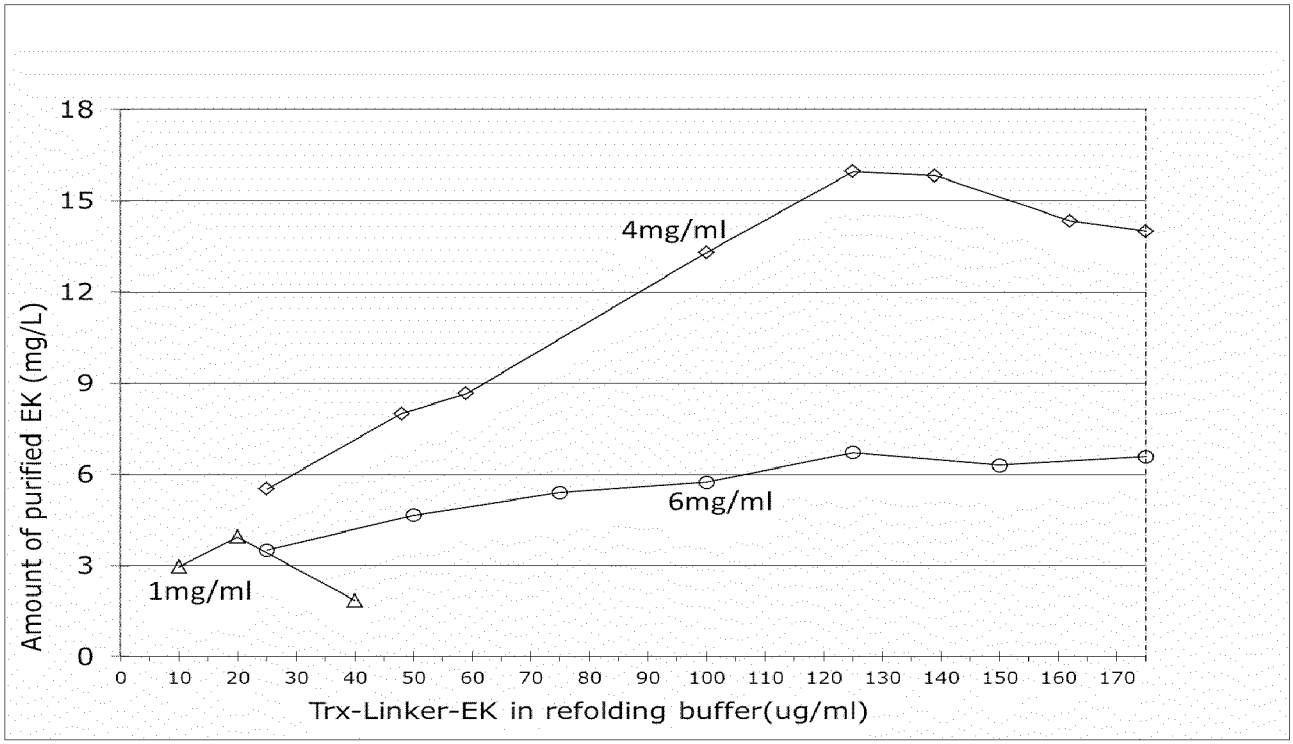
Figure 2

Fig. 2/11


30
25
20
15
10
5
0
Refolding yeild (%)
1mg/ml
4mg/ml
6mg/ml
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170
Trx-Linker-EK$_{LM}$ in refolding buffer (ug/ml)

3A

Fig. 3A/11


18
15
12
9
6
3
0
Amount of purified EK (mg/L)
4mg/ml
6mg/ml
1mg/ml
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170
Trx-Linker-EK in refolding buffer(ug/ml)

3B

Fig. 3B/11

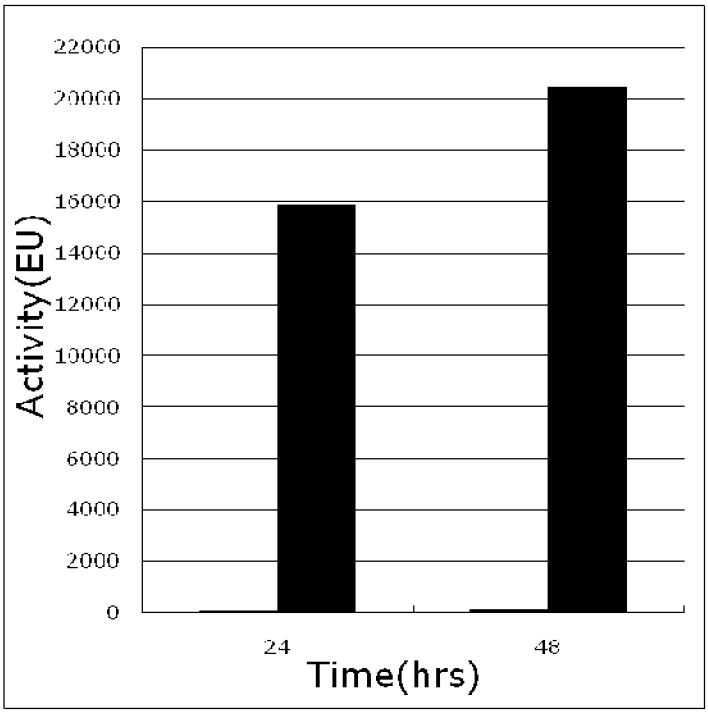


Figure 4

Fig. 4/11

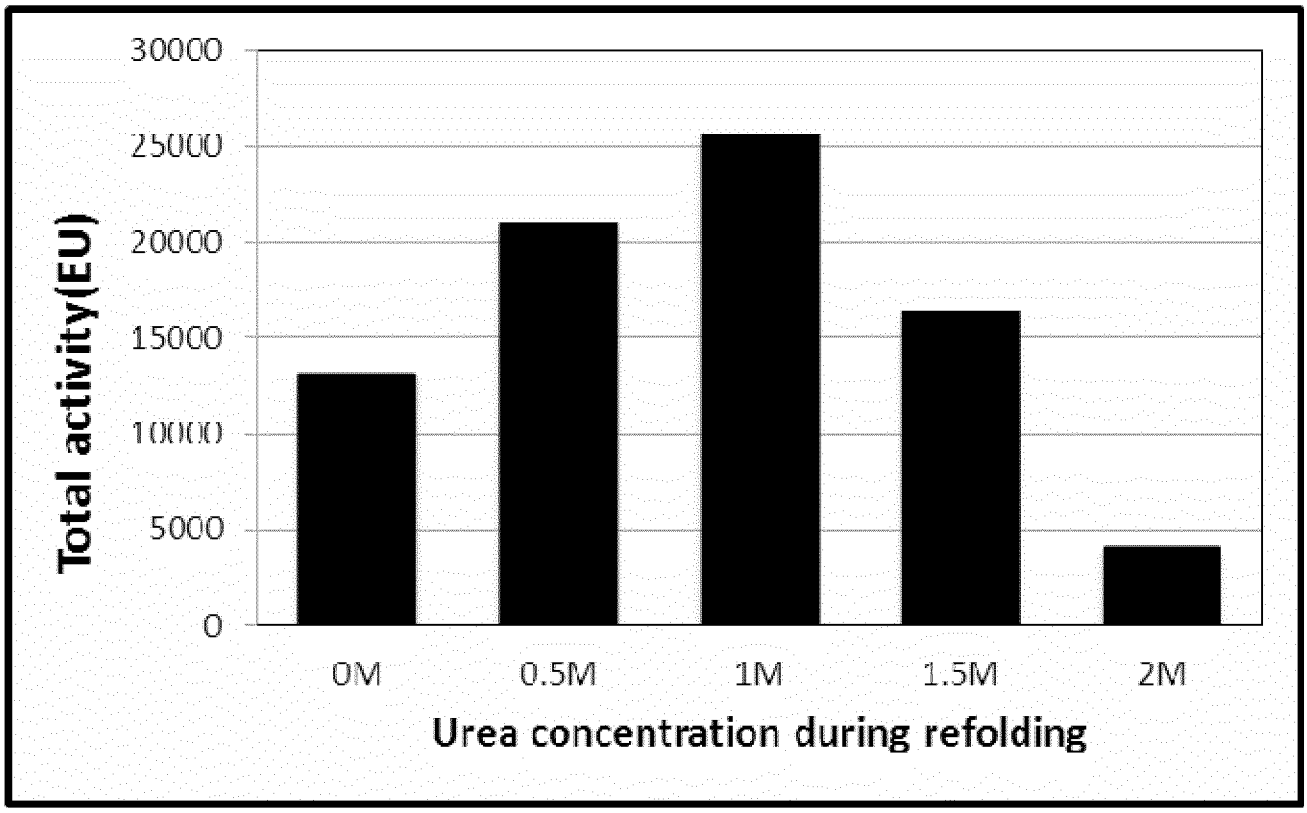

30000
25000
20000
15000
10000
5000
0
Total activity(EU)
0M
0.5M
1M
1.5M
2M
Urea concentration during refolding

Fig. 5/11

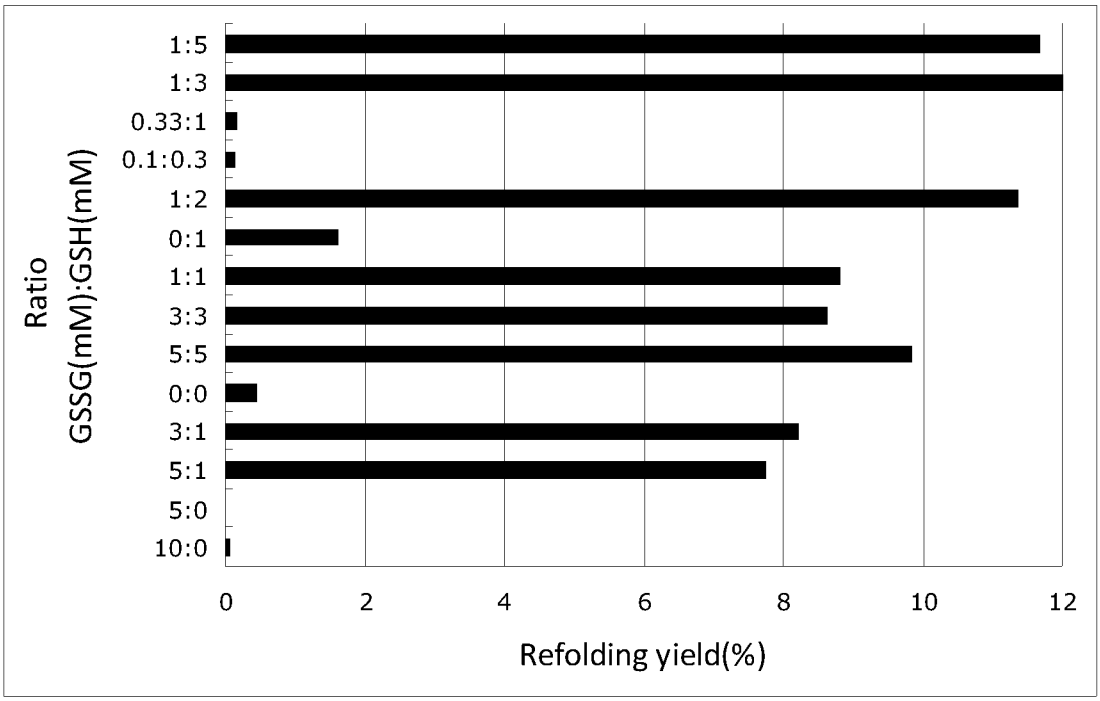


Figure 6

Fig. 6/11

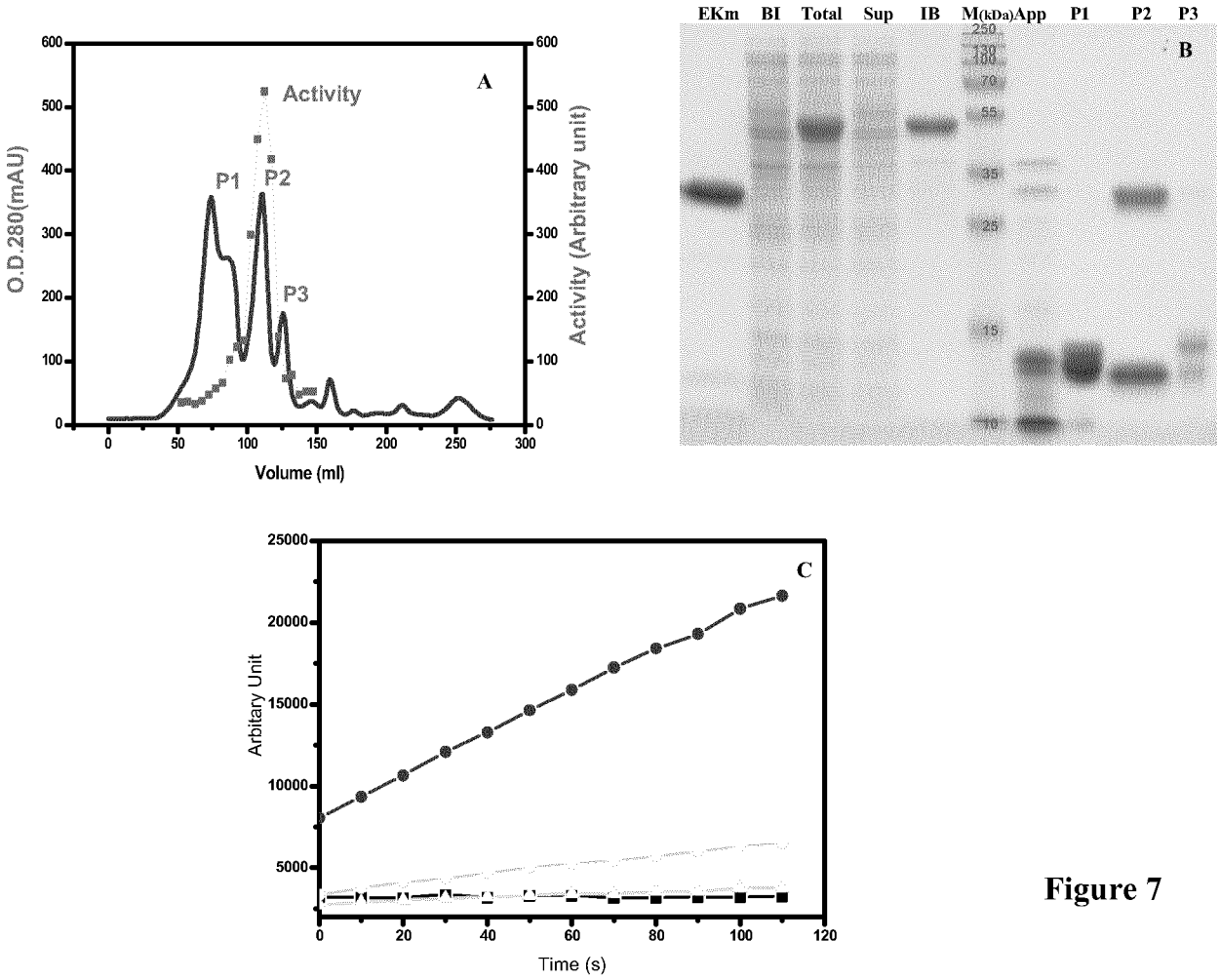


Figure 7

Fig. 7/11

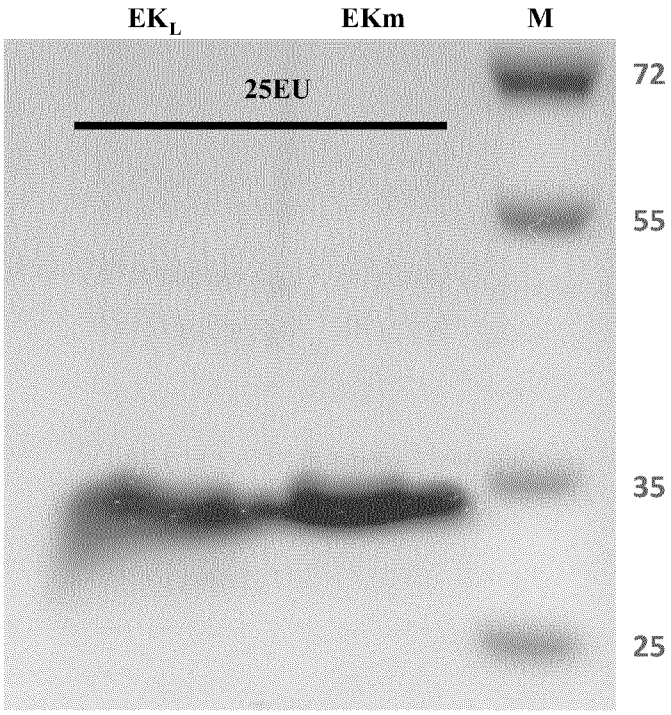


Figure 8

Fig.8/11

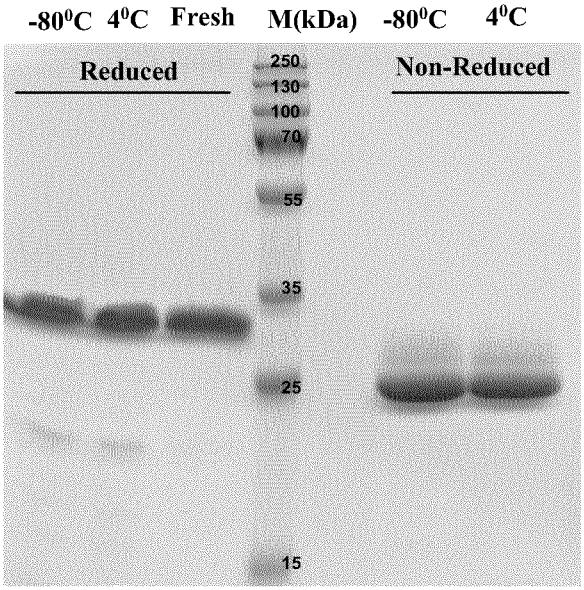


Figure 9

Fig. 9/11

```
                1                                                    50
Trx-Linker-EKLM MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY
       Trx-EKLM MSDKIIHLTD DSFDTDVLKA DGAILVDFWA EWCGPCKMIA PILDEIADEY

               51                                                   100
Trx-Linker-EKLM QGKLTVAKLN IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL
       Trx-EKLM QGKLTVAKLN IDQNPGTAPK YGIRGIPTLL LFKNGEVAAT KVGALSKGQL

              101                                                   150
Trx-Linker-EKLM KEFLDANLAG SGSGHMHHHH HHSSGLVPRG SGMKETAAAK FERQHMDSPD
       Trx-EKLM KEFLDANLAG SGSG...... .......... .......... ..........

              151                                                   200
Trx-Linker-EKLM LGTDDDDKIV GGSDSREGAW PWVVALYFDD QQVCGASLVS RDWLVSAAHC
       Trx-EKLM .GTDDDDKIV GGSDSREGAW PWVVALYFDD QQVCGASLVS RDWLVSAAHC

              201                                                   250
Trx-Linker-EKLM VYGRNMEPSK WKAVLGLHMA SNLTSPQIET RLIDQIVINP HYNKRRKNND
       Trx-EKLM VYGRNMEPSK WKAVLGLHMA SNLTSPQIET RLIDQIVINP HYNKRRKNND

              251                                                   300
Trx-Linker-EKLM IAMMHLEMKV NYTDYIQPIA LPEENQVFPP GRICSIAGWG AKKYQGSTAD
       Trx-EKLM IAMMHLEMKV NYTDYIQPIA LPEENQVFPP GRICSIAGWG AKKYQGSTAD

              301                                                   350
Trx-Linker-EKLM VLQEADVPLL SNEKCQQQMP EYNITENMVC AGYEAGGVDS CQGDSGGPLM
       Trx-EKLM VLQEADVPLL SNEKCQQQMP EYNITENMVC AGYEAGGVDS CQGDSGGPLM

              351                                                393
Trx-Linker-EKLM CQENNRWLLA GVTSFGYQCA LPNRPGVYAR VPRFTEWIQS FLH
       Trx-EKLM CQENNRWLLA GVTSFGYQCA LPNRPGVYAR VPRFTEWIQS FLH
```

Fig. 10/11

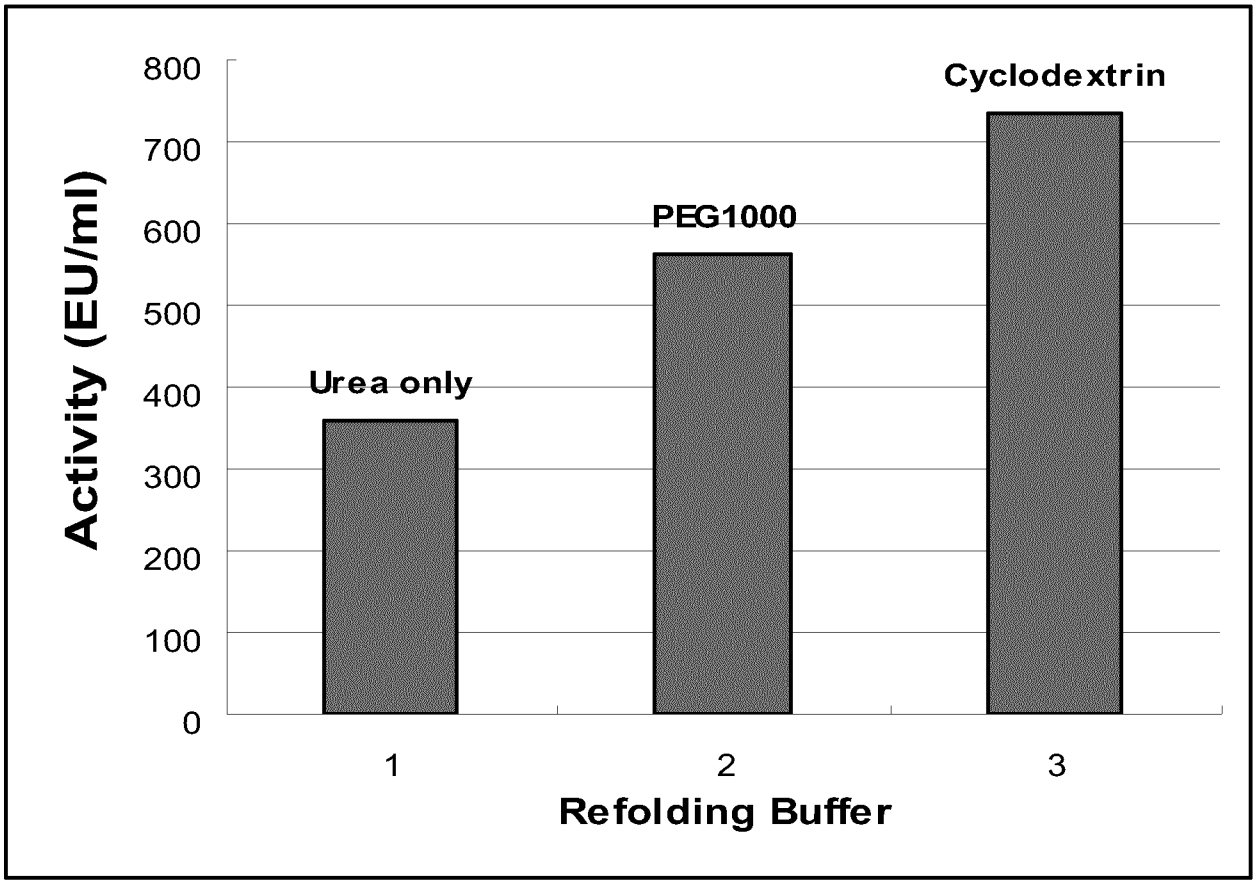

800
700
600
500
400
300
200
100
0
Activity (EU/ml)
Urea only
PEG1000
Cyclodextrin
1
2
3
Refolding Buffer

Fig. 11/11

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/076372

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/64 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)
C12N

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TAN ET AL: "Purification and refolding optimization of recombinant bovine enterokinase light chain overexpressed in Escherichia coli", PROTEIN EXPRESSION AND PURIFICATION, ACADEMIC PRESS, SAN DIEGO, CA, vol. 56, no. 1, 3 October 2007 (2007-10-03), pages 40-47, XP022284711, ISSN: 1046-5928, DOI: 10.1016/J.PEP.2007.07.006	10
Y	page 42, column 1 page 43, column 1	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" documentwhich may throw doubts on priority claim(s) orwhich 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
18 March 2013	03/04/2013

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	Mabit , Hel ène

1

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.
X	GASPARIAN M E ET AL: "Expressi on, puri fi cati on, and characteri zati on of human enteropepti dase catalyti c subuni t in Escheri chi a col i", PROTEIN EXPRESSION AND PURI FICATION, ACADEMIC PRESS, SAN DI EGO, CA, vol . 31, no. 1, 1 September 2003 (2003-09-01) , pages 133-139 , XP004454325 , ISSN: 1046-5928, D0I: 10. 1016/51046-5928(03)00159-1	10
Y	page 133 page 134, col umn 2, paragraph 2 - page 135, col umn 1, paragraph 1	1-15
Y	DATABASE WPI Week 200428 Thomson Sci enti fi c, London, GB; AN 2004-296039 XP002693849 , -& CN 1 470 634 A (LIU J) 28 January 2004 (2004-01-28) abstract	1-15
Y	HAARIN CHUN ET AL: "Desi gn and effi ci ent producti on of bovi ne enteroki nase light chai n wi th higher speci fi city in", BIOTECHNOLOGY LETTERS, SPRINGER NETHERLANDS, DORDRECHT, vol . 33, no. 6, 18 February 2011 (2011-02-18) , pages 1227-1232 , XP019903577 , ISSN: 1573-6776, D0I: 10. 1007/S10529-011-0562-3 page 1227	1-15
Y	LENA AAGREN ET AL: "Hydrophobi city engi neeri ng of chol era toxi n Al subuni t in the strong adjuvant fusi on protei n CTA1-DD", PROTEIN ENGINEERING, OXFORD UNIVERSITY PRESS, SURREY, GB, vol . 12, no. 2, 1 February 1999 (1999-02-01) , pages 173-178, XP008146027 , ISSN: 0269-2139 , D0I: 10. 1093/PR0TEIN/12 .2. 173 page 173, col umn 2, paragraph 3; tabl es 1, 2	1-15

-/--

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LU D ET AL: "Crystal structure of enteropepti dase light chai n compl exed wi th an anal og of the trypsi nogen acti vati on pepti de", JOURNAL OF MOLECULAR BIOLOGY, ACADEMIC PRESS, UNITED KINGDOM, vol . 292, no. 2, 17 September 1999 (1999-09-17) , pages 361-373 , XP004462288, ISSN: 0022-2836, DOI: 10. 1006/JMBI .1999.3089 the whol e document -----	1-15
Y	Z0U Z ET AL: "Hyper-aci dic protei n fusi on partners improve sol ubi l i ty and assi st correct fol di ng of recombi nant protei ns expressed i n Escheri chi a col i", JOURNAL OF BIOTECHNOLOGY, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol . 135, no. 4, 31 July 2008 (2008-07-31) , pages 333-339 , XP022939268, ISSN: 0168-1656, DOI: 10.1016/ J.JBI0TEC.2008.05.007 [retri eved on 2008-05-27] f i gure 1 -----	1-15
Y	SU Y ET AL: "The aci di ty of protei n fusi on partners predomi nantly determi nes the effi cacy to improve the sol ubi l i ty of the target protei ns expressed i n Escheri chi a col i", JOURNAL OF BIOTECHNOLOGY, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol . 129, no. 3, 1 May 2007 (2007-05-01) , pages 373-382 , XP026862966, ISSN: 0168-1656 [retri eved on 2007-04-13] f i gure 1 -----	1-15
Y	DATABASE Geneseq [Onl i ne] 1 September 2011 (2011-09-01) , "Targeted sol ubl e protei n TrxHi s, SEQ ID:5 .\ XP002693798, retri eved from EBI accessi on no. GSP:AZJ84253 Database accessi on no. AZJ84253 sequence the whol e document -----	1-15

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2012/076372

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
CN 1470634 A	28-01-2004	NON E	