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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 November 2008 (13.11.2008)

PCT

(10) International Publication Number
WO 2008/136014 A1

(51) **International Patent Classification:**
***C07K 14/435* (2006.01)**

(21) **International Application Number:**
PCT/IN2007/000277

(22) International Filing Date: 6 July 2007 (06.07.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
869/MUM/2007 7 May 2007 (07.05.2007) IN

(71) **Applicant** *(for all designated States except US):* **USV LIMITED** [IN/IN]; B. S. D. Marg, Govandi, Maharastra, Mumbai 400 088 (IN).

(72) Inventors; and

(75) Inventors/Applicants *(for US only):* **MAITI, Dipanwita** [IN/IN]; B. S. D. Marg, Govandi, Maharastra, Mumbai 400 088 (IN). **CHUTKE, Anjali** [IN/IN]; B. S. D. Marg, Govandi, Maharastra, Mumbai 400 088 (IN). **KHATRI, Priti** [IN/IN]; B. S. D. Marg, Govandi, Maharastra, Mumbai 400 088 (IN). **AHMED, Monsur** [IN/IN]; B. S. D. Marg, Govandi, Maharastra, Mumbai 400 088 (IN). **MUNESHWAR, Praveen** [IN/IN]; B. S. D. Marg, Govandi, Maharastra, Mumbai 400 088 (IN). **THAKAR, Niren** [IN/IN]; B. S. D. Marg, Govandi, Maharastra, Mumbai 400 088 (IN). **SANKARANKUTTY,** K. **Meera** [IN/IN]; B. S. D. Marg, Govandi, Maharastra, Mumbai 400 088 (IN). **MISHRA, Shrikant** [IN/IN]; B. S. D. Marg, Govandi, Maharastra, Mumbai 400 088 (IN).

(74) Agents: LAKSHMIKUMARAN, Varadachari et al.; Lakshmikumaran & Sridharan, B-6/10, Safdarjung Enclave, New Delhi 110029 (IN).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, **AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, 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, **PG, PH, PL, PT, RO, RS, RU, SC, SD,** SE, **SG, SK, SL,** SM, **SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, **DK,** EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, **RO,** SE, SI, SK, TR), OAPI (BF, **BJ,** CF, CG, CI, CM, GA, GN, GQ, GW, ML, **MR,** NE, SN, TD, TG).

Declaration under Rule 4.17:
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(Ui))*

Published:
— *with international search report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

(54) Title: CLONING AND EXPRESSION OF ENTEROKINASE AND ITS PROCESS FOR PRODUCTION

(57) **Abstract:** The present invention relates to cloning and expression of a novel nucleic acid sequence encoding Enterokinase, wherein the expression has been carried out in the protease positive background. The present invention further provides an improved process for production of Enterokinase. The method disclosed in the invention produces enterokinase protein as inclusion bodies in an active form. The specific activity of the enterokinase protein produce by the method disclosed in the present invention is comparable to or better than the commercially available Enterokinase.

WO 2008/136014 A1

CLONING AND EXPRESSION OF ENTEROKINASE AND ITS PROCESS FOR PRODUCTION

FIELD OF INVENTION

The present invention relates to cloning and expression of a novel nucleic acid sequence encoding enterokinase and an improved process for production of Enterokinase.

BACKGROUND OF THE INVENTION

The serine protease entrokinase (EK), also known as enteropeptidase, a heterodimeric glycoprotein, is a mammalian enzyme which catalyzes the conversion of trypsinogen into active trypsin. EK is highly specific for the substrate sequence (Asp) 4-Lys, where it selectively cleaves after lysine. EK isolated from bovine duodenal mucosa exhibits a molecular weight (MW) of 150,000 and a carbohydrate content of 35%. 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): 1677-1683 (1979)). The function of the heavy chain is to anchor the enzyme. The light chain acts as the catalytic subunit.

Several recombinant serine proteases including *enterokinase* light chain, trypsin and HIV protease have been expressed in *E. coli.* (EP0679189, US20030 157634, (Taylor A et al, Microbiol. Biotechnol, 1992, 37 (2) 205-210; Herber W. K et al, 1991, Appl. Microbiol. Biotechnol 36 (2) 149-152). The recombinant enterokinase has been successfully expressed in many systems including *E. coli,* yeast, *Saccharomyces cerevisiae, Pichia pastoris* and filamentous fungus *Aspergillus* (Collins Racie et al, 1995, Biotechnol 13, 982-987, Vozza et al, 1996, Biotechnol. 14, 77-81). In the case of prokaryotic systems like *E. coli, enterokinase* expression has been tried out with a fusion partner in the cytoplasm with thioredoxin, glutathione *transferase* and maltose binding protein as reported in WO 94/16083.

Specific recognition site other than DDDDK of EK in cleaving fusion peptide has been used purifying many recombinant proteins at laboratory scale. Hydrolysis at the unexpected site located towards the carboxyl end (Oi Wah Liew et al., 2005, Prot. Exp. Pur. 41: 332-340) or any other location on rDNA proteins significantly affects downstream purification process.

In *E. coli,* it is difficult to get EK_L in active form. In all cases reported so far the active protein has been successfully recovered only when it has been fused either to the thioredoxin protein or as a fusion construct with DsbA (E.R La Vallie et al 1993, J.Biol. Chem 31, 2331 1-23317, Huang He et al, 2004, Chin. Med. J, 117, 286-290). In all these *enterokinase* examples it is expressed in a background of a protease negative host with a fusion construct. (L.D Yuan et al, 2002, Prot. Express. Purif, 25, 300-304).

In expressing EK_L in *Saccharomyces cerevisiae, Pichia pastoris* or in *Aspergillus niger* the Kex2 protease cleavage expression of bovine *enterokinase* has been attempted and used as a fusion partner.

In the case of expression of Enterokinase light chain (EKL) in *E. coli* (Collin Racie et al, Biotechnology (N.Y.) 1995, 13, 982-987) in order to avoid formation of inclusion bodies and get active EK_L the cDNA was fused in frame to the 3' end of the coding sequence of DsbA (a thioredoxin homolog) with the two domains of the protein separated by an enterokinase recognition site. However most of the products were found to be non active and yields were low.

In another independent attempt (Yuan et al, 2002, Prot. Expression Purif., 25, 300-304) the *E. coli.* EK_L was cloned into a plasmid pET-32 and fused downstream to a fusion partner thioredoxin (Trx). In this case the protein was expressed mostly in the soluble form but recoveries reported were 4.3mg/100ml broth much higher than those reported earlier.

Another attempt reported by a Chinese group (Huang He et al, 2004, Chin. Med. J, 117, 286-290) mention an economical method of making EK_L was included following the prepro-secretion signal of the alpha mating factor of *Saccharomyces cerevisiae.*

In the case of expression of EK_L in *S. cerevisiae* the EK_L sequence was fused with a mammalian serine protease (PACE) as a secretory partner (Collin Racie Rehemetulla, 1993, J. Biol. Chem. 268 (31) 2331 1-23317) .Yields of 6.3mg/L were obtained in *E. coli* in these constructs.

In the expression of P-type ATPase in *E. coli,* plasma membranes from *Synechocystis* special feed were required to maintain construct stability and obtain reasonable expression levels (M, Gieseler, 1998, Biological Procedures Online VoI 1 no l.May 1, 70-80).

In the expression of recombinant streptokinase by continuous culture in *Escherichia coli* as also in the case of recombinant trypsin, recombinant protein activity can seriously damage host cells. So trypsin was expressed as the insoluble precursor trypsinogen in the form of inclusion bodies harvested and renatured by suitable methods (US 2003/0157634).

It was found that accumulation of toxic active streptokinase resulted in loss of cell viability. But changing the feed nature helped in reducing instability of the construct and helped expression of streptokinase (S S Yazdani, 2002, Bioprocess Biosyst Eng., 24, 341-347). It is postulated that in expressing toxic recombinant proteins the metabolic burden associated is far higher than normal proteins and therefore higher rates of plasmid instability and cell viability loss are observed. Therefore often high density cultures fail to enhance yields. In carrying out enterokinase fermentation it was found that the process required induction at very low cell densities at very early stages of fermentation. Induction at later stages did not lead to protein expression. This is unlike other recombinant *E. coli* fermentations, where induction was carried out after sufficient cell build-up. The fermentation was also found to be adversely affected by the levels of glucose at the beginning of the fermentation. Presence of trace amounts of glucose was found to be detrimental to expression, however once the culture attained logrithamic growth phase the glucose supplied was rapidly utilised and did not affect enterokinase expression. The feed rate to the culture was adjusted in a manner that prevented the accumulation of glucose in the medium, and thereby acetate accumulation was prevented.

The prior art shows that heterologous expression of the catalytic domain of bovine *Enterokinase* (EK_L) in *E. coli* and subsequent refolding to obtain pure soluble protein is not easily achieved. As disclosed (LaVallie et al., 1995, Biotechnology 13, 982-987), intracellularly expressed EK_L was inactive and attempts to produce EKL with various fusions did not give active enzyme. However, the same constructs when expressed as soluble secreted products, provided active enzyme after partial purification of the fusion protein. A similar observation was reported by Huang et al (Chin Med J; 117(2) 286-290, 2003), EK_L is fused to DsbA in *E. coli.*

US Patent No. 6,746,859 discloses the expression of catalytic subunit of enterokinase fused to *E. coli* thioredoxin (Trx) gene to provide better solubility. The fusion protein needs to be partially purified prior to initiation of autocatalytic cleavage. However, there is no mention

of the efficiency of this process or the final yield of *enetrokinase.* Gasparian et al reported expression of fusion protein of Trx and a synthetic gene for human EK_L in *E. coif* (Prot Exp and Purif, 31, Indian asianl33-139, 2003). No EK_L activity was recovered from soluble cytoplasm expressed TIX-EK_L even after partial purification. However, activity was recovered from protein expressed as inclusion bodies after a complex process of solubilization, dialysis and refolding.

SUMMARY OF THE INVENTION

The present invention relates to cloning and expression of a novel nucleic acid sequence encoding enterokinase and an improved process for production of *Enterokinase.* The method disclosed in the invention produces an active form of enterokinase.

The present disclosure relates to a process for cloning and expression of an Indian Asian EK_L-EK_L protein was expressed as inclusion bodies in a protease positive background and active protein has been recovered successfully. The construct does not contain a fusion tag unlike the earlier reports of enterokinase expression. The protein goes completely into inclusion bodies which are amenable to refolding. The specific activity (55,000 IU/mg) is comparable to or better than commercially available enterokinase. This enterokinase was utilized in downstream purification process of several specific rDNA generated protein of interest and compared with the performance of commercial EK_L. EK_L purified as described in the present invention has application in purification of chimeric recombinant proteins with inbuilt EK_L sensitive sites.

One aspect of the present invention provides an isolated nucleic acid comprising a polynucleotide sequence as shown in SEQ ID NO: 5, wherein said nucleic acid encodes for an enterokinase having an amino acid sequence as set forth in SEQ ID NO: 6.

Another aspect of the invention provides the recombinant vector comprising the polynucleotide sequence as set forth in the SEQ ID NO: 5 coding for the amino acid sequence as set forth in SEQ ID NO: 6.

Another aspect of the invention provides a host cell comprising the recombinant vector comprising the polynucleotide sequence as set forth in the SEQ ID NO: 5 coding for the amino acid sequence as set forth in SEQ ID NO: 6.

Further aspect of the present invention is to provide an improved process for production of enterokinase having the amino acid sequence as set forth in SEQ ID NO: 6 encoded by the polynucleotide sequence as set forth in SEQ ID NO: 5, wherein said method comprising

- culturing the host cells in the growth medium comprising inducer, wherein the host cell comprising the polynucleotide sequence as set forth in SEQ ID NO: 5 encoding the amino acid sequence as set forth in SEQ ID NO: 6;
- recovering the enterokinase;
- refolding of the enterokinase ; and
- purifying the enterokinase

Still another aspect of the present invention provides an improved process for production of enterokinase having the amino acid sequence as set forth in SEQ ID NO: 6 encoded by the polynucleotide sequence as set forth in SEQ ID NO: 5, wherein said method comprising:

- adding polynucleotide sequence as set forth in SEQ ID NO: 5 to the cell free transcription and translation system to produce enterokinase having amino acid sequence as set forth in SEQ ID NO: 6;
- isolating the enterokinase; and
- purifying the enterokinase

BRIEF DESCRIPTION OF ACCOMPANYING **DRAWINGS**

Figure 1 shows map of the recombinant vector pRA-EK

Figure 2 shows the effect of glucose levels in the medium on the expression of *enterokinase* in *E. coli.* .

Figure 3 shows the SDS-PAGE profile expression of His EK_L in the arabinose induced samples No HJS-EK_L protein is visible in the supernatant (lane 4).

Figure 4 shows the SDS-PAGE profile of the affinity purification of His-EKL by expanded bed chromatography on streamline chelating beads

Figure 5 shows the SDS-PAGE profile of refolded STI purified EK_L

Figure 6 shows SDS-PAGE profile of purified EK_L compared to commercially available EK_L

Figure 7 shows the kinetic assay comparison of purified EK_L versus the commercially available EK_L

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to cloning and expression of a novel nucleic acid sequence encoding enterokinase and an improved process for production of *Enterokinase.* The method disclosed in the invention produces an active form of enterokinase.

The present disclosure relates to a process for cloning and expression of an Indian Asian EK_L. EK_L protein was expressed in a protease positive background, protein expressed as inclusion bodies and active protein has been recovered successfully. The construct does not contain a fusion tag unlike the earlier reports of enterokinase expression. The protein goes completely into inclusion bodies which are amenable to refolding. The specific activity (55,000 IU/mg) is comparable to or better than commercially available enterokinase. This enterokinase was utilized in downstream purification process of several specific rDNA generated protein of interest and compared with the performance of commercial EK_L. EK_L purified in the present invention has application in purification of chimeric recombinant proteins with inbuilt EK_L sensitive sites.

The present invention provides isolation of a polymorphic nucleic acid sequence from *Bos taurus* of western Indian Asian origin which was expressed in *E. coli* and has *enterokinase* activity in its non-glycosylated form. EK_L was purified from the inclusion bodies by using a novel downstream purification process and refolded to its specific activity compared to commercial ones.

One embodiment of the present invention relates to an isolated nucleic acid comprising a polynucleotide sequence as shown in SEQ ID NO: 5, wherein said nucleic acid encodes for an enterokinase having an amino acid sequence as set forth in SEQ ID NO: 6.

Another embodiment of the present invention provides a recombinant vector comprising the nucleic acid having a polynucleotide sequence as set shown in SEQ ID NO: 5 operably linked to an inducible promoter.

Still another embodiment of the invention provides the recombinant vector comprising the nucleic acid comprising a polynucleotide sequence as shown in SEQ ID NO: 5 operably linked to an inducible promoter, wherein said inducible promoter is selected from a group consisiting of AraB, T7, trp, lac, tac.

Further embodiment of the invention provides a host cell comprising the recombinant vector comprising the polynucleotide sequence as shown in the SEQ ID NO: 5 coding for the amino acid sequence as shown in SEQ ID NO: 6.

Still yet another embodiment of the invention provides the host cell comprising the recombinant vector comprising the polynucleotide sequence as set forth in the SEQ ID NO: 5 coding for the amino acid sequence as shown in SEQ ID NO: 6, wherein the host cell is selected from a group consisting of *E .coli,* CHO, yeast and eukaryotic cell lines.

Present invention also provides an improved process for production of enterokinase having the amino acid sequence as shown in SEQ ID NO: 6 encoded by the polynucleotide sequence as shown in SEQ ID NO: 5, wherein said method comprising

- culturing the host cells in a growth medium wherein the host cell comprising the polynucleotide sequence as shown in SEQ ID NO: 5 encoding the amino acid sequence as shown in SEQ ID NO: 6;
- recovering the enterokinase;
- refolding of the enterokinase ; and
- purifying the enterokinase

Further embodiment of the present invention relates to the process for production of enterokinase, wherein host cells were cultured in the growth medium comprising sources of carbon and nitrogen in a ratio of 1:1 to 2:1.

Additional embodiment of the present invention provides the growth medium for culturing the host cells wherein the medium further comprises an inducer selected from a group consisting of Isopropyl-beta-D-thiogalacto pyranoside (IPTG), lactose, maltose, arabinose and arabino galactan.

One embodiment relates to the growth medium used for the culturing the host cells comprising arabinose as an inducer.

Another embodiment of the present invention relates to incubation temperature of the host cells. The host used for the production of enterokinase were incubated at a temperature ranging from 30-37^{0}C.

The process for production of enterokinase disclosed in the present invention comprises refolding of the enterokinsae, wherein the refolding carried out by sulfitolysis followed by oxidative refolding in red-ox glutathione environment.

The present invention discloses the process for production of enterokinase, wherein enterokinase is expressed in a protease positive background.

Further the present invention provides an improved process for production of enterokinase having the amino acid sequence as set forth in SEQ ID NO: 6 encoded by the polynucleotide sequence as set forth in SEQ ID NO: 5, wherein said method comprising:

- adding polynucleotide sequence as set forth in SEQ ID NO: 5 to the cell free transcription and translation system to produce enterokinase having amino acid sequence as set forth in SEQ ID NO: 6;
- isolating the enterokinase; and
- purifying the enterokinase

Nucleic acid forms such as genomic DNA (gDNA), complementary DNA (cDNA), a synthetic DNA prepared by chemical synthesis as well as DNA with deletions or

mutations, allelic variants and sequences that hybridize thereto under stringent conditions are also within the scope of the present invention as long as they encode polypeptides having enterokinase activity as defined in the present invention.

Association of nucleic acid sequences provided by the invention with homologous or heterologous species expression control sequences, such as promoters, operators, regulators, and the like, allows for *in vivo* and *in vitro* transcription to the corresponding mRNA which, in turn, allows translation of proteins and related poly-and oligo-peptides, in large quantities, having enterokinase activity.

In a presently preferred expression system of the invention, enterokinase encoding sequences are operatively associated with a promoter sequence allowing for transcription and translation in a eukaryotic cell system to provide e.g., enterokinase polypeptides having protease activity.

The novel nucleic acid sequences encode the light chain of enterokinase which surprisingly provides enterokinase activity. The enterokinase activity of the invention may be generated from one or more expression vector(s) each comprising one or more portions of the enterokinase activity, or, alternatively, the enterokinase activity can be generated from one or more expression vector(s) contained in one or more cell lines, each of which express all or a portion of the enterokinase activity.

The nucleotide sequence encoding for enterokinase activity may be inserted into an expression vector by conventional method. Incorporation of these recombinant vectors comprising these sequences into prokaryotic and eukaryotic host cells by standard transformation and transfection processes is also within the scope of the present invention and is expected to provide useful enterokinase in quantities greatly in excess of those obtainable from tissue sources. Appropriate host cells include *E. coli*, CHO, yeast, and lepidoptera cells.

The non-glycosylated protein products of the present invention include those having the primary structural conformation (i.e. amino acid sequence) of enterokinase comprising the sequence substantially as set forth in SEQ ID NO: 6 encoded by the polynucleotide sequence as shown in SEQ ID NO: 5 and having enterokinase protease activity.

One embodiment of the invention provides the amino acid sequence substantially as set forth in SEQ ID NO: 6

The present invention provides detailed description of the invention which includes numerous illustrative examples of the practice of the invention reference being made to the Sequence Listing wherein: SEQ ID NO: 5 provides the polynucleotide sequence of the bovine enterokinase sequence and SEQ ID NO: 6 provides the predicted amino acid sequence of the entire catalytic domain (light chain) of bovine enterokinase.

The total RNA was isolated from bovine tissue by a method known in the art (DNA Cloning 1: a practical approach, 2nd Edition, 1995, Glover and Hames). The method involves isolation of total RNA using guanidium thiocyanate (Ausuble et al., 1987, Current protocols in Molecular biology VoI 1 4.1 John Wiley and Son Inc). Poly (A+) RNA was purified by affinity chromatography on Oligo(dT) cellulose. The mRNA was separated by Oligo(dT) cellulose affinity chromatography from the total RNA, which was homogenized. The resulting mRNA was then used for generating l st-strand by AMV transcriptase (Maniatis et al, 1989). Non-specific interactions between strands were reduced using RNase H and RNase A, followed by purification of the I st-strand DNA (20 ng/ul) using phenol/chloroform/isoamyl alcohol and 70% ethanol wash. PCR amplification was done using *Pfu* polymerase to generate double-stranded cDNA prior to cloning the gene of interest into an appropriate choice of host-vector system *E. coli* or yeast preferably *E. coli.* This 732 bp fragment of interest was then directly cloned into an inducible expression vector. Transformants were selected against ampicillin selection; inserted fragment was screened by restriction enzyme analysis of the DNA fragment using suitable inducers. Details are provided in Example 1. Expression of enterokinase in transformants was checked in. Details are provided in Example 2.

Culturing transformants in medium density culture condition (Figure 2), the expressed protein accumulates in the inclusion bodies (Figure 3), and refolded into its native state (Figure 4).

This putative DNA sequence may encode both the heavy and light chains, or only light chain, or truncated light chain or a variant with mutations in amino acids in the light chain

with biologically active mature peptide sequence encoding enterokinase activity without carrying its heavy chain.

The invention discloses a homologous mature light chain of enterokinase cDNA fragment with the affinity tag which shows the mutations at the amino acid level at six eight sites with the corresponding wobble changes at the cDNA level at six eight places and subsequent wobble changes at the cDNA level at additional six places without altering the amino acids of native EKL. The EKL of the present invention as shown by polynucleotide sequence of SEQ. ID No. 5 encoding amino acid sequence of SEQ. ID No. 6 is unique in having the following mutations: S5N, R82P, N90D, N103T, P12IS, A133T, D176E and T227A with respect to the published EKL sequence. The expressed and purified light chain variant shows full enterokinase specific activity efficiency at the level of commercial EK. This example encounters the possibility of genetic polymorphism in bovine EK light chain. This is encoded by the cDNA of 732 bp sequence (SEQ ID NO: 5).

The polynucleotide sequence (SEQ ID NO: 5) coding for polymorphic bovine enterokinase was cloned into an inducible expression vector. The use of *E. coli* platform makes it possible for high yield purification of enterokinase with desired specific activity.

The invention also relates to the culturing of the selected clone expressing enterokinase in a fermentation medium containing a suitable carbon and nitrogen source in an appropriate ratio. The range of carbon sources that can be used could include glucose, glycerol and other sugars which could be used in appropriate concentrations. A range of nitrogen sources have been used to carry out fermentation ranging from yeast extracts to tryptones and peptones of both plant and animal origins. An appropriate ratio of carbon to nitrogen during fermentation helps in optimal growth, product formation and recovery. The induction of the protein can be carried out using a suitable inducer, in this case arabinose was used as an inducer and the time at which induction was carried out was found to greatly affect expression of the desired protein. Therefore induction was carried out at different time points during the lag and log phase of the culture and it was found that the protein expression was limited to very early lag phase of growth when the fermentation was carried out in a batch model. Induction during fed batch mode of growth did not lead to protein induction at higher cell densities. The presence and concentration of carbon source used was found to adversely affect the expression of the proteins and this limited

the use of glucose to lower percentages although higher glucose feeding lead to higher cell densities. The fermentation was found to generate minimum acid during growth as the carbon feeding to the culture was restricted and only moderate cell densities were achieved. The detailed procedure is described in Example 3.

The host cell containing the recombinant expression vector comprising the novel polynucleotide sequence of the enterokinase as shown in SEQ ID NO: 5 was grown in the culture medium containing the appropriate antibiotics and the recombinant protein expressed under appropriate conditions. The cells are then harvested using conventional techniques like centrifugation at 5000 x g for 15 min. The culture supernatant can be drained off and the cells resuspended in buffer of choice, prior to disruption. Alternatively use of tangential flow filtration with appropriate membranes porosity can be used to harvest, wash and concentrate the cells thereby removing the impurities present in culture media prior to disruption (Example 4, 5 and 6).

To obtain the expressed protein, the cells have to be disrupted. Various methods are available in prior art that describes means to disrupt cells to obtain inclusion bodies. Preferred methods include sonication, bead milling, french press or high pressure disruption. In the present invention, the cell suspension in buffer was passed through a two stage high pressure homogenizer operating at 800-900 bars to disrupt cells. The disruption cycle was carried out thrice, with cooling in between passes to ensure maximal breakage of cells and reduction in viscosity due to shearing of nucleic acids released from the broken cells (Example 7).

Isolation of inclusion bodies (IB's) from the crude cell lysate is a well-defined process and is usually achieved by low speed centrifugation at 10,000 x g for 10-15 minutes. Inclusion bodies having higher density tend to settle faster and can be separated from soluble and other particulate matter in the crude cell lysate. The protein pellet obtained after centrifugation can be dispersed in suitable buffer containing low concentrations of denaturant e.g., 2 M urea in addition to having non-ionic detergents like *2%* Triton-X 100. This combination helps to solubilize most of the protein impurities that pellet down with the inclusion bodies and a similar centrifugation step as above yields inclusion bodies that have substantially higher content of the protein of interest. Alternatively the IB's could be washed with buffer containing low concentrations of ionic detergent e.g., N-laurly sarkosyl

to remove most of the impurities. The purified IB's are then solubilized in high concentration of denaturant like 6 M Guanidine hydrochloride or 8 M urea in appropriate buffer and pH, the protein of interest can be purified by affinity chromatography (Example 8).

In present invention solid urea crystals were added to the crude cell lysate crystals to achieve 8 M final concentration and stirring the solution for 14-20 hrs at room temperature (20-25^{0}C). This enhanced the solubilization of *enterokinase* IB's producing homogenous mixture of soluble His- tagged *enterokinase* and other cellular proteins. This crude mixture was then pumped into a streamline chelating sepharose column run in expanded mode. Under appropriate conditions only His- tagged *enterokinase* was bound to the column matrix and was purified from all the cellular impurities. Elution with a higher concentration of imidazole, preferably at 100-250 mM released the His-tagged Enterokinase from the matrix. This allowed one step purification of the His- tagged Enterokinase to >85% with very little cellular impurities Example 9.

The purified tagged *enterokinase* solution in 8 M urea and buffer needs to be buffer exchanged to remove salt and imidazole present from the earlier step. This is easily achieved on a Sephadex G-25 column. Alternatively tangential flow filtration with appropriate Molecular weight cut off column membrane can be used in diafiltration mode to remove salt and imidazole to acceptable level (Example 10).

Sulfitolysis reaction is a process that is well reported in literature since 1960s. It involves a simple chemical reaction by which disulfide bonds in protein molecules are reduced and reversibly protected in the form of S-sulfonates. This is especially useful in case of recombinant proteins that have disulfide bonds and are expressed as inclusion bodies. In such a case there is high probability of improper disulfide bond formation both intra as well as inter-molecular, leading to misfolded protein structures with altered or no activity. Sulfitolysis process reversibly reduces all the disulfide bonds in protein molecules and allows for proper disulfide bond formation during the refolding process. The process involves reacting proteins, either in native state or denatured form with molar excess of sodium sulfite or bisulfite and in combination with sodium tetrathionate or thio bis-nitrobenzoic acid. Alternatively proteins can be reduced with other reducing agents like

dithiothreitol (DTT) and then oxidized to form refolded moiety. Details are given in Example 11.

Refolding process is an intrinsic property of any protein needed to achieve its functional property. This complex process occurs *in vivo* and is often an assisted procedure wherein other protein molecules known as chaperones help in attaining the correct structure and function of a protein. Although large amounts of heterologous proteins could be expressed in *E. coli,* more often than not it is in the form of inactive aggregates called as inclusion bodies. These need to be isolated from other cellular proteins, often achieved by cell disruption and centrifugation, solubilized in a buffer of choice often containing high concentration of denaturants like guanidine hydrochloride or urea. Other modes of solubilization include high or low pH buffer, or detergents or organic solvents. Once solubilized the proteins need to refold back to their native form. This is often achieved by diluting the protein solution in buffer that allow for proper refolding and disulfide bridge formation in protein. The choice of buffer constituents is highly dependent on the protein to be refolded and cannot be generalized. The use of glutathione redox system for refolding of proteins is very popular and innumerable reports exist of its use in achieving the desired result in a large variety of proteins expressed in *E. coli.* Other systems used are those containing cysteine-cystine and reduced and oxidized DTT. Copper sulfate in very minute concentrations has also been reported to facilitate formation of proper disulfide bond formation during refolding in proteins. Both glutathione as well as cysteine-cystine system were used and observed that glutathione system provides optimal refolding results and activity of recombinant bovine Enterokinase. For detailed procedure please see Example 12.

Post refolding the purification process must involve a step that quickly concentrates the active protein, stabilizes it and removes most of the components used in refolding process. Concentration using Tangential flow filtration provides a rapid method of reducing volume and removal of refolding additives could be achieved by diafiltration into appropriate buffer. The refolded enterokinase was concentrated using a 5 kDa MWCO membrane. Other method includes use of anion exchanger to bind the refolded enterokinase after dilution of the refolded solution to bring the conductivity down to acceptable levels. Bound protein was eluted from the column using sodium chloride gradient. Active fraction was

pooled and taken to next step of purification. Concentrated enterokinase was finally purified to homogeneity by affinity chromatography on either soybean trypsin inhibitor agarose or benzamidine agarose. Both are inhibitors of serine proteases that include trypsin, chymotrypsin and *enterokinase.* These ligands have strong and specific binding affinity to the catalytic domain of serine proteases. Only the properly folded *enterokinase* having the correct catalytic domain conformation will bind these ligands and the unfolded or misfolded molecules will not bind to the matrix. Elution of the bound molecules can be achieved by low pH buffer or organic solvent containing buffer. Soybean trypsin inhibitor agarose was used and eluted the bound refolded and active enterokinase with low pH glycine buffer. The pH of the eluted protein was adjusted with unbuffered Tris solution assayed for activity and protein concentration and stored as 50% glycerol mixture to provide favorable storage condition and shelf life (Example 13).

Kinetic assay and colorimetric assay of enetrokinase enzyme activity is described in detail in Example 14, Example 15 respectively.

The gene-specifc primers were used on total RNA isolated from Indian bovine duodenum to get the RT-PCR product, 762 bp cDNA, of the catalytic chain of Enterokinase (EK_L).

A 732 bp of cDNA encoding catalytic domain (SEQ ID NO: 5) of bovine enterokinase (EK_L) was cloned into *E. coli* expression vector.

This EK_L CDNA sequence is unique at its start of the open reading frame which contains 24 nucleotides coding eight amino acids extra at the starting of N-terminus which are IVWGFADK compared to native bovine enterokinase amino acid sequence.

Nucleotide sequence of cloned cDNA of EKL contains a 235 amino acid polypeptide that does not share a complete match with the published bovine Enterokinase light chain sequence.

a) 8 amino acids those are different in the present EK polypeptide sequence which is a novel sequence.

b> 6 amino acids which have shown their differences in their wobble sequence on the triplet at the cDNA level of EK_L

Expression of enterokinase was localized in the inclusion bodies of *E. coli* or in the secreted material of yeast. This enterokinase has shown to be an enzymatically active protease which cleaves specifically after (Asp)-4-Lys (DDDDK) sequence in various numbers of fused protein products between affinity tag and the mature cDNA translated product and shown 100% efficiency in its enzymatic activity.

These modifications of amino acids in the light chain of *Enterokinase* (EK_L) are novel which perhaps evolved through geographical polymorphism.

A process for preparing mutant enterokinase protein in *E. coli* cells transformed with a plasmid carrying the *enterokinase* gene and an inducible promoter by low cell density fermentation involving batch and fed batch stages and isolation and purification of the expressed protein from the cultures.

Refolding process for bovine catalytic subunit of enterokinase contributed in the expression of the enterokinase in the form of inclusion bodies in recombinant *E. coli* wherein use of sulfitolysis followed by refolding on glutathione redox system.

While the invention is broadly as defined above, it will be appreciated by those persons skilled in the art that it is not limited thereto and that it also includes embodiments of which the following description gives examples.

EXAMPLES

It should be understood that the following examples described herein are for illustrative purposes only and that various modifications or changes in light will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and the scope of the appended claims.

Example 1

Total RNA Isolation and synthesis **of first strand** of **cDNA**

Total RNA was isolated from Bovine duodenum by the method known in the art. About 2 g of tissue was processed for RNA isolation. Total RNA pellet was resuspended in ImI nuclease-free water and stored at -70 ^{0}C. About 4 ng/ml RNA was obtained.

Polyadenylated RNA was pooled with the help of Oligo (dT) 8+12 beads from total RNA isolated as described above. 10 ng of total RNA was used for each reaction. Polyadenylated RNA was reverse-transcribed into single-stranded complementary DNA (1st-strand cDNA template) using AMV reverse transcriptase at 42^{0}C for 1 hr. About 20ng/ul lst-strand cDNA was obtained.

PCR amplification of putative light chain cDNA of bovine enterokinase (EK_L)

Preparation of frafiment 1 and fragment 2

A 762 bp enterokinase catalytic chain cDNA was amplified using the bovine duodenum 1st-strand cDNA as the template. Two gene-specific primers P1 and P2 having *Drd* I and *EcoR* I restriction sites respectively were used PCR amplification. These primers were designed from the bovine enterokinase polynucleotide sequence reported by La Value ER *et al.* (J Biol Chem, 268 (31), 23311-17, 1993). The nucleotide sequence of the primers used is provided in SEQ ID NO: 1 and SEQ ID NO: 2.

P1 : 5'-AAAAAATTAGCGACAAGATTGTCGGAGGAA-S' SEQ ID NO: 1

P2: 5-GTCTGGGAATTCTTTAATGTAGAAAACTTTGTAT-S 'SEQ ID NO: 2

Polymerase chain reaction was carried out by using 100 ng of 1st-strand cDNA in each 50µl of reaction. 50 pmoles of primers were used for each 50 µl of PCR reaction. PCR was performed in 4 stages for total of 40 cycles. Denaturation at 94^{0}C for 2 min; annealing at 46°C for 2 min; extension at 72^{0}C for 2 min (stage 1); denaturation at 94^{0}C for 1 min, annealing at 49^{0}C for 2 min, extension at 72°C for 2 min (stage 2); Final extension at 72^{0}C for 10 min (Stage 3). *Pfu* polymerase was used in presence of MgSO4. 10 µl of PCR reaction was loaded into each lane from each 50 µl total reaction mix (Stage 4).

A 762 bp Enterokinase catalytic chain cDNA was obtained and amplified PCR fragment was checked on 1% agarose gel. The amplified product was purified and digested with *Ndel* restriction enzyme. Digestion of the PCR product with *Ndel* gave rise to two fragments of a 220 bp and a 542 bp size fragments. 220 bp DNA fragment thus obtained was further digested with *Drdl* to obtain a DNA fragment of around 202 bp size designated as fragment 1. The second DNA fragment of 542 bp was digested with *EcoRl* and a

fragment of 532 bp designated as fragment 2 was purified through column. Both the fragments were purified through column purification process.

Preparation of fragment 3

pRA vector having *Sspl* restriction sites at the positions 928 and 3695 was digested with *Sspl* restriction enzyme to obtain 1292 and 2767 bp DNA fragments. Fragment of 1292 bp size was purified from the gel and used as the template for the subsequent PCR.

PCR amplification was carried out with the template DNA of 1292 bp size using the primers P3 and P4 having polynucleotide sequence as shown in SEQ ID NO: 3 and SEQ ID NO: 4.

P3: 5'-GCGTCTTTTACTGGCTCTTCTCGCTAA-S' SEQ ID NO: 3

P4: 5'-GGATCCCCAGACAATCTTGTCGTCGTCGTCGT-S 'SEQ ID NO: 4

A PCR product of 342 bp was thus obtained was purified from gel digested with *BshTl (Agel)* and *Drdl* to obtain a 296 bp DNA fragment and is designated as fragment 3.

Ligation and Transformation

The pRA vector used for EKL expression includes the arabinose inducible promoter. The expression level is otherwise suppressed by the repressor control system which is the internal/external glucose level of the cell itself. The expression level of the cloned ORF (EKL) is enhanced several times by the presence of L-arabinose in a culture medium for 4-6 hrs.

Expression vector pRA was digested with *BshTl (Age I)* and *EcoR* / to obtain 3717 bp fragment designated as fragment 4. The digested vector DNA was gel purified.

Fragment 1 and fragment 2 as described above were ligated to obtain the 732 bp DNA fragment of light chain of the bovine enterokinase gene and is designated as fragment 5. This 732 bp DNA fragment was analyzed by sequencing. The polynucleotide sequence of final cDNA of 732 bp thus obtained consists of the 24 nucleotides at the 5'-end of the light chain 732 bp cDNA. The polynucleotide sequence of this 732 bp cDNA is shown in the SEQ ID NO: 5 and is designated as fragment 5. The amino acid sequence of the

polypeptide encoded by the polynucleotide sequence as shown in SEQ ID NO: 5, is shown in the SEQ ID NO: 6.

The ligation mixture containing 20 ng-50 ng DNA of fragment 1 and 2 was incubated overnight at 4^0C to increase the number of transformants. This gave rise a 732 bp fragment (fragment 5) of cDNA EKL. Further ligation of fragment 5 and fragment 3 (296 bp, vector portion) was carried out to obtain the 1030 bp size fragment designated as fragment 6. The fragment 6 (1030 bp) was then ligated to the fragment 4 (3717 bp, pRA vector backbone), to obtain final recombinant vector comprising the 732 bp of bovine enterokinase. This recombinant vector was designated as pRA-EKG100 (Figure 1). The ligation reaction was incubated at 4^0C for overnight. The ligation mixture was transformed in competent cells TOP 10 strain of *E. coli.* At the heat shock step ligation mix and cells were incubated at 42^0C for 2 min. Later plating was done on LB agar plates containing 50 µg/ml ampicillin. Plates were incubated at 37^0C in the incubator for overnight. More than 100-150 colonies were obtained.

Verification of the putative clone by restriction enzyme **map** analysis

To determine the unique sites on the ORF as well as in the vector construct, restriction enzyme analysis with *BsrG\, SaWXJiol, NcollEcoRl* was done on DNA isolated from the putative clone comprising the 732 bp fragment of bovine enterokinase gene by alkaline lysis method as described by Maniatis *et* «/.1989. Plasmid DNA isolated from the putative clone was digested with these enzymes at the final volume of 50 µl using 1µl of 10 Units of enzyme. After short spin at 13,000 g, reaction mix was incubated at 37^0C for 1 ½ hrs. This reaction mixes were then loaded onto 1% agarose gel for confirmation of the respected band of our interest. Positive clones containing 732 bp EKL insert were selected.

Example 2

Expression of Enterokinase **in** *E. coli*

Positive clones (4-5 clones) containing the 732 bp of EK_L insert were inoculated in the 2X YT medium (Yeast extract containing tryptone medium) comprising yeast extract, tryptone, NaCl, ampicillin (50 µg/ml) and incubated for 16 hr at 37^0C. Fresh inoculum (10%) from the cell culture was inoculated in 500 ml of 2X YT liquid media containing

50μg/ml ampicillin. Culture conditions were monitored by checking O.D. in Spectrophotometer till the O.D. reached at 0.5. Immediately inducer L-arabinose was added to a final concentration of 0.2%. Induction was continued for 6 hr before pelleting down the culture for further processing at the downstream level.

Determination of time of induction

One of the selected clone pRA EK was taken up for further studies. The clone pRA EK was inoculated into 2XYT medium containing 50-100 μg/ml ampicillin and incubated at 37^0C temperature. Inducer L-arabinose at different time points from 0 hour, 2 hours and 8 hours was added to this culture. The expressed protein was then checked isolated as described above and checked for activity. The results of expression levels are given in Table 1, (O.K.). Figure 3 shows the SDS-PAGE profile expression of His EKL in the arabinose induced sample.

Example 3

Fermentation of **pRA EK**

A loopful of recombinant *E. coli* comprising the plasmid pRA-EK containing EK_L fragment was plated on 2X YT agar plates and incubated overnight at 37^0C. A single colony was picked up and inoculated in 2X YT containing 50 -500 ug/ml ampicillin and kept on a rotary shaker for 12 hours at 37^0C. This was used to inoculate 100 ml of medium of the same composition. The inoculated medium was kept on a rotary shaker for 8-16 hours. The seed generated in this manner was used to inoculate a fermentor B. Braun of 1.5 -5 liter capacities and pH was controlled by automatic addition of ammonia at pH 7.0. The speed of the impeller was increased from 360-1200 rpm in 8 hours. Stirrer speed was increased gradually over initial eight hours and maintained at maximum thereafter. The dissolved oxygen set point was 30%. Airflow to the fermentor was between 5 to 10 litres per minute preferably 10 litres per minute. Fermentation was carried out for 24 hours and the samples were taken at hourly intervals and used for optical density determination and expression studies. At the end of the fermentation the culture was centrifuged and pellet was used for expression and isolation of enterokinase. The 2X YE medium comprises yeast extract 2.5 -5%, NaCl 0.5 -1%, $MgSO_4$ 2 -5 uM and ampicillin 50-500 ug/ml. The pH of the medium was adjusted to 7.4-7.6

Example 4

Expression optimization using antibiotics

The inoculum from seed medium obtained as in Example 3 was added to 2XYE medium containing antibiotic at concentrations ranging from 200-1000 mg/1. The fermentation was carried out at one liter scale in a fermentor B. Braun of 1.5 (2.0) litre capacity with temperature at 37^0C and pH was controlled by automatic addition of ammonia at pH 7.0. The speed of the impeller was increased from 360-1200 rpm in 8 hours. The dissolved oxygen set point was 30%. Airflow to the fermentor was between 5-10 L/m preferably 10L/m. The culture was initially maintained in a carbon deficient medium but the inducer L-arabinose was maintained at 1-5% concentration from the start of fermentation. The fermentation medium consists of 20 % yeast extract and 15 % glucose in 600 ml with 375 mg of ampicillin. The effect of ampicillin concentrations are shown in Table 2 (O.K.).

Example 5

Expression optimization using glucose

The inoculum from the seed medium obtained as in Example 3 was added to 2X YE medium containing antibiotic at a concentration of 300-400 mg in 600 ml in the feed. Fermentation was carried out using glucose between 0-25 % concentration both in the initial medium and in the feed medium. Figure 2 (O.K.) gives the effect of glucose on the expression of enterokinase in E. coli.

Example 6

Tangential flow filtration

E. coli culture expressing recombinant enterokinase is transferred to the sample reservoir of TFF system and the cells harvested using 0.45μ hollow fiber module. The concentrated broth was washed with 50 mM Tris, pH 8.0/ 150 mM NaCl buffer. The washed cells were drained from the reservoir and transferred to a pre-chilled stainless steel container kept on ice and allowed to chill to ~10°C.

Example 7

Pressure Disruption

Cell suspension as obtained in example 6 was disrupted using Niro-Soavi pressure disrupter. The disrupter was first washed with 0.22 μ water and then with lysis buffer (5OmM Tris, 15OmM NaCl). For cell disruption the instruments power setting was preset at 50 Hz. The chilled cells were disrupted at -900 bar total pressure, (-800-850 bar at 1st stage and -80-100 bar at 2nd stage). The lysate is collected in pre chilled stainless steel container kept on ice and the system washed with 50 ml of chilled lysis buffer. The cell lysate was allowed to cool to ~10°C before subjecting to next cycle of disruption. Three isothermal cycles were carried to ensure maximal cell disruption. The volume of lysate is approximately 500-550 ml at the end of this step.

Example 8

Solubilization

To the lysate, urea (8 M) was added. 390 g of urea was added to every 500 ml of cell lysate to give 8 M final concentration. This was slowly mixed at room temperature for - 2 hrs till urea solubilized into solution. For complete solubilization, the solution was kept at 20-25°C for 12-15 hrs. Imidazole was added to give 30 niM final concentration.

Example 9

Expanded bed chromatography

Affinity chromatography was carried out using Streamline chelating beads in expanded bed mode. A streamline 25 column was used with 30 ml bed volume (bed height of 6 cm) of matrix charged with $NiSO_4$. The column was equilibrated with equilibration buffer (50mm Tris pH 8.0/ 15OmM NaCl/ 30 mM Imidazole/ 8 M urea) at flow rate of 10 ml / min in expanded form and the bed expanded to 2-2.5 times the settled bed height. The lysate, under constant gentle stirring was loaded onto the column at a flow rate of 3 ml/min to maintain the same expanded bed height and the unbound material was washed away with equilibration buffer, sequentially using

1) 100 ml at 3 ml/ min expanded mode;

2) 100 ml at 10 ml/min in expanded mode and finally; and

3) 100 ml at 10 ml/min in settled mode.

The flow through and wash was collected in a separate container for further analysis. Elution was carried out in settled bed mode using elution buffer (50mm Tris pH 8.0/ 15OmM NaCl/ 250 niM Imidazole/ 8 M urea). About 3 CV (column volume) was needed to elute out all the bound proteins. However the first 30 ml was discarded and about 90 ml of eluate containing the protein of interest was collected. Figure 4 shows the SDS-PAGE of the affinity purification of His EKL by expanded bed chromatography on streamline chelating beds.

Example 10

Gel filtration chromatography

Gel filtration of the eluate obtained above was carried out using Sephadex G-25 beads packed in a XK 50/20 column, bed volume of 250 ml. The column was equilibrated with 50 mM Tris, pH 8.0/ 8 M urea at 20 ml/min flow rate. After 3 CV of equilibration, sample is loaded onto the column at the same flow rate. The protein was eluted with the same equilibration buffer and the protein peak collected till conductivity reaches ~2-3 mS/cm.

Example 11

Sulfitolysis

Protein concentration of the gel filtration eluate was determined using Bradford's assay method. Protein concentration was adjusted to 0.5 mg/ml using buffer (20 mM Tris, pH 8.0/ 8 M urea). To this solution 25 mM sodium sulfite and 1 mM sodium tetrathionate was added and the solution was stirred slowly at room temperature for 4 hrs to complete the sulfitolysis reaction.

Example 12

Refolding

Refolding was carried out in redox buffer system comprising 0-100 mM Tris, 0.3-3 mM EDTA, 0.2-2 mM Reduced Glutathione and 0.5-5 mM Oxidised Glutathione, 0-0.1% Triton X 100, 0-0.5% Tween-20, 0-0.5M arginine and 0-0.1M glycine. The Sulfitolysed sample 5-50 mg sample (100 ml of Sulfitolysed sample) obtained as described above was added to the 900 ml refolding buffer. During addition, constant stirring was maintained to ensure rapid dilution of the sample in refolding buffer. Flow rate for addition is adjusted so that the entire addition takes about 20 min (e.g., for 20 ml of sample addition, flow rate is 1 ml/min; for 40 ml sample flow rate is 2 ml/min). The refolding solution is transferred to 4°C and kept under constant stirring for 60 hrs to achieve maximal activity.

Example 13

STI chromatography

After 60 hrs of refolding, the solution was concentrated using a 5 kDa MWCO hollow fiber module (5 kDa Molecular weight cut off membrane module can also be used). The solution is concentrated to >100X. A slightly turbid solution was obtained after concentration and this is further clarified by passing through 0.22 μ TFF cassette. Sodium phosphate was added to the filtrate at pH 8.0 to give 50 mM final concentration and NaCl to 500 mM final concentration. This was loaded onto a soybean trypsin inhibitor column (STI)-agarose column (10 ml bed volume, 10 cm bed height), previously equilibrated with 5 CV of 50 mM sodium phosphate, pH 8.0 / 500 mM NaCl at 2 ml/min flow rate. The unbound material was washed with 5 CV of equilibration buffer till UV absorption at 280 nm reached baseline levels. Bound protein was eluted with 100 mM sodium formate, pH 3 / 200 mM NaCl. Peak fraction was collected (~20-25 ml) and immediately neutralized by addition of 2 M Tris. Equal volume of glycerol was added to achieve 50% (v/v) concentration and slowly mixed to obtain a homogenous solution. This was stored at -20°C. Enterokinase activity was assayed in all the samples using kinetic assay as well as colorimetric end point assay method using commercially available EK as control sample. A summarized purification data with respect to total protein (g), total activity (units), yield

(%) and specific activity (units/mg) is shown in Table 3. Figure 5 shows the SDS-PAGE profile of the refolded STI purified EK_L.

Example 14

Kinetic assay **of** enzyme activity

The assay is set up using 225 µl of substrate (NαCBZ L-lysine Thiobenzyl ester, 1 mM stock in water), 26 µl of Dithionitrobenzoic acid (DTNB, 5 mM stock in 100 mM sodium phosphate, pH 8.0), 50 µl of 1 M Tris-Cl, pH 8.0, 199 µl water and 10 µl of enzyme solution containing 0.1-1 units. This is mixed thoroughly and the rate of 3-carboxy-4-nitrophenoxide production is recorded at 412 nm for 5-10 min at ambient temperatures. The rate of change in absorbance (Δ Abs@ 412 nm) is directly proportional to enzyme activity.

Figure 6 shows SDS-PAGE profile of purified EKL to commercially available EKL. Figure 7 shows the kinetic assay comparison of purified EKL versus the commercially available EKL.

Example 15

Colorimetric assay **of** enzyme activity

This assay protocol describes a colorimetric end point assay for determination of enterokinase activity based on the cleavage of a specific ligand containing cleavage site for the enzyme. The assay was set up with the desired units of Enterokinase (0.1 to 1 units) in 200 µl of assay buffer (10 mM Tris-Cl, pH 8.0 / 10 mM $CaCl_2$). To this 2 µl of substrate (Gly-(-Asp-)$_4$-Lys-Naphthalamide. Na^+) in assay buffer, was added to give a final concentration of 5 mM. This was incubated at 37^0C for 12-16 hours. The reaction was stopped by addition of 100 µl of 2N hydrochloric acid. To this 100 µl of 0.2% sodium nitrite solution in water was added and mixed. After 3 min 200 µl of 0.5% ammonium sulfate solution in water was added and mixed. After 3 min 400 µl of 0.05% N-(-l-napthyl) ethylenediamine solution in 95% ethanol was added, reaction solution mixed and incubated for 45 min at 22-25 ^{0}C. Colour developed was read on a spectrophotometer at 595 nm and corrected against appropriate blank. A standard curve was generated using

increasing concentrations of enterokinase. The present invention relates to cloning expression and purification of a polymorphic bovine duodenal enterokinase light chain. This enzyme has usefulness in enzymatic processing of recombinant proteins containing the EK cleavage site.

Table 1: Effect of time induction on EK_L expression

Time of induction in hours	Enterokinase expression percentage
0	26
1	24.6
2	20.6
3	16.2
4	9.0
No inducer added	No expression

Table 2: Effect of antibiotic & concentrations on EK_L recovery

Concentration in µgs of antibiotic used per ml in feed medium	Enterokinase units recovered per 10 mgs refolded protein	Final optical density achieved in one liter at 600 nms
50	5000	40
125	6700	40
500	2000	47
650	12913	48
1000	10395	47

Table 3: Purification of EK from 500 ml culture broth

Purification step	Total Protein (g)	Total activity Units	Yield %	Specific activity (U/mg)
Cell lysis	16	-	-	-
Affinity purified	0.233	-	100	
Buffer exchange	0.227	-	97.4	-
Refolding & STI purification	0.013	718945	5.78	55303

What is claimed is:

1. An isolated nucleic acid comprising a polynucleotide sequence as shown in SEQ ID NO: 5, wherein said nucleic acid encodes for an enterokinase having an amino acid sequence as set forth in SEQ ID NO: 6.

2. A recombinant vector comprising the nucleic acid of claim 1 operably linked to an inducible promoter.

3. The recombinant vector as claimed in claim 2, wherein said inducible promoter is selected from a group consisting of AraB, T7, trp, lac, tac.

4. A host cell comprising the recombinant vector comprising the polynucleotide sequence as set forth in the SEQ ID NO: 5 coding for the amino acid sequence as set forth in SEQ ID NO: 6.

5. The host cell as claimed in claim 4, wherein the host cell is selected from a group consisting of *E .coli,* CHO, yeast and eukaryotic cell lines.

6. An improved process for production of enterokinase having the amino acid sequence as set forth in SEQ ID NO: 6 encoded by the polynucleotide sequence as set forth in SEQ ID NO: 5, wherein said method comprising

    a. culturing the host cells in the growth medium, wherein the host cell comprising the polynucleotide sequence as set forth in SEQ ID NO: 5 encoding the amino acid sequence as set forth in SEQ ID NO: 6;

    b. recovering the enterokinase;

    c. refolding of the enterokinase ; and

    d. purifying the enterokinase

7. The process as claimed in claim 6, wherein said medium comprising sources of carbon and nitrogen in a ratio of 1:1 to 2:1.

8. The process as claimed in claim 6, wherein said medium comprising inducer selected from a group consisting of Isopropyl-beta-D-thiogalacto pyranoside (IPTG), lactose, maltose, arabinose and arabino galactan.

9. The process as claimed in claim 8, wherein said inducer is arabinose.

10. The process as claimed in claim 6, wherein culturing the host cells was carried out at temperature is in the range of about 30-37^{0}C.

11. The process as claimed in claim 6, wherein the refolding comprising sulfitolysis followed by oxidative refolding in red-ox glutathione environment.

12. The process as claimed in claim 6, wherein said enterokinase is expressed in a protease positive background.

13. An improved process for production of enterokinase having the amino acid sequence as set forth in SEQ ID NO: 6 encoded by the polynucleotide sequence as set forth in SEQ ID NO: 5, wherein said method comprising:

a. adding polynucleotide sequence as set forth in SEQ ID NO: 5 to the cell free transcription and translation system to produce enterokinase having amino acid sequence as set forth in SEQ ID NO: 6;

b. isolating the enterokinase; and

c. purifying the enterokinase

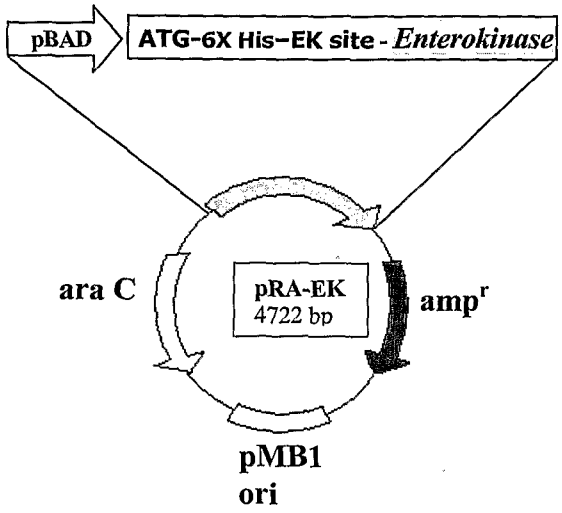

pBAD
ATG-6X His-EK site - Enterokinase
ara C
pRA-EK
4722 bp
amp^r
pMB1
ori

Figure 1

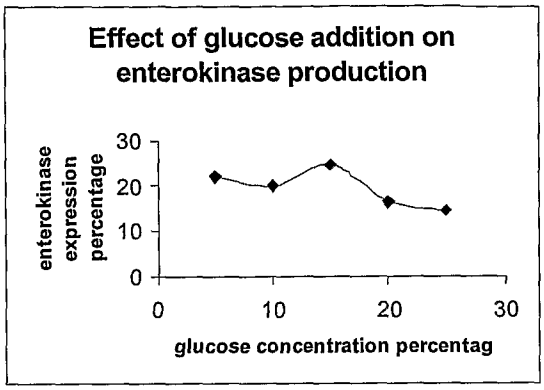
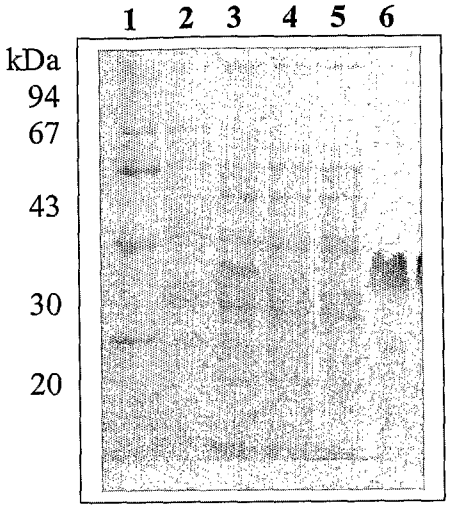
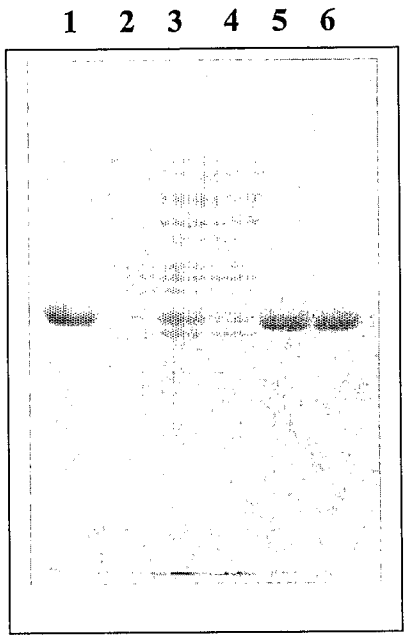

Effect of glucose addition on enterokinase production
enterokinase expression percentage
30
20
10
0
0
10
20
30
glucose concentration percentag

Figure 2



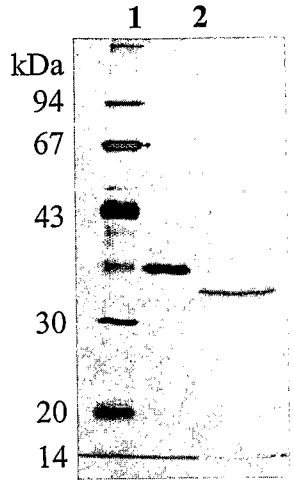
Lane 1: LMW markers
Lane 2: Uninduced Total protein
Lane 3: Induced Total Protein
Lane 4: Supernanant
Lane 5: Pellet Wash
Lane 6: Purified IB's

Figure 3



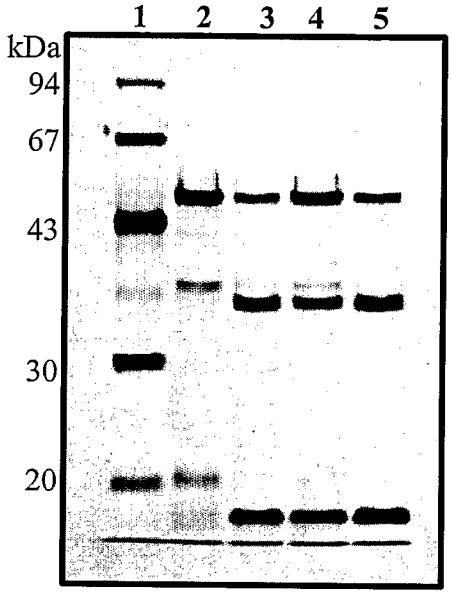
Lane 1: His-EK_L
Lane 2: Uninduced TP
Lane 3: Induced TP, SM for EBA
Lane 4: FT & wash of EBA
Lane 5: Elution of EBA
Lane 6: GF of Elution

Figure 4



Lane 1: LMW Markers
Lane 2: Ni-NTA purified His-EK
Lane 3: STI-Purified EK

Figure 5



Lane 1: LMW Markers
Lane 2: Cleavage control protein (Novagen)
Lane 3: CCP cleaved with EK (Novogen)
Lane 4: CCP cleaved with USV-EK (diluted)
Lane 5: CCP cleaved with USV-EK (concentrated)

Figure 6

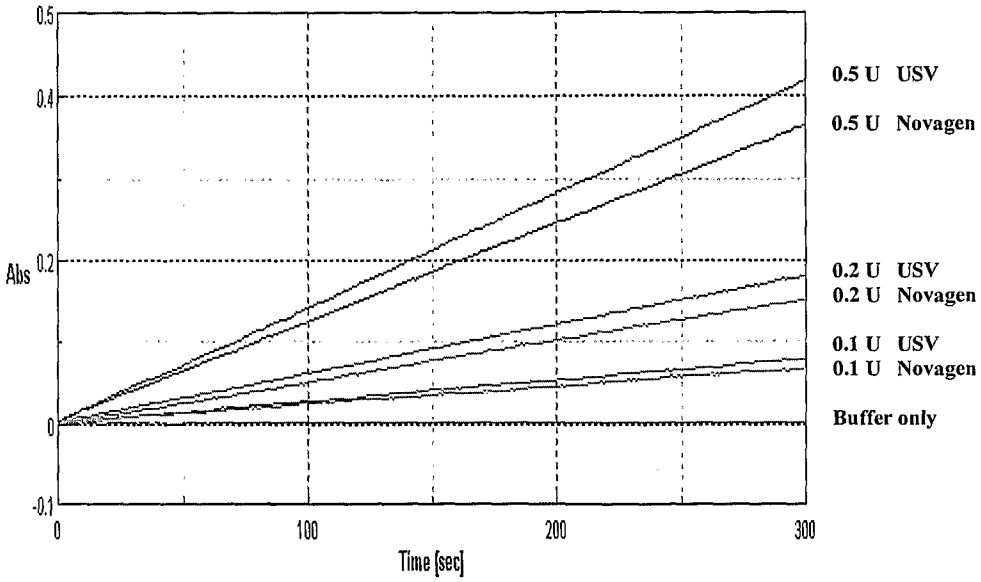

0.5
0.4
0.2
0
-0.1
Abs
0
100
200
300
Time [sec]
0.5 U USV
0.5 U Novagen
0.2 U USV
0.2 U Novagen
0.1 U USV
0.1 U Novagen
Buffer only

Figure 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/IN2007/000277

A CLASSIFICATION OF SUB,JECT MATTER
INV . C07K14/435

According Io International Palenl Classification (IPC) or Io bolh national classilicalion and IPC

B. FIELDS SEARCHED

Minimum documenlation searched (classilicalion system lollowed by classification symbols)
C07K

Documentation searched olher than minir ation Io 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 practical search terms used)
EPO-Internal, WPI Data, PAJ, BIOSIS, EMBASE, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category'	Citation ol document with indication where appropriate ot the relevant passages	Relevanl to claim No
X	CHOI ET AL.: "Recombinant enterokinase light chain with affinity tag: Expression from Saccharomyces cerevisae and its utilities in fusion protein technology" BIOTECHNOLOGY AND BIOENGINEERING, vol . 75, no. 6 , 20 December 2001 (2001-12-20), pages 719-724, XP002475276 the whole document ----- -/-	1-13

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

[X] See patent family annex

' Special calego πes of cited documents

'A' document defining the general state ol the art which is not considered Io be of particular relevance

'E' earlier documenl but published on or after the international tiling dale

•L' document which may Ihrow doubts on pπoπty claιm(s) or which is cited Io establish the publication dale ol another citation or other special reason (as specified)

'O" document referring Io an oral disclosure use exhibition or other means

"P' document published prior Io the international filing date bul later than the pπority date claimed

[1]T' later document published after the international tiling date or pnority date and not in conflict with the application bul cited Io understand the principle or theory underlying lhe invention

•X[1] documenl of particular relevance the claimed invention cannol be considered novel or cannot be considered Io involve an inventive slep when lhe documenl is taken alone

"Y' document of particular relevance the claimed invention cannot be considered Io involve an inventive step when (he document is combined with one or more other such docu¬ments such combinalion being obvious Io a person skilled in the an

'&' documenl member ol the same patent lamily

Dale ol the actual completion ol the international search	Dale ot mailing ol lhe international search report
7 Apri l 2008	28/04/2008
Name and mailing address of the ISA/ European Patent Office P B 58IB Patentlaan 2 NL - 2280 HV Ri|swιk Tel (+31-70) 340-2040 Tx 31 651 epo nl Fax (+31-70) 340-3016	Authorized officer Kel ler , Yves

Form PCT/ISA/2 10 (second sheβl) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/IN2007/000277

C(Contlnuatlon). DOCUMENTS CONSIDERED TO BE RELEVANT

Category'	Citation of document with indication where appropriate ol the relevanl passages	Relevant to claim No
X	Y. KITAMOTO ET AL.: "Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of distinctive assortment of domains" P.N.A.S., vol. 91, August 1994 (1994-08), pages 7588-7592, XP002475277 abstract figure 1 figure 2 page 7589, column 1, paragraph 4 - page 7590, column 1, paragraph 2	1-13
Y	CN 1 332 248 A (UNIV NANJING [CN]) 23 January 2002 (2002-01-23) abstract; claims 1-4	1-13
Y	US 5 665 566 A (LAVALLIE EDWARD R [US]) 9 September 1997 (1997-09-09) claims 1-20 examples 1-7	1-13
Y	US 6 342 362 Bl (MYTELKA DANIEL S [US]) 29 January 2002 (2002-01-29) abstract column 6, paragraph 1	1-13
Y	WO 2006/001027 A (USV LTD [IN]; MAITI DIPANWITA [IN]; MISHRA SHRIKANT [IN]; RAO SRINIVAS) 5 January 2006 (2006-01-05) page 7, line 3 - line 29	1-13

INTERNATIONAL SEARCH REPORT

Information on patent lamlly members

International application No
PCT/IN2007/000277

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
CN 1332248 A	23-01-2002	NONE	
US 5665566 A	09-09-1997	AT 242330 T	15-06-2003
		AU 675506 B2	06-02-1997
		AU 6089894 A	15-08-1994
		CA 2153254 Al	21-07-1994
		DE 69432787 Dl	10-07-2003
		EP 0679189 Al	02-11-1995
		JP 3149185 B2	26-03-2001
		JP 8502176 T	12-03-1996
		WO 9416083 Al	21-07-1994
US 6342362 Bl	29-01-2002	NONE	
WO 2006001027 A	05-01-2006	AU 2004320986 Al	05-01-2006
		CA 2558760 Al	05-01-2006
		CN 1953992 A	25-04-2007
		EP 1758933 Al	07-03-2007
		JP 2008503233 T	07-02-2008
		US 2007128622 Al	07-06-2007