

## Efficient Expression and Purification of Recombinant Human Enteropeptidase Light Chain in *Escherichia coli*

Li-Xi Niu, Jia-Yue Li, Xue-Xue Ji and Bin-Sheng Yang\*

Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education; Shanxi University; Taiyuan Shanxi - China

### ABSTRACT

Human enterokinase (synonym: enteropeptidase, EC 3.4.21.9) light chain (hEK<sub>L</sub>) gene was designed and artificially synthesized with built-in codon bias towards *Escherichia coli* codon preference. The synthetic hEK<sub>L</sub> gene was cloned into prokaryotic expression vector pMAL-s and transferred into the expression strain *E. coli* BL21 (DE3). Recombinant hEK<sub>L</sub> protein with a maltose binding protein (MBP) tag was expressed at high levels in soluble form, which yielded about 42% of the total cellular protein. The target protein was then purified to the homogeneity (> 95%) by affinity chromatography. The peptide substrate GST-Melittin with enterokinase recognition site was completely cleaved by the purified MBP-hEK<sub>L</sub> at the molar ratio of 1:5000 (enzyme:substrate). Tricine SDS-PAGE analysis showed that the activity of MBP-hEK<sub>L</sub> was approximately seven times that of bovine enterokinase catalytic subunit (EKMax<sup>TM</sup>, Invitrogen). From 1 L flask culture, 206 mg pure active MBP-hEK<sub>L</sub> was with specific activity of  $1.4 \times 10^4$  U/mg.

**Key words:** Human enterokinase light chain (hEK<sub>L</sub>), Fusion expression, Affinity purification, Activity analysis, Enzymatic cleavage

### INTRODUCTION

Fusion expression can often achieve efficient expression and purification of the target protein. But sometimes, the target protein can not be used in the form of fusion protein. Therefore, the specific protease recognition site is designed between the target protein and the label protein. Bioactive target protein can be obtained through special cleavage. At present, commonly used proteases include thrombin, enterokinase, and factor Xa etc, among which enterokinase is regarded as an ideal enzyme to prepare target protein by sequence specific cleavage from fusion protein (Shahrvan et al. 2008; Mikhailova et al. 1998). Enterokinase, which consists of two chains connected by a disulfide bond, is a heterodimeric serine protease from the mammalian duodenum.

The heavy chain (115 kDa) anchors enterokinase in the intestinal membrane and strongly influences macro-molecular substrate recognition and inhibitor specificity. The light chain (35 kDa) is the catalytic subunit, which consists of a chymotrypsin-like serine protease domain (Mikhailova et al. 1998; Huang et al. 2007). Enterokinase recognizes Asp-Asp-Asp-Asp-Lys sequence and cleaves the C-terminal peptide bond of the lysine residue. Due to broad reaction scope (temperature range: 4-45°C; pH range: 4.5-9.5) and mild reaction conditions, enterokinase is considered a powerful tool and widely applied in biochemistry and biotechnology (Gasparian et al. 2006). The source of natural enterokinase is limited due to high cost of extraction and separation, and the enzyme derived from the animals is always contaminated by other protease,

\* Author for correspondence: yangbs@sxu.edu.cn

which lead to degrade the protein (Vozza et al. 1996). Therefore, efforts have been made to develop enterokinase using genetic engineering methods.

There are reports on the use of prokaryotic or eukaryotic expression system to produce recombinant bovine (Kitamoto et al. 1994; Collins-Racie et al. 1995; Huang et al. 2007) and mice (Sun et al. 2006) enterokinase light chain protein with biological activity. However, there is lack of study on the recombinant human enterokinase light chain (Gasparian et al. 2006). The human enterokinase light *hEK<sub>L</sub>* gene has been expressed in yeast cells (Pepeliaev et al. 2011). However, expression was very low. In another attempt, a recombinant protein fused to thioredoxin (Trx) was produced at high levels in *E. coli*, but the protein was mainly insoluble and occurred in inclusion bodies (Gasparian et al. 2003). Thus, problems such as the insoluble/inclusion bodies or low expression level have plagued previously reported recombinant expression systems. The aim of this work was to study the synthesis of the *hEK<sub>L</sub>* gene using the *E. coli* preferred codon bias and the expression of the protein in soluble form.

## MATERIALS AND METHODS

### Materials

Host strain *E. coli* BL21 (DE3) and plasmid pMAL-s were stored in the laboratory. Fusion protein GST-Melittin, which contained the enterokinase recognition sequence. Plasmid pMAL-s was derived from the pMAL-p2X (NEB). Factor Xa recognition sequence was substituted by human rhinovirus 3C protease recognition sequence with gene splicing by overlap extension. Bovine enterokinase catalytic subunit (EKMax<sup>TM</sup>) was purchased from Invitrogen. The plasmid pRSS7, which contained human enterokinase gene was purchased from GenScript. Amylose resin and Gly-Asp-Asp-Asp-Asp-Lys-β-naphthylamide (GD<sub>4</sub>K-na) were purchased from New England Biolabs. Restriction enzymes, DNA polymerase and T4 DNA ligase were purchased from TaKaRa.

### Construction of expression vectors

To improve the expression level of hEK<sub>L</sub>, the *hEK<sub>L</sub>* coding sequence (GenBank Accession No. U09860.1) was optimized using OptimumGene<sup>TM</sup> algorithm that replaced the rare codons and could

optimize other aspects of mRNA structure. The codon-optimized synthetic DNA clone was provided by GenScript (NJ, USA). In total, the *hEK<sub>L</sub>* gene was amplified from pRSS7 plasmid by the polymerase chain reaction (PCR) using the following oligonucleotide primers. FW (hEK<sub>L</sub>): 5'-CGGGATCCATTG TTGGAGGAAGTAAT-3' and RV (hEK<sub>L</sub>): 5'-GCAAGCTTCTAATGT AGAAAACCTTTG-3' were added *Bam*H I and *Hind* III restriction sites, respectively. The conditions used for PCR reaction were 30 cycles of (98°C for 10 s, 60°C for 10 s, 72°C for 1 min) and extension of 72°C for 10 min. The amplified product was ligated into the vector pMAL-s at *Bam*HI and *Hind*III sites, resulting in recombinant plasmid designated as pMAL-s-*hEK<sub>L</sub>*. The recombinant plasmid pMAL-s-*hEK<sub>L</sub>* was sequenced to verify the authenticity of the cloned synthetic *hEK<sub>L</sub>* coding region and the reading frame.

### Optimization of the expression conditions

The recombinant plasmid pMAL-s-*hEK<sub>L</sub>* was transformed into competent *E. coli* BL21 (DE3) by a 40 s heat shock at 42°C. Transformed cells were then spread on Luria-Bertani (LB) agar plates, supplemented with 100 mg/mL ampicillin. Positive clones of *E. coli* BL21 (DE3) containing recombinant plasmid then grown overnight in LB medium containing 100 mg/mL ampicillin at 37°C at 180 rpm. When the absorbance reached to 0.6-0.8 at 600 nm, the culture was divided into different tubes and IPTG was added to each subculture with different final concentrations (0, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0 and 5.0 mM). All the subcultures were continued to incubate at 37°C for an additional 5 h for the optimization of IPTG concentration. For the optimization of expression temperature, the tubes were continued to incubate at 16, 20, 25, 30 and 37°C, respectively, with a final 0.2 mM IPTG concentration. For the optimization of induction duration, the cultures were incubated for an additional 2, 3, 4, 5 and 6 h with a final concentration of IPTG at 0.2 mM. Ten microliters of each re-suspended cell culture was used for SDS-PAGE analysis. Each treatment was repeated at least two times.

### Large-scale expression and purification of MBP-hEK<sub>L</sub>

Recombinant cell cultures were performed in a 3 L flask containing 1 L LB medium under the optimum cultivation and induction conditions.

Cells were harvested by centrifugation. After washing twice with 20 mM Tris-HCl buffer (containing 500 mM NaCl, pH 8.0), the precipitate was re-suspended in the same buffer and disrupted through sonication at 20 kHz for a period of 5 min (SONICS Vibra Cell VC 455, 230 V, Germany) on ice. Debris were removed by centrifugation at 10 000 r/min at 4°C for 30 min.

The supernatant was applied onto an amylose affinity column, rinsed with column buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 8.0). Fusion protein was then eluted with 10 mM maltose. The collected fractions were concentrated by PEG20000 and dialyzed in 20 mM Tris-HCl buffer (containing 50 mM NaCl, 2 mM CaCl<sub>2</sub>, 0.1% Tween-20, pH 7.4). The purified protein was stored in 20 mM Tris-HCl buffer (containing 200 mM NaCl, 2 mM CaCl<sub>2</sub>, 50% glycerol, pH 7.4) at -20°C for use. Protein concentration was determined by the Bradford method and protein purity was analyzed on 10% SDS-PAGE.

#### Enzyme assays

The specific activities of the recombinant hEK<sub>L</sub> were determined by a modified procedure (Grant and Hermon-Taylor, 1979). Assays were performed in 1.0 mL of 25 mM Tris-Cl (pH 8.0), 10% DMSO, 10 mM CaCl<sub>2</sub> and 0-1.0 mM GD<sub>4</sub>K-na. The reaction was started by adding MBP-hEK<sub>L</sub> (0.5 nM) and monitored continuously for 5 min by the increasing absorbance of generated β-naphthylamine at 420 nm using an exciting wavelength of 337 nm. Enterokinase from Invitrogen (EKMax™) with specific activity of 2 000 U/mg and concentration of 0.5 mg/mL was used as standard. For the specific activity calculation of the fusion protein MBP-hEK<sub>L</sub>, only the enteropeptidase part was used on the basis of molar mass in order to facilitate comparison.

#### Cleavage of Glutathione S-transferase (GST) fusion protein by MBP-hEK<sub>L</sub>

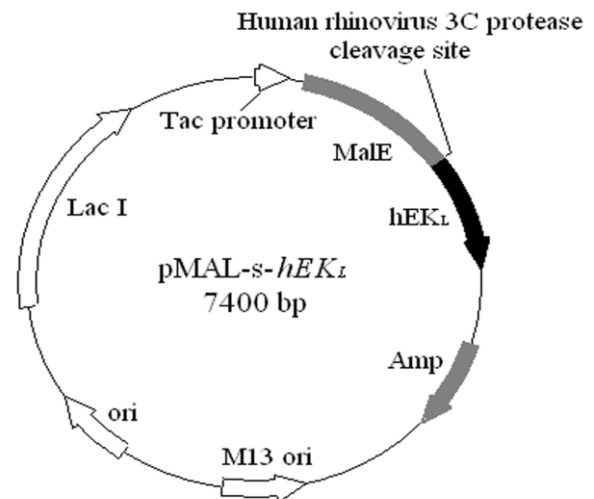
The protein cleavage activity of MBP-hEK<sub>L</sub> was determined by using the substrate GST-Melittin fusion protein containing DDDDK enterokinase recognition sequence between the GST-tag and Melittin protein. GST-Melittin was dissolved to a concentration of 1.0 mg/mL in 20 mM Tris-HCl (containing 50 mM NaCl, 2 mM CaCl, 0.1% Tween-20, pH 7.4). Substrate solution was mixed (90 μL) with different amounts of MBP-hEK<sub>L</sub> and placed in the water bath at 25°C for 20 h. The

samples were analyzed on 12% Tricine SDS-PAGE.

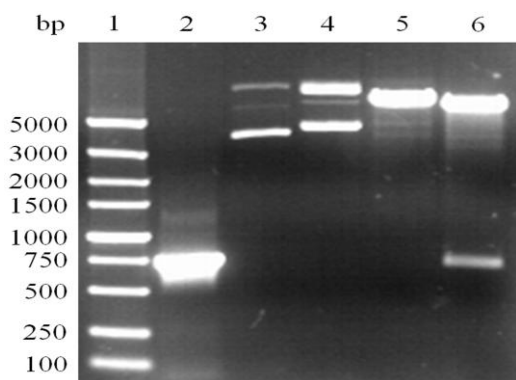
## RESULTS AND DISCUSSION

### Construction of the prokaryotic expression vector pMAL-s-hEK<sub>L</sub>

To create a modified hEK<sub>L</sub> gene with enhanced translation efficiency compared to the native gene, a codon-optimized version of hEK<sub>L</sub> gene was designed. This modified sequence retained the identical amino acid coding sequence of the native gene; however, the overall G+C content of synthetic hEK<sub>L</sub> DNA sequence was increased to 54%. The restriction enzymes sites for BamHI and HindIII were introduced at the 5'- and 3'-ends of the hEK<sub>L</sub> gene, respectively to provide convenient restriction sites for cloning into the prokaryotic expression vector. A 705 bp hEK<sub>L</sub> gene fragment was amplified by PCR from plasmid pSSR7 and inserted downstream of the MBP-tag sequence in the prokaryotic expression vector pMAL-s under the control of Tac promoter to construct the expression plasmid pMAL-s-hEK<sub>L</sub> (Fig. 1). The recombinant plasmid pMAL-s-hEK<sub>L</sub> was identified by PCR and double enzyme digestion (Fig. 2). The plasmid was then sent to Shanghai Sangon for sequencing, and the results matched the expected sequence completely.



**Figure 1** - A Schematic diagram of the expression vector pMAL-s-hEK<sub>L</sub>. hEK<sub>L</sub> was expressed as a fusion protein with the MBP-tag. Human rhinovirus 3C protease cleavage site was designed in the middle of MBP and hEK<sub>L</sub> to facilitate MBP-tag removal from the fusion protein. Amp was an abbreviation of Ampicillin.



**Figure 2** - Identification of recombinant plasmid by digestion. Lane 1 DNA maker DL 5000; Lane 2 PCR products of *hEK<sub>L</sub>*; Lane 3 pMAL-s; Lane 4 pMAL-s-*hEK<sub>L</sub>*; Lane 5 pMAL-s-*hEK<sub>L</sub>* digested by *Bam*HI; Lane 6 pMAL-s-*hEK<sub>L</sub>* digested by *Bam*HI and *Hind*III.

### Optimization of expression condition of *hEK<sub>L</sub>* gene in *E. coli* BL21 (DE3)

Thin layer scanning of SDS-PAGE profile showed the percentage of expressed product in the total bacterial protein reached maximum when induced at 37°C with 0.2 mM IPTG (data not shown). The SDS-PAGE profile of the expression level of MBP-*hEK<sub>L</sub>* fusion protein from *E. coli* BL21 (DE3) transformant in various induction durations is shown in Figure 3. When the liquid culture grew to OD<sub>600</sub> of 0.6-0.8, IPTG was added to a final concentration of 0.2 mM and the expression products was not increased after 5 h of additional incubation (Fig. 3). Based on the above optimization experiments, it could be concluded that the best expression conditions for the fusion protein were when the culture was incubated to an OD<sub>600</sub> of 0.6-0.8, IPTG was added to a final concentration of 0.2 mM and then was continued to incubate for an additional 5 h at 37°C.

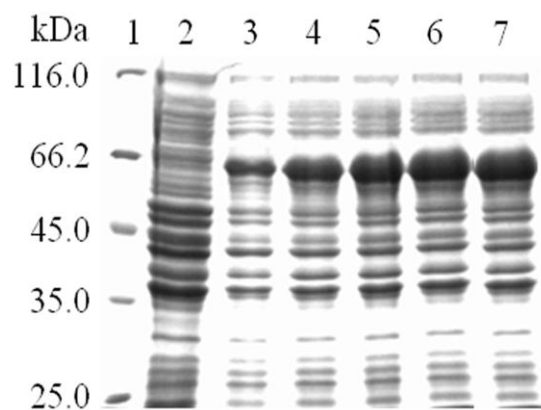
### Fusion protein expression and purification

As shown in Fig. 4 (lane 4), *E. coli* BL21 (DE3) harboring plasmid pMAL-s-*hEK<sub>L</sub>* exhibited no target protein on electrophoresis gel without induction. MBP-*hEK<sub>L</sub>* was expressed under IPTG induction. Comparing the un-induced cells, there was clearly an extra protein band (Fig. 4, lane 5), consistent with the expected molecular weight (MW) of 66 kDa, which was just the sum of MBP (40 kDa) and *hEK<sub>L</sub>* (26 kDa), after IPTG induction analyzed by SDS-PAGE. The fusion protein was expressed approximately 42% of the total soluble protein as assessed by densitometric scanning (Fig.

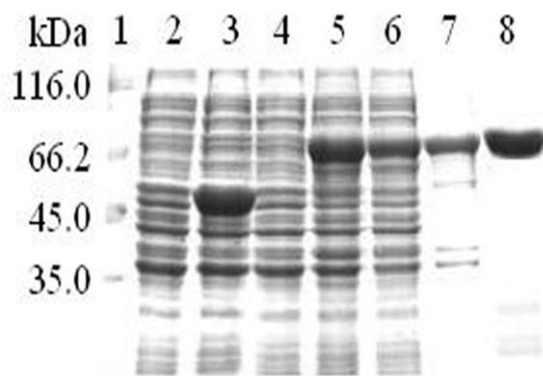
4, lane 6). The soluble MBP-*hEK<sub>L</sub>* was purified to homogeneity to >95% purity by amylose affinity chromatography, affording 206 mg MBP-*hEK<sub>L</sub>* from 1 L cell culture (Fig. 4, lane 8).

### Enzymatic assay of MBP-*hEK<sub>L</sub>*

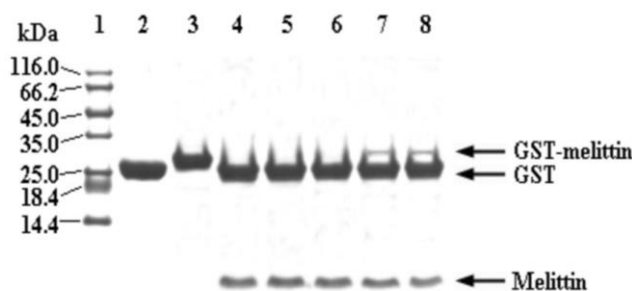
The activity of MBP-*hEK<sub>L</sub>* was compared to the activity of the standard EKMax<sup>TM</sup> (Invitrogen). Fluorometric assay showed that the specific activity of MBP-*hEK<sub>L</sub>* was approximately  $1.4 \times 10^4$  U/mg when Fusion protein GST-Melittin containing an inter-domain enterokinase recognition site was used to characterize the protein cleavage activity. Upon cleavage of GST-Melittin, the GST fragment and the Melittin fragment produced a theoretical MW of 26 kDa (GST) and 2.8 kDa (Melittin), respectively (Fig. 5). Various molar ratios of enzyme to substrate were tested. In case of 1:1000, 1:3000, 1:5000, the fusion protein was cleaved completely by MBP-*hEK<sub>L</sub>* and even in the case of 1:7000, more than 90% of substrate was cleaved. The commercially available enzyme was seven times less effective and the 90% cleavage was achieved at the ratio of 1:1000. The above results showed that through genetic engineering methods, bioactive recombinant human enterokinase light chain protein could be obtained.



**Figure 3** - SDS-PAGE analysis for the expression of MBP-*hEK<sub>L</sub>* in *Escherichia coli* with different induction time. Lane 1 Middle molecular weight protein marker; Lane 2 Whole cell lysate of non-induced *E. coli* BL21(DE3) cells containing the plasmid pMAL-s-*hEK<sub>L</sub>*; Lanes 3-7 Whole cell lysate of *E. coli* BL21(DE3) cells containing the plasmid pMAL-s-*hEK<sub>L</sub>* obtained 2,3, 4,5 and 6 h respectively after induction with 0.2 mM IPTG.



**Figure 4** - SDS-PAGE (12%) analysis of fusion protein MBP-hEK<sub>L</sub> expression and purification. *Lane 1* Middle molecular weight protein marker; *Lane 2* Whole cell lysate of *E. coli* BL21(DE3) cells containing the empty vector pMAL-s without IPTG induction; *Lane 3* Whole cell lysate of *E. coli* BL21(DE3) cells containing the empty vector pMAL-s with IPTG induction; *Lane 4* Whole cell lysate of non-induced *E. coli* BL21(DE3) cells containing the plasmid pMAL-s-hEK<sub>L</sub>; *Lane 5* Whole cell lysate of *Escherichia coli* BL21(DE3) cells containing the plasmid pMAL-s-hEK<sub>L</sub> obtained 5 h after induction with 0.2 mM IPTG; *Lane 6* Supernatant of cell extract of pMAL-s-hEK<sub>L</sub>/BL21(DE3) induced; *Lane 7* Pellet from ultrasonic disruption of the cell precipitate of pMAL-s-hEK<sub>L</sub>/BL21(DE3) induced; *Lane 8* Eluted sample from amylose column with 10 mM maltose.



**Figure 5** - Enzyme cleavage of GST-Melittin by Tricine SDS-PAGE. *Lane 1* Middle molecular weight protein maker; *Lane 2* GST; *Lane 3* GST-Melittin; *Lanes 4-7* Corresponding to the molar ratio of MBP-hEK<sub>L</sub> to GST-Melittin at 1:1000, 1:3000, 1:5000 and 1:7000; *Lane 8* Corresponding to the molar ratio of EKMMax<sup>TM</sup> to GST-Melittin at 1:1000.

In Summary, there have been a large number of reports about the bovine enterokinase (LaVallie et

al. 1993; Kitamoto et al. 1994; Collins-Racie et al. 1995), but only a few reports about human enterokinase. The ratio  $K_{cat}/K_m$  of human enterokinase light chain (hEK<sub>L</sub>) as 10-fold for the bovine enterokinase light chain (bEK<sub>L</sub>) (Gasparian et al. 2003) indicated that the catalytic efficiency of hEK<sub>L</sub> was high.

Among the previous reports, the highest activity of recombinant bEK<sub>L</sub> was reported by Peng et al. (2004), reaching 4.5 times that of EKMax<sup>TM</sup>. The expression and purification of hEK<sub>L</sub> was achieved in *Pichia pastoris* with 3.8 mg/L, and the enzyme activity was three times that of EKMax<sup>TM</sup> (Pepeliaev et al. 2011). hEK<sub>L</sub> was also expressed as inclusion bodies in *E. coli* and 10 mg purified enzyme with five times activity of EKMax<sup>TM</sup> was obtained from 1 L cell culture after denaturation, renaturation and purification (Gasparian et al. 2003). hEK<sub>L</sub> contains nine cysteine residues (which form four disulfide bonds and one remains unpaired), which make mis-folding of the expressed protein in cytoplasm easily. MBP is an effective fusion partner for increased solubility and expression yields for a variety of recombinant over-expressed proteins, especially eukaryotic protein (Kapust et al. 1999). In the present work, hEK<sub>L</sub> was expressed in soluble form by *E. coli* for the first time. After affinity chromatography purification, 206 mg fusion protein MBP-hEK<sub>L</sub> was obtained from 1 L culture medium. The enzymatic activity of the fusion protein was reached  $6.0 \times 10^5$  U/ $\mu$ M, which was seven times that of EKMax<sup>TM</sup>. It could be concluded that the specific activity reported in this work was the highest among the data currently available. The enzyme could be easily removed to prevent further degradation of the target protein after cleavage of fusion protein using amylose affinity chromatography. The fusion protein sequence also contained a human rhinovirus 3C protease cleavage recognition sequence between MBP-tag and hEK<sub>L</sub> for removal of the fusion handle. The majority of studies have shown that the recombinant bEK<sub>L</sub> protein from yeast expression system with glycosylation modification displayed higher enzymatic activity than that from prokaryotic expression system without glycosylation modification (Yuan et al. 2002; Peng et al. 2004). Present results showed that the glycosylation of human enterokinase light chain protein did have not positive connection with its enzymatic activity.

## ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation of Shanxi Province in China (No. 2011021035-5).

## REFERENCES

- Collins-Racie LA, McColgan JM, Grant KL, DiBlasio-Smith EA, McCoy JM, LaVallie ER. Production of recombinant bovine enterokinase catalytic subunit in *Escherichia coli* using the novel secretory fusion partner DsbA. *Biotechnol.* 1995; 13(9): 982-987.
- Gasparian ME, Ostapchenko VG, Dolgikh DA, Kirpichnikov MP. Biochemical characterization of human enteropeptidase light chain. *Biochem (Mosc).* 2006; 71(2): 113-119.
- Gasparian ME, Ostapchenko VG, Schulga AA, Dolgikh DA, Kirpichnikov MP. Expression, purification, and characterization of human enteropeptidase catalytic subunit in *Escherichia coli*. *Protein Expr Purif.* 2003; 31(1): 133-139.
- Grant DA, Hermon-Taylor J. Hydrolysis of artificial substrates by enterokinase and trypsin and the development of a sensitive specific assay for enterokinase in serum. *Biochim Biophys Acta.* 1979; 567(1): 207-215.
- Huang L, Ruan H, Gu W, Xu Z, Cen P, Fan L. Functional expression and purification of bovine enterokinase light chain in recombinant *Escherichia coli*. *Prep Biochem Biotechnol.* 2007; 37(3): 205-217.
- Kapust RB, Wangh DS. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 1999; 8(8): 1668-1674.
- Kitamoto Y, Yuan X, Wu Q, McCourt DW, Sadler JE. Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of a distinctive assortment of domains. *Proc Natl Acad Sci.* 1994; 91(16): 7588-7592.
- LaVallie ER, Rehemtulla A, DiBlasio EA, Ferenz C, Grant KL, Light A, et al. Cloning and functional expression of a cDNA encoding the catalytic subunit of bovine enterokinase. *J Biol Chem.* 1993; 268(31): 23311-23317.
- Mikhailova AG, Rumsh LD. Structural characteristics providing for high specificity of enteropeptidase. *Bioorg Khim.* 1998; 24(4): 282-287.
- Peng L, Zhong X, Ou J, Zheng S, Liao J, Wang L, et al. High-level secretory production of recombinant bovine enterokinase light chain by *Pichia pastoris*. *J Biotechnol.* 2004; 108(2): 185-192.
- Pepeliaev S, Krahulec J, Černý Z, Jílková J, Tlustá M, Dostálová J. High level expression of human enteropeptidase light chain in *Pichia pastoris*. *J Biotechnol.* 2011; 156(1): 67-75.
- Shahravan SH, Qu X, Chan IS, Shin JA. Enhancing the specificity of the enterokinase cleavage reaction to promote efficient cleavage of a fusion tag. *Protein Expr Purif.* 2008; 59(2): 314-319.
- Sun L, Wu XL, Wang YM, Yu YL, Wang LY. Expression of recombinant mice enterokinase light chain in *E. coli*. *J Jilin Univ Med Edit.* 2006; 32(3): 360-363.
- Voza LA, Wittwer L, Higgins DR, Purcell TJ, Bergseid M, Collins-Racie LA, et al. Production of a recombinant bovine enterokinase catalytic subunit in the methylotrophic yeast *Pichia pastoris*. *Biotechnol.* 1996; 14(1): 77-81.
- Yuan LD, Hua ZC. Expression, purification, and characterization of a biologically active bovine enterokinase catalytic subunit in *Escherichia coli*. *Protein Expr Purif.* 2002; 25(2): 300-304.

Received: May 04, 2014;  
Accepted: August 03, 2014.