J Appl Biotechnol Rep. 2021 December;8(4):395-405

Journal of
Applied Biotechnology Reports

doi 10.30491/JABR.2021.250014.1293

Original Article

# Cloning, Expression, and Purification Strategies for Enhanced Production of Enterokinase using TrpE fusion tag in Bench Scale Bioreactor

**Santhosh Nagaraj Nanjundaiah [1,2], Jayasri MA [2*], Sunilkumar Sukumaran [1], Ganesh Sambasivam [1]**

[1] Discovery Biology, Anthem Biosciences Pvt Ltd, Bommasandra, Bangalore, India
[2] School of Biosciences and Technology, Vellore Institute of Technology, Vellore, India

**Corresponding Author:** Jayasri MA, PhD, Professor, School of Biosciences and Technology, Vellore Institute of Technology, Vellore, India. Tel: +91-9739090010, E-mail: santhoshnagaraj1@gmail.com

**Received** September 28, 2020; **Accepted** January 10, 2021; **Online Published** December 30, 2021

**Abstract**

**Introduction:** Enterokinase (EK) is an enzyme of the serine protease family which is widely used in protein purification. EK acts at the C terminus of the DDDDK site in a protein chain. The enzyme has gained commercial importance in recent times owing to its specificity in biosimilar processing and also while removing the fusion tags in the course of protein purification.

**Materials and Methods:** The commercial production of EK has faced several challenges and demands the usage of novel strategies. This research shows the construction of a vector using TrpLE1413 (TrpE) as a fusion tag that pushes the produced EK inside the cell towards the inclusion body fraction and produced more of the desired protein in the BL21 (DE3) strain of *Escherichia coli*. The inclusion bodies produced by fed-batch fermentation were solubilized, refolded, activated, and purified by a single step of anion exchange chromatography.

**Results:** We purified 241 mg/L of recombinant EK, and its purity confirmed by RP-HPLC was greater than 97%. However, the maximum EK yield reported by other researchers is only 106 mg/L.

**Conclusions:** Overall, our results demonstrate the potential of the TrpE fusion tag along with novel expression and purification strategies to increase the enzyme yield by 2-2.5 times when compared to the yield achieved using traditional methods. Hence, this study has paved the way for the industrial production of EK in an economically viable manner. The same strategy could possibly be implemented on the expression of other industrially important recombinant enzymes depending on the protein characteristics.

**Keywords:** Recombinant enzyme, *E. coli*, enterokinase, overproduction, fermentation

**Citation:** Nanjundaiah SN, Jayasri MA, Sukumaran S, Sambasivam G. Cloning Expression and Purification Strategies for Enhanced Production of Enterokinase using TrpE fusion tag in Bench Scale Bioreactor. J Appl Biotechnol Rep. 2021;8(4):395-405. doi:10.30491/JABR.2021.250014.1293

## Introduction

Enterokinase (EK) (EC 3.4.21.9) is a mammalian intestinal serine protease that catalyzes the conversion of trypsinogen into trypsin by cleaving at the C terminus of the DDDDK site. Owing to its wide reaction temperature of 4.0-45.0 °C and pH of 4.5-9.5, the enzyme is considered as a potent tool in protein engineering.[1] It is a heterodimer and consists of a heavy chain (115 kDa), which helps in enzyme anchorage as well as substrate recognition, and a light chain (35 kDa), which is paramount in mediating the catalytic activity.[1] Until recent times, EK has been purified from different sources such as porcine,[2] bovine,[3,4] and human intestine;[5] however, with the recognition of its role in fusion protein engineering, recombinant EK is being extensively produced. Evidently, *E. coli* remains to be an ideal recombinant protein expression system because of its rapid growth rate and amenability to fermentation processes.[6] Moreover, its genome is easy to manipulate as per the needs, facilitating high expression levels of recombinant proteins.[7]

Proteins have so far been overexpressed in *E. coli* by using conventional fusion tags such as thioredoxin,[8] glutathione S-transferase [GST],[9] and histidine.[10] The most commonly used one is the 6-histidine residual tag, which serves the dual purpose of enhancing protein expression and functioning as a useful tool in purification.[11] Moreover, proteins have also been expressed in the soluble form by using signal peptides such as OmpA and MalE for secretion into the periplasm.[12] Nonetheless, both the approaches have been reported to result in a substantial reduction in the protein quantity and enzyme activity due to lower expression levels and the protein to tag ratio, which compromises the commercial viability of the process. There have been reports that have demonstrated the difficulty in expressing EK in prokaryotes like *E. coli* and several eukaryotes due to the autocatalytic activation of the enzyme leading to degradation in case of eukaryotes. This is mainly observed is case of autocatalytic enzymes. EK is one such autocatalytic enzyme that can cleave and make other inactive EK molecules active. Wolfgang Skala et al. have tried to express EK with a

histidine fusion tag where they claim of getting 1.7 mg folded protein per litre of culture and the expression pattern images seem to be promising, therefore histidine EK construct was designed in order to check out the expression levels in our hands.[13] Gasparian et al. expressed EK with thioredoxin fusion tag to purify 1 mg from 100 ml of culture.[14] Haidong Tan et al. when expressed EK with GST fusion tag yielded 106 mg/L of active folded protein.[15] Thus, based on the above statistics, it could be inferred that 106 mg/L is the maximum EK yield obtained in *E. coli* with fusion tag or with signal sequences. It also makes a way for the researchers to make a commercially viable process for production at a cheaper cost. A range of eukaryotic hosts have also been exploited. For instance, when expressed in *Aspergillus niger*, *Pichia pastoris* and *Saccharomyces cerevisiae* produced 1.9, 6.3 and 1 mg/L of culture respectively.[16-18]

EK plays a leading role in biosimilar processing and also in removing the fusion tags from proteins. For the expression of tough to express proteins or enzymes, fusion tag such as TrpLE needs to be exploited. A deleted outer region of TrpL and initial peptide region of TrpE operon constitute the TrpLE fusion tag.[19] In the past few years, over 30 or more proteins and peptides have been expressed successfully using TrpE fusion tag.[20] TrpE fusion has been fused and successfully expressed with proteins such as somatostatin,[21] Insulin A and B chain[22] and Mature TGF alpha.[20] TrpE mainly prevents the proteolytic action of the host proteases on the heterologous protein by pushing it towards inclusion body fraction. As the TrpE fusion tag does not consist of cysteine groups, it is well understood that the tag would not interfere in refolding any protein fused to it. The cleavage of the fusion tag is easier as the fusion tag is short and is exposed for cleavage from outside. As the expression of an enzyme like EK desires to be expressed in an inclusion body form, TrpE fusion tag was selected for the current study. The innovativeness of the current study was to produce EK with a novel fusion tag and develop a cost-effective and non-time consuming fermentation process with a new purification strategy to increase the yield of EK. This in turn reduced the cost of biosimilar production.

## Materials and Methods

### *Strains, Plasmids, Reagents and Media*

*E. coli* DH5α cells were used to maintain the plasmid and *E. coli* BL21 (DE3) was used as the expression host. The histidine EK gene was synthesized and supplied by GeneArt. The TrpE and T7 oligonucleotides were procured from Sigma Aldrich, and pET20b (+) vector was sourced from Novagen. Plasmid isolation and gel elution kits were purchased from Qiagen. The PCR clean-up kit was purchased from Sigma. Standard EK was acquired from Invitrogen and roche. InsTAclone PCR cloning kit and Pierce bicinchoninic acid assay kit were obtained from Thermo Fisher. Anti-6X His tag antibody was procured from Abcam. Restriction enzymes and polymerases required were sourced from New England Bio Labs. Z-lys-SBZL substrate, ampicillin, and isopropyl β-D-1-thiogalactopyranoside (IPTG) were from Sigma. The ion exchange resins were from Bio-Rad, and the bioreactors were from Applikon and Sartorius.

### *Design of Oligonucleotides, PCR Amplification, and Construction of Expression Vector*

The codon-optimized DNA sequence encoding a histidine fusion tag followed by a DDDDK cleavage site with an EK light chain at its C terminus overall composed of 746 base pairs was synthesized and supplied by GeneArt (ABB 44/pMA-T). The synthetic gene was designed with *NdeI* and *BamHI* sites, which served as flanking regions in the pMA-T vector. After assessing the lower expression levels of histidine EK construct in the pET20b (+) vector when expressed in *E. coli* BL21 (DE3), the same gene was used as a template for generating the TrpE histidine EK (TrpE H EK) construct through overlapping PCR amplification. The vector map of ABB 44/pMAT obtained from GeneArt is shown in Figure 1. TrpFP1, TrpFP2, and TrpFP3 as well as EK RP oligonucleotides were synthesized to add TrpE fusion tag base pairs to the histidine EK gene at its 5′ end. The amino acid sequence of the fusion tag is shown in Table 1.

**Table 1:** Amino Acid Sequence of the TrpE Fusion Tag

| Construct | Sequence |
|---|---|
| TrpE 1413 tag amino acid sequence | KAIFVLKGSLDRDPEF |

Note: *Start Codon is not Included in the Amino Acid Sequence

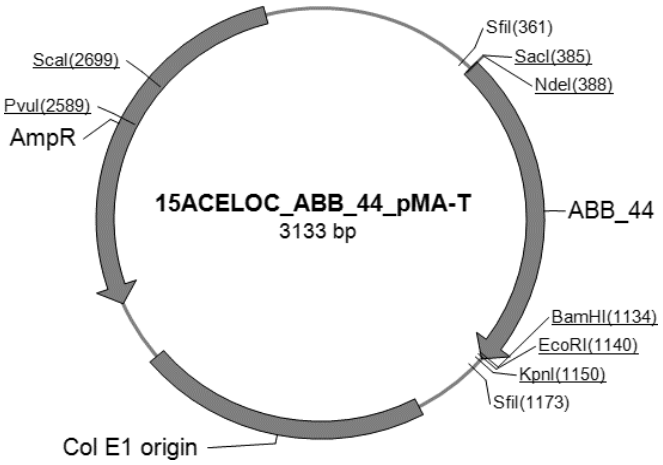


**Figure 1.** Vector Map of the Synthesized Histidine EK (+) Gene.

The PCR was performed in three steps, and every oligonucleotide had around 14-18 overlapping nucleotides at both ends except TrpFP1-bearing the *NdeI* restriction site at its 5′ end. The EK RP oligonucleotide was designed to encode the EK enzyme at the 3′ end, followed by the *BamHI* site 5′-**GGATCC**TTAATGCAGAAAGCTCTGAATCCATTCG-3′

(*NdeI* and *BamHI* sites are in bold letters). The first set of oligonucleotide (TrpFP3-5′-GGATCGTGATCCGGAATTT CATCATCATCATCATCACGATGA-3′ and EK RP) was designed to encode six nucleotides of the TrpE fusion tag, whereas the rest overlapped with the histidine gene of histidine EK. The second set (TrpFP2-5′-TTTTGTTCTGAA AGGTAGCCTGGATCGTGATCCGGA-3′ and EK RP) was used to add 21 nucleotides to the product obtained from the previous PCR reaction. However, the third set (TrpFP1-5′-**CATATG**AAAGCAATTTTTGTTCTGAAAGG-3′ and EK RP) added 21 nucleotides to the second product of PCR in addition to the *NdeI* restriction site at the 5′ end of the desired gene. The diagrammatic representation of the amplification process is shown in Figure 2.

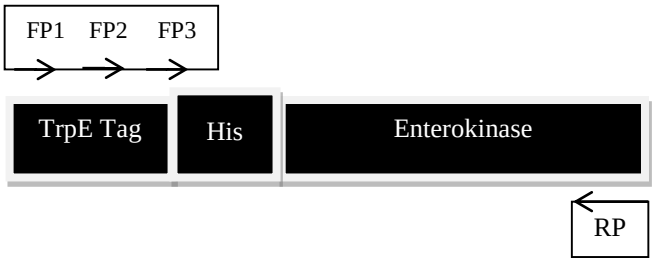


**Figure 2.** Diagrammatic Representation of TrpLE 1413 Fusion Tag Using 3 Primers to Histidine-tagged EK Synthesized Gene.

In all the PCR procedures, 1.2 pmol of primers and 100-300 ng of template DNA were employed for 100 µl of the reaction mixture. The used conditions were 95 ºC for 5 minutes; 30 cycles at 95 ºC for 60 seconds, 54-58 ºC for 45-60 seconds (primer annealing), and 72 ºC for 60 seconds; and an extension cycle for 10 minutes at 72 ºC. The primer annealing temperature was in the range of 54-58 ºC for all three reactions. After agarose gel electrophoresis, the obtained amplicons were purified by the Qiagen gel extraction kit. The purified DNA obtained from the third amplification reaction of TrpE H EK was ligated into the pTZ57R/T vector and then sub cloned into the pET20b (+) vector. Five micrograms each of TrpE H EK/pTZ57R/T vector and pET20b (+) vector were digested with *NdeI* and *BamHI* sites and incubated for 3 hours at 37 °C temperature. The samples were electrophoresed on an agarose gel, and the desired digested bands (801 bp and 3627 bp) were excised and taken ahead for gel extraction. The insert and vector concentration was estimated and the ligation reaction was set as per NEBiocalculator software. The ligation mixture was transformed into DH5α cells using the polyethylene glycol (PEG) method with the supplementation of 100 µg/ml ampicillin. Eight colonies were selected after ligation and colony PCR was performed with the help of T7 primers. The positive clones among them were screened by choosing particular restriction sites within the vector and insert. After 3 hours of digestion, samples were run on 1.2% agarose gel. As clone 2 and 3 plasmid showed prerequisite pattern, both were sent for sequencing using T7 forward and reverse primers.

### *Small-scale Expression and Induction Studies*

Initially, six clones of histidine EK were selected from the transformed plate of histidine EK/pET20b (+) in BL21 (DE3). These clones were inoculated into 10 ml super broth containing 100 µg/ml ampicillin and cultivated at 37 ºC. Based on the lower expression levels observed with histidine EK, TrpE H EK/pET20b (+) plasmid was generated and transformed into *E. coli* BL21 (DE3). The above shake flask studies were carried out by inducing with 1 mM IPTG concentration. In order to optimize IPTG concentration, we induced the cells with IPTG ranging from 0.25 to 2 mM concentration. Once 0.5 mM IPTG showed maximum level of expression, induction temperature studies (37-42 ºC) were carried out keeping IPTG concentration constant. The grown culture was induced with 0.5 mM IPTG when the $OD_{600}$ was in the range of 0.6-0.8 and incubated at different temperatures of 37-42 ºC for 16 hours. The temperature of 39 ºC showed better expression levels than our previous experiments. The cells were grown and induced with the optimized temperature and IPTG concentration for induction. After overnight induction, 0.5 $OD_{600}$ cells were centrifuged and the sample were prepared using PBS (Phosphate buffer saline) and SSB (Sample solubilization buffer) and boiled at 100 °C for 10 minutes. Uninduced and induced cells (20 µl) were loaded in separate wells of 12.5% SDS PAGE.[23] Based on the highest expression levels determined by Biorad EZ imager quantitation tools, clone five was carried forward for bioreactor batch on a bench scale.

### *Western Blotting*

The protocol for western blotting was as described by Mahmood T et al.[24] After running the TrpE histidine EK induced samples on SDS PAGE, the presence of protein tagged to histidine was exploited for the detection of EK. Antibodies raised against histidine residues were used as primary antibody and an alkaline phosphatase-conjugated antibody was treated as a secondary antibody. The nitrocellulose membrane was used for blot transfer. The blotting process was carried at 70 Volts for 90 minutes followed by blocking with 5% BSA for 1 hour. The blot was incubated in 1X TBST (Tris Buffer Saline Tween 20) buffer with the primary antibody added at a ratio of 1:1000 with shaking for 1 hour at room temperature. Alkaline phosphatase (ALP) conjugated secondary antibody (1:5000) incubation was for 1 hour with shaking followed by the development of blot with nitro-blue tetrazolium chloride-5-bromo-4-chloro-3'-indolyl phosphate p-toluidine salt (NBT-BCIP) substrate for color development. Once the blot development was completed, the membrane was thoroughly washed with water to terminate the blot development.

### *Fed-batch Fermentation*

Various media compositions were tried out at shake flask levels and the best media composition was carefully chosen for bioreactor studies. By varying the concentrations in the range of ± 20% the final fermenter media composition was developed. The same media was used during the fermentation batches. The fermentation process was divided into four stages: Seed, batch phase, fed-batch phase and induction phase. For this purpose, 1 ml of glycerol stock was inoculated into seed medium and incubated for 4-6 hours at 37 ºC. Once its $OD_{600}$ reached 3.0, 37.5 ml was transferred into a 2-L bioreactor containing 750 ml of production medium consisting of the following: 9 g/L of magnesium sulfate, 35 g/L of yeast extract, 20 g/L of dextrose monohydrate, 4.8 g/L of ammonium sulfate, 10 g/L of di-potassium hydrogen orthophosphate, and 20 ml/L of trace salts. The batch phase was terminated once the pH started spiking, and the fed-batch phase was commenced with the addition of feed containing 400 g/L of carbon source and 90 g/L of nitrogen source at regular intervals.[25] During optimization trials, Wet Cell Weight (WCW) and induction time studies were carried out with 200, 300 and 350 g/L WCW and 4, 6, 8 and 10-hour induction time. After optimization, 300 g/L WCW proved to be the best among the others. Whereas, during induction with 300 g/L WCW, 6-hour induction showed better levels of protein expression and further increase in time led to the degradation of protein. So, once a WCW of 300 g/L was attained, the culture was induced with 0.5 mM of IPTG based on the broth volume during induction. The feeding rate after induction was adjusted in such a way that it prevented acetate production as well as the accumulation of glucose. The temperature during the growth phase was 37 °C, and it was adjusted to 39 °C during the induction phase to stimulate enhanced protein production. The samples collected every 2 hours after induction till 6-hour induction were electrophoresed on 12.5% SDS PAGE.

### *Cell Lysis and Recovery of Inclusion Bodies*

One liter of broth was centrifuged at 10000 g for 20 minutes at 4 °C. The pelleted biomass was subjected to lysis with a buffer composed of 100 mM Tris and 10 mM ethylenediamine tetra-acetic acid (EDTA) (pH 8.0). For every 1 g of the cell, 9 g of lysis buffer was added. A high-pressure homogenizer with 1000 bar pressure was used, and the cells were passed thrice for enhanced lysis efficiency.[26] All the centrifugation steps were carried out at 10000 g for 20 minutes at 4 °C. The supernatant was discarded, and the inclusion bodies were separated from the cell debris by washing with 100 mM Tris, 10 mM EDTA (pH 8.0), and 0.5% Triton X-100. The ratio of washing buffer addition was 1:8 based on the pellet weight, and the last washing step was repeated with buffer lacking Triton X-100 to remove traces of the detergent.

### *Solubilization and Refolding of EK*

The solubilization buffer was prepared by using 6 M guanidine HCl with 100 mM Tris (pH 9.0). The inclusion body was diluted with solubilization buffer in a ratio of 1:60 (w/v) and was kept under gentle stirring for 4 hours. The solubilized protein mixture was exposed to 5 mM DTT for reduction and stored under anaerobic conditions with gentle stirring for 45 minutes. The reduced inclusion body was added in a pulsated manner to the pre-cooled (10 °C) refolding buffer such that it was diluted up to 40 times the initial volume. The novel refolding step was carried out at 10 °C using 0.5 M arginine, 50 mM Tris, 20 mM calcium chloride, 1 mM EDTA containing 1 mM cysteine set to pH 8.5, and 3 mM cystine. Disulfide bond formation was induced by incubating the protein mixture with gentle stirring at 4 °C for 72 hours. It was ensured that the refolding pH was maintained in the range of 8.5 ± 0.2 for 72 hours. The refolded protein solution was concentrated with a refolding buffer by using 5 kDa ultrafiltration membranes and centrifuged at 8000 rpm for 30 minutes at 4 °C. The pellet was discarded, and the supernatant was mixed with a pre-chilled activation buffer.

### *Activation of EK*

The novel activation mixture with the supernatant was prepared by using 50 mM Tris, 50 mM sodium chloride, and 4 mM calcium chloride of pH 7.4 maintained under continuous stirring at 10 °C for 12 ± 2 hours. The activation of EK was initiated by adding a unit of standard EK to the protein mixture in the activation buffer and incubated at 37 °C for 1 hour. The activation pH was also maintained in a range of 7.4 ± 0.2.

### *Purification and Storage*

The removal of the fusion tag from the mixture using anion exchange chromatography was performed at 4 °C. Sepharose Q high-performance column was equilibrated with four volumes of equilibration buffer composed of 50 mM Tris, pH 8.0, which would be used as a buffer A during protein elution. The activated protein mixture was loaded onto the column, followed by washing with four column volumes of 50 mM Tris and 100 mM sodium chloride of pH 8.0 to eliminate the unbound proteins. The protein was eluted with a seven-column volume with a gradient of 0-100% buffer B (200 mM Tris–HCl (pH 8.0) and 1 M sodium chloride) and Buffer A. Multiple fractions of 3-5 ml were collected after an increase in $A_{280}$ was observed. The fractions were diluted 1:1 with glycerol to obtain a final concentration of 100 mM Tris–HCl (pH 8.0), sodium chloride of 500 mM with 50% glycerol which was utilized for the storage of EK. The eluted fractions were pooled and absorbance was measured at 280 nm.

### *Estimation of Protein Concentration*

The protein concentration was estimated using the BCA (Bicinchoninic acid assay) kit along with the BSA (Bovine

serum albumin) standards provided in the kit. The samples were diluted to obtain protein concentrations falling within the range of the standards.

### *EK Activity Assay*

The activity of the final product of EK was estimated by using the Z-L-Lys-SBZL hydrochloride substrate. The assay buffer consisted of 50 mM Tris, 150 mM sodium chloride, and 10 mM calcium chloride set to pH 7.5. Fifty micromoles of SBZL were prepared in 1 mM dimethyl sulfoxide, and 200 µM DTNB was prepared in 1 mM DMSO. Sixty microliters of SBZL, 200 µl of DTNB, and 1 ml of assay buffer constituted the reaction mixture. The reaction was initiated by the addition of 1 µl of in-house EK. The same protocol was followed for the standard too. The absorbance values were noted at 405 nm, and the kinetic measurements were made at an interval of 15 seconds.

### *HPLC for EK*

Enzyme sample was injected into the reverse phase high pressure liquid chromatography (RP-HPLC) instrument, using ZORBAX column for analysis. The used mobile phase was 0.2 M Sodium Phosphate, pH 7.0, sample flow rate was 1 ml/minute, and ambient temperature was maintained during analysis. The Detection was at 214 nm wavelength.

### *Substrate Reaction Assay Using In-house EK*

TrpE fusion tag-infused protein containing the DDDDK cleavage site was used to test the cleavage efficiency of the in-house enzyme. Initially, 15 micrograms of the tagged protein were used per reaction. One unit of in-house EK was added to cleave the fusion tag in 20 µl of the reaction mixture (50 mM Tris, 1 mM calcium chloride, and 50 mM sodium chloride set to pH 8.0), which was incubated at 37 ºC for 10 hours. In order to mimic the same reaction set with in-house EK, another reaction was set with a unit of standard EK into 15 µg of substrate with a reaction volume of 20 µl. Substrate blank without any enzyme was used as a substrate control. During optimization, a unit of EK was kept constant for all the reactions by varying the quantity of protein taken for assay and time taken for complete cleavage of the substrate. The sample before and after cleavage was loaded on Tricine SDS PAGE for estimating the cleavage efficiency of in-house EK and standard EK. The cleavage percent analysis was estimated from SDS PAGE using the Biorad gel doc EZ imager quantitation tools.

## Results

### *Cloning and Construction of the Expression Vector*

Histidine EK gene was cloned into pET20b (+) vector using *NdeI* and *BamHI* restriction sites. Based on the lower expression levels obtained, the TrpE tag was fused to histidine EK with the help of primers. Later TrpE H EK gene was cloned into pET20b (+) vector and transformed into *E. coli* DH5 alpha cells. Two positive clones derived from colony PCR were subjected to restriction digestion with three different combinations of enzymes (*EcoRV* + *BamHI*, *BlpI* and *NdeI* + *EcoRI*) and screening displayed a prerequisite pattern (Figure 3). Clone three after sequencing with T7 forward and reverse primers matched with the exact sequence of the TrpE H EK gene. Clone three plasmid was selected for the BL21 (DE3) transformation

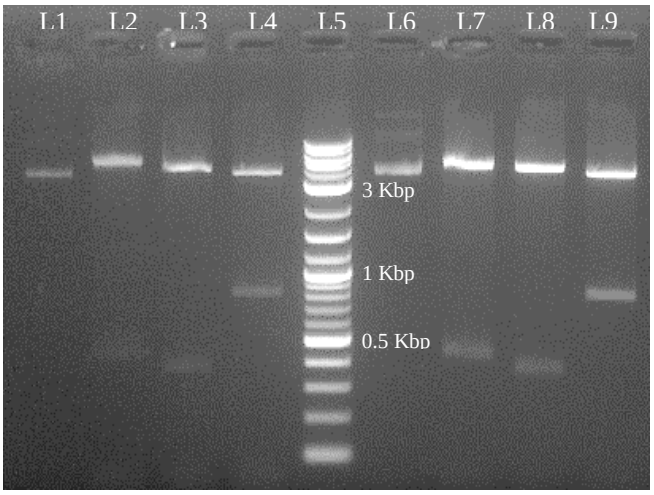


**Figure 3.** Restriction Digestion Pattern for TrpE H EK. Lane 1: TrpE H EK #2 uncut (4428 bp); Lane 2: TrpE H EK #2 *EcoRV* + *BamHI* digested (435 + 3993 bp); Lane 3: TrpE H EK #2 *BlpI* digested (374 + 4054 bp); Lane 4: TrpE H EK #2 *NdeI* + *EcoRI* digested (807 + 3621 bp); Lane 5: Marker (Quick-Load Purple 2-Log DNA Ladder); Lane 6: TrpE H EK #3 uncut (4428 bp); Lane 7: TrpE H EK #3 *EcoRV* + *BamHI* digested (435 + 3993 bp); Lane 8: TrpE H EK #3 *BlpI* digested (374 + 4054 bp); Lane 9: TrpE H EK #3 *NdeI* + *EcoRI* digested (807 + 3621 bp).

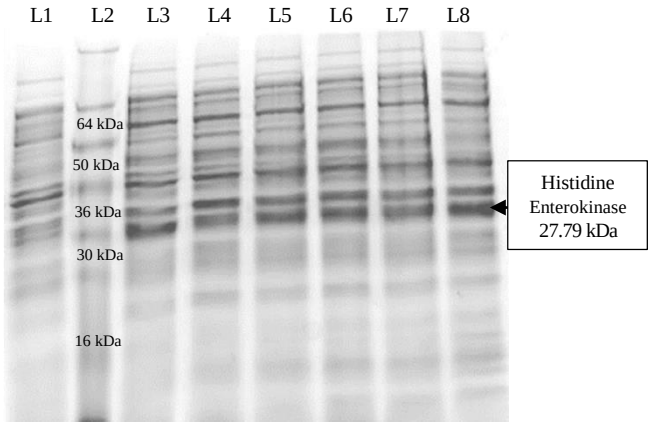


**Figure 4.** Histidine Enterokinase Expression in *E.coli* BL21 DE3 Cells Transformed with Enterokinase/pET20b (+) Plasmid. Lane 1: Clone 1 uninduced sample; Lane 2: See blue prestained protein ladder; Lane 3 to Lane 8: 6 induced clones of histidine enterokinase.

### *Small-scale Induction and Expression Studies*

Six clones of histidine EK were grown and induced in super broth with 100 µg/ml ampicillin in it. During the screening

of the clones on SDS PAGE, expression levels obtained at 27.79 kDa was less (Figure 4). As a counterpart, TrpE H EK was amplified and cloned into pET20b (+) vector and transformed into *E. coli* BL21 (DE3). Eight clones of TrpE H EK/pET20b (+) in BL21(DE3) were grown in super broth for induction studies and 0.1 OD cells after overnight induction were electrophoresed on 12.5% SDS PAGE (Figure 5). Expression of TrpE H EK was observed at ~29.89 kDa in the induced samples and no leaky expression of the desired protein was seen in the uninduced samples. During the optimization of temperature and IPTG concentration for induction, 39 °C and 0.5 mM was found to be effective in increasing the desired protein concentration in the samples loaded on SDS PAGE. Once the induced samples were quantitated using Biorad gel doc EZ imager, we found that the desired protein percentage varied from 60-70%, while the rest was host cell protein. The ratio contributed by the inclusion body comprised a major portion in the cells.

Based on the results of the shake-flask experiments, the TrpE fusion tag was found to be suitable for the expression of EK. The fifth clone (#5) was selected for optimized enzyme production in the bioreactor based on the Biorad gel doc EZ imager quantitation tools.

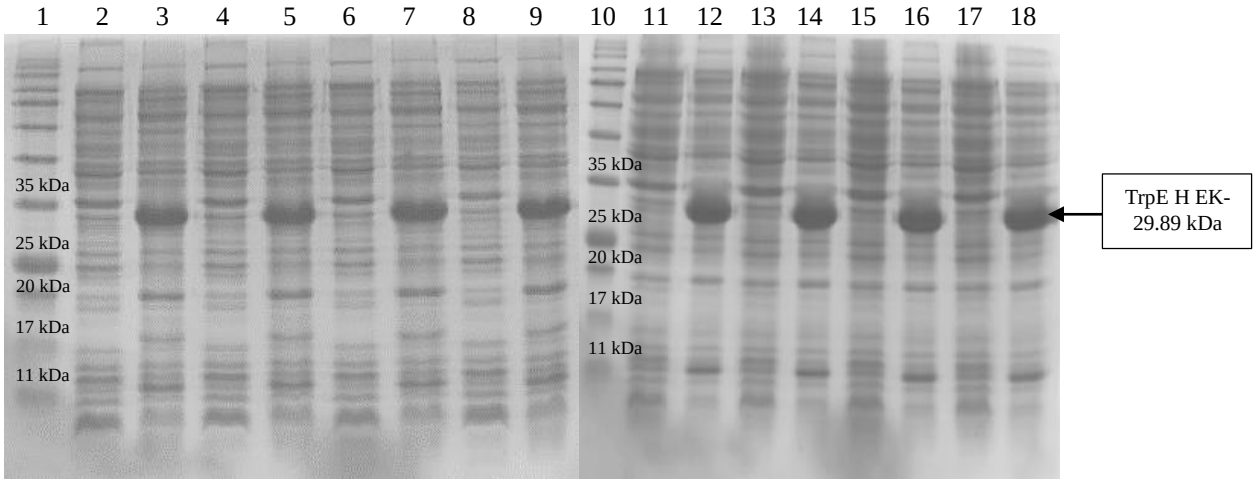


**Figure 5.** TrpE H EK Expression at Shake-flask Levels with 0.5 mM IPTG and 39 °C Induction Temperature. Lane 1 and 10: Pre-stained protein ladder from Himedia; Lanes 2, 4, 6, 8, 11, 13, 15, and 17: Uninduced clones 1-8 of TrpE H EK; Lanes 3, 5, 7, 9, 12, 14, 16, and 18: Induced clones 1-8 of TrpE H EK.

### *Western Blotting*

Complete transfer of protein was observed with the transfer of pre-stained protein ladder bands onto the membrane and no bands were observed on SDS PAGE after transfer. After the development of blot with NBT BCIP substrate, a brown color band appeared parallel to TrpE H EK band on the blot without any non-specific bands. This result signifies that the samples contain the protein of interest with a histidine tag in it (Figure 6).

### *Fed-batch Fermentation*

Once the optimum parameters such as media composition, induction temperature and IPTG concentration were determined from the shake-flask experiments, further studies were performed using a 2-L bench-scale bioreactor to validate the expression levels. Based on the gel image electrophoresed post-fermentation batch, it was inferred that the expression levels constantly increased from the second hour until the end of fermentation (Figure 7).

Even though we tried optimizing the fermenter media composition, the previous media worked better in all the aspects except with WCW during induction, IPTG concentration, induction temperature, induction time and the

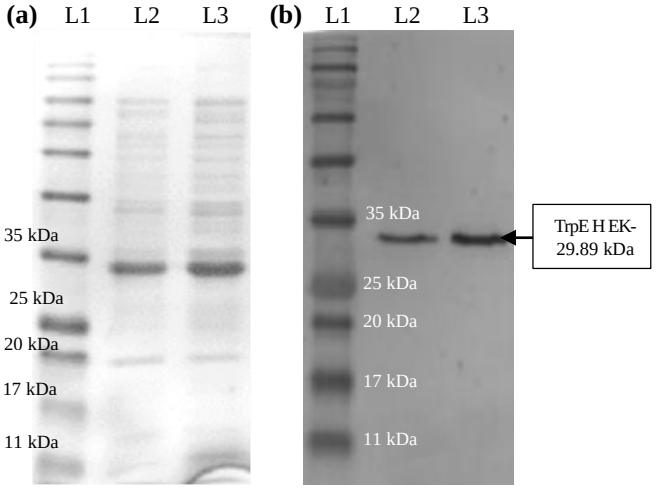


**Figure 6. (a)** SDS PAGE for Induced Samples of TrpE Histidine Enterokinase; **(b)** TrpE Histidine Enterokinase Expression Confirmed by Western Blotting on Nitrocellulose Membrane. Lane 1: Hi media prestained protein ladder; Lane 2: induced sample 1; Lane 3: induced sample 2.

refolding conditions. A 0.5 mM IPTG concentration and 39 °C temperature showed better expression both at shake flask and bioreactor levels. Reduction in IPTG concentration drastically reduced the cost of EK production, as half of the

media cost is accommodated by IPTG. The WCW study was carried out with 200, 300 and 350 g/L, but 300 g/L proved to be the best for producing EK. Whereas the two other WCW trials showed a reduced inclusion body percentage. Optimization of IPTG induction time was evaluated by harvesting small amount of broth at 4, 6, 8 and 10 hours. The desired protein concentration reduced after 6 hour induction. Consequently, 6 hour induction was concluded for enhanced protein production.

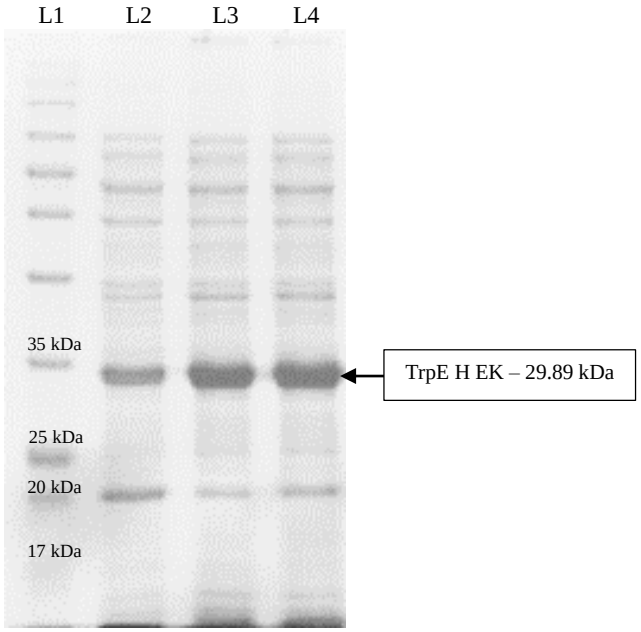


**Figure 7.** TrpE H EK Expression Levels from Bioreactor Studies. Lane 1: Himedia pre-stained protein ladder; Lane 2: 2 hour of induction; Lane 3: 4 hour of induction, Lane 4: 6 hour of induction.

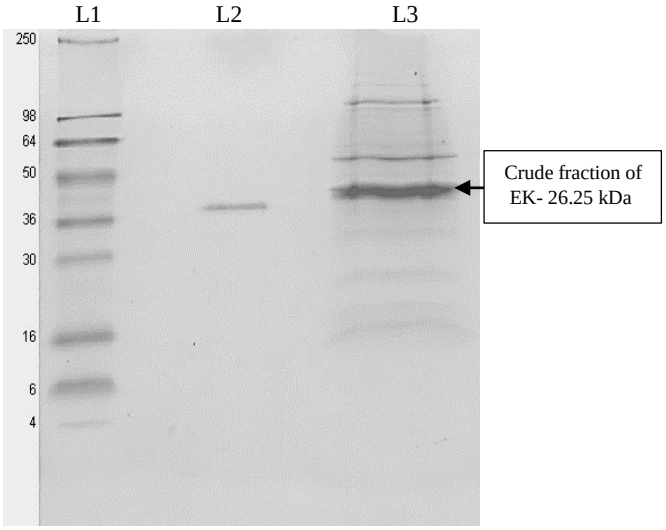


**Figure 8.** Tricine SDS PAGE Profile of the Crude and Purified Sample. Lane 1: See blue prestained protein ladder; Lane 2: Protein fraction eluted after Anion exchange chromatography; Lane 3: Crude protein after EK digestion to be loaded into ion exchange chromatography column.

### *Cell Lysis and Recovery of the Inclusion Bodies*

From the fermentation batches of TrpE H EK, 34.3 g/L of inclusion bodies were obtained. The inclusion body percentage obtained after cell lysis was found to be 11.43%. No loss occurred during subsequent washing steps; hence, the process was efficient.

### *Solubilization and Refolding of EK*

After solubilizing the inclusion bodies with a higher concentration of denaturants at a ratio of 1:60 (w/v), the solubilized protein was not subjected to Ni-NTA purification even though histidine tag was added during gene construction as most of the HCP's were eliminated with the efficient washing steps. The complete process of refolding was performed at ≤10 ºC to prevent the activation of EK. During the refolding step, the concentration of protein was maintained between 0.6-1 mg/ml for increasing the efficiency of refolding. The protein mixture was diluted to 1:40 ratio with novel refolding buffer and was allowed to refold at pH 8.5 for 72 hours at 4 ºC, enhanced the refolding efficiency, reduced the amount of early activation, and resulted in maximum recovery during the activation steps. Refolding efficiency was completely uplifted by pH maintenance in a range of 8.5 ± 0.2 for 72 hours. However, implementation of refolding conditions with pH above 8.5 proved to be effective in avoiding auto activation during refolding and reduced the degradation of protein. Refolding for 72 hours to form disulfide bonds also influenced in increasing the refolded protein concentration.

### *Activation of EK*

Novel activation buffer composition defined for activation of EK resulted in complete removal of TrpE fusion tag that could be observed on Tricine SDS PAGE (Figure 8).[27] The refolded mixture was thoroughly mixed with the activation buffer for 12-14 hours to achieve maximum activation of EK. The addition of EK led to auto-activation of the enzyme, and the fusion tag was removed. Almost 90% of EK was devoid of fusion tag after the activation step. As the size of the fusion tag was too small (< 2.5 kDa), proper resolution of the bands was not possible and the difference could not be observed on SDS PAGE. The crude mixture containing cleaved TrpE and EK was electrophoresed on Tricine SDS PAGE as per the protocol from Schagger et al. is shown in figure 8. The EK samples after refolding seems to run higher than the desired molecular weight due to the disulfide bond formation whereas the samples before refolding ran according to its molecular weight. The activation buffer pH maintained at a range of 7.4 ± 0.2 helped in maximum activation of the protein and removal of fusion tag after 1 hour of incubation.

### *Purification and Storage*

Purification involving Sepharose Q anion exchange

chromatography was performed to selectively bind the protein and elute it. The samples were collected when the A280 peak rose at around 50% and the collection was stopped when the peak started descending after adding 70% of buffer B. Ion exchange chromatography led to the complete removal of host cell proteins and other contaminating proteins. The fractions eluted with 200 mM Tris–HCl (pH 8.0) and 1 M sodium chloride was diluted with 1:1 of sterile 100% glycerol to obtain 100 mM Tris–HCl (pH 8.0), sodium chloride of 500 mM with 50% glycerol that was used for storage of EK at -20 °C. The purified sample of EK (26.25 kDa) eluted after the chromatography step was loaded on Tricine SDS PAGE (Figure 8).

### *Estimation of Protein Concentration*

The samples showing absorbance at 280 nm were pooled, and their final concentration was estimated to be 1.102 mg/ml using the BCA method by running them against BSA standards. The final concentration of the enzyme was 22.04 mg/ml (1:20) after taking the dilution factor into consideration. The volume of the sample obtained after purification was 10.94 ml, leading to 241 mg of EK production per liter of the broth. The recovery and yield percentage obtained at each stage of purification are shown in Table 2.

### *EK Activity Assay*

The conversion rate of the EK in-house sample, which was noted at intervals of 15 seconds, increased with time and the linearity difference ranged between 0.008 and 0.012. The average difference obtained for the standard and sample was 0.0163 and 0.01125 respectively. The in-house EK activity was 2.37 Units/µl upon comparison with the invitrogen standard (3.43 Units/µl). The increased values after enzyme addition were lined up and R-squared value obtained was 0.99, revealing the maintenance of enzyme efficiency for the initial 5 minutes.

### *HPLC for EK*

EK in-house sample was injected into the RP-HPLC column eluted EK at 9.45 minutes that matched with the retention time of standard EK. One µl of protein was injected into the column irrespective of the concentration. 2-3% non-specific protein peaks were observed in case of in-house sample (Figure 9). As the achieved purity is comparably better than standard, it could be considered that our process leads to the production of EK with higher purity.

**Table 2.** Recovery and Yield Obtained at Each Stage of Purification

| Process Steps | Recovery (%) | Loss (%) | Protein(g/L) | Total Enterokinase yield (%) |
|---|---|---|---|---|
| Inclusion bodies | 100 | 0 | 34.300 | 100.00 |
| Solubilization | 7 | 93 | 2.401 | 7.00 |
| Refolding | 47 | 43 | 1.128 | 3.29 |
| Concentration | 78 | 22 | 0.880 | 2.57 |
| Enzyme reaction for fusion tag removal | 54 | 46 | 0.475 | 1.39 |
| After Ion Exchange Chromatography | 59 | 41 | 0.280 | 0.82 |
| Overall Recovery (Storage buffer after dialysis) | 86 | 14 | 0.241 | 0.70 |

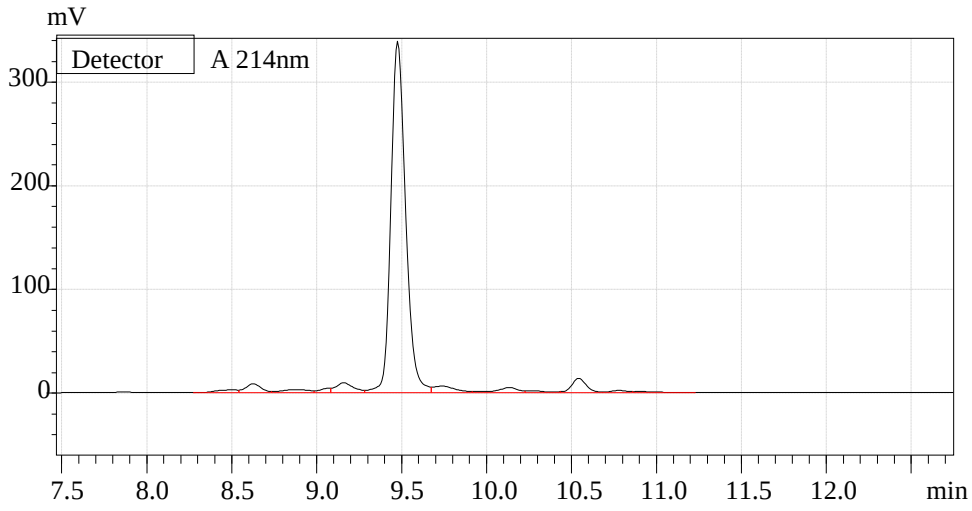


**Figure 9.** Figure Represents Pooled Fraction of Purified In-house Enterokinase HPLC Profile at 9.45 Minutes.

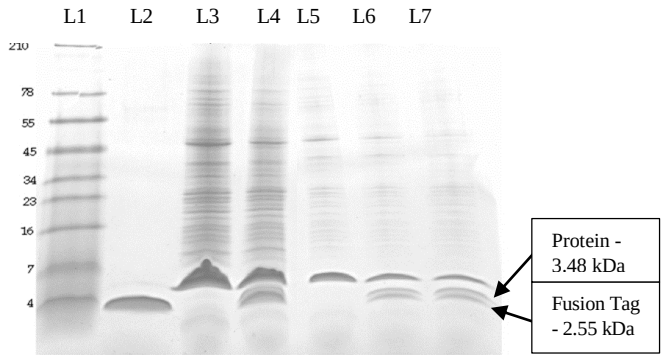


**Figure 10.** Tricine-SDS PAGE Profile of the Substrate Cleavage Assay for Inhouse Enterokinase. Lane 1: See Blue pre-stained protein ladder; Lane 2: Cleaved Protein standard; Lane 3: Substrate control - 15 µg protein (without digestion- 6.03 kDa); Lane 4: Substrate digested sample- 15 µg protein (1 unit) (Not optimized); Lane 5: Control- 5 µg Substrate protein (without digestion - 6.03 kDa); Lane 6: Digested substrate after optimization- 5 µg protein (16 hours digestion) & Lane 7: Digested substrate after optimization- 5 µg protein (20 hours digestion).

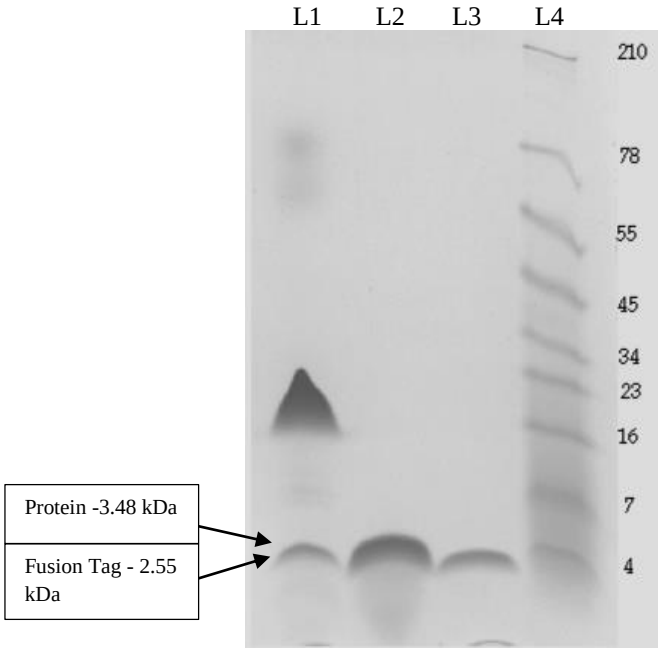


**Figure 11.** Tricine-SDS PAGE Profile of the Substrate Cleavage Assay for Standard Enterokinase. Lane 1: Substrate digested with standard (1 unit)- 15 µg protein; Lane 2: Cleaved protein standard (3.48 kDa)- 15 µg; Lane 3: Cleaved protein standard - 5 µg; Lane 4: See Blue pre-stained protein ladder.

### *Substrate Reaction Assay for the In-house Enzyme*

After substrate cleavage assay studies, it was concluded that almost 70-80% cleavage was observed when five microgram of the tagged protein was used per reaction with an incubation period of 16-20 hours at 37 °C for complete cleavage efficiency of the enzyme on the substrate. As the fusion tag (2.55 kDa) and protein (3.48 kDa) are of equimolar ratio (3 ± 0.6 kDa), an equal quantity of cleaved tag and protein was observed on Tricine SDS PAGE. Optimization of the activation buffer composition, quantity of protein to be taken, units of EK to be added, and the time of incubation has augmented the efficiency of the inhouse enzyme and cleaved the maximum amount of the protein (Figure 10). About 10-25% ratio of cleaved protein has been observed when the same substrate was cleaved with a unit of standard EK. In case of standard both the bands are merged, due to the lower concentration of protein being cleaved. The upper band without cleavage has run higher than required molecular weight due to the salt concentration in the buffer used for the standard as per the protocol mentioned with the standard (Figure 11). Standard EK (1 unit) was able to cleave around 25-30% of the protein, whereas in house EK was able to cleave 40-50% of the protein when one unit is used to cleave 15 µg of protein.

## Discussion

Enzymes play a major role in the human body and in the production of major end products by their specific action on the substrate and product formation. Demand for a biosimilar is rapidly increasing not mainly because of the increased population using the biosimilar, but due to the costly reagents used in the processing of biosimilars like trypsin, EK and carboxypeptidase B. Although these enzymes could be isolated from the natural host, regulatory issues linked to the viral load in the final product prevents them from doing so. Unlike other organisms, *E. coli* is known for its faster growth and fermentation with inexpensive media. It should also be noted that the enzymes without fusion tags in *E. coli* are either prone to degradation or expression levels are reduced. Fusion tags >10 kDa leads to a reduction in the yield of the final product as the molar ratio of protein to tag plays a major role in the final recovery of protein. Histidine fusion tag yielded less EK as described by Wolfgang Skala et al. with their process achieving around 2.6 grams of inclusion bodies per liter of broth leading to the production of 1.7 mg/L of EK at shake flask levels in the end. The same fusion tag was exploited for our process with a novel fermentation process with fed-batch strategy which led to the production of 78 mg/L of EK proving the fermentation and purification process to be more beneficial for an increased production of the recombinant enzyme. Previously we had expressed trypsin using TrpE fusion tag that showed better expression levels than before.[28] Taking the advantage and disadvantages of the above strategies into consideration to increase the final yield of protein, we expressed EK with TrpE fusion tag to push the protein towards the inclusion body in *E. coli* as an expression host and produce a commercially viable clone. The obtained high level of expression was mainly due to the hydrophobic core of the fusion tag that boosted the production of inclusion bodies. The present study involved a reduction in the concentration of IPTG from 1 mM to 0.5 mM, the expression of TrpE H EK almost remained unchanged even after a reduction in

IPTG concentration, thereby reducing the cost of the fermenter run. The Dissolved Oxygen (DO) level maintained above 20% throughout the fermentation and helped in achieving higher biomass within a short span of time (8 to 10 hours). The use of carbon and nitrogen feed as a mild acid of pH-6, encouraged the growth of cells by reducing the consumption of acid. The induction time of 6 hours showed better expression levels, when compared to 8-hour induction that is a usual induction time used in *E. coli* Bioreactor batches. Based on the gel images of shake flask and bioreactor studies, the protein expression at the bench-scale bioreactor level was superior when compared with that obtained at the shake-flask level which was confirmed by estimating the protein by Biorad EZ gel imager. The growth and induction phases were completed within 16 ± 2 hours, and the higher induction levels in bioreactor enhanced the cost-effectiveness of the process. The reduction in the number of hours for fermentation and IPTG addition led to a reduction in cost per fermenter run by several folds. Finally, 300 g/L WCW during induction with 0.5 mM IPTG at 39 °C temperature with 6-hour induction and refolding at pH 8.5 produced a yield of EK higher than the levels achieved never before. For the production of EK, optimization of conditions during induction played a major role, when compared to media and growth of the cells.

The pH needs to be maintained in a range of ± 0.2 during refolding and activation steps to obtain highly refolded and completely cleaved EK with better quality. Novel buffer formulations used for refolding, activation of EK and chromatography along with optimization in bioreactor parameters helped in increasing the yield of the desired protein from 78 mg/L (Histidine EK) to 241 mg/L (TrpE histidine EK) in our lab. The yield has increased drastically when compared with the conventional methods in which the maximum yield achieved was only 106 mg/L. An enzymatic assay using standard protocol revealed better activity by showing a constant difference between the values over time inferring the enhanced efficiency of the enzyme. Anion exchange chromatography based on its charge allowed specific protein binding to the column and removed all other impurities. EK after purification was used to cleave the fusion tag attached to the protein in order to estimate the efficiency of the enzyme. In-house EK was able to cleave almost 70-80% of the fusion tag from the protein. The in-house enzyme matches almost with the standard in case of activity. Thus, our process sets a benchmark for the production of enzymes in the future at a commercial scale.

## Conclusion

In this study, new methods have been adopted to enhance the production of EK. Suitable fusion tag, fermentation process, and buffer compositions were exclusively developed for this study and the use of anion exchange chromatography yielded 241 mg/L of active EK. We observed around 2-2.5-fold increase in the enzyme titer when compared to traditionally used fusion tags. The findings from this research could aid in cost reduction, enabling the production of EK in a commercially viable manner. Therefore, this fusion tag as well as the fermentation strategies could possibly be explored for the production of other industrially relevant recombinant enzymes to achieve an improved yield.

## Acknowledgements

The authors wish to thank Anthem Biosciences Pvt Ltd for providing a suitable platform to carry out this research.

## Authors' Contributions

SNN conceived the original idea, planned, carried out the experiments and took the lead in writing the manuscript. JMA contributed to designing the framework and implementation of the research. SS developed the theory, analysis of data and performed the computations. GS investigated and supervised the findings of this study. All the above authors provided inputs and helped to shape the research, analysis and manuscript.

## Conflict of Interest Disclosures

The authors of this paper declare that they have no conflicts of interest with any financial organization regarding the material discussed in the manuscript.

## References

1. Niu LX, Li JY, Ji XX, Yang BS. Efficient expression and purification of recombinant human enteropeptidase light chain in *Escherichia coli*. Braz Arch Biol Technol. 2014;58:154-65. doi:10.1590/S1516-8913201400094
2. Baratti J, Maroux S, Louvard D, Desnuelle P. On porcine enterokinase. Further purification and some molecular properties. Biochim Biophys Acta Enzymol. 1973;315(1): 147-61. doi:10.1016/0005-2744(73)90138-1
3. Anderson LE, Walsh KA, Neurath H. Bovine enterokinase. Purification, specificity, and some molecular properties. Biochemistry. 1977;16(15):3354-60. doi:10.1021/bi00634a011
4. Liepnieks JJ, Light A. The preparation and properties of bovine enterokinase. J Biol Chem. 1979;254(5):1677-83. doi:10.1016/S0021-9258(17)37826-2
5. Magee AI, Grant DA, Hermon-Taylor J. Further studies on the subunit structure and oligosaccharide moiety of human enterokinase. Clin Chim Acta. 1981;115(3):241-54. doi:10.1016/0009-8981(81)90238-2
6. Ahsan N, Aoki H, Watabe S. Overexpression in *Escherichia coli* and functional reconstitution of anchovy trypsinogen from the bacterial inclusion body. Mol Biotechnol. 2005;30(3):193-205. doi:10.1385/MB:30:3:1 93
7. Joseph BC, Pichaimuthu S, Srimeenakshi S, Murthy M, Selvakumar K, Ganesan M, et al. An overview of the parameters for recombinant protein expression in *Escherichia coli*. J Cell Sci Ther. 2015;6(5):1000221. doi:10.4172/2157-7013.1000221
8. McCoy J, LaVallie E. Expression and purification of thioredoxin fusion proteins. Curr Protoc Mol Biol. 1994;28(1):16-8. doi:10.1002/0471142727.mb1608s28
9. Harper S, Speicher DW. Expression and purification of

GST fusion proteins. Curr Protoc Protein Sci. 2008;52(1): 6-6. doi:10.1002/0471140864.ps0606s52
10. Trigoso YD, Evans RC, Karsten WE, Chooback L. Cloning, expression, and purification of histidine-tagged *Escherichia coli* dihydrodipicolinate reductase. PloS one. 2016;11(1): e0146525. doi:10.1371/journal.pone.0146525
11. Booth WT, Schlachter CR, Pote S, Ussin N, Mank NJ, Klapper V, et al. Impact of an N-terminal polyhistidine tag on protein thermal stability. ACS Omega. 2018;3(1): 760-8. doi:10.1021/acsomega.7b01598
12. Rosano GL, Ceccarelli EA. Recombinant protein expression in Escherichia coli: advances and challenges. Front microbiol. 2014;5:172. doi:10.3389/fmicb.2014.00 172
13. Skala W, Goettig P, Brandstetter H. Do-it-yourself histidine-tagged bovine enterokinase: a handy member of the protein engineer's toolbox. J Biotechnol. 2013;168(4): 421-5. doi:10.1016/j.jbiotec.2013.10.022
14. 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):133-9. doi:10.1016/S1046-5928(03)00159-1
15. Tan H, Wang J, Zhao ZK. Purification and refolding optimization of recombinant bovine enterokinase light chain overexpressed in *Escherichia coli*. Protein Expr Purif. 2007;56(1):40-7. doi:10.1016/j.pep.2007.07.006
16. Svetina M, Krasevec N, Gaberc-Porekar V, Komel R. Expression of catalytic subunit of bovine enterokinase in the filamentous fungus *Aspergillus niger*. J Biotechnol. 2000;76(2-3):245-51. doi:10.1016/S0168-1656(99)0019 1-1
17. Vozza 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*. Nat Biotechnol. 1996;14(1):77-81. doi:10.1038/nbt0196-77
18. Choi SI, Song HW, Moon JW, Seong BL. Recombinant enterokinase light chain with affinity tag: expression from *Saccharomyces cerevisiae* and its utilities in fusion protein technology. Biotechno Bioeng. 2001;75(6):718-24. doi:10.1002/bit.10082
19. Hwang PM, Pan JS, Sykes BD. Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in *Escherichia coli*. FEBS lett. 2014;588(2):247-52. doi:10.1016/j.febslet.2013.09.028
20. Derynck R, Roberts AB, Winkler ME, Chen EY, Goeddel DV. Human transforming growth factor-α: precursor structure and expression in *E. coli*. Cell. 1984;38(1):287-97. doi:10.1016/0092-8674(84)90550-6
21. Itakura K, Hirose T, Crea R, Riggs AD, Heyneker HL, Bolivar F, et al. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. Science. 1977;198(4321):1056-63. doi:10.1126/science. 412251
22. Goeddel DV, Kleid DG, Bolivar F, Heyneker HL, Yansura DG, Crea R, et al. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. Proc Natl Acad Sci U S A. 1979;76(1):106-10. doi:10.1073/ pnas.76.1.106
23. Al-Tubuly AA. SDS-PAGE and Western Blotting. Diagnostic and Therapeutic Antibodies. Humana Press, Totowa, NJ. 2000;391-405.
24. Mahmood T, Yang PC. Western blot: technique, theory, and trouble shooting. N Am J Med Sci. 2012;4(9):429-34. doi:10.4103/1947-2714.100998
25. Kim MD, Lee WJ, Park KH, Rhee KH, Seo JH. Two-step fed-batch culture of recombinant *Escherichia coli* for production of *Bacillus licheniformis* maltogenic amylase. J Microbiol Biotechnol. 2002;12(2):273-8.
26. Ramanan RN, Tey BT, Ling TC, Ariff AB. Classification of pressure range based on the characterization of *Escherichia coli* cell disruption in high pressure homogenizer. Am J Biochem Biotechnol. 2009;5(1):21-9. doi:10.3844/ ajbbsp.2009.21.29
27. Schagger H. Tricine–sds-page. Nat Protoc. 2006;1(1):16-22. doi:0.1038/nprot.2006.4
28. Nanjundaiah SN, Ma J, Sukumaran S, Sambasivam G. Development of clone with novel TrpE fusion tag in *E. coli* for overexpression of trypsin in a bench-scale bioreactor. Prep Biochem Biotechnol. 2021;51(2):144-52. doi:10.1080/10826068.2020.1799392